Simulated-annealing real-space refinement as a tool in model building

Andrei Korostelev,[a,b] **Richard Bertram**[c,b] **and Michael S. Chapman**[a,c*]

[a]Department of Chemistry and Biochemistry, Florida State University, Tallahassee, FL 32306-4380, USA, [b]Kasha Laboratory of Biophysics, Florida State University, Tallahassee, FL 32306-4380, USA, and [c]Department of Mathematics, Florida State University, Tallahassee, FL 32306-4380, USA

Correspondence e-mail: chapman@sb.fsu.edu

Methods have been developed that further automate the building of macromolecular structures into electron-density maps. The software supports molecular-dynamics real-space refinement of an atomic model to local regions of a map within the context of a popular molecular-modeling program, *O* [Jones *et al.* (1991), *Acta Cryst.* A**47**, 110–119]. It is implemented as a module to the *CNS* refinement package [Brünger *et al.* (1998), *Acta Cryst.* D**54**, 905–921], controlled by a graphical user interface and commands executed directly through the molecular-graphics package. The method is illustrated with examples of the building and rebuilding of protein and nucleic acid models in which laborious manual adjustments are avoided. The resulting models show improved convergence during subsequent reciprocal-space refinement. The novel feature of the *RSRef*2000 software is the combination of simulated-annealing optimization with local real-space refinement, allowing several local minima to be explored quickly and automatically within the context of interactive model building.

1. Introduction

Real-space refinement was successfully applied to the determination of some of the first protein structures (Diamond, 1971). Nevertheless, it is reciprocal-space methods that are now more commonly used. In conventional reciprocal-space refinement, the objective function that is optimized includes the least-squares target,

$$E_{\text{xtal}}(F) = \sum_{\mathbf{h}} [|F_{\text{calc}}(\mathbf{h})| - |F_{\text{obs}}(\mathbf{h})|]^2, \qquad (1)$$

where F_{calc} and F_{obs} are the calculated and observed structure factors for reflection $\mathbf{h}$, respectively. The advantage of this reciprocal-space target is that it is independent of poorly determined experimental phases. Shifts in atomic parameters such as coordinates and thermal factors are made by refinement programs while minimizing the objective function

$$E_{\text{total}} = E_{\text{geom}} + wE_{\text{xtal}}, \qquad (2)$$

where E_{geom} is a stereochemical (geometrical) target comprised of a number of semi-empirical terms restraining the model towards well established parameters for the covalent bond lengths and angles, torsion angles and several non-bonded atomic interactions. The crystallographic target E_{xtal} is weighted relative to the geometrical target through the weight w. There are several variants of the targets in common use. A

maximum-likelihood target is now commonly substituted for (1) (Murshudov *et al.*, 1997; Pannu & Read, 1996). E_{geom} may be specified as a collection of *ad hoc* harmonic restraints (Hendrickson, 1985) or as a semi-empirical force field (Brünger *et al.*, 1987).

In real-space refinement, the target function (1) is replaced by another defined as the squared difference of the observed and calculated electron-density values,

$$E_{xtal}(\rho) = \int_V [|\rho_{calc}(\mathbf{x})| - |\rho_{obs}(\mathbf{x})|]^2 d\mathbf{x}. \qquad (3)$$

Here, ρ_{obs} is the current electron-density map at position $\mathbf{x}$, calculated as the sum of contributions from neighboring atoms. Most implementations have followed Diamond (1971) in using Gaussian functions to approximate the density for each atom. Chapman (1995) introduced an atomic density function that was the Fourier transform of the atomic scattering factor. This has the advantage of accounting rigorously for the effects of resolution limits on the observed electron-density map. In contrast to reciprocal-space methods, in which phases may but need not be used, real-space methods have an implicit requirement for phases because they are needed for the synthesis of the ρ_{obs} electron-density map. The phases can be computed from the model or, preferably, derived from experimental data: isomorphous replacement, multiple wavelength anomalous dispersion (MAD), symmetry averaging *etc.* One of the potential benefits of real-space refinement is that implicit inclusion of phases could result in improved conditioning through an increased data-to-parameter ratio. At about 3.0 Å resolution the number of experimental diffraction amplitudes does not greatly exceed the number of refined parameters.

