

Analytical ultracentrifugation as a tool in the studies of aggregation of the fluorescent marker, Enhanced Green Fluorescent Protein

Aleksandra Dawidziak-Pakula, Joanna Krasowska and Beata Wielgus-Kutrowska✉

Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland

Enhanced green fluorescent protein (EGFP) is a fluorescent marker used in bio-imaging applications, including as an indicator of folding or aggregation of a fused partner. However, the limited maturation, low folding efficiency, and presence of non-fluorescent states of EGFP can influence the interpretation of experimental data. To measure aggregation associated with *de novo* folding of EGFP from a high GdnHCl concentration, the analytical ultracentrifugation method was used. Absorption detection at 280 nm allowed to monitor the presence of monomers and aggregated forms. Fluorescence detection enabled the observation of only properly folded molecules with a functional chromophore. The results showed intensive aggregation of EGFP in low concentrations of GdnHCl with a continuous distribution of aggregated forms. The properly folded monomers with mature chromophore were fluorescent, while the conglomerates of EGFP molecules were not. These facts are essential for a proper interpretation of data obtained with EGFP labelling.

Key words: aggregation; analytical ultracentrifugation; fluorescent marker; green fluorescent protein; protein folding

Received: 02 December, 2019; revised: 16 February, 2020; accepted: 27 February, 2020; available on-line: 19 March, 2020

✉e-mail: Beata.Wielgus-Kutrowska@fuw.edu.pl

Acknowledgements of Financial Support: The described experiments were conducted in NanoFun laboratories co-financed by the ERDF Project POIG.02.00-00-025/09 and in the Laboratory of Biopolymers, ERDF Project POIG.02.01.00-14-122/09. This work was also supported by the Ministry of Science and Higher Education, Republic of Poland (BST-73300/BF) and the University of Warsaw (PhD grant DSM-118000/IFD).

Abbreviations: wtGFP, wild type green fluorescent protein; EGFP, enhanced GFP; GdnHCl, guanidine hydrochloride; AUC, analytical ultracentrifugation with absorbance detection system; F-AUC, analytical ultracentrifugation with fluorescence detection system; FRET, Förster Resonance Energy Transfer; SV, sedimentation velocity method; $s_{20,w}$, sedimentation coefficient value corrected to water at 20°C as a solvent; c(s), continuous sedimentation coefficient distribution; c, the population of macromolecules; FP, fluorescent protein

INTRODUCTION

Wild type green fluorescent protein (wtGFP) is a small, 27 kDa globular protein, isolated from the Pacific jellyfish, *Aequorea victoria* (Shimomura *et al.*, 1962). The native structure consists of 11 β -strands forming a β -barrel and a single α -helix running inside the molecule. In the centre of the protein a unique chromophore is formed after protein folding upon cyclisation, dehydration, and oxidation of residues Ser65-Tyr66-Gly67 (Reid & Flynn, 1997; Bartkiewicz *et al.*, 2018). This chromo-

phore is responsible for wtGFP fluorescence in the range of visible light. After pioneering experiments in which the fluorescent protein was expressed in nematode cells (Chalfie *et al.*, 1994), wtGFP became a source of a high number of mutants with altered spectral and biophysical properties. They developed into widespread biomolecular markers. Enhanced GFP (EGFP, S65T/F64L-GFP) belonged to this group of mutants, with better fluorescent properties than wtGFP (i.e. 35-times higher fluorescence intensity, quantum yield of 0.6) (Cormack *et al.*, 1996; Tsien, 1998).

EGFP fluorescence is used in bio-imaging *in vitro* and *in vivo* (Chalfie & Kain, 2006), in a number of biotechnological and biophysical applications, including observation of gene expression, protein-protein interactions, localisation and migration of proteins (e.g. Tsien, 1998; Skaar *et al.*, 2015; Tan *et al.*, 2018). It can also be an indicator of folding (Chang *et al.*, 2005) or aggregation of a fused partner (Cabantous *et al.*, 2008; Gregoire & Kwon, 2012; Higgins *et al.*, 2018). Data based on green emission of EGFP, used as an aggregation indicator, can be interpreted in two ways. On the one hand, a decrease in fluorescence was interpreted as aggregation of fused partners (Cabantous *et al.*, 2008; Gregoire & Kwon, 2012), on the other hand the observed green emission in aggregates was proof of the target protein's presence (Higgins *et al.*, 2018). In the first case, it was assumed that the decrease in fluorescence was directly related to the aggregation of the fused partner, not GFP. In the second case, it was not obvious that lack of fluorescence indicated that the target protein aggregates were not present. In both cases, the data describing the behaviour of the fused partners could be affected by the photophysical and biochemical properties of EGFP. There are still unanswered questions about how the measured parameters (such as fluorescence intensity and diffusion coefficient) can be properly interpreted considering the behaviour of EGFP, which is dependent on the environment (e.g. pH, ionic strength, buffer composition).

