focus

Inhibiting virus-capsid assembly by altering the polymerisation pathway

Peter E. Prevelige Jr

Virus capsids assemble through the repeated interaction of well-defined protein subunits in a highly specific process. Basic research into the mechanism of protein polymerisation and virus assembly suggest that inhibition of the protein–protein interactions necessary for assembly is a valid therapeutic strategy. Computer models of virus-capsid assembly have located vulnerable stages in assembly, and small-molecule inhibitors of virus assembly have been identified. The challenge will be identifying agents that block assembly with the required specificity.

Viral disease continues to cause significant human suffering and economic loss. Vaccination is an extremely successful strategy for those cases in which a vaccine exists or can be developed, and can be administered prophylactically. However, the need for antiviral agents to treat established infections will remain. To be effective, an antiviral drug must target a biochemical event unique to the virus and not shared by the host. Existing drugs target the enzymes involved in replication and viral protein processing. The success of these strategies results from a detailed understanding of mechanistic enzymology. However, drug targets need not just be enzymes – the protein–protein interactions required for virus assembly are both unique and highly specific, and, as such, represent a potential target. To date, they remain relatively unexplored, in part because the principles of protein polymerisation and virus assembly are not widely appreciated.

Nanometer-scale biological structures, such as microtubules, actin filaments, flagella and viral capsids, are built through the controlled polymerisation of similarly folded protein subunits using a small number of well-defined bonding contacts. This results in polymers that are both regular in appearance and biologically functional. Aberrant polymerisation can induce pathogenesis, such as that caused by the polymerisation of sickle-cell hemoglobin, β-amyloid protein or PrP^{Sc} in the prion diseases. This is distinct from aggregation in that, in aggregation, both the conformation of the subunits and the bonding contacts may be variable. The potential for selectively altering protein polymerisation by altering the subunit conformation or association holds clinical promise. Therapeutic agents may interfere with polymerisation or depolymerisation in a variety of ways. The anticancer chemotherapeutic agent taxol stabilises microtubles and blocks cell division[1], and it has been suggested that the antiviral WIN compounds might block viral uncoating[2]. Other agents may work either by completely blocking polymerisation or, alternatively by altering the structure of the polymer formed. This article will review the principles of controlled protein polymerisation, specifically with a view towards the inhibition of viral capsid assembly.

P. E. Prevelige Jr (prevelig@uab.edu) is at the Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294, USA.

Form determination

For a structure to be biologically functional, it must be stable and well formed. It must achieve this in an appropriate time interval and, generally, in response to a trigger signal. The concept of self assembly suggests that the information required for form determination is a property built into the subunits themselves. The repeated interaction of identical subunits by well-defined bonding interactions will necessarily result in a symmetrical structure[3]. Deviations from symmetry arise from alterations in either the subunit structure or the bonding contacts. Helical symmetry, as seen in microtubules, flagellae and rod-shaped viruses, is one of the most common forms of biological symmetry and arises from simple translation and rotation operations. In a helical polymer, every subunit, with the exception of those at the ends, is in an identical environment.

Although some viruses have helical symmetry, a large number of medically significant viruses have icosahedral symmetry. An icosahedron composed of identical subunits, each forming identical bonds to its neighbors, must have 60 subunits, arranged with 5,3,2 symmetry (Fig. 1). However, in order to form shells large enough to package the required quantity of nucleic acid, most

Copyright © 1998, Elsevier Science Ltd. All rights reserved. 0167 – 7799/98/$19.00. PII: S0167-7799(97)01154-2

viruses build shells that contain more than 60 subunits. This is accomplished by packing mutiple copies of the subunits together to form a new, larger building block for the icosahedron. Because these building blocks form the fundamental structural unit of the icosahedron, there are 60 of them, and the number of subunits packed together is known as the *T* number. As a consequence of this packing, the subunits now make different types of contacts with their nearest neighbors in the structural unit, and are termed quasiequivalent[3]. The most striking of these different interactions are the hexameric and pentameric clusters of subunits, termed capsomeres. In order to form these variable interactions, the subunits undergo conformational switching during their assembly, which requires plasticity in the subunit as well as a mechanism by which the subunit can determine the conformation to adopt[4]. Although it is apparent that the information required for successful assembly is encoded in the subunits, the exact manner in which this switching is controlled is unknown.

