Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A systematic case study on using NMR models for molecular replacement: p53 tetramerization domain revisited

Yu Wai Chen[a]* and G. Marius Clore[b]

[a]Centre for Protein Engineering and Cambridge University Chemical Laboratory, Medical Research Council Centre, Hills Road, Cambridge CB2 2QH, England, and [b]Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Health, Bethesa, Maryland 20892-0510, USA

Correspondence e-mail: ywc@mrc-lmb.cam.ac.uk

Received 22 May 2000
Accepted 31 August 2000

Molecular replacement using search models derived from nuclear magnetic resonance (NMR) spectroscopy has often proved problematic. It has been known for some time that the overall differences in atomic positions (r.m.s.d.) between the crystalline and the solution states of the same protein are of the order of 1–2 Å and approach the limit of molecular replacement. In most cases, this structural difference is a result of calculating the NMR structure with insufficient data, yielding an NMR structure of limited accuracy. A systematic case study was performed to investigate the use of NMR models for molecular replacement on the p53 tetramerization domain: NMR search models of varying degrees of accuracy were employed to solve phases for the 1.5 Å X-ray diffraction data. An approximate correlation was found between the accuracy of the NMR search model and the clarity and quality of the molecular-replacement solution. It was found that ensemble models perform better than single averaged models and have a larger tolerance in model inaccuracy. Also, distance-derived *B* factors can improve the performance of single models.

1. Introduction

Molecular replacement (MR) is a very powerful method for solving the phase problem in X-ray crystallography (Rossmann, 1972). Numerous examples in the literature reveal, not surprisingly, that the higher the sequence homology between the search model and the macromolecule to be solved, the higher the likelihood of success. However, molecular replacement using search models determined by nuclear magnetic resonance spectroscopy (NMR) has often proved to be difficult, even if the search model has 100% sequence identity with (*i.e.* is the same protein as) the structure to be solved. In the past, this puzzling phenomenon has generally been attributed to apparent structural differences between the solution and crystalline states.

In recent years, a number of new methods have been introduced into NMR structure calculations (reviewed in Clore & Gronenborn, 1998). NMR structures solved recently which make use of these restraints are more accurate (see, for example, Bewley *et al.*, 1998) and lead to a smooth solution of the corresponding crystal structure by molecular replacement (Yang *et al.*, 1999). Kuszewski *et al.* (1999) pointed out that the widely observed looser packing of NMR structures compared with crystal structures does not represent a real difference between the two states, but rather a deficiency in the nature of the NMR data. To address this problem, a pseudo-potential for the radius of gyration was incorporated into the NMR

© 2000 International Union of Crystallography
Printed in Denmark – all rights reserved

structure calculation and this led to substantially improved agreement with the crystal structure.

It follows that the difficulty in solving molecular-replacement problems from NMR models may be the consequence of insufficient accuracy of the NMR model. In this work, we studied how the accuracy of an NMR model affects the success of molecular replacement using that search model. We performed a systematic study with a well documented case, namely the tetramerization domain of p53 tumour suppressor. This is a good system for study for two reasons. (i) There are seven structures independently determined by four groups (Table 1). Among these, four NMR structures are available sampling a wide range (backbone r.m.s.d. of 0.4–1.9 Å) of model accuracy. (ii) Two crystal structures were solved by molecular replacement with different NMR search models. It is interesting to find out if difficulties in molecular replacement are related to the accuracy of the search model.

The highest resolution crystal structure, X6, was solved by molecular replacement with great difficulty using N2 as the search model. According to the authors, it would have been impossible to solve this structure without knowledge of the 222 symmetry of the tetrameric molecule (Mittl *et al.*, 1998). The correct molecular-replacement solution had to be identified by screening conditions consistent with the known internal symmetry of the tetramer. In addition, they had to use a tetrameric ensemble search model to find the monomer in the asymmetric unit (tetragonal space group). X5, on the other hand, was solved rather smoothly using the better refined N4 as the search model. The whole tetramer in the asymmetric unit of the trigonal space group was found readily using a tetrameric search model (Miller *et al.*, 1996). X5 and X6, however, are of different crystal forms and space groups (Table 1) and the difficulties encountered in each case cannot be compared.

