COMMUNICATION

Application of $^1$H NMR Chemical Shifts to Measure the Quality of Protein Structures

Michael P. Williamson$^1$, Jun Kikuchi$^2$ and Tetsuo Asakura$^2$

We have developed a program that can calculate proton NMR chemical shifts for proteins, using a set of co-ordinates provided for example from an X-ray or NMR structure. When applied to NMR structures, agreement between calculated and observed shifts is generally of the same quality as that for crystal structures of resolution between 2.0 and 3.0 Å. There is a rather weak correlation between standard deviation (SD) and the number of NMR constraints per residue, but none with the root-mean-square deviation of one NMR structure from another. Where minimised averaged structures are present, they have about the same SD as the population from which they were taken. Refinement methods such as energy minimisation and the use of relaxation matrices and back calculation produce little or no improvement in SD. The calculation has several applications, particularly as an independent means of measuring the quality of a structure (either in the crystal or in solution), and in identifying possible assignment errors.

Keywords: NMR; chemical shift; protein structure; precision; accuracy

Proton chemical shifts are probably the most easily measured, and certainly the most precise, NMR parameter that can be obtained for proteins. They have long been recognised to carry conformational information, for example on secondary structure (Wishart et al., 1991). There have recently been several attempts to develop methods for calculating chemical shifts (for a review, see Szilágyi, 1995), stimulated by the large number of proteins with essentially complete $^1$H NMR assignments. Here, we describe how the comparison of experimentally observed chemical shifts with those calculated from a set of co-ordinates can be used to provide novel information about the quality (i.e. the precision and accuracy) of the co-ordinates.

Our first aim was to study the relationship between chemical shifts and protein structure quality for structures whose quality is known, i.e. for crystal structures. We have used two measures of the quality of a crystal structure: its resolution, and its stereochemical quality. Although in comparisons of crystal structures with NMR data, $R$-factors were found to be important additional indicators of goodness of fit (Bartik et al., 1993), we have chosen simply to use resolution as a readily quantifiable measure. For stereochemical quality we used the program PROCHECK (Laskowski et al., 1993).

Figure 1 shows the results for a selection of crystal structures, which show a good overall correlation of standard deviation of chemical shifts with resolution; similar results were obtained for stereochemical quality (data not shown). These correlations are pleasing, considering the uncertainties surrounding the calculation, which include positional errors in the crystal structure, the inherent difference between crystal and solution structures, and deficiencies in the calculation.

Several structures in Figure 1 are identified by name because they have either much better or much poorer values for standard deviation (SD) than expected from the “quality” of the crystal structure. The structures with better SD values than expected are 1ROP, 1RCB and 2INT. These are, respectively, the rop DNA-binding domain (Banner et al., 1987) and two structures of interleukin 4 (Wlodawer et al., 1992; Walter et al., 1992). They have in common that they are almost entirely $\alpha$-helical. The smaller chemical shift dispersion of $\alpha$-helical proteins (Williamson, 1990) tends to lead to a lower SD. The proteins with larger SD values than expected are more interesting. Two structures (1LYZ and 1LYM) are of hen lysozyme (Diamond, 1974; Rao et al., 1983). These are both very early structures, and 1LYZ has had no refinement at
all. 1LYM is the monoclinic crystal form, which has a very high protein density and so is particularly prone to crystal packing distortions. It is likely that most of the problems with both structures stem from their lack of suitable refinement; in other words, they are not really as good as implied by the quality measures used here.

The reason for the poor SD of 1PGB is different. This is a structure of protein G domain B1 (Gallagher et al., 1994), for which there are several published assignments. Using the assignment list from the Brookhaven Protein Data Bank (PDB) entry 1GB1 (Gronenborn et al., 1991), one residue (Thr18) stands out as having a very large difference between calculated and observed shifts. On comparison of the different assignment tables for domain B1 and the similar domain B2 (Lian et al., 1991, 1992; Gronenborn et al., 1991; Orban et al., 1992), there was found to be a discrepancy of over 1 p.p.m. at this position. The erroneous shift at this position was a typographical error in the assignment table and has since been corrected (A. M. Gronenborn, personal communication). If the assignments of Lian et al. (1991) are used, another discrepancy appears, at residue Glu32 (equivalent to Glu27 in the numbering of 1PGB). This was an assignment error and has also been corrected (Lian et al., 1992). It should be noted that this study was not conducted with the aim of searching for such errors. Rather, the errors forced themselves upon our attention during the course of the work. It is therefore entirely possible that there may be other assignment or typographical errors still present in the literature. This work implies that chemical shift calculations can be used to identify likely assignment errors, by looking at the difference between calculated and observed shift for each residue.