Real-space and reciprocal-space refinements become nearly equal if (i) appropriate weighting is used in (2), (ii) the phases are calculated from the current atomic model and (iii) the integration volume V in (3) is the entire asymmetric unit. Although real-space and reciprocal-space refinements can be nearly equivalent, this work exploits the differences that exist under conditions that commonly apply during macromolecular structure determination. Particularly important is the advantage that the real-space refinement need not be a 'global' optimization of the whole structure, but can be a 'local' refinement of a small part of a structure to the neighboring electron density.

In tests over several years, a substantial body of experience has led to the following guidelines on the use and practical limitations of real-space refinement.

(i) Final refinement is best performed in reciprocal space with the final R_{free} typically 0.01 lower than in the most favorable applications of real-space refinement. This is a consequence of the limitations of even the most accurate MAD phases (Chen *et al.*, 1999).

(ii) Initial refinement is usually best started in real space, as judged by reduced overfitting, and progress before the first manual intervention is required. This is true even when compared with maximum-likelihood reciprocal-space methods (Chen *et al.*, 1999). It is also true for all qualities of experimental phasing from MIR (Chapman & Blanc, 1997) to MAD (Chen *et al.*, 1999). In most cases, there is a point at which it is advantageous to switch from real- to reciprocal-space methods, but the point at which real-space methods cease to help depends on the quality of the experimental phases. The reduced overfitting that results from starting in real space can be rationalized in terms of the local nature of real-space refinement. Unlike reciprocal-space refinement, one region cannot adjust to compensate for the inadequacies of a remote region to restore agreement with the diffraction data (Hodel *et al.*, 1992).

(iii) Real-space refinement seems to do no harm. Experience indicates that a good strategy involves refining as much as possible in real space before switching to reciprocal space. Crystallographers have long embodied an approximation to this strategy by alternating a laborious 'manual' fitting of the model to electron density followed by automatic refinement in reciprocal space.

(iv) Real-space refinement of viruses. The very high noncrystallographic symmetry present in icosahedral viruses makes real-space refinement the method of choice for solving such structures, largely owing to the very high quality of phases resulting from high-order averaging applied to the observed map (Jones & Liljas, 1984a,b). In addition, real-space refinement is much faster than reciprocal-space methods since it is carried out on a non-crystallographic asymmetric unit which often contains 60-fold fewer atoms than the crystallographic asymmetric unit used by reciprocal-space methods (Xie & Chapman, 1996).

The goals of this work were to substantially automate the 'manual' part of the rebuilding process in the hope that the labor of structure determination would be significantly reduced and that improved input to reciprocal-space refinement would lead to improved final model quality.

There have been prior implementations of local real-space refinement within model-building programs. The *RSR* module of *FRODO* (Jones, 1978; Jones & Liljas, 1984a) moves small rigid fragments, searching for the best fit. Good stereochemistry must be restored subsequently. Oldfield has re-parameterized to a torsion-angle optimization that is particularly appropriate for broad searches of conformational space within procedures designed to automatically build atomic models (Oldfield, 2001). The routine *MAID* (Levitt, 2001) uses a real-space fitting routine with a procedure for building an initial model into an electron-density map. Our own methods (Blanc & Chapman, 1997) have focused on obtaining the very best fit of model to map. They were the first to use a model electron-density function that explicitly accounts for the resolution limits (Chapman, 1995). They also perform simultaneous refinement of the fit to the map and stereochemistry, thereby eliminating non-productive oscillating shifts that can occur when fit and stereochemistry are alternately refined.

