Supporting information for Kentsis *et al.* (January 15, 2002) *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.012317299.

Supporting Methods and Analysis

Analytical Ultracentrifugation. For both SV and sedimentation equilibrium, sample absorbance was measured at 280 nm for 100 μM samples and at 230 nm for dilute preparations where the XL-I D_2 lamp has its flux maximum. Step-size and point averaging were varied to achieve data sampling optimal for the speed of sedimenting species, with starting values of 0.005 cm per step and four points per step, respectively. Data were never collected beyond the 7.18-cm point of the cell to avoid optical artifacts associated with edge scattering. Data sets were examined for incompleteness (sedimenting boundaries that cleared the meniscus before initial scan), optical artifacts (boundary discontinuities caused by window scratches), and velocity instability (variable time interval between scans). Qualified data sets were analyzed using the direct boundary fitting method of Fujita and MacCosham and its approximation of the Lamm equation, as improved by Philo and implemented in SVEDBERG 6.39 (1). Sequential scans were subtracted from each other to minimize systematic noise present in the baseline (2), and the pair-subtracted scans were fit independently to both sedimentation and diffusion, because initial values for both coefficients could be determined from dynamic light-scattering and gel-filtration measurements. The number of sedimenting species to be fit were determined from gel-filtration measurements and visual inspection. Data sets were fit without significant parameter cross-correlations ($r < 0.56$) and with residuals that were normally distributed (root mean square errors < 0.01, degrees of freedom > 2,600). Using this approach, it was routinely possible to obtain reliable boundary plateau concentrations for the sedimenting species as well as distributions of their sedimentation and diffusion coefficients. Because scans acquired late in the run have broad boundaries, artificial boundary spreading late in the run was checked using the van Holde--Weischet time extrapolation method (3), as implemented by Demeler in ULTRASCAN 4.1 (4). Sedimentation coefficients were corrected to standard temperature and pressure according to Tanford:

$$s_{20,w} = s_{T,b}\left(\frac{\eta_{T,b}}{\eta_{20,w}}\right)\left(\frac{1-\nu\rho_{20,w}}{1-\nu\rho_{T,b}}\right)$$

where η, ν, and ρ are viscosity, partial specific volume, and density, respectively (5). Molecular weights were determined according to the Svedberg relation:

$$M = \frac{sRT}{D(1-\nu\rho)}$$

where D and R are diffusion and gas constants, respectively (6). Because determination of molecular weight by SV depends on shape, rugosity, and hydration as reflected in the SV coefficient, it biases molecular weight determination toward molecules that are spherical, smooth, and unhydrated. Z bodies are spherical as observed by EM (Fig. 1*c*) and do not exhibit unusual hydrodynamics as observed by gel filtration (Fig. 2*c*). By

using Philo's improvements of the approximate solutions of the Lamm equation, it is possible to estimate molecular weights to within ±2% (1). This is indeed the case as verified independently by SV and gel filtration (Fig. 2 *c* and *d*, supporting Table 2).

Assembly on similar time scales as sedimentation results in transport, where a fast-moving sedimentation boundary that contains high molecular weight bodies is repopulated by freshly assembled species that are delayed in their sedimentation, thereby leading to its broadening. This is represented by the disagreement between sedimentation and diffusion coefficients (4), as reflected in the direct boundary fitting method (1). Alternatively, diffusion coefficients that are disproportionately small compared to sedimentation coefficients could be an artifact of aggregation, where a polydisperse population contains heterogeneously aggregated species with a range of diffusion and sedimentation coefficients (supporting Table 2). Discrimination between these two possible causes of disproportionately broad boundaries, therefore, must come from other measurements, which in our case are provided by EM and gel filtration.

Dynamic Light Scattering. Measured intensity autocorrelation spectra were converted to field correlation spectra and analyzed by single-value decomposition (7), as implemented in DYNALS 1.35 (Protein Solutions, Lakewood, NJ). This allows accurate measurements of sample polydispersity with three scattering species and reliable diffusion coefficient determination of low polydispersity samples. Because the amplitude of each phase of the correlation spectrum is a convolution of both concentration and molecular weight, only qualitative assessments of relative sizes of each population can be made. For steady-state measurements, data were collected in 10-s intervals, rejecting correlation functions with signal/noise < 10, and box-averaged. For kinetic measurements, data were collected in 5-s intervals with a signal/noise threshold of 3. In both cases, the sum of squares error was less than 9, with usual values around 6.

