Where freedom is given, liberties are taken

Structure 15 June 1995, **3**:535–540

Many people who are interested in biology are still under the impression that the beautiful pictures of macromolecules presented in this journal and others are to be believed down to the most intricate detail [1]. Some years ago, Brändén and one of us (TAJ) [2] attempted to explain to the non-specialist how some aspects of protein crystallography require a subjective interpretation of the diffraction results that, in the worst cases, can lead to serious errors, which, in the best cases, get published in our most prestigious scientific journals. The crystallographic community has since then attempted to learn from these mistakes by, for example, developing new methods for building and refining structures, and by trying to recognize errors in models (for a recent review, see [3]). The adverse publicity probably helped to ensure that individual groups were more careful in assessing the accuracy of their work before pressing for publication. After all, few of us like to look foolish.

Here, we argue that, overall, the community is still doing a poor job in its treatment of structures whose crystals diffract poorly. In the worst cases, even if there are no 'errors', biological results are being interpreted with a precision that is not warranted by the information contained in the diffraction data. In the best of cases, a low R-factor (a measurement of how well a model fits the measured diffraction data) is being waved around as proof of the correctness of the structure. In their publication [2], Brändén and Jones warned that structures with R-factors around 25% or higher could indicate problem structures.

Errors in a model

Let us recap the kind of errors that can occur in a crystallographic study, in order of severity. We have rarely seen the publication of a structure that has a totally wrong fold, although the preliminary structure of asparaginase/glutaminase (first published in 1988 [4] and later corrected by the same authors [5]) came close, having only ~20% of its residues in approximately the correct position. A less severe error occurs when structural units such as α-helices and β-strands are correctly identified but are joined up incorrectly. It is our impression that this kind of error is less common than it used to be.

A much more likely error occurs when the sequence gets out of register with the electron density for parts of the structure. When this happens, if one is lucky, a second error brings the sequence back into register with the density. This error is easy to make and has occurred in a few published structures. At the next level of error, the structure may be locally incorrect in either the main chain or the side chains or in both. These kinds of errors are easy for the non-specialist to understand. To follow the next discussion, it is necessary to understand how structures are improved after the first model has been made and this is best done with a short history lesson.

History and some jargon

If one has the coordinates of a structure, it is a straightforward calculation to see how well it matches the experimental diffraction data. The R-factor is defined as $\Sigma||F_o|-|F_c||/\Sigma|F_o|$, where F_o is the observed structure factor derived from the diffraction experiment, and F_c is the calculated structure factor obtained from the atomic model. Crystallographic refinement refers to a method where a model is changed to reduce the discrepancy between F_o and F_c, and, thereby, the R-factor. This has been an area of active research since the pioneering work of Jensen and co-workers, more than 20 years ago [6]. In their original work, each atom was described by a coordinate and a temperature factor that indicates how much the atom moves around its equilibrium position. No attempt was made to impose knowledge of stereochemistry so that when viewed on a computer graphics system, even flat aromatic rings looked distorted.

A totally different approach was taken by Huber's laboratory [7] using a program developed by Diamond [8]. In this program, changes to the structure were made by bond rotations, keeping the stereochemistry fixed. During the 1970s and 1980s, various programs were developed that used methods somewhere in between these two approaches. In some programs, such as PROLSQ [9] and EREF [10], stereochemical information could be used as part of the function being minimized. A model generated by such programs would not have exactly standard bond lengths, angles, flat planes, chiral centres etc. Deviations from standard values could then be indicated by stating root mean square deviations (rmsds) from ideal values. In such programs, stereochemistry is maintained by applying restraints to these preferred values.

In a different approach, the program CORELS [11] allowed individual residues to move as rigid bodies and with bond rotations, while restraints could be used between neighbouring residues to keep good stereochemistry. Another widely used program, TNT [12], allows a mixture of restrained and constrained groups to be refined. The word 'constraint' is now used to indicate that the group of atoms has exact stereochemistry, usually conforming to some set of standard values, which are often obtained from small-molecule crystallography. How to weigh the relative contributions of the geometrical and crystallographic components became a matter that was discussed with almost religious zeal and intensity

(see, for example, the Collaborative Computational Project, No. 4 [CCP4] proceedings [13]). Should the stereochemical restraints be tight, slack, or somewhere in between, reflecting the variation seen in small molecules? How should temperature factors be treated? Should one use just one per molecule, one per atom, or one per residue; should neighbouring atoms have similar values, etc.? Why are these questions important?

