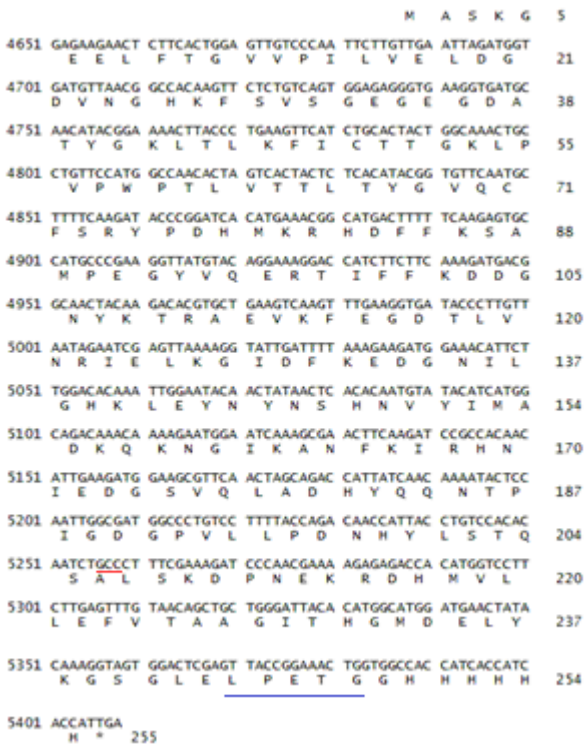
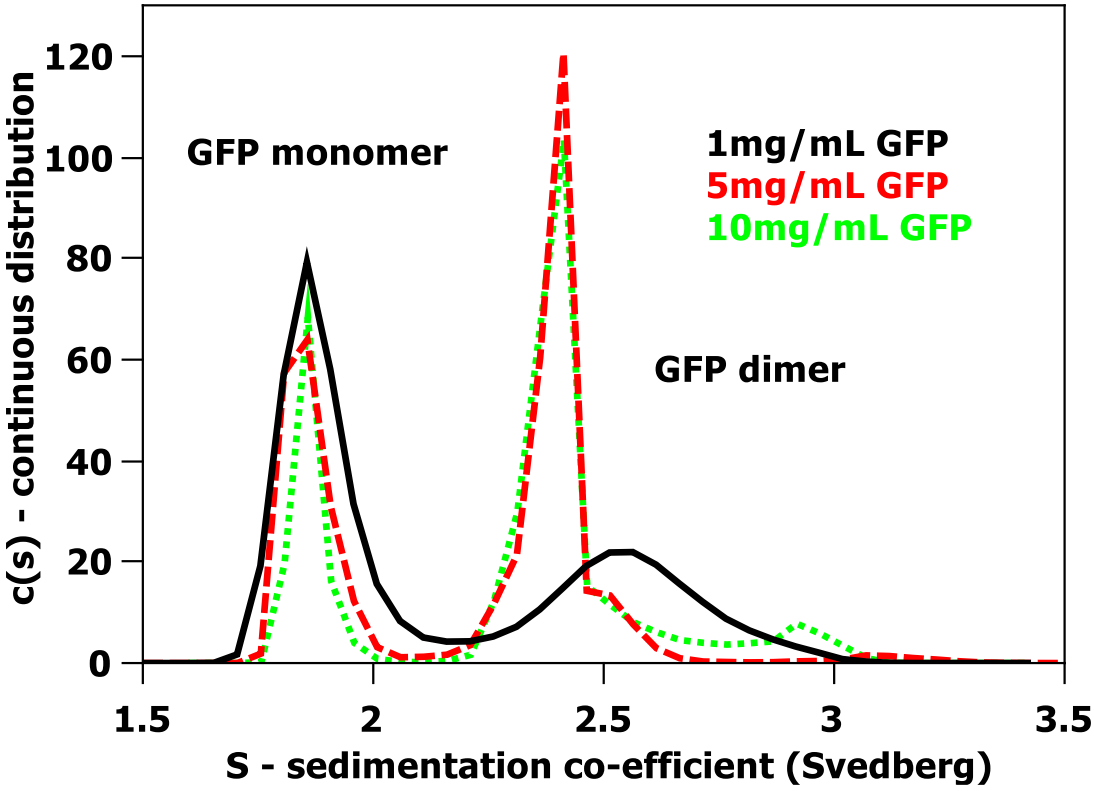
Supplementary Material