In this context, the fluorescence properties of EGFP are important parameters that serve as a direct measure of protein behaviour. All EGFP molecules are assumed to emit a signal in the visible range of electromagnetic radiation. However, fluorescent proteins can exist in non-fluorescent states caused by the photophysical properties of the chromophore. For example, only the anionic form of EGFP chromophore is fluorescent. Therefore, it is sensitive to pH of environment and its brightness (defined as a product of extinction coefficient and quantum yield) is lowering with decreasing pH value. This is significant when this biomarker is found in different cell compartments, especially those with acidic

pH (Patterson *et al.*, 1997). Additionally, low folding efficiencies, reported for EGFP (Krasowska *et al.*, 2014), including incomplete chromophore maturation, can influence interpretation of the data obtained for fused proteins (Dunsing *et al.*, 2018). Thus, detailed knowledge of EGFP fluorescent properties is necessary to properly interpret data obtained when EGFP was used as a marker.

The analytical ultracentrifugation method was used in this study to measure the aggregation accompanying *de novo* folding of EGFP, obtained from inclusion bodies and dissolved in a buffer with 6 M guanidine hydrochloride (GdnHCl). Absorption detection at 280 nm allowed for the observation of the monomeric and aggregated protein. The more sensitive fluorescent detection (excitation at 488 nm, emission in the range 505–565 nm) enabled identification of the molecules with a functional chromophore that was formed after the proper folding of EGFP. The obtained results show that these techniques can be useful as a tool in the studies of aggregation of enhanced green fluorescent protein and a proof of the existence of its non-fluorescent states.

MATERIALS AND METHODS

EGFP preparation. The protein was expressed in *E. coli* (BL21) strains transformed with pRSET B plasmid containing the EGFP gene with a 6xHis-Tag. 50% of EGFP was found in inclusion bodies. In contrast to soluble fluorescent fraction, this protein had no chromophore. It was purified in the unfolded form in a phosphate buffer at pH 8.0 with 6 M GdnHCl and 300 mM NaCl using IMAC method with a gradient of imidazole, as described earlier (Krasowska *et al.*, 2010). The imidazole was then removed by washing the sample with buffer without imidazole on Amicon filters, and the protein was concentrated to ~1mM. It was stored in a phosphate buffer pH 8.0 with 6 M GdnHCl and 300 mM NaCl. The folding of such EGFP was initiated by diluting the sample at least 100x in a 50 mM phosphate buffer, pH 8.0, containing 300 mM NaCl, 14 mM β -mercaptoethanol (Krasowska *et al.*, 2014) and a desired concentration of denaturant to get the final concentration in the range from 0.06 (residual concentration of GdnHCl) to 2.5 M. The measurement conditions varied from those in which the protein folds/aggregates immediately to those in which it should remain unfolded (with high concentrations of GdnHCl). Dilution was initiated by the rapid injection of a buffer into the tube containing the small volume of concentrated protein.

The protein found in the soluble fraction was in a folded form with a mature fluorescent chromophore and was purified, washed and concentrated analogously to the unfolded EGFP in a 50 mM phosphate buffer at pH 8.0 and 300 mM NaCl, but without GdnHCl.

Absorption and fluorescence characteristics of EGFP. The absorption spectrum of the purified EGFP sample was measured on a Cary 100 UV-Vis spectrophotometer (Agilent Technologies, Australia) using a quartz cuvette. The concentration of the protein was calculated using an extinction coefficient of $21\,900\text{ M}^{-1}\text{cm}^{-1}$ at 280 nm for folded EGFP (Mach *et al.*, 1992) and $19\,800\text{ M}^{-1}\text{cm}^{-1}$ at 280 nm for denatured protein (Gill & von Hippel, 1989), and in stationary absorption measurement it was 8.4 μM and 8.2 μM , respectively. Stationary fluorescence spectra were recorded using an LS-55 spectrofluorometer (Perkin-Elmer, UK). The quartz cuvette (Hellma, Germany) had a path length of 2 mm for excitation and 5 mm for emission. Samples with concentra-

tion 0.1 μM (folded and denatured EGFP) were excited at 295 nm or 489 nm and spectra were recorded at a bandwidth of 2.5 nm both for excitation and emission in the range of 300–550 nm and 500–550 nm, respectively. The measurements were performed at 20°C, in 50 mM phosphate buffer pH 8.0 with 300 mM NaCl for native fluorescent protein, and for unfolded EGFP the buffer contained additionally 6 M GdnHCl.

Analytical ultracentrifugation with absorbance (AUC) or fluorescence (F-AUC) detection system. Sedimentation velocity (SV) experiments were conducted in an Optima XL-I analytical ultracentrifuge equipped with the absorbance (Beckman Coulter, USA) and fluorescence (Aviv Biomedical, Inc., USA) detection systems in a four- or eight-position AN-Ti rotor. Fluorescence detection system contains a 10 mW laser emitting a wavelength of 488 nm. The protein samples were loaded into a double-sector 1.2 cm cell with an epon charcoal centrepiece and either quartz or sapphire windows. For absorbance measurements, a sample solution (390 μl) and a reference buffer (400 μl) were loaded into the right and the left sector, respectively. In the fluorescence measurements, no reference solution was required, hence both sectors were used for the protein samples, with up to 14 samples per run in an eight-hole rotor.