Proteins destined to polymerise into supramolecular structures have been observed to fold *in vivo* into assembly-incompetent forms that require activation by conformational change for assembly, providing a biological control mechanism. The subunits are activated upon association with and incorporation into the growing structure in a process that has been termed autostery[5]. The most compelling demonstration of this phenomenon is the polymerisation of bacterial flagellin. Purified flagellin monomers do not spontaneously polymerise, but fragments of assembled flagella can act as seeds to promote association. The kinetics of polymerisation can be described by Michealis–Menten kinetics, with the seed analogous to the enzyme and the flagellin monomer analogous to the substrate[6]. The flagellin monomer is bound by the seed and the N- and C-terminal residues, which are disordered in solution, fold into α-helices. Following this conformational change, whose rate is ~3 molecules sec^{-1}, the growing edge is capable of activating another flagellin monomer. The polymerisation process thus consists of repeated cycles of subunits binding at the growing edge, being activated by the growing edge into undergoing conformational change, and subsequently binding and switching fuller subunits from solution.

The rate of spontaneous switching in solution is very low. For example, the protein subunits that form the tube of bacteriophage tails have no tendency to aggregate in solution. However, if purified from preformed tails, they can be isolated in the switched conformation, which will spontaneously associate[7].

In the examples described above, conformational switching provides the temporal and positional control necessary to ensure biological function. In the case of icosahedral-virus assembly, switching is also a critical element in form determination. The conformational changes upon polymerisation in helical structures are likely to be identical in every subunit, whereas in the case of icosahedral capsids, variable conformational switching is required to account for quasi-equivalence. Crystallographic studies of the *T*=3 (i.e. 180 subunit) RNA-containing plant viruses reveal two distinct subunit conformations, the difference arising from a disorder-order transition[8]. Of these subunits, 60 have intertwined arms, which are located at the 20 icosahedral threefold axis and are triplexes, composed of residues contributed by the amino termini of three protein subunits; in the remaining 120 subunits, these amino acid residues are disordered. The intertwined arms can only fold during assembly, and it has been suggested that formation of these intertwined arms nucleates the assembly. Similarly, polyoma virus is a *T*=7 (420 subunit) lattice composed of 72 pentamers. The carboxyl end of the protein subunits within each pentamer contact the neighboring pentamer in a variable fashion. Prior to assembly, the conformation of these tails is likely to be random, with folding into the appropriate conformation coupled to assembly[9].

Alterations in subunit structure or intersubunit bonding interactions can induce profound changes in the polymer produced. For helical polymers, this is most evident in the case of bacterial flagella, where alterations in the pH and ionic strength alter the helical parameters[10]. For icosahedral viral capsids, the most striking example of this polymorphism is the ability of the major coat protein of polyoma virus, VP1, to polymerise *in vitro* into a wide variety of closed shells,

