Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

To free or not to free?

Alexandre Urzhumtsev

Centre for Integrative Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS–INSERM-UdS, 1 rue Laurent Fries, Illkirch, BP 10142, 67404, France, and Université de Lorraine, Faculté des Sciences et Technologies, Vandoeuvre-les-Nancy, BP 239, 54506, France

Correspondence e-mail: sacha@igbmc.fr

Refinement of an atomic model is the optimization step that makes the model accurate enough for its structural interpretation. As for any optimization problem, the choice of the target is one of three key decisions to take (two others are the model parameterization and optimization method).

From the early days (Hughes, 1941) a straightforward choice was minimizing the least-squares (LS) discrepancy between the experimental structure-factor amplitudes $\{F^{\rm obs}_{hkl}\}$, $(hkl) \in S$, and the corresponding values $\{F^{\rm calc}_{hkl}\}$ calculated from the atomic model, where S is the complete set of diffraction data. However, macromolecular models may contain errors that cannot be corrected by any choice of parameter values; one example is atoms missing from the model, which is always the case for macromolecules. LS-based refinement of incomplete models may make them worse.

The errors that cannot be removed by changing model parameters can be accounted for statistically. This comes down to the question: how large is the probability (statistical likelihood) to obtain calculated amplitudes equal to the observed ones after necessary corrections are made randomly, for example after the model has been completed randomly with 'atoms' to compensate for the missing ones? In crystallographic refinement, the likelihood is usually approximated by the product of probabilities corresponding to individual structure factors, implying their independence. Defining the maximum-likelihood (ML) target as a negative logarithm of the likelihood, results in a sum over reflections where each term is expressed through the corresponding $F^{\rm obs}_{hkl}$, $F^{\rm calc}_{hkl}$ and two parameters, α_{hkl} and β_{hkl} (or a single parameter $\sigma_{{\rm A},\,hkl}$) that reflect model errors. For the purpose of refinement one needs first to obtain the shape of the ML target, *i.e.* to find the parameter values that define the target, and then search for a better model that minimizes this target.

The values of the parameters α_{hkl}, β_{hkl}, or $\sigma_{{\rm A},\,hkl}$ may be assumed to be the same for reflections in thin resolution shells and can be obtained from comparison of $\{F^{\rm obs}_{hkl}\}$ with $\{F^{\rm calc}_{hkl}\}$ corresponding to the available model. This approach gives accurate parameter estimates for models that have never been refined but not for those refined previously. The reason for this is that structure factors $\{F^{\rm calc}_{hkl}\}$ are calculated through atomic parameters, refinement of which against the full set of structure-factor amplitudes makes the calculated structure factors in this data set S mutually dependent.

A solution was to estimate α_{hkl}, β_{hkl}, or $\sigma_{{\rm A},\,hkl}$, using $\{F^{\rm calc}_{hkl}\}$ from a subset of diffraction data excluded from refinement (Lunin & Skovoroda, 1995; Pannu & Read, 1996; Murshudov *et al.*, 1997). The original goal to split the whole data set into *work* and *test* subsets (Brünger, 1992) was to use the former for refinement and the latter only for model validation. The test subset $S_{\rm test} \subset S$ usually contains reflections chosen randomly and uniformly in space, 5–10% of the total set, and the crystallographic R factor calculated over $S_{\rm test}$ is called $R_{\rm free}$ factor. Structure factors that belong to $S_{\rm test}$, even when calculated from the same model, are much more statistically independent than those in $S_{\rm work}$. Their quasi-independence allows them to be used to better estimate ML-target parameters, a goal that has nothing to do with the original goal of validation.

Pražnikar & Turk (2014) started their article from examples showing that, if a relatively small test set is used, the estimates of α_{hkl} and β_{hkl} may depend on the particular choice of test set reflections and thus make the corresponding refinement unstable. The problem becomes even more important at low resolution where the total number of reflections may be insufficient to extract a representative test data set. This gives rise to the question of how to estimate the ML parameters other than by using the 'free' reflections.

Pražnikar & Turk (2014) suggest an answer to this question. To estimate these parameters, one needs a never-refined model so that all structure factors calculated from it are

