

# Energy landscapes, self-assembly and viruses

D. J. WALES\*

Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, UK

Phenomena such as protein folding, crystallisation, self-assembly, and the observation of magic number clusters in molecular beams are all the result of non-random searches. Analysis of the underlying potential energy surface may provide a unifying framework to explain how such events occur as the result of a guided exploration of the landscape. In particular, icosahedral shells composed of 12 pentagonal pyramids are found to be thermodynamically favourable and kinetically accessible when the pyramids are not too spiky and not too flat. Hence, viruses with icosahedral capsids not only minimise the genetic material required to encode the repeated subunits, but may also utilise the favourable properties of a potential energy surface that effectively directs self-assembly.

*Keywords:* Energy landscapes; Self-assembly; Virus capsid

*Msc:* 82-06; 82D99; 92B99

## 1. Introduction

The structure, dynamics and thermodynamics of any molecular system are determined at a fundamental level by the interatomic or intermolecular forces. The interplay of these forces governs the form of the potential energy surface (PES), including the geometries of the local minima, their distribution in energy and the magnitude of the barriers that separate them. The study of energy landscapes is largely concerned with how such characteristics of the PES determine observable properties [1]. For example, the observation that “structure-seekers”, such as magic number clusters, proteins and crystals, can reliably locate particular low energy states, whilst glasses avoid crystallisation, indicates a qualitative difference in how the corresponding landscapes are organised [1–5, 29]. The spontaneous self-assembly of complex structures from smaller building blocks may be governed by similar principles, and many such processes are known, both in nature and in the laboratory. In particular, reversible self-assembly has been demonstrated in the laboratory for a number of different virus capsids, and the thermodynamics of such processes is an active area of research [6–8]. However, it is important to realise that self-assembly requires both a thermodynamic driving force and a kinetically accessible target [1].

A simple intermolecular potential has recently been developed to describe the effective interaction between pentameric building blocks of a virus capsid [9].

The underlying PES was found to possess all the characteristics necessary for efficient relaxation to an icosahedral arrangement, providing the constituent pentagonal pyramids exhibit an appropriate morphology. This analysis suggests that self-assembly in such systems may be achieved in the same way as relaxation in other “structure-seekers”, and that such processes are fundamentally connected by the topology of the PES. For capsid assembly some specific predictions can be made concerning the shape of the capsomers. Characterisation of the PES may therefore provide an opportunity for theory to guide the design (or disruption) of self-assembling materials.

## 2. Model and results

Experiments have identified two levels of organisation within most virus capsids, where the structure can be viewed as an assembly of capsomers, and each capsomer is composed of five or six protein subunits. Both the subunit–subunit and capsomer–capsomer bonds are noncovalent, and there is some experimental evidence that the capsomer–capsomer bonds are weaker [10,11]. Satellite tobacco necrosis virus provides an example of the simplest capsid structure, involving just 12 pentameric units and 60 protein subunits in total. To model this system at a coarse-grained level, an effective pentamer–pentamer potential was developed to establish the minimal

\*Tel.: +44-01223-336354. Fax: +44-01223-336362. Email: dw34@cam.ac.uk

characteristics required for efficient relaxation of 12 such units to an icosahedral global minimum [9]. Each pentamer was represented by a rigid pentagonal pyramid of height  $h$  and radius  $r$ . To provide a compact but flexible representation of the capsomer shape a pairwise additive potential was defined in terms of six sites corresponding to the vertices of each pyramid.

The resulting model differs from most previous treatments [12] through its formulation as an intermolecular potential (or potential of mean force), including a steep repulsive wall at short separations, an attractive well, and a dissociation limit [9]. In particular, the formation of a closed shell must be encoded implicitly within the site–site form, since no constraints are applied to the geometry beyond the rigid-body approximation for the pentamers. This treatment therefore enables the global PES, including open shells, to be examined in terms of a small number of essential parameters.

The height of the repulsive site above the base of the pyramid,  $h$ , turns out to be the most important factor in determining the overall form of the PES. Without this site the pyramids simply aggregate with their pentagonal faces together as far as possible, thereby maximising the number of attractive interactions. Non-zero values of  $h$  endow the pentamers with a three-dimensional shape; if  $h$  is too small then the potential energy gradient, and hence the driving force towards a closed structure, is diminished.