We have arrived at our current point through several developmental steps. Our first implementation (Blanc & Chapman, 1997) was as a module for *TNT* (Tronrud *et al.*, 1987), whose gradient-descent optimizer limited the search to

a single local minimum. Our second implementation was as a module for *X-PLOR* (Brünger, 1992), demonstrating the power of an optimizer that can explore other local minima in real space (Chen *et al.*, 1999), but in an implementation suitable only for 'global' refinement of an entire asymmetric unit. In this, a third implementation as a module for *CNS* (Brünger *et al.*, 1998), support is added for refinement of 'local' regions of a model, as are interfaces between refinement and the popular *O* molecular-graphics package (Jones *et al.*, 1991). The result is our first system in which simulated annealing (Brünger *et al.*, 1987) can be applied as an aid in model building.

2. Methods

2.1. Software design and implementation

The local real-space refinement software package *RSRef*2000 is implemented as a module for *CNS* (Brünger *et al.*, 1998). The reciprocal-space energy term (1) and its derivatives are replaced with a real-space term (3) and its derivatives adapted from Chapman (1995). Thus, the refinement is restrained by the real-space target as well as the stereochemical energies that are native to *CNS* (Brünger *et al.*, 1998). The new module is implemented as a subprogram of *CNS*, requiring minor modifications to the parent *CNS* code that can be made with a patch script. With this *CNS* implementation, potential energy minimization can be performed with either conjugate-gradient (CG) or simulated-annealing (SA) methods, the latter either using Cartesian or torsion-angle dynamics (Rice & Brünger, 1994).

Although the method can refine an entire molecule, the emphasis of the current work is on 'local' refinement. The region to be refined can be specified as a 'zone' of residues or as a spherical region. Optionally, a margin of surrounding amino acids can be defined that will contribute to the energy terms but will remain fixed. Neighbors generated by crystallographic or non-crystallographic symmetry can be generated automatically. Atoms outside the refining region and margin are excluded from calculations, greatly reducing the computation time.

A user-friendly graphical interface is used for quick and simple entry of control parameters. A refinement control program is written in Python (Lutz, 1996) with a GUI editor written using the Tk graphics library (Raines & Tranter, 1999). This prepares a script that runs the refinement. The GUI editor supports specification of (i) optimization method (CG or SA with either slow cooling or rapid-quenching), (ii) coordinate system (Cartesian or torsion angle), (iii) temperature settings for SA, (iv) regions to be refined, (v) size of margin, (vi) name of file with symmetry operators and (vii) weight w for the experimental X-ray term. A portion of the GUI menu is shown in Fig. 1(*a*).

The progress of refinement is assessed with statistics that are displayed every five or ten cycles or at every SA temperature change (Fig. 1*b*). These statistics include a real-space R factor and correlation coefficient (Zhou, Wang *et al.*,

1998) that are modified for improved sensitivity compared with those used in *O* (Jones *et al.*, 1991). Root-mean-square deviations for covalent geometry are also reported at the end of refinement.

An interface to the molecular-modeling program *O* (Jones *et al.*, 1991) was also written using *O* macros and integrated with the Python control program. This provides an optional alternative means of program control. The refinement region can be selected by picking atoms that are displayed with three-dimensional graphics. Options are offered for selecting several consecutive residues or a sphere centered at an atom. The example in Fig. 1(*c*) reflects what a user sees on the *O* screen if a zone of residues is refined. The interface allows the user to quickly try alternative refinement parameters to obtain a better fit, such as annealing temperature.

Molecular dynamics can be computationally intensive. Molecular graphics are usually run on single-user stereographic workstations that may not warrant configuring with high-end computational capabilities. Acceptable turnaround for interactive use was achieved by supporting the optional running of the refinement on a remote computer. The control program automatically handles the file transfer and runs the necessary scripts on the remote computer using a secure shell.

