BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

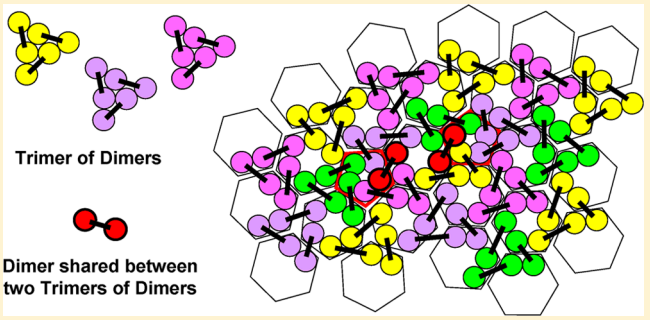
A Trimer of Dimers Is the Basic Building Block for Human Immunodeficiency Virus-1 Capsid Assembly

Manuel Tsiang,* Anita Niedziela-Majka,* Magdeleine Hung, Debi Jin, Eric Hu, Stephen Yant, Dharmaraj Samuel, Xiaohong Liu, and Roman Sakowicz

Gilead Sciences, 333 Lakeside Drive, Foster City, California 94404, United States

Supporting Information

ABSTRACT: Human immunodeficiency virus-1 (HIV-1) capsid protein (CA) has become a target of antiviral drug design in recent years. The recognition that binding of small molecules to the CA protein can result in the perturbation of capsid assembly or disassembly has led to mathematical modeling of the process. Although a number of capsid assembly models have been developed using biophysical parameters of the CA protein obtained experimentally, there is currently no model of CA polymerization that can be practically used to analyze in vitro CA polymerization data to facilitate drug discovery. Herein, we describe an equilibrium model of CA polymerization for the kinetic analysis of in vitro assembly of CA into polymer tubes. This new mathematical model has been used to assess whether a triangular trimer of dimers rather than a hexagonal hexamer can be the basic capsomere building block of CA polymer. The model allowed us to quantify for the first time the affinity for each of the four crucial interfaces involved in the polymerization process and indicated that the trimerization of CA dimers is a relatively slow step in CA polymerization in vitro. For wild-type CA, these four interfaces include the interface between two monomers of a CA dimer (K_D = 6.6 μM), the interface between any two dimers within a CA trimer of dimers (K_D = 32 nM), and two types of interfaces between neighboring trimers of dimers, either within the same ring around the perimeter of the polymer tube (K_D = 438 nM) or from two adjacent rings (K_D = 147 nM). A comparative analysis of the interface dissociation constants between wild-type and two mutant CA proteins, cross-linked hexamer (A14C/E45C/W184A/M185A) and A14C/E45C, yielded results that are consistent with the trimer of dimers with a triangular geometry being the capsomere building block involved in CA polymer growth. This work provides additional insights into the mechanism of HIV-1 CA assembly and may prove useful in elucidating how small molecule CA binding agents may disturb this essential step in the HIV-1 life cycle.



While current combination antiretroviral therapy (ART) provides long-term effective suppression of viral replication in HIV-infected patients, no definite cure is in sight. Limitations of current ART such as the emergence of resistance or chronic adverse effects and the desire to expand options for additional single-tablet regimens continue to drive the discovery efforts toward new antiretroviral targets to complement the existing treatment strategies. Human immunodeficiency virus-1 (HIV-1) capsid protein (CA) is such a target and has become a focus of antiviral research efforts in recent years.[1] Several classes of small molecule inhibitors targeting CA with potent in vitro CA polymerization modulatory activity and antiviral potency were recently reported in the literature.[2−4]

In the past decade, tremendous advances have been made in the understanding of the structure of HIV-1 CA and how CA molecules are organized in the polymerized structure of the viral capsid. The CA protein is composed of an N-terminal domain (NTD) and a C-terminal domain (CTD), each independently folded and connected to the other domain via a flexible linker.[5−7] The current model of a mature conically shaped HIV-1 capsid is based on a curved two-dimensional lattice of ~250 CA hexamers that is closed by the asymmetric incorporation of five pentamers at one end of the cone and seven pentamers at the opposite end.[8−10] Biochemical and genetic experiments suggest the existence of NTD−CTD interactions essential for capsid formation.[9,11] A high-resolution pseudoatomic model of full-length CA was constructed by docking the high-resolution structures of the two CA domains into the cryo-electron micrographic (cryo-EM) density map of a two-dimensional crystal.[12] This model is consistent with a hexameric arrangement of CA monomers and revealed intermolecular contacts, including NTD−CTD and NTD−NTD interactions of neighboring subunits as well as CTD−CTD interactions that serve as bridges between hexamers. These intermolecular interactions were further confirmed by an X-ray structure of a CA hexamer obtained by intersubunit disulfide cross-linking of mutant A14C/E45C/W184A/M185A.[13] A subsequent study of the cryo-EM structure of assembled CA in

Received: January 12, 2012
Revised: April 26, 2012
Published: May 7, 2012

   dx.doi.org/10.1021/bi300052h | *Biochemistry* 2012, 51, 4416−4428

tubular form and solution NMR structure of the CTD dimer showed a very good fit of the CTD structure into the EM density map, indicating the preservation of the CTD−CTD interaction in the assembled capsid.[14] An atomic-level modeling of the HIV capsid was finally made possible by the determination of the X-ray structure of a CA pentamer that completes the structural determination of the components of the fully assembled capsid.[15]

The availability of structural information about viral capsids in general[16] and the HIV capsid in particular has encouraged efforts to understand the mechanisms of capsid assembly that remain unclear because of the rapidity of the capsid self-assembly and the difficulty of experimentally monitoring such a dynamic process with molecular-level detail. One way to gain further insight into this self-assembly process is to use experimentally gathered structural, biochemical, and biophysical information to build mathematical models that can simulate capsid assembly. These models are largely of two types: (i) coarse-grained molecular dynamics models[17−20] and (ii) kinetic models based on differential equations.[21−27] Herein, we report the construction of a mathematical model for in vitro HIV-1 CA protein polymerization that can be used to quantify the binding interactions between CA proteins during their self-assembly into tubes. Our modeling effort was motivated by the desire to understand how small molecule CA binding agents identified during the drug discovery process or mutations in CA protein may disturb these binding interactions, leading to inhibition or acceleration of the kinetics of polymerization. In particular, we are interested in the quantification of the relative changes in equilibrium constants induced by system perturbations. Our equilibrium model has unveiled the binding constants for four types of interfaces involved in CA polymerization and indicates that the capsomere building block for HIV-1 capsid assembly is composed of a trimer of dimers with a triangular geometry.

EXPERIMENTAL PROCEDURES

Construction of Wild-Type (WT) and Mutant CA Expression Plasmids. A fragment encoding the 231 residues of wild-type HIV-1 CA protein was amplified via polymerase chain reaction (PCR) from the infectious HIV-1 cDNA plasmid, pLAI.2 obtained from J. M. Becket. The 5′ PCR primer contains an NdeI cloning site and a codon for N-terminal methionine. The 3′ PCR primer contains a stop codon and an XhoI cloning site. The amplified fragment was digested with NdeI and XhoI, gel-purified, and cloned into the NdeI and XhoI sites of the pET30a vector (catalog no. 70781, EMD BioSciences, La Jolla, CA). The double mutant, 2Mut CA (A14C/E45C), was modified from wild-type CA by site-directed mutagenesis. The DNA sequence for the quadruple mutant, 4Mut CA (A14C/E45C/W184A/M185A), was obtained by reverse translation from the amino acid sequence associated with Protein Data Bank (PDB) entry 3H47,[13] codon-optimized for *Escherichia coli* expression, and contract synthesized with DNA2.0. The insert fragment was cloned in the pJexpress 411 vector harboring a T7 promoter (DNA2.0).

Expression and Purification of Wild-Type and Mutant CA. Plasmids carrying wild-type and mutant CA were used to transform One Shot Chemically Competent BL21 DE3 cells (catalog no. C6000-03, Invitrogen, Carlsbad, CA). A single colony of each construct was grown in 2×YT medium containing 50 μg/mL kanamycin and induced for protein expression by the addition of IPTG to a final concentration of 0.5 mM. The purification protocols for wild-type and mutant CA were similar but with minor buffer composition differences that will be described elsewhere (manuscript in preparation). Briefly, cells expressing the wild-type CA protein were lysed through a microfluidizer. Soluble CA protein in the clarified lysate was concentrated by precipitation from 30% saturated ammonium sulfate. CA protein was redissolved in 50 mM Tris (pH 8.0) and functionally purified by the addition of sodium chloride to a final concentration of 2.5 M. After two rounds of functional purification, the wild-type CA protein was resuspended in 50 mM sodium phosphate buffer (pH 7.5) and dialyzed against the same buffer. The dialyzed sample was further purified by a subtractive anion exchange chromatography step using a Q-HP HiTrap column (catalog no. 17-1154-01, GE Healthcare, Piscataway, NJ).

Purification of 2Mut and 4Mut CA mutants followed the same protocol that was used for wild-type CA except that 200 mM β-mercaptoethanol was included in all buffers throughout the process. The functionally purified capsid mutant proteins were redissolved in 50 mM Tris (pH 7.5) and 40 mM β-mercaptoethanol prior to dialysis in the same buffer, and a subsequent purification with subtractive anion exchange chromatography was performed.

Preparation of the Cross-Linked A14C/E45C/W184A/M185A CA Hexamer. Soluble cross-linked CA A14C/E45C/W184A/M185 hexamers were prepared as described previously.[13] The hexamer of A14C/E45C/W184A/M185A CA was assembled by sequential dialysis of 15−20 mg/mL protein against a buffer containing 50 mM Tris-HCl (pH 8.0), 1 M NaCl, and 200 mM β-mercaptoethanol (β-ME) followed by a second dialysis against the same buffer with 0.2 mM β-mercaptoethanol and a final dialysis against 50 mM Tris-HCl (pH 8.0). Each dialysis step was performed at 4 °C for at least 12 h. After dialysis, cross-linked hexamers were separated from the trace amounts of monomers, dimers, and higher-order oligomers with a size exclusion Shodex KW2003 column (Showa Denko America, New York, NY) in the presence of 50 mM Tris-HCl (pH 8.0). The purity of the final hexamer product was determined to be >93% by capillary electrophoresis and analytical ultracentrifugation.

