Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Iterative-build OMIT maps: map improvement by iterative model building and refinement without model bias

Thomas C. Terwilliger,[a]* Ralf W. Grosse-Kunstleve,[b] Pavel V. Afonine,[b] Nigel W. Moriarty,[b] Paul D. Adams,[b] Randy J. Read,[c] Peter H. Zwart[b] and Li-Wei Hung[a]

[a]Los Alamos National Laboratory, Los Alamos, NM 87545, USA, [b]Lawrence Berkeley National Laboratory, One Cyclotron Road, Building 64R0121, Berkeley, CA 94720, USA, and [c]Department of Haematology, University of Cambridge, Cambridge CB2 0XY, England

Correspondence e-mail: terwilliger@lanl.gov

Received 30 October 2007
Accepted 12 February 2008

A procedure for carrying out iterative model building, density modification and refinement is presented in which the density in an OMIT region is essentially unbiased by an atomic model. Density from a set of overlapping OMIT regions can be combined to create a composite 'iterative-build' OMIT map that is everywhere unbiased by an atomic model but also everywhere benefiting from the model-based information present elsewhere in the unit cell. The procedure may have applications in the validation of specific features in atomic models as well as in overall model validation. The procedure is demonstrated with a molecular-replacement structure and with an experimentally phased structure and a variation on the method is demonstrated by removing model bias from a structure from the Protein Data Bank.

1. Introduction

Model bias is a continuing problem in macromolecular crystallography. It results from using an atomic model to calculate crystallographic phases, in which case the resulting electron-density map will tend to have the features present in the model even if they are not actually present in the structure (Ramachandran & Srinivasan, 1961; Read, 1986; Bhat, 1988; Hodel *et al.*, 1992; Adams *et al.*, 1999; Kleywegt, 2000). Once an atomic model has been refined, model bias can be indirect as well as direct because the positions and other parameters describing correctly placed atoms are adjusted during refinement in order to compensate for the incorrectly placed atoms. Consequently, even if the incorrectly placed atoms are removed from the model before the calculation of phases, a memory of their positions can remain and the resulting map can retain incorrect features. Refinement of a model omitting incorrectly placed atoms should reduce this indirect bias, but there remains the question of how extensive the refinement must be in order to reverse all compensating adjustments. Model bias can make the interpretation of electron-density maps difficult, particularly in cases in which molecular replacement (Rossmann, 1972) is used to solve a structure.

There are many ways that model bias can be defined. In this work, model bias refers to the situation where a map has peaks of density resembling atomic density that arise only from atoms in the working model used in phasing or density modification and not from the presence of atoms in the real structure.

© 2008 International Union of Crystallography
Printed in Singapore – all rights reserved

Many methods have been developed to reduce the effects of model bias. These fall into two general classes. The first class consists of methods to remove the model bias after it has been introduced. The second class consists of methods in which a model is never introduced in a particular location of a map, so that in that location there is never any model bias.

The reason for the use of an approach in which model bias is introduced and then removed is that the process of building and refining a model greatly improves the overall accuracy of the phase information and this is often required in order to obtain a structure at all. Once an initial structure has been determined, it becomes important to know which details are correct and the removal of model bias can be important for this. A well established method for reducing model bias is to calculate a map using σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ coefficients (Read, 1986; Lunin *et al.*, 2002), typically using coefficients m and D calculated using a test set of reflections (Urzhumtsev *et al.*, 1996; Pannu & Read, 1996). The σ_A procedure yields high-quality maps, but they can retain some model bias because the procedure is based on the assumption of random errors in the model, while the actual errors after model refinement are typically correlated (Read, 1997). Simulated-annealing OMIT maps (Hodel *et al.*, 1992; Brünger *et al.*, 1998), 'kicked' OMIT maps (Guncar *et al.*, 2000) and model rebuilding with randomization (Zeng *et al.*, 1997; Reddy *et al.*, 2003) can all reduce model bias by removing, at least to some extent, the memory of some or all of the atoms in the model. Prime-and-switch density modification (Terwilliger, 2004) reduces the effects of model bias in a different way. This approach uses a model to calculate phases, yielding a map that is biased but nearly correct, and then using a characteristic of this map that is relatively uncorrelated with the model, such as the flatness of the solvent region, to calculate less-biased phases. Although these procedures can reduce the effects of model bias, they all have the disadvantage that it is not possible to know for certain that model bias has been removed. Additionally, techniques that involve a solvent mask have some potential for model bias as the mask can be influenced by the model.

The second class of methods are those that never introduce model bias and can consequently yield a higher degree of confidence in electron density that is in the same location as an atom in a model. The simplest method in this class is an OMIT map in which the atoms in a particular region of the map are never included in phase calculation or in the refinement of the other atoms in the structure (Bhat, 1988). In a molecular-replacement structure determination, this type of OMIT map can easily be calculated early in the process, after the molecule has been placed in the correct location in the unit cell but before any atomic refinement has occurred. In this case, the memory of the presence of the omitted atoms is likely to be minimal and the resulting map is unbiased in the OMIT region. The disadvantage of OMIT maps is that they are typically very noisy and consequently difficult to interpret. Additionally, once the structure (with all atoms) has been refined, OMIT maps are of less utility because of the indirect model bias described above.

Another approach that avoids introducing model bias is to carry out the usual process of iterative model building and refinement but to avoid building a particular part of the model. The part of the model deliberately not built might be a ligand, a side chain or any other part of the model. This approach is commonly used for poorly defined portions of electron-density maps and has been used in some cases specifically to obtain unbiased information (James *et al.*, 1980). Such poorly defined regions are typically not interpreted until the improvement in the rest of the model is sufficient to yield clear electron density for a model in those regions. This approach has recently been extended into a systematic procedure ('ping-pong refinement') that allows each of the side chains of a structure to be built into density that is unbiased while gradually building up a complete atomic model (Hunt & Deisenhofer, 2003).

Here, we combine the ideas of OMIT maps and ping-pong refinement in an 'iterative-build OMIT' procedure for obtaining a partial or complete composite electron-density map that is essentially free from model bias yet that benefits from the power of iterative model building and refinement.