2.2. Testing and test systems

Three representative structure determinations were used for testing. The structure of arginine kinase had previously been refined at 1.9 Å (Zhou, Somasundaram *et al.*, 1998) by alternating real-space and reciprocal-space refinements (Zhou *et al.*, 1999) until the R and R_{free} were 0.196 and 0.224, respectively. The real-space refinement methods described here were applied to the extension of this model to 1.2 Å resolution that became possible with the collection of additional synchrotron data. The 1.9 Å model was refined into the initial map available for extension, synthesized from amplitudes to 1.2 Å resolution and phases calculated from the 1.9 Å refined model. Test reflections were not included in the calculation of the maps for both arginine kinase and the test RNA structure (below). Arginine kinase is a 42 kDa enzyme, crystallized as a transition-state analog complex in space group $P2_12_12_1$, with unit-cell parameters $a = 65.4$, $b = 70.9$, $c = 80.4$ Å.

The second test involved the initial structure determination of an 18-mer RNA duplex (Lima *et al.*, 2002) crystallized as a bromine derivative in space group $P6_322$, with unit-cell parameters $a = 45.3$, $b = 45.3$, $c = 95.1$ Å. Phases were determined by MAD to 1.6 Å resolution. The initial model was a canonical A-helix of the appropriate sequence that was manually fitted as a rigid body to the electron density. This unrefined model had R and R_{free} of 0.53 and 0.56, respectively. Real-space refinement against the 1.6 Å MAD map was performed in parallel in both local and global modes, yielding similar results. It was the global mode that allowed calculation of the reciprocal-space statistics (R factors) described here. Real-space refinement was followed by phase-restrained maximum-likelihood (Murshudov *et al.*, 1997; Pannu & Read, 1996)

reciprocal-space refinement. The combined real-space and reciprocal-space refinements were compared with exclusive use of reciprocal-space refinement (from the initial model) as a control.

To test a medium-resolution refinement, an example from the tutorial for O was used that will be familiar to many: the P2 myelin protein and its 3.0 Å MIR electron-density map (Jones *et al.*, 1988). A seven-residue α-helix was fitted as a rigid group into the map and best fitting database side-chain rotamers (Ponder & Richards, 1987) were found in O. This local zone was then real-space refined.

Assessments of speed were made using workstations that would be examples of lower-end and mid-to-upper-end workstations prevalent in crystallographic laboratories at the time of publication. The lower-end machine is a 200 MHz Silicon Graphics (SG) workstation with 96 Mbyte memory. The upper-end machine is a dual-processor Digital (Compaq) Alpha server with 2 × 533 MHz CPUs and 512 Mbyte memory. Various combinations of these were tested in which one machine acted as the local graphics workstation and either the same or the other computer hosted the refinement. Disks were not cross-mounted, so times for data transfer constitute a worst-case scenario with 100 Base T network communication.

3. Results and discussion

3.1. Improvement of models

3.1.1. Arginine kinase. Visual comparison of the arginine kinase 1.9 Å model and the 1.2 Å map showed several regions that could likely be improved involving both side-chain and main-chain atoms. Improvable side chains were grouped into two categories. The first had errors in the selection of rotamer; the second had smaller errors. Both CG and SA refinements corrected the small errors of group 2 when applied to the offending residues and their two nearest neighbors. The first group was more challenging, as exemplified in Fig. 2(*a*). These regions did not yield to CG refinement (Fig. 2*a*), regardless of the weight used for fitting. They all yielded to *RSRef*2000 SA refinement (Fig. 2*a*), albeit at a variety of annealing temperatures. Statistics of these refinements are reported in Table 1.

Figure 1
(*a*) A portion of the graphical user interface for entry of refinement parameters. Many of the parameters are passed directly from O and need not be entered through the GUI. (*b*) Statistics reported during real-space refinement. (*c*) Interactive control with the *RSRef*2000–O interface. Information in bold is entered by a user. '#' signifies places where items on the molecular-graphics window are selected by mouse-picking.