After equilibration at 20°C and 3000 or 5000 rpm, the radial calibration was performed both for absorbance and fluorescence measurements. In the fluorescence measurements, photomultiplier voltage and gain were adjusted for each cell, and focus scans were conducted for all samples. An appropriate focusing depth was selected to maximise the signal, typically around 5000 μm . After initial procedures, the rotor was stopped, and the temperature was equilibrated to 20°C. The ultracentrifuge was then accelerated to 50000 rpm and radial absorption at 280 nm or fluorescence scans of protein-concentration profiles in the cell were collected. The samples were measured about three hours after sample dilution (see EGFP preparation), because of the time needed for temperature equilibration. After each SV experiment, the samples were stored at 20°C in the AUC cells and measured again the next day.

Concentration of protein in AUC measurements.

In statistical measurements of the sedimentation coefficient for folded, fluorescent EGFP, concentration of 5 μM was used. It was a compromise between the detection accuracy of absorption (acceptable S/N ratio) and fluorescence (too high signal intensity for higher concentrations of fluorescent protein) (see Fig. 2).

De novo folding of EGFP was initiated by dilution of the unfolded protein solution containing 6M GdnHCl to a protein concentration of 5 μM (fluorescence detection) or 10 μM (absorption detection) in 50 mM phosphate buffer with 300 mM NaCl and 14 mM β -mercaptoethanol and a final concentration of GdnHCl in the range of 0–2.5 M. In these measurements, part of the aggregated protein dropped to the bottom of the cell during ultracentrifuge speeding. Thus, the initial EGFP concentration of 5 μM was too low to monitor the sedimentation in absorption detection system and the signal in statistical measurements was much noisier than for folded EGFP. Aggregation depends on protein concentration and could be higher for higher protein concentration, but our main goal was to check if we can observe aggregation of the enhanced green fluorescent protein in the analytical ultracentrifugation experiments and if aggregates formed during EGFP folding are fluorescent or not.

Data analysis. Density and viscosity of buffers with different GdnHCl concentrations and partial specific volume of EGFP (0.72689 cm³/g, from amino acid composition) were calculated using the SednTerp programme (Hayes *et al.*, 1995; Hayes *et al.*, 2006).

Sedimentation coefficient was calculated with HYDROPRO programme (Garcia De La Torre *et al.*, 2000) using the structure of EGFP from PDB, entry ID 2Y0G (Royant & Noirclerc-Savoie, 2011), and the following parameters: viscosity of the solvent: 0.010575 P; EGFP partial specific volume: 0.72689 cm³/g; and solvent density: 1.0178 g/cm³.

The sedimentation velocity profiles were analysed using the Sedfit programme with the continuous sedimentation coefficient distribution $c(s)$ model (Schuck, 2000; Lebowitz *et al.*, 2002). Meniscus positions and frictional ratios were treated as adjustable parameters in the non-linear regression of $c(s)$.

A sedimentation coefficient distribution $c(s)$ can be defined according to the Equation 1:

$$a(r,t) = \int c(s) X(s, D(s), r, t) ds + \epsilon \quad (1)$$

where $a(r,t)$ – experimentally observed signal, with an error of measurement ϵ ,

$c(s)$ – the concentration of species with sedimentation coefficients between s and $s+ds$,

$X(s, D(s), r, t)$ – the Lamm equation solution.

The integration of $c(s)$ in the range between s and $s+ds$ gives the total value of signal corresponding to the number of molecules with average sedimentation coefficient s . This value is called population of molecules, ν , for calculated sedimentation coefficient. This value is expressed in absorption or fluorescence units, depending on the detection system.

All s -values were corrected to water at 20°C as a solvent (Equation 2).

$$s(20,w) = s \frac{(1 - \nu \rho(20,w))}{(1 - \nu \rho_b,T)} \cdot \frac{\eta_{b,T}}{\eta(20,w)} \quad (2)$$

s – measured sedimentation coefficient value

$\rho_{b,T}$ – the buffer density

$\eta_{b,T}$ – the buffer viscosity

$\rho_{20,w}$ – the water density in standard solution conditions at 20°C

$\eta_{20,w}$ – the water viscosity in standard solution conditions at 20°C

ν – the partial-specific volume of the protein.

RESULTS AND DISCUSSION

Steady-state measurements

The enhanced green fluorescent protein (EGFP) found in inclusion bodies occurred in an unfolded form without a functional fluorescent chromophore. UV-Vis absorption and fluorescence spectra of this form confirmed that the chromophore was not present in the EGFP structure. The absorption spectra measured in the 240–550 nm range showed only one maximum at 280 nm resulting from the presence of aromatic amino acids (1 Trp and 11 Tyr, Fig. 1A). The fluorescence spectrum of the protein excited at 295 nm had only one maximum at about 350 nm (Fig. 1B, inset).

