Journal of Biomolecular NMR, **25:** 225–234, 2003.
KLUWER/ESCOM

The precision of NMR structure ensembles revisited

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

[a]*Centre for Molecular and Biomolecular Informatics, University of Nijmegen, The Netherlands;* [b]*Department of NMR-spectroscopy, University of Utrecht, The Netherlands;* [c]*Department of Biophysical Chemistry, University of Nijmegen, The Netherlands*

Received 17 October 2002; Accepted 1 November 2002

Key words: accuracy, NMR structures, precision, structure refinement, structure validation

Abstract

Biomolecular structures provide the basis for many studies in research areas such as structure-based drug design and homology modeling. In order to use molecular coordinates it is important that they are reliable in terms of accurate description of the experimental data and in terms of the overall and local geometry. Besides these primary quality criteria an indication is needed for the uncertainty in the atomic coordinates that may arise from the dynamic behavior of the considered molecules as well as from experimental- and computational procedures.

In contrast to the crystallographic B-factor, a good measure for the uncertainty in NMR-derived atomic coordinates is still not available. It has become clear in recent years that the widely used atomic Root Mean Square Deviation (RMSD), which is a measure for the precision of the data, overestimates the accuracy of NMR structure ensembles and therefore is a problematic measure for the uncertainty in the atomic coordinates.

In this study we report a method that yields a more realistic estimate of the uncertainty in the atomic coordinates by maximizing the RMSD of an ensemble of structures, while maintaining the accordance with the experimentally derived data. The results indicate that the RMSD of most NMR structure ensembles can be significantly increased compromising neither geometric quality nor NMR data. This maximized RMSD therefore seems a better estimate of the true uncertainty in the atomic coordinates.

Introduction

The precision and accuracy of NMR structure ensembles have been subject of a long-standing debate in the field of biomolecular structure determination by NMR-spectroscopy. In an elaborate discussion Zhao and Jardetzky (Zhao and Jardetzky, 1994) have addressed the fundamental aspects of the problem and concluded that the accuracy of NMR structure ensembles is at best of the order of 1 to 2 Å. Although the relevance of their analysis has been criticized (e.g., see Chalaoux et al., 1999), their main point, i.e., the importance to distinguish between precision and accuracy, still stands.

*To whom correspondence should be addressed. E-mail: vriend@cmbi.kun.nl

The accuracy is a measure of closeness of the structures to the true structure and can only be obtained when a 'gold standard' is available, like in the case of a simulated data set. For real NMR structures the accuracy is often calculated with respect to a reference X-ray structure of the same molecule (e.g., see Kuszewski et al., 1999; Linge and Nilges, 1999; Sprangers et al., 2000; Tjandra et al., 2000). Although such a reference structure is not a 'gold standard' the comparison provides at least some measure of the accuracy. In fact, examples are known where the X-ray structure satisfies independently measured NMR observables better than the NMR ensemble that was refined without the inclusion of these observables (e.g., see Spronk et al., 2002).

The precision, expressed as the coordinate RMSD of the ensemble, is commonly used as an indication of

how well the structures have been refined. Normally, ensembles are generated by selection of a number of low energy structures from a larger ensemble that fulfills acceptance criteria based on the experimental data. This approach is designed to find a set of structures that represents the global energy minimum of the molecule. The RMSD of the ensemble therefore depends not only the quality and amount of available data but also on the procedure used for selection of the structures.

It has become clear from several studies using both real- and simulated NMR data that the precision normally exceeds the accuracy (Havel and Wüthrich, 1985; Brünger et al., 1993; Clore et al., 1993; Zhao and Jardetzky, 1994; Gronenborn and Clore, 1995; Chalaoux et al., 1999). For real NMR data the accuracy that can be achieved is at best equal to the precision of the ensemble of structures. Therefore, in order to get a good impression of the accuracy it is required to determine the minimum precision (i.e. maximum RMSD) of the ensemble given the experimental restraints.

There are strong indications that current procedures for structure calculation and selection result in ensembles that underestimate the fluctuations and do not reflect the true conformational freedom (Torda et al., 1990; Scheek et al., 1995; Pfeiffer et al., 1997; Horstink et al., 2000). An accurate representation of a solution structure should not only describe the global minimum but also reflect the inherent dynamics of the molecule and the uncertainty in the experimentally derived NMR data. Therefore, we describe an approach in which iterative re-sampling and refinement of an ensemble is used to assess the minimum precision. The method is designed to improve the sampling and representation of the conformational space that is defined by the experimental restraints. The procedure, which we will refer to as 're-sampling', was tested on five different data sets: A simulated data set for the protein Crambin (Jelsch et al., 2000) and experimentally derived data sets for Ubiquitin (Cornilescu et al., 1998), the protein-peptide complex PAH2-Mad1 (Spronk et al., 2000), the immunoglobulin binding domain of streptococcal protein G (Gronenborn et al., 1991) and the scorpion toxin chlorotoxin (Lippens et al., 1995). The results show that for all test cases the RMSD of the original ensembles can be increased substantially, while fitting the experimental data well, and maintaining good local and overall geometric quality of the structures.

Materials and methods

Re-sampling NMR ensembles

The basic idea behind our method to find the lower limit of the precision of NMR structure ensembles is to iteratively increase the RMSD of the ensemble by randomly generating structures around a set of experimental input structures and subsequent refinement and fitting to the experimental data (Figure 1).

Generation of structures

Generation of structures was done using the program CONCOORD (version 1.2), which was originally developed to probe the conformational freedom of proteins (de Groot et al., 1997). CONCOORD generates random protein structures that fulfill a set of upper and lower distance limits. These distance limits are derived from the distances measured in the experimental structures, secondary structure elements (generated by DSSP (Kabsch and Sander, 1983)) and hydrophobic and electrostatic interactions. The size of the upper and lower distance limits can be scaled using the CONCOORD damp factor, which is used to control the spread in the newly generated ensemble. For further details on the program CONCOORD please see de Groot et al. (1997) and http://www.mpibpc.gwdg.de/abteilungen/071/bgroot/concoord.html.

