Recent developments for the efficient crystallographic refinement of macromolecular structures

Axel T Brünger*†‡, Paul D Adams† and Luke M Rice†

Macromolecular crystallographic refinement has recently been made more efficient by the use of cross-validated maximum likelihood targets and torsion-angle molecular dynamics simulated annealing. In combination with automated model building methods, the amount of manual intervention required to complete and refine a structure is greatly reduced.

Addresses
*Howard Hughes Medical Institute and †Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511, USA
‡e-mail: brunger@laplace.csb.yale.edu

Current Opinion in Structural Biology 1998, **8**:606–611

http://biomednet.com/elecref/0959440X00800606

Introduction

The ultimate goal of refinement is to simultaneously optimize the agreement of an atomic model with both observed diffraction data and *a priori* chemical information. The target function used for this optimization normally depends on many parameters and, most importantly, on atomic coordinates. The large number of adjustable parameters makes the target function very complicated and results in the multiple minima problem — the target function contains many local minima in addition to the global minimum. This complexity tends to defeat gradient-descent optimization techniques, such as conjugate gradient or least-squares methods [1]. The challenges of crystallographic refinement arise not only from the high dimensionality of the parameter space, but also from the phase problem. For new crystal structures, initial electron density maps must be computed from a combination of observed diffraction amplitudes and experimental phases. Frequently, these phases are of a poorer quality and/or lower resolution than the observed amplitudes. A different problem arises when structures are solved by molecular replacement, [2,3] a technique that uses a similar structure as a search model to obtain the initial phases. In this case, the resulting electron density maps can be severely 'model biased', that is, they sometimes seem to confirm features of the search model without providing clear evidence of actual differences between it and the true crystal structure. In both cases, initial atomic models usually contain significant errors and require several cycles of refinement interspersed with manual intervention in order to correct gross errors. Here, we review the recent advances in the field of crystallographic refinement that decrease manual intervention and increase efficiency, arriving at more correct models in a shorter amount of time.

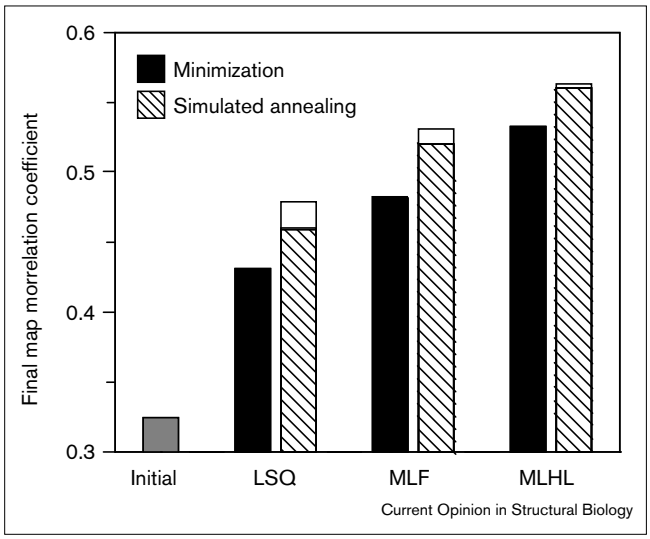
Simulated annealing [4] is an optimization technique that is particularly well suited to overcoming the multiple minima problem, thereby reducing the manual intervention required during refinement [5]. Unlike gradient-descent methods, simulated annealing can cross barriers between minima and, thus, can explore a greater volume of the parameter space in order to find better models (deeper minima). Many examples have shown that simulated annealing refinement, starting from initial models obtained by standard crystallographic techniques, produces significantly better final models compared to those produced by least-squares or conjugate-gradient minimization [5–9]. Since its introduction [5], crystallographic refinement has undergone major improvements in three principal areas: the measure of model; quality, the search of the parameter space; and the target function.

For crystallographic refinement, the introduction of cross-validation in the form of the free *R*-value [10] has significantly reduced the danger of overfitting the diffraction data. Cross-validation also produces more realistic coordinate error estimates based on the Luzzati or σ_A methods [11]. The complexity of the conformational space can be reduced by the introduction of torsion-angle refinement methods [9,12], which decreases the number of adjustable parameters that describe a model by approximately 10-fold. The target function has been improved using a maximum likelihood approach that takes into account model error, model incompleteness and errors in the experimental data [13,14•,15••]. Cross-validation of the parameters for the maximum likelihood target function is essential in order to obtain better results than conventional target functions [13,16•,17•].

