Supporting Information for "The impact of conformational fluctuations on self-assembly: Cooperative aggregation of archaeal chaperonin proteins"

Stephen Whitelam[1,*], Carl Rogers[2], Andrea Pasqua[3], Chad Paavola[4], Jonathan Trent[4], Phillip L. Geissler[3,5,†]
[1] *Systems Biology Centre, University of Warwick, Coventry CV4 7AL, UK*
[2] *Biophysics Graduate Group, University of California at Berkeley, Berkeley, CA 94720*
[3] *Department of Chemistry, University of California at Berkeley, Berkeley, CA 94720*
[4] *Bioengineering Branch, NASA Ames Research Center, Mail Stop 239-15, Moffett Field, CA, 94035*
[5] *Materials Sciences, Physical Biosciences, and Chemical Sciences Divisions, Lawrence Berkeley National Laboratory, Berkeley, CA 9472*

Materials and Methods

1. *Rosettasome purification* The gene for the wild-type beta subunit from S. shibatae was previously cloned and sequenced [1]. Proteins were expressed from plasmids introduced into E. coli BL21CodonPlus(DE3)-RIL competent cells (Stratagene) and the recombinant proteins were purified by heat treatment (75-60 Celsius, 30 min), and ion exchange chromatography using a Mono-Q column on FPLC (Pharmacia) with a linear gradient from 0 to 400 mM NaCl in 25mM HEPES at pH 7.5. Protein fractions were pooled and concentrated using Centriprep YM30 centrifugal concentrators (Millipore) and buffer was exchanged three times with 25mM HEPES at pH 7.5 in the centrifugal concentrator.

2. *Electron microscopy* Protein samples were attached to ultra-thin carbon-coated formvar substrates (Ladd Research Industries) by floating substrates on protein solutions for 2 minutes, wicking away excess with filter paper. Samples were stained immediately with 0.22 um filtered 2% uranyl acetate for 3 min, rinsed with water and air dried at room temperature. The grids were viewed using a LEO 912 AB transmission electron mictoscope with a tungsten filament at 60kV. Images were recorded with a MegaView digital camera using ANALYSIS 3.5 software.

Model Parameters

We may set an upper bound on the angular tolerance (i.e. inverse specificity) of the polar interaction in our simulations, $\sigma_{\rm pol}$, by noting that bundled chaperonin strings in experiments appear to grow in register, via the addition of monomers to each string tip, rather than via the recruitment of monomers shared between neighboring strings. In simulations, the angular component of the polar energy of interaction between two aligned model chaperonins [i.e., $\phi = 0$, see Figure 1 (main text)] is (Equation (3))

$$v_{\rm pol}(\theta) = -\epsilon_{\rm pol} \exp\left(-2\sigma_{\rm pol}^{-2}[\cos\theta - 1]^2\right), \tag{1}$$

where θ is the (small) angle between the orientation directors of the units and the vector joining their centers. To set an upper bound on $\sigma_{\rm pol}$ we require that in simulations a bundle of two perfectly-aligned strings finds it energetically favorable to recruit monomers preferentially to the tips of each of its two filaments, rather than by sharing monomers between filaments. We therefore require that $v_{\rm pol}(0) < 2v_{\rm pol}(\pi/6)$, or

$$\sigma_{\rm pol}^2 < \frac{2}{\ln 2}\left(\cos(\pi/6) - 1\right)^2, \tag{2}$$

giving $\sigma_{\rm pol} \lesssim 0.23$. We choose $\sigma_{\rm pol} = 0.12$ to strongly disfavor off-register growth of bundled strings.

