PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

Phil. Trans. R. Soc. A (2005) **363**, 357–377
doi:10.1098/rsta.2004.1497
Published online 13 December 2004

The energy landscape as a unifying theme in molecular science

By David J. Wales

University Chemical Laboratories, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK (dw34@cam.ac.uk)

The potential energy surface (PES) underlies most calculations of structure, dynamics and thermodynamics in molecular science. In this contribution connections between the topology of the PES and observable properties are developed for a coarse-grained model of virus capsid self-assembly. The model predicts that a thermodynamically stable, kinetically accessible icosahedral shell exists for pentameric building blocks of the right shape: not too flat and not too spiky. The structure of the corresponding PES is probably common to other systems where directed searches avoid Levinthal's paradox, such as 'magic number' clusters, protein folding and crystallization.

Keywords: energy landscape; self-assembly; protein folding; magic numbers

1. Introduction

The purpose of this overview is to illustrate how the potential energy surface (PES) provides a common ground for theories of structure, dynamics and thermodynamics in molecular science (Wales 2003). For a system of N atoms the potential energy, V, depends on $3N$ Cartesian coordinates, $X_1, X_2, \ldots, X_{3N}$. The most interesting features of this high-dimensional function are usually the stationary points, where the gradient vanishes. At such points the leading terms in a Taylor expansion of V are quadratic, and the potential energy can be written as

$$\frac{1}{2}\sum_{\alpha=1}^{3N}\omega_\alpha^2 Q_\alpha^2$$

in terms of $3N$ normal coordinates, Q_α, correct to second order. Here ω_α^2 is an eigenvalue of the mass-weighted second derivative matrix, and a displacement in the corresponding coordinate Q_α raises or lowers the energy if ω_α^2 is positive or negative, respectively. The number of negative eigenvalues, or Hessian index, can therefore be used to further classify stationary points of the PES.

At a local minimum $\omega_\alpha^2 \geqslant 0$ for all α, and small displacements must raise the energy, unless they correspond to normal coordinates where $\omega_\alpha^2 = 0$, such as overall rotations and translations of an isolated molecule. Geometrically, a transition state

One contribution of 17 to a Discussion Meeting 'Configurational energy landscapes and structural transitions in clusters, fluids and biomolecules'.

is defined as a stationary point of Hessian index one (Murrell & Laidler 1968), and the unique negative Hessian eigenvalue is associated with displacements that lower the energy. The corresponding Hessian eigenvector corresponds to the reaction coordinate.

Empirical observations and theoretical arguments suggest that the number of local minima with distinct structures, $n_{\min}(N)$, will generally grow exponentially with N, i.e. $n_{\min}(N) = \exp(aN)$ (Doye & Wales 2002; Stillinger & Weber 1984). Similarly, the number of transition states, $n_{\rm ts}(N)$, is predicted to increase as $n_{\rm ts}(N) = aN\exp(bN)$, where both a and b are system-dependent constants. The number of higher index saddles increases combinatorially (Wales & Doye 2003), and additional factors, depending factorially on the number of atoms, must be included to account for permutational isomerism (Wales 2003).

For small molecules local minima of the PES can often be identified with distinct isomers, which are generally treated as separate free energy minima. Similarly, geometrical transition states can be used to define dividing surfaces of dimension $3N-1$ that represent dynamical bottlenecks, or free energy maxima, between local minima. Statistical rate theories can then be formulated to calculate the corresponding rate constants (Evans & Polanyi 1935; Eyring 1935; Pelzer & Wigner 1932).

For larger systems the connection between local minima and transition states of the free energy and stationary points of the underlying PES is generally more complicated. In discussing free energy surfaces it is always necessary to define order parameters of some sort, and in this process the dimensionality of the $3N$ atomic degrees of freedom is usually reduced to just one or two coordinates. Minima and transition states of the free energy then subsume many stationary points of the PES. This connection may be quantified using the superposition approach to thermodynamics (Wales 1993, 2003), in which the total partition function is written as a sum over local minima of the PES. A similar philosophy has recently been used in the discrete path sampling (DPS) approach to dynamics (Wales 2002).

The necessity to think of ensembles of pathways between free energy minima in large systems is included in various theoretical models (Bryngelson & Wolynes 1989; Bryngelson *et al.* 1995; Camacho & Thirumalai 1993; Dill *et al.* 1995; Doye & Wales 1996; Gō 1983; Onuchic *et al.* 2000), and in the notion of a 'folding funnel' as a set of kinetically convergent pathways (Leopold *et al.* 1992). The latter definition was first introduced in the context of free energy surfaces, but can also be realized in terms of pathways on the PES, as described in § 3 and § 4. The term 'energy landscape' was probably first used for free energy surfaces (Bryngelson & Wolynes 1987, 1989, 1990; Bryngelson *et al.* 1995; Dobson *et al.* 1998; Leopold *et al.* 1992; Onuchic *et al.* 1995, 1997), and subsequently extended to potential energy surfaces (Wales 2003). In the protein-folding literature the concept of an ensemble of pathways became known as the 'new view', and was contrasted with the 'old view' in terms of 'an ordinary chemical reaction that has a defined series of intermediates and a single rate-limiting step' (Baldwin 1994). In fact, we can still continue to think of a reaction path for large systems if we simply extend the definition to include an ensemble of related paths, all leading to some given free energy minimum (Lazaridis & Karplus 1997; Wales 2003).

The following sections provide illustrations of how stationary points of the PES can be used to visualize the potential energy landscape, and to calculate thermodynamic and dynamic properties of interest. In particular, some new results for a coarse-

grained model of virus capsid self-assembly are used to illustrate how landscapes with a 'palm tree' topology lead to efficient structural relaxation. This topology provides a fundamental connection between the non-random searches that result in 'magic number' clusters, crystallization, protein folding and self-assembly. In each case the form of the underlying PES results in a directed search, thereby circumventing Levinthal's paradox (Levinthal 1969).

2. Locating stationary points and pathways

Calculating thermodynamic and dynamic properties from stationary point data can provide access to much longer time-scales than conventional molecular dynamics. Ergodic results can also be achieved for systems where relevant portions of configuration space are separated by large potential energy barriers (Doye & Calvo 2002; Doye *et al.* 1998, 1999*a*). Such calculations are feasible because large steps can be taken when sampling stationary points, and the geometry optimization techniques involved are not generally subject to trapping in deep potential wells. All the vibrational degrees of freedom are then treated using standard tools from statistical mechanics and rate theory, which introduces additional approximations into the calculations. In particular, treatments based upon samples of stationary points often employ harmonic densities of states, which are usually straightforward to calculate, but can result in systematic errors. However, various techniques can be used to allow for anharmonicity, sampling errors and quantum effects (Calvo *et al.* 2001*a*,*b*; Clary 2001; Doye & Wales 1995; Miller & Clary 2003; Miller *et al.* 1999; Wales 1993).

The two principal tools employed for searching the potential energy surface in the present research were GMIN for global optimization and OPTIM for general geometry optimization. Global optimization is often a useful first step before investigating more general properties of a PES, since the global potential energy minimum will become the global free energy minimum at sufficiently low temperatures. The GMIN program, for which a public domain version is available on the World Wide Web (Wales et al. 2004), is based on the basin-hopping approach (Li & Scheraga 1987, 1988; Wales & Doye 1997). OPTIM, on the other hand, can locate local minima, transition states and characterize steepest-descent pathways using a wide variety of different algorithms.

The simplest geometry optimization problem is the location of a local minimum, and the most efficient method currently available seems to be Nocedal's L-BFGS algorithm (Liu & Nocedal 1989; Nocedal 1980). Finding transition states is significantly harder, because the potential energy has a local maximum in one degree of freedom, but is a minimum in all the other $(3N-1)$ degrees of freedom. The system is therefore balanced on a knife-edge along the reaction coordinate. Characterizing such regions away from local minima therefore presents more difficult challenges to theory, while experiments such as high-resolution spectroscopy of tunnelling states (Pugliano & Saykally 1992; Saykally & Blake 1993) and ultrafast electron diffraction (Thomas 2004; Vigliotti *et al.* 2004) can now provide detailed information for comparison.

Many different schemes have been suggested to find transition states (Schlegel 2003; Wales 2003), and they can usefully be divided into single-ended and double-ended approaches. Single-ended methods locate a transition state given a single starting point geometry, while double-ended methods seek a transition state or a more

general path between two end points. In OPTIM all the single-ended transition state search strategies are based upon eigenvector-following (Cerjan & Miller 1981; Pancíř 1975; Wales 1994), which provides a systematic way to move uphill in one direction, while minimizing with respect to all the other degrees of freedom. Hybrid eigenvector-following schemes enable transition states to be located without diagonalizing the second derivative matrix, or without calculating second derivatives at all (Kumeda *et al.* 2001; Munro & Wales 1999). Double-ended searches in OPTIM now employ the doubly nudged elastic band (DNEB) algorithm (Trygubenko & Wales 2004).