Capsid Polymerization Assay. Polymerization reactions for wild-type and mutant capsid proteins (A14C/E45C/W184A/M185A CA hexamer and A14C/E45C CA) were performed using a rapid dilution method similar to that described in a previous report.[28] CA was diluted to a 2× final concentration (0−80 μM) in buffer containing 50 mM $NaPO_4$ (pH 7.5) and 0.005% (v/v) Antifoam 204. Fifty microliters of diluted CA was loaded into a 96-well plate (EIA/RIA Plate, Half Area, Flat Bottom, Costar catalog no. 3695) and incubated for 5 min at 25 °C. Polymerization was initiated by addition of 50 μL of buffer containing 50 mM sodium phosphate (pH 7.5), 4 M NaCl, and 0.005% (v/v) Antifoam 204. This plate was covered with a plate sealer (Thermal Adhesive Sealing Film, Fisher-Brand catalog no. 08-408-240) and CA polymerization monitored by measuring the sample absorbance at 350 nm at 20 s intervals for 12 h at 25 °C in an M5 plate reader (Molecular Devices, Sunnyvale, CA). The final reaction conditions were 50 mM $NaPO_4$ (pH 7.5), 2 M NaCl, and 0.005% (v/v) Antifoam 204. In the case of the CA A14C/E45C mutant, 0.1 M β-mercaptoethanol was present as a carryover from protein storage buffer. β-Mercaptoethanol prevents unwanted cross-linking of surface-exposed cysteine residues in CA A14C/E45C during purification and storage.

Model Simulation and Curve Fitting. CA polymerization model simulation and curve fitting were performed using a

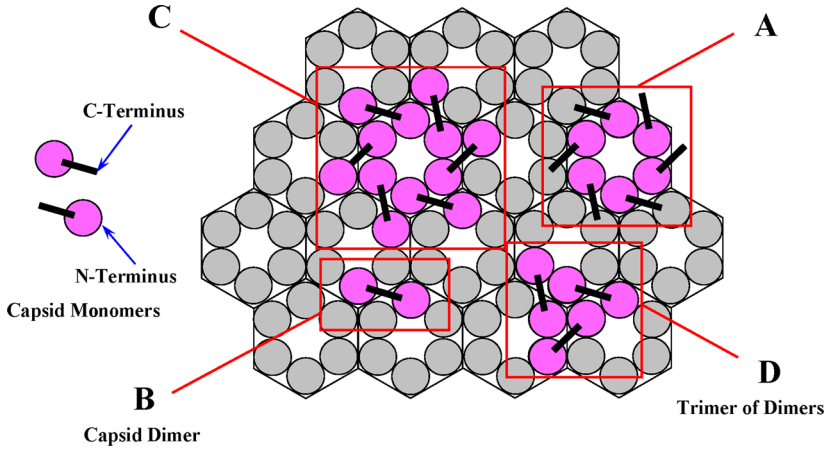


Figure 1. Representation of oligomeric forms of HIV-1 CA in a polymerized lattice. A CA monomer is represented by a lollipop, where the N-terminus resides on the side of the disk and the C-terminus at the opposite end on the side of the stick. (A) Hexagonal capsid hexamer. (B) Capsid dimer. (C) Hexamer of capsid dimers. (D) Triangular trimer of capsid dimers.

general purpose differential equation solver [Berkeley Madonna version 8.0 developed by R. Macey and G. Oster (http://www.berkeleymadonna.com)]. The change in sample absorbance at 350 nm during polymerization was fit to the equation $A_{350} = S \times \text{Poly} = S(60R + 6P_6)$. This equation relates the magnitude of light scattered by polymerized CA (measured as absorbance at 350 nm) to the concentration of polymerized species in our model. S is a factor that converts the concentration of polymerized CA expressed as monomer equivalents in units of micromolar into absorbance units (see Results for the derivation of the equation for the polymerized species, Poly). The value of S was determined during curve fitting. We have also independently measured the concentration of CA in the polymerized material (Bradford assay) for different initial CA concentrations and related this value to the sample absorbance at 350 nm after the reaction had reached the steady state. A linear relationship was observed between both quantities, with a slope comparable to the value of parameter S obtained from curve fitting the time course of polymerization with our model.

RESULTS

Stoichiometric Considerations for Modeling in Vitro Capsid Polymerization. To develop a kinetic model of in vitro capsid polymerization, we first identified the reactant and product species to be included in such a model. The obvious place to start was the capsid monomers and how they can assemble on a hexagonal lattice of hexamers as seen in cryo-EM reconstruction of tubular structures[14] and two-dimensional X-ray crystals of cross-linked hexamers.[13] To help in the illustration of such a model, we adopted a simple lollipop graphic representation of the capsid monomer where the C-terminus is at the loose end of the stick and the N-terminus is on the disk at the opposite end (Figure 1). It is unlikely that monomers can come together to form a hexamer where the C-terminal domains (CTDs) are not involved in any intermolecular contact (Figure 1A), because monomers in which the CTD is mutated (W184A/M185A) cannot dimerize and polymerize (Figure S1B of the Supporting Information), implying the necessity of intact CTDs. In fact, in the absence of salt, capsid monomers are in equilibrium with capsid dimers formed through the interaction of the CTDs from two monomers with an equilibrium dissociation constant of ~18 μM[29] (Figure S1A of the Supporting Information). This suggests that capsid dimers (Figure 1B) should be used to build higher-order oligomers when in vitro polymerization is initiated with the addition of salt. A hexamer of dimers (Figure 1C) is not likely, because it does not form a nonoverlapping self-repeating unit in the hexamer lattice. On the other hand, a trimer of dimers (Figure 1D) is the most compact structure that can be formed with dimers and is a nonoverlapping self-repeating unit in the hexamer lattice. In fact, this trimer of dimers can be seen from the two-dimensional X-ray structure of cross-linked capsid hexamers, in which three N-terminal helices form a central helix bundle with the three neighboring hexamers being connected via a dimer interface (Figure 2A,B).[13,14]

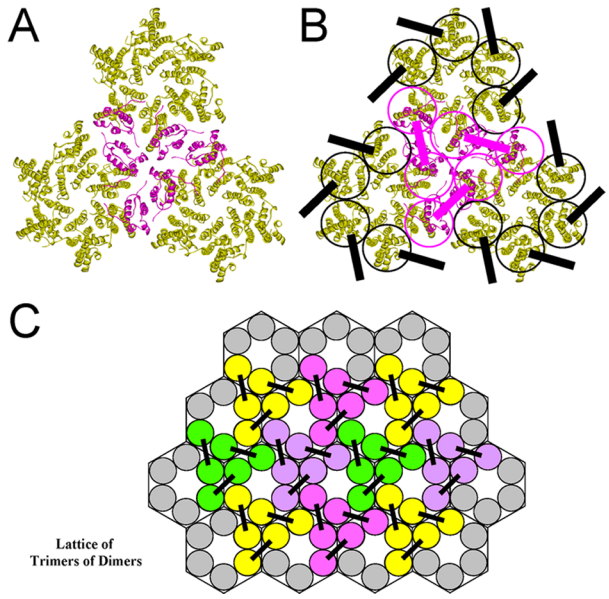


Figure 2. Polymer lattice of a trimer of dimers. (A) CTD view of the structure of three CA hexamers in a sheet of a crystal. Three crystal structures of the CA hexamer (yellow, PDB entry 3H47) were assembled by superposition of three NMR structures of the CA CTD dimer (magenta, PDB entry 2KOD). Pymol 1.2 and Discovery Studio 3.1 were used for the superposition. (B) Lollipop representation of CA hexamers overlaid on their crystal structure. The trimer of dimers is highlighted in magenta. The magenta stick represents the CTD−CTD interaction within a dimer. (C) Diagram showing how a polymerized hexagonal hexamer lattice can be populated with triangular trimers of dimers.

dx.doi.org/10.1021/bi300052h | *Biochemistry* 2012, 51, 4416−4428

For these reasons, we propose to use the trimer of dimers (triangular hexamer) as the next-level oligomer in our modeling (Figure 2C).

The next-level assembly would involve successive additions of either monomers, dimers, or trimers of dimers to the initial intermediate of the trimer of dimers, resulting in the formation of a lattice of trimers of dimers (triangular hexamers). Because there may be different ways of conceiving the growth of polymer tubes, we sought to use the simplest way that is consistent with the architecture of the polymer tube grown in vitro.[14] Hence, these stepwise additions can be simplified as the self-assembly of multiple trimers of dimers coming together to form a triangular hexameric lattice that grows as a tube. From the cryo-EM image, the polymer tube has a cross section with a diameter of ~40 nm and each hexamer unit a diameter of ~9.25 nm. The arrangement of the hexamers on the tube can be interpreted as the stacking of successive slightly slanted rings of hexamers. The diameters of the polymer tube and the hexamer unit allow 10−12 hexamers to be positioned along the tube perimeter to form a ring structure. On the basis of these dimensions, we used a simple way of modeling polymer tube growth in which 10 trimers of dimers first form a ring and additional trimers of dimers will use this ring as a foundation to stack successive rings of 10 trimers of dimers that lengthen the polymer tube.

Simplifications of the Polymerization Steps. We set out to build an equilibrium model of in vitro capsid polymerization as a system of differential equations each describing the rate of change of a species of the reactant or product. Because the number of species of reactants and products is practically infinite as the polymerization reaction progresses, some simplification is necessary to reduce the number of differential equations in the system. We resorted to two simplification methods: (i) condensation of a finite number of successive similar steps into one and (ii) substitution of an infinite number of polymerized species with their monomeric component. To illustrate the condensation method, let us consider a simple dimerization equilibrium reaction (scheme A). In this case, no condensation of steps is necessary because it is already a single-step reaction with a dissociation constant ($K_D = k_{-1}/k_1$) expressed in micromolar (k_1 and k_{-1} are the on rate and off rate constants, respectively).

$$A + A \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} A_2 \qquad \text{(A)}$$

In the case of a trimerization reaction, we have two identical reactant concentration-dependent addition steps (for linear trimers) followed by a final concentration-independent ring closure step (for trimers with a radial symmetry) (scheme B). In reality, the second and final steps may not be distinguishable and may actually constitute a single step with the addition of the third subunit.

$$A + A \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} A_2 + A \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} AAA \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} A_3 \qquad \text{(B)}$$

Scheme B can be simplified as scheme C by condensing the three steps into a single step with a composite $K_D = k_{off}/k_{on} = (k_{-1}k_{-2}k_{-3})/(k_1k_2k_3)$ expressed in μM^2.

$$A + A + A \underset{k_{off}}{\overset{k_{on}}{\rightleftarrows}} A_3 \qquad \text{(C)}$$

Higher-order oligomerization can be simplified in a similar manner with composite K_D values expressed in units of μM^3, μM^4, and μM^5 for tetramerization, pentamerization, and hexamerization, respectively. To illustrate the substitution method of simplification, let us consider a situation in which monomers polymerize into strings of different lengths (Figure 3A). It can be seen that the number of

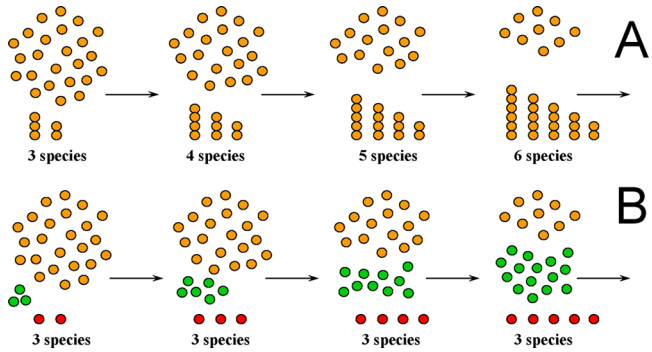


Figure 3. Substitution of an infinite number of polymerized species with their monomeric component. (A) As polymerization progresses, the total number of species increases from three to six in the time interval illustrated here. (B) Substitution of each polymerized species by a monomer base unit (red) and its polymerized monomeric component (green) reduces the total number of species to three irrespective of the progress of polymerization. This is an important simplification scheme used during the construction of our CA polymerization model.

polymerized species increases in a given time interval. Each species would require a rate equation for the accounting of all the masses in the system. In the substitution method of simplification, all the polymerized species are broken down into two species for the purpose of accounting: base unit and polymerized monomer (Figure 3B). This allows for the reduction of a potentially infinite number of equations to just three that capture both the free and polymerized monomers, together with the base unit. Both condensation and substitution are used in the construction of our model for in vitro CA polymerization.