Given the specificity of the polar interaction we can put an approximate lower bound on its strength $\epsilon_{\rm pol}$ by requiring that in thermal equilibrium we observe chains of (at least) hundreds of units (assuming a sufficiently large system). Chaperonin strings of many hundreds of units are observed in experiment [2]. Strings are quasi-one-dimensional structures (each particle in the bulk binds to only two neighbors) and so a numerically large binding energy is

* Current address: Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
† Author to whom correspondence should be addressed.
E-mail: geissler@berkeley.edu

required to offset the unfavorable entropy of polymerization that results from the growth of long strings. We can estimate how large this binding energy must be using the Wertheim thermodynamic perturbation theory [3] outlined in Ref. [4]. This theory provides an expression for the free energy of associating particles possessing two identical binding sites (such as an identically sticky north and south pole, as here), and so may be used to calculate the extent of string formation for a collection of model chaperonins. From the results of Ref. [4] we find that chaperonin string-formers of number density $\rho = N/V$ form strings whose equilibrium length distribution $\rho(l)$ is

$$\rho(l) = \rho \ell_{\rm eq}^{-2} \left(1 - \ell_{\rm eq}^{-1}\right)^{l-1}. \tag{3}$$

Here $\ell_{\rm eq}$ is the average string length in equilibrium, and satisfies

$$2\ell_{\rm eq} = 1 + \sqrt{1 + 8\rho\Delta}. \tag{4}$$

The parameter Δ is related to the second virial coefficient of our model system. It is calculated by integrating the Mayer f function for two chaperonin units over all angles and distances for which they interact:

$$\Delta = 4\pi \int d\boldsymbol{\omega} dr\, r^2\, g_{\rm HS}(r) \langle e^{-\beta U(\boldsymbol{r})} - 1 \rangle_{\boldsymbol{\omega}}. \tag{5}$$

In this expression $\boldsymbol{r}$ is the vector joining the two units, and r is its magnitude; $\boldsymbol{\omega}$ represents two sets of angles (we assume one chaperonin to be fixed at the origin with orientation along $\hat{\boldsymbol{z}}$); $g_{\rm HS}(r)$ is a (reference) pair correlation function for hard spheres (for this we use the approximation of Ref. [4]); $\langle\cdot\rangle_{\boldsymbol{\omega}}$ represents an average over the angles $\boldsymbol{\omega}$; and U is the energy of interaction of the two units. We plot in Figure 1 the average string length in equilibrium (for two chaperonin concentrations) and the string length distribution (for two attraction strengths), assuming $\sigma_{\rm pol} = 0.12$. We find that $\epsilon_{\rm pol}$ must exceed roughly 15 $k_{\rm B}T$ (to the nearest 5 $k_{\rm B}T$) in order that chains formed in simulations may approach hundreds of units in length. In the main text we take $\epsilon_{\rm pol} = 16\, k_{\rm B}T$ for stringon units.

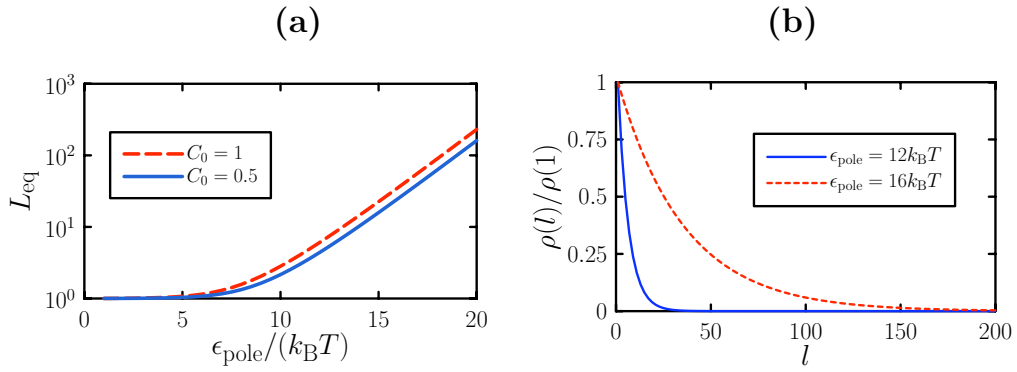

FIG. 1: Predictions of Wertheim thermodynamic perturbation theory [3, 4] applied to a collection of our model stringons (with inverse specificity parameter $\sigma_{\rm pol} = 0.12$ and no equatorial coupling). (a) Average string length $\ell_{\rm eq}$ in equilibrium as a function of $\epsilon_{\rm pol}$ for two concentrations; C_0 is the concentration corresponding to 1500 units in a cubic box of length $64a$, as considered in the main text. (b) String length l distribution $\rho(l)$ for two binding strengths. Our conclusion from these estimates is that $\epsilon_{\rm pol}$ must exceed roughly 15 $k_{\rm B}T$ in order that the longer strings in simulations approach hundreds of units in length (as is seen in experiment).