Armed with this library of structures (Table 1) all of the same protein with slightly different numbers of residues at the N and C termini, we have now re-examined this case using the 1.5 Å resolution X-ray data set (Mittl *et al.*, 1998) and attempted molecular replacement with the four NMR structures (N1, N2, N4 and N7) of varying degrees of accuracy. For each NMR structure, we studied the efficiency of using a single averaged model *versus* an ensemble model and also the usefulness of assigning artificial B factors. We have used a molecular-replacement protocol employing typical values of search parameters. It was our intention to keep these search parameters unchanged and we did not go through the exhaustive exercise of parameter fine-tuning. The standard protocol therefore served as a screen for the usefulness of various models. Of the 12 molecular-replacement calculations we performed, nine led to a correct solution (Table 2 and Fig. 1).

2. Materials and methods

The coordinates for all structure models and the structure factors of X6 (PDB code r1aiesf) were obtained from the Protein Data Bank (PDB; Berman *et al.*, 2000). The crystal structure X6 belongs to the space group $P422$ and there is one monomer in the asymmetric unit. The X6 tetramer exhibits 222 symmetry and is generated by the crystallographic twofold axes. All molecular-replacement calculations were performed with the program *AMoRe* (Collaborative Computational Project, Number 4, 1994; Navaza, 1994; Navaza & Saludjian, 1997). A standard protocol was adopted for all calculations so that the results could be compared directly. The monomer model of X6 was first rotated and translated arbitrarily to a new position. All search models were moved with their molecule A aligning with this 'offset' X6 model. In some cases when the correct peak is not the highest intensity peak, this can help identifying its rank. For the cross rotation-function (RF) calculations, we used data in the resolution range 10–3.5 Å and an integration sphere with a radius of 13 Å. The top 99 solutions were employed for the translation search, also using data in the resolution range 10–3.5 Å. We extracted the top five translation-function (TF) peaks from each RF solution input and ranked all TF output by correlation coefficient. The sorted TF solutions were then subjected to rigid-body (RB) refinement, using first 10–3.5 Å and subsequently 10–3.0 Å data. Models that used distance-derived pseudo-B factors were prepared with a script, *rmsdB.pl* (available upon request from YWC), implementing the protocol of Wilmanns & Nilges (1996), using a multiplying factor of 1. The pseudo-B factors of these models were increased by 10 $Å^2$ in *AMoRe* because many atoms of the NMR models had pseudo-B factors that were unrealistically low (less than 2 $Å^2$). Each ensemble model was loaded into *AMoRe* as a single coordinate file. Owing to the limitations (on maximum number of atoms input) of our version of *AMoRe*, all ensemble models were prepared such that any side chains apart from glycines and alanines were truncated to serine using a script written by Gerard J. Kleywegt (ftp://alpha2.bmc.uu.se/pub/gerard/omac/multi_probe). We also limited the number of models in a given ensemble to 19–21; this has proved to be adequate in all four cases. Except for models employing pseudo-B factors, all models used uniform B factors of 15 $Å^2$. For structure comparison, all structures were aligned with residues 326–354 which comprised the well defined core of the protein. Backbone r.m.s.d. calculations were performed with the program *LSQKAB* (Collaborative Computational Project, Number 4, 1994) for equivalent C, C^{α}, O and N atoms of the main chain. All molecular-replacement solutions were verified to be correct with the program *xpack.pl* (Fu & Chen, 1996).