For every transition state of the PES two distinct steepest-descent paths are defined in terms of a first-order differential equation (Wales 2003). These paths, like all the stationary points themselves and their properties, are independent of mass, temperature and the coordinate system (Banerjee & Adams 1992; Wales 2000). The two steepest-descent paths usually terminate at two local minima, which can be located by energy minimization following small displacements parallel and antiparallel to the reaction vector. Such calculations generally deviate from the true steepest-descent path, and can occasionally lead to different local minima. More accurate characterization of the steepest-descent paths is certainly possible but is not usually necessary for simply determining the connections. These connections between local minima, mediated by transition states, provide the starting point for visualization of the global PES using disconnectivity graphs, as described in the next section.

3. Disconnectivity graphs

The problem of visualizing the $3N$-dimensional potential energy function, without misrepresenting the connectivity and topology, has been addressed in several ways (Ball *et al.* 1996; Becker & Karplus 1997; Berry & Breitengraser-Kunz 1995; Kunz & Berry 1995). Here we will follow the disconnectivity graph approach of Becker & Karplus (1997), who first applied it to a database of minima and transition states calculated for the tetrapeptide isobutyryl-(ala)$_3$-NH-methyl by Czerminski & Elber (1990). Many different systems have since been analysed in this way (Wales 2003), and graphs based upon free energy have also been described (Evans & Wales 2003; Krivov & Karplus 2002).

Given a database of local minima, transition states and pathways from the PES, a disconnectivity graph is constructed as follows. For a given total energy, E, the local minima can be divided into disjoint sets, or 'superbasins'. The members of each set can all be interconverted by one or more rearrangements without exceeding the threshold energy, while the minimum energy required to interconvert members of different sets is greater than E. The superbasin analysis is performed for specified energies, $E_1 < E_2 < E_3 < \cdots$, and each distinct superbasin is represented as a point, or node, on the horizontal axis. The vertical axis corresponds to energy, and lines are drawn upwards starting from the potential energy of each local minimum, V_a, to the node corresponding to the appropriate superbasin for the closest $E_i > V_\mathrm{a}$. Lines are also drawn to connect nodes at adjacent energies if they correspond to the same superbasin or to superbasins that merge together.

Some examples are shown in figure 1. Graphs such as these are termed trees, and this nomenclature suggests a helpful way to describe the various patterns that emerge (Wales *et al.* 1998). A PES with a well-defined global minimum and relatively small downhill barriers, which is associated with efficient relaxation, resembles a 'palm

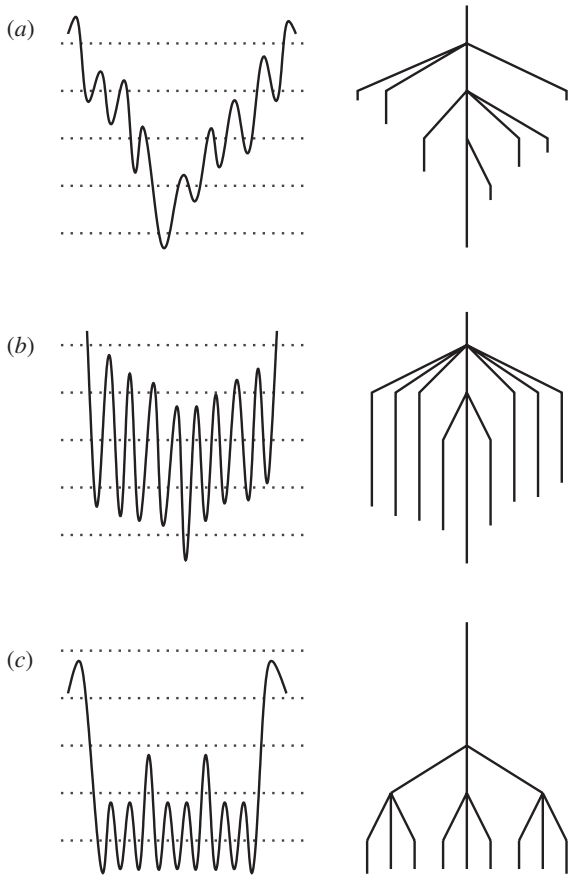

Figure 1. One-dimensional potential energy functions (left) and the corresponding disconnectivity graphs (right). The dotted lines indicate the energies at which a superbasin analysis was performed. (*a*) 'Palm tree'; (*b*) 'willow tree'; (*c*) 'banyan tree'.

tree' (figure 1*a*). This motif has been identified as a potential energy 'funnel' in previous work, in accord with the notion of a 'folding funnel' for a protein as a set of kinetically convergent pathways (Leopold *et al.* 1992). The second, 'willow tree', pattern, in figure 1*b*, results when the barriers between successive local minima are larger; the long dangling branches result from the higher activation energies involved. This pattern has been identified in C_{60} fullerenes (Kumeda & Wales 2003; Wales *et al.* 1998); efficient relaxation to the global minimum is still possible if the system has sufficient energy to overcome the higher barriers. However, for the final, 'banyan tree', pattern, in figure 1*c*, the energy barriers are large compared with the separation of connected local minima, and the two distinct barrier heights produce a hierarchical structure. In contrast to the palm and willow motifs, cutting a single edge in this graph separates a number of local minima from the rest of the tree. Relaxation to the global minimum for such surfaces is expected to be much slower, and this structure has been identified in the $(H_2O)_{20}$ cluster modelled by a rigid water monomer empirical potential (Wales 2003; Wales *et al.* 1998, 2000). In this system the hierarchical structure probably results from the interplay between centre-

of-mass and orientational degrees of freedom, which also makes global optimization significantly harder in water clusters than for most atomic clusters of comparable size (Wales & Hodges 1998).

The calculation of thermodynamic and dynamic properties from databases of stationary points generally requires additional information, such as vibrational frequencies, in order to evaluate partition functions and rate constants (Wales 2003). However, some of the resulting features can be predicted quite easily from the corresponding disconnectivity graphs, without detailed calculation. A particularly important case occurs when the PES supports more than one 'palm tree'-like feature. These separate potential energy funnels generally result in rapid relaxation to one of the corresponding minima. However, relaxation to the true global minimum may be much slower if the system first falls into one of the other funnels, which act as kinetic traps. Distinct palm tree features in the disconnectivity graph can therefore be associated with competition between different morphologies, and a separation of time-scales for relaxation. Many examples of such competition have been identified, including binary salt clusters (Doye & Wales 1999*a*,*b*) and the three-colour, 46-bead model protein (Evans & Wales 2003; Honeycutt & Thirumalai 1990, 1992; Miller & Wales 1999).

The 38-atom cluster bound by pairwise Lennard-Jones interactions (Jones & Ingham 1925) provides a particularly clear example of a double-funnel PES (Calvo *et al.* 2000; Doye *et al.* 1999*a*; Frantz 2001; Neirotti *et al.* 2000; Wales 2003; Wales *et al.* 1998). The two funnels correspond to local minima based upon either icosahedral or face-centred-cubic (FCC) packing, and the global potential energy minimum is the FCC truncated octahedron (figure 2). At low temperature the global free energy minimum is also the truncated octahedron, but icosahedral structures are favoured by entropy, and a solid–solid transition occurs at higher temperature (Calvo *et al.* 2000; Doye *et al.* 1999*a*; Frantz 2001; Neirotti *et al.* 2000). This transition is marked by a noticeable feature in the heat capacity (Calvo *et al.* 2000; Neirotti *et al.* 2000), which is somewhat exaggerated in a harmonic superposition calculation (figure 3). The two-stage relaxation is clear from figure 4, where the occupation probabilities of the icosahedral and FCC minima are plotted as a function of time starting from a high-energy non-equilibrium distribution (Doye & Wales 1998; Doye *et al.* 1998). The first direct phase of relaxation corresponds to a time of around 10^9 reduced units, but equilibrium is only achieved on a time-scale of about 10^{14} units. The slower component corresponds to initial relaxation into the icosahedral region of configuration space, followed by subsequent escape into FCC structures.

4. Self-assembly of a virus capsid

Reversible self-assembly has been demonstrated for a number of different virus capsids (Fraenkel-Conrat & Williams 1955; Kiselev & Klug 1969), and the corresponding thermodynamics is an active area of research (Bruinsma *et al.* 2003; Ercolani 2003). Crick & Watson (1956) initially proposed that virus capsids might consist of identical subunits, and Caspar & Klug (1962) provided a geometrical theory for the construction of capsids with icosahedral symmetry. Electron microscopy, X-ray diffraction and electron-diffraction experiments have since provided direct evidence for such packing, although other structures are certainly possible (Ehrlich *et al.* 2001; Ganser *et al.* 1999). This work has also identified two levels of structure within most virus