2. Methods

2.1. Calculation of an iterative-build OMIT map for a single OMIT region

Obtaining an iterative-build OMIT map for a single region of the asymmetric unit of a crystal is in principle quite straightforward. Firstly, the OMIT region is defined as a contiguous region representing part of the asymmetric unit. A border (typically 2 Å thick) is then added to this region. Finally, an iterative model-building, refinement and density-modification procedure is carried out in a standard fashion (Terwilliger, 2003), except that any atom that is located within the OMIT or border regions is given zero occupancy in all calculations. We call this overall process the 'iterative-build OMIT procedure'. In this procedure, all atoms in the OMIT region are included in geometric restraint calculations and are included in all rebuilding steps. In this way, the geometry of the model is retained. Owing to their zero occupancy values, however, the atoms in the OMIT region do not contribute to the structure-factor calculation. This prevents direct model bias. This procedure also prevents indirect model bias in the density calculated within the OMIT region, as the parameters of atoms outside the OMIT region are never adjusted to compensate for electron density of atoms in the OMIT or border regions.

To ensure that model bias is not indirectly introduced through the application of NCS, in the *RESOLVE* density-modification steps no NCS-based target electron density is transferred into the OMIT or border regions. This is accomplished by defining the boundaries of the OMIT region and specifying that no NCS information is to be transferred into this region. As density modification with NCS is performed point by point, using the density from $N - 1$ copies as a target for density modification for the remaining copy (Terwilliger,

2002), it is straightforward to leave out NCS information for all points in the OMIT region. Once the final electron-density map has been obtained from the iterative-build OMIT procedure, the OMIT region (but not the boundary region or other parts of the map) will have essentially no bias arising from structure-factor contributions from the atomic model. As the parameters describing the NCS relationships are refined, there is in principle some possibility of model information being transferred between NCS regions. As in the case of rigid-body refinement of an MR model, however, it is not likely that significant information about the density in a particular location in the map is going to be transmitted through the very small number of parameters refined in this step. In the standard *PHENIX* model-building procedure in the presence of NCS, all building is performed independently for all copies and in a specific step the structure of each NCS copy is transformed to match each other one and the best parts of the structure from each NCS copy are kept. This step is not carried out when an OMIT map is constructed, so that no information about the structure within the OMIT region is transmitted between the NCS copies. NCS restraints are applied during model building with the OMIT procedure. The effect of this is that the (zero-occupancy) atoms in the OMIT region may be placed in incorrect positions because of the NCS restraints from copies outside the OMIT region.

Other potential sources of model bias are the bulk-solvent correction and geometric restraints. In the procedure carried out here, a bulk-solvent model and geometric restraints are applied throughout, as not applying them would lead to a poorer atomic model in the regions outside the OMIT boundaries and would therefore result in an OMIT map with greater artefacts and less utility.

We note that there are some circumstances in which the density within an OMIT region can be affected by the positions of atoms in the model inside the OMIT region. For example, an atom inside an OMIT region may be bonded to an atom outside the OMIT region, so that the positions are correlated. Similarly, the position of an atom inside the OMIT region could affect the placement of a solvent molecule outside the OMIT region or the allowed conformations of a side chain outside the OMIT region. Furthermore, during automated model building, large units (helices or strands) may be placed based on density both inside and outside the OMIT region. In all these cases, however, this coupling between atoms inside and outside the OMIT region is unlikely to lead to density at the positions of the atoms inside the OMIT region. Consequently, there is unlikely to be any model bias (density at the coordinates of atoms in the model owing to the presence of those atoms in the model) in the resulting maps.

In our procedure, OMIT regions are constructed so that they tile to fill the asymmetric unit. Normally, approximately 10–20 OMIT regions are used to cover the asymmetric unit, but more (as many as 132 in our tests) may be chosen so as to have a minimal impact on the density-modification procedure. As a consequence of choosing the OMIT region in this way, there may be some OMIT regions that contain no atoms from the macromolecule and others with many atoms. Those OMIT regions that contain many atoms typically have poor electron density compared with those with few atoms, as density from the atoms in the OMIT region is excluded from contributing to the density-modification procedure. It may be possible to improve the procedure by defining variable-sized OMIT regions that contain more equal numbers of atoms.

The reason for adding a border region around the OMIT region is that a peak of density in electron-density maps calculated from a model containing a particular atom has a substantial contribution within a radius that may depend on the resolution of the data and atomic displacement factors. We use a value of 2 Å for the thickness of the border region based on previous experience with composite OMIT maps in *CNS* (Brünger *et al.*, 1998).

The approach of never including atoms in an OMIT region is applicable in a straightforward way to cases in which a model is being built without reference to an existing model. In cases where molecular replacement is used, the atoms in the OMIT region should ideally be omitted from the very start of the procedure, so that the placement of the molecule and any rigid-body refinement carried out are not affected by the atomic positions of these atoms. In practice, however, this is probably unnecessary as the number of parameters being refined in the placement of the molecule is so small that little information about the positions of specific atoms in the OMIT region can be retained.

We note that the OMIT procedure described here has a kind of negative model bias. In the OMIT region a model is built but occupancies are set to zero. No solvent atoms are placed at the locations of these atoms and no bulk-solvent model is placed there. Consequently, there is low (zero) density near the locations of atoms in the model within the OMIT region. It seems possible that the procedure could be improved by setting the density in these locations to an intermediate value rather than to zero.

Although this procedure is straightforward in concept, it is somewhat less simple in implementation as all programs that operate on models that are built during iterative model-building, density modification and refinement need to keep track of which atoms are in the OMIT and border regions. These steps have been implemented in the *PHENIX* (Adams *et al.*, 2002) *AutoBuild Wizard* by using the *RESOLVE* OMIT box-generation procedure to specify the OMIT region and its boundary and to identify which atoms in a model are within these regions. Additionally, all density-modification procedures with *RESOLVE* are called with the specification of the OMIT region and boundary so that that no model-based information is transmitted into these regions through the application of NCS. Standard procedures for iterative model building, density modification and refinement are used as implemented in the *PHENIX AutoBuild Wizard*.