3. Results

3.1. Monomer *versus* tetramer models

When a monomer (chain A of the tetrameric NMR structure) is used as the search model, we could not obtain a molecular-replacement solution even with the best-refined NMR structure (N7), no matter whether an ensemble model or a single averaged structure was employed, with or without use of pseudo-B factors (results not shown). However, when a tetrameric model is used in the rotation-function search, as suggested by Mittl *et al.* (1998), huge unambiguous peaks came

Table 1
Published p53 structures with varying degrees of accuracy used in this study.

AU stands for asymmetric unit. Bold PDB codes represent NMR ensemble models.

Model	Crystal symmetry information	PDB code	Experimental data	Accuracy of model (backbone r.m.s.d. with X6) (Å)		Reference
				Monomer	Tetramer	
N1	—	1olg, **1olh**	3268 restraints	0.6	1.2	Clore *et al.* (1994); Clore, Omichinski *et al.* (1995)
N2	—	1pes, **1pet**	1980 restraints	1.1	1.9	Lee *et al.* (1994)
X3	*P*422, one monomer in AU	1c26	3815 reflections (6–1.7 Å)	0.15	0.16	Jeffrey *et al.* (1995)
N4†	—	**1sae–1saj**, 1sak, 1sal	4472 restraints	0.5	0.6	Clore, Ernst *et al.* (1995)
X5	$P3_121$, one tetramer in AU	Not deposited	4722 reflections (8–2.5 Å)	Not reported	0.5 (C^{α})‡	Miller *et al.* (1996)
X6	*P*422, one monomer in AU	1aie	5355 reflections (8–1.5 Å)	0.0	0.0	Mittl *et al.* (1998)
N7	—	**3sak**	Same as N4, plus two pseudo-potentials	0.4	0.4	Kuszewski *et al.* (1999)

† 1sal was used as ⟨N4⟩; the N4 ensemble model, $\{N4\}_{21}$, was comprised of 21 models taken from 1saf. ‡ Data as reported in the publication.

up as the highest intensity peaks when using the two best-refined NMR structures N4 and N7. Presumably, a monomer model has limited self-Patterson vectors compared with cross-Patterson vectors and leads to a low signal-to-noise ratio in molecular-replacement calculations. The number of self-Patterson vectors in a tetramer model is dramatically increased, thus producing much clearer results. Mittl *et al.* (1998) also noted that the more spherical shape of the tetramer leads to a better discrimination between self- and cross-Patterson vectors.

3.2. More is better? Ensemble *versus* averaged models

In all four structures tested, ensemble models are more successful than single averaged models. For N1 and N2, only ensemble models led to correct solutions; for N4 and N7, while both ensemble and single models led to correct solutions, ensemble models generated clearer results: higher correlation coefficients and lower *R* factors (Table 2 and Fig. 1). Contrary to the report of Mittl *et al.* (1998), we could obtain a correct molecular-replacement solution without knowledge of the orientation of the internal 222 symmetry of the tetramer structure when starting with the ensemble search model $\{N2\}_{19}$. It is not necessary to pre-process the search model by aligning the molecule's internal twofold symmetry axis with the crystal's cell edge and to look for solutions that are consistent with the 222 symmetry of the tetramer. Single models ⟨N1⟩ and ⟨N2⟩ failed to arrive at a solution, which agrees with previous findings (Mittl *et al.*, 1998). It is not at all clear why the single models ⟨N2⟩ and $\langle N2\rangle_B$ did not produce a solution despite the correct RF peaks being clearly identified and being much better than the corresponding RF peak of $\{N2\}_{19}$ (Fig. 1). From this comparison, it appears that ensemble models produce more consistent results and are more likely to arrive at the correct solution.