We next applied the calculations to NMR structures, as a way of estimating how close NMR structures are to the "true" solution structure. The aim of this section of the work has been to use chemical shifts as an independent measure of the quality of a structure, looking at various aspects of NMR solution structures. The value of SD calculated for an NMR structure is a measure of how close the structure is to the true time-averaged solution structure. Therefore, both the average SD for a family of structures and the SD for the minimised average of a family of structures reflects accuracy much more than precision, since it is a value obtained for a

---

**Figure 1.** The dependence of the standard deviation of calculated to observed C'H shift on resolution for 72 different crystal structures. The line is a best-fit line. A complete list of the proteins used and a table of results can be obtained from the authors. Structures were taken from the Brookhaven Protein Data Bank (Abola et al., 1987). The equations and parameters for chemical shift calculation have been described (Williamson & Asakura, 1993), and a Fortran program is available from M. Williamson@Sheffield.ac.uk, or via anonymous ftp from directory pub/uni/academic/I-M/mbb in ftp.shef.ac.uk. Chemical shift differences from random coil are calculated as a sum of ring current shifts (calculated using the Haigh-Mallick equation), bond magnetic anisotropies (calculated using C = O and C'-N bonds separately) and electric field effects.
Communication

543

structure that is close to the average of the family. By contrast, the SD obtained for a single member of the family is a function both of accuracy and precision, since we have previously shown (Williamson et al., 1992) that a 0.15 Å random variation in atomic positions can increase SD by at least 0.24 p.p.m., implying that precision has a large effect on SD. We have compared the SD of chemical shifts to the two most commonly used measures of the quality of an NMR structure, namely the number of constraints per residue and the root-mean-square deviation (r.m.s.d.) between structures.

Figure 2 presents a comparison of the SD of chemical shifts with the measures of quality described above, applied to the average value from families of structures. It is apparent that there is a correlation, albeit a weak one. The correlation between SD and the number of constraints per residue is better than that between SD and r.m.s.d., suggesting that the number of constraints per residue may be a better guide than r.m.s.d. to the real quality (i.e. accuracy) of a structure. It is worth pointing out, as shown in Figure 2(c), that the correlation between r.m.s.d. and number of constraints per residue is not particularly good either. It is clear that other guides to the quality of an NMR structure are required in addition to these standard measures, one of which can be provided by chemical shifts. This measure has the advantage that it uses a parameter that (to date) has not been used at all in the structure calculation, and therefore unlike stereochemical quality is not correlated to the type of constraints used (Morris et al., 1992).

Figure 2(a) shows that the better NMR structures have an SD of around 0.3 p.p.m. By comparison with Figure 1, it is concluded that this is equivalent to the SD obtained from a crystal structure with a resolution of about 2.0 Å. However, most structures are worse than this, with many structures having SD values equivalent to structures of 3 Å or worse. This is broadly in line with the conclusions of Clore et al. (1993) and MacArthur & Thornton (1993), who concluded that the quality of the core parts of NMR structures was in the range of that of a 2.0 to 2.3 Å crystal structure.

NMR structures are generally calculated and presented as a family of structures. A single representative structure is often derived from the family by calculating a mean structure and energy minimising it. There has been some discussion of how representative such a structure is of the family as a whole, from which the conclusion was that it is better to study the ensemble as a whole (Sutcliffe, 1993). We have compared the SD for the minimised average structure to that for the ensemble (Table I). The minimised average structure generally has a SD close to the mean of the ensemble. These chemical shift results therefore imply that the minimised average structure is probably quite a good representation of the family as a whole, if such a representative is required. They also support the contention (Zhao & Jardetzky, 1994) that precision is not related to accuracy.

We have also carried out chemical shift calculations on structures refined in different ways. The structure 1BUS, which was the first protein NMR structure
Table 1

Standard deviation of observed and calculated chemical shift of NMR structures for families and for minimised average structures