Several residues refined well with an annealing temperature of 2000 K and a cooling rate of 500 K per cycle. At the other extreme, some residues required annealing at 35 000 K and a cooling rate of 10 000 K per cycle to shake the side chain out of a local energy minimum. Such strategies would be drastic in a global refinement because other regions of the model would be disrupted. It is an advantage of the local real-space approach that radical strategies can be used to escape from a

trapped configuration without disturbing the refinement of the rest of the structure. Integration with molecular graphics allows quick screening of possible refinement parameters to select the best strategy for each local region.

There were several regions in the 1.9 Å arginine kinase model that needed corrections for main-chain atoms. The required changes were modest. Both simulated-annealing and conjugate-gradient refinements were able to bring the backbone region into the density. Thus, in real-space refinement we see the same effect long seen in reciprocal space that small random errors are removed by either conjugate-gradient or simulated-annealing methods, but that larger errors require simulated annealing.

3.1.2. RNA duplex. Several nucleotides of the initial canonical A-helix required adjustment to the density. For each, simulated-annealing refinement was able to correct the structure (Fig. 2b), while conjugate-gradient refinement was not able to dislodge the model from a local minimum. Statistics are included in Table 1. Temperatures of the start of simulated annealing were between 2000 and 4000 K depending on the region. A cooling rate of 500 K per cycle was sufficient for every nucleotide.

Global and local real-space simulated-annealing refinements of the RNA generated similar results as the optimal
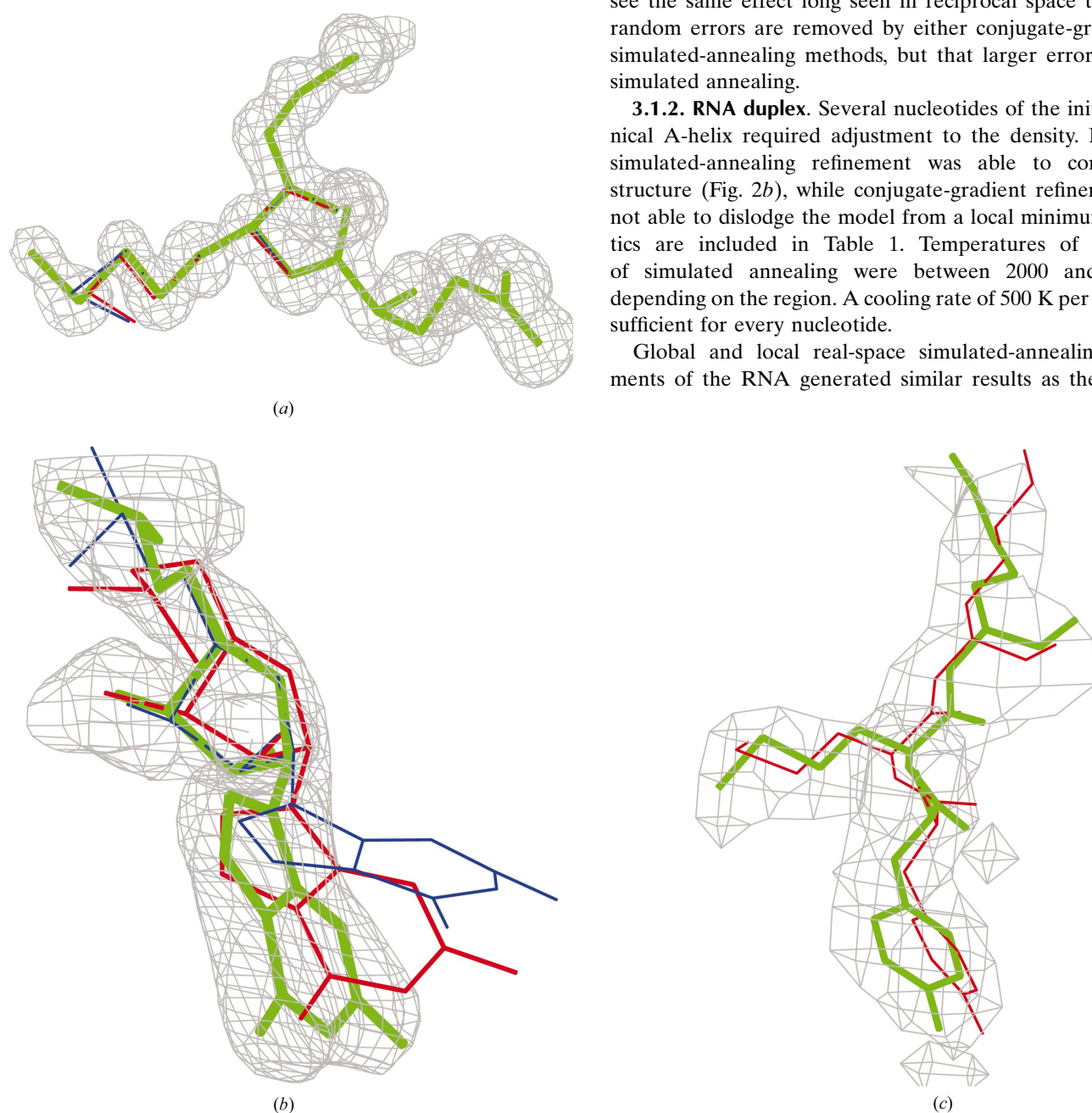
Figure 2
Examples of the refinement of models into the electron density. (a) Three residues of arginine kinase refined against a 1.2 Å resolution map: red, the starting model, refined at 1.9 Å resolution with an incorrect side-chain conformation; blue, the conjugate-gradient real-space refined model; green, simulated-annealing real-space refined model (starting temperature 35 000 K, dropping 5000 K per cycle). (b) One of the 18 nucleotides of the RNA duplex. Red, the molecular-replacement model; blue, the reciprocal-space refined model; green, the simulated-annealing real-space refined model (starting temperature 4000 K, dropping 500 K per cycle). (c) Three residues of P2 myelin protein. Red, the initial model built in O with database fragments; green, the real-space simulated-annealing refined model.