The protein present in the soluble fraction was folded and with mature fluorescent chromophore. The absorption spectrum revealed an additional band at 489 nm (Fig. 1A). In the emission spectrum upon 295 nm excita-

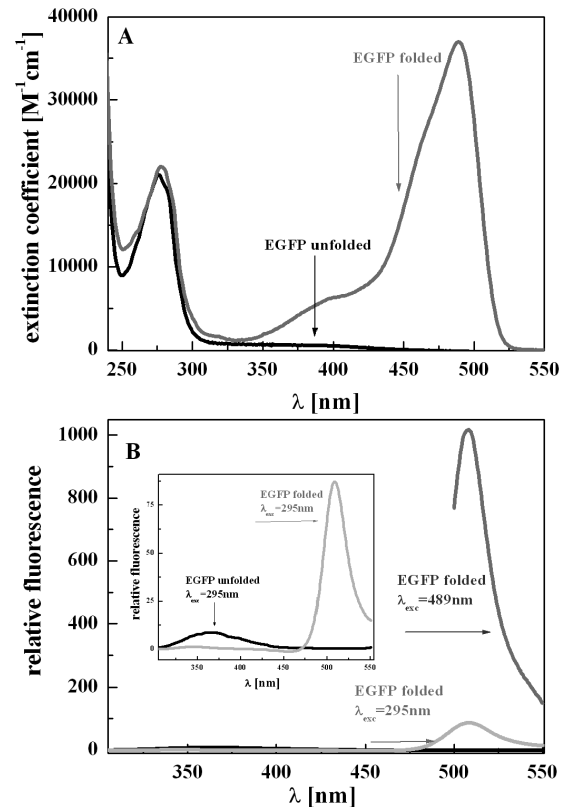


Figure 1. Absorption and fluorescence spectra of different forms of EGFP in pH 8.0.

(A) Absorption spectrum (normalized to the extinction coefficient values) of unfolded protein, without mature chromophore (black line), and folded, with mature chromophore (gray line) EGFP. (B) Fluorescence spectrum of unfolded protein, excited at 295 nm (black line), and folded, excited at 489 nm (gray line) and 295 nm (light gray line).

tion, the tryptophan fluorescence was suppressed, most probably because of Förster resonance energy transfer (FRET) between tryptophan and the chromophore in the β -barrel, due to the small overlap of tryptophan emission spectra of and absorption of EGFP chromophore (compare Figs. 1A and 1B). When the Trp is excited at 295 nm in the denatured protein, the emission of Trp can be observed with maximum fluorescence at about 365 nm, while for the native protein in the same conditions the Trp peak is negligible and the chromophore emission dominates (Fig. 1B, inset). The green fluorescence of the chromophore was observed both for excitation at 295 nm (Trp absorption region) and 489 nm (maximum of chromophore absorption, Fig. 1B).

Statistical measurements of EGFP sedimentation coefficient

The sedimentation coefficient of the folded, fluorescent EGFP monomer has been measured by several labs. The obtained value varies in the range 2.52 S–2.65 S (Vámosi *et al.*, 2016; Zhao *et al.*, 2013). For His-tagged EGFP, two values were reported: 2.81 S for absorption and 2.73 S for fluorescence detection systems (MacGregor *et al.*, 2004).

To confirm these results and to compare the accuracy and repeatability of measurements made using the SV method with absorption or fluorescence detection, experiments with His-tagged EGFP were conducted (see example in Fig. 2). Each of the six measured samples

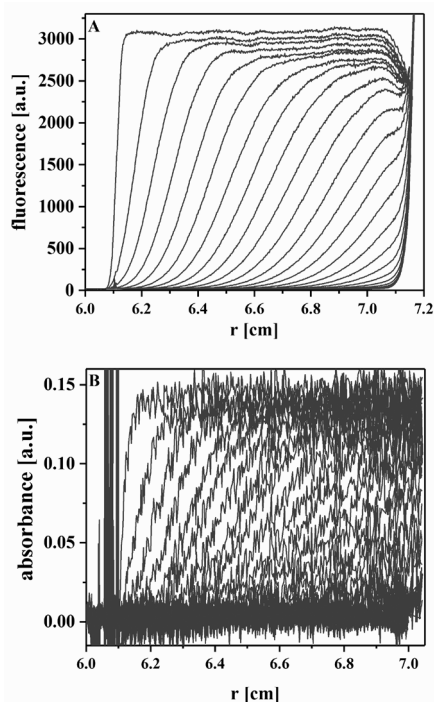


Figure 2. The concentration distributions obtained in the sedimentation velocity experiment for 5 μM native EGFP, properly folded with mature fluorescent chromophore, (A) in fluorescence, or (B) in absorbance detection system, as a function of radial distance (r).

contained native, properly folded and fluorescent protein at 5 μM concentration. The measurements were performed at 20°C and 50 000 rpm.

The sedimentation coefficient obtained using both detection systems was equivalent in the range of the errors: 2.63 ± 0.02 S for absorption, and 2.64 ± 0.08 S for fluorescence. The sedimentation coefficient calculated with HYDROPRO as 2.56 S (PDB ID: 2Y0G) was slightly lower than that obtained from the experiments.

The population of EGFP inferred from the distribution plot in the absorbance measurements showed that monomers represented $81 \pm 10\%$ of absorbing objects in solution. However, the fluorescence detection system showed $93.7 \pm 1.1\%$ of monomers in the mixture of fluorescent molecules. In the fluorescence detection system, a laser emitting at a wavelength of 488 nm was used. Therefore, only the molecules with a correctly formed chromophore were visible during fluorescence measurements. On the other hand, the absorption detection at 280 nm allowed observation of all the molecules that had aromatic amino acids in the primary structure. For this reason, molecules without functional chromophore were also detected.