2.2. Calculation of a composite iterative-build OMIT map

A composite iterative-build OMIT map can be calculated by dividing the asymmetric unit of the crystal into a set of OMIT regions, calculating an iterative-build OMIT map for

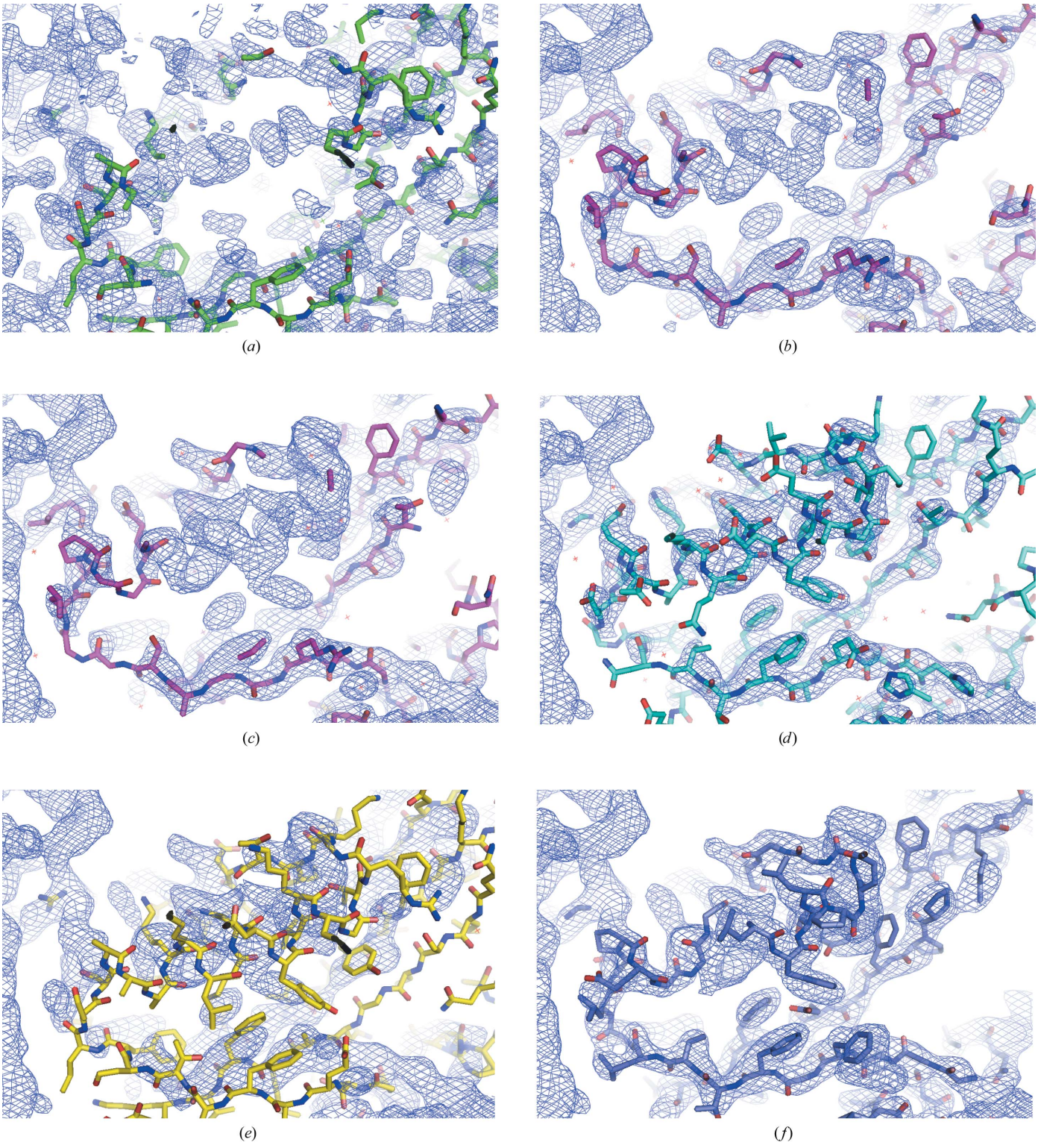

Figure 1
Iterative-build OMIT and composite iterative-build OMIT maps for the molecular-replacement solution of 1hp7 (Kim *et al.*, 2001). Maps are contoured at 1σ. (*a*) OMIT map calculated with σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ coefficients (Read, 1986) after refinement of the molecular-replacement model, omitting all atoms in one OMIT region. The atoms in the structure that were not omitted are shown. (*b*) Iterative-build OMIT map for the region shown in (*a*) after ten cycles of iterative model building, density modification and refinement. Shown is the model that was built outside of the OMIT region. (*c*) Composite iterative-build OMIT map constructed by combining all OMIT regions obtained as in (*b*). The model is the same as shown in (*b*). (*d*) Composite iterative-build OMIT map as in (*c*) with the refined structure 1hp7 superimposed. (*e*) Composite iterative-build OMIT map shown in (*c*) and (*d*) with the MR starting model superimposed. (*f*) Standard iterative-build density-modified map and model built starting from the MR starting model after removing all the atoms that are omitted in (*a*).

each region and its boundary as described above and then simply combining the OMIT regions of all the iterative-build OMIT maps. This method of combining OMIT maps is similar to that used to create composite OMIT maps in *CNS* (Brünger *et al.*, 1998). The resulting iterative-build composite OMIT map has the property that the density at each point in the map has never been affected by the presence of a model atom near that point (or near any NCS-related point). It should be noted that owing to the way that such a map is constructed, it can potentially have discontinuities at the boundaries between OMIT regions, although we have not noticed any in the maps that we have examined.

The calculation of an iterative-build OMIT map can be carried out automatically using the 'omit_type=composite_omit' keyword in version 1.3b or higher of the *PHENIX AutoBuild Wizard* (available from http://www.phenix-online.org; Adams *et al.*, 2002).

3. Results and discussion

3.1. Iterative-build composite OMIT map for antitrypsin with molecular replacement

We tested the use of an iterative-build composite OMIT map by applying it to repeat the structure solution of antitrypsin (Kim *et al.*, 2001; PDB code 1hp7) by molecular replacement using the structure of antichymotrypsin (PDB code 1as4; Lukacs *et al.*, 1998) as a search model. We used the *PHENIX AutoMR Wizard* which calls *Phaser* (McCoy *et al.*, 2005) to obtain an initial molecular-replacement solution. The *PHENIX AutoBuild Wizard* was then used to edit the sequence of the search model to match that of antitrypsin, truncating side chains at the C^{α} or C^{β} atoms if the remainder of the atomic positions were unknown. The *AutoBuild Wizard* was then used to define 128 OMIT regions covering the asymmetric unit of the crystal.