Cross-validation

Cross-validation plays a fundamental role in all the maximum likelihood target functions described to date. A few remarks about this method are therefore warranted here. For cross-validation, the diffraction data are divided into two sets — a large working set (typically comprising 90% of the data) and a complementary test set (comprising the remaining 10%). The diffraction data in the working set are used in the normal crystallographic refinement process, whereas the test set data are not. The cross-validated (or 'free') *R*-value computed using the test set data is a more faithful indicator of model quality. It provides a more objective guide during the model building and refinement process than the conventional *R*-value. It also indicates whether the introduction of additional parameters (e.g. water molecules, the relaxation of noncrystallographic symmetry restraints or multiconformer models) improves the quality of the model or, instead, increases overfitting [18•].

Figure 1



Simulated annealing produces better models than extensive conjugate gradient minimization. Map correlation coefficients were computed before and after refinement against the native penicillopepsin diffraction data [39] for a polyalanine model derived from *Rhizopuspepsin* [40]. Correlation coefficients are between σ_A-weighted maps calculated from each model and the published penicillopepsin structure. The observed penicillopepsin diffraction data were in space group C2, with cell dimensions a = 97.37 Å, b = 46.64 Å, c = 65.47 Å and β = 115.4°. All refinements were carried out using diffraction data from the lowest resolution limit of 22.0 Å up to 2.0 Å. The MLHL refinements used single isomorphous phases from a $K_3UO_2F_5$ derivative of the penicillopepsin crystal structure, which covered a resolution range of 22.0 Å to 2.8 Å. The simulated annealing refinements were repeated five times with different initial velocities. The numerical averages of the map correlation coefficients for the five refinements are shown as the hashed bars. The best map correlation coefficients from simulated annealing are shown as the white bars.

Target functions

Crystallographic refinement is a search for the global minimum of the target:

$$E = E_{\text{chem}} + w_{\text{xray}} E_{\text{xray}} \qquad (1)$$

where E is a function of the parameters of an atomic model. E_{chem} comprises empirical information about chemical interactions; it is a function of all atomic positions, describing covalent (bond lengths, bond angles, torsion angles, chiral centers and planarity of aromatic rings) and nonbonded (intramolecular as well as intermolecular and symmetry related) interactions [19]. E_{xray} is related to the difference between the observed and calculated data and w_{xray} is a weight appropriately chosen to balance the gradients (with respect to atomic parameters) arising from the two terms.

Maximum likelihood refinement targets

The traditional form of E_{xray} consists of the crystallographic residual LSQ, defined as the sum of the squared differences between the observed ($|\mathbf{F}_o|$) and calculated ($|\mathbf{F}_c|$) structure-factor amplitudes for a particular atomic model:

$$E_{\text{xray}} = \text{LSQ} = \sum_{hkl \in \text{working set}} (|\mathbf{F}_o| - \kappa|\mathbf{F}_c|)^2 \qquad (2)$$

where hkl are the indices of the reciprocal lattice points of the crystal and κ is a relative scale factor.

Minimization of the LSQ can improve the atomic model, but can also lead to the accumulation of systematic errors in the model by fitting noise in the diffraction data [20]. The least-squares residual is a limiting case of the maximum likelihood theory and is only justified if the model is nearly complete and error free. These assumptions are usually violated during the initial stages of refinement. Improved targets for the refinement of incomplete, error-containing models have been obtained using the more general maximum likelihood formulation [13,14•,15••]. The goal of this method is to determine the likelihood of the model given estimates of the errors in the model and the measured intensities [15••,17•].

In general, an empirical approach is used to estimate errors in the model. The σ_A distribution for the model is a good source of such error estimates [21,22]. As the model improves, however, the σ_A distribution changes and new error estimates must be obtained. Refinement methods that improve the model $\mathbf{F}_c$ structure factors will therefore have a beneficial effect on error estimates and this will then enhance the next refinement cycle. Thus, powerful optimization methods and maximum likelihood targets are expected to interact in a synergistic fashion (Figure 1).

