Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A new software routine that automates the fitting of protein X-ray crystallographic electron-density maps

David G. Levitt

Department of Physiology, University of Minnesota, Minneapolis, MN 55455, USA

Correspondence e-mail: levitt@dcmir.med.umn.edu

Received 23 January 2001
Accepted 16 April 2001

The classical approach to building the amino-acid residues into the initial electron-density map requires days to weeks of a skilled investigator's time. Automating this procedure should not only save time, but has the potential to provide a more accurate starting model for input to refinement programs. The new software routine *MAID* builds the protein structure into the electron-density map in a series of sequential steps. The first step is the fitting of the secondary α-helix and β-sheet structures. These 'fits' are then used to determine the local amino-acid sequence assignment. These assigned fits are then extended through the loop regions and fused with the neighboring sheet or helix. The program was tested on the unaveraged 2.5 Å selenomethionine multiple-wavelength anomalous dispersion (SMAD) electron-density map that was originally used to solve the structure of the 291-residue protein human heart short-chain L-3-hydroxyacyl-CoA dehydrogenase (SHAD). Inputting just the map density and the amino-acid sequence, *MAID* fitted 80% of the residues with an r.m.s.d. error of 0.43 Å for the main-chain atoms and 1.0 Å for all atoms without any user intervention. When tested on a higher quality 1.9 Å SMAD map, *MAID* correctly fitted 100% (418) of the residues. A major advantage of the *MAID* fitting procedure is that it maintains ideal bond lengths and angles and constrains φ/ψ angles to the appropriate Ramachandran regions. Recycling the output of this new routine through a partial structure-refinement program may have the potential to completely automate the fitting of electron-density maps.

1. Introduction

Recent advances in synchrotron-radiation sources (Hendrickson, 1991; Ogata, 1998) combined with the use of selenomethionine cloned proteins (Doublie, 1997; Hendrickson *et al.*, 1990) has markedly reduced the time required to acquire X-ray intensities and preliminary phase information. These advances, fueled by recent funding programs such as the Protein Structure Initiative (Smaglik, 2000; Abbott, 2000) which are designed to stimulate the production of thousands of protein structures, have increased the incentive for automating the steps involved in protein structure determination.

In going from crystals to the final refined protein structure, one of the most time-consuming steps is the initial fitting of the known amino-acid sequence into the phase-modified selenomethionine multiple-wavelength anomalous dispersion (SMAD) electron-density map. The classical approach to this initial fitting is to use an interactive graphics workstation and a program such as *O* (Jones *et al.*, 1991). As these approaches require substantial manual intervention, this can involve days to weeks of a skilled investigator's time. Currently, there are at least four programs that are being developed to automate this fitting: (i) *wARP* (Perrakis *et al.*, 1999), (ii) *MAIN* (Turk & Guncar, 1999), (iii) two routines (*X-Powerfit* and *X-Autofit*) that are part of the *QUANTA* package from Molecular Simulations, Inc (Oldfield, 2001) and (iv) *MAID*, the program described here. The generally accepted leader in this field is *wARP*, which starts with a preliminary SMAD map and then uses *ab initio* procedures to extend the map to high resolution and position the protein atoms. This program has demonstrated some remarkable successes. For example, without any manual intervention it was able to completely solve the structure of a 4189-atom protein using a 1.2 Å resolution map starting with only the locations of the co-crystallized U atoms (Tame, 2000). However, since the most powerful features of *wARP* require atomic resolution data, this program will be much more limited when applied to the lower resolution maps (2.5 Å or less) that are frequently encountered. The program described here (*MAID*) should be applicable to these lower resolution maps.

The main program development has been based on application to a 2.5 Å unaveraged SMAD map that was originally used to solve the structure of human heart short-chain L-3-hydroxyacyl-CoA dehydrogenase (SHAD; Barycki *et al.*, 1999). A major advantage of *MAID* is that it returns a very accurate fit to the final refined structure while maintaining ideal bond lengths and angles. This should simplify the subsequent refinement steps. For example, using as input just the SMAD map and the amino-acid sequence, without any user intervention, *MAID* was able to fit 80% of the SHAD structure with a 0.43 Å r.m.s. error for the main-chain atoms (C^{α}, N, C and O) and a 1.0 Å r.m.s. error for all atoms.