To examine the properties of OMIT maps, the starting molecular-replacement model was then used to calculate electron-density maps in two ways. Firstly, the occupancies of all atoms in one OMIT region were set to zero and the entire structure was refined with the *PHENIX* refinement package *phenix.refine* (Afonine *et al.*, 2005*b*) without rebuilding manually or with *RESOLVE*. The standard refinement procedure used included three macrocycles of refinement with automated estimation of parameters in a bulk-solvent model and placement of solvent molecules (Afonine *et al.*, 2005*a*), individual atomic coordinate shifts and isotropic atomic displacement parameters. This refinement yielded an R and $R_{\rm free}$ of 0.41 and 0.48, respectively, and a $\sigma_{\rm A}$-weighted $(2mF_{\rm o} - DF_{\rm c})\exp(i\varphi_{\rm c})$ OMIT map was calculated (Read, 1986). This $\sigma_{\rm A}$-weighted OMIT map (Fig. 1*a*) has some features corresponding to a helix in the omitted region, but is very difficult to interpret.

The second type of map calculated was an iterative-build OMIT map (Fig. 1*b*) in which the occupancies of all atoms in the same OMIT region as in Fig. 1(*a*) were maintained at a value of zero for ten cycles of iterative model building, density modification and refinement. This map has very clear features of helical density in the omitted region, despite the fact that density was never calculated using a model for any atoms in this region. Figs. 1(*c*) and 1(*d*) show that OMIT regions such as that illustrated in Figs. 1(*a*) and 1(*b*) can be joined together to form a composite iterative-build OMIT map that has clear electron density for much of the structure. Fig. 1(*e*) shows the starting model superimposed on the final composite iterative-build OMIT map. It can be seen in Fig. 1(*e*) that the helix in the starting model is offset from the final position of the helix density by about 2 Å. Fig. 1(*f*) shows that for this OMIT region a map similar to the OMIT map can also be obtained by deleting all the atoms that are in the OMIT region completely and then carrying out a standard iterative-build procedure (with no OMIT regions).

The reason why the density in the iterative-build OMIT map in Fig. 1(*b*) is so much improved over the standard OMIT map in Fig. 1(*a*) is that the model outside the omitted region was much more accurate after the iterative model-building

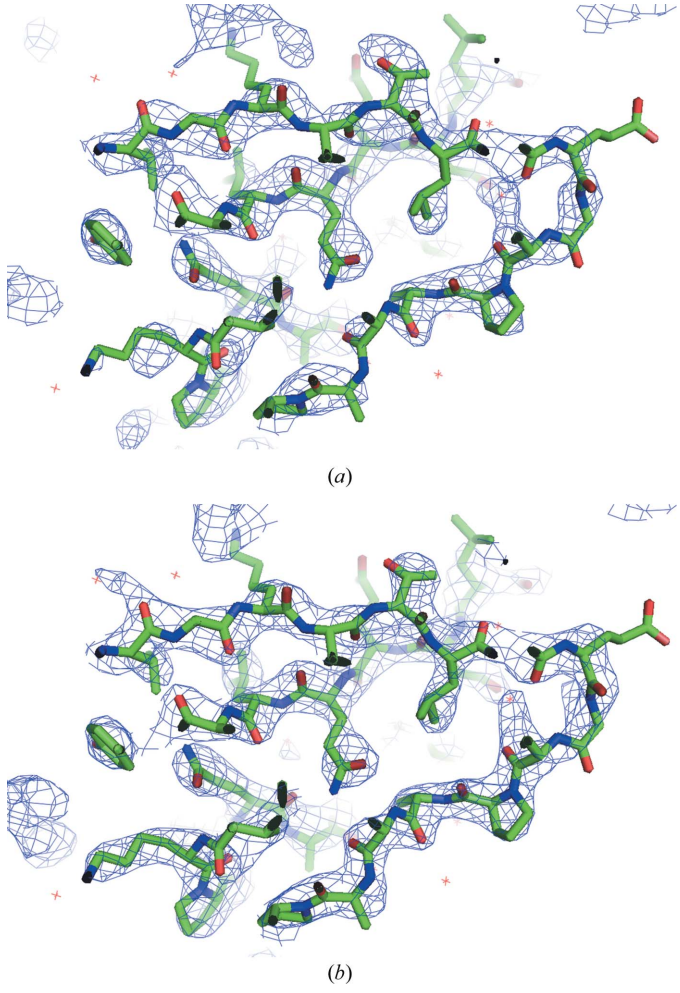

Figure 2
RESOLVE density-modified and composite iterative-build OMIT maps for the SAD experimental phasing solution of 1vqb (Skinner *et al.*, 1994). Maps are contoured at 1.5σ. (*a*) *RESOLVE* density-modified SAD-phased map (Terwilliger, 2000). (*b*) Iterative-build OMIT map for the region shown in (*a*). The model shown is the refined structure of 1vqb.

process and this model is used as a source of information for density modification. For the 128 OMIT procedures carried out, the mean final R and R_{free} (including only the part of the model outside the omitted region) were 0.29 and 0.34, respectively (compared with the starting R and R_{free} of 0.41 and 0.48, respectively). The range of R factors was from 0.25 to 0.38 and the range of free R factors was from 0.30 to 0.44 (the low R factors are paired with low free R factors). The range of both sets of R factors is quite large and the higher R factors typically correspond to OMIT regions containing larger numbers of atoms.

3.2. Iterative-build composite OMIT map for SAD-phased gene 5 protein

We also tested the use of an iterative-build composite OMIT map for a case where experimental phases were available. It seemed possible that a density-modified map could be created that had no potential for model bias but that benefited from the use of iterative model building, density modification and refinement. Experimental phases were obtained using *SOLVE* SAD phasing (Terwilliger & Berendzen, 1999) with single-wavelength anomalous dispersion (SAD) from gene 5 protein in this test. The structure of gene 5 protein had been determined previously by MAD (PDB code 1vqb; Skinner *et al.*, 1994); in this test only the data corresponding to the peak wavelength were used, simply to yield a poorer starting set of phases.