A major advance in refinement methodology took place with the advent of minimizers that made use of molecular dynamics protocols to increase the radius of convergence [14,15]. Molecular dynamics was not being used to study dynamic properties of the molecule, but as a means to an end, to get a low R-factor. The program X–PLOR, developed by Brünger, rapidly gained in popularity amongst the community. In 1992, Brünger [16] made another important advance: he divided the diffraction data into two pots, one used by the minimizers (the working set, usually consisting of 90–95% of the data), and the other used to evaluate the progress of the refinement by generating a second reliability index which he called the free R-factor, R_{free}. At the time, one of us (TAJ) was not terribly impressed by the use of R_{free} to refine a structure with a folding error, as it did not seem greatly to assist in recognition of the error. The importance of R_{free} has now become clearer to us in that it can be used to guide the refinement protocol, whether it is the optimum choice of weights [17], temperature-factor model, or most anything else. One of the most important findings is that R_{free} is highly correlated with the accuracy of the model (more precisely, with the mean phase error).

One final bit of jargon concerns non-crystallographic symmetry (NCS). When crystals are formed, the repeating (or asymmetric) unit may consist of just one macromolecule or many (ranging from two to sixty for some virus structures). If more than one copy of a molecule is present in the repeating unit, the structure is said to contain non-crystallographic symmetry. NCS is very common; we estimate that it occurs in roughly half the solved low-resolution protein structures. Whereas the repeating units within a crystal are by definition identical, the separate (NCS) units may not be. Real differences may exist due to a domain movement caused by different crystal contacts, for example, and loops and side chains on the surface of the molecule may differ for the same reason. Different ways of treating NCS-related molecules during refinement exist: they can all be kept identical ('constrained'), kept similar ('restrained'), or set free from one another. Even experienced crystallographers seem to confuse the use of constrained and restrained in this context. One referee recently asked us if "the NCS restraints were applied to side-chain as well as main-chain atoms ?" although our text clearly stated that we used strict NCS constraints.

Today's standard refinement protocol

Today's 'average structure' is built with FRODO [18] or O [19]. It is then refined with a simulated annealing protocol in X-PLOR, and subjected to a number of cycles of rebuilding at the graphics workstation (this is needed to make the kind of changes to the model that are outside the radius of convergence of the refinement program) followed by more minimization. Whatever the resolution, individual atomic temperature factors will have been refined. If there is NCS, it will most likely have been ignored. The molecular geometry will be restrained, with rmsds of 0.01–0.02 Å on bond lengths, 2–4° on bond angles and 2–4° on improper dihedrals. If the best currently available force field parameters were used [20], the rmsds are likely to be on the low side of the ranges quoted. For structures solved at 2.5–3 Å resolution, the R-factor will be 15–20%. Whatever the resolution, there will be water molecules, accounting for ~5–10% of the atoms in the model. If there is NCS, the similarities between the NCS-related molecules may be described by the rmsd based on the Cα atoms. Differences between such molecules may be discussed with their relevance to biological function.

Few quality checks will be provided except for the rmsds on the geometrical restraints. A caring minority will mention how good the stereochemistry of the structure is when compared with other structures solved at the same resolution using PROCHECK [21] — it will not be worse than average. If the structure is published in *Nature*, *Science* or *Cell* there will be no Ramachandran plot, but there will be one if it is published in *Acta Crystallographica D*. If the paper describes both a native structure and a ligand complex, the native structure will have been solved at higher resolution than the complex (or *vice versa*, if the complex formation stabilizes the protein and improves the diffraction quality of the crystals). They will have the same R-factor, and if it is possible to evaluate the refinement of the complex from the paper, it will have been done with the same protocol as the native structure.

Some examples

So what's wrong with this protocol? Maybe nothing, but it depends on the crystallographer, the resolution (i.e. how well the crystal diffracts and hence how much diffraction data have been collected), and how good the starting model was. To illustrate this, some refinement statistics for models of two closely related proteins, cellular retinoic acid binding proteins (CRABP) I and II [22], are given in Table 1. A *Nature/Science/Cell* paper, would probably describe model X of CRABP II something like: "The structure has been refined to an R-factor of 21.4% with better than average stereochemistry when compared to other structures solved at this resolution. Only a few residues deviate from allowed regions of the Ramachandran plot. A conservative number of waters have been added to the model." This is in fact a structure which, after it had been solved correctly at 1.8 Å, was intentionally traced backwards, i.e. the N-terminal residue has been built at the C terminus, and the correct C-terminal residue has been built at the N terminus. In such a model, every residue is incorrectly placed. We have described such a structure previously, refining it to the

Table 1. Comparison of some model and refinement statistics of CRABP II (intentionally traced backwards; model X), CRABP I refined conservatively (with NCS constraints and grouped temperature factors; model Y), and the CRABP I structure refined according to today's 'standard' refinement practices (model Z).

