TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt

Lehrstuhl für Chemie der Biopolymere

Characterization of protein stoichiometry via fluorescence anisotropy and GFP fusion proteins

Philipp Johann Heckmeier

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigten Dissertation.

Vorsitzender: Prof. Dr. Martin Klingenspor

Prüfer der Dissertation:

1. Prof. Dr. Dieter Langosch

2. Prof. Dr. Don C. Lamb

Die Dissertation wurde am 26.09.2019 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 06.02.2020 angenommen.

Zusammenfassung

Eine große Anzahl löslicher und Membran-assoziierter Proteine bildet multimere Strukturen mit zwei, drei oder mehreren identischen Untereinheiten. Diese Protein-Oligomere sind an zahlreichen zellulären Prozessen beteiligt, einschließlich der Signalübertragung und der Genregulation. Die Frage, ob ein Protein ein Dimer, Trimer oder höheres Oligomer ausbildet, d.h. seine Stöchiometrie, ist für das Verständnis seiner Funktionalität von wesentlicher Bedeutung.

Es gibt eine Vielzahl von Methoden, um die Stöchiometrie von Proteinen zu charakterisieren. Die Bestimmung des Oligomerisierungs-Grades von Proteinen stellt jedoch immer noch eine Herausforderung dar. Es gibt Verfahren, die dafür den Förster-Resonanzenergietransfer zwischen homotypischen Fluorophoren (homo-FRET) nutzen. Für diese Methoden ist es erforderlich, dass die oligomerisierenden Untereinheiten mit einer fluoreszierenden Sonde markiert werden. Kommt es zur Oligomerisierung, führt homo-FRET zwischen diesen Sonden zur teilweisen Depolarisation des emittierten Lichts – und das in Abhängigkeit von der Stöchiometrie des Protein-Clusters. Das Ausmaß der gemessenen Depolarisation wird durch die Fluoreszenzanisotropie quantifiziert.

In dieser Studie wurden superhelikale Peptide in Fusionsproteinen mit einem grün fluoreszierenden Protein (GFP) fusioniert. Zur Bestimmung des Oligomerisierungs-Grades dieser Peptide über homo-FRET wurden zwei Ansätze auf ihre Anwendbarkeit getestet, die beide die Fluoreszenzanisotropie im Fließgleichgewicht nutzen. Einer der Ansätze erfordert Daten aus dem Fließgleichgewicht und zusätzlich Parameter, die durch zeitaufgelöste Experimente ermittelt wurden. Ein anderer Ansatz basiert ausschließlich auf der Veränderung der Anisotropie im Fließgleichgewicht nach sukzessivem Photobleichen der GFP-Domänen. Für die letztgenannte Technik wurde ein geeignetes theoretisches Modell erstellt, um den Oligomerisierungs-Grad zu berechnen. Mit Hilfe dieser Methoden konnte die Stöchiometrie von mehreren Modellproteinen korrekt bis zum Trimer bestimmt werden. Diese Arbeit präsentiert damit den theoretischen Rahmen für die Auswertung experimenteller homo-FRET Messungen zur zuverlässigen Unterscheidung von Monomer, Dimer, Trimer und höherem Oligomer.

Bei der Untersuchung der gebleichten GFP-Fusionsproteinen haben wir ferner festgestellt, dass die Bestrahlung von GFP mit blauem Licht hoher Intensität zum Abbau des Proteins führt. Dieser Licht-induzierte Abbau wurde bei zwei verschiedenen GFP-Varianten beobachtet. Zwei Spaltstellen konnten genauer identifiziert werden, am klarsten eine dem Fluorophor benachbarte Fragmentierungsstelle am C_α Atom an Position 65. Es ist zu erwarten, dass die in dieser Studie gewonnenen Erkenntnisse die Anwendung des GFP als fluoreszierender Reporter erweitern. Zudem ermöglichen sie eine neue Perspektive auf das Verhalten des GFP-Fluorophors bei extrem intensiver Bestrahlung.