2. Overview

The approach used by *MAID* is basically an automation of the steps that a skilled investigator would use in the classical fitting approach. The program involves a complex branching tree consisting of more than 100 different routines. Fig. 1 shows a flow chart that summarizes the major steps in the program. Each of these steps will be briefly described here and then discussed in detail below.

The first step is selecting the region of the electron-density map that should be fitted. In order to minimize the solution time, one would like to pick a region that covers just one subunit. However, since this is not usually possible, the procedure adopted in *MAID* is to allow the user to start with a large region that may overlap parts of different subunits. Then, in the last step, *MAID* looks for symmetry operators between the different subunits and uses the partial fits to synthesize complete subunits.

The next step is to find the secondary structure (β-sheets and α-helices). *MAID* uses the skeletonized bones as trial points for the initial positioning of a helix or sheet which is then extended as far as the density allows. An overriding imperative in the design of the program is to avoid incorrect assignments, because they are difficult to correct. For example, a helix extended one residue too far will be impossible to connect to its neighbor sequence. Thus, a number of routines have been developed to test these helix and sheet 'fits' and to remove any questionable residues.

The next step is to try and connect these 'fits' through the loop regions. Fitting the loops is a much more difficult problem because φ/ψ constraints are not as limited and the map quality is poorer. The following procedure was adopted after trying a large number of different approaches. Fitting these loop regions for low-resolution maps becomes much easier if the amino-acid sequence has been previously assigned. This has two advantages. Firstly, one can use the specific side chains to help guide the extension and, most importantly, one knows which additions are glycines, with their corresponding unlimited φ/ψ constraints, or prolines, with the possibility of a *cis* bond. Thus, the first step in the loop extension is to assign the sequence. This is performed by sliding the known amino-acid sequence along the fits and looking at the side-chain density (Jones *et al.*, 1991). In order to make an unequivocal assignment, one must usually have 15–20 consecutive residues. Since most of the fits are not this long, it is necessary to first use the bones to find which fits are connected and then use these two (or three, if necessary) connected fits to make the sequence assignment.

This sequence assignment is then used to connect these two fits. For each residue that is added, a large set of starting φ/ψ values and constraints are tried and any addition that meets some minimal standard is allowed. A set of up to nine different possible chain extensions are followed. The chain that make the closest connection to the neighboring helix or sheet is defined as the best extension and it is 'fused' with this neighbor. This assigned extended fit is extended in both directions until it reaches a region where the density is so poor that a connecting fit cannot be found. *MAID* then goes back and tries to make a new sequence assignment, repeating this process until all the fits have been tested. Finally, in the last step, *MAID* uses symmetry operators between the different subunits to try and synthesize a complete subunit.

3. Computation details

3.1. Hardware

Two different programs have been developed. (i) A batch program that runs the autofitting routines. It is written in C++ and, although only used on SGI machines, should be easily adaptable to other platforms. This is the only version that is needed to run the fitting routine. (ii) A graphics program that is used for software development and as an optional procedure to set up the conditions and files used by the batch program. It uses Motif and OpenGL and runs only on the SGI platform. The use of this custom graphics program was essential for the development process because it allowed all the different steps to be directly visualized. All of the figures used in this paper (except for the Ramachandran plot) are based on 'snapshots' of the screen using this graphics routine.

These routines represent a complete revision of an older version of *MAID* (Levitt & Banaszak, 1993).

3.2. Test maps

Two different SMAD maps were used for program development. Most of the early testing used the map from SHAD (Barycki *et al.*, 1999). This data was collected on the 19-ID beamline at the Argonne National Laboratory's Advanced Photon Source. The map was obtained by finding the initial phases using *SOLVE* (Terwilliger, 1997) and then processing through one round of density modification using *DM* (Cowtan & Zhang, 1999) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Although there were two subunits per asymmetric unit, the map was not averaged because there was significant contact asymmetry. Since the map is unaveraged, each subunit region provides a unique test for *MAID*. When the program was nearly completed, another round of development was carried out using the 1.9 Å SMAD map for fumarylacetoacetate hydrolase (FAH; Timm *et al.*, 1999). This is a very high quality easily interpretable map that provides a case in which one would expect *MAID* to find a nearly complete fit. FAH also has two subunits per asymmetric unit and was not averaged.

3.3. Skeletonization

A custom skeletonization routine for making the 'bones' was developed based on the algorithm of Greer (1974). Core tracing (Swanson, 1994) was also tried, but it did not provide as close a fit to the refined coordinates. The 'bones' are used in two different steps. They are first used to pick initial starting positions for the helices and sheets. They are also used to find which 'fits' are connected and to guide the extension through the loop regions.