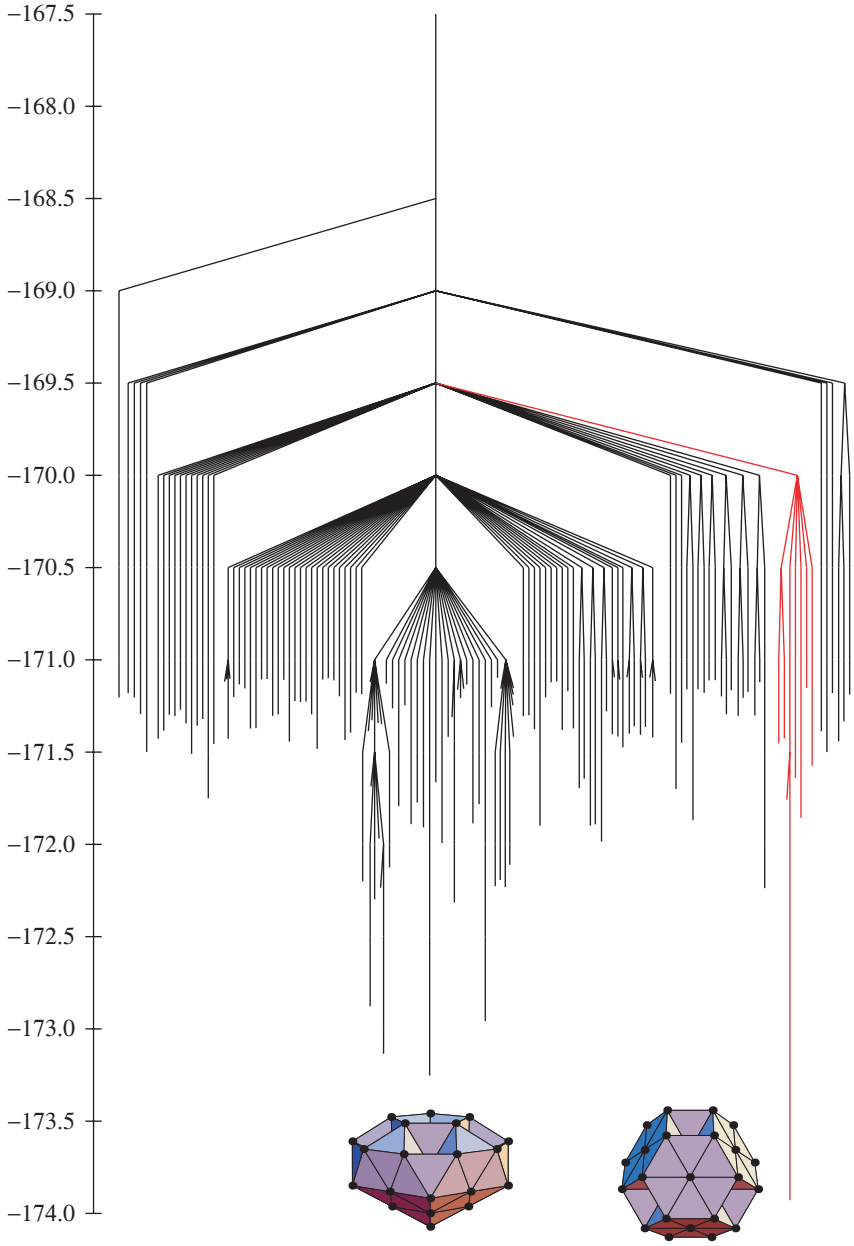

Figure 2. Disconnectivity graph for the LJ_{38} cluster including the lowest 200 local minima. The global potential energy minimum lies at the bottom of the 'palm tree' feature highlighted in red. The energy is in reduced units for the Lennard-Jones potential.

capsids in terms of capsomers, which are each composed of five or six subunits. The intermolecular forces both between the capsomers and between the subunits are non-covalent, but the capsomer–capsomer interactions are probably weaker (Davis *et al.* 1980; Eiserling & Dickson 1972). Hence, the present model considers rigid pentameric units as the fundamental building blocks.

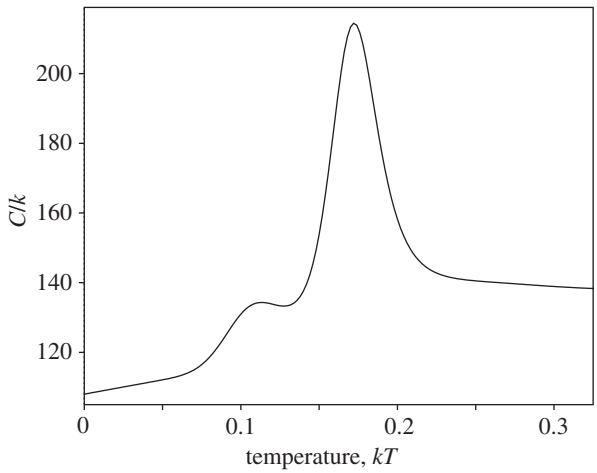

Figure 3. Canonical heat capacity of the LJ_{38} cluster calculated using the harmonic superposition approximation. kT is measured in reduced units for the Lennard-Jones potential.

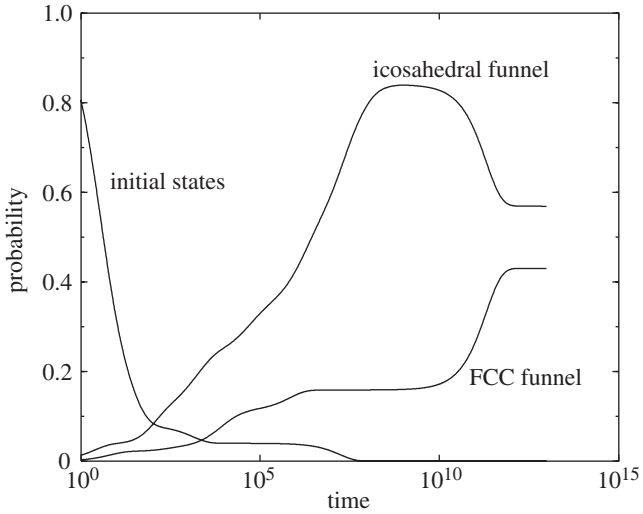

Figure 4. Relaxation of LJ_{38} from high-energy minima at a total energy of -160, showing the fast and slow contributions to the final probability of the FCC funnel. All quantities are in reduced units for the Lennard-Jones potential.

The simplest capsid structure corresponds to 12 pentamers in an icosahedral arrangement, as exemplified by satellite tobacco necrosis virus. The first objective of this preliminary study was to determine minimal requirements on the interaction potential between the pentameric building blocks for efficient relaxation to an icosahedral global minimum. Each pentamer was therefore treated as a rigid pentagonal pyramid of height h and radius r, as shown in figure 5. The total potential energy was defined in terms of pairwise additive terms between sites fixed within each rigid body. The objective was to provide the simplest possible representation of the pentamer's shape, and the total potential energy for each pair was therefore written

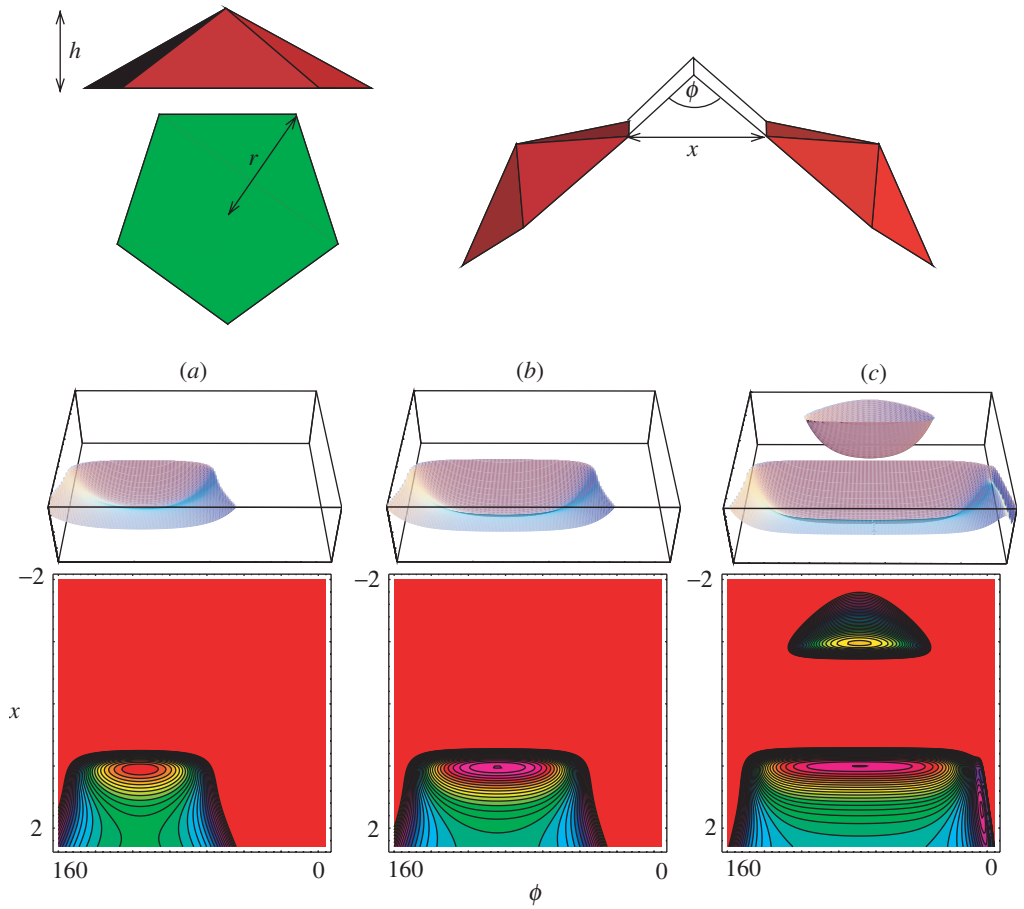

Figure 5. Definition of the geometrical parameters h and r that define the pentagonal pyramids used to model pentameric capsomers: (*a*) $h = 0.35r$; (*b*) $h = 0.50r$; (*c*) $h = 0.75r$. The potential energy for two such pyramids is illustrated in both surface and contour plots as a function of the angle ϕ and edge-to-edge distance x defined on the right. This PES varies systematically as the height of the repulsive site changes, as shown in the lower panels. For the largest value of h an unphysical secondary minimum appears corresponding to interpenetrating pyramids.

