Genetics helping molecular dynamics

Analogy with natural evolution is the guiding principle of an essay in molecular dynamics that successfully predicts the structure of C_{60}, or buckminsterfullerene.

THE growing company of researchers in the physical sciences with an interest in solving problems in biology has frequently been applauded for its daring and ingenuity, as in the modelling of the cell cycle, the exploration of the dynamics of the self-assembly of molecular systems (the spontaneous construction of microtubules from tubulin, for example) and the specification of neural networks as analogues of how circuitry in the brain may work. Less attention has been paid to the occasions when biology has something substantial to offer physics.

Of course, nanotechnologists have already been stimulated to interesting rumination and experiment by the ubiquity of self-assembly in biological systems, at synapses for example. But here is something more tangible — the use of evolutionary principles to avoid the worst pitfalls in the technique of molecular dynamics, the now familiar way of telling the equilibrium configuration of a system of atoms by choosing that among all possible configurations whose energy is the least.

The obvious difficulty is that the number of possible configurations increases with the number of atoms in the system too quickly for the comfort of any finite computer; for a system of N atoms, all distinguishable, the number of ways in which they can be rearranged among each other will be of the order of $N!$. The more subtle, but also well known, difficulty is that many physical systems are distinguished by a large number of local energy minima in configuration-space. Computer programs designed to find the global energy minimum by means of small variations of the parameters of the atomic system will then be at a loss to tell whether they have found the true global minimum or some other.

The standard technique for escaping this pitfall, called 'simulated annealing', is figuratively to add a finite amount of kinetic energy to the system and then to take it away again in small decrements. If a false minimum is separated from the global minimum by a modest potential barrier, the computer model of the system may find its way to the true minimum, and its equilibrium configuration will be determined. There are two snags. Simulated annealing is a great thief of computer time, and there is no objective way of testing whether the energy minimum eventually found is truly the global minimum.

This is where biology may help. Molecular dynamics seeks to find an optimum configuration, one whose free energy is the least. Evolution by natural selection is also a means by which the consequences of the genetic attributes of a species (its phenotype) are optimized to the local environment. But natural selection has the advantage over simulated annealing that sexual reproduction is forever offering novel genetic arrangements on which selection can be allowed to work. So why not use the analogy of evolution to find energetic optima for an atomic system?

The idea is far from new, but D. M. Deaven and K. M. Ho, from the US Department of Energy's Ames Laboratory in Iowa, have now applied it to the 'prediction' of the well known truncated icosahedral structure of molecular C_{60}, otherwise buckminsterfullerene (*Phys. Rev. Lett.* **75**, 288–291; 1995). They say attempts to solve this structure by conventional molecular dynamics have been unsuccessful.

There is more to the procedure than putting 60 carbon atoms in a box, specifying a field of force between them and then seeing what configuration emerges in the course of time. Because the procedure is meant to simulate evolution, there has to be a population of trial molecules at each stage of the enterprise. And, crucially, pairs of molecules must be able to 'mate' with each other to produce a 'child'.

How? Deaven and Ho follow a simple rule: they draw the same randomly chosen imaginary plane through the centre of gravity of the two partners to the mating, stick the top half of one and the bottom half of the other together to form the child of the mating and then put the child through conventional molecular dynamics. Because the two halves may not have 30 carbon atoms each, the program will displace the midplanes by equal but opposite distances until the child contains just 60 atoms.

Various refinements make the program manageable. If all possible matings were allowed, the number of the candidate molecules in the 'population' would be squared at each new generation, but good sense demands that it should be constant. That is done by choosing the partners for each mating by a numerical rule embodying a bias in favour of potential 'parent' molecules whose energy is towards the low end of the whole range. (Deaven and Ho use as a weighting rule a simple Boltzmann negative exponential in which the numerator of the exponent is the energy of the putative parent and the denominator is the total range of the energy of the parents.)

That rule is the program's analogue of natural selection; low-energy configurations will contribute preferentially to the next generation, but the chance that high-energy configurations will have an influence is not excluded. Evidently, that is how the algorithm concentrates on the evolution of the putative molecular structure towards a particular minimum while taking account of regions of configuration space far away from what may be a possibly spurious local minimum.

The outcome is pleasing. Starting with a variety of initial configurations, the authors reach the buckyball configuration after 5,500 mating generations. Interestingly, the algorithm appears to hang up on various incorrect intermediate structures; in one illustration, a form of C_{60} in which two pentagons are adjacent to each other was converted into true buckminsterfullerene only after 2,000 mating generations. The calculated binding energy is 9.4 eV per atom, not far from reality.

One snag with the approach is that, while C_{60} apparently gives an unambiguous result, superficially simpler molecules such as C_{20} do not; the simple evolutionary algorithm yields a structure in which 20 carbon atoms form a simple circular ring. So the authors describe a process called 'mutation' in which the positions of the atoms of an evolving cluster are periodically altered, in a predetermined fraction of an offspring generation, according to one of other of two recipes. (Either all atoms are sent on a randomly specified brownian random walk, or some configurations are moved against the force gradient into a neighbouring watershed of potential.)

That way, the authors can be sure of including the influence of radically different structural classes, which will come to the fore only if they are energetically favourable. For C_{20}, the authors thus recapture the cap-like structure formed when six benzene rings are cemented together around a pentagon to form one end of a buckyball.

It will be interesting to see what can be made of more complicated molecules. The technique seems ideally suited to that old biological question of protein folding. The mating procedure would seem well designed to make the best of circumstances in which a protein evolving towards its native condition is represented by well-folded half-chains in separate members of some generation. But even as things are, computers are probably not powerful enough to say very much. **John Maddox**