3.3. Model flexibility: ensemble *versus* distance-derived pseudo-*B* factors

The temperature factor (*B* factor) is a very important quantity derived from crystal structures. Atomic *B* factors describe the relative flexibility (and thus reliability) of different parts of a structure. For molecular replacement, it is important to weight up rigid or well defined regions and weight down flexible or imprecisely determined regions in a search model. NMR structures do not have *B* factors and therefore require some manipulation to describe the relative reliability of atomic positions. For this purpose, two approaches have been introduced. (i) Artificial *B* factors are derived from the atomic r.m.s.d. to the mean structure (Wilmanns & Nilges, 1996). This can be easily adopted for use with a single minimized averaged model (see, for example, Wenk *et al.*, 1999). (ii) An ensemble model (Brünger *et al.*, 1987; Leahy *et al.*, 1992; Kleywegt *et al.*, 1994; Müller *et al.*, 1995) is used to represent the relative diversities of atomic positions in the conformational space. Here, we compared the two methods using the four NMR structures. The $\{N1\}_{20}$ and $\{N2\}_{19}$ ensembles produced outstanding top peaks that were easiest to interpret at every stage (Fig. 1). On the other hand, a single model with distance-derived pseudo-*B* factors did not facilitate a solution to be obtained with $\langle N2\rangle_B$. The advantages of using pseudo-*B* factors were best demonstrated when ⟨N1⟩ and $\langle N1\rangle_B$ were compared. $\langle N1\rangle_B$ did give a clear top solution after RB refinement (Table 2 and Fig. 1) despite the fact that the RF solution ranked ninth (Fig. 1) and the subsequent TF solution ranked second (data not shown). In the two more accurate structures, the single models without pseudo-*B* factors, ⟨N4⟩ and ⟨N7⟩, are good enough to produce the correct solutions, but the use of pseudo-*B* factors led to improvements in performance (Table 2). Both $\langle N4\rangle_B$ and $\langle N7\rangle_B$ offered better signal-to-noise discrimination than the respective ensemble models in the combined TF/RB calculations (Fig. 1).

3.4. More accurate search models lead to better solutions

Search models of different degrees of accuracy (as reflected by the backbone r.m.s.d. to the crystal structure X6) met varying degrees of difficulty in the process of MR. Using the less accurate models N1 and N2, whether one could solve the structure by molecular replacement depended critically on

how the model was prepared, as discussed above. In the case of N2, use of single averaged models ⟨N2⟩ and $\langle N2 \rangle_B$ did give very prominent RF solutions but then failed in the subsequent TF search and RB refinement. By way of contrast, when $\langle N1 \rangle_B$ was used, although there was no obvious RF solution (the correct solution ranked ninth), a correct molecular-replacement solution could be obtained after the TF search and the final RB refinement. The two more accurate models, N4 and N7, produced molecular-replacement results that were easiest to interpret, no matter whether a single averaged model or an ensemble model was used.

There is an approximate correlation between the quality of the final solution (as reflected by the *R* factor of the model after final RB refinement, Table 2) and the accuracy of the search model (as indicated by the backbone r.m.s.d. to X6, Table 1). From the least accurate model N2 to the most accurate N7, these led to corresponding molecular-replacement solutions having decreasing *R* factors ranging from 53 to 42%. The best model resulted in a molecular-replacement solution with an *R* factor of 42.3% for data from 10–3.0 Å, which is a very good starting point for subsequent crystallographic refinement.