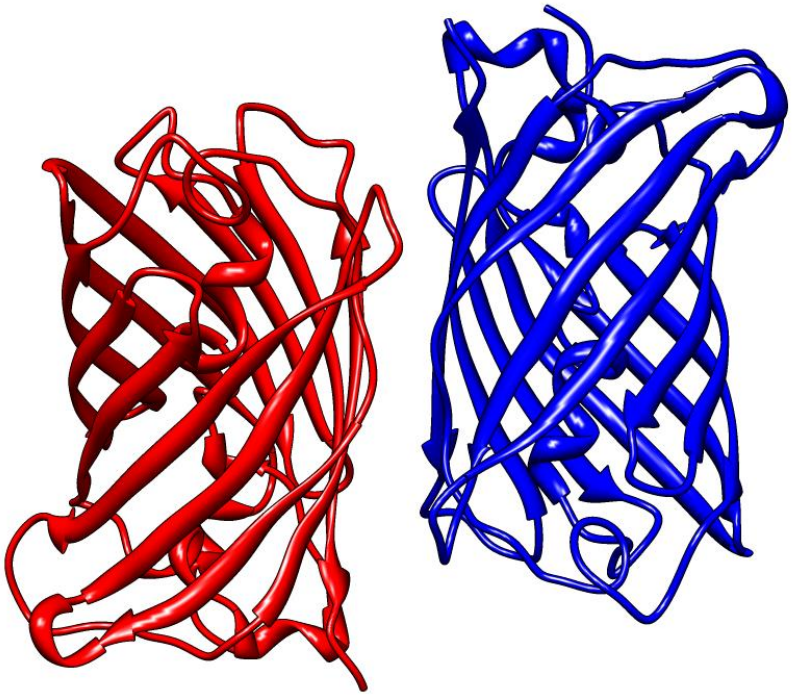


SM Figure 1 – The DNA (top line) and amino acid (bottom line) coding sequence of eGFP with His-tag attached. For m-EGFP, residue 206 (red underlined) has been mutated from an alanine (A – GCC) to a Lysine (K – AAA).

This A206K mutation is known to disrupt eGFP's dimerisation interface, reducing the dimerisation binding affinity from around 100µM in eGFP, to 74 mM in m-eGFP. From SM Table 1 we can see that for the concentrations used for SANS measurement, the protein is now over 99% monomeric. His-Tag cleaved using Sortase A enzyme at LPXTG motif (blue line) before experiments.



SM Figure 2 – A sedimentation velocity AUC experiment using Rayleigh interference optics of hydrogenous eGFP and at 3 different concentrations (1 mg/mL (Black line), 5 mg/ml (Red – dash line) and 10 mg/mL (Green –dot line).
The sedimentation co-efficient distribution was obtained using SEDFIT Analysis. The sample was ran in a 20 mM phosphate, 150 mM NaCl in ddH_2O at 20 ˚C and 129,024 g (RCF).



SM Figure 3 – The crystal structure of the Green Fluorescent Protein (GFP) dimer – Pdb:1GFL [15]

Supplementary Material :- Equation 1) Theoretical determination of monomer / dimer percentages

Non-covalent homo-dimerisation can be described by the following equation:

SM Equation 1) $$M + M \underset{K_D}{\Leftrightarrow} D$$

Where M = monomers and D equals dimers, respectively. The equilibrium dissociation constant (K_D) is the ratio of monomer to dimer.

The binding affinity (KD) can therefore be described as:

SM Equation 2) $$K_D = \frac{[M]^2}{[D]}$$

Where, [M] and [D] are the molar concentrations (mol/L) of monomer and dimer, respectively. The total amount of protein can be expressed as :

SM Equation 3) $$[M]_T = [M] + 2[D]$$

Where [M]T is the total molar concentration (mol/L) of protein. This can be re-arranged as:

SM Equation 4) $$[D] = \frac{[M]_T - [M]}{2}$$

Substituting Equation 4) into Equation 2) we get:

SM Equation 5) $$K_D = \frac{2[M]^2}{[M]_T - [M]}$$

By knowing the total protein concentration ($[M]_T$) and binding affinity (K_D), we can solve to determine the concentration of monomers. By subtraction of the monomer concentration we get the dimer concentration, and these values can be converted into percentages for the system.

	Mg/mL	Total, µM	Monomer, µM	Dimer, µM	Theoretical monomer (%)	Theoretical dimer (%)	AUC Experimental monomer (%)	AUC Experimental dimer (%)
eGFP	1	37.17	24.85	12.43	66.80	33.20	61	39
	5	185.87	74.59	111.28	40.13	59.87	40	60
	10	371.75	113.61	258.14	30.56	69.44	30	70
m-eGFP	1	37.17	37.16	0.01	99.97	0.03	100	0
	5	185.87	184.95	0.92	99.51	0.49	100	0
	10	371.75	368.09	3.66	99.02	0.98	100	0

SM Table 1 – The theoretical (SM Equation 1) and AUC experimental (Figure 3 and SM Figure 2) monomer – dimer percentages of eGFP and m-eGFP at the 3 concentrations (1, 5 and 10 mg/mL) utilised for SAXS and SANS experiments.

