Supporting Materials: Protein Oligomerization Monitored by Fluorescence Fluctuation Spectroscopy: Self-Assembly of Rubisco Activase

Manas Chakraborty, Agnieszka M. Kuriata, J. Nathan Henderson, Michael E. Salvucci, Rebekka M. Wachter and Marcia Levitus.

Table of contents:

Figure S1. A model of the closed ring tobacco Rca apo-AAA+ hexamer (alternating cyan and blue protomers; PDB ID 3ZW6) from cryo electron microscopy data with the creosote recognition domain x-ray structure (red; residues 250-351; PDB ID 3THG) superimposed. Shown in yellow are the locations of the Tobacco AAA+ fragment C-termini. **A)** View down the 6-fold axis of the hexameric model. **B)** Side-on-view with the model rotated up 90°.

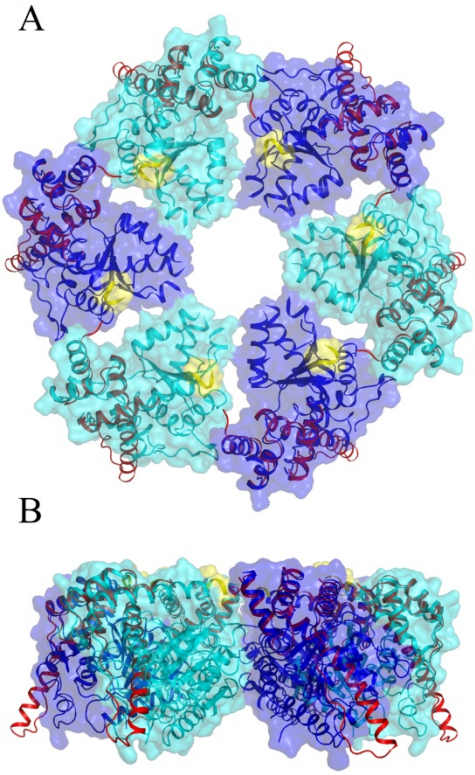
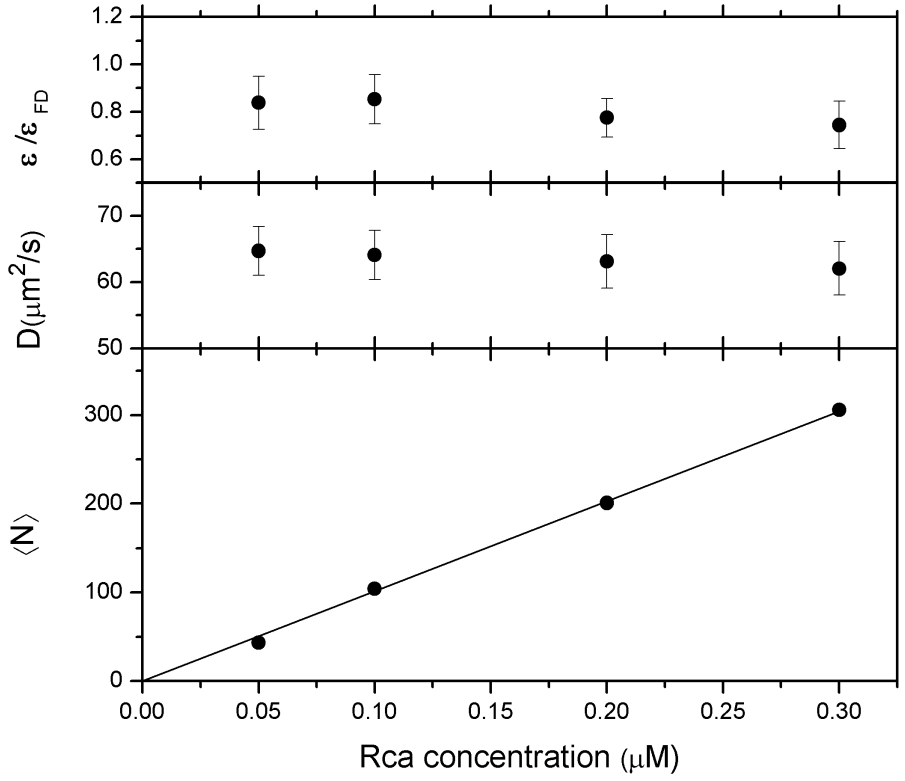


Figure S2: Rca is a monomer below 300 nM. Top: Ratio of the molecular brightness measured in experiments with labeled Rca (ε) and free Alexa 546 dye (ε_{FD}) as a function of Rca concentration. Middle: Diffusion coefficients obtained from fitting the FCS decays measured with labeled Rca in the 50-300 nM range. Bottom: mean number of diffusing particles obtained from the amplitude of the FCS decays measured with labeled Rca (Eq. 2). Put together, these results are consistent with a monomeric form of Rca.