Geometrical considerations suggest that closed shells with icosahedral symmetry are kinetically accessible [13], and if high symmetry is imposed then such structures are indeed obtained from models involving disks or pentagons constrained to the surface of a sphere [14,15]. Structures based upon icosahedral packing in three dimensions also provide a common structural motif for atomic and molecular clusters, since they optimise the number of nearest-neighbour contacts while minimising the surface energy [1]. A more detailed simulation approach for virus capsids involves subunits that shift between binding and non-binding conformations and interact according to a permitted set of patterns for a given target structure if they pass sufficiently close together in an appropriate orientation [16,17]. Possible assembly intermediates have also been predicted using association energies for subunit proteins and restricting the structure to an icosahedral surface lattice [18,19]. Studies that focus upon aggregation have considered the formation of a dodecahedron from 12 pentagonal subunits using a contact potential to define the free energy [20,21]. Bruinsma *et al.* [6] have previously treated capsomers as disks with adhesive edges, together with a harmonic term to introduce curvature via a hinge angle. For the model based upon pentagonal pyramids it is the height of the disks,  $h$ , which primarily determines the optimal hinge angle for a pair of pentamers, and the curvature is close to the optimal value for icosahedral packing when  $h = r/2$  [9].

The model based upon rigid pyramids has been analysed to examine the conditions under which an

icosahedral global minimum is both thermodynamically favoured and kinetically accessible. The global PES was characterised by calculating databases of local minima and the transition states that connect them, followed by the construction of disconnectivity graphs [1,6,22]. Knowing the potential energy of the pathway with the lowest barrier between any two minima enables us to partition these minima into disjoint sets at a regularly spaced series of total energies,  $E$ . Two minima are in the same set if they are connected by a pathway whose highest transition state lies below  $E$ . For minima in different sets no such path exists. As the threshold energy increases, new superbins appear as higher energy minima become accessible. However, superbins that are distinct at lower energy merge together when  $E$  lies above the highest transition state on the lowest pathway between them. Hence the bottom of every branch in the graph corresponds to the potential energy of a local minimum, and these branches join together when the corresponding minima lie in the same superbins. The spacing on the horizontal axis is simply chosen to clarify the structure of the graph.

The stationary points used to construct the disconnectivity graphs are local minima and transition states of the potential energy. These points, and the graph itself, are independent of temperature, mass and coordinate system [1]. Thermodynamic properties, such as free energies, as well as relaxation dynamics, can also be calculated using these stationary points and their properties. In particular, a local density of states is needed for each structure to calculate the total partition function and to apply statistical rate theory for the rate constants corresponding to each rearrangement [1,29]. In fact, such detailed calculations are not necessary for the graph shown in figure 1, since it has a standard form [1,23]. The figure reveals a well-defined icosahedral global minimum with a steep potential energy gradient, and there are no low energy alternative morphologies, which would appear as deep side-funnels and act as kinetic traps. The corresponding graph corresponds to the “palm tree” form identified in previous work [23]. This pattern is characteristic of systems for which efficient relaxation to the global minimum occurs over a wide range of temperature, such as magic number clusters observed in molecular beam studies. In this temperature range the global potential energy minimum is also the global free energy minimum, and the system has sufficient thermal energy to overcome the potential energy barriers involved in relaxation from higher energy minima. The “palm tree” graph therefore corresponds to the “funnelling” associated with minimal frustration in previous work [5]. The global potential energy minimum lies sufficiently below the entropically favoured ensemble of defective structures to remain the free energy minimum at temperatures where the downhill barriers are comparable to  $kT$ . Dissociation of one or more pentamers was not considered in this study, so the present analysis begins once 12 such capsomers become associated.