as

$$V = \varepsilon_{\mathrm{rep}}\left(\frac{\sigma}{R_{\mathrm{ax}}}\right)^{12} + \varepsilon\sum_{i=1}^{5}\sum_{j=1}^{5}(\mathrm{e}^{\rho(1-R_{ij}/R_{\mathrm{e}})} - 2)\mathrm{e}^{\rho(1-R_{ij}/R_{\mathrm{e}})},$$

where R_{ax} is the distance between the apices of the two pyramids and R_{ij} is the distance between basal sites i and j in different pentamers. ε and R_{e}, the pair well depth and equilibrium separation for each of the Morse contributions in the sum, will henceforth be used as the units of energy and distance. The Morse form was chosen for the attractive terms because it includes a steep repulsive wall and an attractive well, and allows the units to dissociate. The parameter ρ governs the range of the Morse interaction (Braier *et al.* 1990; Wales 2003; Wales *et al.* 2000), and was fixed

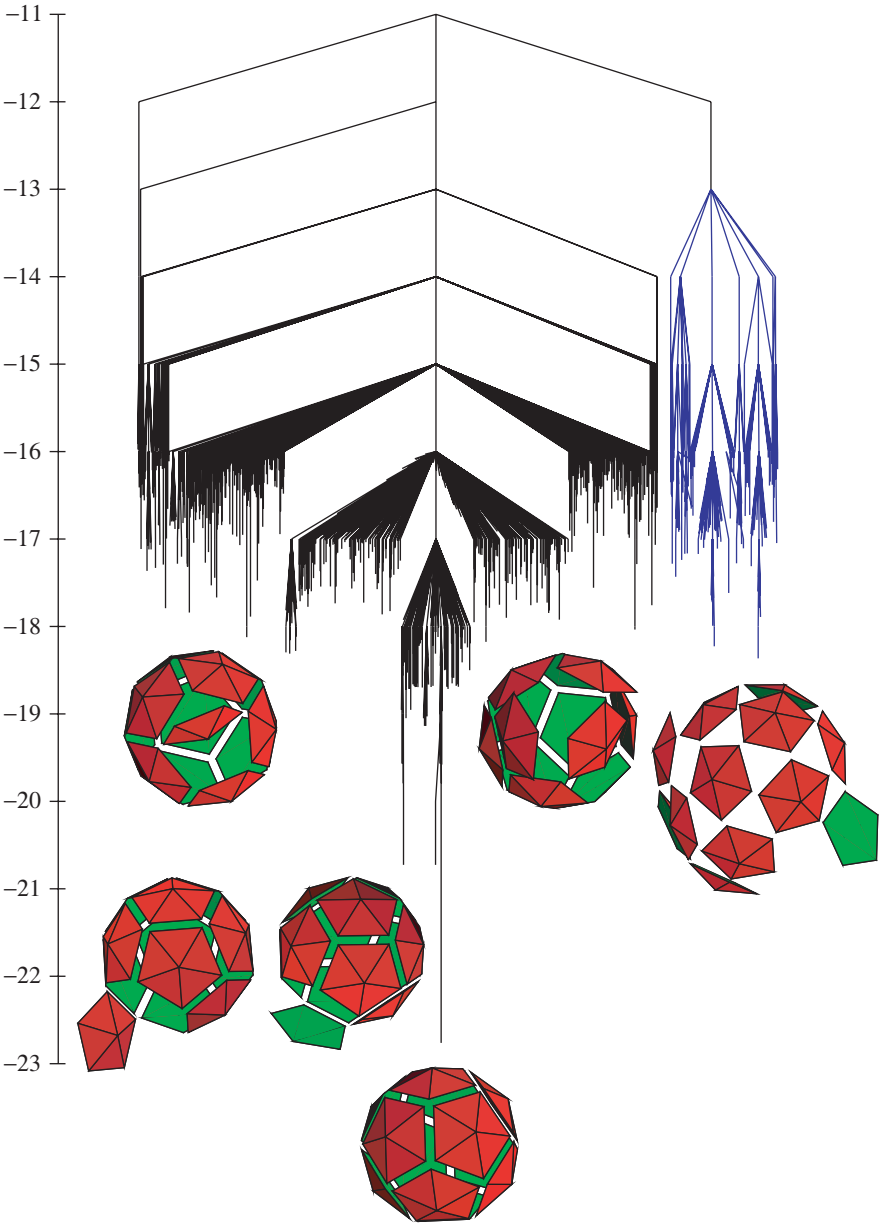

Figure 6. Disconnectivity graph for 12 pentagonal pyramids with $h = 0.35r$. A side-funnel associated with more open shells is highlighted on the right in blue.

at $\rho = 3$ for the present calculations, while the radius was fixed at $r = 5$ and $\varepsilon_{\mathrm{rep}}$ was fixed at $\varepsilon/2$. Similar results were obtained for different values of these three parameters, while the height of the pentamer, h, has more interesting effects.

Without the axial repulsive site the pentamers would simply aggregate with their faces together. However, the presence of the repulsive site at a non-zero height endows the pentamers with a three-dimensional shape. For two pentamers arranged as shown

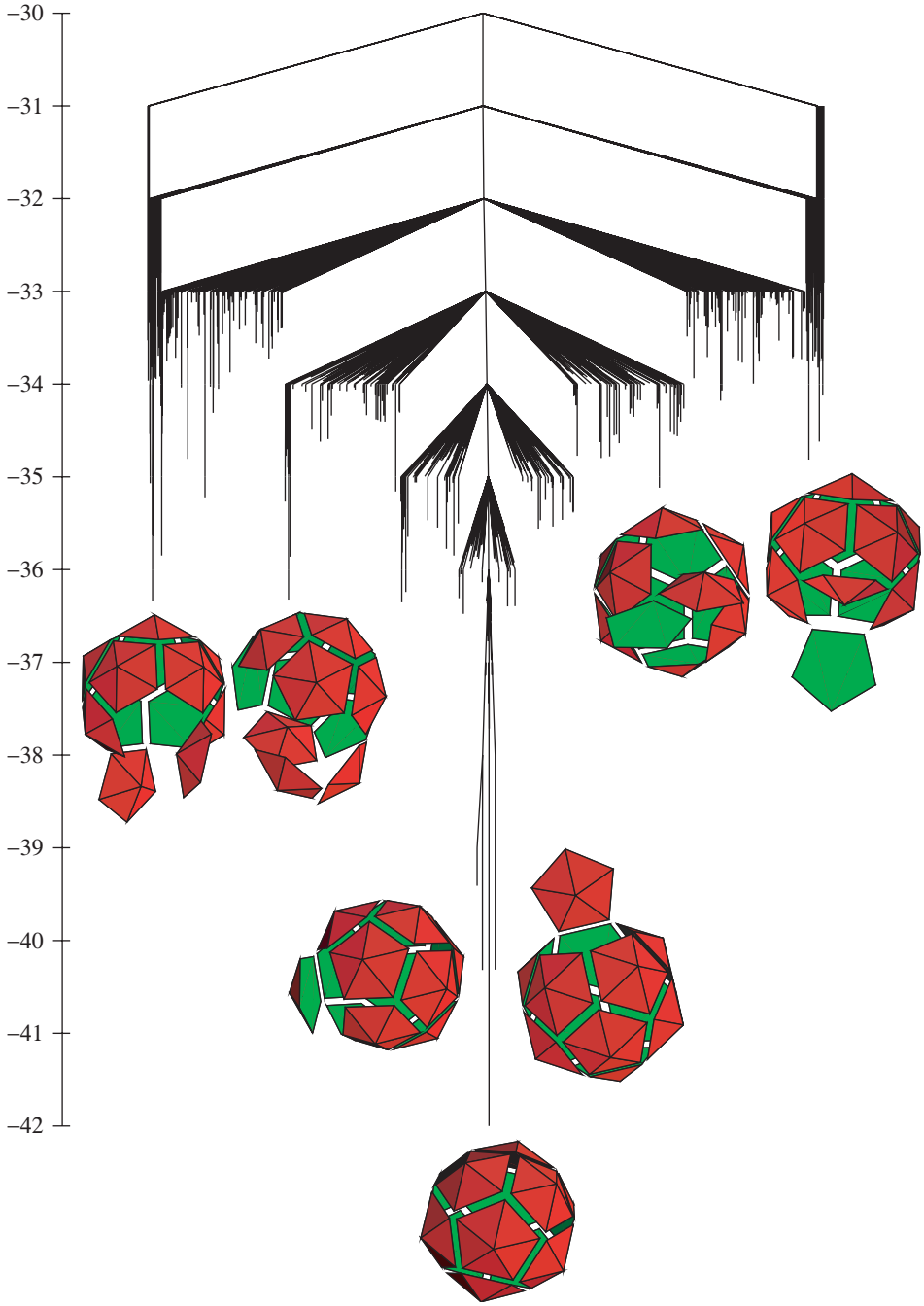

Figure 7. Disconnectivity graph for 12 pentagonal pyramids with $h = 0.5r$.