The EGFP *de novo* folding *in vitro* and aggregation

Protein folding, chromophore maturation, and aggregation of EGFP are processes of a different time scale. Folding proceeds in microseconds, maturation takes hours, and aggregations takes days (Krasowska *et al.*, 2014; Krasowska *et al.*, 2010). Therefore, sedimentation velocity experiments were conducted over the span of a few days to monitor the aggregation that accompanied EGFP *de novo* folding as a function of decreasing concentration of GdnHCl. The existence of molecules with functional fluorescent chromophore was monitored by

observing the 505–565 nm range emission for excitation at 488 nm. The presence of all forms of protein, including non-fluorescent or misfolded monomers and aggregates, was identified by measuring absorption at 280 nm.

AUC measurements in the absorption detection system

The dynamic process of aggregates formation began after mixing unfolded EGFP with appropriate buffer without or at low concentration of denaturant. The aggregates appeared and grew during the measurements. Thus, a one-day experiment specifying the statistics of this process using the same conditions in six cells was performed. A small, highly concentrated unfolded EGFP sample was rapidly diluted to a concentration of 10 μM in a buffer with 14 mM β -mercaptoethanol and without GdnHCl.

The $s_{(20,w)}$ in most cases was higher than for properly folded EGFP. The mean values of $s_{(20,w)}$ and c (the population of macromolecules) were calculated from the six measurements with the standard deviations, as follows:

$$s_{(20,w)} = 2.78 \pm 0.10 \text{ S}$$

$$c = 0.027 \pm 0.011$$

Standard deviation obtained for c was not satisfactory as it was reported at 40%. This was likely due to the rapid protein aggregation, which was not reproducible. Thus, aggregation was an accidental process and it was not possible to accurately predict the creation efficiency of the correct EGFP form in every cell. In addition, we do not see the largest aggregates that fall to the bottom of the cell during the centrifuge acceleration.

The value of the standardised sedimentation coefficient $s_{(20,w)}$ was higher than the sedimentation coefficient of properly folded EGFP, which was 2.63 S. In the absorption system, the molecules that possess the aromatic amino acids, which absorb at 280 nm, were observed. The presence of mature, folded chromophore is not obvious. The peak obtained from data analysis interpreted as a monomer may not represent an individual state of the EGFP molecule, but a dynamic mixture of the different forms of protein. Probably, not only folded monomers, but also initial forms of aggregates – misfolded molecules prone to aggregation – are hidden under this one peak. Sedfit programme could not separate these individuals.

Next, seven-day folding experiments were conducted, in which unfolded EGFP was initially rapidly diluted to a concentration of 10 μM in buffers with different concentrations of GdnHCl (0–2.5 M).

The population of protein remaining in the solution is dependent on the denaturant concentration. In all samples the initial concentration of protein was the same. The high concentration of GdnHCl is a hindrance to folding and aggregation of EGFP. Consequently, in the presence of GdnHCl, a high population of protein remained in the solution (for 2M of GdnHCl see Fig. 3E). With a decrease in GdnHCl concentration, an acceleration in the folding and accompanying aggregation was observed. The huge conglomerates of EGFP aggregates dropped to the bottom of the cell during ultracentrifuge speeding. The results presented in Fig. 3 show that the population of monomeric EGFP remaining in solution was much lower for low concentration of GdnHCl than for 2M GdnHCl (Fig. 3A, C). As expected, the monomer sedimentation coefficient (s) depended on the buffer density, viscosity and stage of the protein unfolding and