In the current implementation CONCOORD distance bounds were generated using default parameters, whilst retaining all hydrogen-atoms and varying the CONCOORD damp factor (see below). CONCOORD structures were generated for each individual conformer of the input ensemble. This procedure was repeated until after refinement two accepted structures were obtained per input structure. In order to slowly increase the RMSD of the ensemble the CONCOORD distance limits were scaled with the CONCOORD-damp factor starting at 0.25. This factor was incremented by 0.25 (to a maximum of 1.25) whenever the backbone RMSD of an ensemble of accepted structures did not increase more than 5% in two subsequent cycles. The re-sampling procedure was automatically aborted when less than 5% increase in the backbone RMSD was observed in 5 iterations.

Structure refinement

The structures generated by CONCOORD were optimized and fitted to the experimental data by a short

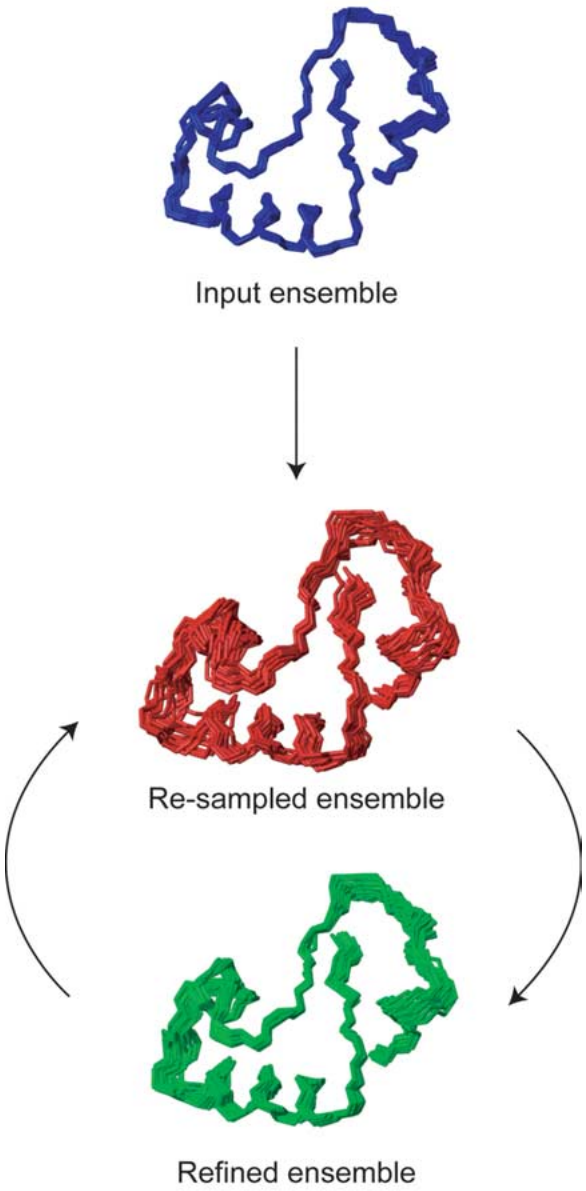

Figure 1. Schematic drawing of the re-sampling procedure. The input ensemble is re-sampled using CONCOORD and subsequently refined and re-sampled iteratively. Figure generated with MOLMOL (Koradi et al., 1996).

restrained Molecular Dynamics refinement in explicit solvent using ARIA (version 1.1) (Nilges et al., 1997; Linge and Nilges, 1999; Linge et al., 2001) and CNS (version 1.1) (Brünger et al., 1998) with the following modifications: (1) For the sake of speed the MD heating, high temperature and cooling stages were shortened to 1.25, 0.25 and 2.5 ps, respectively. (2) The distributions of the covalent geometry parameters were improved by changing the force constants for the bond lengths, bond angles, omega angles and improper dihedral angles in the cooling stage to 130 kcal/(mole Å^2) and 80, 18, 180 kcal/(mole rad^2), respectively. An additional improper dihedral angle with a force constant of 35 kcal/(mole rad^2) was defined for the planarity of the atoms bonded to the backbone carbonyl atom. The force field parameters used in the water-refinement were those from PARALLHDG5.2, using the OPLS parameters for the non-bonded interactions (Jorgensen and Tirado-Rives, 1988) and the TIP3P water model (Jorgensen et al., 1983).

Selection of structures

Although any selection criteria can be used, we have chosen the widely used protocol to accept structures without violations of distance and dihedral angle restraints larger than 0.5 Å and 5°, respectively. Of the ensemble of accepted structures half of the structures with the lowest experimental restraint energy were used for analysis.

Two different re-sampling procedures were tested. In the first run half of the accepted structures with the lowest restraint energy were used as the input for the next round of re-sampling (referred to as low energy selection). In the second run half of the accepted structures with the highest pair-wise RMSD were used as input for the next round of re-sampling (subsequently referred to as high RMSD selection).

Structure validation

An important aspect of the re-sampling procedure is the validation of the structures in the resulting ensembles. It is essential that these structures still fit the NMR data well within the experimentally determined error bounds while increasing the RMSD during re-sampling. This is done using the abovementioned selection criteria for the structures. The fit of the accepted structure ensemble to the data is expressed as the RMS deviation of the calculated to the input restraints.