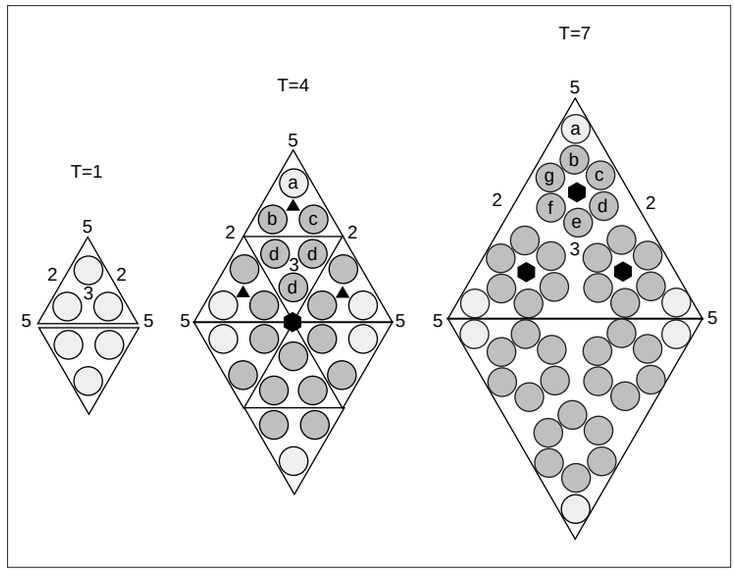

Figure 1

The geometry of icosahedral capsids. Protein subunits with a similar size can be packed into a number of lattices that display overall icosahedral symmetry. The triangular faces represent two of the 20 faces of an icosahedral lattice. In all cases, similar faces meet at the pentameric vertices (the vertices of the triangular faces). The true fivefold, threefold and twofold symmetry axes are labeled numerically for the top face; the local symmetry axes are labeled geometrically. The *T*=1 lattice is a true icosahedron, containing 60 identical subunits in identical environments. The *T*=4 lattice contains a total of 240 (60 × 4) subunits, and the subunits labeled a, b, c and d are in distinct symmetry-related positions. The *T*=7 lattice contains 420 (60 × 7) subunits. Subunits in different symmetry-related positions are labeled a–g. The *T*=7 lattice is handed, and the left-handed form is represented here. The darker subunits participate in hexameric interactions. *T* numbers represent the triangulation number. The total number of subunits corresponds to 60 × *T*, and there are *T* different subunit environments.

depending upon the salt and pH conditions. These forms include smaller-than-normal icosahedra and even octahedra[11]. The conformation of the VP1 subunits in the capsomeres appears to be unchanged but the relationship between the capsomeres is altered. A similar situation can be seen in bacteriophage P22, in which, in the presence of the scaffolding protein (a protein that assists in assembly) P22 assembles particles ~600 Å in diameter containing 420 subunits with high fidelity. In the absence of scaffolding protein, 240 Å particles containing 240 subunits are produced in addition to the 600 Å particles. The conformation of the subunits in the two structures is similar, the difference resulting primarily from the manner in which the subunits are packed in the lattice. The smaller icosahedra and octahedra produced would probably be unable to package the viral DNA, and would not be expected to be able to mature into infectious viruses. These observations suggest that subtle alterations in bonding angle or subunit conformation may be sufficient to block the formation of infectious virus. Conformational changes of this magnitude are apparently relatively easy to induce (as can be seen from the effects of pH and salt), possibly through the binding of low molecular weight compounds.

Viral capsid geometry and capsid assembly

The architectural complexity of viral capsids has important implications for the inhibition of capsid assembly. The subunit switching required by quasi-equivalence introduces design constraints on the subunits – they must be capable of adopting multiple conformations that are stable in the context of the viral lattice. Physical–chemical data suggest that the protein subunits are very flexible until incorporated into the capsid[12,13]. The conformational switching required to realise quasi-equivalent subunit packing results in pathway-dependent assembly[4]. Although topologically closed shells are expected to be more thermodynamically stable than open forms and could, in principle, form by a random-search mechanism, assembly appears to proceed along a well-defined pathway[14]. In this way, capsid assembly resembles protein folding.

Berger and colleagues have recently developed a local-rules-based computer model for virus capsid assembly[15]. This model, consistent with available biological data, posits that all the protein subunits adopt distinct conformations depending upon their position within the capsid, and that all the information required to control this switching is available to the subunit locally. They were then able to use these rules to simulate the process of virus-capsid assembly and explore the tolerance margins that would result in successful assembly. When the tolerance limits were exceeded even once during assembly, spiral malformations were induced (Fig. 2). These spiral malformations consumed large numbers of subunits that would otherwise be capable of productive assembly. This suggests that a therapeutic approach that induced spiraling might be more effective than one designed to completely inhibit subunit association. The prevalence of spiral malformations in viral infections suggests that assembly is poised near the tolerance margins and might be easily disrupted.