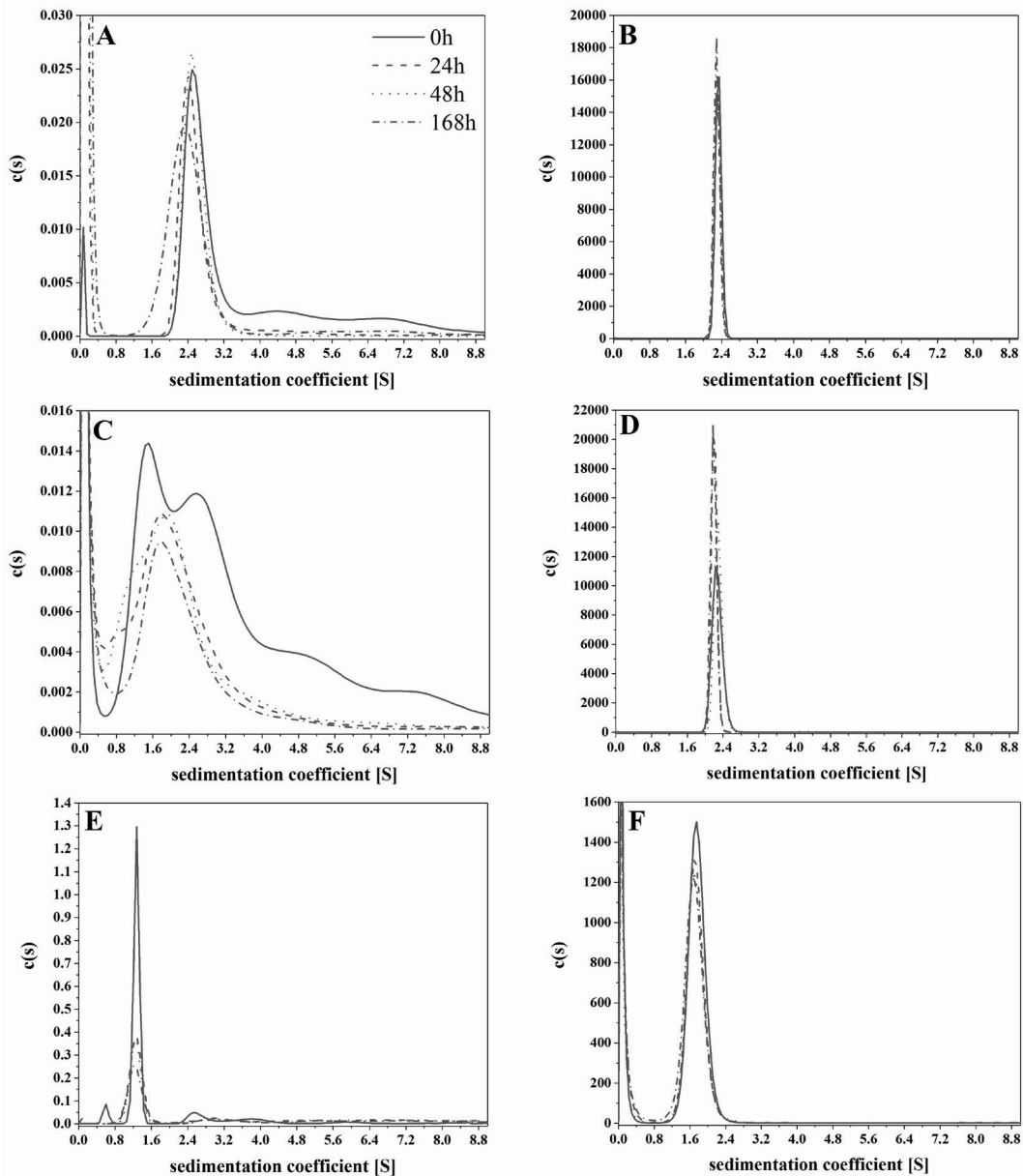


Figure 3. The results of analytical ultracentrifugation measurements performed during observation of *de novo* folding and aggregation of EGFP for: A–B, 0 M; C–D, 0.5 M; and E–F, 2 M GdnHCl obtained in absorption (A, C and E) and fluorescence (B, D and F) detection system for 0 (solid line), 24 (dashed line), 48 (dotted line) and 168 (dash-dotted line) hours of experiment. The sharp peak visible in the fluorescence experiment for high GdnHCl concentration (panel E) probably comes from small fluorescent molecules with concentration undetectable by absorption.

varied with the concentration of GdnHCl. It is noteworthy that for absorbance detection system, the molecules with a lower sedimentation coefficient were observed in several, but not all samples, which could be interpreted as unfolded molecules in the mixture. There were also individuals with higher sedimentation coefficients, probably dimers and larger aggregates of EGFP with immature, non-fluorescent chromophore (Fig. 3A, C, and E).

AUC measurements in the fluorescence detection system

The fluorescence detection system showed mainly the stable monomers with mature chromophore, while the non-fluorescent conglomerates of EGFP molecules were not visible.

The analysis of data obtained for different GdnHCl concentrations showed that the fluorescent molecules were mainly monomers (99% of observed molecules, Fig. 3B, D, F for 0, 0.5, 2 M GdnHCl, respectively). There was a trace population of fluorescent protein aggregates (likely dimers). In contrast with absorption measurements, the calculated $s_{(20,\mu)}$ for fluorescent monomers was similar to that for folded EGFP with mature chromophore. For example, for protein in 2 M GdnHCl $s_{(20,\mu)}$ was in the range of 1.78–1.83 S for absorption detection and 2.40–2.49 S for fluorescence (Table 1).

The folded monomers were very stable, and the sedimentation coefficient changed negligibly over time (Fig. 4A), but as the concentration of GdnHCl in-

Table 1. The s [S] and $s_{(20,w)}$ [S] obtained for EGFP folding and aggregation in the sedimentation velocity measurements as a function of GdnHCl concentration for absorbance (A) or fluorescence (F) detection system. The unit [S] (Svedberg) is equal to 10^{-13} s.

[GdnHCl]	0 M	0.2 M	0.5 M	1 M	1.3 M	2 M							
Time [h]	Detector	s	$s_{(20,w)}$	s	$s_{(20,w)}$	s	$s_{(20,w)}$	s	$s_{(20,w)}$	s	$s_{(20,w)}$	s	$s_{(20,w)}$
0	A	2.51	2.76	2.51	2.83	2.54	2.93	1.89	2.34	1.67	2.14	1.29	1.83*
	F	2.33	2.59	2.11	2.38	2.27	2.65	2.14	2.66	1.87	2.49	1.76	2.49
24	A	2.41	2.66	2.54	2.87	1.82	2.10	1.80	2.21	1.68	2.15	1.26	1.78*
	F	2.31	2.58	2.09	2.38	2.20	2.57	2.07	2.57	1.79	2.32	1.73	2.45
48	A	2.49	2.74	2.53	2.85	1.93	2.23	1.78	2.20	1.67	2.14	1.27	1.80*
	F	2.31	2.57	2.13	2.38	2.29	2.68	2.07	2.57	1.85	2.39	1.73	2.45
168	A	2.31	2.69	2.51	2.83	1.81	2.09	1.96	2.42	1.94	2.49	1.28	1.82*
	F	2.29	2.55	2.11	2.38	2.19	2.55	2.07	2.55	1.86	2.40	1.69	2.40

*Probably EGFP present in unfolded form.