Analytical data on Alexa-labeled Rca preparations.

To calculate the labeling stoichiometry of Alexa-derivatized Rca, protein preparations were analyzed by reverse phase HPLC (Figure S3). Eluting protein fractions were collected and absorbance scans were measured (Figure S4). The molecular masses of labeled Rca and control reactions in the presence of DTT were determined by MALDI (Figure S5). For experimental details, see Methods.

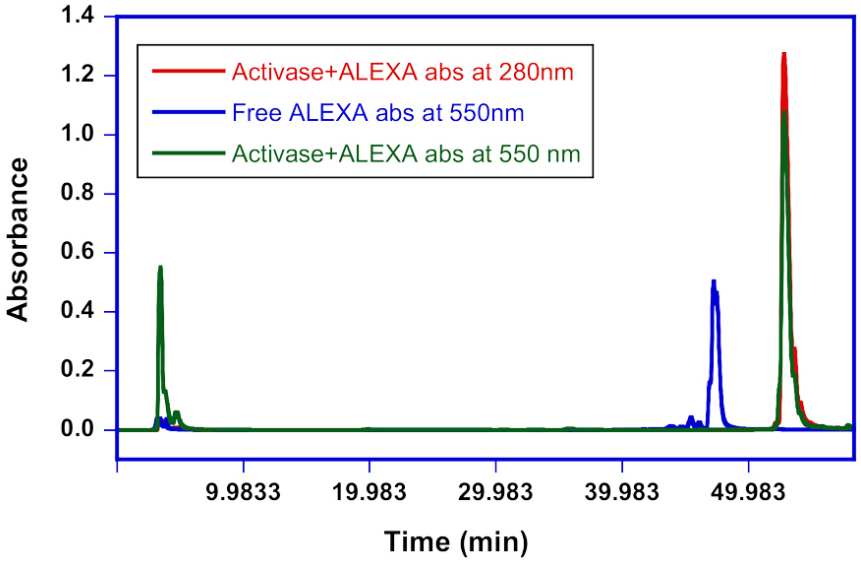


Figure S3. HPLC chromatograms. Reverse-phase HPLC chromatograms of Alexa-labeled Rca monitored at 280 nm (red) and at 550 nm (green), and free Alexa dye (blue). The chromatograms indicate that labeled Rca preparations do not contain any traces of free Alexa dye.

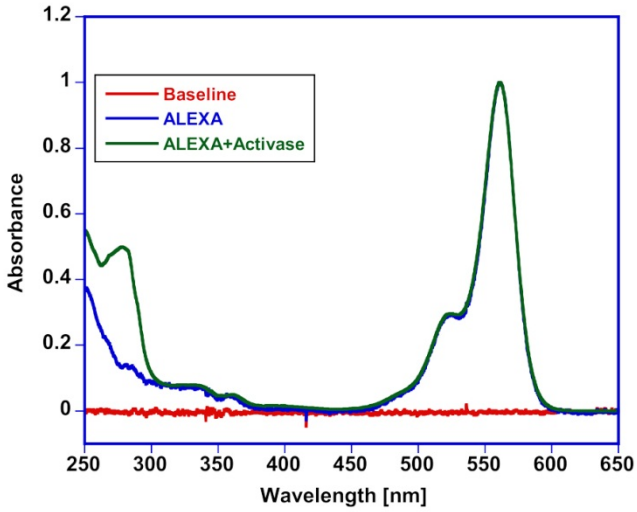


Figure S4. Absorbance spectra. The 55-min peak eluting from the HPLC column (labeled protein) was collected and analyzed by absorbance. The molar ratio of Alexa:Rca in covalently labeled preparations was determined by absorbance as described in the Methods section.