The *AutoSol Wizard* was used to resolve the structure of gene 5 protein, yielding a starting density-modified SAD-phased electron-density map with a correlation coefficient of 0.71 to the model electron-density map calculated from the refined gene 5 structure (Fig. 2*a*). The *AutoSol Wizard* uses the initial density-modified map for model building and does not perform iterative model building and refinement. It built 64 of 87 residues of the protein and (correctly) docked nine residues into sequence. The R and R_{free} for this model were 0.46 and 0.47, respectively. The *AutoBuild Wizard* was then used to calculate an iterative-build composite OMIT map (Fig. 2*b*). During refinement, the experimental phases were used as restraints (Pannu *et al.*, 1998). The mean R and R_{free} values for the final models built to construct the 132 OMIT regions were 0.26 and 0.34, respectively. The composite iterative-build map has an improved correlation coefficient of 0.82 to the model electron-density map calculated from the refined gene 5 structure. This improvement relative to the starting density-modified SAD-phased map comes from including model information from outside each OMIT region in the phase calculation for that OMIT region. In effect, both the map in Fig. 2(*a*) and the map in Fig. 2(*b*) can be thought of as density-modified maps. They differ in that the density modification used in Fig. 2(*b*) includes model information, while that in Fig. 2(*a*) does not; both are essentially free of model bias.

We note that iterative-build OMIT maps may normally be unnecessary for models autobuilt into experimental electron-density maps. The procedures typically used in autobuilding have a high cutoff for density and incorrectly placed atoms are normally removed in subsequent cycles, so that it is un-

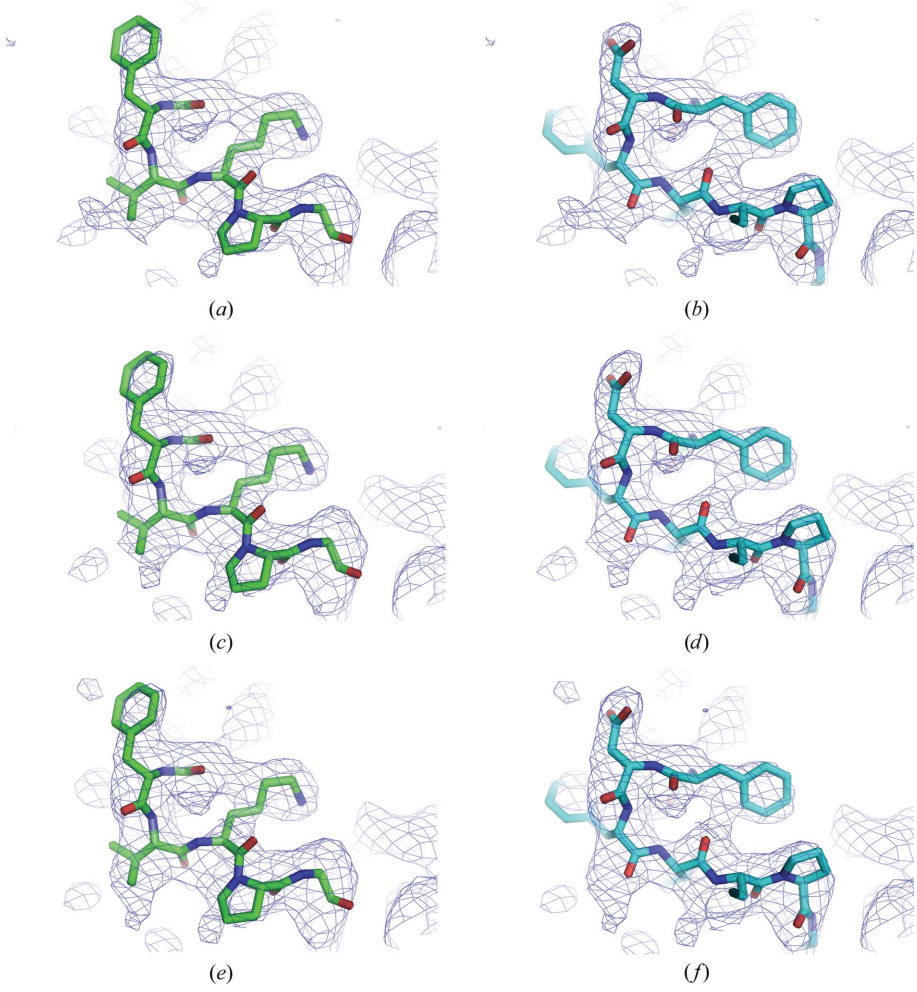

Figure 3
σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ and OMIT maps for 1zen (Cooper *et al.*, 1996) compared with the structure 1zen and that of chain *A* from 1b57 (Hall *et al.*, 1999) superimposed on the structure 1zen. Maps are contoured at 1σ. (*a*, *b*) σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map (Read, 1986) calculated after refinement of the 1zen structure with *phenix.refine* (Afonine *et al.*, 2005*b*), compared with structure of 1zen (*a*) and with chain *A* from 1b57 (*b*). (*c*, *d*) As in (*a*) and (*b*), except that the atoms in the 1zen structure were moved randomly by an r.m.s. distance of 1 Å ('shake' procedure) and then refined for six cycles with *phenix.refine*. (*e*, *f*) As in (*c*) and (*d*), except that solvent water molecules were removed after the shake procedure and ten cycles of refinement including simulated annealing were carried out.

common for atoms to be repeatedly placed in very incorrect positions, as would be required to introduce model bias. A map for this structure can also be calculated by carrying out the iterative-building procedure with no OMIT regions; this map is slightly better than the OMIT map, with a correlation coefficient with the model map of 0.85.