With polar strength and specificity determined by equilibrium considerations, we find that nucleation and growth of filaments proceeds rapidly. Sheet formation, by contrast, is relatively sluggish: the requirement that self-assembly of sheets proceed without the development of substantial kinetic traps imposes stringent limits on the strength of equatorial interactions (given their specificity) [5]. Based on our previous work we choose for planons the parameters $\sigma_{\rm eq} = 0.3$ and $\epsilon_{\rm eq} = 6.5 k_{\rm B}T$ to ensure that planons self-assemble into sheets with a low density of defects even at the largest concentrations we consider.

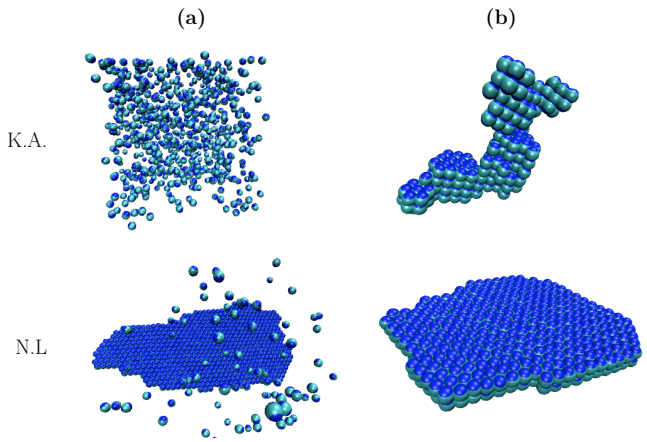

FIG. 2: Large regions of the phase diagram of Figure 2 (main text) are dominated by nonequilibrium effects. (a) At state $(\epsilon_{\rm eq}, \epsilon_{\rm pol}) = (6k_BT, 0k_BT)$, kinetically accessible (K.A.) configurations, i.e., those generated by an algorithm designed to mimic a natural dynamics over experimental time scales, primarily comprise unassociated units. However, nonlocal (N.L.) moves (similar to those employed in [6]) reveal that close-packed sheets are the thermodynamically preferred state. (b) Kinetically frustrated aggregates at state $(\epsilon_{\rm eq}, \epsilon_{\rm pol}) = (7, 8)\, k_{\rm B}T$ are replaced by structures with few defects and with a very different form under nonlocal moves. These examples indicate the considerable difference between what is kinetically accessible and what is thermodynamically preferred in this example of self-assembly.

Kinetics versus thermodynamics for fixed-conformation self-assembly

In Figure 2 we contrast the products of aggregation *dynamics* seen in Figure 2 of the main text with configurations typical of true thermal equilibrium (which might be reached dynamically on time scales too long to witness even in the laboratory). Equilibrium structures were obtained using a Monte Carlo algorithm that samples the Boltzmann distribution much more efficiently than does the system's natural dynamics [6]. In these simulations a nonlocal move set dramatically accelerated the annealing of defects [7, 8] by shuttling particles between the interaction regions of other particles.

General considerations of stringon versus planon self-assembly

Molecular structures extending in one dimension abound in the natural world, from DNA to amyloid fibrils to microtubules. Those involving noncovalently linked proteins in their native states are much less common, primarily fulfilling cytoskeletal duties *in vivo*. Nonetheless, the basic requirements for one-dimensional growth of nanoparticle chains would appear to be mild. Any freely-diffusing species with (at least) bipolar attractions should be capable of forming such structures readily. While there is no thermodynamic phase transition associated with the development of filamentous structure, there is for any binding strength a concentration of unassociated monomers above which the average growth rate of a filament is positive. Furthermore, the statistical independence of each link in a stiff chain ensures that little can go wrong with this mode of assembly. These considerations suggest that filaments form readily under most conditions for which they are thermodynamically stable.