In order to achieve improvement over the least-squares residual (Equation 2), cross-validation was found to be essential [13,16•] for the estimation of model incompleteness and errors. Maximum likelihood refinement without cross-validation gives much poorer results, as indicated by higher free R-values, higher $R_{\text{free}} - R$ differences and higher phase errors compared to the final refined structure [16•]. It should be noted that when using the cross-validated maximum likelihood target, the final normal R-value is, in general, increased compared to refinements using the least-squares target. This is a consequence of the reduction in the overfitting achieved by the maximum likelihood target.

For many structures, some initial experimental phase information is available from either isomorphous heavy-atom replacement or multiwavelength anomalous diffraction methods. These phases represent additional observations that can be incorporated in the refinement target. The maximum likelihood formulation naturally extends itself to the incorporation of this information [14•,15••]. Tests have shown that the addition of experimental phase information greatly improves the results of refinement [23••,24••].

Pannu and Read [13] have developed an efficient Gaussian approximation in the case of structure-factor

amplitudes with no prior phase information, termed the 'MLF' target function. In the limit of a perfect model, MLF reduces to the traditional least-squares residual (Equation 2) with $1/\sigma^2_O$ weighting. In the case where prior phase information is included, integration over the phase angles is carried out numerically and is termed the 'MLHL' target [23••]. A maximum likelihood function that expresses the probability distribution in terms of observed intensities has also been developed and is termed 'MLI' [13].

Real-space refinement

In real-space refinement, the target that is minimized is the agreement between the current electron density map and the electron density calculated from the model. This method has a long history in macromolecular refinement [12], but has become less frequently used since the introduction of efficient reciprocal-space refinement methods [19]. Interest in the method has been renewed, however, after recent improvements to the real-space target function [25•] and the increasing availability of high quality experimental phases from multiple wavelength anomalous diffraction phasing or noncrystallographic symmetry averaging. Tests have shown the method to be useful in the initial stages of refinement, especially when alternated with reciprocal-space refinement [26•].

Simulated annealing

Annealing denotes a physical process wherein a solid is heated until all the particles randomly arrange themselves in a liquid phase and then is slowly cooled so that all the particles arrange themselves in the lowest energy state. By formally defining the target E (Equation 1) to be the equivalent of the potential energy of the system, one can simulate such an annealing process [4]. There is no guarantee that simulated annealing will find the global minimum [27]. Compared to conjugate-gradient minimization, however, whereby search directions must follow the gradient, simulated annealing achieves more optimal solutions by allowing motion against the gradient. The likelihood of uphill motion is determined by a control parameter referred to as temperature. The higher the temperature, the more likely it is that simulated annealing will overcome barriers. It should be noted that the simulated annealing temperature normally has no physical meaning and merely determines the likelihood of overcoming barriers of the target function (Equation 1). The annealing schedule can, in principle, be any function of the simulation step (or 'time' domain). The two most commonly used protocols are linear slow cooling or constant temperature followed by quenching. An advantage is obtained with slow cooling [28]. The duration of the annealing schedule is another parameter. Too short a protocol does not allow for sufficient sampling of the conformational space. Too long a protocol may waste computer time since it is more efficient to run multiple trials as opposed to one long refinement protocol [29••].

Torsion-angle dynamics

Although Cartesian (i.e. flexible bond lengths and bond angles) molecular dynamics places restraints on bond lengths and bond angles (through E_{chem}, see Equation 1), one might want to implement these restrictions as constraints, that is, fixed bond lengths and bond angles. This is supported by the observation that deviations from ideal bond lengths and bond angles are usually small in macromolecular X-ray crystal structures. Indeed, fixed-length constraints have been applied to crystallographic refinement using least-squares minimization [12]. It is only recently, however, that efficient and robust algorithms have become available for molecular dynamics in torsion-angle space [9,30–32]. An approach is chosen that retains the Cartesian coordinate formulation of the target function and its derivatives, allowing them to be calculated in a relatively straightforward manner that can be applied to any macromolecule or complex [9]. In this formulation, the expression for the acceleration becomes a function of positions and velocities. Iterative equations of motion for constrained dynamics in this formulation can be derived and solved by finite difference methods [33]. This method is numerically very robust and has a significantly increased radius of convergence in crystallographic refinement compared to Cartesian molecular dynamics [9].