Equilibrium Model for in Vitro Capsid Polymerization. We assumed an equilibrium between capsid monomers and dimers in the absence of salt and that addition of salt not only alters the equilibrium dissociation constant between monomers and dimers but also introduces new equilibrium constants between dimers and trimers of dimers and between trimers of dimers and higher-order polymer species. Essentially, our model treats in vitro capsid polymerization as the process by which the system regains a new equilibrium after the addition of salt. This model is represented by scheme D, where C is the capsid monomer, C_2 is the capsid dimer, C_6 is the trimer of dimers, R is the ring of 10 trimers of dimers, and P_6 is the polymerized trimer of dimers. In this model, a new ring R can be generated only from the assembly of $10C_6$ on their own. Once a ring R is formed, it can serve as a base unit for the polymerization of additional free C_6. In the last leg of scheme D,

$$\begin{array}{ccccccccc} & & & & 10(P_6) + R & & & & \\ & & & & k_{-4} \uparrow\downarrow k_4 & & & & \\ 10(6C) & \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} & 10(3C_2) & \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} & 10(C_6) & \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} & R & & \\ & & & & + & & + & & \\ & & & & R & \underset{k_3}{\overset{k_{-3}}{\rightleftarrows}} & 10(C_6) & \underset{k_2}{\overset{k_{-2}}{\rightleftarrows}} & 10(3C_2) \underset{k_1}{\overset{k_{-1}}{\rightleftarrows}} 10(6C) \\ & & & & k_4 \uparrow\downarrow k_{-4} & & & & \\ & & & & R + 10(P_6) & & & & \end{array} \qquad \text{(D)}$$

dx.doi.org/10.1021/bi300052h | Biochemistry 2012, 51, 4416−4428

with on and off rate constants of k_4 and k_{-4}, respectively, R is neither produced nor consumed, and the designation P_6 is merely a mathematical construct for the accounting of the free C_6 species that have been polymerized onto R to lengthen the polymer tube. Because of simplifications by condensations and substitution, the model has only five species described by system I of five differential equations.

$$\begin{aligned}
\frac{dC}{dt} &= 60\left[k_{-1}C_2 - k_1\left(\frac{C}{2}\right)^2\right]\\
\frac{dC_2}{dt} &= 30\left[k_1\left(\frac{C}{2}\right)^2 - k_{-1}C_2 - k_2\left(\frac{C_2}{3}\right)^3 + k_{-2}C_6\right]\\
\frac{dC_6}{dt} &= 10\left[k_2\left(\frac{C_2}{3}\right)^3 - k_{-2}C_6 - k_3\left(\frac{C_6}{10}\right)^{10} + k_{-3}R\right.\\
&\quad \left. - k_4R\left(\frac{C_6}{10}\right)^{10} + k_{-4}R\right]\\
\frac{dR}{dt} &= k_3\left(\frac{C_6}{10}\right)^{10} - k_{-3}R\\
\frac{dP_6}{dt} &= 10\left[k_4R\left(\frac{C_6}{10}\right)^{10} - k_{-4}R\right]
\end{aligned} \quad \text{(I)}$$

The starting CA monomer−dimer equilibrium mixture is a clear solution that does not scatter light. Upon initiation of polymerization by the elevation of the ionic strength, larger oligomer and polymer species gradually form, and this polymerization process is accompanied by an increase in the turbidity of the solution as measured by an increased absorbance. To analyze the kinetic data that result from light scattering on polymerized capsid proteins measured as absorbance, we divide the five species in system I into two categories: the nonpolymerized species (NPoly) that do not scatter light, represented by eq II, and the polymerized species (Poly) that are light scattering, represented by eq III. C_6 like C and C_2 does not scatter light because purified cross-linked hexamers have no absorbance.

$$\text{NPoly} = C + 2C_2 + 6C_6 \quad \text{(II)}$$

$$\text{Poly} = 60R + 6P_6 \quad \text{(III)}$$

Both NPoly and Poly are expressed as capsid monomers in units of micromolar. Equation III was used to fit the time course data of polymerization as monitored by absorbance measurement. The average polymer tube length (APTL) expressed as the number of rings per tube at any given time point can be calculated using eq IV.

$$\text{APTL} = \text{Poly}/(60R) \quad \text{(IV)}$$

The model described by system I has four sets of rate constants describing four binding equilibria: dimerization of monomers, trimerization of dimers, decamerization of hexamers (i.e., formation of rings made of trimers of dimers), and polymerization of hexamers (i.e., assembly of successive rows of hexamers onto a ring to lengthen the polymer tube). The four corresponding equilibrium dissociation constants that can be determined by curve fitting of the time course data are as follows: $K_{dimer} = k_{-1}/k_1$ (μM), $K_{trimer\_of\_dimers} = k_{-2}/k_2$ (μM^2), $K_{decamer} = k_{-3}/k_3$ (μM^9), and $K_{pol} = k_{-4}/k_4$ (μM^{10}), respectively. These equilibrium constants define four types of interfaces with affinity constants expressed in units of μM, which can be calculated {e.g., $K_{dimer\ interface} = K_{dimer} = k_{-1}/k_1$, $K_{trimer\ interface} = (k_{-2}/k_2)^{1/2}$, $K_{decamer\ interface} = (k_{-3}/k_3)^{1/9}$, and $K_{pol\ interface} = (k_{-4}/k_4)]^{1/10}$.

Simulations Using the Capsid Polymerization Model. To illustrate the performance of our model and to show the distribution of various CA species during polymerization, we ran simulations using values of affinity constants derived from curve fits for the four types of interfaces with a K_{dimer} of 16 μM, a $K_{trimer\ interface}$ of 27 nM, a $K_{decamer\ interface}$ of 470 nM, and a $K_{pol\ interface}$ of 130 nM, three initial concentrations of capsid monomers ($[C]_o$ = 50, 40, and 30 μM), and the initial concentrations of all the remaining species set to zero ($[C_2]_o = [C_6]_o = [R]_o = [P_6]_o = 0$ μM) (Figure 4). The time courses of [Poly] (Figure 4A, red curves) all start with a lag phase that is longer for lower initial $[C]_o$ values and end with a plateau that is higher for higher initial $[C]_o$ values when equilibrium is reached. The time courses of [NPoly] (Figure 4A, blue curves) are descending curves that are mirror images of the time courses of [Poly]. The sum of the two curves ([Poly] + [NPoly]) at any time point always equals the initial monomer concentration $[C]_o$, showing conservation of mass. The time courses of free hexamers, [C_6] (magenta curves), and of polymerized hexamer, [P_6] (green curves), are shown in Figure 4B. [C_6] reaches a peak and then decreases as the free hexamers are used to form rings and build polymer tubes. The time courses of [P_6] closely mirror those of [Poly] because they are the main constituent of the polymerized tubes. The time courses of the monomer, [C] (brown curves), and dimers, [C_2] (dark cyan curves), are shown in Figure 4C. The levels of monomers rapidly decrease to a new equilibrium concentration that is independent of the initial concentration and is consistent with the presence of a critical concentration characteristic of nucleation-limited polymerization processes.[30−32] In this model simulation, the critical concentration represented by the new equilibrium level is roughly equal to 10 μM (Figure 4C). The levels of dimers rise initially and decrease to a new equilibrium concentration as they are being used to build the higher-affinity complex of trimers of dimers, [C_6], in this simulation. The time courses of [R] display the same type of initial lag phase as those of [P_6], but then quickly rise to a peak and fall to a plateau (Figure 4D). It should be noted that the rise of [R] lags behind the rise of [C_6] and the rise of [P_6] lags behind the rise of [R]. As polymer tubes started to grow using some rings as base units, other rings could dissociate into free hexamers and be used to grow existing polymer tubes instead of serving as base units to grow a new tube. This is consistent with the polymer interface having a higher affinity than the decamer interface in this simulation and explains the decrease in [R].

Wild-Type Capsid Polymerization. Wild-type capsid polymerization kinetics using six different total capsid concentrations (15.6, 20.8, 26.0, 31.2, 36.4, and 41.6 μM) were first analyzed using our model by curve fitting with system I. Initially, we performed our analysis by assuming that all the capsid proteins were present as monomers before the addition of salt (Figure 5A and Table 1). Using this curve fit, we were able to determine at 2 M NaCl the affinities of the dimer interface (K_{dimer} = 4.8 ± 3 μM), the trimer of dimer interface ($K_{trimer\ interface}$ = 32 ± 9 nM), the decamer interface ($K_{decamer\ interface}$ = 460 ± 30 nM), and the polymer interface ($K_{pol\ interface}$ = 137 ± 5 nM), where the errors are the standard deviations from the mean of six determinations using six CA