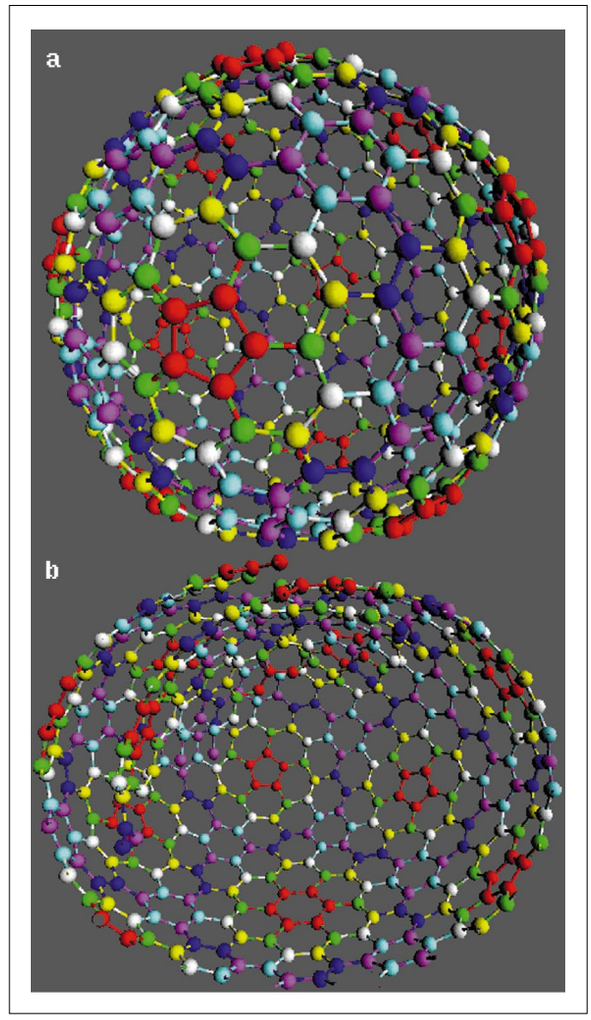

Figure 2

Molecular models of a properly (**a**) and aberrantly (**b**) assembled *T*=7 (60 × 7 subunits) capsid. The models were built using a computer simulation based on the local rules of Berger *et al.*[15] The *T*=7 rule set results in proper assembly (**a**) but when assembly nucleates from a hexamer rather than a pentamer, a spiral structure is generated (**b**). The nucleating hexamer can be seen in red at the 6 o'clock position.

Thermodynamics of polymerisation

The driving force for polymerisation is the formation of favorable bonding interactions as free protein subunits are incorporated into the growing polymer. Structural analysis suggests that, while the driving force for association is hydrophobic, the specificity is provided by shape complementarity and directional bonding[16]. These favorable interactions must be sufficient to compensate for the loss of translational, rotational and conformational entropy. The surface area buried when protein subunits interact is typically on the order of 1500 Å^2 and contains at least ten hydrogen bonds; calculations suggest that this is sufficient to overcome the entropy penalty for subunit association[17]. Frequently,

significant conformational changes in the subunits accompany assembly, and excess binding energy can be used to drive these changes. Given the relatively large size of intersubunit interfaces, one question is whether small-molecule compounds can bind tightly enough to block interactions effectively. Clackson and Wells have recently demonstrated that, although approximately 30 amino acid residues are involved in the binding of human growth hormone to the extracellular domain of its receptor, a central hydrophobic region dominated by two tryptophan residues contributes more than three quarters of the binding energy[18]. Similarly, for the bacteriophage P22 viral procapsid, Prevelige *et al.* have shown that a single tryptophan residue contributes the bulk of the entropic stabilisation[19].

Kinetics of polymerisation

Polymerisation can be divided into two kinetically and thermodynamically distinct phases: the nucleation phase and the growth phase, the nucleation phase being thermodynamically unfavorable compared with the growth phase. This 'nucleation barrier' arises because the entropic cost of immobilising the subunits and the conformational switching required to form the nucleus outweighs the intersubunit bonding energy gained during its formation (Fig. 3)[20]. These factors make it difficult to populate the intermediates required to form the nucleus and results in the appearance of a critical concentration, below which polymerisation does not occur. Once the nucleus has formed, the subsequent addition of subunits becomes favorable, because the subunits can form multiple bonding interactions, and growth occurs. The overall progress curve of the reaction shows a lag phase followed by a growth phase.