Thermodynamics Methods. For all thermodynamic measurements, samples were equilibrated to their solution conditions for 36 h by dialysis. Because SV data contains information about relative concentrations of each species from plateau absorbance values as well about the order of the assembly from the species' sedimentation coefficients, it is possible to calculate equilibrium constants and changes in free energy for each step of the assembly pathway according to Stafford's formalism:

$$K_i = \frac{\alpha_i}{c_i^{i-1}(1-\alpha_i)^i}; \quad \alpha_i = \frac{s_w - s_{i-1}}{s_i - s_{i-1}}$$

where i is the order of the assembly, c_i is the plateau concentration, α_i is the weight fraction of the ith polymer, s_w is the weight average sedimentation coefficient, and s_i and s_{i-1} are the sedimentation coefficients for the ith and $i-1$th assembly species (8). Stafford's formalism is analogous to the representation of assembly equilibria in terms of molar fractions (Eqs. **1** and **2**), replacing f_b, $f_,$, and f_m with c_b, c_t, and c_m, respectively, and accounting for the order of assembly using sedimentation coefficients (supporting Table 2), which are measures of molecular weight. The validity of application of the SV

boundary plateau concentration analysis depends on two assumptions: requirements that assembly be reversible and that its time scale be well separated from the sedimentation time. Reversibility of assembly was judged from the lack of hysteresis during guanidine-induced assembly and disassembly (Fig. 3*a*). Assembly kinetics were measured directly by rapidly diluting unfolded and monomeric Z into low guanidine concentrations and monitoring its equilibration to the native condition. These measurements show that Z requires more than 15 h to assemble (Fig. 3*c*), much longer than the 30 min required for sedimentation of bodies (Fig. 2*d*).

Thermodynamics Analysis. Guanidine denaturation profiles were fit to the following expression (9), describing the dependence of observed signal of the fully assembled body on the total and oligomer protein concentrations, [total] and [oligomer], degree of assembly of the oligomer, n, and linear dependence, m^o, of the change in Gibbs free energy ΔG^o on guanidine hydrochloride concentration [GdmHCl]:

$$S_{obs} = \frac{S_{oligo}[\text{oligomer}]}{[\text{total}]} + \frac{S_{body} n[\text{oligomer}]^n}{\exp\left(\frac{-\Delta G^o - m^o[\text{GdmHCl}]}{\text{R}T}\right)[\text{total}]}$$

Thus, discrimination between hierarchical and concerted thermodynamic models of assembly was accomplished by comparing the congruence of fit between models of guanidine denaturation and ultracentrifugation data due to the information degeneracy of SV measurements, insofar as observed plateau concentrations and sedimentation coefficients can be recovered from a variety of models depending on fit parameters. Success of discrimination, therefore, depends on the relative instability of the tetrameric intermediate, allowing a direct measurement of the equilibrium between unfolded monomers and assembled bodies (Eq. **1**), which sets the upper energetic limit for the assembly process. Thus, comparing the change in free energy in assembling unfolded monomers into assembled bodies, ΔG°_{ub}, as derived from guanidine denaturation, with the energetics of assembly of monomers into bodies, ΔG°_{mb}, as derived from SV boundary plateau concentration analysis for the hierarchical and concerted assembly models allows their discrimination.

For van't Hoff analysis, samples of Z were equilibrated to temperatures between 4 and 40°C, and their partitioning among the assembly forms was followed by plateau concentration analysis of SV data. This allows the calculation of equilibrium constants of assembly as a function of temperature. Thermodynamic parameters of assembly were calculated, as described by

$$\Delta H^o = -\text{R}\frac{d(\ln K)}{d(1/T)}$$

where the slope of the function is ΔH° and its y intercept is ΔS°, with deviations from linearity being caused by changes in heat capacity. In this form, van't Hoff analysis of the

thermodynamics of assembly may underestimate the contribution of changes in heat capacity, thereby overestimating the entropic component of temperature-driven assembly. Approximation of such an error suggests that potential changes in heat capacity for the tetramer-body step do not exceed 0.13 ± 0.24 kcal/mol/K, as ascertained from the coefficient of the quadratic term in the fit, which represents the second order dependence of K on temperature, and is equivalent to the first order dependence, or change in heat capacity, of ΔH. CD measurements indicate that temperature titration in this range does not lead to unfolding (Fig. 2*a*). Completeness of temperature equilibration was ascertained by comparing them with a 40°C sample equilibrated for 2 weeks.