in figure 5, the repulsive interaction is minimized for $h = r/2$ when

$$\phi = 2\sin^{-1}\sqrt{\frac{2}{5-\sqrt{5}}} = 116.57^{\circ},$$

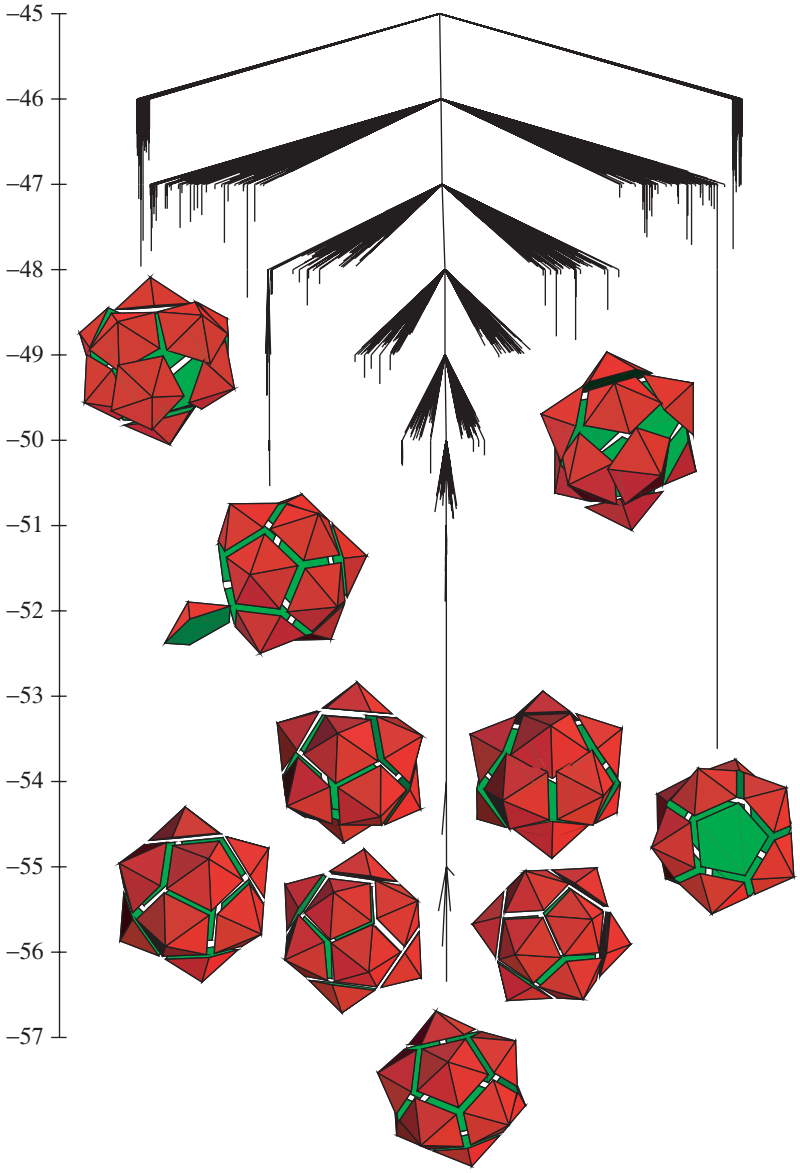

Figure 8. Disconnectivity graph for 12 pentagonal pyramids with $h = 0.5r$.

which corresponds to icosahedral packing. The separation of the repulsive sites is then

$$\frac{R}{R_{\mathrm{e}}} = 1 + r\sqrt{\frac{\sqrt{5}+5}{2}},$$

where r is measured in units of R_{e}, and the length parameter σ was fixed at this value.

Previous work suggests that closed shells with icosahedral symmetry are kinetically accessible (Wales 1987), and such structures have been found for discs or pentagons

constrained to move on the surface of a sphere (Marzec & Day 1993; Tarnai *et al.* 1995). Other simulations include models where subunits can shift between binding and non-binding conformations (Berger *et al.* 1994; Schwartz *et al.* 1998), as well as treatments involving a contact potential (Endres & Zlotnick 2002; Zlotnick *et al.* 1999), a lattice representation (Horton & Lewis 1992; Reddy *et al.* 1998), and discs with adhesive edges interacting via a harmonic angular potential (Bruinsma *et al.* 2003).

To construct disconnectivity graphs for the present model a general scheme for rigid bodies involving isotropic site–site potentials was implemented in both the GMIN and OPTIM programs. This scheme simply requires the distance derivatives of the site–site potentials along with the site positions in a reference orientation for each rigid body. All the formulae for the energy and derivatives can then be written in a general way, as described in the appendix.

Databases of stationary points were created using systematic transition state searches (Wales 2003; Wales *et al.* 2000) for a range of different height parameters. Results for three values of h are illustrated in figures 6–8; only the lowest 2000 local minima are shown in each case, although all the minima were included in each superbasin analysis. The graph for $h = r/2$, where the repulsive interaction between adjacent pentamers is minimized in the icosahedral global minimum, exhibits a well-defined 'palm tree' structure. We therefore expect efficient relaxation to the icosahedron over a wide range of temperature (Wales 2003). The icosahedron is also the global potential energy minimum for $h = 0.35r$ and $0.75r$, but in each case the PES possesses features that will hinder self-assembly. For $h = 0.35r$ (flatter pyramids) the potential energy gradient and driving force towards closed shell structures are diminished. There is also a side-funnel containing structures with more loosely packed pyramids, which might act as a kinetic trap. For smaller values of h the global potential energy minimum, located using basin-hopping (Li & Scheraga 1987, 1988; Wales & Doye 1997; Wales & Scheraga 1999), is non-compact. For $h = 0.75r$ an additional minimum appears on the two-dimensional cut through the six-dimensional dimer surface shown in figure 5. The resulting disconnectivity graph now contains low-lying local minima with perturbed icosahedral geometries where one or more pyramid bases overlap.

These results indicate that the icosahedral geometry will be thermodynamically stable and kinetically accessible for pentameric building blocks that are not too flat and not too spiky, which provides sufficient conditions for self-assembly. Of course, this is a very simplistic model of a virus capsid, but it does illustrate how self-organization can arise under relatively weak conditions.

5. Conclusions

It is the 'palm tree' topology of the underlying PES in figure 7 that permits efficient self-assembly, and it seems likely that this structure underpins many other non-random searches in molecular science. For example, 'magic numbers' in mass spectra, protein folding, and even crystallization, all involve the location of a few particular low potential energy structures from an exponentially large number of alternatives (Ball *et al.* 1996). The palm tree topology enables Levinthal's paradox to be circumvented in each case, by effectively guiding the system towards the target structure. We therefore identify this topology with nucleation and growth in a crystal,

with the principle of 'minimal frustration' (Bryngelson & Wolynes 1987, 1989; Bryngelson *et al.* 1995; Onuchic *et al.* 1997), and with a free energy surface that exhibits funnelling properties (Leopold *et al.* 1992). The potential energy funnel represented by the palm tree graph would produce a free energy surface involving a systematic decrease in the vicinity of the global potential energy minimum over a wide range of temperature. The corresponding order parameter could be any reasonable metric that measures distance from the global minimum. The palm tree topology is also consistent with a large value for the ratio $T_{\mathrm{f}}/T_{\mathrm{g}}$, where T_{f} is the 'folding' temperature, below which the potential energy and free energy global minima coincide, and T_{g} is the 'glass' temperature, where relaxation slows beyond some given time-scale (Socci & Onuchic 1994; Socci *et al.* 1996). If $T_{\mathrm{f}}/T_{\mathrm{g}}$ is large, then the surface possesses a global free energy minimum that is also kinetically accessible (Bryngelson & Wolynes 1987; Goldstein *et al.* 1992; Karplus & Šali 1995).

The main purpose of the simple model described in § 4 is to show how the 'palm tree' motif can provide a unifying basis for self-assembly, 'magic numbers', protein folding and crystallization. However, it is also interesting to note that Crick & Watson (1956) originally proposed that virus capsids might be composed of repeated subunits by considering how a limited amount of genetic information could be used most efficiently. The present analysis suggests that viruses may also exploit special properties of the energy landscape for subunits of the right form, in order to realize such structures.

I am very grateful to Professor Chris Dobson for his suggestion to investigate the energy landscape of a model virus capsid.