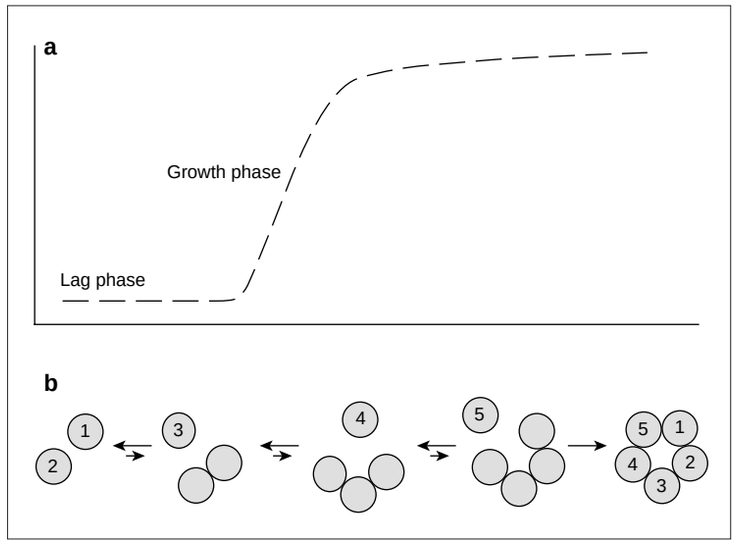

Figure 3

(**a**) A kinetic progress curve showing the lag and growth phases. The duration of the lag phase is extremely dependent upon the protein concentration. (**b**) The origin of nucleation-limited behavior. The addition of subunits 1–4 is unfavorable, because the entropy loss is only compensated by a single bonding interaction. The incorporation of the fifth subunit is favorable because two bonding interactions are gained. Subsequent subunit addition is favorable because adding subunits can form multiple interactions with adjacent subunits.

Nucleation-limited polymerisation is seen in a wide variety of systems including microtubules[20], actin[21], flagellin[6] and the polymerisation of both helical and icosahedral viral capsids[14]. Nucleation of polymerisation is also the rate-determining step in the polymerisation of β-amyloid (in Alzheimer's disease) and PrP (in scrapie)[22].

The inhibition of subunit interactions leading to nucleation represents a potentially useful target for drugs for two reasons – the thermodynamics of polymerisation suggest that weaker drugs will be more effective against nucleation than against growth, and a slight decrease in the concentration of active subunits can dramatically delay polymerisation. For example, with a nucleus size of 8, a fivefold decrease in the effective concentration will increase the lag time from one hour to 45 years[23]. A yeast actin mutant with decreased ability to nucleate has been isolated, although the growth phase of polymerisation is relatively unaffected. The mutations, two amino acid substitutions in a region thought to be involved in nucleation, result in a yeast strain that is incapable of growth at the normal physiological temperature of 37°C, indicating the critical role of nucleation in complex cellular processes[21]. In these mutants, Congo red inhibits the nucleation of PrP polymerisation and blocks the conversion of PrP into PrP^{Sc} in cell culture[24].

For icosahedral viral capsids, the best-studied assembly system is that of bacteriophage P22; this first assembles a procapsid composed of an outer shell of coat protein and an inner core of scaffolding protein whose presence plays a key role in controlling the interactions of the coat protein subunits to form a properly dimensioned closed shell. The coat and scaffolding protein can both be purified in biologically active form and used in *in vitro* assembly experiments. Assembly of P22 procapsids is mediated by hydrophobic interactions between the coat-protein subunits, and is nucleation limited. The nucleation event requires the formation of a pentameric cluster of coat-protein subunits, which will ultimately form the vertex of the nascent icosahedron[14].

Teschke *et al.* demonstrated that the low molecular weight hydrophobic dye, 1,1′-bi(4-anilino)naphthalene-5,5′-disulfonic acid (BisANS) bound to the coat protein subunit in a 1:1 ratio and was also capable of inhibiting capsid assembly *in vitro*. Characterisation of the protein–ligand complex demonstrated that the dye bound relatively weakly (K_d = 7 μM) and did not dramatically alter the secondary or tertiary structure of the protein subunit. The mechanism of action appears to be that BisANS binding induces dimerisation of the coat protein, the dimers being incapable of nucleation[25,26]. These results demonstrate the feasibility of directly inhibiting the assembly of viral capsids using low molecular weight compounds. However, BisANS binding to coat protein is not specific, and the challenge will be to discover agents that bind specifically to the protein of interest. Although the complexity and extent of the conformational changes accompanying polymerisation present opportunities, they can also be expected to pose challenges for structure-based approaches to drug design. A key will be the development

of a detailed understanding of the character of these dynamic changes using techniques such as crystallography, electron microscopy and spectroscopy.