Next, the structures are compared to a reference database of well-refined high-resolution X-ray structures using the program WHAT IF (Vriend, 1990; Hooft et al., 1996) in order to address the local and overall geometric quality of the structure ensembles. WHAT IF provides so-called structure Z-scores for the packing quality, Ramachandran plot, χ-1/χ-2 rotamer distribution and backbone conformation. It is

important to realize that a structure Z-score equals the number of standard deviations away from the mean of the database and is therefore a normality score and not a quality score (Hooft et al., 1997). We consider structure Z-scores to be within acceptable ranges when they are between −3 and +3. Structure Z-scores outside these ranges do not necessarily mean that the structures are bad, but are less likely to be correct. It is however worrisome that average Z-scores in NMR structure ensembles are found to be around −4 for all indicators, which is a significant deviation from the database of reliable high-resolution X-ray structures (see Table 1).

Further, WHAT IF analyzes bond lengths and bond angles, omega angles, chirality and side chain planarity in the structures. The values are compared to the internal WHAT IF database, which contains the parameters for bond lengths and bond angles described by Engh and Huber (1991), for omega angles as described by MacArthur and Thornton (1996), for side chain planarity as derived from the Cambridge Small molecule Database (Allen et al., 1983; Hooft et al., 1996) and for chiralities derived from high-resolution X-ray structures. From these comparisons RMS Z-scores are calculated for each parameter. An RMS Z-score is an indicator for the variance in each set of the parameters. An RMS Z-score is equal to 1.0 if the distribution has the same average and variance as the reference distribution. Values lower or higher than 1 indicate that the parameter has a lower or higher variance, which in the case of protein structures indicates too tight or too loose restraining of the geometry (see also http://www.cmbi.kun.nl/gv/pdbreport/checkhelp/intro.html). For completeness it should be noted that RMS Z-scores are meaningless if the averages of the compared distributions are not the same.

Finally, the number and size of the interatomic bumps and the presence of unsatisfied buried hydrogen-bond donors and acceptors are analyzed in order to assess the packing quality and energetics of the resulting structures after re-sampling.

Data sets

The simulated data of the protein Crambin (46 residues) were derived from the atomic resolution X-ray structure (0.54 Å, pdb-entry 1EJG (Jelsch et al., 2000)). For this set an artificial perfect set of 1497 NOE distance restraints was calculated containing all proton–proton distances smaller than 5 Å, upon which 20% error bounds were added (Spronk et al., 2002). For Ubiquitin (76 residues) 2727 distance, 27 hydrogen bond and 98 dihedral angle restraints were taken from pdb-entry 1D3Z (Cornilescu et al., 1998). For PAH2-Mad1 (98 residues) we used 2176 distance-, 27 hydrogen bond and 26 dihedral angle restraints as described previously (Spronk et al., 2000). For these three data sets input structures for the re-sampling method were calculated using the CHARMM22 water-refinement protocol (Spronk et al., 2002).

In addition, we tested the method on two structure ensembles and their corresponding data sets obtained from the BioMagResBank: The immunoglobulin binding domain of streptococcal protein G (pdb-entry 1GB1, 56 residues, 854 distance, 68 hydrogen bond and 94 dihedral angle restraints (Gronenborn et al., 1991)) and the scorpion toxin chlorotoxin (pdb-entry 1CHL, 36 residues, 183 distance and 13 dihedral angle restraints (Lippens et al., 1995)). Prior to re-sampling, these two ensembles were refined using a full ARIA1.1 water-refinement scheme, which gives similar improvements as the CHARMM22 water-refinement (Linge et al., 2003). The only differences with the refinement scheme deployed in the refinement of the CONCOORD structures are the lengths of the MD heating, high temperature and cooling stages, which were now set to 10, 2.5 and 25 ps, respectively.

The input ensembles of Crambin, Ubiquitin and PAH2-Mad1 consisted of 20 structures that contained no violations of input distance and dihedral angle restraints larger than 0.5 Å and 5°, respectively. The original ensembles of 1GB1 and 1CHL consisted of 60 and 7 structures, respectively. After water-refinement (prior to the re-sampling procedure) none of these structures contained 0.3 Å distance or 5° dihedral angle restraints violations.

For all data sets the number of accepted structures generated in each cycle was set to two times the number of input structures. Half of the accepted structures were then used for subsequent analysis and as input for the next iteration. All validation analyses were done including all residues for the different test proteins. RMSD calculations included the residues used for RMSD calculations in the original papers for 1GB1 (all residues), Ubiquitin (residues 1–70), 1CHL (residues 2–4 and 13–35) and the PAH2-Mad1 complex (residues 5–41 and 55–80 of PAH2 and all residues of the Mad1 helix). All 46 residues were included for Crambin since for this structure we used a simulated complete data set with all proton–proton distances within 5 Å.

Table 1. Structure quality indicators of X-ray and NMR structures

	X-ray[a] (WHAT IF database)	NMR[b] (released 2002)
Structure Z-scores:		
1st generation packing quality	-0.2 ± 1.0	-4.0 ± 2.2
2nd generation packing quality	-0.0 ± 1.8	-4.1 ± 2.3
Ramachandran plot appearance	0.3 ± 0.9	-4.5 ± 1.8
χ-1/χ-2 rotamer normality	0.4 ± 0.9	-3.4 ± 1.9
Backbone conformation	0.1 ± 1.0	-4.3 ± 3.2
Inter-atomic bumps:		
No. bumps	26 ± 33	53 ± 84
No. bumps/100 res.	7.3 ± 5.8	70 ± 66
Sum of bumps	2.0 ± 3.1	7.6 ± 16.7

The analysis included: [a]489 structures from the internal WHAT IF data base and [b]97 NMR ensembles (1980 individual structures) of protein structures released in 2002.