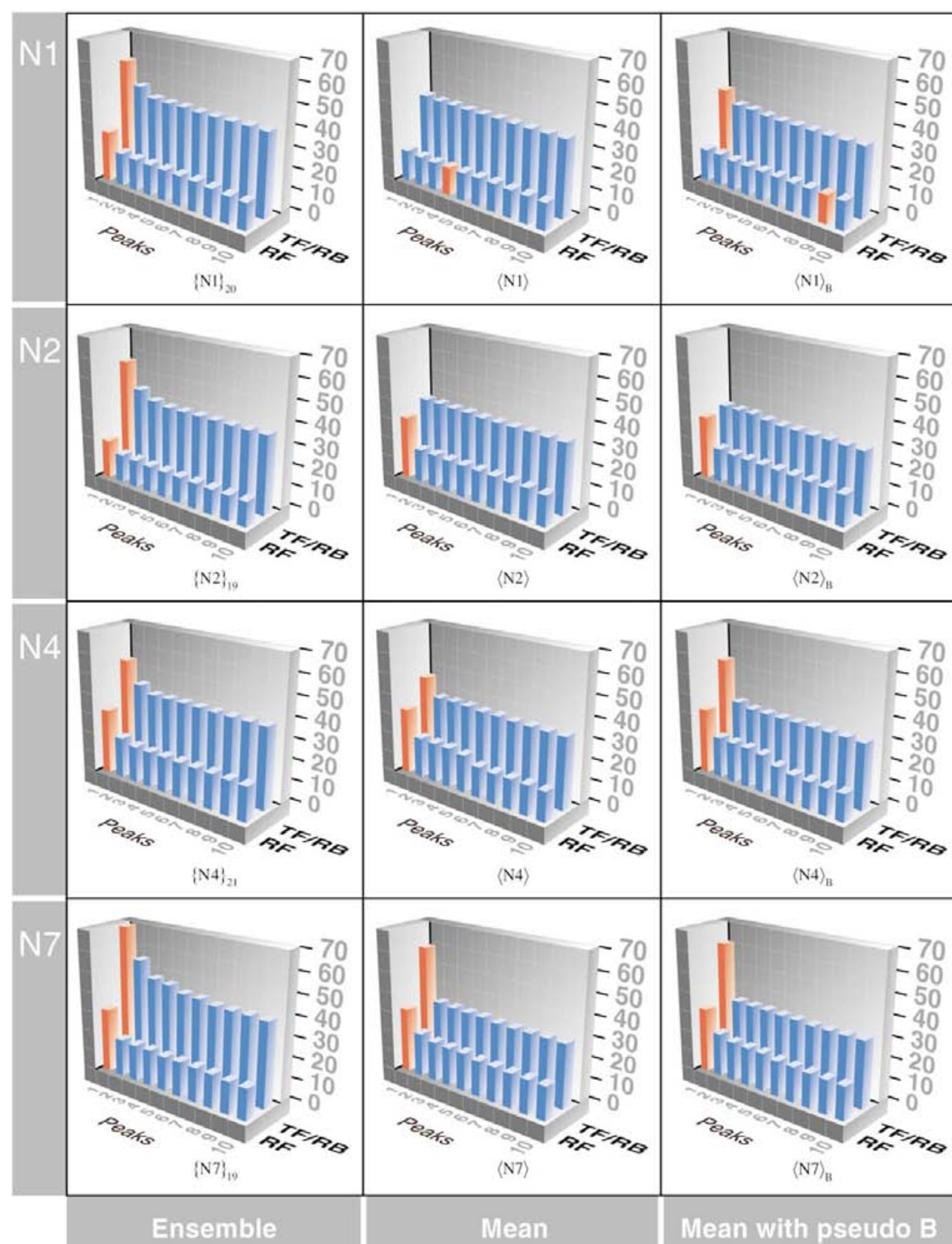

Figure 1
Molecular-replacement results comparing signal-to-noise ratios using four different NMR models (N1, N2, N4 and N7). Vertical values are correlation coefficients (CC). Shown here in each plot are the top ten peaks in each search, in descending order to the right. Rotation-function results are labelled 'RF' and results from the combined translation-function search with rigid-body refinement are labelled 'TF/RB'. Red bars are the correct solutions; blue bars are noise peaks. No TF/RB solutions were found when using ⟨N1⟩, ⟨N2⟩ and $\langle N2 \rangle_B$ as search models.

4. Discussion

It is interesting to observe that a set of search models of varying quality produces a spectrum of molecular-replacement experiences, from very straightforward cases to failed cases. For the three models that did not lead to a solution we cannot conclude that these cases are unsolvable. We can only state that they failed under the conditions specified in the standard protocol, bearing in mind that we used very typical molecular-replacement search parameters. From these results, we can observe some trends and extract useful recommendations for molecular-replacement experiments using NMR search models in general.