dx.doi.org/10.1021/bi300052h | Biochemistry 2012, 51, 4416−4428

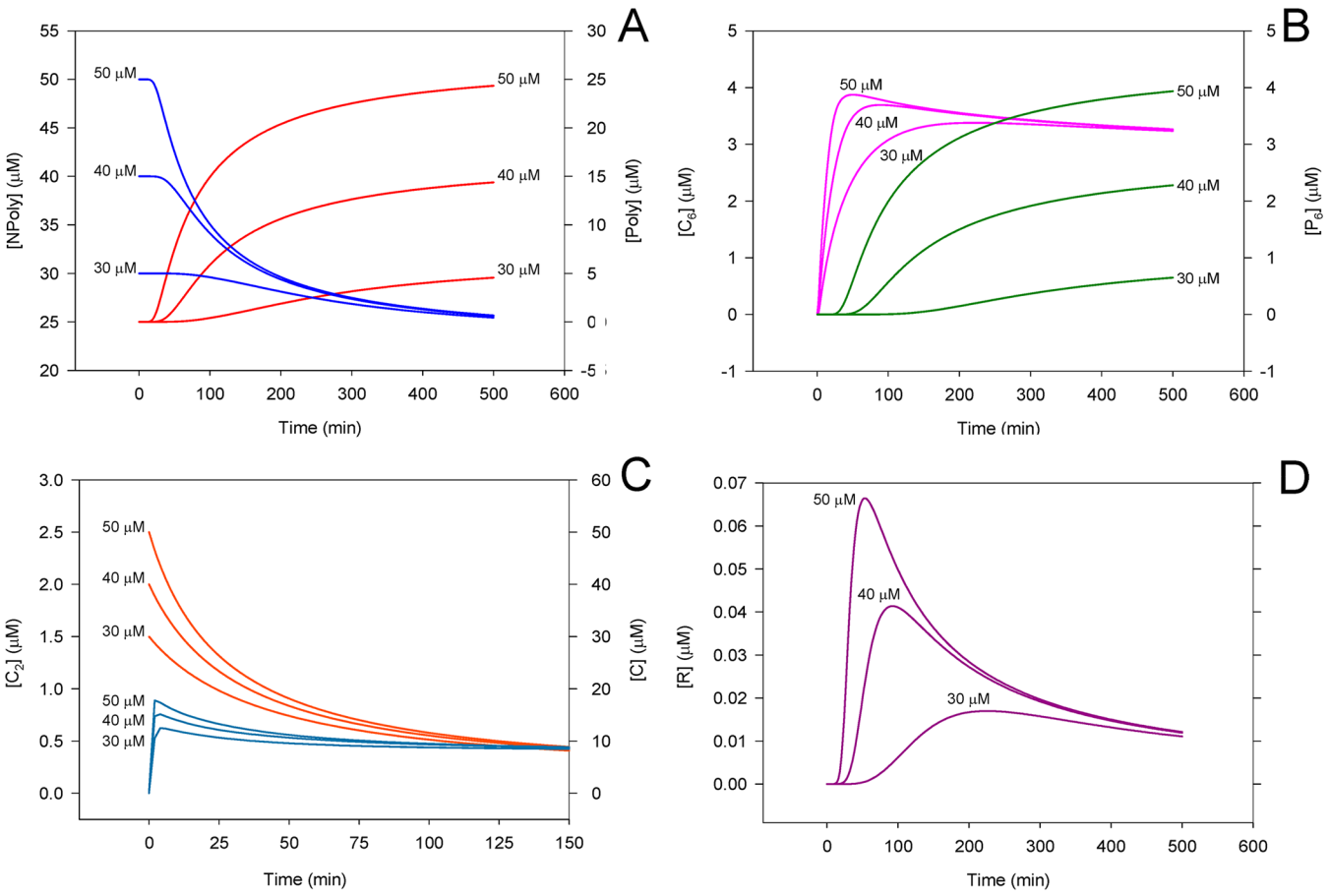


Figure 4. Simulation of the time course of the five species described in our CA polymerization model. (A) Total nonpolymerized species, [NPoly] (blue lines), and total polymerized species, [Poly] (red lines). Both [NPoly] and [Poly] are expressed as the concentration of their monomer components. (B) Free hexamers, [C_6] (magenta lines), and polymerized hexamers, [P_6] (green lines). (C) Monomers, [C] (brown lines), and dimers, [C_2] (cyan lines). (D) Decamer ring, [R] (purple lines). Simulations were performed using three different total CA concentrations (30, 40, and 50 μM) using realistic dissociation constants for the four types of interfaces (K_{dimer} = 16 μM, $K_{trimer\ interface}$ = 27 nM, $K_{decamer\ interface}$ = 470 nM, and $K_{pol\ interface}$ = 130 nM).

concentrations. This result shows that the trimer of dimer interface has the highest affinity of the four interfaces and suggests that trimers of dimers are the capsomere building blocks for higher-order oligomer structures. The observation that the decamer interface has a 3.4-fold weaker affinity than the polymer interface is consistent with a preference for polymer tube growth over decamer base unit formation. A closer examination of the curve fit (Figure 5A and Figure S4A of the Supporting Information) showed that although the fit was generally good, they displayed a longer initial lag phase then the data. In fact, before the addition of salt, capsid proteins already existed as an equilibrium mixture of monomers and dimers, and the preexisting dimers could give polymerization a head start. To account for the preexistence of dimers before the addition of salt, we refit the data by allowing the model to determine the presalt equilibrium between monomers and dimers (Figure 5B and Figure S4B of the Supporting Information). With this simple modification, the curve fit became nearly perfect and we obtained in addition a presalt affinity for the dimer interface of 14 μM, which is similar to the previously reported dimer dissociation constant of 18 μM.[29] The postsalt dimer interface affinity slightly increased to 6.6 ± 0.9 μM, while the values for the other three interfaces remained almost the same, 32 ± 17, 438 ± 48, and 146 ± 13 nM for trimer, decamer, and polymer interfaces, respectively. The fact that incorporation of dimers preformed before the addition of salt into the curve fit can reduce the initial lag phase to allow for a better curve fit suggests that dimerization might be rate-limiting after the addition of salt according to our model. This is consistent with the observation that the dimerization on rate constant ($k_1 = 7 \times 10^{-5}\ \mu M^{-1}\ min^{-1}$) is the slowest of the four on rate constants (Table 1). Using the determined dissociation constants, we calculated the corresponding free energies of dissociation (Table 2). For example, the free energy of dimer dissociation is 6.99 kcal/mol for the wild-type capsid. Because there is one interface per dimer, the dissociation energy per dimer interface is also 6.99 kcal/mol. For a trimer of dimers, the overall free energy of dissociation into component dimers is 20.23 kcal/mol. Because there are three interfaces between dimers in a trimer, the dissociation energy per interface is 20.23/3 = 6.74 kcal/mol for the wild-type capsid. Using the same principle, we also calculated a dissociation free energy of 7.72 kcal/mol between two adjacent trimers of dimers within a decamer ring and a dissociation free energy of 9.22 kcal/mol between two adjacent trimers of dimers, from two contiguous decamer rings.

Effect of Switching Affinities between Interfaces. To examine whether the affinities assigned by curve fitting to the four interfaces are the only possible assignments, we simulated four polymerization time courses by switching the affinities

dx.doi.org/10.1021/bi300052h | *Biochemistry* 2012, 51, 4416−4428

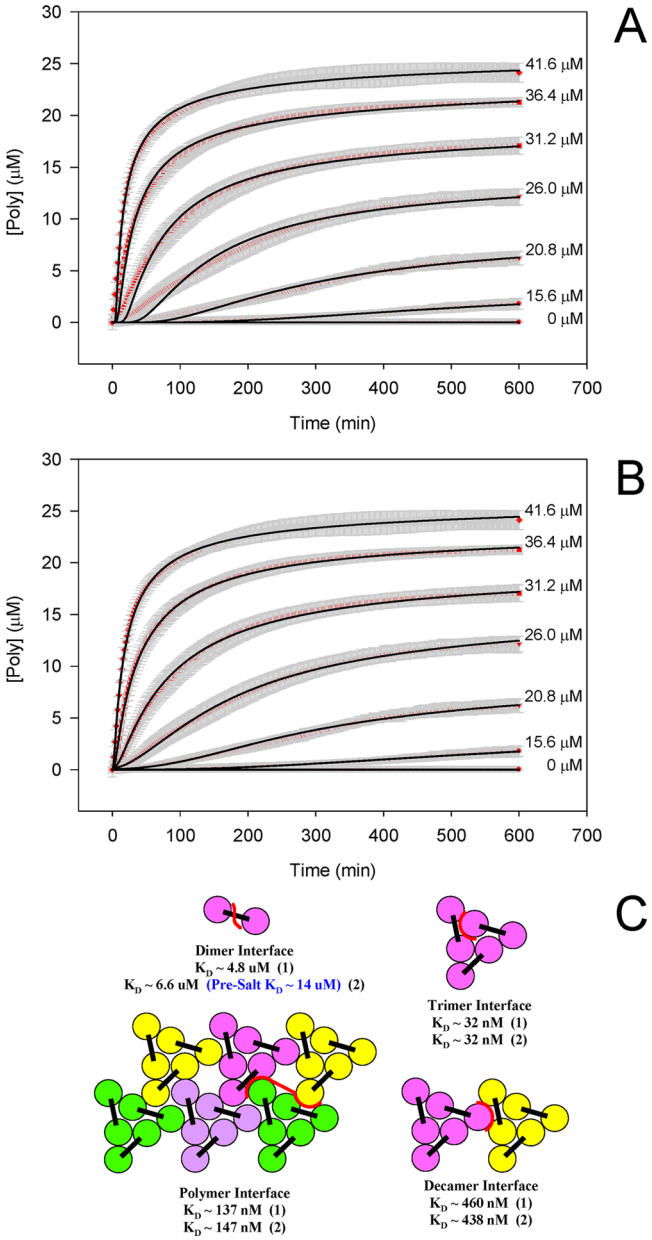


Figure 5. Wild-type CA polymerization kinetics. (A) Curve fitting with the model using the assumption that all CA proteins are monomeric at time zero [CV(rmsd) = 3.35%]. (B) Curve fitting with the model using the assumption that there is equilibrium between monomers and dimers at time zero [CV(rmsd) = 1.73%]. The data are plotted as the means of triplicate runs (red circles), with the standard deviations shown as gray bars. The curve fits are shown as black lines (see also Figure S4 of the Supporting Information). Six total concentrations of CA expressed as concentrations of monomer components were tested (15.6, 20.8, 26, 31.2, 36.4, and 41.6 μM). (C) Determined affinities for the four types of interfaces. The numbers 1 and 2 denote results derived from curve fits in panels A and B, respectively. Each type of interface is highlighted with a thick red line.

determined by curve fitting between the four types of interfaces (Figure 6). The red curve was simulated using essentially the four affinities determined by curve fitting in their original assignment to the four types of interfaces. These affinities resulted in an average polymer tube length of ~34 rings near the end of polymerization, consistent with the length (551 ± 208 nm or ~55 rings) of CA polymer tubes observed via EM. We found that the affinity of the dimer interface cannot be switched with that of any other interface without drastically weakening polymerization. In the example shown where the dimer interface affinity was switched with the decamer interface affinity (blue curve), formation of the decamer ring became so slow that all polymerization activity was spent on elongating the tube, resulting in an average tube length of ~35000 rings. When the decamer interface affinity was switched with that of the polymer interface (dark yellow curve), overall polymerization was enhanced by ~50%, but because of increased decamer interface affinity, most polymerization activity was spent on decamer ring formation rather than tube elongation. As a result, the average tube length remains at the one-ring stage even after 500 min. When the trimer interface affinity was switched with the polymer interface affinity, overall polymerization was enhanced by ~5% with a steeper slope. The increase in polymer interface affinity dramatically increased the average tube length to ~1400 rings. In conclusion, this simulation validates the assignment of affinities derived from curve fitting with our model to the four types of interfaces, because the switching of affinities between interfaces produces simulated curves that cannot be superimposed on each other and average tube lengths that are either too short or too long to be realistic. In addition, a sensitivity analysis was performed to determine the lower and upper limits for each of the eight rate constants, showing that k_1 and k_{-1} are the most sensitive and k_4 and k_{-4} are the least sensitive, while k_2, k_{-2}, k_3, and k_{-3} have intermediate sensitivity (Table S1 and Figure S5 of the Supporting Information). This could be due to a larger role played by the dimerization step in shaping the degree of sigmoidicity with a fast change in rate in the early phase of the time course, whereas later steps of ring formation and tube elongation bring the time course to a plateau with a slower change in rate. The similarity between the intervals of each pair of rate constants suggests some level of correlation between the on and off rate constants of each step. However, the affinity constants that are of primary interest to us are relatively independent of each other as demonstrated by the affinity switching simulations.