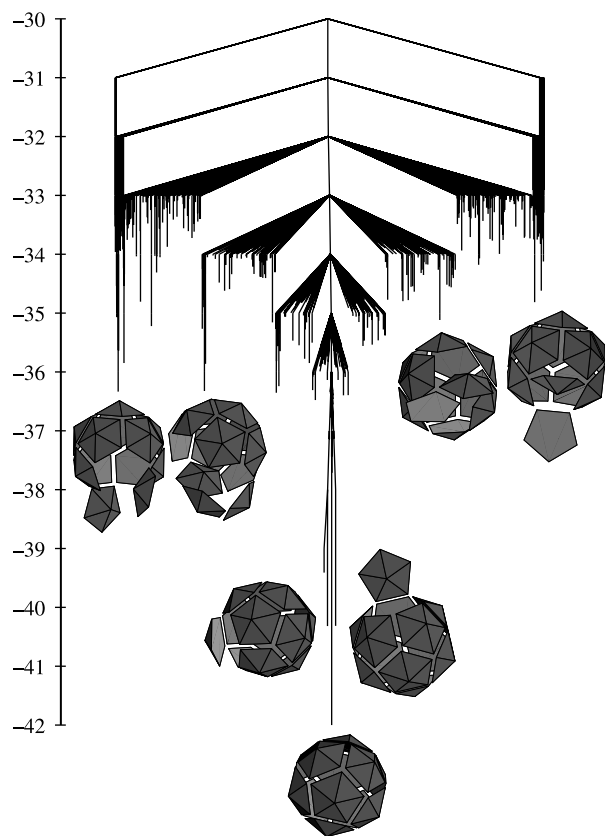


Figure 1. Disconnectivity graph for 12 pentagonal pyramids with  $h = 0.5 \times r$  [9]. To simplify the graph, only the lowest 2000 local minima are shown, although all the minima were included in each superbasin analysis.

The graphs for pentagonal pyramids that are too flat or too spiky were found to possess additional features that would hinder relaxation. Flatter pyramids produce a smaller potential energy gradient towards the icosahedron, and relatively deep side-funnels were observed corresponding to shells where the pyramids are packed more loosely [9]. The global minimum also becomes non-compact when  $h$  falls below a certain threshold. At the other extreme, when  $h$  is too large, local minima begin to appear with overlapping pyramids or pyramids pointing into the shell. The corresponding landscapes also contain kinetic traps. Nevertheless, there exists a relatively wide range for  $h$  where the landscape possesses favourable properties for self-assembly.

### 3. Discussion

These results for the coarse-grained model based upon rigid pentagonal pyramids indicate that an icosahedral shell will self-assemble from such building blocks, so long as they are not too flat and not too spiky [9]. This shape represents a first approximation to the pentameric capsomers found in picornaviruses, where five copies of the protein VP1 bind around the five-fold axis to form a protrusion on the surface. A hypothetical pentameric

assembly intermediate has also been suggested for the  $\phi 6$  double-stranded RNA phage [24]. Incomplete shells with one pentagonal pyramid remaining to swing into place are the lowest-lying local minima above the icosahedron for  $h = 0.35r$  and  $r/2$  (figure 1). However, two minima are observed where the pyramid in question points either into or out of the shell. There is some evidence that scaffolding proteins in bacteriophage  $\Phi X-174$  help to ensure a correct two-fold contact between pentameric units [25]. Mechanisms such as this, and perhaps conformation changes in subunit proteins during assembly [24], could improve efficiency by removing kinetic traps from the PES, thus producing a more ideal potential energy funnel.

The observation of magic numbers in mass spectra, the folding of proteins to their native states, and even crystallisation [3], depend upon non-random searches of configuration space to avoid Levinthal's famous paradox [26]. The organisation of the underlying PES ultimately governs whether efficient relaxation can occur to a particular structure. The "palm tree" motif in figure 1 effectively guides such systems towards the global minimum, and provides a graphical representation of a landscape with "funnelling" properties [2] and minimal frustration [5,27], Onuchic *et al.* 1997). Similarly, the PES for an icosahedral shell of pentagonal pyramids possesses the characteristics required for self-organisation when the shape of the interaction potential satisfies relatively weak conditions. Hence the construction of icosahedral virus capsids from a set of identical building blocks not only optimises the use of genetic information [28], but may also exploit the special properties of the energy landscape for such structures.