4.1. On models

Brünger *et al.* (1987) first demonstrated that NMR ensembles can be used as search models in molecular replacement. Since then, about two dozen crystal structures have been solved with NMR models: the majority using single NMR mean structure models (see, for example, Baldwin *et al.*, 1991; Anderson

Table 2
Molecular-replacement results of various p53 NMR models.

$\langle N1 \rangle$ and $\langle N1 \rangle_B$ represent the minimized averaged structure of N1 and that with distance-derived B factors; $\{N1\}_{20}$ represents an ensemble of N1 with 20 models. RF stands for rotation function, TF for translation function and RB for rigid-body refinement. The results of cross-rotation function searches are given in correlation coefficient (CC) and in multiples of σ (r.m.s.d. from mean density of search map).

Search model	Cross RF search (10–3.5 Å), CC (multiples of σ)		Combined TF search and RB (10–3.5 Å), CC		Final RB (10–3.0 Å), CC	Final R factor (%)
	Correct peak	Highest noise	Correct peak	Highest noise		
$\langle N1 \rangle$	14.1 (2.8σ)†	15.6 (3.1σ)	No solution	—	No solution	—
$\langle N1 \rangle_B$	14.1 (2.8σ)‡	16.2 (3.2σ)	43.5	37.6	43.9	52.6
$\{N1\}_{20}$	25.2 (4.7σ)	16.4 (3.1σ)	58.1§	47.0	57.5	51.2
$\langle N2 \rangle$	31.0 (5.6σ)	16.4 (3.0σ)	No solution	—	No solution	—
$\langle N2 \rangle_B$	31.2 (5.7σ)	17.0 (3.1σ)	No solution	—	No solution	—
$\{N2\}_{19}$	19.0 (3.8σ)	13.9 (2.8σ)	56.1§	43.6	54.5	52.7
$\langle N4 \rangle$	32.6 (6.2σ)	20.0 (3.8σ)	45.6	37.0	47.9	50.4
$\langle N4 \rangle_B$	32.4 (6.2σ)	20.2 (3.9σ)	54.4§	34.8	53.0	48.4
$\{N4\}_{21}$	32.0 (5.8σ)	20.0 (3.6σ)	54.2§	43.5	55.5	49.1
$\langle N7 \rangle$	31.9 (5.5σ)	21.4 (3.7σ)	59.7§	33.7	59.5	44.6
$\langle N7 \rangle_B$	31.9 (5.5σ)	21.4 (3.7σ)	61.6§	34.3	61.9	42.7
$\{N7\}_{19}$	31.6 (5.8σ)	18.3 (3.4σ)	70.9§	55.1	68.9	42.3

† Ranked fourth. ‡ Ranked ninth. § Highest correlation coefficient accompanied by lowest R factor.

et al., 1996; Chirgadze *et al.*, 1999; Wenk *et al.*, 1999) because they are easier to manipulate. In recent years, there have been several cases reporting the success of ensemble models in molecular replacement over single models (see for example, Müller *et al.*, 1995; Dennis *et al.*, 1998; Hoedemaeker *et al.*, 1999). Our results are in line with this observation. We recommend using an NMR ensemble (of about 20 structures) as a search model for molecular replacement rather than a single minimized averaged structure, unless the single mean structure is of excellent quality and high accuracy. In this study, we found that ensemble models have a larger tolerance of initial structural difference: a model with a backbone coordinate accuracy of 1.9 Å ($\{N2\}_{19}$) can still lead to success. However, for a single minimized averaged model with pseudo-B factors, the limitation in accuracy lies somewhere between 1.2 Å ($\langle N1 \rangle_B$, successful) and 1.9 Å ($\langle N2 \rangle_B$, failed).