Results and discussion

We have applied the two different types of the re-sampling procedure, i.e., low energy structure selection and high RMSD structure selection (see Materials and methods), to an ensemble of 20 Ubiquitin structures. In total 13 and 14 iterations were performed for the low energy structure selection and the high RMSD structure selection, respectively, and the effect on the various validation parameters is shown in Figure 2. The large effect on the quality indicators in the first cycle, which is also observed for the Crambin and PAH2-Mad1 data sets (see below and Table 2), is caused by the different force fields used for calculating the original input structures and the refinement in the modified ARIA1.1 protocol. Surprisingly, the low energy selection procedure resulted in Ubiquitin structure ensembles with increased backbone RMSD, while simultaneously the heavy atom RMSD decreased (Figure 2a, left panel). It appears that this effect is correlated with the tighter fit of the structure ensembles to the experimental restraints in the course of the re-sampling procedure (Figure 2b, left panel). A possible explanation for the decrease in heavy atom RMSD and increase in backbone atom RMSD is that more experimental restraints are found for side chains than for the backbone atoms.

In the high RMSD selection run we observe that both the heavy and backbone atom RMSDs increase during re-sampling and a higher maximum is reached for both parameters. Inspection of the quality indicators shows that, disregarding the first cycle in which the change of force field is dominant, the RMS deviation of the distance and dihedral angle restraints increases from 0.0081 ± 0.0006 Å to 0.0087 ± 0.0007 Å and from $1.20 \pm 0.05°$ to $1.31 \pm 0.08°$. It is important however, that the fit to the data is still considerably better than for the input structures, while the RMSDs of the ensemble is increased from 0.5 Å to 0.8 Å for the backbone atoms and from 1.2 Å to 1.4 Å for the heavy atoms. The fact that the fit to the experimental data is still better after re-sampling is because the CHARMM22 force field is much softer than the ARIA force field.

Inspection of the other quality indicators shows that in the first cycle, due to the change in force field, Z-scores are higher for all indicators except for the backbone conformation. During re-sampling the structure Z-scores change slightly, although they all remain well within the ranges found in the reference database of structures. The main difference for the two types of re-sampling is seen for the χ-1/χ-2 rotamer normality, which, in the case of the low-energy structure selection, becomes slightly more positive and in the high RMSD selection slightly more negative. More positive values indicate a closer fit to 'ideal', i.e., corresponding to energetic minima, values in the database (see also Hooft et al., 1997). Thus, it appears that the better fit of the structures to the experimental data during re-sampling correlates with lower heavy atom RMSDs and distributions of χ-1/χ-2 rotamers closer to the ideal values.

Inspection of the number of inter-atomic bumps and presence of unsatisfied buried hydrogen bond-donors and -acceptors shows that mainly the num-