Abstract

A large number of soluble and membrane-associated proteins form multimeric structures with two, three, or more identical subunits. These protein oligomers are involved in numerous cellular processes, including signal transmission and gene regulation. The question of whether a protein interacts as a dimer, trimer, or higher oligomer, i.e. its stoichiometry, is essential to understand its functionality.

There is a variety of methods to characterize the stoichiometry of proteins. However, the determination of the oligomeric state of proteins is still challenging. There are methods that use Förster resonance energy transfer between homotypic fluorophores (homo-FRET). For these techniques, it is necessary that oligomerizing proteins are labelled with a fluorescent tag. If the proteins form oligomers, homo-FRET between fluorescent labels leads to the partial depolarization of the emitted light, in dependence of the stoichiometry. The extent of measured depolarization is quantified by the fluorescence anisotropy.

In this study, we fused α-helical coiled-coil peptides in fusion proteins with a green fluorescent protein (GFP) moiety. To determine the oligomeric state of the coiled-coil fusion proteins via homo-FRET, we tested two approaches for their applicability, both making use of the fluorescence anisotropy in the steady-state. One of the approaches requires steady-state data and additional parameters gained from time-resolved experiments. The other approach is based solely on the steady-state anisotropy upon fractional photobleaching of the GFP moieties. For the latter technique, a suitable theoretical model was generated to calculate the oligomeric state. With the help of these methods, the stoichiometry of a number of model proteins could be determined accurately up to the trimer. This thesis thus provides a framework to evaluate experiments which reliably differentiate between monomer, dimer, trimer, and higher oligomer via homo-FRET.

When studying photobleached GFP fusion proteins, we further found that the irradiation of GFP with blue high-intensity light leads to the degradation of the protein. This light-induced degradation was observed for two different GFP variants. Two cleavage sites could be identified more specifically, most clearly a fragmentation site vicinal to the fluorophore, at the C_α atom of residue 65. It is expected that insights gained in this study broaden future

applications of GFP and provide a new perspective on the behavior of the GFP fluorophore under extremely intense irradiation.

Acknowledgements

Throughout my work on the present doctoral dissertation, I have received support and assistance from many persons whom I want to thank explicitly, here.

I am deeply grateful to my doctoral advisor Prof. Dr. Dieter Langosch for the great supervision during this work. I also want give thanks to my second supervisor Dr. Mark Teese for his help.

Then, I would like to thank Prof. Dr. Don C. Lamb to examine this thesis and Prof. Dr. Martin Klingenspor for being the chair of the examining committee.

In the course of my work, I could rely on the support of collaboration partners and people who provided me with expertise, material, and equipment. Hereby, I want to express my gratitude and appreciation to: Prof. Dr. Don C. Lamb and Ganesh Agam (Chair of Physical Chemistry I, LMU Munich) for the time-resolved fluorescence measurements and the close collaboration; Dr. Martin Haslbeck (Chair of Biotechnology, TUM) and Walter Stelzer (Chair of Biopolymer Chemistry, TUM) for the mass spectrometric measurements; Dr. Ajitha Cristie-David and Prof. Dr. Neil Marsh (University of Michigan) for kindly providing me with their plasmids; Prof. Dr. Siegfried Scherer and Michaela Kreitmeier (Chair for Microbial Ecology, TUM) for puting their FPLC system at my disposal; and Dr. Christina Scharnagl (Chair of Physics of Synthetic Biological Systems, TUM) for her expertise.

I would like to thank my doctoral mentor Dr. Florian Bolze, on one hand, and the students that I could supervise during my time as a PhD candidate, on the other hand: Maximilian Ehrenfeld, Fenna Kolbe, Michael Lauber, and Katja Salbaum (in alphabetical order).