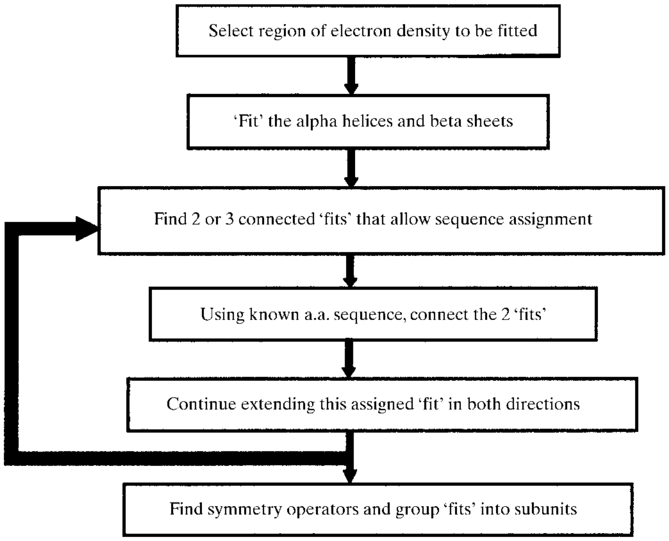
Select region of electron density to be fitted
'Fit' the alpha helices and beta sheets
Find 2 or 3 connected 'fits' that allow sequence assignment
Using known a.a. sequence, connect the 2 'fits'
Continue extending this assigned 'fit' in both directions
Find symmetry operators and group 'fits' into subunits

Figure 1
Flow chart illustrating the major steps in the program.

3.4. Refinement

At each step in the program, the fits are refined using a brief simulated-annealing real-space molecular-dynamic routine. The molecular dynamics uses the 'torsional dynamics' algorithm of Bae & Huang (1987). In most cases, only the last 3–6 residues are included in this dynamics. The side chains are built by first using a rotomer library (Dunbrack & Karplus, 1993) to choose the best rotomer, which is then refined by torsional dynamics. The appropriate φ/ψ constraints are imposed during the torsional dynamics to keep the φ/ψ angle in, for example, the favored α-helix region. The 'fusion' of the overlapping fits is also refined using this procedure. This refinement guarantees that the final *MAID* structure has ideal bond lengths and angles and that the φ/ψ angles are constrained to the appropriate Ramachandran regions.

3.5. Fitting the α-helices and β-sheets ('trace' routine)

The auto fitting is carried out in two major steps. The first step is to fit the α-helices and β-sheets. Because of their limited φ/ψ constraints, these regions have characteristic features that make them relatively easy to identify. The minimum condition that a region will be tested for a helix or sheet is that there is continuous bone trace at least three residues long. The graphics version of *MAID* allows the user to look at bone traces with different minimum densities (0.8–2.0 standard deviations) and choose the one that is most appropriate. This is the only adjustable parameter in the entire program. For good maps (*e.g.* FAH) a minimum value of 1.4 was used, while for the poorer SHAD map a minimum value of 1.2 was used. The graphics version of *MAID* allows the user to move and size a set of spheres that determines the region of the map that will be used as a starting point for locating the sheets or helices.

For each bone trace that is three residues long, an ideal helix or sheet (depending on the geometry of the bone trace) is positioned and refined. If the refined structure fits the map density satisfactorily, then it is extended in both directions as far as possible (keeping the helix or sheet constraint), using a round of real-space dynamics refinement after each addition.

It is critical that errors are not introduced at this step. It is fairly common for one of these fits to be extended one residue too far so that, for example, the main chain follows what is actually the side chain. To prevent this, a number of routines were created to check the geometry of these fits and delete bad residues. One of the most useful approaches is to look at the bone trace at the end of the fit. This usually can identify when, for example, a main chain is erroneously following a side chain. A surprisingly difficult problem is to determine the correct direction of a β-sheet. The fitting routine is always repeated for both directions and the direction that has the best fit to the density is chosen.

Fig. 2 shows an example of the output of the trace routine for SHAD for a neighboring sheet and helix. The black trace is the final refined main chain and the red trace is the *MAID* main chain. There is a three-residue gap in the loop between these two *MAID* fits. The blue lines are the refined side chains.