Model	X	Y	Z
Resolution range (Å)	8.0–3.0	8.0–2.9	6.0–2.9
R	0.214	0.251	0.169
R_{free}	0.617	0.320	0.323
Rmsd bond lengths (Å)	0.009	0.009	0.010
Rmsd bond angles (°)	2.1	1.6	1.6
Temperature factor model	B_{iso}	grouped	B_{iso}
Average temperature factor (Å^2)	13.4	49.2	43.5
Rms ΔB bonded atoms (Å^2)	4.1	–	6.9
Rmsd NCS-related Cα atoms (Å)	–	0.0	0.50
Rmsd all NCS atoms (Å)	–	0.0	0.99
Rms ΔB NSC-related Cα atoms (Å^2)	–	0.0	10.5
Rms ΔB all NCS atoms (Å^2)	–	0.0	12.0
Rms Δφ NCS-related residues (°)	–	0.0	27.8
Residues with $\|\Delta\phi\|>10°$ (%)	–	0.0	52.2
Rms Δψ NCS-related residues (°)	–	0.0	30.2
Residues with $\|\Delta\psi\|>10°$ (%)	–	0.0	54.4
Ramachandran plot, most favoured areas (%)	42.7	81.6	77.5
Additional allowed areas (%)	36.3	16.0	20.1
Generously allowed areas (%)	12.1	1.6	1.6
Disallowed areas (%)	8.9	0.8	0.8
Unusual peptide orientations (%)	24.1	2.2	2.6
Non-rotamer side-chain conformations (%)	29.2	7.4	11.8
Overall G-factor [21]	–0.4	+0.1	+0.06
Overall DACA score [34]	–2.6	–0.4	–0.5

Table 2. Model and refinement statistics for the structures of GST and Rubisco solved both at low and high resolution. (See the text for details.)

Model	1GUH	'ALEX'	9RUB	5RUB
Resolution (Å)	2.6	2.0	2.6	1.7
R/R_{free}	0.229/–	0.196/0.245	0.199/–	0.180/–
Temperature factor model	grouped	B_{iso}	B_{iso}	B_{iso}
Av. temperature factor (Å^2)	35.1	25.5	19.3	29.3
Rms ΔB bonded atoms (Å^2)	–	2.7	1.4	1.0
Rmsd all NCS atoms (Å)	0	0.57	2.31	1.25
Rms ΔB all NCS atoms (Å^2)	0	4.2	7.8	5.3
Rmsd core Cα atoms (Å)	0	0.09	0.95	0.89
Rms ΔB core Cα atoms (Å^2)	0	2.1	7.6	5.1
Rms Δφ NCS-related residues (°)	0	3.0	45.9	18.3
Residues with $\|\Delta\phi\|>10°$ (%)	0	2.3	65.3	18.3
Rms Δψ NCS-related residues (°)	0	3.0	45.8	20.0
Residues with $\|\Delta\psi\|>10°$ (%)	0	0.9	67.2	19.2
Residues with $\|\Delta$ Cα-Cα-Cα$\|>5°$ (%)	0	1.4	51.0	12.5
Residues with $\|\Delta$ Cα-Cα-Cα-Cα$\|>10°$ (%)	0	0.9	39.6	12.5
Ramachandran plot, most favoured areas (%)	91.9	90.9	74.1	91.0
Additional allowed areas (%)	8.1	8.4	19.5	8.2
Generously allowed areas (%)	0	0.8	4.3	0.6
Disallowed areas (%)	0	0	2.2	0.3
Secondary structure (%)	69.2	69.0	58.2	63.2
Non-rotamer side-chain conformations (%)	11.8	10.0	20.0	11.4
Unusual peptide orientations (%)	1.8	2.0	6.8	2.6
Overall G-factor [21]	0.0	+0.4	–1.3	–0.4
Overall DACA score [34]	–0.7	–0.6	–1.5	–0.7

25% barrier [19]. Now, we have broken the barrier, dropping the R-factor into the realm of respectability by using X-PLOR, Engh and Huber parameters, judicious choice of weights, and freely refining atomic temperature factors.