Further, I am deeply indebted to Prof. Dr. Dieter Langosch, Dr. Mark Teese, Fabian Schmidt, Fenna Kolbe, and Florian Heckmeier for revising sections of this thesis. At last, I want to express my appreciation to my colleagues at the Chair of Biopolymer Chemistry for the pleasant work environment.

Content

Zusammenfassung I

Abstract III

Acknowledgements V

Content VII

1 Introduction 1

1.1 Studying protein stoichiometry via homo-FRET 1

1.2 Green fluorescent protein – an appropriate fluorophore for a homo-FRET reporter system 7

2 Motivation 13

3 Theory 15

3.1 Fluorescence anisotropy 15

3.2 Homo-FRET and time-resolved anisotropy 16

3.3 Homo-FRET and steady-state anisotropy 17

3.4 Fractional photobleaching and steady-state anisotropy 18

3.5 Molecular diffusion 21

4 Methods 22

4.1 Plasmid design and construction 22

4.2 Protein expression and purification 31

4.3 Biochemical protein analysis 35

4.4 Fluorescence measurements 43

4.5 Mass Spectrometry 49

5 Results 51

5.1 Fluorescence anisotropy of GFP fusion proteins 51

5.2 Determining the oligomeric state of non-covalently linked GFP fusion proteins via fluorescence anisotropy 67

5.3 GFP photobleaching and photodegradation 81
6 Discussion 91
6.1 Characterization of protein stoichiometry via fluorescence anisotropy 91
6.2 Photodegradation of green fluorescent protein 99
6.3 Outlook 110
7 References 111
8 List of Figures 123
9 List of Tables 126
10 List of Abbreviations 127
11 Appendix 132
11.1 Amino acid sequences of proteins in this thesis 132
11.2 Raw data and data from collaboration partners 135
About the author 142

1 Introduction

1.1 Studying protein stoichiometry via homo-FRET

1.1.1 Protein oligomerization

Most of the membrane-bound proteins in cells are associated as symmetrical oligomeric complexes with two or more subunits that are involved in signal transduction, the regulation of gene expression, antigen-antibody binding, and various other processes [1, 2]. Around 78% of oligomers from 2 to 12 subunits in *Escherichia coli* exist as homooligomers, protein assemblies of identical subunits [1].

The evolutionary driving force for homooligomer formation is assumed to lie in minimizing translational errors, in a most effective way to genetically "store" the structural information of proteins on DNA level, and in a dynamic spatial and temporal regulation of protein function [1].

The strength and the duration of oligomer formation can vary. There are proteins that are solely being found in their associated state with dissociation constants in a nanomolar range [3, 4]. Other protein-protein interactions have a weak or transient character, mainly existing as monomers and dimers with dissociation constants in micromolar dimensions [3].

1.1.2 Coiled-coil protein domains

Numerous protein oligomers are based on the interaction of α-helical coiled-coil domains, including α- and β-tubulins, the D0 domain of flagellins, the G protein subunit alpha I3, or members of the heat shock protein (Hsp70) family [5]. Coiled-coil domains are characterized by their repetitive sequence pattern, a heptad motif (*abcdefg*) with hydrophobic residues on position *a* and *d* which spans the α-helix (3.6 residues per turn) in a variable number of repeats [6, 7]. The contact interface between the hydrophobic residues of at least two coiled-coil domains enhances a strong protein-protein association, further supported by electrostatic interactions of adjacent residues at heptad position *e* and *g* [8, 9].

The most thoroughly studied coiled-coil domain is GCN4-p1. It is responsible for the dimerization of the leucine-zipper GCN4, a transcription factor in yeast (**Figure 1A**) [6].

GCN4-p1 is forming strong dimers with a very low dissociation constant of 8 nM [10, 11]. Residue variations at the hydrophobic interface of GCN4-p1 (**Figure 1B**) alter the protein stoichiometry: isoleucine residues at position *a* and *d* instead of leucine yield trimers (GCN4-pII) while leucine residues at *a* and isoleucine residues at *d* lead to a tetrameric complex (GCN4-pLI) [12].