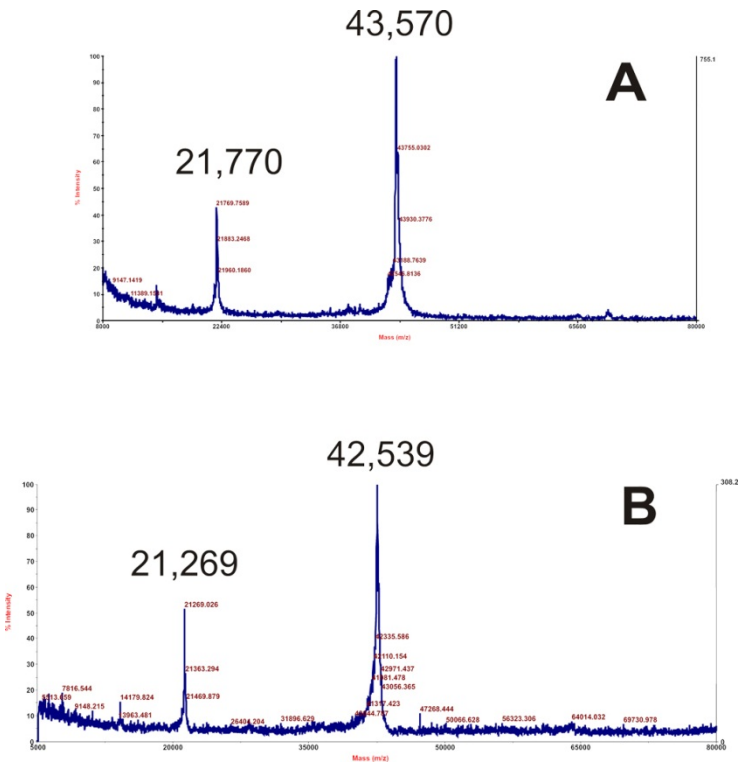


Figure S5. MALDI spectra. Spectra were collected on labeled and unlabeled Rca preparations, and compared to their theoretical values. (A) Labeled protein spectra exhibit main peaks with m/z = 43,570 (z=1), and m/z = 21,770 (z=2). The calculated mass is 43,493 g/mol. (B) Unlabeled protein spectra exhibit main peaks with m/z = 42,539 (z=1), and m/z = 21,269 (z=2). The calculated mass is 42,459 g/mol. The error of the instrument is estimated to be about 90 Da for a protein of this size.

Diffusion coefficient of the monomer and oligomers

In this section we analyze the different approximations made in the calculation of the diffusion coefficients of the monomer and oligomers, and their impact in the interpretation of the experimental results.

Effect of molecular shape.

We have assumed that the different oligomeric states of Rca are approximately spherical. In order to evaluate the possible effect of molecular shape on the values of the ratios D_k/D_1, we modeled the protein as an ellipsoid. The diffusion coefficient of a non-spherical particle is smaller than the diffusion coefficient of a sphere of the same volume, and for an ellipsoid, the diffusion coefficient can be expressed in terms of the ratio of the axial and equatorial semiaxis of the molecule (p) in terms of Perrin's friction factors (1).

$$\frac{D_{sphere}}{D_{ellipsoid}} = \frac{p^{2/3}\xi}{\operatorname{atan}\xi} \quad \text{(oblate ellipsoid)} \qquad \text{(Eq.S1)}$$

$$\xi = \frac{\sqrt{\left|p^2 - 1\right|}}{p} \qquad \text{(Eq.S2)}$$

The structure of the Rca hexamer published by Stotz *et al.* (2) indicates that the particle can be modeled as an oblate ellipsoid with dimensions of about 5.6 nm x 13.5 nm x 13.5 nm, which gives an axial ratio of p = 0.42 (Figure S1). The diffusion coefficient of a particle with this shape is 94% of the value one would measure for a spherical particle of the same volume ($D_{sphere}/D_{ellipsoid}$ = 1.07, Eqs. S1 and S2). The influence of molecular shape in the absolute values of the diffusion coefficients is therefore small, and it is expected to be insignificant in the estimated values of the ratios D_k/D_1.

Effect of hydration water.

In addition, we considered the effect of hydration water, which increases the radius of the diffusing particle and therefore affects the diffusion coefficient. The volume of the hydrated protein (V_h) is related to the volume of the dry molecule (V_{dry}) and the volume of hydration water (V_{water}) as

$$V_h = V_{dry} + V_{water} = \frac{4}{3}\pi R^3 + \frac{M_w}{N_{av}\delta_{water}}h \qquad \text{(Eq.S3)}$$

where R is the radius of the dry sphere, M_w is the molecular weight of the protein, N_{av} is Avogadro's number, δ_{water} is the density of water and h is the mass of hydration water expressed as grams of water per gram of protein (typically in the range 0.2-0.6, (3)).