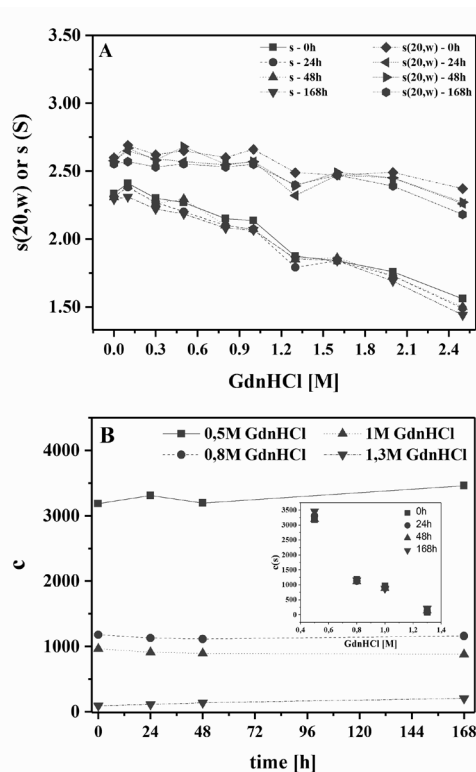


Figure 4. (A) Sedimentation coefficient s and standard sedimentation coefficient $s_{(20,w)}$ obtained from analytical ultracentrifugation measurements with fluorescence detection of folding of EGFP with respect to increasing concentration of GdnHCl.

Measurements were performed for GdnHCl concentration of 0–2.5 M and as a function of time (0–168 hours).

(B) c – population of EGFP monomers for different GdnHCl concentrations as a function of time.

Inset shows c as a function of GdnHCl.

creased, a decrease in the fluorescent monomer content was observed (inset in Fig. 4B).

CONCLUSIONS

Analytical ultracentrifugation with a fluorescence detection system is useful for monitoring events during

FPs folding. Only mature, properly folded EGFP can be detected by this system, while unfolded or largely unfolded protein is not fluorescent but is detectable by absorption. The unfolded protein tends to form aggregates over time. The huge aggregates fall to the bottom of the cell when the centrifuge accelerates. Therefore, sedimentation velocity (SV) in combination with absorbance detection report the s -value distribution of the total protein remaining in the solution, while SV with fluorescence detection shows only the sedimentation behaviour of the folded, active fluorophore. A combination of these two detection methods would possibly allow determining the fraction of an EGFP sample, which is active, the fraction of monomeric unfolded EGFP and the fraction of aggregates in solution.

Our results showed a slow aggregation of EGFP in the solution with 1.6–2 M GdnHCl concentration. Aggregation accelerated at low concentration of denaturant. The sedimentation coefficient $s_{(20,w)}$ changed from 1.66 (for unfolded protein) to 2.8 (for folded protein) in the 50 mM phosphate buffer at pH 8.0, with 300 mM NaCl and GdnHCl concentrations in the range of 0–2.5 M.

This study has shown that in experiments where EGFP was used as a non-interacting fluorescent biomarker, the existence of states without mature emitting chromophore must be considered for a proper interpretation of data obtained with EGFP labelling. In addition, measurements, where the folding of an FP partner is monitored by fluorescence, which disregard non-fluorescent FP forms, could lead to misinterpretation of the data. Thus, the aggregation of the marker itself also should be considered.

Acknowledgements

BWK thanks prof. Patricia Clark from University of Notre Dame, Department of Chemistry and Biochemistry for initializing work with EGFP and previous cooperation.

REFERENCES

Bartkiewicz M, Kazazić S, Krasowska J, Clark PL, Wielgus-Kutrowska B, Bzowska A (2018) Non-fluorescent mutant of green fluorescent protein sheds light on the mechanism of chromophore forma-