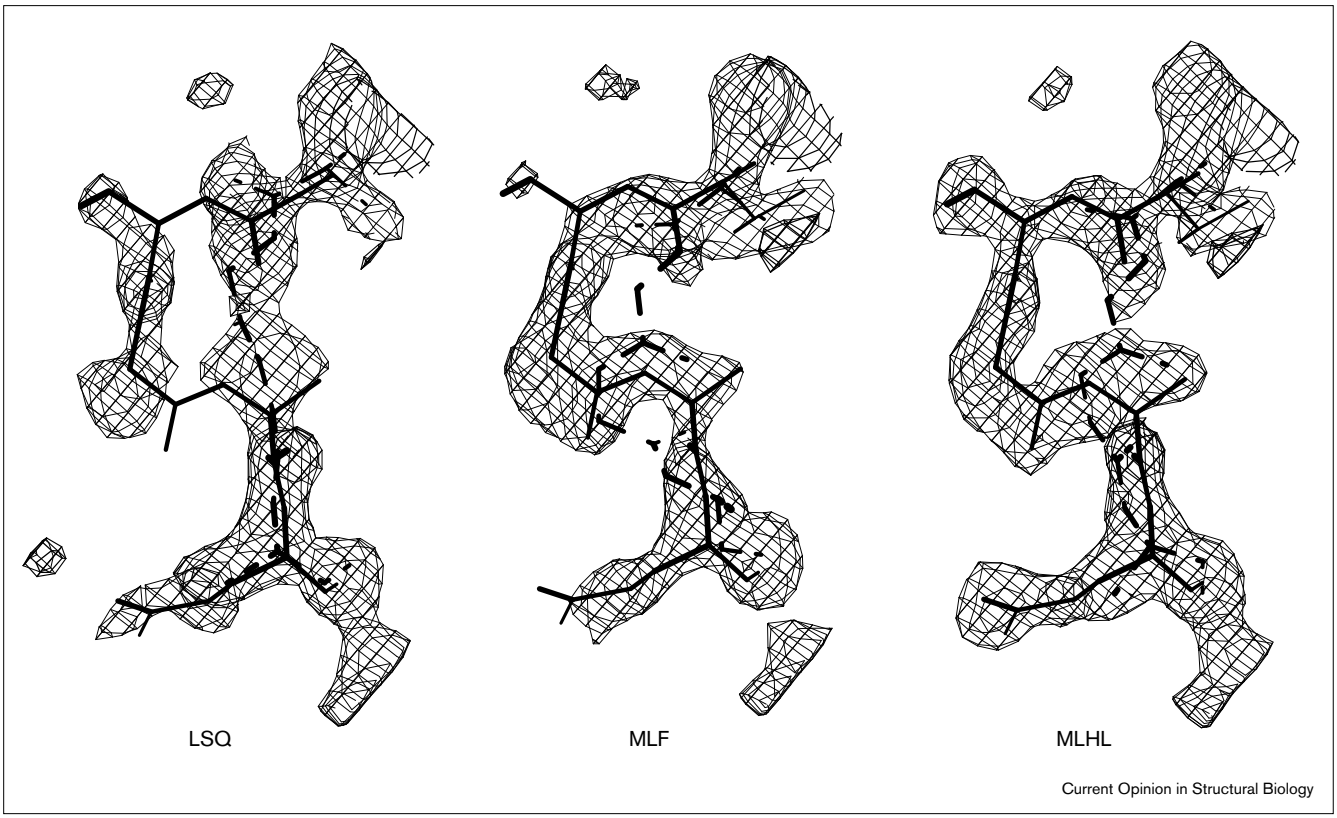
Combined torsion-angle dynamics simulated annealing and maximum likelihood targets

In a realistic test case [24••], a series of models of the aspartic proteinase penicillopepsin was generated from homologous structures present in the Protein Data Bank. The sequence identity among these structures ranged from 100 to 25%, thus providing a set of models with increasing coordinate error when compared to the refined structure of penicillopepsin. These models, after the truncation of all residues to alanine, were each used as search models in molecular replacement against the native penicillopepsin diffraction data. In all cases, the correct placement of the model in the penicillopepsin unit cell was found.

Both conjugate-gradient minimization and simulated annealing were carried out in order to compare the performances of LSQ (the least-squares residual), MLF (the maximum likelihood target using amplitudes) and MLHL (the maximum likelihood target using amplitudes and experimental phase information). In the case of MLHL, phases from single isomorphous replacement were used. A very large number of conjugate-gradient cycles were carried out in order to make the computational requirements equivalent for both minimization and simulated annealing. The conjugate-gradient minimizations were converged, that is, there was no change when further cycles were carried out.