### References

- [1] Wales, D.J., 2003, *Energy Landscapes* (Cambridge: Cambridge University Press).
- [2] Leopold, P.E., Montal, M. and Onuchic, J.N., 1992, *Proc. Natl Acad. Sci. USA*, **89**, 8721.
- [3] Ball, K.D., Berry, R.S., Kunz, R.E., Li, F.-Y., Proykova, A. and Wales, D.J., 1996, *Science*, **271**, 963.
- [4] Doye, J.P.K. and Wales, D.J., 1996, *J. Chem. Phys.*, **105**, 8428.
- [5] Onuchic, J.N., Luthey-Schulten, Z. and Wolynes, P.G., 1997, *Annu. Rev. Phys. Chem.*, **48**, 545.
- [6] Bruinsma, R.F., Gelbart, W.M., Reguera, D., Rudnick, J. and Zandi, R., 2003, *Phys. Rev. Lett.*, **90**, 248101.
- [7] Ercolani, G., 2003, *J. Phys. Chem. B*, **107**, 5052.
- [8] Zandi, R., Reguera, D., Bruinsma, R.F., Gelbart, W.M. and Rudnick, J., 2004, *Proc. Natl Acad. Sci. USA*, **101**, 15556.
- [9] Wales, D.J., 2005, *Phil. Trans. R. Soc. A*, **363**, 357.
- [10] Eiserling, F.A. and Dickson, R.C., 1972, *Annu. Rev. Biochem.*, **41**, 467.
- [11] Davis, B.D., Dulbecco, R., Eisen, H.N. and Ginsberg, H.S., 1980 *Microbiology*, 3rd ed. Harper and Row.
- [12] Zlotnick, A., 2004, *Proc. Natl Acad. Sci. USA*, **101**, 15549.
- [13] Wales, D.J., 1987, *Chem. Phys. Lett.*, **141**, 478.
- [14] Marzec, C.J. and Day, L.A., 1993, *Biophys. J.*, **65**, 2559.
- [15] Tarnai, T., Gáspár, Z. and Szalai, L., 1995, *Biophys. J.*, **69**, 612.
- [16] Berger, B., Shor, P.W., Tucker-Kellogg, L. and King, J., 1994, *Proc. Natl Acad. Sci. USA*, **91**, 7732.
- [17] Schwartz, R., Shor, P.W., Prevelige, P.E. and Berger, B., 1998, *Biophys. J.*, **75**, 2626.

- [18] Horton, N. and Lewis, M., 1992, *Protein Sci.*, **1**, 169.
- [19] Reddy, V.S., Giesing, H.A., Morton, R.T., Kumar, A., Post, C.B., Brooks, C.L. and Johnson, J.E., 1998, *Biophys. J.*, **74**, 546.
- [20] Zlotnick, A., Johnson, J.M., Wingfield, P.W., Stahl, S.J. and Endres, D., 1999, *Biochemistry*, **38**, 14644.
- [21] Endres, D. and Zlotnick, A., 2002, *Biophys. J.*, **83**, 1217.
- [22] Becker, O.M. and Karplus, M., 1997, *J. Chem. Phys.*, **106**, 1495.
- [23] Wales, D.J., Miller, M.A. and Walsh, T.R., 1998, *Nature*, **394**, 758.
- [24] Bamford, D.H., 2003, *Phil. Trans. R. Soc. Lond. A*, **361**, 1187.
- [25] Liljas, L., 1999, *Curr. Opin. Struct. Biol.*, **9**, 129.
- [26] Levinthal, C., 1969, In: P. DeBrunner, J. Tsibris and E. Munck (Eds) *Mössbauer Spectroscopy in Biological Systems, Proceedings of a Meeting Held at Allerton House, Monticello, Illinois* (Urbana: University of Illinois Press), p. 22.
- [27] Bryngelson, J.D., Onuchic, J.N., Socci, N.D. and Wolynes, P.G., 1995, *Proteins: Struct. Funct. Gen.*, **21**, 167.
- [28] Crick, F. and Watson, J.D., 1956, *Nature*, **177**, 473.
- [29] Wales, D.J., Doye, J.P.K., Miller, M.A., Mortenson, P.N. and Walsh, T.R., 2000, *Adv. Chem. Phys.*, **115**, 1.