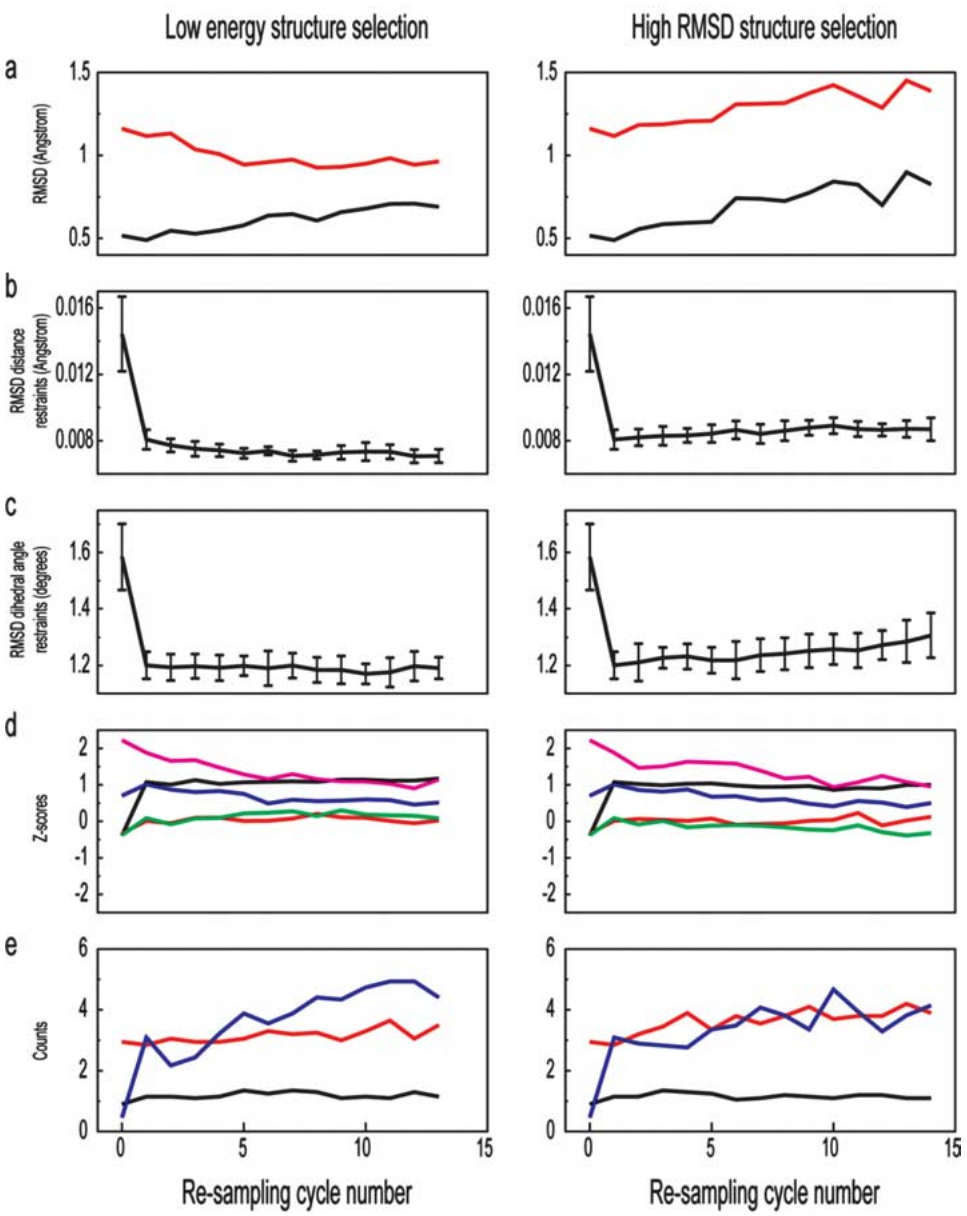

Figure 2. Validation results for the two types of re-sampling procedures for Ubiquitin. Standard deviations have been omitted for clarity in (a), (d) and (e). Average standard deviations for these panels are listed below in parentheses: # low energy structure selection, § high RMSD structure selection. (a) Backbone (black, # 0.17, § 0.19) and heavy atom (red, # 0.23, § 0.22) RMSDs. (b) RMS deviations for distance restraints. (c) RMS deviations for dihedral angle restraints. (d) Structure Z-scores for: 1st generation packing quality (black, # 0.22, § 0.24), 2nd generation packing quality (red, # 0.30, §0.36), Ramachandran plot appearance (blue, # 0.37, § 0.43), χ-1/χ-2 rotamer normality (green, # 0.40, § 0.43) and backbone normality (magenta, # 0.45, § 0.50). (e) Number of unsatisfied buried hydrogen bond acceptors (black, # 0.39, § 0.40), hydrogen bond donors (red, # 1.21, § 1.36) and inter-atomic bumps per 100 residues (blue, # 1.75, § 1.93). Note that the first cycle is identical for the two types of structure selection in the re-sampling procedure.

Table 2. Results of re-sampling for Crambin and PAH2-Mad1

	Crambin						PAH2-Mad1					
	Input (0)[c]		Low energy[a] (11)[c]		High RMSD[b] (14)[c]		Input (0)[c]		Low energy[a] (14)[c]		High RMSD[b] (21)[c]	
Rmsd backbone atoms	0.51	± 0.11	0.74	± 0.24	0.94	± 0.30	0.78	± 0.12	1.31	± 0.34	1.96	± 0.61
Rmsd heavy atoms	0.74	± 0.12	0.87	± 0.22	1.10	± 0.29	1.32	± 0.11	1.63	± 0.36	2.49	± 0.63
Rmsd dihedral angle restraints	–		–		–		0.4	± 0.2	0.5	± 0.1	0.6	± 0.2
Rmsd distance restraints	0.013	± 0.007	0.0082	± 0.0004	0.0091	± 0.0005	0.019	± 0.002	0.0117	± 0.0006	0.014	± 0.001
Distance restraint violations												
>0.1 Å	5.8	± 4.8	1.1	± 0.2	1.3	± 0.6	24.5	± 5.2	7.7	± 2.3	11.4	± 2.9
>0.2 Å	1.3	± 1.7	0.5	± 0.5	0.7	± 0.5	2.8	± 1.3	0.1	± 0.2	1.0	± 0.9
>0.3 Å	0.9	± 1.5	0		0		0.4	± 0.6	0		0.1	± 0.3
>0.4 Å	0.4	± 0.8	0		0		0		0		0	
Structure Z-scores:												
1st generation packing quality	−0.8	± 0.3	0.1	± 0.1	0.0	± 0.2	−1.6	± 0.2	−1.4	± 0.2	−1.8	± 0.3
2nd generation packing quality	0.1	± 0.4	0.4	± 0.2	0.3	± 0.3	−0.4	± 0.4	−0.7	± 0.4	−1.3	± 0.6
Ramachandran plot appearance	−1.0	± 0.7	0.3	± 0.4	0.3	± 0.5	−1.4	± 0.6	−0.2	± 0.4	−0.9	± 0.7
χ-1/χ-2 rotamer normality	0.2	± 0.5	−0.8	± 0.3	−0.7	± 0.3	0.0	± 0.6	−0.8	± 0.3	−1.2	± 0.4
Backbone conformation	1.0	± 0.3	0.4	± 0.4	0.4	± 0.7	−0.5	± 0.7	−1.4	± 0.7	−3.0	± 1.1
Inter-atomic bumps:												
No. bumps	1.1	± 0.9	1.2	± 0.6	1.6	± 0.8	1.2	± 1.1	3.9	± 1.5	4.3	± 1.8
No. bumps/100 res.	2.3	± 1.9	2.6	± 1.3	3.4	± 1.8	1.2	± 1.2	4.0	± 1.6	4.4	± 1.9
Sum of bumps	0.1	± 0.1	0.0	± 0.0	0.1	± 0.1	0.1	± 0.1	0.2	± 0.1	0.2	± 0.1
Unsatisfied H-bond donors	2.7	± 1.0	2.2	± 0.8	2.4	± 0.8	5.8	± 2.1	6.6	± 2.4	9.5	± 2.7
Unsatisfied H-bond acceptors	0		0		0		0.2	± 0.4	0.2	± 0.4	0.2	± 0.4

Validation results for two types of re-sampling procedures of Crambin and PAH2-Mad1. Shown are the values for the input structures and the structures in the last re-sampling cycle for each type of re-sampling:
[a]Low energy structure selection.
[b]High RMSD structure selection (see text for details).
[c]Numbers in parenthesis indicate the re-sampling cycle used for analysis.
All RMSDs are given in Ångstroms, except for dihedral angle restraints where the RMSD is in degrees.

ber of bumps increases during re-sampling. A slight increase is also seen for the number of unsatisfied hydrogen bond donors. It can be argued that these changes indicate deterioration of the quality of the structure ensembles during re-sampling. However, it should be mentioned that the values found for the number of bumps per 100 residues and the average size of the bumps are much lower than what is typically found in NMR-structure ensembles and even in the high-resolution X-ray structures of the WHAT IF database (Table 1 and data not shown).