Appendix A. Optimization in rigid-body coordinates

The method adopted in the present work to provide a general framework for geometry optimization of rigid bodies interacting via isotropic site–site potentials exploits *Chasles's theorem* and the *rotation formula* (Goldstein 1980). Chasles's theorem states that the most general displacement of a rigid body is a translation plus a rotation. The rotation formula tells us that the position vector of a point $\boldsymbol{X}^0$ after rotation through an angle α about an axis defined by the unit vector $\hat{\boldsymbol{p}}$ is

$$\boldsymbol{X} = \boldsymbol{X}^0 \cos\alpha + \hat{\boldsymbol{p}}(\hat{\boldsymbol{p}} \cdot \boldsymbol{X}^0)(1 - \cos\alpha) + \boldsymbol{X}^0 \wedge \hat{\boldsymbol{p}} \sin\alpha, \tag{A 1}$$

where $\wedge$ denotes the vector product. The six degrees of freedom for a (nonlinear) rigid body can therefore be defined in terms of the centre-of-mass coordinates, X, Y and Z, and the three components of a vector, $\boldsymbol{p} = (n, l, m)$, which specify both a rotation axis through the centre of mass, and the magnitude of rotation, $\alpha = \sqrt{n^2 + l^2 + m^2}$. For a given set of rigid-body coordinates X, Y, Z, n, l and m, the site at $\boldsymbol{R}^0$ in the reference geometry is mapped to

$$\boldsymbol{R} = (X, Y, Z) + \boldsymbol{R}^0 \cos\alpha + \hat{\boldsymbol{p}}(\hat{\boldsymbol{p}} \cdot \boldsymbol{R}^0)(1 - \cos\alpha) + \boldsymbol{R}^0 \wedge \hat{\boldsymbol{p}} \sin\alpha. \tag{A 2}$$

If we denote the coordinates of two different rigid bodies using the subscripts I and J, and the sites within each rigid body by subscripts i and j, then for isotropic site–site potentials the total energy is

$$V = \sum_I \sum_{J<I} \sum_{i \in I} \sum_{j \in J} f_{ij}(R_{ij}), \tag{A 3}$$

where $R_{ij} = |\boldsymbol{R}_i - \boldsymbol{R}_j|$, and $f_{ij}(R)$ is the pair potential between sites i and j. R_{ij} is a function of the 12 rigid-body coordinates of molecules I and J, as well as the reference site positions $\boldsymbol{R}_i^0$ and $\boldsymbol{R}_j^0$. If ζ represents a rigid-body coordinate of molecule I, then the first derivative of the potential energy is

$$\frac{\partial V}{\partial \zeta} = \sum_{J \neq I} \sum_{i \in I} \sum_{j \in J} f'_{ij}(R_{ij}) \frac{\partial R_{ij}}{\partial \zeta}, \tag{A 4}$$

where $f'_{ij}(x) = \mathrm{d}f_{ij}(x)/\mathrm{d}x$. Second derivatives can also be derived using the chain rule in a straightforward way.

Using equation (A 2), the rigid-body derivatives of R_{ij} can all be programmed once and for all as functions of the reference site positions and the rigid-body coordinates. The MATHEMATICA program (Wolfram 1996) was used to construct these analytical derivatives. It is then only necessary to provide the position vectors of the sites in the reference geometry, $\boldsymbol{R}_i^0$, and the derivatives $f'_{ij}(x)$ and $f''_{ij}(x)$ to code both first and second derivatives of the potential energy. This procedure could also be generalized to deal with anisotropic site–site potentials.

References

Baldwin, R. L. 1994 Protein-folding: matching speed and stability. *Nature* **369**, 183–184.

Ball, K. D., Berry, R. S., Kunz, R. E., Li, F.-Y., Proykova, A. & Wales, D. J. 1996 From topography to dynamics on multidimensional potential energy surfaces of atomic clusters. *Science* **271**, 963–966.

Banerjee, A. & Adams, N. P. 1992 On coordinate transformations in steepest descent path and stationary point locations. *Int. J. Quant. Chem.* **43**, 855–871.

Becker, O. M. & Karplus, M. 1997 The topology of multidimensional potential energy surfaces: theory and application to peptide structure and kinetics. *J. Chem. Phys.* **106**, 1495–1517.

Berger, B., Shor, P. W., Tucker-Kellogg, L. & King, J. 1994 Local rule-based theory of virus shell assembly. *Proc. Natl Acad. Sci. USA* **91**, 7732–7736.

Berry, R. S. & Breitengraser-Kunz, R. 1995 Topography and dynamics of multidimensional interatomic potential surfaces. *Phys. Rev. Lett.* **74**, 3951–3954.

Braier, P. A., Berry, R. S. & Wales, D. 1990 How the range of pair interactions governs features of multidimensional potentials. *J. Chem. Phys.* **93**, 8745–8756.

Bruinsma, R. F., Gelbart, W. M., Reguera, D., Rudnick, J. & Zandi, R. 2003 Viral self-assembly as a thermodynamic process. *Phys. Rev. Lett.* **90**, 248101.

Bryngelson, J. D. & Wolynes, P. G. 1987 Spin-glasses and the statistical-mechanics of protein folding. *Proc. Natl Acad. Sci. USA* **84**, 7524–6915.

Bryngelson, J. D. & Wolynes, P. G. 1989 Intermediates and barrier crossing in a random energy-model (with applications to protein folding). *J. Phys. Chem.* **93**, 6902–6915.

Bryngelson, J. D. & Wolynes, P. G. 1990 A simple statistical field-theory of heteropolymer collapse with application to protein folding. *Biopolymers* **30**, 177–188.

Bryngelson, J. D., Onuchic, J. N., Socci, N. C. & Wolynes, P. G. 1995 Tunnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167–195.

Calvo, F., Neirotti, J. P., Freeman, D. L. & Doll, J. D. 2000 Phase changes in 38-atom Lennard-Jones clusters. II. A parallel tempering study of equilibrium and dynamic properties in the molecular dynamics and microcanonical ensembles. *J. Chem. Phys.* **112**, 10350–10357.

Calvo, F., Doye, J. P. K. & Wales, D. J. 2001*a* Characterization of anharmonicities on complex potential energy surfaces: perturbation theory and simulation. *J. Chem. Phys.* **115**, 9627–9636.

Calvo, F., Doye, J. P. K. & Wales, D. J. 2001*b* Quantum partition functions from classical distributions. Application to rare gas clusters. *J. Chem. Phys.* **114**, 7312–7329.

Camacho, C. J. & Thirumalai, D. 1993 Kinetics and thermodynamics of folding in model proteins. *Proc. Natl Acad. Sci. USA* **90**, 6369–6372.

Caspar, D. L. D. & Klug, A. 1962 Physical principles in the construction of regular viruses. *Cold Spring Harbour, Symp. Quant. Biol.* **27**, 1–24.

Cerjan, C. J. & Miller, W. H. 1981 On finding transition states. *J. Chem. Phys.* **75**, 2800–2806.

Clary, D. C. 2001 Torsional diffusion Monte Carlo: a method for quantum simulations of proteins. *J. Chem. Phys.* **114**, 9725–9732.

Crick, F. & Watson, J. D. 1956 Structure of small viruses. *Nature* **177**, 473–475.

Czerminski, R. & Elber, R. 1990 Reaction path study of conformational transitions in flexible systems—applications to peptides. *J. Chem. Phys.* **92**, 5580–5601.

Davis, B. D., Dulbecco, R., Ersen, H. N. & Ginsberg, H. S. 1980 *Microbiology*, 3rd edn. Harper and Row.

Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Lee, D. P., Thomas, P. D. & Chan, H. S. 1995 Principles of protein folding: a perspective from simple exact models. *Prot. Sci.* **4**, 561–602.

Dobson, C. M., Sali, A. & Karplus, M. 1998 Protein folding: a perspective from theory and experiment. *Angew. Chem. Int. Ed.* **37**, 868–893.

Doye, J. P. K. & Calvo, F. 2002 Entropic effects on the structure of Lennard-Jones clusters. *J. Chem. Phys.* **116**, 8307–8317.

Doye, J. P. K. & Wales, D. J. 1995 Calculation of thermodynamic properties of small Lennard-Jones clusters incorporating anharmonicity. *J. Chem. Phys.* **102**, 9659–9672.

Doye, J. P. K. & Wales, D. J. 1996 On potential energy surfaces and relaxation to the global minimum. *J. Chem. Phys.* **105**, 8428–8445.

Doye, J. P. K. & Wales, D. J. 1998 Thermodynamics of global optimization. *Phys. Rev. Lett.* **80**, 1357–1360.

Doye, J. P. K. & Wales, D. J. 1999*a* The dynamics of structural transitions in sodium chloride clusters. *J. Chem. Phys.* **111**, 11070–11079.

Doye, J. P. K. & Wales, D. J. 1999*b* Structural transitions and global minima of sodium chloride clusters. *Phys. Rev.* B **59**, 2292–2300.

Doye, J. P. K. & Wales, D. J. 2002 Saddle points and dynamics of Lennard-Jones clusters, solids, and supercooled liquids. *J. Chem. Phys.* **116**, 3777–3788.

Doye, J. P. K., Wales, D. J. & Miller, M. A. 1998 Thermodynamics and the global optimization of Lennard-Jones clusters. *J. Chem. Phys.* **109**, 8143–8153.

Doye, J. P. K., Miller, M. A. & Wales, D. J. 1999*a* The double-funnel energy landscape of the 38-atom Lennard-Jones cluster. *J. Chem. Phys.* **110**, 6896–6906.