Model Y of CRABP I has been refined by what we consider to be a good protocol. This crystal form has two molecules in the asymmetric unit and they have been kept identical (constrained) throughout the refinement process. Only two temperature factors have been refined per residue, one for the main-chain and one for the side-chain atoms. The model does not break the 25% barrier. Note that judged by 'conventional criteria', model Y looks even worse than model X. However, Brünger's R_{free} comes to our rescue and clearly shows up the wrong structure (naturally, the maps didn't look very good in the case of model X either).

What happens if we apply today's standard protocol to model Y? Setting free the NCS constraints and temperature factors immediately causes the number of variables in the system to shoot up and thanks to X-PLOR's powerful minimizer, the R-factor immediately drops to 16.9% (model Z in Table 1). But R_{free} increases slightly, indicating that the drop is meaningless: in fact, we have produced a less accurate model. The two molecules are now different (rmsds of 0.5 Å and 1.0 Å for Cα and all atoms, respectively) but these differences are just artefacts. One cannot conclude that the two NCS-related molecules are identical; however, our diffraction data (a poor 2.9 Å dataset) does not allow us to model any differences. Any information that may be present in the data about the differences apparently drowns in the noise of the experimental error.

The lesson is that when it comes to refinement, if freedom is given, liberties will be taken, or to quote Hendrickson: "That which is not restricted will take its liberties" [13]. This maxim also holds true for model building and re-building. The O project [19] was started in part to demonstrate that one could build structures using just groups of atoms, selected from main-chain and side-chain databases [23].

A second example (Table 2) compares two sets of refinements at low and high resolution where there is non-crystallographic symmetry. Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) from *Rhodospirillum rubrum* has been solved in the native state and with various ligands. The Protein Data Bank (PDB) coordinate set 5RUB [24] is native Rubisco refined to a resolution of 1.7 Å, whereas set 9RUB [25] is a complex with the substrate (ribulose-1,5-bisphosphate) refined at 2.6 Å. They were both refined by traditional protocols, using (mostly) PROLSQ and both contain two molecules in the asymmetric unit. Human alpha class glutathione transferase (GST) was first solved at 2.6 Å resolution (set 1GUH; [26]) and contains two dimers in the asymmetric unit, which were refined with constrained, fourfold NCS. A

mutant complex (set 'ALEX'; [27]) crystallizes in a closely related unit cell, diffracts to higher resolution, and has been refined with restrained twofold NCS. Both GSTs were refined with what we consider to be good protocols, although R_{free} was not used with 1GUH (this was before we appreciated its usefulness).

The Cα atoms of the two molecules of the high-resolution Rubisco model fit with an rmsd of 0.9 Å, while the two molecules in the low-resolution model have a considerably worse fit, ~1.8 Å. To show that this is not the result of conformational changes on ligand binding, the differences in their main-chain dihedral angles have been plotted (Fig. 1). In such a $\Delta\phi/\Delta\psi$ plot [28], spikes correspond to different peptide orientations (what we call pep-flips [19,29]) and other large changes may indicate local conformational changes. Figure 1a clearly shows that in the high-resolution Rubisco structure, the differences between the two NCS-related molecules are small even though they were not forced to be so during the refinement. In the low-resolution Rubisco model (Fig. 1b), there is clearly more noise in the plot as well as some large spikes. The noise is artefact, while some of the spikes may be real.

The same plot for ALEX, (Fig. 2a), shows small variations; and when it is compared with 1GUH, there are three pep-flips (where the carbonyl oxygen has been moved to point in the opposite direction) and a satisfyingly small variation (Fig. 2b). We predict, therefore, that the low-resolution Rubisco model was over-fitted. Since R_{free} correlates with the mean phase error in the system, a better re-refinement protocol may lower R_{free}, and produce improved electron density in the active site. Whether this is worth doing in this particular case is not known.