Table 1
Local real-space refinements of two different regions from arginine kinase (which required correction of a side-chain rotamer) and a nucleotide from the 18-mer RNA duplex.

	SA refinement				CG refinement			
	Real-space R		Correlation coefficient		Real-space R		Correlation coefficient	
Refining region	Start	End	Start	End	Start	End	Start	End
2 amino acids	0.49	0.34	0.927	0.955	0.49	0.38	0.927	0.941
2 amino acids	0.69	0.49	0.837	0.903	0.69	0.52	0.837	0.893
1 nucleotide	0.50	0.29	0.459	0.839	0.50	0.49	0.459	0.466

Table 2
Elapsed run times for various combinations of graphics front-ends and computer servers.

No. of refined amino acids	No. of amino acids in margin	Total atoms	Run time for SA refinement (s)			Run time for CG refinement (s)		
Local (graphics) computer			Alpha	SG	SG	Alpha	SG	SG
Host for refinement			Alpha	Alpha	SG	Alpha	Alpha	SG
2	4	53	23	55	155	20	47	102
3	4	62	25	67	203	23	51	116
4	8	92	40	74	230	31	57	126
5	11	130	53	88	278	43	67	153

annealing temperatures for each residue did not vary greatly. For global refinement, we used a starting temperature of 4500 K that was reduced by 500 K per cycle. Refinement on a remote Alpha server took 190 s. All the nucleotides in the resulting structure had an excellent fit to the electron density that was comparable to that of locally refined regions. Real-space simulated-annealing refinement reduced overfitting, yielding a drop in the reciprocal-space R and R_{free} from 0.53 and 0.56 to 0.42 and 0.42, respectively. Subsequent reciprocal-space refinement, with maximum-likelihood phase restraints, led to an R and R_{free} of 0.250 and 0.285, respectively.