In Table 2 the results of the two different types of re-sampling for the Crambin and PAH2-Mad1 data sets are listed. Table 3 shows the effect of water-refinement and 1 type of re-sampling on 1GB1 and 1CHL. The trends for all four data sets are similar to those of Ubiquitin, although we did not observe the simultaneous increase of backbone RMSD and decrease of heavy atom RMSD. The ratio between the heavy atom and backbone atom RMSD decreases in all data sets except for 1CHL, where the ratio actually increased during re-sampling. Often the backbone omega angle is too tightly restrained in NMR structure ensembles (Doreleijers et al., 1999; Spronk et al., 2002), which will also lead to tighter clustering of the backbone of the ensemble than is actually justified by the experimental and database data (MacArthur and Thornton, 1996; Wilson et al., 1998). Even though this is not a problem in the input structures used in this study, in which the omega angles were properly restrained (RMS Z-score ∼1), our results indicate that the variability of the protein backbone is commonly underestimated.

It is clear that even for the Crambin data set, which contained all proton-proton distances smaller than 5 Å, there is room for a moderate increase of the RMSD

Table 3. Results of re-sampling for 1GB1 and 1CHL

	1GB1						1CHL					
	Input (0)[c]		Refined[a] (0)[c]		High RMSD[b] (14)[c]		Input (0)[c]		Refined[a] (0)[c]		High RMSD[b] (17)[c]	
Rmsd backbone atoms	0.38	± 0.05	0.57	± 0.11	0.89	± 0.25	0.97	± 0.13	0.91	± 0.13	1.14	± 0.41
Rmsd heavy atoms	0.91	± 0.09	1.06	± 0.10	1.36	± 0.26	1.66	± 0.23	1.62	± 0.26	2.28	± 0.89
Rmsd dihedral angle restraints	0.14	± 0.03	0.29	± 0.08	0.40	± 0.09	0.4	± 0.3	0.01	± 0.03	0.02	± 0.06
Rmsd distance restraints	0.094	± 0.007	0.0155	± 0.0008	0.018	± 0.002	0.033	± 0.002	0.030	± 0.004	0.035	± 0.006
Distance restraint violations												
>0.1 Å	19.4	± 1.7	6.3	± 1.8	9.1	± 2.8	4.7	± 1.4	5.6	± 3.3	5.8	± 2.0
>0.2 Å	6.4	± 1.3	0.0	± 0.1	0.2	± 0.5	1.3	± 0.5	0		0.9	± 1.1
>0.3 Å	2.9	± 1.1	0		0		0		0		0	
>0.4 Å	2.7	± 1.0	0		0		0		0		0	
Structure Z-scores:												
1st generation packing quality	−1.2	± 0.2	1.2	± 0.2	0.8	± 0.2	−6.8	± 0.3	−4.2	± 0.6	−4.7	± 0.3
2nd generation packing quality	−2.3	± 0.4	0.9	± 0.4	0.6	± 0.4	−4.3	± 0.6	−2.5	± 0.3	−3.1	± 0.5
Ramachandran plot appearance	−3.9	± 0.5	−0.9	± 0.4	−0.9	± 0.5	−6.7	± 0.6	−3.5	± 0.9	−4.9	± 1.0
χ-1/χ-2 rotamer normality	−3.6	± 0.4	−0.6	± 0.6	−1.3	± 0.5	−4.4	± 0.5	−1.5	± 1.4	−2.0	± 0.4
Backbone conformation	0.7	± 0.4	−0.1	± 0.6	−0.5	± 0.5	−7.5	± 2.0	−6.1	± 0.7	−7.6	± 1.1
Inter-atomic bumps:												
No. bumps	26.4	± 3.0	2.3	± 1.2	3.4	± 1.5	29.9	± 5.7	2.7	± 1.5	4.1	± 0.9
No. bumps/100 res.	47.1	± 5.3	4.1	± 2.1	6.1	± 2.7	82.9	± 15.8	7.5	± 4.2	11.5	± 2.5
Sum of bumps	3.0	± 0.4	0.1	± 0.0	0.2	± 0.1	4.2	± 0.8	0.2	± 0.1	0.3	± 0.1
Unsatisfied H-bond donors	3.0	± 1.0	2.1	± 1.0	2.0	± 1.1	4.9	± 1.7	2.0	± 0.8	4.7	± 1.5
Unsatisfied H-bond acceptors	0		0		0		0		0		0.1	± 0.4

Validation results for re-sampling procedures of 1GB1 and 1CHL using the high RMSD structure selection scheme. Shown are the values for the original structures, the original structures after water-refinement, and the structures in the last re-sampling cycle:
[a]Input structures after water-refinement in ARIA.
[b]Structures after re-sampling.
[c]Numbers in parenthesis indicate the re-sampling cycle used for analysis.
All RMSDs are given in Ångstroms, except for dihedral angle restraints where the RMSD is given in degrees.

without significantly compromising the quality of the structures. On the other hand, for the PAH2-Mad1 data set, we can achieve a backbone and heavy atom RMSD of up to 2 and 2.5 Å, respectively, after 21 cycles of re-sampling. This is more than a factor of 2 increase for the backbone RMSD while the quality indicators stay within normal ranges compared to those from the X-ray structures. It should be mentioned though that there is a clear trend towards more negative values for all Z-scores, indicating some deterioration in the quality of the structures. The results obtained for 1GB1 and 1CHL show strong and moderate increases of the RMSD, respectively, and confirm the results obtained for the other data sets. In addition, we also examined the changes in the RMSD of the disordered regions in the ensembles superimposed only on the well-defined regions (see data sets). As expected we also find small to large increases in heavy atom RMSD in these relatively disordered regions: 1CHL (residues 1, 5–12 and 36): 5.0 Å to 5.2 Å, Ubiquitin (residues 71–76): 5.7 Å to 7.1 Å and PAH2-Mad1 (residues 1–4, 42–54 and 81–84): 7.3 Å to 10.8 Å.