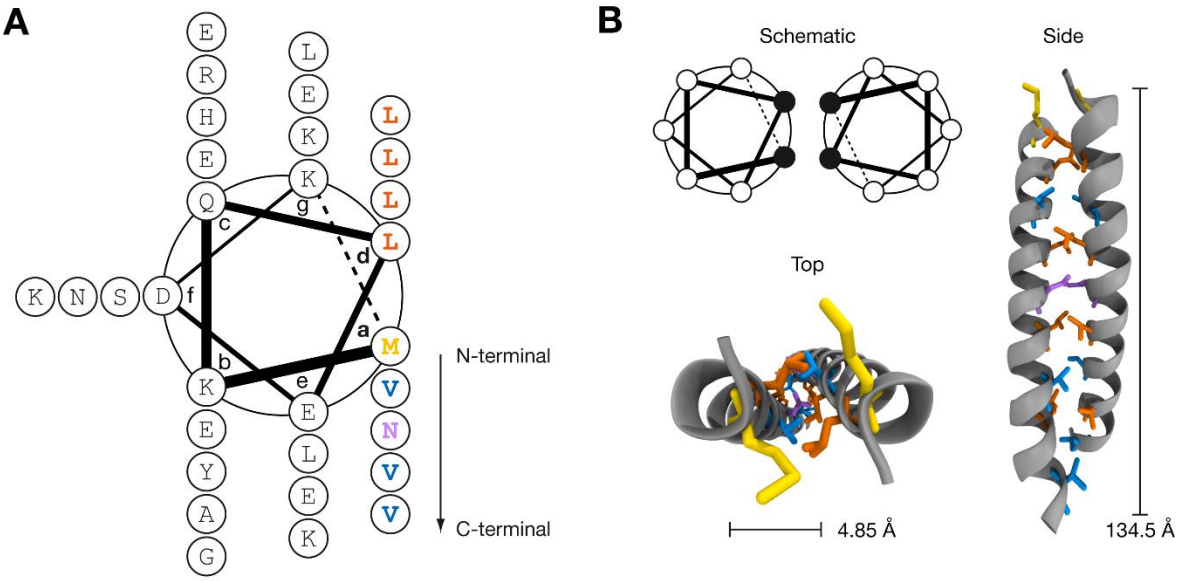


Figure 1: Schematic structure of GCN4-p1, the most thoroughly studied coiled-coil peptide. A) Helical wheel projection of α-helical GCN4-p1. The interface forming residues at heptad position *a* and *d* are highlighted in residue-specific colors (Met: yellow; Leu: orange; Val: blue; Asn: purple). **B)** Schematic overview of the GCN4-p1 interface. Hydrophobic side-chains form the characteristic zipper-like structure. Next to the schematic peptide structure, the coiled-coil radius (distance between the center of the coiled-coil and the center of the α-helix) and the pitch (length of the interacting helices) are indicated [13].

Over the past two decades, the Woolfson laboratory generated a group of *de novo* coiled-coil peptides with various oligomeric states (CC-Di, CC-Tri, CC-Tet, CC-pent, etc.) [14, 15]. Besides this group of *de novo* coiled-coil oligomers, Woolfson et al. further presented the publicly accessible tool CCbuilder to create coiled-coils from sequence-to-structure [16, 17]. The stoichiometry of several of these *de novo* peptides was experimentally examined as their coiled-coil sequences were integrated in fusion proteins [18]. In their study, the majority of the oligomeric states of coiled-coil sequences in chimeric proteins surprisingly did not coincide with the crystal structures of their isolated form. Therefore, it is essential to experimentally evaluate the stoichiometry of fusion protein containing embedded coiled-coil sequences.