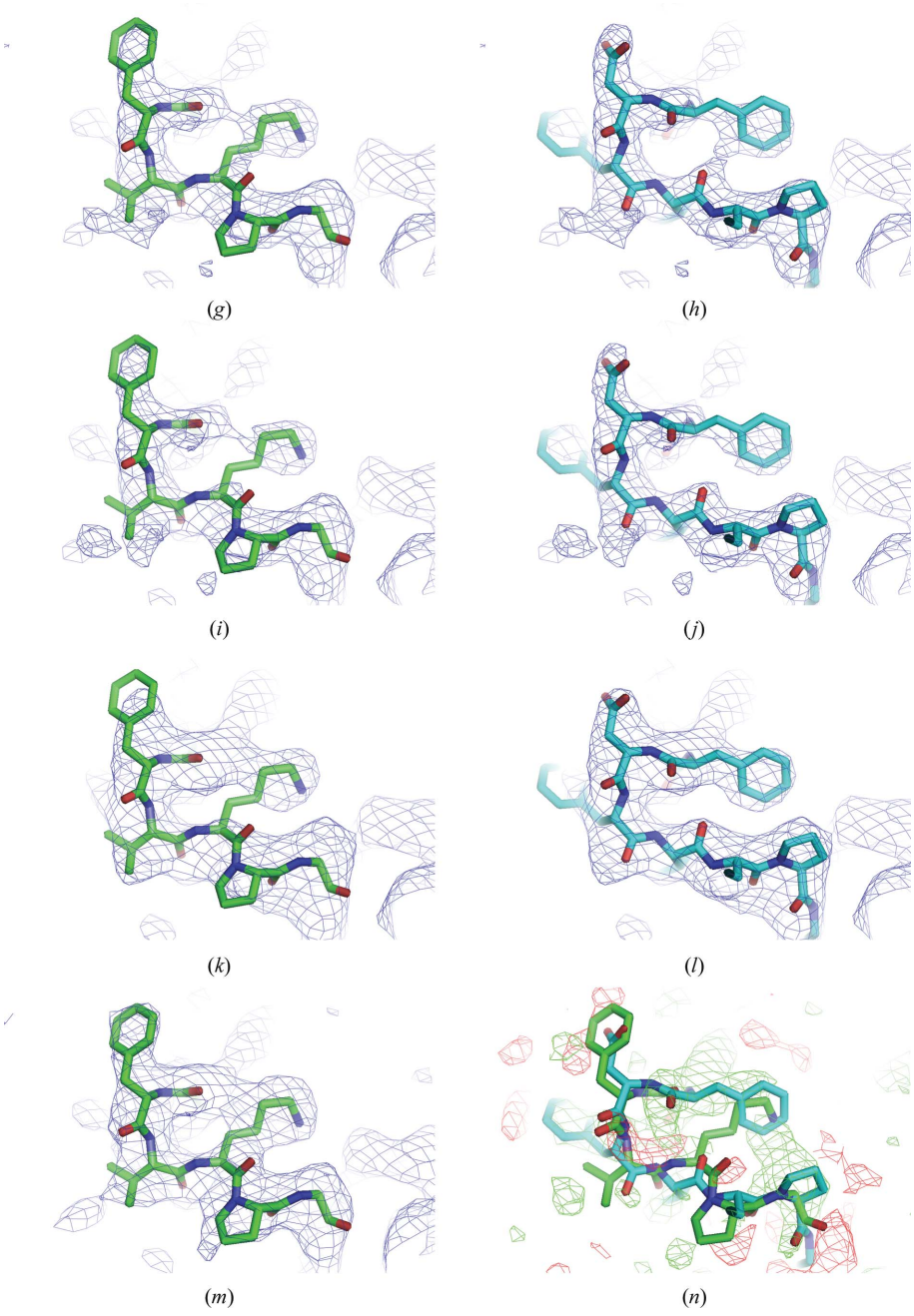

Figure 3 (continued)
(g, h) As in (a) and (b), except the map shown is a simple OMIT map calculated by omitting all atoms in an OMIT box with edges parallel to the cell edges and 4 Å from any atom in residues 5–9 of 1zen (setting their occupancies to zero), refining the resulting structure and calculating a σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map. (i, j) Map calculated as in (g) and (h) except the map is a simulated-annealing OMIT map in which the refinement step in (c) and (d) is replaced by simulated-annealing refinement (Brünger *et al.*, 1998). (k, l) Iterative-build OMIT map calculated as in (g) and (h) except that the 1zen structure was iteratively rebuilt using the rebuild-in-place option of the *PHENIX AutoBuild Wizard*, always setting the occupancies of all atoms in the OMIT box to zero during the procedure. (m, n) Maps downloaded from the EDS density server (Kleywegt *et al.*, 2004). (m) σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map and model for 1zen. (n) σ_A-weighted $(mF_o - DF_c)\exp(i\varphi_c)$ map for 1zen contoured at $\pm 2\sigma$ with coordinates of 1zen (green) and 1b57 (blue) superimposed.

3.3. Using iterative-build OMIT maps to remove existing model bias

Although the principal intent of the iterative-build OMIT procedure described here is to avoid model bias entirely, it seemed possible that the process of rebuilding a model outside an OMIT region might also be useful in removing existing model bias. In particular, we would expect that extensive rebuilding should effectively remove adjustments to atoms in the rest of the model that compensate for incorrectly placed atoms. To test this idea, we identified an entry in the PDB with some features that were likely to be incorrect but which remained in a σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map (Read, 1986) calculated after refinement of the structure. PDB entry 1zen (Cooper *et al.*, 1996) was such a structure (G. Kleywegt, personal communication). It was obtained at a resolution of 2.5 Å and the closely related structure 1b57 (Hall *et al.*, 1999), determined later at a resolution of 2 Å, differs in the sequence register of residues 6–16 by one residue.

Figs. 3(a) and 3(b) show a σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map calculated after re-refinement of 1zen (without omitting any atoms) with *phenix.refine* using deposited structure factors and regenerating a test set of reflections for refinement (the original test set was not available). The final R and R_{free} after refinement were 0.25 and 0.29, respectively. Fig. 3(a) shows a map of residues 5–10 from 1zen, centered on residues Phe6 and Lys8, which are likely to be misaligned by one residue in this structure. Fig. 3(b) shows a map of residues 3–10 of chain A from PDB entry 1b57 after superimposing this chain from 1b57 onto the structure from 1zen by least squares. The σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map has features in common with both the structure used to generate the phases for this map (1zen) and with the structure derived

from the higher resolution model (1b57), including several features that appear to be examples of model bias in the map. In particular, the map shows density for the side chain of Lys8 from the 1zen structure even though the more likely 1b57 structure only has a carbonyl O atom pointing towards this location. Overall, in the neighborhood of residues 5–10 of 1zen, the σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map is somewhat more similar to a model map calculated from the higher-resolution structure 1b57 than to a model map calculated from 1zen. This is shown numerically by local map correlation coefficients, which are summarized in Table 1 for this map and the maps discussed below.

It seemed possible that the model bias found in the σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map shown in Figs. 3(*a*) and 3(*b*) could arise in part from the fact that we did not have access to the original test set used in refinement. To examine this possibility, we carried out a second refinement in which the atoms in the structure were displaced by an r.m.s. distance of 1.0 Å with the 'shake' procedure (Brünger *et al.*, 1998; Guncar *et al.*, 2000), yielding a starting R and R_{free} of 0.48 and 0.47, respectively, followed by six cycles of refinement with *phenix.refine*, leading to a final R and R_{free} of 0.25 and 0.30, respectively. The resulting σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map is shown in Figs. 3(*c*) and 3(*d*). It appears to have slightly less model bias than the map calculated after standard refinement, but overall the maps are very similar.