Because the diffusion coefficient is inversely proportional to the hydrodynamic radius,

$$\frac{D(h_1)}{D(h_2)} = \left(\frac{V_h(h_2)}{V_h(h_1)} \right)^{1/3} \qquad \text{(Eq. S4)}$$

where h_1 and h_2 are two values of hydration water. Using R = 6.35 10^{-7} cm for the hexamer (as described in materials and methods) and M_w = 257.3 kDa, we obtain

$$\frac{D(h=0.6)}{D(h=0.2)} = 0.96$$

This shows that different amounts of bound water in this range change the absolute values of the diffusion coefficients by about 4%. If the amount of bound water is of the same order for the different oligomers, we expect that the effect in the relative values of diffusion coefficients (D_k / D_1 ratios) will be insignificant.

Comparison between different estimates of the diffusion coefficients.

As discussed in the manuscript, we considered two alternative approaches to estimate the diffusion coefficients of the different oligomeric forms of Rca. In the first approach, we use the relationship $D_k = D_1/k^{1/3}$ with the experimentally determined value of D_1 (64.7 $\mu m^2 s^{-1}$). In the second approach, we used the published structure of the reconstructed closed-ring hexamer of tobacco Rca to obtain approximate models for the monomer, dimer, trimer and tetramer (see Eqs. 3 and 4 in the manuscript)

The ratios D_k/D_1 obtained using these two procedures are compared below:

	Method 1 ($D_k = D_1/k^{1/3}$)	Method 2 (from structural data)
D_2/D_1	0.79	0.80
D_3/D_1	0.69	0.66
D_4/D_1	0.63	0.59
D_6/D_1	0.55	0.54

Because both approaches bear significant approximations, it is critical to demonstrate that our interpretation of the experimental results does not depend on the approximations used in the analysis of the data.

To illustrate this point, Fig. S6 shows a comparison of the results obtained using the D_k values estimated from the structure of the reconstructed closed-ring hexamer of tobacco Rca (green line), and using the formula $D_k = D_1/k^{1/3}$ with $D_1 = 64.7\ \mu m^2 s^{-1}$ (red line).

This comparison demonstrates that the dissociation constants reported in the manuscript do not depend significantly on the assumptions behind the estimation of the diffusion coefficients.

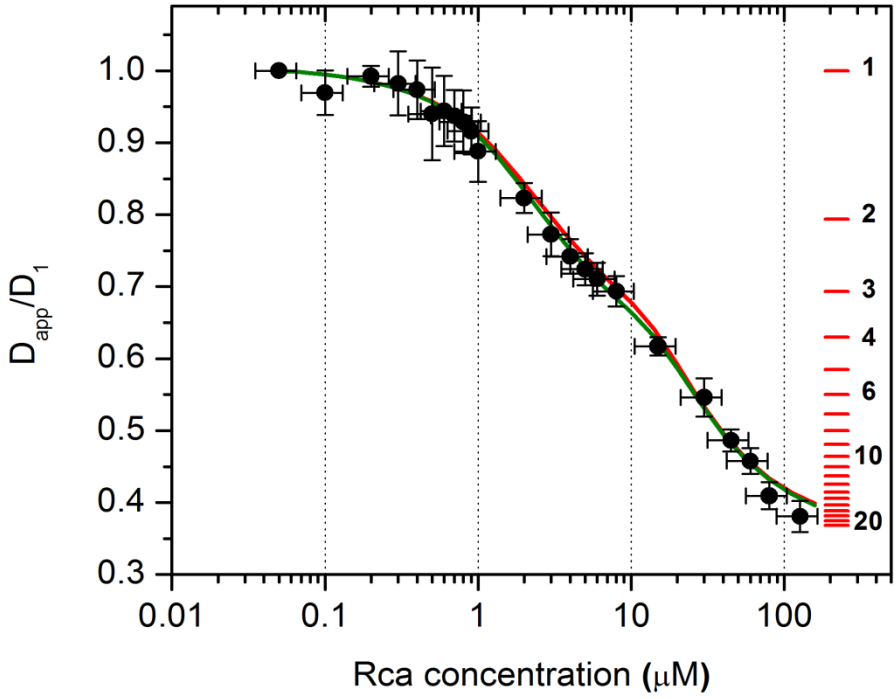


Figure S6. Results of modeling according to model 3 using K_{d1} = 3.5 μM, K_{d2} = 1 μM K_{d3} = 1 μM and K_{d4} = 25 μM^3. The line in green is identical to the one presented in Figure 1D in the manuscript, and has been calculated using the diffusion coefficients calculated from the structure of the reconstructed closed-ring hexamer of tobacco Rca. The line in red has been calculated using $D_k = D_1/k^{1/3}$ with a value of D_1 equal to the diffusion coefficient determined at 50 nM Rca.