1.1.3 Techniques resolving protein stoichiometry

There are numerous techniques to investigate the stoichiometry of proteins. Sedimentation equilibrium- and sedimentation velocity analytical ultracentrifugation are two of the oldest methods to determine the mass and the stoichiometry of soluble macromolecular assemblies [19-22]. They are both technically demanding and require a considerable extent of expertise [23]. Size-exclusion chromatography (SEC) or gel filtration is another common technique that enables the separation of oligomeric protein complexes in solution by their size [12, 18, 24]. By using native polyacrylamide gel electrophoresis (PAGE), assemblies of folded and associated proteins can be separated, too, as proteins migrate through a polyacrylamide gel to the anode solely driven by their intrinsic charge (colorless-native) or by the charge of an associated dye, i.e. Coomassie Brilliant Blue-G250 (blue-native) [25-27]. Other approaches involve chemical crosslinking, for instance via copper(II) (1,10-phenanthroline) [23, 28, 29], dynamic light- and X-ray scattering [30-32] and the determination of high-resolution structures, namely by nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography [33-37].

1.1.4 Fluorescence-based methods, homo-FRET, and fluorescence anisotropy

A further alternative to investigate the stoichiometry of interacting proteins are Förster resonance energy transfer (FRET) based methods where the complex-forming proteins have to be labeled with fluorophores of overlapping emission/absorption spectra [38-43]. Only if the donor and the acceptor fluorophore are in spatial proximity (< 10 nm) FRET can occur [44]. By the beginning of the 1990s, the usage of FRET-based techniques to resolve protein-protein interactions was declining as high-resolution structure elucidation became more and more significant. With the expansive adoption of fluorescent proteins into biochemical practice, however, FRET-based techniques were revived [40]. This could be mainly ascribed to an effective "*en passant*" labeling of proteins by expressing them in fusion proteins together with a fluorescent protein tag and to new possibilities to observe protein-protein interaction *in vivo*.

For characterization of the oligomeric state of a protein cluster, FRET between identical fluorophores (homo-FRET) can be applied. The usage of homo-FRET facilitates the labeling process as the cluster forming proteins have to be labeled with only one kind of fluorophore.

Contrasting classical FRET between heterogeneous fluorophores, homo-FRET can be detected by exploiting differences in the fluorescence anisotropy (**Figure 2**). The fluorescence anisotropy is defined by the extent of the emission that is polarized parallel and perpendicular to the excitation direction. The degree to which the emission is depolarized depends on the transition dipole moment change between excitation and emission, thus on molecular rotation [45] and the number of fluorophores interacting via homo-FRET [46]. With an increased number of interacting fluorophores, the fluorescence anisotropy decreases [46]. The fluorescence anisotropy can be quantitatively determined by resolving its decay after excitation at a nanosecond time scale, referred to as time-resolved anisotropy, or by integrating parallel- and perpendicular-polarized emission for a way longer time interval of hundreds of milliseconds, called steady-state anisotropy [47]. Time-resolved anisotropy analysis further yields insight in the rotational correlation times of the observed fluorophore, as well in homotransfer rates under certain circumstances [48].