These results were compared with those using a conventional methodology. This started with a manual fit of the initial model to the electron density, which took several days and brought the model to an R and R_{free} of 0.414 and 0.433, respectively. Reciprocal-space refinement of the manually improved model led to an R and R_{free} of 0.256 and 0.284, respectively, similar to our prior results (Fig. 3).

For further comparison, we used reciprocal-space refinement directly on the molecular-replacement model without any adjustment to fit the electron density (Fig. 2b). Simulated-annealing refinement yielded an R and R_{free} of 0.284 and 0.328, respectively, demonstrating that an initial fitting of the model to the electron density was required for the best refinement, performed either manually or with the real-space annealing procedure described here.

3.1.3. P2 myelin protein. This system was chosen to exemplify the use of simulated-annealing refinement within the context of typical model building. Of the seven-residue region fitted with database fragments, a segment of three amino acids was selected for refinement that had poor fit to the 3.0 Å MIR electron-density map. Local real-space simulated annealing started with a temperature of 10 000 K which was reduced by 3000 K per cycle. Fig. 2(c) shows a clear difference between *RSRef*2000 real-space refined and the database-fitted models and how a high-quality model can be quickly built by crude fragment fitting followed by fast local simulated-annealing refinement.

3.2. Speed and interactivity

In a world used to instant response over the Internet, the times reported in Table 2 correspond to noticeable pauses. However, it should be remembered that the 'manual' adjustment required to achieve similar results takes a skilled

investigator many minutes per amino acid. The times quoted for separate graphics and refinement computers are over-estimates of what will be achieved in most laboratories, since the file transfer over 100 base T networking accounts for about half of the run time and this will not be required when disks are cross-mounted. Even the ~1 min times achieved in the authors' computing environment correspond to about 20 s per amino acid, a substantial improvement over manual rebuilding. The difference between the 200 MHz SG and the 2×533 MHz Alpha tracks the improvement in computational resources of a typical crystallographic laboratory over a 5 y period. Similar advances over the next 5 y would lead to sub-10 s local refinements – approaching true interactivity.

4. Conclusions

Local simulated-annealing refinement results in improvements in the speed with which models can be built and in their quality. They have the potential to eliminate a labor-intensive bottleneck in macromolecular structure determination. This will be increasingly important in high-throughput structure-determination environments.

Equally important, the quality of the model input into conventional refinement is improved. The real-space refine-
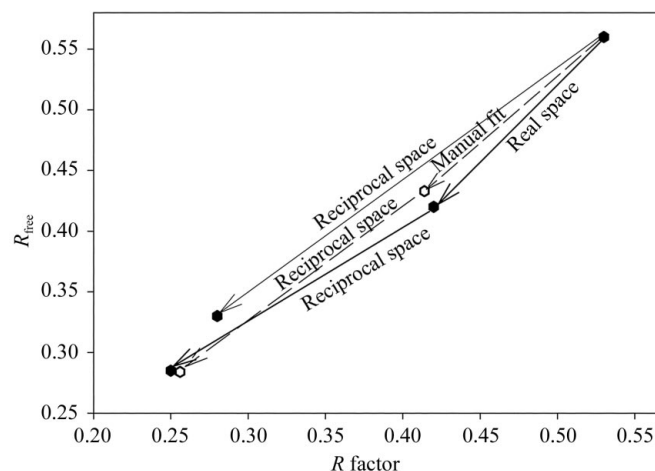
Figure 3
Progress of various strategies in the refinement of the 18-mer RNA structure.

ment consistently generates stereochemically good fits to the electron density that are at least as good as those generated by experienced model builders. Both medium- and high-resolution models may be improved with minimal effort from a crystallographer. Local simulated annealing allows aggressive strategies to be applied locally without spoiling the rest of the structure. Local refinement avoids the introduction of mutually compensating errors, so if the initial improvement is performed in real-space, less overfitting is introduced at the susceptible first stages of refinement. Together, these can have substantial impact on the speed and quality of structure determination, in some cases making the difference between success and failure in refinement.