Diffusion coefficient of a stack of four hexamers.

As described above, the Rca hexamer can be approximated as an oblate ellipsoid with dimensions of about 5.6 nm x 13.5 nm x 13.5 nm, which gives an axial ratio of 0.42. Estimating the diffusion coefficient of a 24-mer formed by stacking 4 hexamers would require knowledge of structural information that we do not have. As a first approximation we assume that a linear stack of four hexamers can be modeled by a prolate ellipsoid with axial ratio of approximately 1.67 (4 × 5.6 nm /

13.5 nm). As described above, for the hexamer we calculate $D_{sphere}/D_{ellipsoid}$ = 1.07. This ratio is expected to decrease towards 1 if two hexamers are stacked because the axial ratio gets closer to one (more sphere-like particles), but will then increase again if more hexamers are stacked because the oligomer becomes a prolate ellipsoid. For a prolate ellipsoid with axial ratio 1.67 Perrin's equations (Eqs. S2 and S4) give $D_{sphere}/D_{ellipsoid}$ = 1.03. This means that in the case of Rca, the shape factors for the hexamer and the stack of 4 hexamers are almost the same, and therefore it is not unreasonable to assume that $D_{24} = D_6/4^{1/3}$ in the absence of any other structural information.

$$\frac{D_{sphere}}{D_{ellipsoid}} = \frac{p^{2/3}\xi}{\text{atanh}\,\xi} \text{ (prolate ellipsoid)} \qquad \text{(Eq.S4)}$$

FCS analysis of polydisperse samples.

In the manuscript we stated that we could obtain adequate fits using a one-component FCS model even in situations where we strongly believe solutions contain two or more oligomeric species in equilibrium. Meseth et al (4) analyzed the issue of resolvability in the FCS decay of two-component solutions, and concluded that the diffusion times of the two species must differ by a factor of ~1.6 for the decay to be experimentally distinguishable from the decay of a one-component solution. Although this value depends on a variety of experimental factors, it provides a reasonable "rule of thumb". When mixtures cannot be resolved, one obtains an apparent good fit using a one-particle model with an intermediate diffusion time. We performed simulations of FCS decays using the software simFCS (developed by Enrico Gratton) to show that it is reasonable to expect that the FCS decays of mixtures of monomers, dimers, etc will not be distinguishable from the FCS of monodisperse samples. To illustrate this point, Figure S7 shows the result of a simulation that would represent a condition close to what we expect for Rca at 10 μM. We set the simulation assuming that dimers, tetramers and hexamers are present at a 1/3 fractional concentration each. In addition, because the concentration of labeled Rca is much lower than the concentration of unlabeled Rca, the probability that an oligomer will contain more than one fluorescent probe is very small, so it is reasonable to assume that the brightness of dimers, tetramers and hexamers are all the same. For the purpose of the simulation, we further assume that the diffusion coefficients of the species are related by $D_2/D_4 = 2^{1/3}$ and $D_2/D_6 = 3^{1/3}$, and we set D_2 at 52 $\mu m^2 s^{-1}$.

Figure S7-A shows the simulated FCS decay of the mixture of dimers, tetramers and hexamers described above (black line), and the result of fitting this decay with a one component model (red line, and residuals below). The apparent diffusion coefficient recovered from the fit is 40.7 $\mu m^2 s^{-1}$. We then performed a simulation where we assumed all particles had the same diffusion coefficient D = 40.7 $\mu m^2 s^{-1}$ (Figure S7-B, black) and fitted the result with a one component model. We recover a diffusion coefficient of 40.3 $\mu m^2 s^{-1}$, which is very close to our input. Interestingly, the qualities of the fits (judged by the residuals and χ^2) are not significantly different. This exercise shows that FCS is not very sensitive to small differences in diffusion coefficients, and that it is expected that all our decays can be fitted with a one-component model.

The quality of our experimental fits, as judged by the residuals, is not as good as the ones obtained in the simulations (Figure S8). This is not surprising giving the complications of experimental work, especially when working with proteins. Yet, we obtain similar fits routinely with short pieces of fluorescently labeled dsDNA and with other proteins that are known to be monomeric, where solutions are in principle monodisperse. In addition, the quality of the fit is similar at Rca concentrations of 50 nM (where we have strong evidence that Rca is a monomer), and at higher concentrations, where we believe two or more oligomeric species co-exist. Using a two-component model improves the quality of the fit as expected, but

this is also true in situations where we know proteins are monomeric. With this in mind, we believe that the fits obtained with the one-component model are good enough for the purpose of estimating the dissociation constants of Rca, and introducing more components in the model would not necessarily reflect the actual physical composition of the solution.