Note: eGFP has a monomer dimer dissociation constant (K_d) of 100 µM, whilst for m-eGFP the monomer : dimer dissociation constant (K_d) is 74 mM, as taken from [17].

Technique	**Concentration**	**Radius of gyration (+/-) 1 σ, Å(ngstrom)**	**Zero angle intensity I(0) (+/-) 1 σ , cm^{-1}**
SAXS	1	19.48	0.04 (0.37)
	5	20.29	0.20 (0.40)
	10	20.49	0.43
Average			
Zero angle intensity I(0) (+/-) 1 σ, cm^{-1}			0.4 (+/-) 0.03
95% confidence range (1.96 σ), cm^{-1}			0.34 – 0.46
SANS	1	17.17	0.06 (0.56)
	5	19.59	0.34 (0.68)
	10	20.80	0.72
Average			
Intensity at zero angle I(0) (+/-) 1 σ, cm^{-1}			0.65 (+/-) 0.08
95% confidence range (1.96 σ), cm^{-1}			0.49 – 0.82
Radius of gyration (+/-) 1 σ, Å(ngstrom)		19.64 (+/-) 1.31	
95% confidence range (1.96 σ), Å(ngstrom)		17.06 – 22.21	

SM Table 2 - Experimental results and statistical error table of the Guinier plots for the SAXS and SANS curves

The Guinier plot radius of gyration (Rg) and zero angle intensity (I(0)) to 2 decimal places at each concentration (1, 5 and 10 mg/mL) for the SAXS and SANS data shown in Figures 5 and 6. The radius of gyration (Rg) is then averaged for the SAXS and SANS data (+/-) 1 σ (Standard deviation) and then a 95% confidence level range (1.96 σ) is given. Values outside the confidence range are deemed significant. The same mathematical treatment is provided for SAXS and SANS data for intensity at zero angle (I(0)). For comparative purposes the intensity at zero angle (I(0)) values given in brackets show the 1 and 5 mg/mL values, multiplied by either 10 or 2, respectively, to give the expected 10 mg/mL value.

Day 1	**1 mg**	**5mg**	**10mg**	**Average**
Radius of gyration (Rg) (+/-) 1 σ, Å	16.86	17.86	18.36	17.69 (+/-) 0.76
Intensity at zero angle I(0) (+/-) 1 σ, cm^{-1}	0.04 (0.42)	0.23 (0.46)	0.45	0.44 (+/-) 0.02
Day 15				
Radius of gyration (Rg) (+/-) 1 σ, Å	17.02	18.92	19.13	18.36 (+/-) 1.16
Intensity at zero angle I(0) (+/-) 1 σ, cm^{-1}	0.04 (0.37)	0.19 (0.38)	0.37	0.38 (+/-) 0.01
Day 30				
Radius of gyration (Rg) (+/-) 1 σ, Å	17.85	18.24	18.99	18.36 (+/-) 0.80
Intensity at zero angle I(0) (+/-) 1 σ, cm^{-1}	0.04 (0.40)	0.23 (0.46)	0.41	0.41 (+/-) 0.08
Average				
Radius of gyration (Rg) (+/-)1 σ, Å	17.24 (+/-) 0.53	18.34 (+/-) 0.54	18.83 (+/-) 0.41	18.14 (+/-) 0.82
95% confidence range (1.96 σ), Å	16.20 – 18.28	16.19 – 18.30	17.22 – 18.05	16.52 – 19.75
Intensity at zero angle I(0) (+/-) 1 σ), cm^{-1}	0.40 (+/-) 0.03	0.42 (+/-) 0.08	0.41 (+/-) 0.04	0.41 (+/-) 0.03
95% confidence range (1.96 σ), cm^{-1}	0.35-0.45	0.32 – 0.48	0.32 – 0.48	0.35 – 0.47

SM Table 3 – Experimental and statistical results table of the Guinier plots for the 30 day time course of hydrogenous m-eGFP in deuterated buffer

The Guinier plot radius of gyration (Rg) and zero angle intensity (I(0)) to 2 decimal places at each concentration (1, 5 and 10 mg/mL) and each time point (1, 15 and 30 days) for the SANS data time course shown in Figure 8. The radius of gyration (Rg) is then averaged for the SANS data at each concentration and time point respectively and as a combined total with (+/-) 1 σ (Standard deviation) and then a 95% confidence level range (1.96 σ) is given. Values outside the confidence range are deemed significant. For comparative purposes the intensity at zero angle (I(0)) values given in brackets show the 1 and 5 mg/mL values, multiplied by either 10 or 2, respectively, to give the expected 10 mg/mL value.