A14C/E45C/W184A/M185A Capsid Polymerization. To confirm that the hexamer formed during WT CA polymerization is indeed a trimer of dimer, we resorted to comparative analysis of affinities for the interfaces formed during polymerization of two mutant CA species with those of WT CA. One of these mutants is the quadruple CA mutant A14C/E45C/W184A/M185A (4Mut) previously used for generating disulfide bridge cross-linked hexamers.[13] The A14C/E45C double mutation in the NTD was introduced to allow disulfide cross-linking between neighboring monomers within a hexamer unit, while the W184A/M185A double mutation in the CTD was designed to abolish dimerization of CA monomers. This dimerization interface can also enhance the contact between neighboring cross-linked hexamers in the polymerized lattice, and its removal allows the production of soluble cross-linked hexamers. Using the previously published method,[13] cross-linked hexamers of this quadruple mutant were purified and tested in our CA polymerization assay at cross-linked hexamer concentrations of 6.7, 5.0, 3.3, 2.5, and 1.7 μM. These 4Mut cross-linked hexamers displayed much faster polymerization kinetics than the wild-type capsid, with the total absence of any detectable initial lag phase (Figure 7A). To analyze the polymerization time course of the 4Mut cross-linked hexamers using our CA polymerization model, we assumed that there is no CA monomer at time zero and that

Table 1. Rate and Equilibrium Constants of Wild-Type and Mutant CA Polymerization

	WT CA[a]		2Mut CA[a]			x-fold change	
rate and equilibrium constants	$[C_2]_o = 0$	$[C_2]_o \neq 0$	$[C_2]_o = 0$	$[C_2]_o \neq 0$	4Mut CA[a]	WT/2Mut[b]	WT/4Mut[b]
k_1 (μM^{-1} min^{-1})	0.00013	0.00007	0.00040	0.00073	–	0.10	–
k_{-1} (min^{-1})	0.00061	0.00047	0.00003	0.00001	–	72.2	–
k_2 (μM^{-2} min^{-1})	0.63	0.55	41.93	37.90	–	0.01	–
k_{-2} (min^{-1})	0.00066	0.00057	0.00004	0.00003	–	22.6	–
k_3 (μM^{-9} min^{-1})	821	771	1590	640	753	1.20	1.02
k_{-3} (min^{-1})	0.78	0.46	0.03	0.02	0.00025	23.7	1851
k_4 (μM^{-10} min^{-1})	514	606	1269	2833	364740	0.21	0.0017
k_{-4} (min^{-1})	1.22×10^{-6}	2.77×10^{-6}	1.76×10^{-6}	1.75×10^{-6}	4.46×10^{-6}	1.58	0.62
pre-salt K_{dimer} (μM)	–	14	–	26	–	0.53	–
K_{dimer} (μM)	4.78	6.59	0.074	0.009	–	731.8	–
$K_{trimer\ of\ dimers}$ (μM^2)	0.0011	0.0010	0.0000011	0.0000007	–	1546	–
$K_{decamer}$ (μM^9)	0.00095	0.00060	0.00002	0.00003	3.31×10^{-7}	19.7	1809
K_{pol} (μM^{10})	2.37×10^{-9}	4.57×10^{-9}	1.39×10^{-9}	6.18×10^{-10}	1.22×10^{-11}	7.40	374
$K_{d\ trimer\ interface}$ (μM)	0.032	0.032	0.0010	0.0008	–	39.3	–
$K_{d\ decamer\ interface}$ (μM)	0.4615	0.4384	0.3033	0.3148	0.1905	1.39	2.30
$K_{d\ pol\ interface}$ (μM)	0.1372	0.1465	0.1301	0.1200	0.0810	1.22	1.81
average tube length	18	11	4	8	8		

[a]Rate and affinity constants represent the means determined from six or seven curves with a decreasing CA concentration. [b]Ratio calculated from parameters obtained with $[C_2]_o \neq 0$.

Table 2. Gibbs Free Energies of Dissociation of the Interfaces of Wild-Type and Mutant CA Polymerization

	WTCA[a]			2Mut CA[a]				
	$[C_2]_o = 0$	$[C_2]_o \neq 0$		$[C_2]_o = 0$	$[C_2]_o \neq 0$		4Mut CA[a]	
Gibbs free energy of dissociation[b] (kcal/mol)	overall	overall	per interface	overall	overall	per interface	overall	per interface
$\Delta G_{0\ dimer}$	7.18	6.99	6.99	9.62	10.86	10.86		
$\Delta G_{0\ trimer\ of\ dimers}$	20.21	20.23	6.74	24.26	24.53	8.18		
$\Delta G_{0\ decamer}$	76.96	77.23	7.72	79.17	78.98	7.90	81.63	8.16
$\Delta G_{0\ pol}$	92.62	92.23	9.22	92.93	93.41	9.34	95.71	9.57

[a]Gibbs free energies represent the means calculated from six or seven curves with a decreasing CA concentration. [b]Gibbs free energies were calculated from dissociation constants for a temperature of 22 °C and therefore have a positive sign.

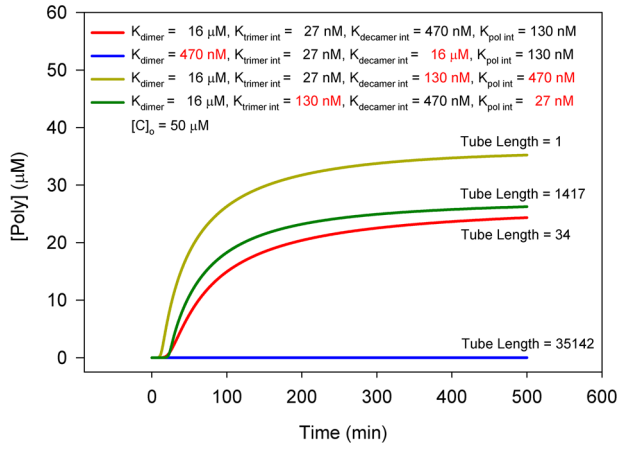


Figure 6. Simulation of the time course of total polymerized species: Effect of swapping dissociation constants among the four types of interfaces. Simulations were performed using a CA concentration of 50 μM, using realistic dissociation constants for the four types of interfaces (K_{dimer} = 16 μM, $K_{trimer\ interface}$ = 27 nM, $K_{decamer\ interface}$ = 470 nM, and $K_{pol\ interface}$ = 130 nM) before dissociation constants were swapped (red line). Swapped dissociation constants for each successive simulation are highlighted in red.

the only species present at time zero is the free cross-linked hexamer. This assumption is supported by the fact that our preparation of hexamers is more than 93% pure based on capillary electrophoresis and analytical ultracentrifugation. In addition, the rate constants k_1, k_{-1}, k_2, and k_{-2} were all set to zero in system I to inactivate the dimerization and trimerization segments of the model before curve fitting. Our model was able to generate a good fit of the data and determine the affinities of decamer and polymer interfaces (Figure 7 and Table 1). Again, the resulting affinity of the decamer interface (190 ± 7 nM) is somewhat weaker than the affinity of the polymer interface (80 ± 8 nM), allowing elongation of the polymer tubes. Because of cross-linking, the dimer and trimer interfaces no longer exist as separate entities, and the hexamers are hexagonal and different from the trimers of dimers likely formed by wild-type CA which are triangular in geometry. Thus, the decamer and polymer interfaces result from the contacts between hexagonal hexamers instead of contacts between trimers of dimers. The product of these two affinities (0.1905 × 0.0810 = 0.0154) is a measure of the capsomere polymerization affinity between the cross-linked hexagonal hexamers of 4Mut CA. Compared with the capsomere polymerization affinity of wild-type CA (0.4384 × 0.1465 = 0.0642), the capsomere polymerization affinity between cross-linked hexagonal hexamers is ~4.2-fold stronger. This stronger affinity is consistent with the fact that these new decamer and polymer interfaces were formerly part of the trimer interface (Figure 7B, dashed line contour) whose affinity was the highest among the four

dx.doi.org/10.1021/bi300052h | Biochemistry 2012, 51, 4416–4428

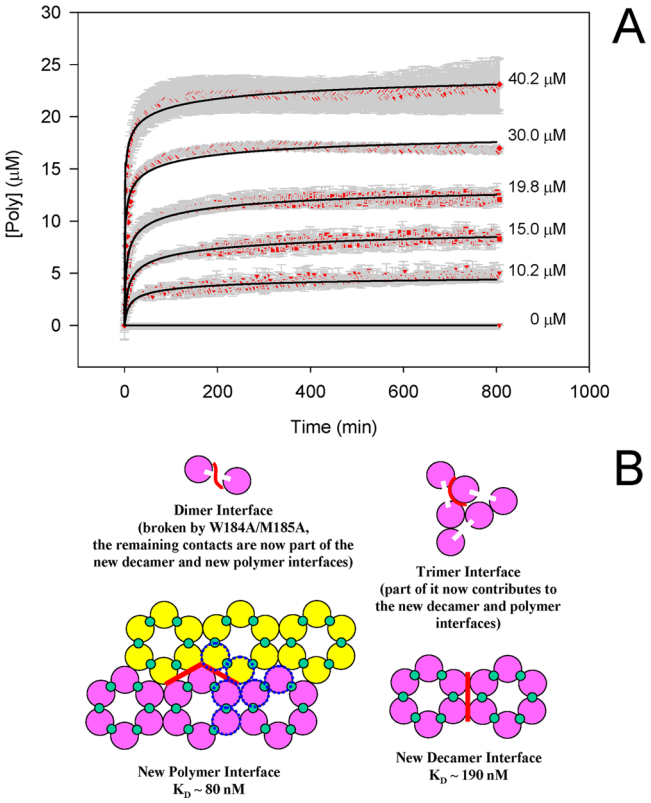


Figure 7. A14C/E45C/W184A/M185A CA cross-linked hexamer polymerization kinetics. (A) Curve fitting with the model using the assumption that all CA proteins are free cross-linked hexamers at time zero. The data are plotted as the means of triplicate runs (red circles) with standard deviations shown as error bars. The curve fits are shown as black lines. Five total concentrations of CA expressed as concentrations of monomer components were tested (10.2, 15, 19.8, 30, and 40.2 μM). (B) Determined affinities for the two types of interfaces between cross-linked hexamers. Each type of interface is highlighted with a thick red line. Dimers and trimers of dimers do not exist for this quadruple mutant. Cross-linking between monomers is shown with small green circles. The contour of a trimer of dimers in the cross-linked hexamer polymer lattice is highlighted with blue dashed lines.

types of interfaces in the wild-type polymer lattice. This differential capsomere polymerization affinity between 4Mut and WT CA constitutes evidence that trimers of dimers (triangular hexamers) are the capsomere building blocks of the WT CA polymer lattice, not the hexagonal hexamers, the existence of which as a separate entity requires engineered cross-linking.