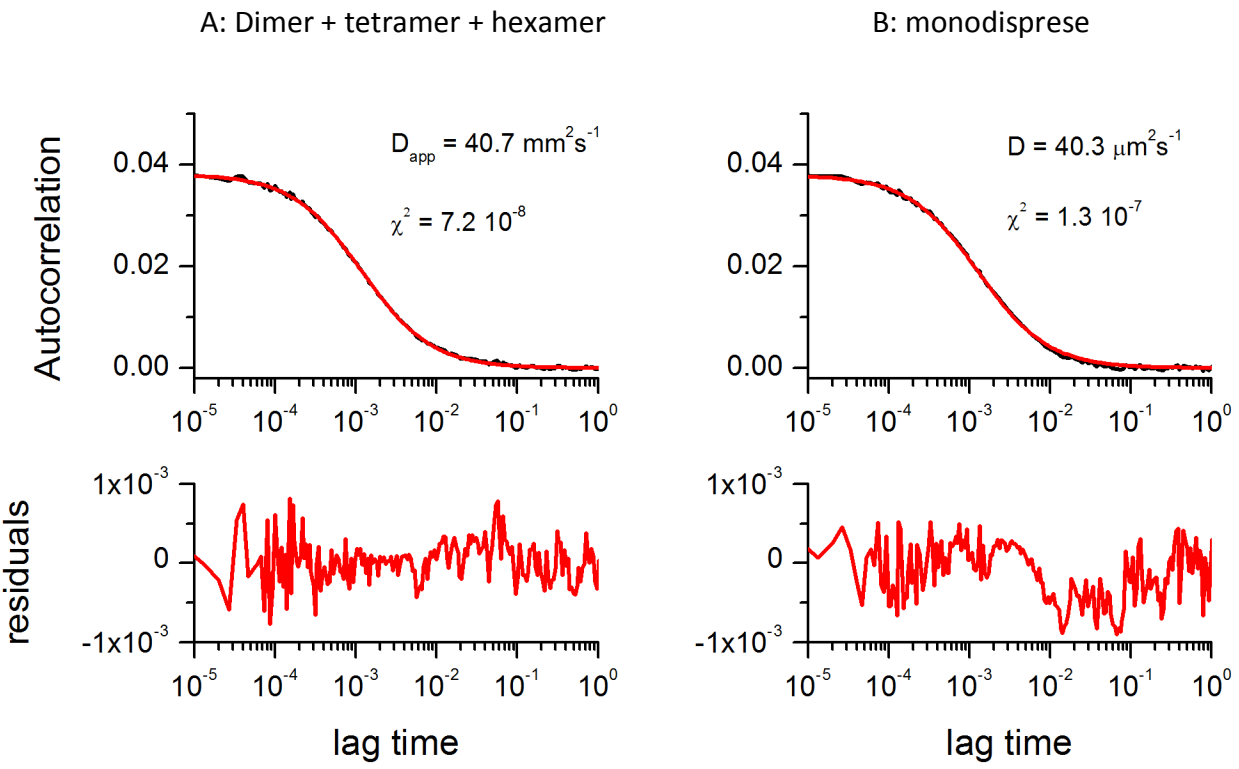


Figure S7. Simulated FCS decays of a mixture of dimer, tetramer and hexamer (A) and a monodisperse sample (B) with the same diffusion coefficient as the apparent diffusion of the mixture. The mixture (fractional concentration of each component = 1/3) produces an FCS decay that can be fitted with a one-component model to give an apparent diffusion coefficient of 40.7 $\mu m^2 s^{-1}$. The quality of the fit, judged by the residuals and χ^2 value, is excellent, and certainly not worse than the fit obtained for the monodisperse sample.