The second dimension of a planar structure can greatly enrich the physical underpinnings of assembly and complicate its requirements. Interactions between the neighbors of a given particle introduce the possibility of defective growth. The imperfect binding of a particle to a cluster, e.g., at a skewed angle, is inevitable in the course of assembly. If such defects are not annealed before further binding events constrain local rearrangements, assembled structures become highly disordered. Thermodynamically, the extended boundary of a sheet, in contrast to the negligible boundary of a filament, allows for nontrivial effects of interfacial tension. As a cluster grows, so too does the unfavorable free energy associated with its edges. Until its interior is large enough to offset this cost, line tension can strongly inhibit growth, giving rise to free energy barriers.

In Figure 3(a) we plot the free energy $G(n)$ of a small sheet as a function of the number n of its constituent planon units, for several concentrations of free planons with no polar couplings. Each of these curves, calculated using umbrella sampling techniques, features a barrier $\Delta G^{\ddagger}$ to nucleation. In Figure 3(b) we show the decay of this barrier

height with increasing planon concentration. In a mixture of stringons and planons, the initial state of the trajectory possesses a free energy barrier to sheet formation on the order of thermal fluctuations. As units are rapidly sequestered into filamentous structures, which possess no barrier to assembly, the free energy barrier to sheet formation increases, making the nucleation of dense sheet structures an increasingly rare event.

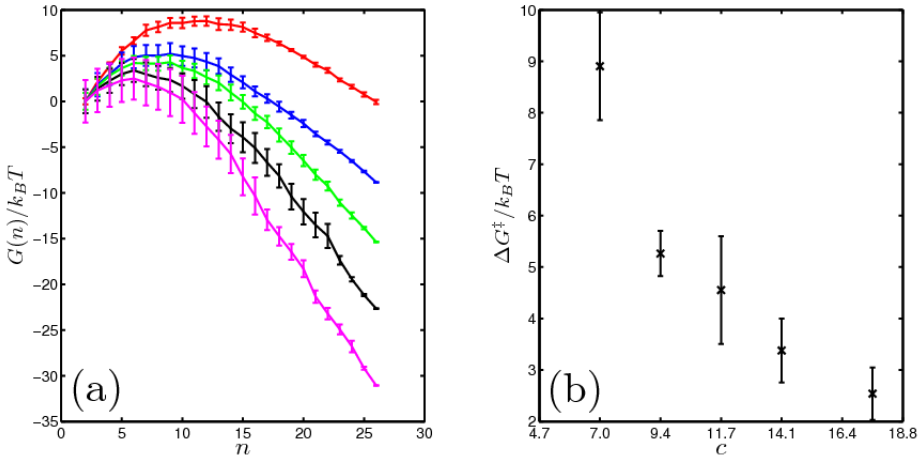

FIG. 3: (a) Free energy of a system of planons that comprise a single sheet-like cluster of n units and $N-n$ unassociated units. Results are shown for several concentrations $c = \frac{N}{V}$, with volume $V = (64a)^3$: $c = 7$ mg/ml (red), $c = 9.4$ mg/ml (blue), $c = 11.7$ mg/ml (green), $c = 14.1$ mg/ml (black), and $c = 17.6$ mg/ml (pink). (b) Barrier height as a function of planon concentration c in mg/ml.