Our results can be explained by considering an NMR ensemble as a set of individual models following a non-Gaussian distribution (because they are equally weighted) from the mean structure. If the structure is highly accurate (N4 and N7), the mean is a better representation of the 'truth' than the ensemble and yields clearer results. However, if a structure is of limited accuracy (N1 and N2), an ensemble offers the 'outlying' (relative to the mean) conformers equal contributions to molecular-replacement calculations. In effect, an ensemble can explore more of the conformation space and thus is more tolerable to model inaccuracy.

It is important to bear in mind how accurate the search model is, especially if it is a multidomain or multisubunit structure. As exemplified in the p53 tetramerization domain study, the intersubunit interactions can be poorly defined compared with the intramolecular interactions, resulting in a relatively accurate monomer structure but poorly defined relative orientations of the monomers. This can lead to failure in a subsequent molecular-replacement trial. The improvement from model N1 to model N4 highlights the importance of obtaining adequate intersubunit nuclear Overhauser effect (NOE) restraints in structure calculation (see below). The work of Kuszewski *et al.* (1999) showed that substantial improvement in the accuracy of the NMR structure can be achieved by employing additional pseudo-potentials for a conformational database and the radius of gyration. The former is independent of experimental data and the latter can either be obtained from small-angle X-ray scattering experiments or be estimated from the number of residues under study.

A molecular-replacement solution cannot warrant success in the subsequent structure refinement. Phases calculated from the molecular-replacement solution are biased towards the search model and can hinder refinement. In extreme cases, the search model and the real crystal structure can differ to an extent that falls outside the reach of refinement programs. It is important to examine the whole NMR ensemble and omit regions that show large conformational variability, *i.e.* those regions that are underdetermined owing to insufficient NMR restraints, before refinement proceeds.

4.2. Difference between NMR models and crystal structures: is it real?

In this study of the p53 tetramerization domain, the structural difference expressed in r.m.s.d. of equivalent backbone atomic positions can be a result of insufficient NMR data. Considering the whole tetramer, the earliest model has a backbone r.m.s.d. of 3.3 Å to X6. This was corrected and replaced by N1 and the backbone r.m.s.d. to X6 reduced to 1.2 Å (Clore, Omichinski *et al.*, 1995). This improvement was a result of correction and addition of a very small number of intersubunit interproton distance restraints which proved to be critical in defining the relative orientation of the *AC* dimer to the *BD* dimer (Clore, Omichinski *et al.*, 1995). Taking the N2 model and comparing this with X6, one may be tempted to believe that there is a genuine structural difference between the NMR structure and the crystal structure, as reflected by a backbone r.m.s.d. of about 1 Å for the monomer, which seems quite acceptable, and a much larger r.m.s.d. of 1.9 Å for the tetramer. Thus, there is a difference in the dimer–dimer interaction in these two structures. One could even argue that crystal packing enforces a strict 222 symmetry on the tetra-

meric molecule, which is relaxed and so is more native-like in solution. As demonstrated by the later NMR structures, this is simply not true. Let us now consider models N1, N4 and N7. There is little improvement in the backbone r.m.s.d. to X6 among monomers (from 0.6 to 0.4 Å), indicating that the monomer structure is well defined in N1. However, the backbone r.m.s.d. to X6 among the tetramers improves substantially, from 1.2 to 0.6 to 0.4 Å. The final structure, N7, has a tetrameric backbone r.m.s.d. of only 0.4 Å to X6, a value that is the same as that of the monomeric structural difference. This demonstrates that the NMR structure of this protein, when fully refined, is the same as the crystal structure, both at the monomer and tetramer levels. The lesson to learn here is that when we compare NMR structures with crystal structures, one should be very cautious if large structural differences are observed. If rigid-body displacement is observed between secondary-structural elements or between subunits, it is important to investigate how much data contribute to defining those interactions and whether these data are correct.

References

Anderson, D. H., Weiss, M. S. & Eisenberg, D. (1996). *Acta Cryst.* D**52**, 469–480.