While with the RNA structure 'blind' automatic procedures worked well, empirical adjustments to the refinement parameters were typically needed for different regions of a protein structure. Thus, the real-space annealing is most effective within the context of a molecular-graphics package. This provides rapid feedback on the choice of annealing temperature and provides complementary tools for manual adjustments and quality control (Jones & Kjeldgaard, 1997).

The *RSRef*2000 software can be downloaded under license from http://www.sb.fsu.edu/~rsref. It will also be available in future releases of the *CNS* package, http://cns.csb.yale.edu/.

References

Blanc, E. & Chapman, M. S. (1997). *J. Appl. Cryst.* **30**, 566–567.

Brünger, A. T. (1992). *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR.* New Haven, USA: Yale University Press.

Brünger, A. T., Adams, P. D., Clore, G. M., Gros, P., Gross-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.

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.

Chapman, M. S. (1995). *Acta Cryst.* A**51**, 69–80.

Chapman, M. S. & Blanc, E. (1997). *Acta Cryst.* D**53**, 203–206.

Chen, Z., Blanc, E. & Chapman, M. S. (1999). *Acta Cryst.* D**55**, 464–468.

Diamond, R. (1971). *Acta Cryst.* A**27**, 436–452.

Hendrickson, W. H. (1985). *Methods Enzymol.* **115**, 252–270.

Hodel, A., Kim, S.-H. & Brünger, A. T. (1992). *Acta Cryst.* A**48**, 851–858.

Jones, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.

Jones, T. A., Bergfors, T., Sedzik, J. & Unge, T. (1988). *EMBO J.* **7**, 1597–1604.

Jones, T. A. & Kjeldgaard, M. (1997). *Methods Enzymol.* **277**, 173–208.

Jones, T. A. & Liljas, L. (1984a). *Acta Cryst.* A**40**, 50–57.

Jones, T. A. & Liljas, L. (1984b). *J. Mol. Biol.* **177**, 735–767.

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

Levitt, D. G. (2001). *Acta Cryst.* D**57**, 1013–1019.

Lima, S., Hildenbrand, J., Korostelev, A., Hattman, S. & Li, H. (2002). Submitted.

Lutz, M. (1996). *Programming Python.* Sebastopol, CA, USA: O'Reilly & Associates.

Murshudov, G., Vagin, A. & Dodson, E. (1997). *Acta Cryst.* D**53**, 240–255.

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

Pannu, N. S. & Read, R. J. (1996). *Acta Cryst.* A**52**, 659–668.

Ponder, J. W. & Richards, F. M. (1987). *Cold Spring Harbor Symp. Quant. Biol.* **52**, 421–428.

Raines, P. & Tranter, J. (1999). *Tcl/Tk in a Nutshell.* Sebastopol, CA, USA: O'Reilly & Associates.

Rice, L. M. & Brünger, A. T. (1994). *Proteins*, **19**, 277–290.

Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* A**43**, 489–501.

Xie, Q. & Chapman, M. S. (1996). *J. Mol. Biol.* **264**, 497–520.

Zhou, G., Somasundaram, T., Blanc, E., Chen, Z. & Chapman, M. S. (1999). *Acta Cryst.* D**55**, 835–845.

Zhou, G., Somasundaram, T., Blanc, E., Parthsarathy, G., Ellington, W. R. & Chapman, M. S. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 8449–8454.

Zhou, G., Wang, J., Blanc, E. & Chapman, M. S. (1998). *Acta Cryst.* D**54**, 391–399.