At this stage, the sequence assignment is not known so that the *MAID* side chains (green) consist simply of carbon–carbon linked atoms.

3.6. Determination of the sequence assignment for connected fits

As discussed above, knowledge of the sequence assignment is essential for accurate extension of the fits through the loops for low-resolution maps. *MAID* uses a routine that looks at all the bone traces starting at, for example, the C-terminal end of a fit and looks for traces that end, for example, on the N-terminal end of another fit. The traces are edited and the shortest connection between the two fits is determined. *MAID* goes through this procedure starting with a map skeletonized using a minimum density of 2.0 standard deviations. If no connection is found at this density, it then tries 1.8, 1.6, 1.4 *etc*. down to a minimum value of 0.8, using the connection at the highest value. If there is no connection at a value of 0.8, then the map is assumed to be of such poor quality in this region that no extension is possible.

Using this bone connection, *MAID* tests if an unambiguous sequence assignment can be made for these two connected fits. The 'gap' between the two fits is estimated from the bone connection. For each residue in these two fits, the best possible fit to the density of each of the 20 possible amino-acid side

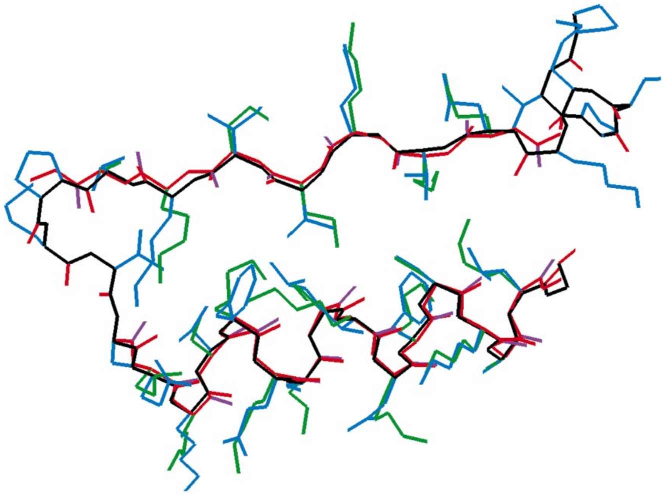

Figure 2
Example of the *MAID* fit (red) to the helix and sheet secondary structure of SHAD. The black line is the final refined structure. The helix and sheet are connected through a three-residue loop.

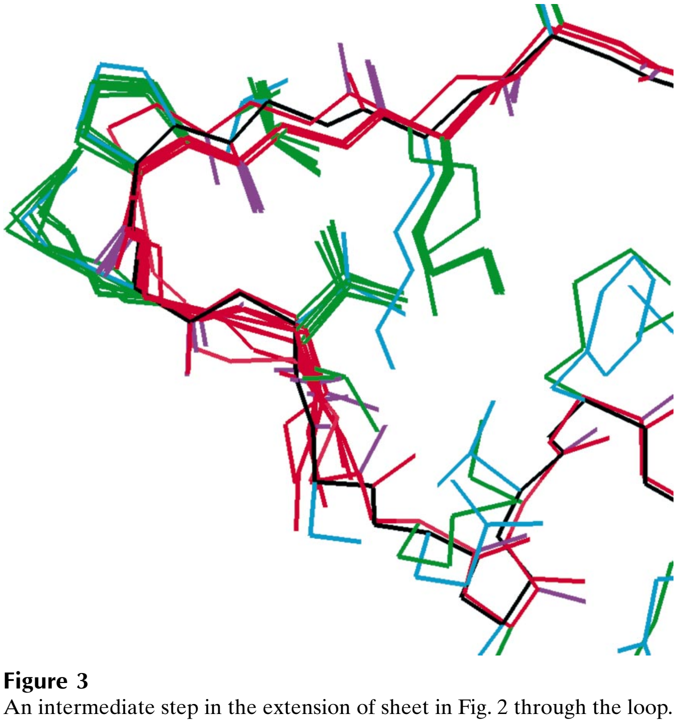

Figure 3
An intermediate step in the extension of sheet in Fig. 2 through the loop. Six trial fits are shown (red) along with the bone connection (pink) that is used to guide the extension.