Baldwin, E. T., Weber, I. T., St. Charles, R., Xuan, J.-C., Appella, E., Yamada, M., Matsushima, K., Edwards, B. F. P., Clore, G. M., Gronenborn, A. M. & Wlodawer, A. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 502–506.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Bewley, C. A., Gustafson, K. R., Boyd, M. R., Covell, D. G., Bax, A., Clore, G. M. & Gronenborn, A. M. (1998). *Nature Struct. Biol.* **5**, 571–578.

Brünger, A. T., Campbell, R. L., Clore, G. M., Gronenborn, A. M., Karplus, M., Petsko, G. A. & Teeter, M. M. (1987). *Science*, **235**, 1049–1053.

Chirgadze, D. Y., Hepple, J. P., Zhou, H., Byrd, R. A., Blundell, T. L. & Gherardi, E. (1999). *Nature Struct. Biol.* **6**, 72–79.

Clore, G. M., Ernst, J., Clubb, R., Omichinski, J. G., Kennedy, W. M. P., Sakaguchi, K., Appella, E. & Gronenborn, A. M. (1995). *Nature Struct. Biol.* **2**, 321–333.

Clore, G. M. & Gronenborn, A. M. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 5891–5898.

Clore, G. M., Omichinski, J. G., Sakaguchi, K., Zambrano, N., Sakamoto, H., Appella, E. & Gronenborn, A. M. (1994). *Science*, **265**, 386–391.

Clore, G. M., Omichinski, J. G., Sakaguchi, K., Zambrano, N., Sakamoto, H., Appella, E. & Gronenborn, A. M. (1995). *Science*, **267**, 1515–1516.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Dennis, C. A., Videler, H., Pauptit, R. A., Wallis, R., James, R., Moore, G. R. & Kleanthous, C. (1998). *Biochem J.* **333**, 183–191.

Fu, T. Y. & Chen, Y. W. (1996). *J. Appl. Cryst.* **29**, 594–597.

Hoedemaeker, F. J., Siegal, G., Roe, S. M., Driscoll, P. C. & Abrahams, J. P. (1999). *J. Mol. Biol.* **292**, 763–770.

Jeffrey, P. D., Gorina, S. & Pavletich, N. P. (1995). *Science*, **267**, 1498–1502.

Kleywegt, G. J., Bergfors, T., Senn, H., Le Motte, P., Gsell, B., Shudo, K. & Jones, T. A. (1994). *Structure*, **2**, 1241–1258.

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

Leahy, D. J., Axel, R. & Hendrickson, W. A. (1992). *Cell*, **68**, 1145–1162.

Lee, W., Harvey, T. S., Yin, Y., Yau, P., Litchfield, D. & Arrowsmith, C. H. (1994). *Nature Struct. Biol.* **1**, 877–890.

Miller, M., Lubkowski, J., Rao, J. K. M., Danishefsky, A. T., Omichinski, J. G., Sakaguchi, K., Sakamoto, H., Appella, E., Gronenborn, A. M. & Clore, G. M. (1996). *FEBS Lett.* **399**, 166–170.

Mittl, P. R. E., Chène, P. & Grütter, M. G. (1998). *Acta Cryst.* D**54**, 86–89.

Müller, T., Oehlenschläger, F. & Buehner, M. (1995). *J. Mol. Biol.* **247**, 360–372.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Navaza, J. & Saludjian, P. (1997). *Methods Enzymol.* **276**, 581–594.

Rossmann, M. G. (1972). Editor. *The Molecular Replacement Method.* New York/London/Paris: Gordon & Breach.

Wenk, M., Baumgartner, R., Holak, T. A., Huber, R., Jaenicke, R. & Mayr, E.-M. (1999). *J. Mol. Biol.* **286**, 1533–1545.

Wilmanns, M. & Nilges, M. (1996). *Acta Cryst.* D**52**, 973–982.

Yang, F., Bewley, C. A., Louis, J. M., Gustafson, K. R., Boyd, M. R., Gronenborn, A. M., Clore, G. M. & Wlodawer, A. (1999). *J. Mol. Biol.* **288**, 403–412.