Conclusions

Detailed studies on protein polymerisation in general and the mechanism of assembly of bacteriophage P22, in particular, have provided a deeper understanding of the process of virus-capsid assembly. In turn, these studies have provided the information necessary to demonstrate the potential of directly inhibiting virus capsid assembly with small-molecule inhibitors. It is now appropriate to consider using this approach on pathogenic viruses. The lifecycle and assembly pathways of different families of viruses have both striking similarities and striking differences. In order for this approach to succeed, it will be necessary to broaden the study of the mechanisms of virus assembly to include representative members of all classes of viruses.

Acknowledgments

I would like to acknowledge Jonathan King, Sean Moore, Matthew Parker and Roman Tuma for their insights and helpful comments on the manuscript. This work was supported in part by grants NIH GM47980 and NSF DBI-9726698 to P. E. P.

References

1 Schiff, P. B. and Horwitz, S. B. (1980) *Proc. Natl. Acad. Sci. U. S. A.* 77, 1561–1565
2 Badger, J. *et al.* (1988) *Proc. Natl. Acad. Sci. U. S. A.* 85, 3304–3308
3 Caspar, D. L. and Klug, A. (1962) *Cold Spring Harbor Symp. Quant. Biol.* 27, 1–24
4 Rossmann, M. G. (1984) *Virology* 134, 1–11
5 Caspar, D. L. (1980) *Biophys. J.* 32, 103–138
6 Asakura, S. (1968) *J. Mol. Biol.* 35, 237–239
7 Poglazov, B. F. and Nikolskaya, T. I. (1969) *J. Mol. Biol.* 43, 231–233
8 Sorger, P. K., Stockley, P. G. and Harrison, S. C. (1986) *J. Mol. Biol.* 191, 639–658
9 Liddington, R. C. *et al.* (1991) *Nature* 354, 278–284
10 Kamiya, R., Hotani, H. and Asakura, S. (1982) *Symp. Soc. Exp. Biol.* 35, 53–76
11 Salunke, D. M., Caspar, D. L. D. and Garcea, R. L. (1989) *Biophys. J.* 56, 887–900
12 Prevelige, P. E., Jr, King, J. and Silva, J. L. (1994) *Biophys. J.* 66, 1631–1641
13 Galisteo, M. L., Gordon, C. L. and King, J. (1995) *J. Biol. Chem.* 270, 16 595–16 601
14 Prevelige, P. E., Jr, Thomas, D. and King, J. (1993) *Biophys. J.* 64, 824–835
15 Berger, B., Shor, P. W., Tucker-Kellogg, L. and King, J. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 7732–7736
16 Chothia, C. and Janin, J. (1975) *Nature* 256, 705–708
17 Janin, J. (1995) *Biochimie* 77, 497–505
18 Clackson, T. and Wells, J. A. (1995) *Science* 267, 383–386
19 Foguel, D., Teschke, C. M., Prevelige, P. E., Jr and Silva, J. L. (1995) *Biochemistry* 34, 1120–1126
20 Erickson, H. P. and Pantaloni, D. (1981) *Biophys. J.* 34, 293–309
21 Buzan, J. M. and Frieden, C. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 91–95
22 Jarrett, J. T. and Lansbury, P. T., Jr (1993) *Cell* 73, 1055–1058
23 Hofrichter, J., Ross, P. D. and Eaton, W. A. (1974) *Proc. Natl. Acad. Sci. U. S. A.* 71, 4864–4868
24 Come, J. H., Fraser, P. E. and Lansbury, P. T., Jr (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 5959–5963
25 Teschke, C. M., King, J. and Prevelige, P. E., Jr (1993) *Biochemistry* 32, 10 658–10 665
26 Stafford, W. F., Liu, S. and Prevelige, P. E. (1995) in *Techniques in Protein Chemistry* (Vol. 6) (J. W. Crabb, ed.), pp. 427–432, Academic Press