Next, we carried out a more extensive re-refinement procedure to try to reduce model bias. The partially randomized model from the 'shake' procedure above was taken as a starting point, all solvent molecules (waters) were removed and ten cycles of refinement and water picking, including two cycles with simulated annealing, were carried out. Once again the starting R and R_{free} were 0.48 and 0.47, respectively, and the final R and R_{free} were 0.25 and 0.30, respectively. The resulting map has less model bias than the starting σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map (Figs. 3*e* and 3*f*) and lowered correlation with the 1zen model (Table 1), but it would still be difficult to decide which of the two models is correct because of the model bias showing density for the entire side chain of Lys8 from the 1zen structure.

Three types of OMIT maps were then calculated using the coordinates and structure factors from 1zen. In each case, the occupancies of all the atoms inside a small OMIT box were set to zero before initial refinement and throughout the procedures. The OMIT box was defined as a region with edges parallel to the cell edges and 4 Å from the nearest atom in residues 5–9 of 1zen. The OMIT maps calculated in these procedures are therefore not based on any density from any atoms in residues 5–9 of 1zen.

The first OMIT map calculated (Figs. 3*g* and 3*h*) was a simple OMIT map in which the structure of 1zen was refined with zero occupancies for the atoms in the OMIT region and a σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ map was calculated. This map still shows model bias from 1zen at the side chain of Lys8 (Fig. 3*g*, Table 1).

The second OMIT map was a simulated-annealing OMIT map (Brünger *et al.*, 1998) calculated with *phenix.refine* (Afonine *et al.*, 2005*b*; Figs. 3*i* and 3*j*). In calculating this map the refinement started at a pseudo-temperature of 5000 K and cooled to a final temperature of 300 K. This map shows substantially less model bias but has relatively weak density for the entire segment, resulting in lower correlations with the density from both models (Table 1).

Table 1
Map correlation coefficients near residues 3–10 of 1zen.

Map (all based on 1zen structure except as noted)	Map correlations† with 1zen model	Map correlations† with 1b57 model
Initial σ_A map, no OMIT	0.68	0.75
σ_A map after 'shake' procedure, no OMIT	0.67	0.74
σ_A map after 'shake', removal of waters, refinement and water picking, no OMIT	0.64	0.75
Simple refined OMIT	0.63	0.71
Simulated-annealing OMIT	0.60	0.71
Iterative-build OMIT	0.65	0.75
1b57 F_{calc} map	0.66	0.98

† Map correlations were calculated with *RESOLVE* (Terwilliger, 2000), including grid points within 2 Å of each atom in the corresponding model. Residues 3–10 from 1zen were chosen because they were largely within the OMIT region and residues 3–10 from 1b57 were selected to match the 1zen fragment.

The third OMIT map was an iterative-build OMIT map (Figs. 3*k* and 3*l*). To create this map, the model, with zero-occupancy atoms in the OMIT region, was rebuilt three times using the *PHENIX* rebuild-in-place algorithm. In this rebuilding procedure the polypeptide chain is rebuilt by iteratively removing a segment and retracing the chain for that segment. The parts of the resulting models that best fit a density-modified map are then combined and side chains are re-fitted into the density. This iterative-build OMIT map shows little model bias from 1zen (Fig. 3*k*) and matches the model from 1b57 well (Fig. 3*l*, Table 1).

We also tested whether a σ_A-weighted $(mF_o - DF_c)\exp(i\varphi_c)$ map calculated from the 1zen model might be as informative as the OMIT maps we have calculated. We downloaded σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ and σ_A-weighted $(mF_o - DF_c)\exp(i\varphi_c)$ maps from the EDS server (Kleywegt *et al.*, 2004). Fig. 3(*m*) shows the $(2mF_o - DF_c)\exp(i\varphi_c)$ map, which as expected is essentially identical to the map calculated by *phenix.refine* and shown in Fig. 3(*a*). Fig. 3(*n*) shows the difference $(mF_o - DF_c)\exp(i\varphi_c)$ map; however, there is no negative (or positive) difference density at the coordinates of the misplaced lysine side chain.

Considering all of these maps, along with the differences between the lower resolution 1zen and higher resolution 1b57 models, the simplest interpretation of these results is that the higher resolution 1b57 model is the more accurate of the two in the region that we have examined and that the iterative-build OMIT map is particularly useful in reducing model bias without much cost to overall map quality.

4. Conclusions

The iterative-build OMIT procedure can be thought of as a type of density modification that involves the use of a model outside the OMIT region. In statistical density-modification

procedures, model density can be used as the expected value of the electron density. Phases are then adjusted to better match the map to this expected density (Terwilliger, 2003). In the iterative-build OMIT procedure, the model is built wherever it exists but only model density outside the OMIT region is used in density modification. In this way, the model density can improve the quality of the crystallographic phases, yet not directly bias the density in the OMIT region.

The iterative-build OMIT procedure can be of substantial use in molecular replacement in situations where initial refinement of the molecular-replacement model yields relatively poor R factors, but in which iterative model building, density modification and refinement yields a greatly improved model with lowered R factors. In such a case a σ_A-weighted $(2mF_o - DF_c)\exp(i\varphi_c)$ OMIT map calculated after initial refinement may be relatively uninformative because the model is not yet good enough to produce phases that lead to a clear map (*e.g.* Fig. 1*a*). On the other hand, after iterative model building, density modification and refinement, the model built outside the OMIT region can be accurate enough to yield phases that clearly show the density inside the OMIT region (*e.g.* Fig. 1*b*). For example, in the case shown in Fig. 1 the starting R factor after refinement of the model was 0.41, but the iterative-building procedure yielded much lower R factors (mean of 0.29) and produced a very clear map for the omitted region.