The results as discussed above, raise a number of important questions. The first relates to how well the procedure samples the conformational space. The efficiency and maximum RMSD that can be reached is clearly dependent on the different parameters in the re-sampling protocol. Further optimization of the different parameters, especially in the refinement part, may lead to even higher RMSD values. This is certainly expected when the more realistic treatment of NMR restraints using time and/or ensemble averaging methods is included.

In this study we have applied generally used acceptance criteria for structure selection. It must be noted that often more stringent criteria are used for structure

selection and the result of the re-sampling depends on these criteria. In case the acceptance criteria are much looser than the actual violations in the input structures the re-sampling procedure will relax the ensemble slightly towards the acceptance criteria (e.g., see Tables 2 and 3). This may occur when the input structures consist of a subset of low-energy structures from a larger ensemble of accepted structures. It is important to realize that the method thus only provides a way to address the conformational variability in an ensemble given the experimental restraints and the acceptance criteria as defined by the researcher.

Another important and yet unresolved matter is that at present it is unclear how to address the quality of NMR structure ensembles in an objective manner. Thus far, validation of NMR structures using programs such as PROCHECK (Laskowski et al., 1993), PROCHECK_NMR (Laskowski et al., 1996) and WHAT IF (Vriend, 1990) has relied on a comparison to a database of high-resolution X-ray structures. Currently there is no better alternative, since the most reliable data on biomolecular structure have been obtained by X-ray crystallography. However, it can be argued that due to the intrinsically higher mobility in solution structures a different comparison set should be used for NMR ensembles. Preferably these comparisons should be made at different temperatures, since thermal motion influences the distributions of, for example, dihedral angles and a larger variation around the energetic minima will be found. At present, building of such reference data bases cannot be reliably done based on existing experimental and structural data, which are often incomplete, contain no information on manually removed inconsistencies and often contain many abnormalities that undoubtedly are errors.

Conclusions and suggestions

The work described in this paper constitutes an essential step towards the definition of the NMR equivalent of the crystallographic B-factor. This NMR B-factor, which should reflect all uncertainties in the data arising from dynamic behavior, experimental procedures and calculation protocols, is of great importance for a good assessment of the quality of structure ensembles. In a first approach, we have focused on finding the limits of the uncertainty in structure coordinates by re-sampling the allowed conformational space under restriction of NMR-restraints. In this approach we chose to use distance- and dihedral angle restraints only, without including other experimentally derived restraints, nor using relaxation data for validation of the variability in the structure ensembles. Further, it should be mentioned that the method does not judge the quality of the input NMR restraints. It only provides a means to sample the conformational space within the limits of the NMR restraints regardless of their quality.

The re-sampling method leads to a systematic increase of the RMSD of protein NMR structure ensembles without significantly compromising the quality of the structures and their fit to experimental input data. The precision, and thus the accuracy, of the input structures is heavily overestimated within the tolerances for violation of experimental data.

Taking into account that restraints are usually not treated as time- or ensemble averages, the RMSDs obtained here should be regarded as a *lower* limit of the uncertainty of the data sets. Furthermore, possible limitations of the method described here will also influence the maximum attainable RMSD for a structure ensemble. Expansion of the technique to include other conformational restraints, such as those derived from residual dipolar couplings is a straightforward extension of the current protocol. Inclusion of or validation against relaxation data is however not a trivial task due to the model-dependency of relaxation data analysis, and remains a challenge for the near future.

To conclude, we would like to stress here that depositors of structures should be aware of the true value of coordinate RMSDs, its relation to the accuracy, and the factors that are of influence (see Zhao and Jardetzky, 1994). We therefore propose that an analysis as described in this paper is used for estimating the precision of NMR structure ensembles, prior to publication or submission of structures to databases. Alongside the deposition of minimized average structures or ensembles representing the global minima using conventional methods, it will be very informative to users of structures to have a separate parameter describing the uncertainty in the structure coordinates after re-sampling. Further, it is of great importance that structures are optimized using sophisticated refinement protocols to remove some of the obvious problems in NMR-structures, such as the relatively high occurrences of inter-atomic bumps and unrealistic charge-distributions (see also Spronk et al., 2002). Such methods are easy to perform and it is important to realize that it is not enough to merely optimize the fit of the structures to the available experimental data.

Inspection of validation reports as obtained from e.g. WHAT IF or PROCHECK_NMR, provides an important means to improve the quality, reliability and thus the usefulness of the deposited NMR structures. Finally, we would like to urge depositors of structures to make all NMR-derived restraints available alongside the coordinates in order to perform validation analyses, including a detailed description of the data handling during structure calculation.

Acknowledgements

The work described in this paper is part of a collaborative effort of a EU network of laboratories supported by the EC 5th Framework program (FP5), contract QLG2-CT-2000-01313 titled 'NMRQUAL. NMR Structure validation: Integration into the structure determination process'. The authors wish to thank all parties involved in NMRQUAL for stimulating discussions. Further, we would like to thank Ad Bax, Cees Hilbers, Michael Nilges and Jurgen Doreleijers for critical reading of the manuscript and Bert de Groot for providing the program CONCOORD and stimulating discussions. The re-sampling method is implemented as a Python-script and is freely available upon request (mail to C.Spronk@cmbi.kun.nl). C. Spronk is financially supported by the Netherlands Organisation for Scientific Research (NWO).

References

Allen, F.H., Kennard, O. and Taylor, R. (1983) *Acc. Chem. Res.*, **16**, 146–153.