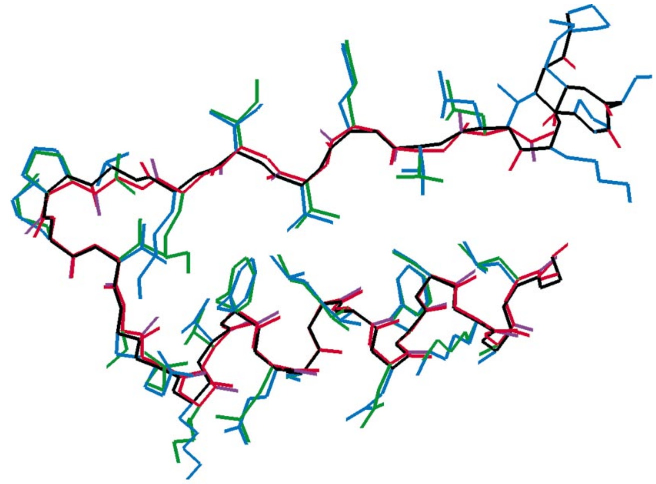

Figure 4
MAID fit (red) after extension of the sheet (Fig. 2) through the loop and fusion with the helix. The black line is the final refined structure.

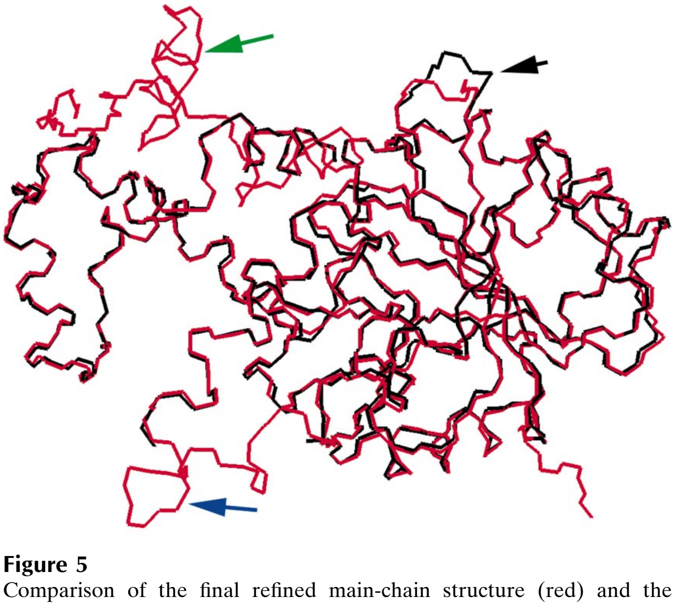

Figure 5
Comparison of the final refined main-chain structure (red) and the structure output by *MAID* (black) for the SHAD map.

chains is determined. This involves first picking the best rotomer from a rotomer library (Dunbrack & Karplus, 1993) and then refining it using torsional Powell minimization (holding the main chain fixed). The average side-chain atom density and the number of atoms in bad density for each amino acid is stored in a lookup table. For each amino acid, 'quality' routines were written that quantified the quality of the amino-acid assignment using the information from all 20 amino acids tested for this residue. For example, if both serine and phenylalanine fit the side-chain density, then serine would be given a poor score and phenylalanine a good score.

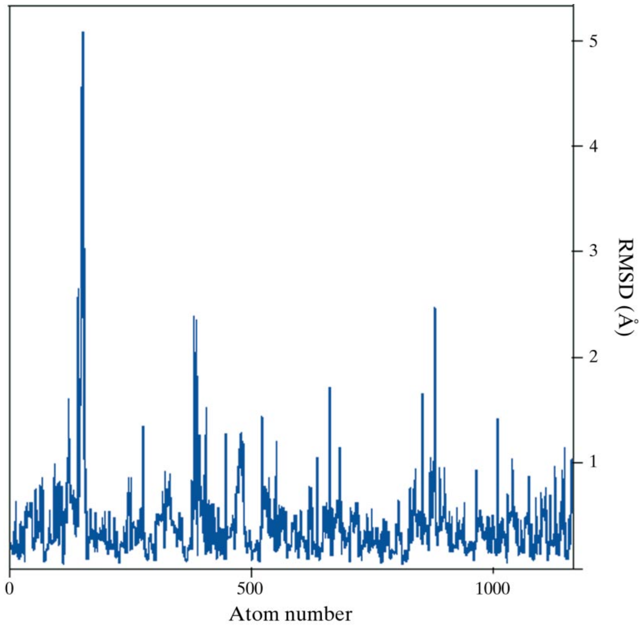

Figure 6
R.m.s.d. error (Å) for the SHAD main-chain atoms fitted by *MAID*.