Contrary to what one might infer from Table 2, 1GUH was solved before ALEX, and 9RUB was solved after 5RUB. 1GUH was refined conservatively with constrained NCS and grouped temperature factors; 9RUB was refined traditionally with no NCS constraints or restraints and with individual atomic temperature factors. Consequently, the low-resolution 1GUH structure has model quality indicators that are as good as, or better than, those of the high-resolution Rubisco structure. The lesson to be learnt from this is that conservative refinement minimizes the chance of introducing artefacts and errors due to over-fitting, whereas traditional refinement more or less invites them.

About 95% of the structures published at 2.5 Å or lower resolution use the standard protocol. We don't mean that like model X in Table 1 they are all wrong, but many suffer from over-fitting like model Z (GJ Kleywegt and TA Jones, unpublished data). If authors publish just a qualitative description of their over-fitted structures, a better refinement protocol would be unlikely to change the final paper very much. However, unless the refinement protocol is a good one, it is not possible to

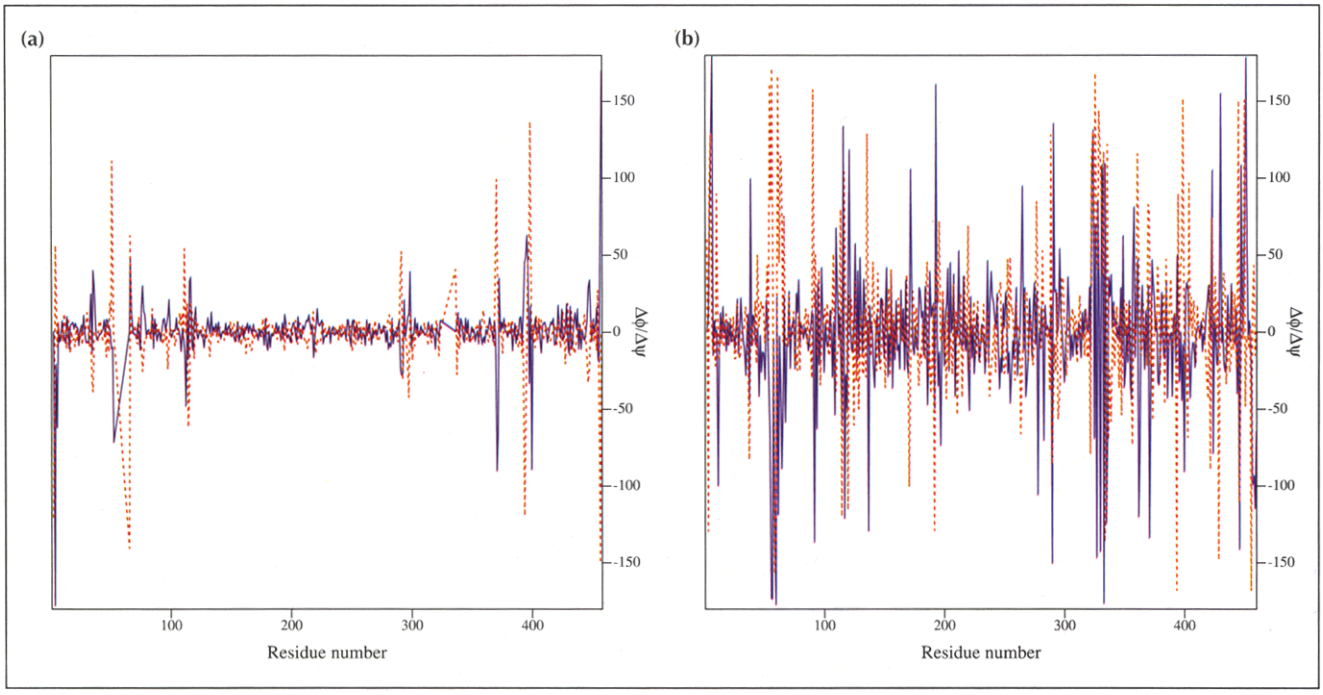

Fig. 1. $\Delta\phi/\Delta\psi$ plots for two Rubisco models with twofold NCS. **(a)** 5RUB solved at 1.7 Å, and **(b)** 9RUB solved at 2.6 Å resolution. The solid blue line shows the difference between the main-chain ϕ angles of NCS-related residues; the dashed red line shows the difference between corresponding ψ angles. Note that there are two gaps in the chain of 5RUB. In the high-resolution model, the two molecules in the asymmetric unit are fairly similar, although isolated variations occur. The (isomorphous) low-resolution model, on the other hand, has clearly suffered from over-fitting, which has introduced a lot more noise in the plot. Based on the low-resolution structure one might be tempted to conclude that the two NCS-related molecules display 'significant' overall differences. However, the high-resolution structure clearly shows that this is not the case.