Ehrlich, L. S., Liu, T. B., Scarlata, S., Chu, B. & Carter, C. A. 2001 IV-1 capsid protein forms spherical (immature-like) and tubular (mature-like) particles *in vitro*: structure switching by pH-induced conformational changes. *Biophys. J.* **81**, 586–594.

Eiserling, F. A. & Dickson, R. C. 1972 Assembly of viruses. *A. Rev. Biochem.* **41**, 467–502.

Endres, D. & Zlotnick, A. 2002 Model based analysis of assembly kinetics for virus capsids or other spherical polymers. *Biophys. J.* **83**, 1217–1230.

Ercolani, G. 2003 A model for self-assembly in solution. *J. Phys. Chem.* B **107**, 5052–5057.

Evans, M. G. & Polanyi, M. 1935 Some applications of the transition state method to the calculation of reaction velocities, especially in solution. *Trans. Faraday Soc.* **31**, 875–894.

Evans, D. A. & Wales, D. J. 2003 Free energy landscapes of model peptides and proteins. *J. Chem. Phys.* **118**, 3891–3897.

Eyring, H. 1935 The activated complex and the absolute rate of chemical reactions. *Chem. Rev.* **17**, 65–77.

Fraenkel-Conrat, H. & Williams, R. C. 1955 Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components. *Proc. Natl Acad. Sci. USA* **41**, 690–698.

Frantz, D. D. 2001 Magic number behavior for heat capacities of medium-sized classical Lennard-Jones clusters. *J. Chem. Phys.* **115**, 6136–6157.

Ganser, B. K., Li, S., Klishko, J. Y., Finch, J. T. & Sundquist, W. I. 1999 Assembly and analysis of conical models for the HIV-1 core. *Science* **283**, 80–83.

Gō, N. 1983 Theoretical-studies of protein folding. *A. Rev. Biophys. Bioengng* **12**, 183–210.

Goldstein, H. 1980 *Classical mechanics*. Reading, MA: Addison-Wesley.

Goldstein, R. A., Luthey-Schulten, Z. & Wolynes, P. G. 1992 Optimal protein folding codes from spin-glass theory. *Proc. Natl Acad. Sci. USA* **89**, 4918–4922.

Honeycutt, J. D. & Thirumalai, D. 1990 Metastability of the folded states of globular proteins. *Proc. Natl Acad. Sci. USA* **87**, 3526–3529.

Honeycutt, J. D. & Thirumalai, D. 1992 The nature of folded states of globular proteins. *Biopolymers* **32**, 695–709.

Horton, N. & Lewis, M. 1992 Calculation of the free energy of association for protein complexes. *Protein Sci.* **1**, 169–181.

Jones, J. E. & Ingham, A. E. 1925 On the calculation of certain crystal potential constants, and on the cubic crystal of least potential energy. *Proc. R. Soc. A* **107**, 636–653.

Karplus, M. & Šali, A. 1995 Theoretical studies of protein-folding and unfolding. *Curr. Opin. Struct. Biol.* **5**, 58–73.

Kiselev, N. A. & Klug, A. 1969 The structure of viruses of the Papilloma–Polyoma type V. Tubular variants built of pentamers. *J. Mol. Biol.* **40**, 155–171.

Krivov, S. V. & Karplus, M. 2002 Free energy disconnectivity graphs: application to peptide models. *J. Chem. Phys.* **117**, 10 894–10 903.

Kumeda, Y. & Wales, D. J. 2003 *Ab initio* study of rearrangements between C_{60} fullerenes. *Chem. Phys. Lett.* **374**, 125–131.

Kumeda, Y., Munro, L. J. & Wales, D. J. 2001 Transition states and rearrangement mechanisms from hybrid eigenvector-following and density functional theory. Application to $C_{10}H_{10}$ and defect migration in crystalline silicon. *Chem. Phys. Lett.* **341**, 185–194.

Kunz, R. E. & Berry, R. S. 1995 Statistical interpretation of topographies and dynamics of multidimensional potentials. *J. Chem. Phys.* **103**, 1904–1912.

Lazaridis, T. & Karplus, M. 1997 'New view' of protein folding reconciled with the old through multiple unfolding simulations. *Science* **278**, 1928–1931.

Leopold, P. E., Montal, M. & Onuchic, J. N. 1992 Protein folding funnels: a kinetic approach to the sequence structure relationship. *Proc. Natl Acad. Sci. USA* **89**, 8721–8725.

Levinthal, C. 1969 How to fold graciously. In *Mössbauer Spectroscopy In Biological Systems, Proc. Meeting held at Allerton House, Monticello, IL* (ed. P. DeBrunner, J. Tsibris & E. Munck), p. 22. Urbana: University of Illinois Press.

Li, Z. & Scheraga, H. A. 1987 Monte-Carlo-minimization approach to the multiple-minima problem in protein folding. *Proc. Natl Acad. Sci. USA* **84**, 6611–6615.

Li, Z. & Scheraga, H. A. 1988 Structure and free energy of complex thermodynamic systems. *J. Mol. Struct. (Theochem)* **179**, 333–352.

Liu, D. & Nocedal, J. 1989 On the limited memory BFGS method for large scale optimization. *Math. Program.* **45**, 503–528.

Marzec, C. J. & Day, L. A. 1993 Pattern formation in icosahedral virus capsids: the papova viruses and nudaurelia capensis beta-virus. *Biophys. J.* **65**, 2559–2577.

Miller, T. F. & Clary, D. C. 2003 Torsional path integral Monte Carlo method for calculating the absolute quantum free energy of large molecules. *J. Chem. Phys.* **119**, 68–76.

Miller, M. A. & Wales, D. J. 1999 Energy landscape of a model protein. *J. Chem. Phys.* **111**, 6610–6616.

Miller, M. A., Doye, J. P. K. & Wales, D. J. 1999 Structural relaxation in atomic clusters: master equation dynamics. *Phys. Rev.* E **60**, 3701–3718.

Munro, L. J. & Wales, D. J. 1999 Defect migration in crystalline silicon. *Phys. Rev.* B **59**, 3969–3980.

Murrell, J. N. & Laidler, K. J. 1968 Symmetries of activated complexes. *Trans. Faraday. Soc.* **64**, 371–377.

Neirotti, J. P., Calvo, F., Freeman, D. L. & Doll, J. D. 2000 Phase changes in 38-atom Lennard-Jones clusters. I. A parallel tempering study in the canonical ensemble. *J. Chem. Phys.* **112**, 10340–10349.

Nocedal, J. 1980 Updating quasi-Newton matrices with limited storage. *Mathematics of Computation* **35**, 773–782.

Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z. & Socci, N. D. 1995 Toward an outline of the topography of a realistic protein-folding funnel. *Proc. Natl Acad. Sci. USA* **92**, 3626–3630.

Onuchic, J. N., Luthey-Schulten, Z. & Wolynes, P. G. 1997 Theory of protein folding: the energy landscape perspective. *A. Rev. Phys. Chem.* **48**, 545–600.

Onuchic, J. N., Nymeyer, H., Garcia, A. E., Chahine, J. & Socci, N. D. 2000 The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. *Adv. Prot. Chem.* **53**, 87–152.

Pancíř, J. 1975 Calculation of the least energy path on the energy hypersurface. *Collect. Czech. Chem. Commun.* **40**, 1112–1118.

Pelzer, H. & Wigner, E. 1932 Über die Geschwindigkeitskonstante von Austauschreaktionen. *Z. Phys. Chem.* **B15**, 445–471.

Pugliano, N. & Saykally, R. J. 1992 Measurement of quantum tunneling between chiral isomers of the cyclic water trimer. *Science* **257**, 1937–1940.

Reddy, V. S., Giesing, H. A., Morton, R. T., Kumar, A., Port, C. B., Brooks, C. L. & Johnson, J. E. 1998 Energetics of quasiequivalence: computational analysis of protein-protein interactions in icosahedral viruses' *Biophys. J.* **74**, 546–558.

Saykally, R. J. & Blake, G. A. 1993 Molecular interactions and hydrogen bond tunneling dynamics: some new perspectives. *Science* **259**, 1570–1575.

Schlegel, H. B. 2003 Exploring potential energy surfaces for chemical reactions: an overview of some practical methods. *J. Computat. Chem.* **24**, 1514–1527.

Schwartz, R., Shor, P. W., Prevelige, P. E. & Berger, B. 1998 Local rules simulation of the kinetics of virus capsid self-assembly. *Biophys. J.* **75**, 2626–2636.

Socci, N. D. & Onuchic, J. N. 1994 Folding kinetics of proteinlike heteropolymers. *J. Chem. Phys.* **101**, 1519–1528.

Socci, N. D., Onuchic, J. N. & Wolynes, P. G. 1996 Diffusive dynamics of the reaction coordinate for protein folding funnels. *J. Chem. Phys.* **104**, 5860–5868.

Stillinger, F. H. & Weber, T. A. 1984 Packing structures and transitions in liquids and solids. *Science* **225**, 983–989.

Tarnai, T., Gaspar, Z. & Szalai, L. 1995 Pentagon packing models for all-pentamer virus structures. *Biophy. J.* **69**, 612–618.