For a given target function, simulated annealing always outperformed minimization (Figure 1). For a given starting model, the maximum likelihood targets outperformed the

Figure 2



Maximum likelihood targets significantly decrease model bias in simulated annealing refinement. σ_A-weighted electron density maps contoured at 1.25σ were produced for models by simulated annealing refinement with different targets. Residues 233 to 237 are shown with the published penicillopepsin crystal structure [39] represented by solid lines and the model with the lowest free *R*-value from five independent refinements in dashed lines.

least-squares residual target for both minimization and simulated annealing, producing models with lower phase errors and higher map correlation coefficients when compared to the published penicillopepsin crystal structure (Figure 1). This improvement is illustrated by σ_A-weighted electron density maps obtained from the resulting models (Figure 2). The incorporation of experimental phase information significantly improved the refinement, despite the ambiguity in the single isomorphous replacement phase probability distributions. Thus, the most efficient refinement will make use of torsion-angle dynamics simulated annealing and prior phase information in the MLHL maximum likelihood target function.

Multi-start refinement and structure-factor averaging

Multiple simulated annealing refinements starting from the same model, termed 'multi-start' refinement, will generally produce somewhat different structures. Even well-refined structures will show some variation, consistent with the estimated coordinate error of the model. More importantly, the poorer the model, the more variation is observed [6]. Some of the models resulting from multi-start refinement may be better than others, for example, as judged by the free *R*-value. Thus, if computer time is available, multi-start refinement has several advantages. A more optimal single model than that produced by a single simulated annealing calculation can usually be obtained. Furthermore, each separate model coming from a multi-start refinement fits the data slightly differently. This could be the result of intrinsic flexibility within the molecule or the result of model building error. Regions in the starting model that contain significant errors often show increased variability after multi-start refinement and a visual inspection of the ensemble of models produced can be helpful in identifying these incorrectly modeled regions.

In order to better identify the correct conformation, structure factors from each of the models can be averaged [29••]. This averaging tends to reduce the effect of local errors (noise), which are presumably different for each member of the family. The average structure factor can produce phases that contain less model bias than phases computed from a single model. It should also produce better estimates of errors in the model for maximum likelihood targets and σ_A-weighted electron density maps because $\mathbf{F}_c$ is used in the computation of these parameters. As it is inherently a noise-reducing technique, multi-start refinement followed by structure-factor

averaging should be most useful in situations in which there is significant noise, namely when the data to parameter ratio is low (e.g. if only moderate resolution diffraction data are available). The program wARP [34•] makes efficient use of structure-factor averaging in the context of phase improvement and automated model completion (see below).

Efficient model building

Efficient crystallographic refinement also requires efficient methods for building an initial model and rebuilding the model during refinement. Programs such as O [35] have removed much of the subjectivity of manual rebuilding by incorporating information from databases of known structures. Attempts are being made, however, to make the process of model building even more automated. In one instance, the use of information from known structures, combined with automated map interpretation, has been used to build an initial model with little or no manual intervention from the user [36•]. These methods have been seen to work even at relatively low resolution (d_{min}~3.0 Å). In other work, refinement in combination with the automated identification and interpretation of potential atomic sites in an electron density map is being used to try to gradually build a more complete model from some small starting fragments [34•,37]. This method is currently limited by the need for relatively high resolution data ($d_{min} < 2.0$ Å).

Conclusions

Simulated annealing has significantly improved the efficiency of crystallographic refinement. A case in point is the combination of torsion-angle molecular dynamics with cross-validated maximum likelihood targets. These two independent developments interact synergistically, producing less model bias than any other method to date. This combined method significantly increases the radius of convergence, allowing the productive refinement of poor initial models, for example, those obtained by weak molecular replacement solutions [9,16•,24••].

In the future, we can look forward to more automated methods for the building and rebuilding of models. These will be combined with powerful maximum likelihood target functions and torsion-angle representations of molecules during refinement. The high quality phases available from multiple wavelength anomalous diffraction phasing will be routinely incorporated into the refinement process, resulting in minimally biased models. The combination of all these methods will result in more accurate models that will be arrived at more rapidly, allowing a more detailed interpretation of the structure and therefore the underlying biology.

Acknowledgements

Many of the computational developments discussed in this review are implemented in the software suite Crystallography and NMR System [38•]. A prerelease of this software suite is available upon request. LMR is a Howard Hughes Medical Institute predoctoral fellow. This work was funded in part by grants from the National Science Foundation to ATB (BIR 9514819 and ASC 93-181159).