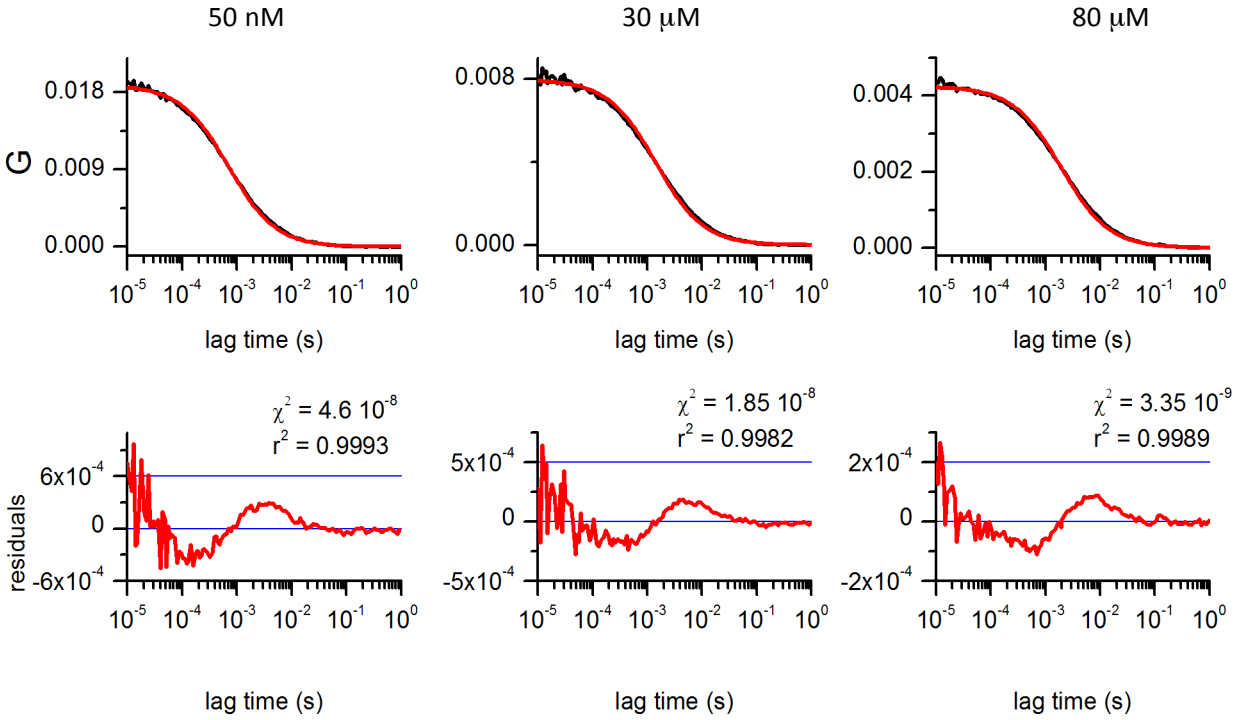


Figure S8. Representative experimental FCS decays (black lines) fitted with a one component model (red lines). The apparent diffusion coefficients extracted from these fits were used to construct figure 1 in the manuscript. The value of χ^2 is calculated as the residual sum of squares divided by the degrees of freedom (N-2, where N is the size of the lag time vector). R^2 is the coefficient of determination. The qualities of the fits are similar at low and high concentrations, where we believe Rca is monomeric (50 nM) or highly polydisperse (high μM).

Uncertainties in the determination of the dissociation constants.

For models 2 and 3, we estimated the range of values of K_{d1} (first step in the assembly) that predict D_{app}/D_1 values consistent with the experimental data. The other dissociation constants were then optimized in each case by keeping the value of K_{d1} fixed.

In the case of model 2, we obtained a good fit to the experimental data using values of K_{d1} in the 3.5-10 μM^2 range (Fig. S9A). This range has been determined arbitrarily by visual inspection considering the error bars in the D_{app}/D_1 and Rca concentration values, and represent only estimates. The values of K_{d2} and K_{d3} used to generate these curves were optimized for the lowest and highest K_{d1} in this range (K_{d2} = 15 μM, K_{d3} = 2 μM^3 for K_{d1} = 3.5 μM^2, and K_{d2} = 1.3 μM, K_{d3} = 70 μM^3 for K_{d1} = 10 μM^2). The concentration profiles on the top panel were obtained with these two sets of K_d values.