Thomas, J. M. 2004 Ultrafast electron crystallography: the dawn of a new era. *Angew. Chem. Int. Ed.* **43**, 2606–2610.

Trygubenko, S. A. & Wales, D. J. 2004 A doubly nudged elastic band method for finding transition states. *J. Chem. Phys.* **120**, 2082–2094.

Vigliotti, F., Chen, S. Y., Ruan, C. Y., Lobastov, V. A. & Zewail, A. H. 2004 Ultrafast electron crystallography of surface structural dynamics with atomic-scale resolution. *Angew. Chem. Int. Ed.* **43**, 2705–2709.

Wales, D. J. 1987 Closed shell structures and the building game. *Chem. Phys. Lett.* **141**, 478–484.

Wales, D. J. 1993 Coexistence in small inert gas clusters. *Mol. Phys.* **78**, 151–171.

Wales, D. J. 1994 Rearrangements of 55-atom Lennard-Jones and $(C_{60})_{55}$ clusters. *J. Chem. Phys.* **101**, 3750–3762.

Wales, D. J. 2000 Potential energy surfaces and coordinate dependence. *J. Chem. Phys.* **113**, 3926–3927.

Wales, D. J. 2002 Discrete path sampling. *Mol. Phys.* **100**, 3285–3305.

Wales, D. J. 2003 *Energy landscapes*. Cambridge University Press.

Wales, D. J. & Doye, J. P. K. 1997 Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J. Phys. Chem.* A **101**, 5111–5116.

Wales, D. J. & Doye, J. P. K. 2003 Stationary points and dynamics in high-dimensional systems. *J. Chem. Phys.* **119**, 12409–12416.

Wales, D. J. & Hodges, M. P. 1998 Global minima of water clusters $(H_2O)_n$, $n \leqslant 21$, described by an empirical potential. *Chem. Phys. Lett.* **286**, 65–72.

Wales, D. J. & Scheraga, H. A. 1999 Global optimization of clusters, crystals and biomolecules. *Science* **285**, 1368–1372.

Wales, D. J., Miller, M. A. & Walsh, T. R. 1998 Archetypal energy landscapes. *Nature* **394**, 758–760.

Wales, D. J., Doye, J. P. K., Miller, M. A., Mortenson, P. N. & Walsh, T. R. 2000 Energy landscapes: from clusters to biomolecules. *Adv. Chem. Phys.* **115**, 1–111.

Wales, D. J., Doye, J. P. K., Dullweber, A., Hodges, M. P., Naumkin, F. Y., Calvo, F., Hernández-Rojas, J. & Middleton, T. F. 2004 The Cambridge cluster database. (Available at http://www-wales.ch.cam.ac.uk/CCD.html.)

Wolfram, S. 1996 *The Mathematica book*, 3rd edn. Wolfram Media/Cambridge University Press.

Zlotnick, A., Johnson, J. M., Wingfield, P. W., Stahl, S. J. & Endres, D. 1999 A theoretical model successfully identifies features of hepatitis B virus capsid. *Biochem.* **38**, 14644–14652.

Discussion

D. C. Clary (*Department of Physical and Theoretical Chemistry, University of Oxford, UK*). Could you summarize the advantages and disadvantages of your approach compared with molecular dynamics (MD) calculations.

D. J. Wales. The main advantage of the discrete path sampling (DPS) approach is that it can provide rate constants for processes that lie beyond the reach of conventional simulations. Discrete paths consisting of local minima and the transition states that connect them can be calculated rapidly, and trapping is not a problem, because transition states for processes corresponding to high barriers are generally no harder to characterize than those with lower barriers. Using appropriate partition functions for the local minima, and statistical rate theory for dynamics, means that the effects of high-frequency vibrational motion, which limit the time-step for MD simulations, are treated implicitly. This coarse-grained approach to the energy landscape in terms of stationary points also provides useful insight into thermodynamic and dynamic properties, since vibrational 'noise' is removed, and the observed values are resolved into contributions from distinct local minima and pathways. The main source of error in the simplest DPS rate calculations stems from the harmonic approximation used for the densities of states. However, once a database of stationary points has been obtained, all the partition functions and individual minimum-to-minimum rate constants can be recalculated using more accurate methods, as required.

J. C. Schön (*Max Planck Institute for Solid State Research, Stuttgart, Germany*). It is very interesting that you have observed for the landscape of the Lennard-Jones

crystal that the major side minima correspond to defects in a basically close-packed atom arrangement. When studying the energy landscape of the three- and two-dimensional noble gases employing a Lennard-Jones-type potential (Schön & Jonsen 1994; Schön *et al*. 1996), we have found the same phenomenon: the structures were always close-packed arrangements of spheres (both HCP and FCC variants), with the neighbour minima corresponding to close packings of spheres that contained one or more holes.

When studying the equilibration among these minima using Markov matrices on multiple-lump tree-graphs of the two-dimensional system (Schön *et al*. 1996), we found that the density of states within the basins can have a certain influence on the order of equilibration among the basins.

Have you observed any influence of the local density of states (beyond the harmonic approximation) on the transition rates as a function of temperature?

D. J. Wales. Previous investigations of defect migration in FCC solids using empirical potentials indicate that $\langle 100\rangle$-split-interstitials are favourable for systems bound by longer-range potentials, and O_h-interstitials are unfavourable, while the converse holds for systems bound by short-range potentials (Wales & Uppenbrink 1994). Equilibration amongst the various local minima is determined by the densities of states of both transition states and local minima in all our calculations. Anharmonicity can certainly have a significant effect (Calvo *et al.* 2001*a*).

J. C. Schön. We have used the threshold algorithm (Schön *et al*. 1996), in order to measure the transition probabilities among the basins in the system MgF_2 (Weavers *et al*. 1999). In particular, we observed that strong entropic barriers can be present that are felt above the lowest saddle points connecting two basins, even sometimes stabilizing regions on the landscape that do not contain local minima, but are shielded from the rest of the landscape by the 'difficulty of finding an exit path' (Schön *et al*. 2003).

What is your experience regarding the contribution of complicated paths to the total transition probability? What about competition between 'simple' paths at high energy and complicated paths at low energy?

D. J. Wales. The coarse-graining employed in our calculations is based upon a rigorous division of the potential energy surface into the basins of attraction of local minima, while minimum-to-minimum transitions are described in terms of transition states that lie on basin boundaries. We have not considered further subdivision of the configuration space within the basin of attraction of a single minimum. However, entropy effects should be accounted for within this picture. The complexity of the underlying mechanism could be considered in two ways, namely the number of discrete paths required to obtain a good estimate of the rate constant, and the average length of these paths. The relative contributions of different discrete paths changes as a function of temperature. At higher temperatures we generally find that shorter discrete paths involving higher potential energy make an increasing contribution. In contrast, at low temperature, where minimizing the potential energy barriers is more important, the dominant contributions come from longer paths with lower barriers. For processes involving at least one extended phase we will generally have to sample over an exponentially large number of possible paths.

P. G. Wolynes (*Department of Chemistry and Biochemistry, Department of Physics and Center for Theoretical Biological Physics, University of California, San*

Diego, USA). For extended systems we usually see the two limits of your trees: funnels for first order transitions, banyans for glass transitions. The fullerene has high barriers even for small changes, this suggests a very non-additive force. Is this connected with the aromaticity of fullerenes? What is the landscape for positive and dipositive ions?'

D. J. Wales. The covalent bonding in fullerenes is indeed non-additive, and rather anisotropic, with aromatic character deriving from the delocalized π system. The high barriers, relative to the energy difference between connected minima, result from disruption of both the σ and π systems, which can be separated to a first approximation in the roughly spherical structures that correspond to local fullerene minima (Stone & Wales 1986). The significant delocalization energy of the π system, which is partly lost in the transition state, certainly accounts for part of the observed barriers.

To the best of my knowledge, the landscape has not yet been analysed for C_{60}^{+} or C_{60}^{2+}. I would expect the lowest energy fullerene for these ions to be based upon icosahedral buckminsterfullerene, but with Jahn–Teller distortions. If the positive charge were localized on the same atoms that move to facilitate fullerene rearrangements, then the barriers may be reduced, since these atoms are undercoordinated in the transition state. This effect would change the willow tree pattern observed for neutral C_{60} towards the palm tree form, which would result in more efficient relaxation to the global minimum. However, it is also possible that the mechanism itself might change.

Additional references

Schön, J. C. & Jonsen, M. 1994 *Ber. Bunsen Ges. Phys. Chem.* **98**, 1541.
Schön, J. C., Putz, H. & Jonsen, M. 1996 *J. Phys. Condens. Matter* **8**, 143.
Schön, J. C., Wevers, M. A. C. & Jonsen, M. 2003 *J. Phys. Condens. Matter* **15**, 5479.
Stone, A. J. & Wales. D. J. 1986 *Chem. Phys. Lett.* **128**, 501–503.
Wales, D. J. & Uppenbrink, J. 1994 *Phys. Rev.* B **50**, 12 342–12 361.
Weavers, M. A. C. *et al*. 1999 *J. Phys. Condens. Matter* **11**, 6487.