- tion. *FEBS Lett.* **592**: 1516–1523. <https://doi.org/10.1002/1873-3468.13051>
- Cabantous S, Rogers Y, Terwilliger TC, Waldo GS (2008) New molecular reporters for rapid protein folding assays. *PLoS ONE* **3**: e2387. <https://doi.org/10.1371/journal.pone.0002387>
- Chalfie M, Kain S eds. (2006) *Green fluorescent protein: properties, applications, and protocols*. 2nd edn, pp 7–10. Wiley-Interscience, Hoboken. <https://doi.org/10.1021/np068247i>
- Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC (1994) Green fluorescent protein as a marker for gene expression. *Science* **263**: 802–805.
- Chang HC, Kaiser CM, Hartl FU, Barral JM (2005) *De novo* folding of GFP fusion proteins: high efficiency in eukaryotes but not in bacteria. *J. Mol. Biol.* **353**: 397–409. <https://doi.org/10.1016/j.jmb.2005.08.052>
- Cormack BP, Valdivia RH, Falkow S (1996) FACS-Optimized mutants of the green fluorescent protein (GFP). *Gene* **173**: 33–38. [https://doi.org/10.1016/0378-1119\(95\)00685-0](https://doi.org/10.1016/0378-1119(95)00685-0)
- Dunsing V, Luckner M, Zühlke B, Petazzi RA, Herrmann A, Chiantia C (2018) Optimal fluorescent protein tags for quantifying protein oligomerization in living cells. *Sci. Rep.* **8**: 10634. <https://doi.org/10.1038/s41598-018-28858-0>
- García De La Torre J, Huertas ML, Carrasco B (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J.* **78**: 719–730. [https://doi.org/10.1016/S0006-3495\(00\)76630-6](https://doi.org/10.1016/S0006-3495(00)76630-6)
- Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**: 319–326
- Gregoire S, Kwon I (2012) A revisited folding reporter for quantitative assay of protein misfolding and aggregation in mammalian cells. *Biotechnol. J.* **7**: 1297–1307. <https://doi.org/10.1002/biot.201200103>
- Hayes D, Laue T, Philo J (1995) *Program Sedinterp: Sedimentation Interpretation Program*. Alliance Protein Laboratories, Thousand Oaks, CA
- Hayes DB, Laue T, Philo J (2006) *Sedimentation Interpretation Program (version 1.09)*. University of New Hampshire, Durham, NH. <https://www.jphilo.mailway.com/download.htm>
- Higgins R, Kabbaj MH, Hatcher A, Wang Y (2018) The absence of specific yeast heat-shock proteins leads to abnormal aggregation and compromised autophagic clearance of mutant Huntingtin proteins. *PLoS ONE* **13**: e0191490. <https://doi.org/10.1371/journal.pone.0191490>
- Krasowska J, Olasek M, Bzowska A, Clark PL, Wielgus-Kutrowska B (2010) The comparison of aggregation and folding of enhanced green fluorescent protein (EGFP) by spectroscopic studies. *Spectroscopy* **24**: 343–348. <https://doi.org/10.3233/SPE-2010-0445>
- Krasowska J, Ugrinov KG, Clark PL, Sienkiewicz A, Bzowska A, Wielgus-Kutrowska B (2014) Spectroscopic properties of two single-cysteine mutants of EGFP: C48S-EGFP and C70S-EGFP. *BSI* **3**: 231–236. <https://doi.org/10.3233/BSI-140070>
- Lebowitz J, Lewis MS, Schuck P (2002) Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci.* **11**: 2067–2079
- MacGregor IK, Anderson AI, Laue TM (2004) Fluorescence detection for the XLI analytical ultracentrifuge. *Biophys. Chem.* **108**: 165–185. <https://doi.org/10.1016/j.bpc.2003.10.018>
- Mach H, Middaugh, CR, Lewis RV (1992) Statistical determination of the average values of the extinction coefficients of tryptophan and tyrosine in native proteins. *Anal. Biochem.* **200**: 74–80
- Patterson GH, Knobel SM, Sharif WD, Kain SR, Piston DW (1997) Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys J.* **73**: 2782–2790. [https://doi.org/10.1016/S0006-3495\(97\)78307-3](https://doi.org/10.1016/S0006-3495(97)78307-3)
- Reid BG, Flynn GC (1997) Chromophore formation in green fluorescent protein. *Biochemistry* **36**: 6786–6791. <https://doi.org/10.1021/bi970281w>
- Royant A, Noirclerc-Savoie M (2011) Stabilizing role of glutamic acid 222 in the structure of enhanced green fluorescent protein. *J. Struct. Biol.* **174**: 385–390. <https://doi.org/10.1016/j.jsb.2011.02.004>
- Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**: 1606–1619. [https://doi.org/10.1016/S0006-3495\(00\)76713-0](https://doi.org/10.1016/S0006-3495(00)76713-0)
- Shimomura O, Johnson FH, Saiga Y (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusa, aequorea. *J. Cell. Comp. Physiol.* **59**: 223–239. <https://doi.org/10.1002/jcp.1030590302>
- Skaar K, Korza HJ, Tarry M, Sekyrova P, Högbom M (2015) Expression and subcellular distribution of GFP-tagged human tetraspanin proteins in *Saccharomyces cerevisiae*. *PLoS ONE* **10**: e0134041. <https://doi.org/10.1371/journal.pone.0134041>
- Tan Y, Zheng J, Liu X, Lu M, Zhang C, Xing B, Du X (2018) Loss of nucleolar localization of NAT10 promotes cell migration and invasion in hepatocellular carcinoma. *Biochem. Biophys. Res. Commun.* **499**: 1032–1038. <https://doi.org/10.1016/j.bbrc.2018.04.047>
- Tsien R (1998) The green fluorescent protein. *Annu. Rev. Biochem.* **67**: 509–544. <https://doi.org/10.1146/annurev.biochem.67.1.509>
- Vámosi G, Mücke N, Müller G, Krieger JW, Curth U, Langowski J, Tóth K (2016) EGFP oligomers as natural fluorescence and hydrodynamic standards. *Sci. Rep.* **6**: 33022. <https://doi.org/10.1038/srep53022>
- Zhao H, Casillas EJr, Shroff H, Patterson GH, Schuck P (2013) Tools for the quantitative analysis of sedimentation boundaries detected by fluorescence optical analytical ultracentrifugation. *PLoS ONE* **8**: e77245. <https://doi.org/10.1371/journal.pone.0077245>