Kinetics Methods. For kinetic measurements of assembly, fully unfolded and monomeric Z or native and fully assembled Z was diluted manually into buffer of appropriate guanidine concentration. Equilibration to denaturing and native conditions was verified by SV and CD, and assembly and disassembly kinetics were monitored by SV, CD, EM, and dynamic light scattering. Dead-time due to dilution, mixing, and temperature equilibration was typically less than 30 s and monitored using a thermocouple and UV absorbance to monitor temperature and protein concentration, respectively. The relatively long data collection times of SV made it an unsuitable method for kinetic studies. The four kinetic probes of SV, CD, EM, and dynamic light scattering, which report on sedimentation coefficient, mean secondary structure content, microscopic size and shape, and mean diffusion coefficient, respectively, were largely coincident (Fig. 3*c*, supporting Fig. 4*e*, and data not shown). Dependence of observed assembly and disassembly rates on protein concentration revealed that assembly rates scale linearly with protein concentration (supporting Fig. 4*d*), indicating a bimolecular rate-limiting step for the observed assembly reaction. Disassembly rates were insensitive to changes in protein concentration in the 50–300 μM range (supporting Fig. 4*d*), which is consistent with monitoring of unimolecular disassembly of bodies into tetramers and monomers that dissociate and unfold on the time scale of seconds. Thus, kinetics were fit either to exponential decays in the case of disassembly and unfolding or to a bimolecular model describing a first-order reaction in which the monomer-tetramer transition reaches fast preequilibrium within the experimental dead-time.

Kinetics Analysis. The exponential form of both assembly and disassembly kinetics suggests that assembly of tetramers into bodies and disassembly of bodies involve a single barrier-crossing process and a unique transition state. As discussed in the text, we characterized the kinetics of assembly using Hammond analysis. Using guanidine as a perturbant provides information on the relative degree of desolvation in the transition state as well as its energetic position relative to reactants and products, i.e., tetramers and bodies. To apply this analysis to self-assembly as it has been applied to simple chemical reactions and protein folding (10, 11), the uniqueness of the barrier for forward and reverse processes was validated by recovering the equilibrium free energy of assembly from the activation free energies of assembly and disassembly. These analyses treat monomeric folding and assembly of tetramers from monomers as fast preequilibria (thus the tetramer-to-body transition is examined) and shows the dependence of the observed rates on guanidine concentration, with the left limb of the chevron showing the decrease

in the assembly rate with increasing guanidine concentrations and the right limb showing the increase in the disassembly rate with increasing guanidine concentrations (Fig. 3*d*):

$$\Delta G^{\ddagger}_{tb} = -\mathrm{R}T \ln 2k_{tb} - m_{tb}[GdmHCl] + \text{constant}$$
$$\Delta G^{\ddagger}_{bt} = -\mathrm{R}T \ln k_{bt} - m_{bt}[GdmHCl] + \text{constant}$$

where k_{tb} and k_{bt} are rates of assembly and disassembly, respectively. Note the bimolecular coefficient in $\Delta G^{\ddagger}_{tb}$. Values of the slopes m_{tb} and m_{bt} of the guanidine dependence are proportional to the surface burial for the assembly tetramer-transition state and disassembly body-transition state steps, respectively. Extrapolation of the two limbs to 0 M guanidine yields values for activation energies that are comparable directly with those derived from equilibrium methods. The free energies of activation were calculated according to Eyring

$$\Delta G^{\ddagger} = \mathrm{R}T(\kappa \mathrm{k_B}T / \mathrm{h} - \ln k)$$

where the transmission factor κ has been set to unity (12).

1. Philo, J. S. (1997) *Biophys. J.* **72,** 435–444.
2. Philo, J. S. (2000) *Anal. Biochem.* **279,** 151–163.
3. van Holde, K. E. & Weischet, W. O. (1978) *Biopolymers* **17,** 1387–1403.
4. Demeler, B., Saber, H. & Hansen, J. C. (1997) *Biophys. J.* **72,** 397–407.
5. Tanford, C. (1961) *Physical Chemistry of Macromolecules* (Wiley, New York).
6. Svedberg, T. & Pedersen, K. O. (1940) *The Ultracentrifuge* (Oxford Univ. Press, London).
7. Pecora, R. (1985) *Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy* (Plenum, New York).
8. Stafford, W. F. I. (1994) in *Modern Analytical Ultracentrifugation*, eds. Schuster, T. M. & Laue, T. M. (Birkhauser, Boston), pp. 119-137.
9. Santoro, M. M. & Bolen, D. W. (1992) *Biochemistry* **31,** 4901-4907.
10. Hammond, G. S. (1955) *J. Am. Chem. Soc.* **77,** 334-338.
11. Matouschek, A. & Fersht, A. R. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 7814-7818.
12. Glasstone, K., Laidler, J. & Eyring, H. (1941) *Theory of Rate Processes* (McGraw-Hill, New York).