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Press WH, Flannery BP, Teukolosky SA, Vetterling WT: **Minimization or maximization of functions.** In *Numerical Recipes*. Cambridge: Cambridge University Press; 1986:498-546.

2. Hoppe W: **Die Faltmolekülmethode - eine neue Methode zur Bestimmung der Kristallstruktur bei ganz oder teilweise bekannter Molekülstruktur.** *Acta Crystallogr* 1957, **10**:750-751.

3. Rossmann MG, Blow DM: **The detection of sub-units within the crystallographic asymmetric unit.** *Acta Crystallogr A* 1962, **15**:24-51.

4. Kirkpatrick S, Gelatt CD, Vecchi MP Jr: **Optimization by simulated annealing.** *Science* 1983, **220**:671-680.

5. Brünger AT, Kuriyan J, Karplus M: **Crystallographic R factor refinement by molecular dynamics.** *Science* 1987, **235**:458-460.

6. Brünger AT: **Crystallographic refinement by simulated annealing: application to a 2.8 Å resolution structure of aspartate aminotransferase.** *J Mol Biol* 1988, **203**:803-816.

7. Kuriyan J, Brünger AT, Karplus M, Hendrickson WA: **X-ray refinement of protein structures by simulated annealing: test of the method on myohemerythrin.** *Acta Crystallogr A* 1989, **45**:396-409.

8. Fujinaga M, Gros P, van Gunsteren WF: **Testing the method of crystallographic refinement using molecular dynamics.** *J Appl Crystallogr* 1989, **22**:1-8.

9. Rice LM, Brünger AT: **Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement.** *Proteins* 1994, **19**:277-290.

10. Brünger AT: **The free R value: a novel statistical quantity for assessing the accuracy of crystal structures.** *Nature* 1992, **355**:472-474.

11. Kleywegt GJ, Brünger AT: **Cross-validation in crystallography: practice and applications.** *Structure* 1996, **4**:897-904.

12. Diamond R: **A real-space refinement procedure for proteins.** *Acta Crystallogr A* 1971, **27**:436-452.

13. Pannu NS, Read RJ: **Improved structure refinement through maximum likelihood.** *Acta Crystallogr A* 1996, **52**:659-668.

14. • Murshudov GN, Vagin AA, Dodson EJ: **Refinement of macromolecular structures by the maximum-likelihood method.** *Acta Crystallogr D* 1997, **53**:240-255.
A description of the implementation and application of a maximum likelihood target for crystallographic refinement within the program REFMAC.

15. •• Bricogne G: **Bayesian statistical viewpoint on structure determination: basic concepts and examples.** *Methods Enzymol* 1997, **276**:361-423.
A comprehensive review of Bayesian theory, which underlies many crystallographic procedures.

16. • Adams PD, Pannu NS, Read RJ, Brünger AT: **Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement.** *Proc Natl Acad Sci USA* 1997, **94**:5018-5023.
This paper describes the first application of maximum likelihood refinement in combination with simulated annealing. The combination of the two methods was shown to be more powerful than either method on its own.

17. • Read RJ: **Model phases: probabilities and bias.** *Methods Enzymol* 1997, **278**:110-128.
A review of the use of structure-factor probability theory in the reduction of model bias. These theories form the basis for many of the maximum likelihood targets developed to date.

18. • Brünger AT: **Free R value: cross-validation in crystallography.** *Methods Enzymol* 1997, **277**:366-396.
A comprehensive review of the use of cross-validation in crystallography.

19. Hendrickson WA: **Stereochemically restrained refinement of macromolecular structures.** *Methods Enzymol* 1985, **115**:252-270.

20. Silva AM, Rossmann MG: **The refinement of southern bean mosaic virus in reciprocal space.** *Acta Crystallogr B* 1985, **41**:147-157.

21. Read RJ: **Improved Fourier coefficients for maps using phases from partial structures with errors.** *Acta Crystallogr A* 1986, **42**:140-149.

22. Read RJ: **Structure-factor probabilities for related structures.** *Acta Crystallogr A* 1990, **46**:900-912.

23. •• Pannu NS, Murshudov GN, Dodson EJ, Read RJ: **Incorporation of prior phase information strengthens maximum likelihood structural refinement.** *Acta Crystallogr D* 1998, in press.
The incorporation of experimental phase information within the framework of maximum likelihood is described and applied to crystallographic refinement. Test results show a significant improvement over the use of amplitude information alone, even when relatively poor experimental phases are used.