A14C/E45C Capsid Polymerization. The second mutant is the double CA mutant A14C/E45C (2Mut). Like quadruple mutant 4Mut, 2Mut contained the two cysteine mutations that allow for potential cross-linking of neighboring monomers into a hexamer but was capable of dimer formation. Purified 2Mut does not contain cross-linked hexamers because of the presence of high concentrations of reducing agent in the protein storage buffer but is an equilibrium mixture of monomers and dimers similar to wild-type CA. Seven concentrations (22.5, 20, 17.5, 15, 12.5, 10, and 7.5 μM) of 2Mut CA were tested in the polymerization assay (Figure 8). 2Mut CA polymerization at high concentrations displayed very fast kinetics similar to that of 4Mut CA and devoid of detectable sigmoidicity. However, at lower CA concentrations (i.e., <15 μM), the sigmoidicity of the time course became apparent, suggesting that the 2Mut CA

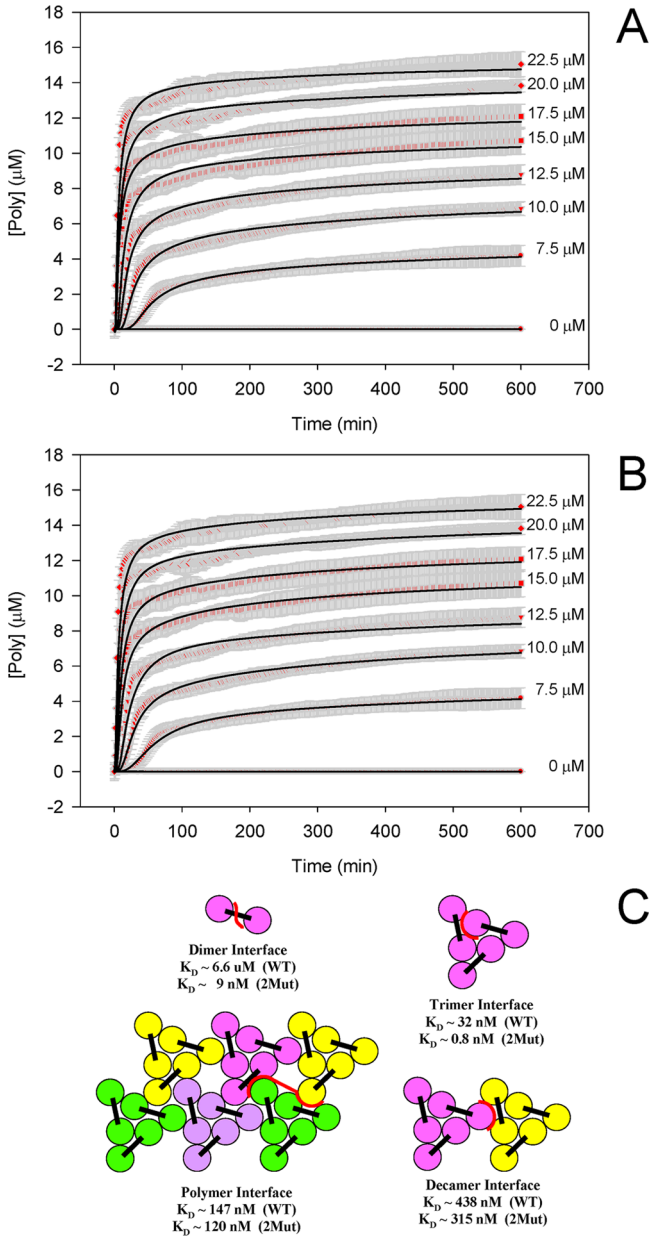


Figure 8. A14C/E45C CA polymerization kinetics. (A) Curve fitting with the model using the assumption that all CA proteins are monomeric at time zero [CV(rmsd) = 3.49%]. (B) Curve fitting with the model using the assumption that there is an equilibrium between monomers and dimers at time zero [CV(rmsd) = 2.88%]. The data are plotted as means of triplicate runs (red circles) with standard deviations shown as error bars. The curve fits are shown as black lines. Seven total concentrations of CA expressed as concentrations of monomer components were tested (7.5, 10, 12.5, 15, 17.5, 20, and 22.5 μM). (C) Determined affinities for the four types of interfaces. The affinities determined from the curve fit in panel B are compared with those derived from wild-type CA (Figure 5B). Each type of interface is highlighted with a thick red line.

polymerization has to go through the same initial steps of dimerization and trimerization as wild-type CA. To determine how the affinities of the four interfaces were affected by the double-cysteine mutation, the kinetic data were analyzed using our CA polymerization model in the same way wild-type CA was analyzed. Curve fitting was performed by assuming either that only monomers were present at time zero or that both

dx.doi.org/10.1021/bi300052h | *Biochemistry* 2012, 51, 4416−4428

monomers and dimers were present as an equilibrium mixture at time zero. The monomer/dimer equilibrium mixture assumption gave overall a slightly better curve fit (Figure 8A,B). While the presalt dimer interface affinity of 2Mut CA (26 μM) remained similar to that of wild-type CA (14 μM), the apparent postsalt dimer interface affinity of 2Mut CA increased by ~731.8-fold (Figure 8C and Table 1). The apparent affinities of the trimer, decamer, and polymer interfaces increased by 39.3-, 1.39-, and 1.22-fold, respectively (Figure 8C and Table 1). These changes translated into an increase in total polymerization affinity for 2Mut CA of (731.8 × 39.3 × 1.39 × 1.22) = 48770-fold over that of wild-type CA.

The four interface affinities determined for 2Mut CA are apparent. In particular, the postsalt K_{dimer} of 9 nM determined here is apparent because the intrinsic K_{dimer} should not change, since the S−S bond does not affect the dimer interface. The increase in the apparent interface affinities of 2Mut CA implies that during the course of polymerization, the steps were altered to gradually favor formation of hexagonal hexamers over that of trimers of dimers (i.e., triangular hexamers) once S−S cross-linking was established in the polymerized 2Mut CA lattice (Figure 9). Another way of looking at the drastic increase in apparent dimer interface affinity is that during the course of 2Mut CA polymerization, S−S cross-linking between neighboring monomers within the polymer effectively brought the off rate "constant" of the dimerization gradually to zero as the dimer entity is supplanted by the newly cross-linked hexagonal hexamer entity in the polymerized lattice, while the on rate constant of dimerization remained unchanged for the free monomers. The same argument can be made for the ~40-fold increase in trimer interface affinity. As for the apparent decamer and polymer interface affinities of 2Mut CA (i.e., 315 and 120 nM, respectively), it is interesting to observe that the capsomere polymerization affinity of 2Mut CA (0.3148 × 0.1200 = 0.0378) is roughly equal to the geometric mean capsomere polymerization affinities between wild-type and 4Mut CA [$(0.0642 \times 0.0154)^{0.5}$ = 0.0314] (Table 1). This latter observation is further evidence that 2Mut CA polymerization proceeded by a mixed mechanism starting with the formation of trimers of dimers (triangular hexamer) and assembly of the latter into a polymer lattice like wild-type CA but undergoing a transition to the assembly of hexagonal hexamers like 4Mut CA once S−S cross-linking is established (Figure 9).

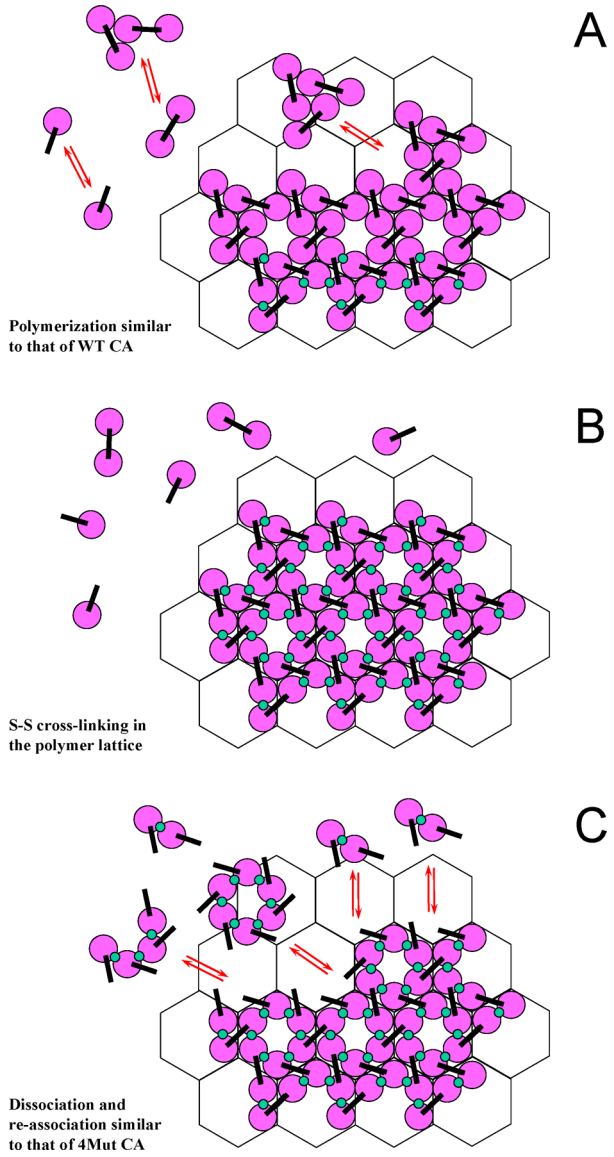


Figure 9. A14C/E45C CA polymerization proceeds with a pathway that is a hybrid between that of wild-type and cross-linked A14C/E45C/W184A/M185A CA. (A) A14C/E45C CA started polymerization in a manner similar to that of wild-type CA with the formation of dimers and trimers of dimers and the assembly of trimers of dimers into a polymer tube. (B) Once assembled into a lattice of trimers of dimers, cysteine 14 from one monomer is brought into the proximity of cysteine 45 of another monomer to allow disulfide bridge formation (green circles), leading to cross-linked hexamers in the polymerized lattice. (C) The cross-linked hexamers can now dissociate and reassociate in a manner similar to that of the cross-linked hexamers of A14C/E45C/W184A/M185A CA, but with a higher affinity because CTD−CTD interaction is preserved.