Brünger, A.T., Adams, P.D., Clore, G.M., Delano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) *Acta Cryst.*, **D54**, 905–921.

Brünger, A.T., Clore, G.M., Gronenborn, A.M., Saffrich, R. and Nilges, M. (1993) *Science*, **261**, 328–331.

Chalaoux, F.R., O'Donoghue, S.I. and Nilges, M. (1999) *Proteins*, **34**, 453–463.

Clore, G.M., Robien, M.A. and Gronenborn, A.M. (1993) *J. Mol. Biol.*, **231**, 82–102.

Cornilescu, G., Marquardt, J., Ottiger, M. and Bax, A. (1998) *J. Am. Chem. Soc.*, **120**, 6836–6837.

de Groot, B.L., van Aalten, D.M., Scheek, R.M., Amadei, A., Vriend, G. and Berendsen, H.J. (1997) *Proteins*, **29**, 240–251.

Doreleijers, J.F., Vriend, G., Raves, M.L. and Kaptein, R. (1999) *Proteins*, **37**, 404–416.

Engh, R.A. and Huber, R. (1991) *Acta Cryst.*, **A47**, 392–400.

Gronenborn, A.M. and Clore, G.M. (1995) *Crit. Rev. Biochem. Mol. Biol.*, **30**, 351–385.

Gronenborn, A.M., Filpula, D.R., Essig, N.Z., Achari, A., Whitlow, M., Wingfield, P.T. and Clore, G.M. (1991) *Science*, **253**, 657–661.

Havel, T.F. and Wüthrich, K. (1985) *J. Mol. Biol.*, **182**, 281–294.

Hooft, R.W., Sander, C. and Vriend, G. (1996) *J. Appl. Cryst.*, **29**, 714–716.

Hooft, R.W., Sander, C. and Vriend, G. (1997) *Comput. Appl. Biosci.*, **13**, 425–430.

Horstink, L.L.M., Nilges, M. and Hilbers, C.W. 2000. *Internal Dynamics in Single-Stranded DNA Binding Proteins*, Ph.D. Thesis, University of Nijmegen, The Netherlands, pp. 67–80.

Jelsch, C., Teeter, M.M., Lamzin, V., Pichon-Pesme, V., Blessing, R.H. and Lecomte, C. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 3171–3176.

Jorgensen, W.L. and Tirado-Rives, J. (1988) *J. Am. Chem. Soc.*, **110**, 1657–1666.

Jorgensen, W.L., Chandrasekhan, J., Madura, J., Impley, R.W. and Klein, M.L. (1983) *J. Chem. Phys.*, **79**, 926–935.

Kabsch, W. and Sander, C. (1983) *Biopolymers*, **22**, 2577–2637.

Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph.*, **14**, 51–55.

Kuszewski, J., Gronenborn, A.M. and Clore, G.M. (1999) *J. Am. Chem. Soc.*, **121**, 2337–2338.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) *J. Appl. Cryst.*, **26**, 283–291.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) *J. Biomol. NMR*, **8**, 477–486.

Linge, J.P. and Nilges, M. (1999) *J. Biomol. NMR*, **13**, 51–59.

Linge, J.P., O'Donoghue, S.I. and Nilges, M. (2001) *Meth. Enzymol.*, **339**, 71–90.

Linge, J.P., Williams, M.A., Spronk, C.A.E.M., Bonvin, A.M.J.J. and Nilges, M. (2003) *Proteins*, **50**, 496–506.

Lippens, G., Najib, J., Wodak, S.J. and Tartar, A. (1995) *Biochemistry*, **34**, 13–21.

MacArthur, M.W. and Thornton, J.M. (1996) *J. Mol. Biol.*, **264**, 1180–1195.

Nilges, M., Macias, M.J., O'Donoghue, S.I. and Oschkinat, H. (1997) *J. Mol. Biol.*, **269**, 408–422.

Pfeiffer, S., Karimi-Nejad, Y. and Rüterjans, H. (1997) *J. Mol. Biol.*, **266**, 400–423.

Scheek, M., van Nuland, N.A.J., de Groot, B.L. and Amadei, A. (1995) *J. Biomol. NMR*, **6**, 106–111.

Sprangers, R., Bottomley, M.J., Linge, J.P., Schultz, J., Nilges, M. and Sattler, M. (2000) *J. Biomol. NMR*, **16**, 47–58.

Spronk, C.A.E.M., Linge, J.P., Hilbers, C.W. and Vuister, G.W. (2002) *J. Biomol. NMR*, **22**, 281–289.

Spronk, C.A.E.M., Tessari, M., Kaan, A.M., Jansen, J.F., Vermeulen, M., Stunnenberg, H.G. and Vuister, G.W. (2000) *Nat. Struct. Biol.*, **7**, 1100–1104.

Tjandra, N., Marquardt, J. and Clore, G.M. (2000) *J. Magn. Reson.*, **142**, 393–396.

Torda, A.E., Scheek, R.M. and van Gunsteren, W.F. (1990) *J. Mol. Biol.*, **214**, 223–235.

Vriend, G. (1990) *J. Mol. Graph.*, **8**, 52–56.

Wilson, K.S., Butterworth, S., Dauter, Z., Lamzin, V.S., Walsh, M., Wodak, S., Pontius, J., Richelle, J., Vaguine, A., Sander, C., Hooft, R.W.W., Vriend, G., Thornton, J.M., Laskowski, R.A., MacArthur, M.W., Dodson, E.J., Murshudov, G., Oldfield, T.J., Kaptein, R. and Rullmann, J.A.C. (1998) *J. Mol. Biol.*, **276**, 417–436.

Zhao, D. and Jardetzky, O. (1994) *J. Mol. Biol.*, **239**, 601–607.