<table>
<thead>
<tr>
<th>Protein</th>
<th>Protein PDB code</th>
<th>SD (p.p.m.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(family/min. avg. struct.*)</td>
<td>Family</td>
</tr>
<tr>
<td>Tendamistat</td>
<td>2AIT/3AIT</td>
<td>0.402 ± 0.040</td>
</tr>
<tr>
<td>Tendamistat</td>
<td>2AIT/4AIT</td>
<td>0.402 ± 0.040</td>
</tr>
<tr>
<td>Interleukin-1β</td>
<td>7IB/6IB</td>
<td>0.342 ± 0.008</td>
</tr>
<tr>
<td>Interleukin-4</td>
<td>1BCN/1BNN</td>
<td>0.357 ± 0.026</td>
</tr>
<tr>
<td>Interleukin-4</td>
<td>1ITI</td>
<td>0.268 ± 0.008</td>
</tr>
<tr>
<td>Interleukin-8</td>
<td>2I18/I118</td>
<td>0.369 ± 0.013</td>
</tr>
<tr>
<td>Protein G (B1)</td>
<td>1GB1/2GB1</td>
<td>0.512 ± 0.009</td>
</tr>
<tr>
<td>Hirudin</td>
<td>2HIR/5HIR</td>
<td>0.383 ± 0.050</td>
</tr>
<tr>
<td>Neurotoxin-I</td>
<td>2SH1/1SH1</td>
<td>0.370 ± 0.047</td>
</tr>
<tr>
<td>PI3K</td>
<td>2PN1/1PNJ</td>
<td>0.416 ± 0.028</td>
</tr>
<tr>
<td>Succ. transferase</td>
<td>1BAL/1BBL</td>
<td>0.231 ± 0.023</td>
</tr>
<tr>
<td>Pyr. dehydrogenase</td>
<td>1LAB/1LAC</td>
<td>0.441 ± 0.027</td>
</tr>
<tr>
<td>Cellobiohydrolase</td>
<td>2CBH/1CBH</td>
<td>0.308 ± 0.029</td>
</tr>
<tr>
<td>hTGFα</td>
<td>3TG1/2TG1</td>
<td>0.429 ± 0.035</td>
</tr>
</tbody>
</table>

*a min. avg. struct., minimised average structure.

b The chemical shift of Thr18 C=H has been corrected (see the text).

(Williamson et al., 1985), received no refinement after distance geometry, and has a very high SD of 0.86 p.p.m. This result is consistent with the general assumption that distance geometry alone is not sufficient to generate accurate structures (Liu et al., 1992), and that further refinement, for example by simulated annealing, is needed.

The structure of tendamistat has been refined by taking one structure out of the family and energy minimising it using either the program AMBER or FANTOM (Billeter et al., 1990). The starting structure had a SD of 0.377 p.p.m., and the AMBER and FANTOM refinements have SD values of 0.357 and 0.431 p.p.m., respectively. Thus, the AMBER

Figure 3. Comparison of the SD of 16 structure models of murine EGF before (open squares) and after (filled squares) energy minimisation. The mean SD values are 0.345 (±0.078) and 0.353 (±0.062) p.p.m., and the structures are 1EGF and 3EGF, respectively (Montelione et al., 1988, 1992).
refinement has produced some improvement in chemical shifts, but the FANTOM refinement has produced a very significant worsening. Plots of shift difference against residue number show that most of this worsening is due to one residue, Trp18. The CH of Trp18 is very close to the ring of Tyr60, and it appears that the FANTOM refinement has moved Tyr60 away from Trp18, in a way which is inconsistent with the chemical shifts. Thus, chemical shifts can be used to limit the conformational range of calculated structures (Kikuchi et al., 1994).

Montelione et al. (1992) have reported a structure for murine epidermal growth factor (EGF) determined using the distance geometry program DIANA with and without minimisation using the ECEPP force-field (3EGF and 1EGF, respectively). The results (Figure 3) show that the minimisation has produced no significant improvement in chemical shifts; in fact, the SD has become marginally worse, from 0.345 (+0.078) to 0.353 (+0.062) p.p.m. A similar refinement for human transforming growth factor α (Moy et al., 1993) produced a slight improvement, from 0.409 (+0.081) to 0.377 (+0.043) p.p.m. Similarly, the structure of α-conotoxin GVIA has been refined using the BKCALC program (Hare Research, Inc.: Pallaghy et al., 1993), which leads to a small but not statistically significant improvement in SD (0.352 (+0.050) to 0.329 (+0.032) p.p.m.).

Most refinements of the type described above aim to reduce nuclear Overhauser effect (NOE) violations. However, Smith et al. (1993) have shown that when a crystal structure of lysozyme was energy minimised using NOE constraints, although the NOE violations decreased substantially, the discrepancy between calculated and observed values for both coupling constants and chemical shifts actually increased. This striking result can be interpreted in several ways, one of the most convincing being that the NOE constraints used were inaccurate, because of motional averaging. Such a situation is expected to be quite general. Thus, a reduction of violations of NOE constraints as a result of refinement merely demonstrates the success of the refinement algorithm, and does not necessarily prove that the resultant structure is better (i.e. more accurate). The best way to demonstrate that the quality of the structure has improved is to use an independent measure of structure quality, such as chemical shift.

Acknowledgements

We thank Jeremy Craven and Pete Artymiuk for help with programming; Iain Campbell, Christina Redfield and Guy Montellione for supplying data; László Szilágyi for providing a preprint; and the Japan Society for the Promotion of Science for financial support.

References


Edited by P. E. Wright

(Received 29 November 1994; accepted 30 January 1995)