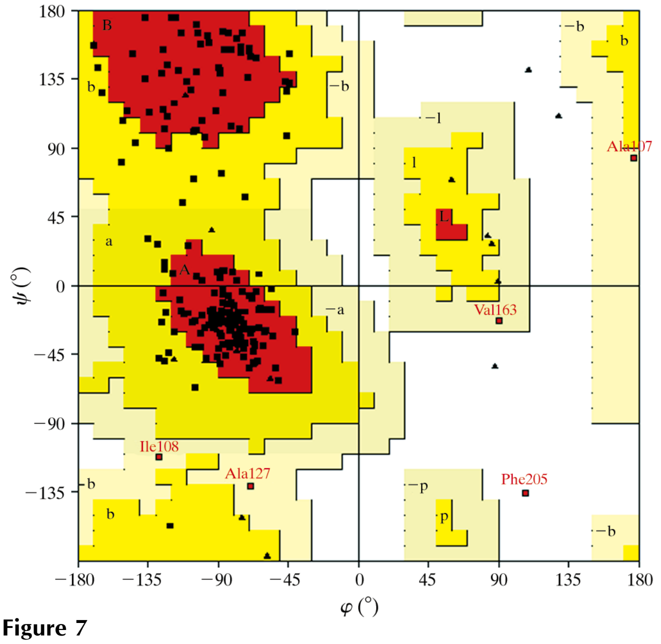

Figure 7
Ramachandran plot for the SHAD structure output by *MAID*.

All possible assignments of the known amino-acid sequence are then tested and quantified and the quality of each possible sequence assignment is determined using the lookup table and the 'quality' routines. It is assumed that the 'gap' determined from the bone connection is accurate to ±1 so that, for the example shown in Fig. 2, gaps of two, three and four are tested in addition to all possible sequence assignments. If the top-ranking assignment is much better than the second-best assignment, then it is assumed that this is the correct assignment. For example, for the two fits shown in Fig. 2, the *MAID* output for the top five sequence assignments was (the larger the tvalue, the worse the fit):

```
seq. = 168(179) gapnum = 3 tvalue = 10.000
seq. = 145(155) gapnum = 2 tvalue = 20.714
seq. = 12(23) gapnum = 3 tvalue = 23.571
seq. = 11(23) gapnum = 4 tvalue = 23.571
seq. = 62(74) gapnum = 4 tvalue = 23.571.
```

It can be seen for this case that the second-best assignment is much worse than the first, allowing an unequivocal (and correct) sequence assignment. If these assignments were not unequivocal, then *MAID* would look for a third connected fit and repeat the process for the three fits. If a sequence can still not be assigned, then the whole procedure is repeated starting with another fit. If a sequence assignment is possible, the specific side chains are built and refined.

3.7. Connecting the two assigned 'fits'

The next step is to 'extend' one fit until it can be fused with the connected fit. For each residue addition, an array of initial trial φ/ψ angles are used. This initial array consists of four φ/ψ angles in the Ramachandran helix region, two in the δ region,

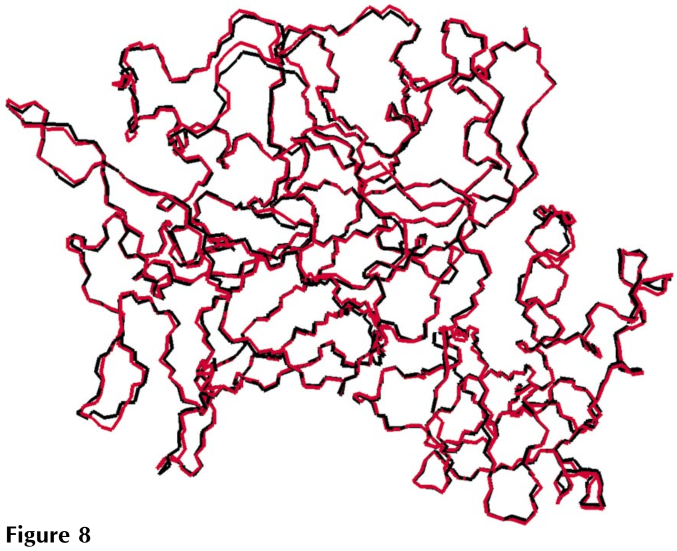

Figure 8
Comparison of the final refined main-chain structure (red) and the structure output by *MAID* (black) for the FAH map.

four in the sheet region, one in the left α-helix region and two in the ε region (Sibanda *et al.*, 1989). If the residue being added is a glycine, an additional set of six φ/ψ angles are tried. If the residue is a proline, then the starting value of φ is set to an angle of -58.0°. Using each of these starting angles and the corresponding constraints, the new residue is refined using torsional dynamics. Additions that fit minimum conditions for atom density and distance from the bone connection are ranked and saved (up to a maximum of nine). If the *trans* proline addition does not fit the density, then a *cis* proline is tried and used if it provides a better fit to the density. All of the saved fits are used for the next addition.