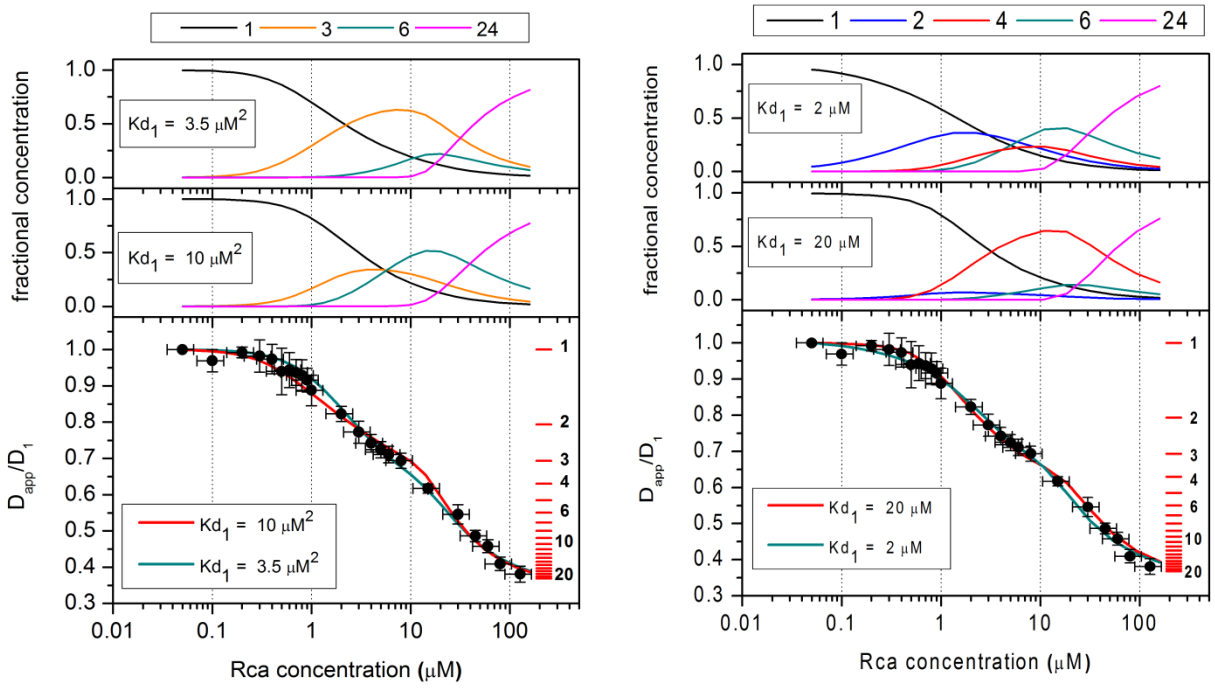


Figure S9. Results of modeling according to models 2 (panel A) and 3 (panel B). The solid circles are the results of experiments, and the red and green solid lines behind them are the results of modeling. **A:** model 2 using K_{d1} = 3.5 μM^2, K_{d2} = 15 μM and K_{d3} = 2 μM^3 (green) and K_{d1} = 10 μM^2, K_{d2} = 1.3 μM and K_{d3} = 70 μM^3 (red). **B:** model 3 using K_{d1} = 2 μM, K_{d2} = 2 μM, K_{d3} = 1 μM and K_{d4} = 20 μM^3 (green) and K_{d2} = 0.03 μM, K_{d3} = 2 μM, K_{d4} = 8 μM^3 for K_{d1} = 20 μM (red). The concentration profiles shown on top of graphs A and B show the fractional concentrations calculated using these sets of K_d values.

In the case of model 3, we can fit the experimental data using values of K_{d1} in the range of 2.0-20 μM (Fig. S9B). This range is larger than in the case of model 2 because model 3 has an additional adjustable

parameter, and the number of parameters in model 2 is already enough to obtain a good fit to the data. The values of of K_{d2} - K_{d4} used to generate these curves were optimized for the values of K_{d1} on each extreme of this range (K_{d2} = 2 μM, K_{d3} = 1 μM, K_{d4} = 20 μM^3 for K_{d1} = 2 μM, and K_{d2} = 0.03 μM, K_{d3} = 2 μM, K_{d4} = 8 μM^3 for K_{d1} = 20 μM).

Supporting References:

1. Perrin, F. 1934. The Brownian movement of an ellipsoid. - The dielectric dispersion of ellipsoidal molecules. J Phys-Paris 5:497-511.
2. Stotz, M., O. Mueller-Cajar, S. Ciniawsky, P. Wendler, F. U. Hartl, A. Bracher, and M. Hayer-Hartl. 2011. Structure of green-type Rubisco activase from tobacco. Nat Struct Mol Biol 18:1366-1370.
3. Squire, P. G., and M. E. Himmel. 1979. Hydrodynamics and Protein Hydration. Arch. Biochem. Biophys. 196:165-177.
4. Meseth, U., T. Wohland, R. Rigler, and H. Vogel. 1999. Resolution of fluorescence correlation measurements. Biophys J 76:1619-1631.