DISCUSSION

We have developed a mathematical model of in vitro CA polymerization. This model takes into account the diameter of the assembled polymer tubes and the stoichiometry of their component capsomeres (i.e., clusters of six monomer subunits) but does not depend on the specific geometric arrangement of the monomer subunits within each capsomere (Figure 1A,D). In fact, the cryo-EM structure of the CA polymer tube can be construed as a string of successive rings of 10 capsomeres each, stacked against each other. There are at least two ways of grouping six monomer subunits into nonoverlapping self-repeating capsomeres within the polymer tube lattice: the more visually obvious grouping into hexamers with a hexagonal geometry and the less noticeable grouping into trimers of dimers with a triangular geometry. There is no mathematical distinction between the two geometric arrangements in our model that is able to accommodate the kinetics of polymerization based on either geometric grouping of six monomer subunits. However, because the physical CA dimer does not reside within a hexagonal hexamer but is shared between neighboring hexamers, the triangular trimer of dimers is a more logical capsomeric building block of the polymer tube than the hexamer, and this was later confirmed by our experimental data.

In our experiments, CA proteins that existed as an equilibrium solution of monomers and dimers were induced to polymerize by dilution with a salt solution. Our model treats this in vitro CA polymerization as a process through which the system reestablishes a new equilibrium. In this equilibrium model, we have made two kinds of simplifications

dx.doi.org/10.1021/bi300052h | *Biochemistry* 2012, 51, 4416−4428

(i.e., condensation of multiple identical steps into one step and substitution of multiple polymerized species with their monomer components) to reduce the total number of reactants to only five. This approach allowed the definition of four types of interaction interfaces (i.e., dimer interface, trimer of dimers interface, interface between trimers of dimers within a decamer ring, and interface between trimers of dimers from two adjacent rings within a polymer tube) without having to specify in detail the many different pathways by which a monomer is incorporated into a polymer tube. It also allowed the determination of the dissociation constants of these four types of interfaces through data curve fitting. Our equilibrium model was able to generate a good fit of the wild-type CA polymerization time course that displays an initial lag phase followed by a sharper growth phase that eventually reached a plateau. The presence of a lag phase is often considered diagnostic of a nucleation-limited behavior,[30−33] and this effect can be generated by an equilibrium model as shown previously[21] and in our study. Even though our model does not explicitly incorporate a nucleation step, our model does display a critical CA concentration consistent with a nucleation-limited behavior (Figure 4C). A model of in vitro CA polymerization analogous to that of tubulin and actin[34−36] with an explicit nucleation step followed by elongation through addition of one subunit at a time to the nucleus can be conceived. However, the number of different interface affinities involved in the addition of each successive subunit becomes quite considerable. Using a single average affinity for all these interfaces may simplify the model but would not be very useful if the goal is to determine the affinities of the major interfaces, as is the case in this study. The affinities that are determined in this study from curve fitting with our model should not be taken as absolute, as they may be model-dependent. There are currently no other data in the literature that can be used for comparison. However, the values derived from our model are valid in that they reflect the relative affinities of interfaces within a polymer tube and between interfaces under various conditions (i.e., WT vs mutant CA), and the model might be able to quantify changes in interface affinity in the presence of small molecule inhibitors of CA and thus help in the elucidation of the mode of action of those molecules.

We determined by curve fitting that the dimerization step after the addition of salt has the slowest on rate constant (k_1 = 7×10^{-5} μM^{-1} min^{-1}). However, dimerization after the addition of salt may not be rate-limiting in practice because of the relatively large amount of dimers present in the presalt monomer−dimer equilibrium which has a dimer dissociation constant of 14 μM. For a CA concentration of 40 μM, this translates to ∼44.2% of the CA existing as dimers in the presalt equilibrium. The trimerization of dimers could be the practical rate-limiting step because it displays the next slowest on rate constant (k_2 = 0.55 μM^{-2} min^{-1}) but the highest affinity of the four interfaces (32 nM). The high affinity for the trimer of dimers interface may be necessary to compensate for the large entropic loss resulting from the formation of the trimers of dimers which could serve as building blocks for the growth of larger polymer species with more entropically favorable additions of subunits. An isolated structure of the trimer of dimers is shown in Figure 10. The much tighter trimer of dimers interface (Figure 10B, black curves) stands in contrast to the more spacious dimer interface (Figure 10B, blue thick line) and is consistent with their respective best fit affinities of 32 nM and 6.6 μM, respectively. The two remaining interfaces

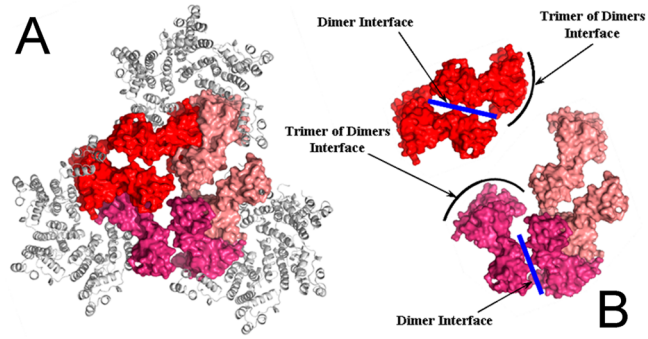


Figure 10. Structure of the trimer of dimers. (A) CTD view of the structure of a trimer of dimers in the context of three crystal structures of CA hexamer (gray, PDB entry 3H47). The three dimers in the trimer of dimers are each colored red, peach, or magenta. (B) View of the same trimer of dimers in which one dimer is taken apart. The black curve is for the trimer of dimers interface. The thick blue line is for the dimer interface.

are between two trimers of dimers within a ring (i.e., decamer interface of 438 nM) and between two trimers of dimers from adjacent rings within a polymer tube (i.e., polymer interface of 147 nM). We showed by simulation that it is not incidental that the polymer interface has a higher affinity than the decamer interface because this arrangement would allow the polymer tube to grow in length without trapping the trimers of dimers into decamer rings and short tubes (Figure 6).

Using the cross-linked A14C/E45C/W184A/M185A CA hexamers that are hexagonal, we were able to determine the affinities of the decamer interface (0.191 μM) and polymer interface (0.081 μM), giving a capsomere polymerization affinity per hexamer of (0.1905 × 0.0810) = 0.0154 μM (Table 1). If wild-type CA also polymerized as hexagonal hexamers, one would expect a similar capsomere polymerization affinity per hexamer. However, the total polymerization affinity per wild-type hexamer is (0.4384 × 0.1465) = 0.0642 μM, which is ∼4.2-fold weaker, suggesting that the capsomere building block for wild-type CA is not the hexagonal hexamer but the triangular trimer of dimers. This interpretation is also consistent with the high-affinity trimer of dimers interface determined for wild-type CA that is partially shared by the interfaces between artificially cross-linked hexamers. A second line of evidence arguing for trimers of dimers as capsomere building blocks of wild-type polymer tubes comes from the polymerization kinetics of A14C/E45C CA. This double-mutant CA is still capable of dimerization like wild-type CA but has the ability to form cross-linked hexagonal hexamers like the quadruple mutant. We showed that because of these properties, the double-mutant CA is able to begin polymerization like wild-type CA with formation of trimers of dimers but gradually transition to dissociation and reassociation of cross-linked hexagonal hexamers from the polymer tube (Figure 9). This behavior results in an apparent capsomere polymerization affinity per hexamer of (0.3148 × 0.1200) = 0.0378 μM, which is almost identical to the geometric mean capsomere polymerization affinity per hexamer of the wild type and the quadruple mutant.

The free energy of dissociation for every binding interface measured for 2Mut CA during its polymerization was larger than that of WT CA (Table 2). Because 2Mut CA is capable of S−S bond formation between monomers, the difference in the free energy of dissociation between 2Mut CA and WT CA must be contributed by S−S bond formation. Because the free energy

dx.doi.org/10.1021/bi300052h | *Biochemistry* 2012, 51, 4416−4428

of breaking an S−S bond is 60 kcal/mol, we calculated the number of S−S bond formations that can be ascribed to this free energy difference between 2Mut and WT CA for various species of oligomers and polymers of 2Mut CA (Table S2 of the Supporting Information, columns 7 and 8). We also calculated the total number of potential S−S bonds for each of the corresponding species of oligomers and polymers of 2Mut CA (Table S2 of the Supporting Information, column 9; Figure S2 of the Supporting Information). We found that the number of S−S bonds that can be ascribed to the difference in free energy between 2Mut and WT CA is ∼5.6% of the total number of potential S−S bonds in all the oligomer and polymer species (Figure S3 of the Supporting Information). Because the effect of S−S bond formation can be detected as a change in apparent dissociation constants only if various oligomer and polymer species having acquired S−S bonds after an initial association are actively dissociating and reassociating (Figure 9), our result implies that ∼5.6% of polymerized species are involved in active dissociation and reassociation at any point during the polymerization process (e.g., a S−S bond formed in the middle of long polymer tubes will not affect the kinetics of polymerization as those regions are less prone to dissociation). In other words, 94.4% of the polymerized structures never dissociate once they are polymerized.

Our finding that trimers of dimers are the capsomere building blocks of polymer tubes formed by in vitro CA polymerization has implications that extend to how closure of the viral capsid can be understood. It is known that closure of capsids requires the presence of pentamers in the polymer lattice, but it is conceptually unclear how pentamers can be induced to form instead of hexamers. Via examination of the arrangements of CA on a closed viral capsid model,[15] it became evident that the viral capsid is assembled with capsomere units made entirely of trimers of dimers (Figure 11). Instead of invoking the formation of pentamers for the closure of the capsid, the curvature necessary for capsid closure can be realized by local irregularities that resulted in the sharing of a single dimer between two trimers of dimers (Figure 11). This sharing of a dimer between two trimers of dimers ultimately resulted in what is perceived as pentamers in the viral capsid.