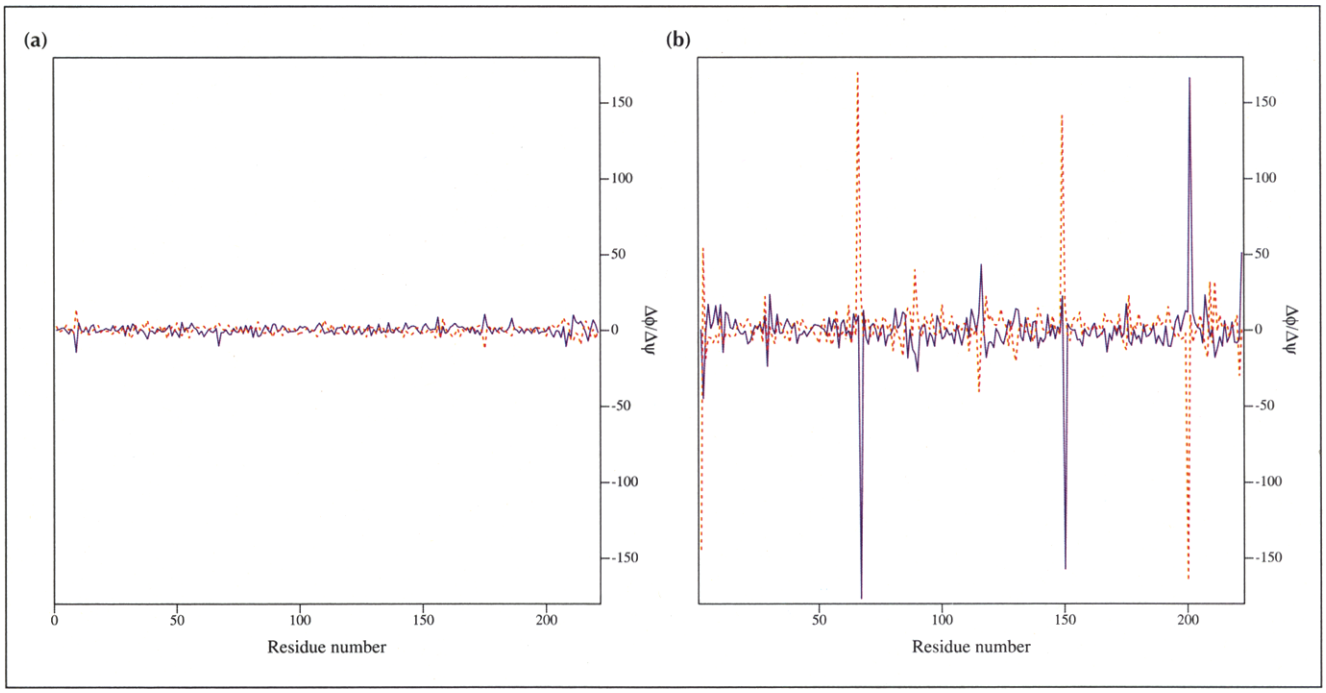

Fig. 2. (a) A $\Delta\phi/\Delta\psi$ plot for the two NCS-related molecules in a 2.0 Å structure of human alpha class glutathione transferase. The two molecules have been restrained to be similar, and during refinement and rebuilding no indications of major deviations from this assumption were encountered. **(b)** A $\Delta\phi/\Delta\psi$ plot comparing a 2.6 Å model of human alpha class glutathione transferase with the non-isomorphous 2 Å model. The low-resolution model was refined with conservative assumptions in order to reduce the number of degrees of freedom (constrained fourfold NCS and grouped temperature factors). Apart from the N-terminal residue and three places where the peptide plane was 'flipped', this has led to a 2.6 Å structure which, despite the limited resolution, is very similar to the 2 Å structure, and with comparable quality indicators (see Table 2). The colour coding is the same as in Figure 1.

know what level of detail is to be believed. In a good refinement protocol, the number of degrees of freedom is kept as low as possible. When it is increased, R_{free} is used to judge if the increase was justified. For example, if bonded atoms are allowed to have very different temperature factors, R_{free} should be used to prove that this yields a more accurate model. When NCS is present, the units should normally be identical, and R_{free} should be used to decide when and if the use of NCS restraints is justified. When strict NCS is used, phases can be improved by cyclic averaging (a number of easy-to-use programs are available to do this), which often yields vastly improved density maps. In addition, refinement and rebuilding of the structure becomes much faster. Hence, paradoxically, one can actually produce a better model faster than one can obtain an over-fitted one!

