

18. REFINEMENT

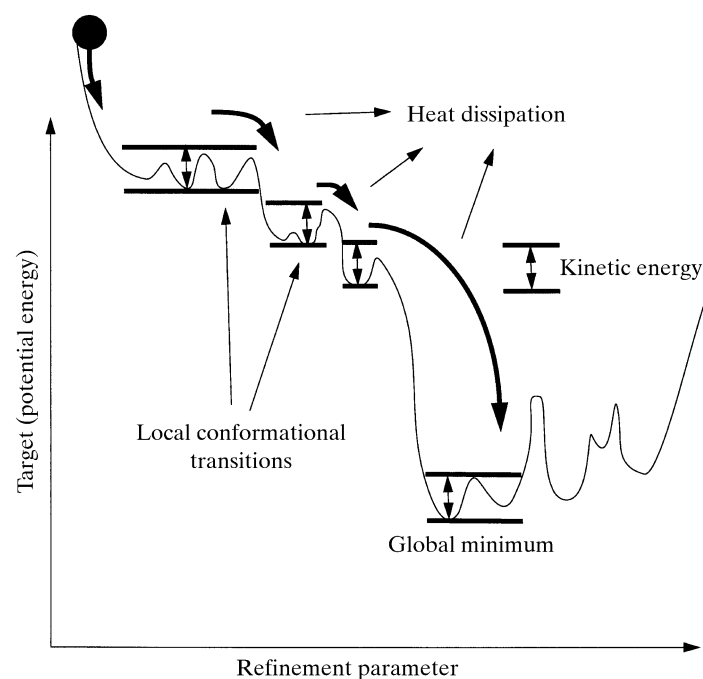


Fig. 18.2.4.1. Illustration of simulated annealing for minimization of a one-dimensional function. The kinetic energy of the system (a 'ball' rolling on the one-dimensional surface) allows local conformational transitions with barriers smaller than the kinetic energy. If a larger drop in energy is encountered, the excess kinetic energy is dissipated. It is thus unlikely that the system can climb out of the global minimum once it has reached it.

The simulated-annealing algorithm requires a mechanism to create a Boltzmann distribution at a given temperature, T , and an annealing schedule, that is, a sequence of temperatures $T_1 \geq T_2 \geq \dots \geq T_l$ at which the Boltzmann distribution is computed. Implementations differ in the way they generate a transition, or move, from one set of parameters to another that is consistent with the Boltzmann distribution at a given temperature. The two most widely used methods are Metropolis Monte Carlo (Metropolis *et al.*, 1953) and molecular dynamics (Verlet, 1967) simulations. For X-ray crystallographic refinement, molecular dynamics has proven extremely successful (Brünger *et al.*, 1987) because it limits the search to physically reasonable 'moves'.

18.2.4.1. Molecular dynamics

A suitably chosen set of atomic parameters can be viewed as generalized coordinates that are propagated in time by the classical equations of motion (Goldstein, 1980). If the generalized coordinates represent the x , y , z positions of the atoms of a molecule, the classical equations of motion reduce to the familiar Newton's second law:

$$m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = -\nabla_i E. \quad (18.2.4.1)$$

The quantities m_i and \mathbf{r}_i are, respectively, the mass and coordinates of atom i , and E is given by equation (18.2.3.1). The solution of the partial differential equations (18.2.4.1) can be achieved numerically using finite-difference methods (Verlet, 1967; Abramowitz & Stegun, 1968). This approach is referred to as molecular dynamics.

Initial velocities for the integration of equation (18.2.4.1) are usually assigned randomly from a Maxwell distribution at the appropriate temperature. Assignment of different initial velocities will generally produce a somewhat different structure after simulated annealing. By performing several refinements with

different initial velocities, one can therefore improve the chances of success of simulated-annealing refinement. Furthermore, this improved sampling can be used to study discrete disorder and conformational variability, especially when using torsion-angle molecular dynamics (see below).

Although Cartesian (*i.e.* flexible bond lengths and bond angles) molecular dynamics places restraints on bond lengths and bond angles [through E_{chem} , equation (18.2.3.1)], one might want to implement these restraints as constraints, *i.e.*, fixed bond lengths and bond angles (Diamond, 1971). This is supported by the observation that the deviations from ideal bond lengths and bond angles are usually small in macromolecular X-ray crystal structures. Indeed, fixed-length constraints have been applied to crystallographic refinement by least-squares minimization (Diamond, 1971). It is only recently, however, that efficient and robust algorithms have become available for molecular dynamics in torsion-angle space (Bae & Haug, 1987, 1988; Jain *et al.*, 1993; Rice & Brünger, 1994). We chose an approach that retains the Cartesian-coordinate formulation of the target function and its derivatives with respect to atomic coordinates, so that the calculation remains relatively straightforward and can be applied to any macromolecule or their complexes (Rice & Brünger, 1994). In this formulation, the expression for the acceleration becomes a function of positions and velocities. Iterative equations of motion for constrained dynamics in this formulation can be derived and solved by finite-difference methods (Abramowitz & Stegun, 1968). This method is numerically very robust and has a significantly increased radius of convergence in crystallographic refinement compared to Cartesian molecular dynamics (Rice & Brünger, 1994).

18.2.4.2. Temperature control

Simulated annealing requires the control of the temperature during molecular dynamics. The current temperature of the simulation (T_{curr}) is computed from the kinetic energy

$$E_{\text{kin}} = \sum_i^n \frac{1}{2} m_i \left(\frac{\partial \mathbf{r}_i}{\partial t} \right)^2 \quad (18.2.4.2)$$

of the molecular-dynamics simulation,

$$T_{\text{curr}} = 2E_{\text{kin}}/3nk_B. \quad (18.2.4.3)$$

Here, n is the number of atoms, m_i is the mass of the atom and k_B is Boltzmann's constant. One commonly used approach to control the temperature of the simulation consists of coupling the equations of motion to a heat bath through a 'friction' term (Berendsen *et al.*, 1984). Another approach is to rescale periodically the velocities in order to match T_{curr} with the target temperature.

18.2.4.3. Annealing schedules

The simulated-annealing temperature needs to be high enough to allow conformational transitions, but not so high that the model moves too far away from the correct structure. The optimal temperature for a given starting structure is a matter of trial and error. Starting temperatures that work for the average case have been determined for a variety of simulated-annealing protocols (Brünger, 1988; Adams *et al.*, 1997). However, it might be worth trying a different temperature if a particularly difficult refinement problem is encountered. In particular, significantly higher temperatures are attainable using torsion-angle molecular dynamics. Note that each simulated-annealing refinement is subject to 'chance' by using a random-number generator to generate the initial velocities. Thus, multiple simulated annealing runs can be carried out in order to increase the success rate of the refinement. The best structure(s) (as determined by the free R value) among a set of refinements using