Time evolution of schizophrenic assembly

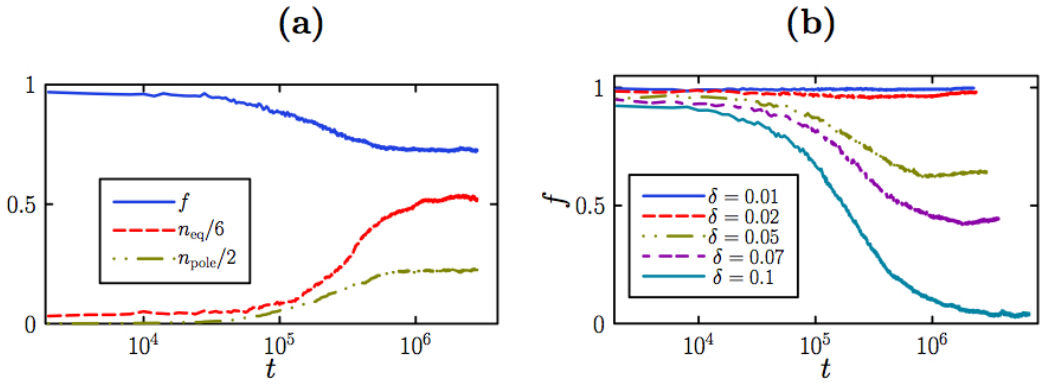

FIG. 4: Assembly kinetics for interconverting units with vanishing minor couplings ($\epsilon_{eq}^{(min)} = 0 = \epsilon_{pol}^{(min)}$). (a) Time-dependence of the fraction of planon units f and the number of optimal equatorial (polar) contacts per particle n_{eq} (n_{pol}) for the trajectory shown in Figure 4 (with $\delta = 0.05$, main text). Rapid string formation lowers the planon-to-stringon ratio so that both strings and sheets proliferate. (b) Time dependence of the fraction of planon units f for single trajectories with different values of the internal bias δ.

In Figure 4 (a) we quantify the kinetics of dichotomous assembly by plotting as a function of time the fraction of planon units for a trajectory possessing the same parameter values as that shown in Figure 4 (main text). We plot also the number of polar and equatorial contacts per particle. The celerity of string growth with respect to sheet nucleation in general causes f to be noticeably diminished from its ideal value of $(1+\delta)^{-1}$. Steady increases in both n_{eq} and n_{pol} at intermediate times reflect the simultaneous growth of corresponding structures. In Figure 4 (b) we

show the time dependence of f for trajectories obtained with several different values of the bias δ. String formation suppresses sheet nucleation almost completely for even modest δ.

In Figure 5 we present the analog of Figure 4 for units with considerable minor couplings. While the rapidity of planon to stringon conversion is influenced by the inter-species interactions, the fundamental kinetic bias in favor of one-dimensional stringons is not.

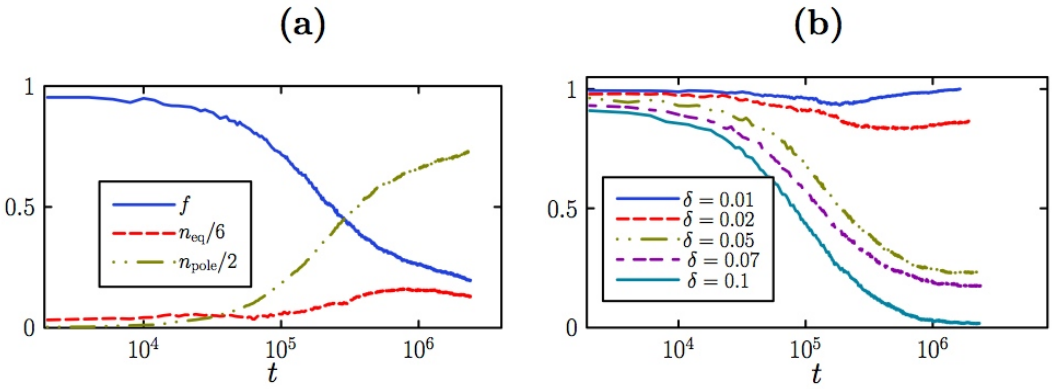

FIG. 5: Analog of Figure 4 for units with considerable minor couplings ($\epsilon_{\rm eq}^{(\rm min)} = 4k_{\rm B}T, \epsilon_{\rm pol}^{(\rm min)} = 5k_{\rm B}T$). (a) Time-dependence of the fraction of planon units f and the number of equatorial (polar) contacts per particle $n_{\rm eq}$ ($n_{\rm pol}$) normalized with reference to optimal value $n_{\rm eq} = 6$ ($n_{\rm pol} = 2$). (b) Time dependence of the fraction of planon units f for single trajectories with different values of the internal energy bias δ. Inter-species interactions modify quantitatively but not qualitatively the bias toward string formation observed in the case of vanishing minor couplings.

Composite structures can arise from planon-stringon interactions

Units' conformational flexibilities can induce interactions between superstructures of different types. The interconversion of units at the surfaces of aggregates can fuse different structures together, or promote nucleation of one kind of structure at the surface of another. We show an example product of sheet-string fusion in Figure 6. In contrast to hybrid structures that proliferate when unit conformations are fixed, different elements of these assemblies tend to maintain their essential shapes even when fused. Figure 7 shows a TEM image of a similar type of composite structure, a sheet with 'hairs' protruding from its edges. When units are free to undergo conformational change in our simulations, the superstructures they form are qualitatively similar to those seen in experiments.