Fig. 3 shows an example for the extension between the two fits shown in Fig. 2. The black line is the final refined main-chain trace, the pink line is the bone connection and the red lines are the six different possible extensions that were found after the addition of three residues. All of these six fits will be used for the addition of the fourth residue. Since the 'gap' was three, this next addition should overlap the other fit. The 'best' fit is then determined by a combination of the quality of the fit to the gap density and the minimum distance of the overlapping atoms. The best fit is 'fused' with the connecting fit using torsional-angle real-space dynamics (Fig. 4). During this dynamics fusion run, a force is imposed that tries to superimpose the overlapping atoms from the two fits. The force is increased slowly and six residues on each side of the fusion position take part in the dynamics, allowing the two chains to adjust to their final position. By necessity, the residue that is fused cannot have completely ideal bond lengths and angles, in contrast to all the other residues.

3.8. Expanding the fits using the symmetry operators and generating the final subunit PDB structure

When the extension routine is completed, the *MAID* structure consists of assigned connected fits along with generic fits whose sequence could not be defined. *MAID* then searches for symmetry operators that can superimpose two fits. All of these symmetry operators are stored and applied to all the fits. If a translated/rotated fit falls on another fit, then the information from the overlapping fits is used to extend both fits to their limits. If it does not overlap a fit, but lies in the original density map, then a new fit is created. This allows *MAID* to use information from different subunits to try and synthesize a complete subunit. This is a particularly useful procedure when there is more than one subunit per asymmetric unit and the map is not averaged.

Finally, *MAID* tries to group these fits into subunits. Firstly, the longest fit is assigned to subunit 1. *MAID* then uses the edited bones to see if this fit can be connected to another fit. (This would be the case if during the 'extend' routine *MAID* found a connection but the extension failed). Finally, *MAID* assumes that the closest assigned fits whose sequences do not overlap with the sequences connected to subunit 1 are also part of subunit 1. The previously determined symmetry operators are sequentially applied to all the fits assigned to subunit 1. Any fits that are overlapped by the application of operator 1 are assigned to subunit 2. This procedure is repeated for each symmetry operator, each time assigning a new subunit.

4. Results

Fig. 5 shows a comparison of the final refined main-chain structure (red lines) and the structure output by *MAID* (black lines) for the 2.5 Å SHAD (Barycki *et al.*, 1999) SMAD map. *MAID* fitted 82% of the residues (241/292) with an r.m.s. deviation of 0.53 Å for the main-chain atoms (C^α, C, N, O) and 1.20 for all atoms. There is a 19-residue C-terminal set of residues that have very poor density that *MAID* could not fit (green arrow) and there is one gap in the *MAID* main chain (blue arrow) where the loop density dropped below 0.8 standard deviations so that *MAID* could not find the connection. In addition, there is one loop where the *MAID* fit was significantly different from the refined fit (black arrow) because the skeletonized connection follows the wrong branch. With the exception of this one loop, all the rest of the main and side chains are accurately fitted. Fig. 6 shows the r.m.s. deviation per main-chain atom (created using the routine *RMSPDB*; Kleywegt, 1999). The region with the large (5 Å) error corresponds to the bad loop (black arrow, Fig. 5). If four residues in this loop are deleted, the r.m.s.d. is 0.43 Å for the main-chain atoms and 1.03 Å for the side-chain atoms. Fig. 7 shows the Ramachandran plot (created by *PROCHECK*; Laskowski *et al.*, 1993) for this *MAID* fit (Fig. 5; four residues deleted). There is only one residue in the disallowed region and this residue (Phe205) is presumably correctly fit, since it has a similar φ/ψ in the refined structure. There are four residues in the generously allowed region and 83% of the residues are in the most favored region. Because of the procedure *MAID* uses to build the fit, all the bond lengths and angles are ideal except for the residue where two fits were fused.