24. •• Adams PD, Pannu NS, Read RJ, Brünger AT: **Extending the limits of molecular replacement through combined simulated annealing and maximum likelihood refinement.** *Acta Crystallogr D* 1998, in press.
A realistic test of optimization methods and crystallographic targets was performed using homologous models for a known structure. It is clearly shown that torsion-angle dynamics simulated annealing in combination with maximum likelihood is the most powerful refinement method currently available.

25. • Blanc E, Chen Z, Chapman MS: **Real-space refinement using Rsref.** In *Direct Methods for Solving Macromolecular Structures*. Edited by Fortier S. Netherlands: Kluwer Academic Publishers; 1998:513-519.
A brief background to real-space refinement methods and the implementation of a target that accounts for the resolution of the experimental data are described.

26. • Chapman MS, Blanc E: **Potential use of real-space refinement in protein structure determination.** *Acta Crystallogr D* 1997, **53**:203-206.
Real-space refinement was applied to a realistic refinement problem. The test calculations suggest ways in which real-space refinement may be best combined with reciprocal-space methods for the refinement process.

27. Laarhoven PJM, Aarts EHL (Eds): *Simulated Annealing: Theory and Applications*. Dordrecht: Reidel Publishing Company; 1987.

28. Brünger AT, Krukowski A, Erickson J: **Slow-cooling protocols for crystallographic refinement by simulated annealing.** *Acta Crystallogr A* 1990, **46**:585-593.

29. •• Rice LM, Shamoo Y, Brünger AT: **Phase improvement by multi-start simulated annealing refinement and structure factor averaging.** *J Appl Crystallogr* 1998, in press.
Multi-start refinement using simulated annealing was combined with structure-factor averaging in order to overcome significant errors in a real refinement problem.

30. Bae D-S, Haug EJ: **A recursive formulation for constrained mechanical system dynamics. Part I. Open loop systems.** *Mech Struct and Mach* 1987, **15**:359-382.

31. Bae D-S, Haug EJ: **A recursive formulation for constrained mechanical system dynamics. Part II. Closed loop systems.** *Mech Struct and Mach* 1988, **15**:481-506.

32. Jain A, Vaidehi N, Rodriguez G: **A fast recursive algorithm for molecular dynamics simulation.** *J Comp Phys* 1993, **106**:258-268.

33. Abramowitz M, Stegun I: *Handbook of Mathematical Functions*. New York: Dover Publications; 1968.

34. • Perrakis A, Sixma TK, Wilson KS, Lamzin VS: **wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models.** *Acta Crystallogr D* 1997, **53**:448-455.
Structure-factor averaging was used in combination with the automated interpretation of electron density maps in order to improve experimentally derived phases.

35. Jones TA, Zou JY, Cowan SW, Kjeldgaard M: **Improved methods for building protein models in electron density maps and the location of errors in these models.** *Acta Crystallogr A* 1991, **47**:110-119.

36. • Crystallographic Computing 7: proceedings from the Macromolecular Crystallography Computing School on World Wide Web URL: http://www.sdsc.edu/projects/Xtal/IUCr/School96/IUCr.html
New methods are described that can be used to automate model building from an electron density map.

37. Lamzin VS, Wilson KS: **Automated refinement of protein models.** *Acta Crystallogr D* 1993, **49**:129-147.

38. • Brünger AT, Adams PD, Clore GM, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS, Read RJ *et al.*: **Crystallography and NMR system (CNS): a new software system for macromolecular structure determination.** *Acta Crystallogr D*, 1998, **54**:905–921.
A description of a new program for structure determination that incorporates many of the advances described in this review.

39. Hsu IN, Delbaere LTJ, James MNG, Hoffman T: **Penicillopepsin from *Penicillium janthinellum* crystal structure at 2.8 Å and sequence homology with porcine pepsin.** *Nature* 1977, **266**:140-145.

40. Suguna K, Bott RR, Padlan EA, Subramanian E, Sheriff S, Cohen GH, Davies DR: **Structure and refinement at 1.8 Å resolution of the aspartic proteinase from *Rhizopus chinensis*.** *J Mol Biol* 1987, **196**:877-900.