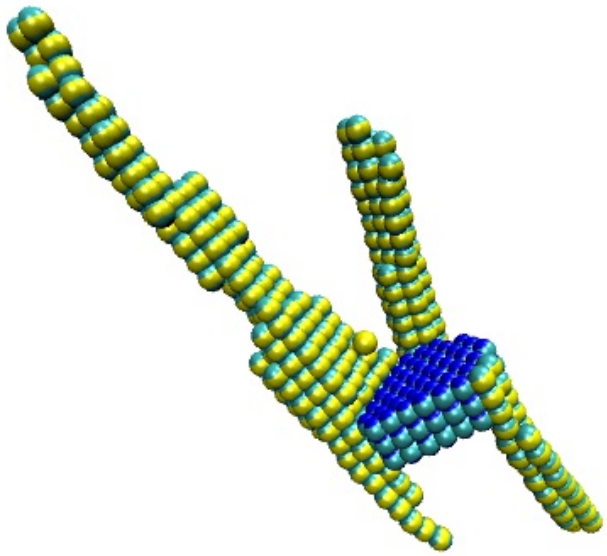

FIG. 6: Structure obtained from a long trajectory of units possessing considerable minor interactions ($\epsilon_{\rm eq}^{(\rm min)} = 4k_{\rm B}T, \epsilon_{\rm pol}^{(\rm min)} = 5k_{\rm B}T$). The interconversion of units near the periphery of sheets and strings permits the fusing together of distinct superstructure types.

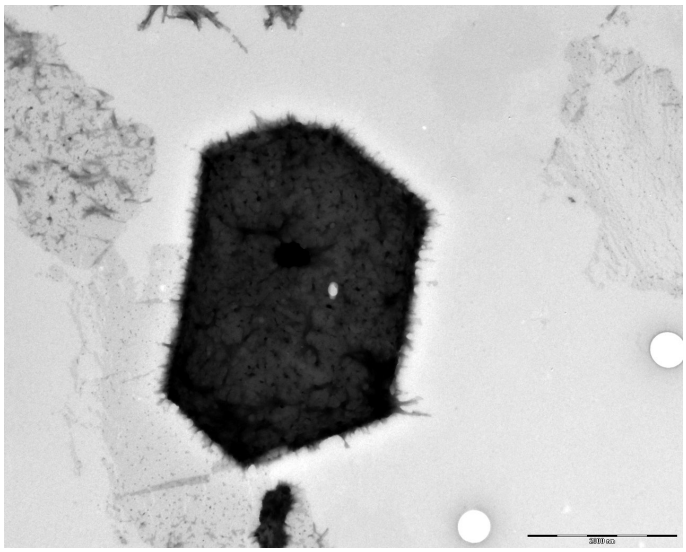

FIG. 7: TEM image of a large-scale composite structure. Sheet-like crystalline order persists over several μm. Boundaries of this assembly, however, are decorated with dangling filaments.

[1] J. Trent, E. Nimmesgern, J. Wall, F. Hartl, and A. Horwich, Nature **354**, 490 (1991).
[2] J. Trent, H. Kagawa, T. Yaoi, E. Olle, and N. Zaluzec, Proceedings of the National Academy of Sciences **94**, 5383 (1997).
[3] M. Wertheim, Journal of Statistical Physics **35**, 19 (1984).
[4] F. Sciortino, E. Bianchi, J. Douglas, and P. Tartaglia, The Journal of Chemical Physics **126**, 194903 (2007).
[5] S. Whitelam and P. Geissler, The Journal of Chemical Physics **127**, 154101 (2007).
[6] B. Chen and J. Siepmann, J. Phys. Chem. B **105**, 11275 (2001).
[7] R. Jack, M. Hagan, and D. Chandler, Physical Review E **76**, 21119 (2007).
[8] D. Rapaport, Arxiv preprint arXiv:0803.0115 (2008).