Concluding remarks

In our own laboratory, we have heard complaints that at low resolution one cannot afford to set aside reflections to use for the R_{free} calculations as there aren't that many reflections anyway. This argument must be ignored because the low-resolution refinements are the ones that suffer most from errors and over-fitting.

A legitimate question is whether R_{free} handles NCS cases correctly. Since NCS leads to relationships between reflections, R_{free} may be lower than it should be, just as in the case of the conventional R-factor [30]. This is clearly seen in the refinement of a complex of a bacteriophage coat protein with RNA (with 10-fold NCS) at 3 Å resolution, where R and R_{free} are almost identical (0.192 and 0.209, respectively) [31]. Such cases probably require new ways of selecting the test set of reflections — for example, by choosing all reflections in a number of thin, randomly chosen resolution shells. In most cases, though, R_{free} remains the best accuracy-indicator that we have at present.

So, should all of the structures deposited at the PDB [32] be re-refined with better protocols? (Given the advances in data collection technologies, should they perhaps be recrystallized?) The answer is that they probably won't be; however, we should make it possible for anyone who cares deeply enough to be able to do it. It is important, therefore, to search out all of our old reflection data sets and send them to the PDB (TAJ confesses negligence in the deposition of reflection data sets, but hopes to redeem himself). In the one case where we thought it necessary to re-do the refinement [33], we found what we believe is a space group error (note that such errors are not so important if a strict refinement protocol is used), and a region of 25 residues that was out of register with the density.

We realize that many people do not like being told how to refine their structures, and no doubt the 'best protocol' does not exist to cover all situations. However, the matter

is important enough that by writing this brief article we are willing to take the risk of offending some members of the crystallographic community. After all, there is no prize at the end of the road, according to *Science* [1].

Acknowledgements: We thank the Editors for allowing us to publish this diatribe, and Dr Alex Cameron for providing us with the coordinates of his unpublished glutathione transferase structure. We are preparing a considerably longer and more technical paper concerning rebuilding and refinement for a forthcoming volume of *Methods in Enzymology*, a preprint of which will be made available via the World Wide Web (http://onyx.bmc.uu.se/).