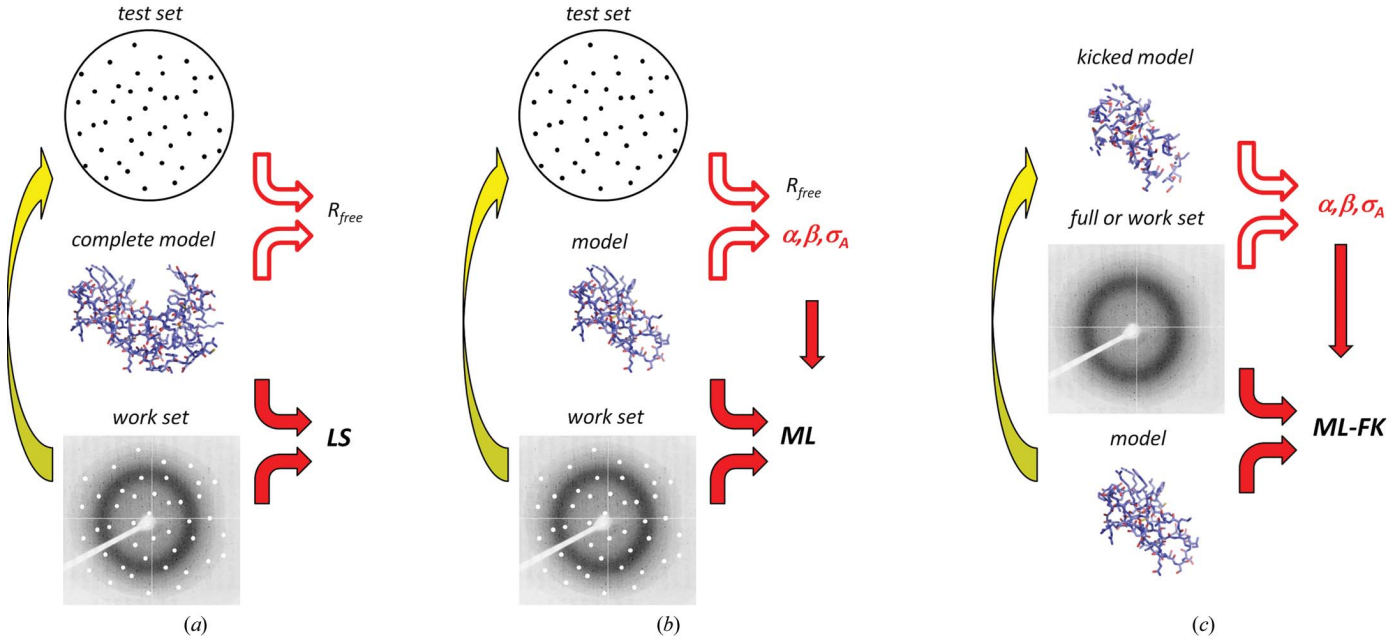

Figure 1
Schematic comparison of different refinement protocols. (*a*) LS refinement of a complete model; LS refinement of an incomplete model may fail. (*b*) ML refinement of an incomplete model with the parameters estimated from the test (free) data set. (*c*) Free-kick ML refinement (Pražnikar & Turk, 2014); their current protocol uses the work set and not the full set of amplitudes – this is due to validation purposes only and not due to refinement requirements.

statistically independent. To obtain an approximation to such a model the authors use the tool that they developed previously to improve crystallographic maps, namely model kicking (Pražnikar *et al.*, 2009). A kicked model is obtained from the initial one by a random shaking ('kicking') of atomic positions. The kicked model is worse than the initial one but this is not important because it is used only to estimate the statistical parameters describing the ML target while the refinement is done using the initial unkicked model (Fig. 1). The delicate question is the size of the kick: too small differences between the kicked and the initial models leave structure factors mutually dependent, whereas too large differences will result in accurate estimates – but for a model irrelevant to the unkicked one, making these estimates useless. The examples given in the article show that finding a good balance is possible, which proves the feasibility of this approach; obviously further studies with data at higher or lower resolutions, with less complete models or with other kinds of model imperfections are required.

Now the two goals above can be separated: model validation can be done, as previously, using the test set $S_{\rm test}$ while estimating the statistical parameters may be done using *any* data set representative enough. If reflections from $S_{\rm test}$ are left out to remove the bias from model validation, then, for example, $S_{\rm work}$ can be taken as such a representative set: being much larger than $S_{\rm test}$ it assures accurate estimates of the statistical parameters, making the ML-refinement protocol robust and the results superior to those from the previous protocols as the authors show. In this sense, the current idea by Pražnikar & Turk presents a significant methodological step forward.

In fact their idea has consequences going much further. Excluding a part of the reflections from refinement means a loss of experimental information. Moreover, $S_{\rm test}$ reflections have to be excluded not only from refinement but also, in principle, from the calculation of the Fourier maps used for model rebuilding; these corrupted maps may be another source of eventual model error (or of a bias in $R_{\rm free}$ if reflections are not excluded). Pražnikar & Turk claim that further development of alternative validation techniques will eventually make calculation of $R_{\rm free}$ unnecessary, being substituted by other approaches. This would mean the possibility of refinement against a *full* set of collected structure-factor amplitudes using the accurate and robust ML-approach suggested in their article.

References

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Hughes, E. W. (1941). *J. Am. Chem. Soc.* **63**, 1737–1752.
Lunin, V. Yu. & Skovoroda, T. P. (1995). *Acta Cryst.* A**51**, 880–887.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Pannu, N. S. & Read, R. J. (1996). *Acta Cryst.* A**52**, 659–668.
Pražnikar, J., Afonine, P.V., Gunčar, G., Adams, P. D. & Turk, D. (2009). *Acta Cryst.* D**65**, 921–931.
Pražnikar, J. & Turk, D. (2014). *Acta Cryst.* D**70**, 3124–3134.