The generation of iterative-build composite OMIT maps can be computation-intensive, particularly in cases in which there is only just enough phase information for the iterative model-building procedure to improve upon the starting model. In such cases, the size of the OMIT regions must be very small or no phase improvement results. Consequently, in some cases many OMIT regions must be constructed and the entire iterative model-building procedure must be carried out many times (132 OMIT regions were combined in the case shown in Fig. 2*b*). The procedure is readily made parallel, so as highly parallel machines and large clusters become increasingly available the procedure may become practical even for very large structures. At present, the procedure can be time-consuming for a very large structure when run on a single processor, as several CPU days can be required for each OMIT region. A use of the iterative-build OMIT procedure that is quicker is to construct an unbiased map for a small region within an electron-density map, such as for the density in the vicinity of a ligand or side chain of interest. In such a case a single OMIT map can often be calculated, setting to zero the occupancies of all atoms in a box containing the region of interest during the iterative-build process. This local application of the iterative-build OMIT procedure may also be of substantial use in checking for errors that may be partially masked by model bias in completed structures, as shown in Fig. 3.

An additional potential use of composite iterative-build OMIT maps is as a source of relatively unbiased phasing information. The density in each OMIT region of these composite maps is not biased by the model within that OMIT region. Consequently, it seems possible that the phase information obtained by using these maps as a target for density modification (Terwilliger, 2001) might be of high quality. It might also have lowered model bias compared with that obtained directly from a model. It would not necessarily be completely unbiased, however, because the inverse Fourier transformation would combine phase information from different OMIT regions and the separation of model information within each OMIT region from the map in that region would therefore no longer be complete. Preliminary experiments indicate that these maps can be improved over standard density-modified maps.

We conclude by noting that increasing use of automated procedures for iterative model building, density modification and refinement (Perrakis *et al.*, 1999; Terwilliger, 2003; DePristo *et al.*, 2005; Ondráček, 2005) has the potential for reducing the effects of model bias and the incidence of significant errors to very low levels, particularly for experimentally phased structures, because a complete check of the fit of model to the electron-density map can be easily carried out during every cycle of automated building. If poorly fitted parts of a model are removed promptly, instead of remaining for many cycles of refinement during which the remainder of the model adjusts to compensate for errors in building, then subsequent electron-density maps will be unlikely to have substantial model bias.

The authors would like to thank the NIH Protein Structure Initiative for generous support of the *PHENIX* project (1P01 GM063210). This work was supported in part by the US Department of Energy under Contract No. DE-AC02-05CH11231. RJR is supported by a Principal Research Fellowship from the Wellcome Trust (UK). The authors would like to thank Gerard Kleywegt for pointing out the differences between PDB entries 1zen and 1b57 and an anonymous reviewer for unusually extensive and insightful questions and comments.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.

Adams, P. D., Pannu, N. S., Read, R. J. & Brunger, A. T. (1999). *Acta Cryst.* D**55**, 181–190.

Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. (2005*a*). *Acta Cryst.* D**61**, 850–855.

Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. (2005*b*). *CCP4 Newslett.* **42**, contribution 8.

Bhat, T. N. (1988). *J. Appl. Cryst.* **21**, 279–281.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Cooper, S. J., Leonard, G. A, McSweeney, S. M., Thompson, A. W., Naismith, J. H., Qamar, S., Plater, A., Berry, A. & Hunter, W. N. (1996). *Structure*, **4**, 1303–1315.

DePristo, M. A., de Bakker, P. I. W., Johnson, R. J. K. & Blundell, T. L. (2005). *Structure*, **13**, 1311–1319.

Guncar, G., Klemenicic, I., Turk, B., Turk, V., Karaoglanovic-Carmona, A., Juliano, L. & Turk, D. (2000). *Structure*, **8**, 305–313.

Hall, D. R., Leonard, G. A., Reed, C. D., Watt, C. I., Berry, A. & Hunter, W. N. (1999). *J. Mol. Biol.* **287**, 383–394.
Hodel, A., Kim, S.-H. & Brünger, A. T. (1992). *Acta Cryst.* A**48**, 851–858.
Hunt, J. F. & Deisenhofer, J. (2003). *Acta Cryst.* D**59**, 214–224.
James, M. N. G., Sielecki, A. R., Brayer, G. D., Delbaere, L. T. J. & Bauer, C.-A. (1980). *J. Mol. Biol.* **144**, 43–88.
Kim, S., Woo, J., Seo, E. J., Yu, M. & Ryu, S. (2001). *J. Mol. Biol.* **306**, 109–119.
Kleywegt, G. J. (2000). *Acta Cryst.* D**56**, 249–265.
Kleywegt, G. J., Harris, M. R., Zou, J., Taylor, T. C., Wählby, A. & Jones, T. A. (2004). *Acta Cryst.* D**60**, 2240–2249.
Lukacs, C. M., Rubin, H. & Christianson, D. W. (1998). *Biochemistry*, **37**, 3297–3304.
Lunin, V. Y., Afonine, P. V. & Urzhumtsev, A. G. (2002). *Acta Cryst.* A**58**, 270–282.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Ondráček, J. (2005). *Acta Cryst.* A**61**, C163.
Pannu, N. S., Murshudov, G. N., Dodson, E. J. & Read, R. J. (1998). *Acta Cryst.* D**54**, 1285–1294.
Pannu, N. S. & Read, R. J. (1996). *Acta Cryst.* A**52**, 659–668.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Ramachandran, G. N. & Srinivasan, R. (1961). *Nature (London)*, **190**, 159–161.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Read, R. J. (1997). *Methods Enzymol.* **278**, 110–128.
Reddy, V., Swanson, S. M., Segelke, B., Kantardjieff, K. A., Sacchettini, J. C. & Rupp, B. (2003). *Acta Cryst.* D**59**, 2200–2210.
Rossmann, M. G. (1972). *The Molecular Replacement Method.* New York: Gordon & Breach.
Skinner, M. M., Zhang, H., Leschnitzer, D. H., Guan, Y., Bellamy, H., Sweet, R. M., Gray, C. W., Konings, R. N. H., Wang, A. H.-J. & Terwilliger, T. C. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 2071–2075.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Terwilliger, T. C. (2001). *Acta Cryst.* D**57**, 1763–1775.
Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 2082–2086.
Terwilliger, T. C. (2003). *Acta Cryst.* D**59**, 1174–1182.
Terwilliger, T. C. (2004). *Acta Cryst.* D**60**, 2144–2149.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Urzhumtsev, A. G., Skovoroda, T. P. & Lunin, V. Y. (1996). *J. Appl. Cryst.* **29**, 741–744.
Zeng, Z.-H., Castano, A. R., Segelke, B. W., Stura, E. A., Peterson, P. A. & Wilson, I. A. (1997). *Science*, **277**, 339–345.