References

1. Hall, S.S. (1995). Protein images update natural history. *Science* **267**, 620–624.
2. Brändén, C.-I. & Jones, T.A. (1990). Between objectivity and subjectivity. *Nature* **343**, 687–689.
3. MacArthur, M.W., Laskowski, R.A. & Thornton, J.M. (1994). Knowledge-based validation of protein structure coordinates derived by X-ray crystallography and NMR spectroscopy. *Curr. Opin. Struct. Biol.* **4**, 731–737.
4. Ammon, H.L., *et al.*, & Roberts, J. (1988). Preliminary crystal structure of *Acinetobacter glutaminasificans* glutaminase-asparaginase. *Proc. Natl. Acad. Sci. USA* **263**, 150–156.
5. Lubkowski, J., *et al.*, & Swain, A.L. (1994). Refined crystal structure of *Acinetobacter glutaminasificans* glutaminase-asparaginase. *Acta Crystallogr. D* **50**, 826–832.
6. Watenpaugh, K.D., Sieker, L.C., Herriott, J.R. & Jensen, L.H. (1973). Refinement of the model of a protein: rubredoxin at 1.5 Å resolution. *Acta Crystallogr. B* **29**, 943–956.
7. Deisenhofer, J. & Steigemann, W. (1975). Crystallographic refinement of the structure of bovine pancreatic trypsin inhibitor at 1.5 Å resolution. *Acta Crystallogr. B* **31**, 238–250.
8. Diamond, R. (1971). A real-space refinement procedure for proteins. *Acta Crystallogr. A* **27**, 436–452.
9. Hendrickson, W.A. & Konnert, J.H. (1980). Incorporation of stereochemical information into crystallographic refinement. In *Computing in Crystallography*. (Diamond, R., Ramaseshan, S. & Venkatesan, K., eds), pp. 13.0–13.23, Indian Academy of Science, Bangalore.
10. Jack, A. & Levitt, M. (1978). Refinement of large structures by simultaneous minimization of energy and R factor. *Acta Crystallogr. A* **34**, 931–935.
11. Sussmann, J.L., Holbrook, S.R., Church, G.M. & Kim, S.H. (1977). A structure-factor least squares refinement procedure for macromolecular structures using constrained and restrained parameters. *Acta Crystallogr. A* **33**, 800–804.
12. Tronrud, D.E., Ten Eyck, L.F. & Matthews, B.W. (1987). An efficient general purpose least-squares refinement program for macromolecular structures. *Acta Crystallogr. A* **43**, 489–501.
13. Machin, P.A., Campbell, J.W. & Elder, M. (eds) (1980). *Refinement of Protein Structures*. SERC Daresbury Laboratory, Warrington, UK.
14. Brünger, A.T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458–460.
15. Fujinaga, M., Gros, P., & van Gunsteren, W.F. (1989). Testing the method of crystallographic refinement using molecular dynamics. *J. Appl. Crystallogr.* **22**, 1–8.
16. Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475.
17. Brünger, A.T. (1993). Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallogr. D* **49**, 24–36.
18. Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11**, 268–272.
19. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119.
20. Engh, R.A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A* **47**, 392–400.
21. Laskowski, R.A., MacArthur, M.W. & Thornton, J.M. (1994). Evaluation of protein coordinate sets. In *From First Map to Final Model*. (Bailey, S., Hubbard, R. & Waller, D.A., eds.), pp. 149–159, SERC Daresbury Laboratory, Warrington, UK.
22. Kleywegt, G.J., *et al.*, & Jones, T.A. (1994). Crystal structures of cellular retinoic acid binding proteins I and II in complex with all-*trans*-retinoic acid and a synthetic retinoid. *Structure* **2**, 1241–1258.
23. Jones, T.A. & Thirup, S. (1986). Using known substructures in protein model building and crystallography. *EMBO J.* **5**, 819–822.
24. Schneider, G., Lindqvist, Y. & Lundqvist, T. (1990). Crystallographic refinement and structure of ribulose-1,5-bisphosphate carboxylase from *Rhodospirillum rubrum* at 1.7 Å resolution. *J. Mol. Biol.* **211**, 989–1008.
25. Lundqvist, T. & Schneider, G. (1991). Crystal structure of activated ribulose-1,5-bisphosphate carboxylase complexed with its substrate, ribulose-1,5-bisphosphate. *J. Biol. Chem.* **266**, 12604–12611.
26. Sinning, I., *et al.*, & Jones, T.A. (1993). Structure determination and refinement of human alpha class glutathione transferase A1-1, and a comparison with the mu and pi class enzymes. *J. Mol. Biol.* **232**, 192–212.
27. Cameron, A.D., *et al.*, & Jones, T.A. (1995). Structural analysis of human alpha class glutathione transferase A1-1, in the apo form and in complexes with ethacrynic acid and its glutathione conjugate. *Structure* **3**, in press.
28. Korn, A.P. & Rose, D.R. (1994). Torsion angle differences as a means of pinpointing local polypeptide chain trajectory changes for identical proteins in different conformational states. *Protein Eng.* **7**, 961–967.
29. Zou, J.Y. & Mowbray, S.L. (1994). An evaluation of the use of databases in protein structure refinement. *Acta Crystallogr. D* **50**, 237–249.
30. Rees, D.C. (1983). Largest likely values for R factors calculated after phase refinement by non-crystallographic symmetry averaging. *Acta Crystallogr. A* **39**, 916–920.
31. Valegård, K., Murray, J.B., Stockley, P.G., Stonehouse, N.J. & Liljas, L. (1994). Crystal structure of an RNA bacteriophage coat protein–operator complex. *Nature* **371**, 623–626.
32. Bernstein, F.C., *et al.*, & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival-file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
33. Kleywegt, G.J. & Jones, T.A. (1995). A re-evaluation of the crystal structure of chloromuconate cycloisomerase. *Acta Crystallogr. D* **51**, in press.
34. Vriend, G. & Sander, C. (1993). Quality control of protein models: directional atomic contact analysis. *J. Appl. Crystallogr.* **26**, 47–60.

Gerard J Kleywegt and T Alwyn Jones, Department of Molecular Biology, Biomedical Centre, Uppsala University, Box 590, S-751 24 Uppsala, Sweden.