MAID also was applied to the high-quality 1.9 Å SMAD for FAH (Timm *et al.*, 1999). As shown in Fig. 8, *MAID* correctly fit all the 418 residues of FAH with a 0.46 r.m.s.d. main-chain error and a 1.00 r.m.s.d. side-chain error.

5. Discussion

MAID was able to accurately fit 82% of SHAD and 100% of FAH using as input just the SMAD map and the amino-acid sequence. However, it should be emphasized that these two maps were used as the basis for program development and it will be necessary to test *MAID* over a wide range of map conditions in order to clearly establish its usefulness. The 2.5 Å SHAD map quality should be representative of the maps that are now being routinely obtained using the SMAD procedure and it is hoped that the SHAD results are typical of what can be expected of *MAID*.

A major advantage of *MAID* over a routine such as *wARP* (Perrakis *et al.*, 1999) is that it does not require high-resolution data. As is shown by the application to SHAD, an average quality 2.5 Å map can be accurately fitted by *MAID*. Although *MAID* has not been tested on map resolutions worse than 2.5 Å, it is expected that *MAID* should be applicable to maps of 2.8 Å or better. Obviously, the better the map, the more successful *MAID* will be. The most demanding step is the extension through the loops. The current version of *MAID* cannot connect two fits if there is some point in the loop where the main-chain map density drops below 0.8 standard deviations. Another limiting feature of this routine is that it is first necessary to determine the sequence assignment of two or three fits before loop extension can be attempted. In regions where the side-chain density is so poor that an unambiguous sequence assignment cannot be made, the output of *MAID* will be limited to residues with generic C—C atom side chains and no extension through the loops.

The next step on the path to a complete automated refinement is to establish protocols for combining this preliminary *MAID* fit with a partial structure-recombination routine to create an improved map that can be used in another round of input to *MAID*. The optimum approach to this has not yet been worked out.

MAID is now freely available from http://www.msi.umn.edu/~levitt. Contact DGL at levitt@dcmir.med.umn.edu for help.

I thank J. Barycki for providing the original SHAD electron-density map, D. Timm for providing the original FAH electron-density map and R. Milius for computational support. Computer software and hardware was supported by a grant from the Minnesota Supercomputing Institute for Digital Simulation and Advanced Computation.

References

Abbott, A. (2000). *Nature (London)*, **408**, 130–132.

Bae, D. & Huang, E. J. (1987). *Mech. Struct. Mach.* **15**, 359–382.

Barycki, J. J., O'Brien, L. K., Bratt, J. M., Zhang, R., Sanishvili, R., Strauss, A. W. & Banaszak, L. J. (1999). *Biochemistry*, **38**, 5786–5798.

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

Cowtan, K. D. & Zhang, K. Y. (1999). *Prog. Biophys. Mol. Biol.* **72**, 245–270.

Doublie, S. (1997). *Methods Enzymol.* **276**, 523–530.

Dunbrack, R. L. Jr & Karplus, M. (1993). *J. Mol. Biol.* **230**, 543–574.

Greer, J. (1974). *J. Mol. Biol.* **82**, 279–301.

Hendrickson, W. A. (1991). *Science*, **254**, 51–58.

Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kleywegt, G. J. (1999). Unpublished program.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Levitt, D. G. & Banaszak, L. J. (1993). *J. Appl. Crystallogr.* **26**, 736–745.

Ogata, C. M. (1998). *Nature Struct. Biol.* **5**(*Suppl.*), 638–640.

Oldfield, T. J. (2001). *Acta Cryst.* D**57**, 82–94.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Sibanda, B. L., Blundell, T. L. & Thornton, J. M. (1989). *J. Mol. Biol.* **206**, 759–777.

Smaglik, P. (2000). *Nature (London)*, **407**, 549.

Swanson, S. M. (1994). *Acta Cryst.* D**50**, 695–708.

Tame, J. R. (2000). *Acta Cryst.* D**56**, 1554–1559.

Terwilliger, T. C. (1997). *Methods Enzymol.* **276**, 530–537.

Timm, D. E., Mueller, H. A., Bhanumoorthy, P., Harp, J. M. & Bunick, G. J. (1999). *Structure Fold. Des.* **7**, 1023–1033.

Turk, D. & Guncar, G. (1999). Am. Crystallogr. Assoc. Abstr.