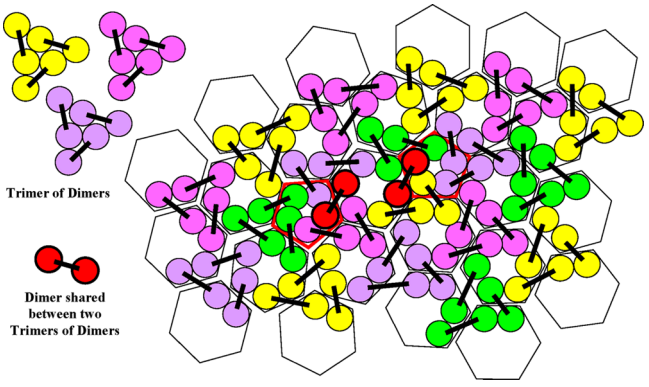


Figure 11. A "pentamer" can be viewed as an aberration in the polymer lattice caused by the sharing of a single dimer between two trimers of dimers. This diagram illustrates the generation of two pentamers (enclosed in red pentagons) through the sharing of a dimer (highlighted in red) between two adjacent trimers of dimers. Two such pairs of trimers of dimers are illustrated (magenta and purple trimers of dimers and yellow and green trimers of dimers). The presence of an odd number of pentamer configurations would necessitate continuous propagation of dimer sharing between successive pairs of trimers of dimers in the polymer lattice. An even number of pentamer configurations ends the propagation of dimer sharing between trimers of dimers in the lattice.

Finally, this mathematical model with the ability to quantify the affinity of four types of interfaces involved in CA polymerization in vitro can be used to assess how small molecule CA binding agents affect the apparent affinities of these interfaces during CA polymerization.

■ ASSOCIATED CONTENT

Supporting Information

Analysis of the continuous sedimentation coefficient distribution, $c(s)$, of CA proteins by sedimentation velocity (Figure S1), counting potential S−S bonds in 2Mut capsid oligomers and polymers (Figure S2), binding free energy contribution of S−S bond formation (Figure S3), wild-type CA polymerization kinetics shown via a detailed view of the early time point (Figure S4), sensitivity analysis of model determination of kinetic constants for WT CA (Figure S5), sensitivity analysis of model determination of kinetic constants for WT CA (Table S1), and numbers of S−S bonds calculated from the free energy difference between 2Mut and WT CA (Table S2). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*A.N.-M.: telephone, (650) 522-1591; fax, (650) 522-5166; e-mail, amajka@gilead.com. M.T.: telephone, (650) 522-5860; fax, (650) 522-5143; e-mail, mtsiang@gilead.com.

Author Contributions

M.T. and A.N.-M. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Prevelige, P. E., Jr. (2011) New approaches for antiviral targeting of HIV assembly. *J. Mol. Biol. 410*, 634−640.

(2) Blair, W. S., Pickford, C., Irving, S. L., Brown, D. G., Anderson, M., Bazin, R., Cao, J., Ciaramella, G., Isaacson, J., Jackson, L., Hunt, R., Kjerrstrom, A., Nieman, J. A., Patick, A. K., Perros, M., Scott, A. D., Whitby, K., Wu, H., and Butler, S. L. (2010) HIV capsid is a tractable target for small molecule therapeutic intervention. *PLoS Pathog. 6*, e1001220.

(3) Shi, J., Zhou, J., Shah, V. B., Aiken, C., and Whitby, K. (2011) Small-molecule inhibition of human immunodeficiency virus type 1 infection by virus capsid destabilization. *J. Virol. 85*, 542−549.

(4) Fader, L. D., Bethell, R., Bonneau, P., Bos, M., Bousquet, Y., Cordingley, M. G., Coulombe, R., Deroy, P., Faucher, A. M., Gagnon, A., Goudreau, N., Grand-Maître, C., Guse, I., Hucke, O., Kawai, S. H., Lacoste, J. E., Landry, S., Lemke, C. T., Malenfant, E., Mason, S., Morin, S., O'Meara, J., Simoneau, B., Titolo, S., and Yoakim, C. (2011) Discovery of a 1,5-dihydrobenzo[b][1,4]diazepine-2,4-dione series of inhibitors of HIV-1 capsid assembly. *Bioorg. Med. Chem. Lett. 21*, 398−404.

(5) Momany, C., Kovari, L. C., Prongay, A. J., Keller, W., Gitti, R. K., Lee, B. M., Gorbalenya, A. E., Tong, L., McClure, J., Ehrlich, L. S., Summers, M. F., Carter, C., and Rossmann, M. G. (1996) Crystal structure of dimeric HIV-1 capsid protein. *Nat. Struct. Biol. 3*, 763−770.

(6) Berthet-Colominas, C., Monaco, S., Novelli, A., Sibaï, G., Mallet, F., and Cusack, S. (1999) Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. *EMBO J. 18*, 1124−1136.

(7) Mortuza, G. B., Haire, L. F., Stevens, A., Smerdon, S. J., Stoye, J. P., and Taylor, I. A. (2004) High-resolution structure of a retroviral capsid hexameric amino-terminal domain. *Nature 431*, 481−485.
(8) Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T., and Sundquist, W. I. (1999) Assembly and analysis of conical models for the HIV-1 core. *Science 283*, 80−83.
(9) Ganser-Pornillos, B. K., von Schwedler, U. K., Stray, K. M., Aiken, C., and Sundquist, W. I. (2004) Assembly properties of the human immunodeficiency virus type 1 CA protein. *J. Virol. 78*, 2545−2552.
(10) Li, S., Hill, C. P., Sundquist, W. I., and Finch, J. T. (2000) Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature 407*, 409−413.
(11) Lanman, J., Lam, T. T., Barnes, S., Sakalian, M., Emmett, M. R., Marshall, A. G., and Prevelige, P. E., Jr. (2003) Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. *J. Mol. Biol. 325*, 759−772.
(12) Ganser-Pornillos, B. K., Cheng, A., and Yeager, M. (2007) Structure of full-length HIV-1 CA: A model for the mature capsid lattice. *Cell 131*, 70−79.
(13) Pornillos, O., Ganser-Pornillos, B. K., Kelly, B. N., Hua, Y., Whitby, F. G., Stout, C. D., Sundquist, W. I., Hill, C. P., and Yeager, M. (2009) X-ray structures of the hexameric building block of the HIV capsid. *Cell 137*, 1282−1292.
(14) Byeon, I.-J. L., Meng, X., Jung, J., Zhao, G., Yang, R., Ahn, J., Shi, J., Concel, J., Aiken, C., Zhang, P., and Gronenborn, A. M. (2009) Structural convergence between cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell 139*, 780−790.
(15) Pornillos, O., Ganser-Pornillos, B. K., and Yeager, M. (2011) Atomic-level modeling of the HIV capsid. *Nature 469*, 424−427.
(16) Reddy, V. S., Natarajan, P., Okerberg, B., Li, K., Damodaran, K. V., Morton, R. T., Brooks, C. L., III, and Johnson, J. E. (2001) Virus Particle Explorer (VIPER), a website for virus capsid structures and their computational analyses. *J. Virol. 75*, 11943−11947.
(17) Hagan, M. F., and Chandler, D. (2006) Dynamic pathways for viral capsid assembly. *Biophys. J. 91*, 42−54.
(18) Elrad, O. M., and Hagan, M. F. (2008) Mechanisms of size control and polymorphism in viral capsid assembly. *Nano Lett. 8*, 3850−3857.
(19) Krishna, V., Ayton, G. S., and Voth, G. A. (2010) Role of protein interactions in defining HIV-1 viral capsid shape and stability: A coarse-grained analysis. *Biophys. J. 98*, 18−26.
(20) Chen, B., and Tycko, R. (2011) Simulated self-assembly of the HIV-1 capsid: Protein shape and native contacts are sufficient for two-dimensional lattice formation. *Biophys. J. 100*, 3035−3044.
(21) Zlotnick, A. (1994) To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes. *J. Mol. Biol. 241*, 59−67.
(22) Zlotnick, A., Johnson, J. M., Wingfield, P. W., Stahl, S. J., and Endres, D. (1999) A theoretical model successfully identifies features of hepatitis B virus capsid assembly. *Biochemistry 38*, 14644−14652.
(23) Endres, D., and Zlotnick, A. (2002) Model-based analysis of assembly kinetics for virus capsids or other spherical polymers. *Biophys. J. 83*, 1217−1230.
(24) Casini, G. L., Graham, D., Heine, D., Garcea, R. L., and Wu, D. T. (2004) In vitro papillomavirus capsid assembly analyzed by light scattering. *Virology 325*, 320−327.
(25) Morozov, A. Y., Bruinsma, R. F., and Rudnick, J. (2009) Assembly of viruses and the pseudo-law of mass action. *J. Chem. Phys. 131*, 155101(1)−155101(17).
(26) Hagan, M. F., and Elrad, O. M. (2010) Understanding the concentration dependence of viral capsid assembly kinetics: The origin of the lag time and identifying the critical nucleus size. *Biophys. J. 98*, 1065−1074.
(27) Moisant, P., Neeman, H., and Zlotnick, A. (2010) Exploring the paths of (virus) assembly. *Biophys. J. 99*, 1350−1357.
(28) Lanman, J., Sexton, J., Sakalian, M., and Prevelige, P. E., Jr. (2002) Kinetic analysis of the role of intersubunit interactions in human immunodeficiency virus type 1 capsid protein assembly in vitro. *J. Virol. 76*, 6900−6908.
(29) Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Worthylake, D. K., Wang, H., McCutcheon, J. P., Sundquist, W. I., and Hill, C. P. (1997) Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science 278*, 849−853.
(30) Oosawa, F., and Asakura, S. (1975) *Thermodynamics of the polymerization of protein*, Academic Press, New York.
(31) Prevelige, P. E., Jr. (1998) Inhibiting virus-capsid assembly by altering the polymerization pathway. *Trends Biotechnol. 16*, 61−65.
(32) Prevelige, P. E., Jr., Thomas, D., and King, J. (1993) Nucleation and growth phases in the polymerization of coat and scaffolding subunits into icosahedral procapsid shells. *Biophys. J. 64*, 824−835.
(33) Hyun, J.-K., Radjainia, M., Kingston, R. L., and Mitra, A. K. (2010) Proton-driven assembly of the Rous sarcoma virus capsid protein results in the formation of icosahedral particles. *J. Biol. Chem. 285*, 15056−15064.
(34) Johnson, K. A., and Borisy, G. G. (1977) Kinetic analysis of microtubule self-assembly in vitro. *J. Mol. Biol. 117*, 1−31.
(35) Tobacman, L. S., and Korn, E. D. (1983) The kinetics of actin nucleation and polymerization. *J. Biol. Chem. 258*, 3207−3214.
(36) Cooper, J. A., Buhle, E. L., Jr., Walker, S. B., Tsong, T. Y., and Pollard, T. D. (1983) Kinetic evidence for a monomer activation step in actin polymerization. *Biochemistry 22*, 2193−202.