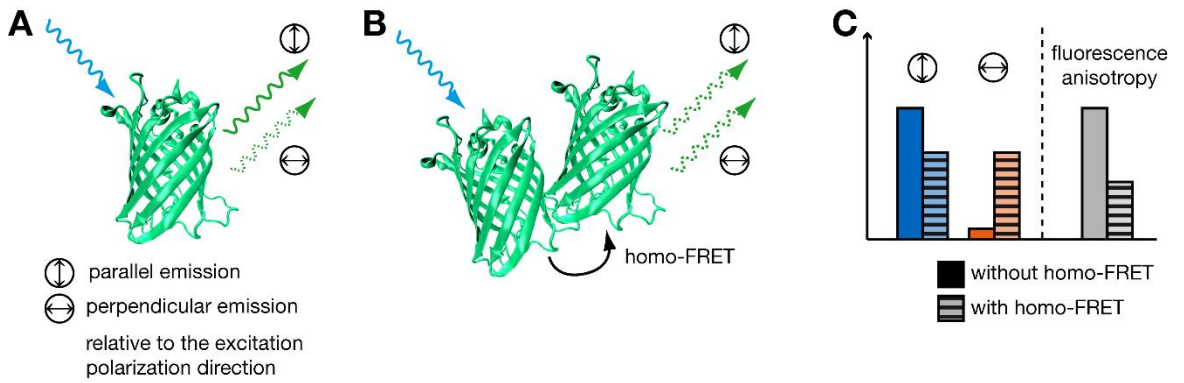


Figure 2: Schematic representation of homo-FRET between two fluorescent proteins and its effect on steady-state anisotropy. A) A fluorescent protein is excited with polarized light (at excitation wavelength λ_{ex}; blue arrow). After being excited, the fluorophore emits light polarized parallel and perpendicular to the excitation polarization direction (at emission wavelength λ_{em}; green arrows). In the absence of homo-FRET, the extent of polarized emission perpendicular to the excitation light (↔) is only depending on the orientation change due to molecular rotation since the depolarization correlates with the transition dipole moment difference between excitation and emission. **B)** If two proximal fluorophores of the same kind interact via homo-FRET, the extent of perpendicular (↔) polarized emitted light is increased. **C)** The ratio between parallel (↕) or perpendicular (↔) polarization orientation of the emitted light is commonly expressed as fluorescence anisotropy. The fluorescence anisotropy is a reporter for homo-FRET. It decreases with the number of fluorophores interacting via homo-FRET [46] and can be determined experimentally.

There were plenty of applications in the past where homo-FRET and fluorescence anisotropy helped to characterize protein oligomerization. In several *in vivo* studies, fluorescence anisotropy microscopy contributed to quantify clusters of glycosylphosphatidylinositol-anchored proteins [49-51]. In other publications, a flow-cytometric method using homo-

FRET characterized the epidermal growth factor receptor as a homodimer with a dynamic transition to higher-order complexes [52]. In an *in vitro* approach, homo-FRET was used to classify the number of fluorescent proteins artificially assembled by an oligonucleotide framework [53, 54]. Furthermore, recording the behavior of fluorescence anisotropy upon progressive photobleaching could determine the formation of higher-order complexes of the membrane-integrated serotonin$_{1A}$ receptor, tested in Chinese hamster ovary cells [55]. More recently, homo-FRET and anisotropy helped to characterize the stoichiometry of lysozyme, labeled with a small organic dye, associated on large unilamellar vesicles [56].

Evaluating protein oligomerization via fluorescence anisotropy can be methodically achieved by time-resolved anisotropy and by steady-state anisotropy analysis. In an intriguing work, Runnels and Scarlata delivered a fundamental model to resolve the stoichiometry of fluorophore-labeled, cluster forming proteins by fluorescence anisotropy [46]. Their model comprises several parameters that can only be determined via time-resolved anisotropy. Unfortunately, this requires measurements with sophisticated instrumentation. To circumvent time-resolved analysis, an alternative approach has been proposed by Yeow and Clayton [57]. It uses the steady-state anisotropy of fractionally fluorophore-labeled proteins to resolve their stoichiometry (**Figure 3A**). In the past, this theoretical framework has been used to characterize the oligomeric state of membrane-associated proteins, for instance lysozyme on the surface of large unilamellar vesicles [56], and integral membrane proteins, such as serotonin$_{1A}$ receptors [55] and epidermal growth factor receptors [58].

Both approaches, the one by Runnels and Scarlata and the one by Yeow and Clayton, can be very useful in order to systematically determine the oligomeric state of fluorescently labelled fusion proteins. Hence, we want to test both theoretical frameworks for their applicability. For this purpose, we use oligomerizing coiled-coil sequences integrated in chimeric proteins, together with a fluorescent protein domain (**Figure 3B**). This will deliver insights into the methods' resolving power, its limitations, or a possible necessity for calibration which then can be translated in order to resolve other proteins' stoichiometry.

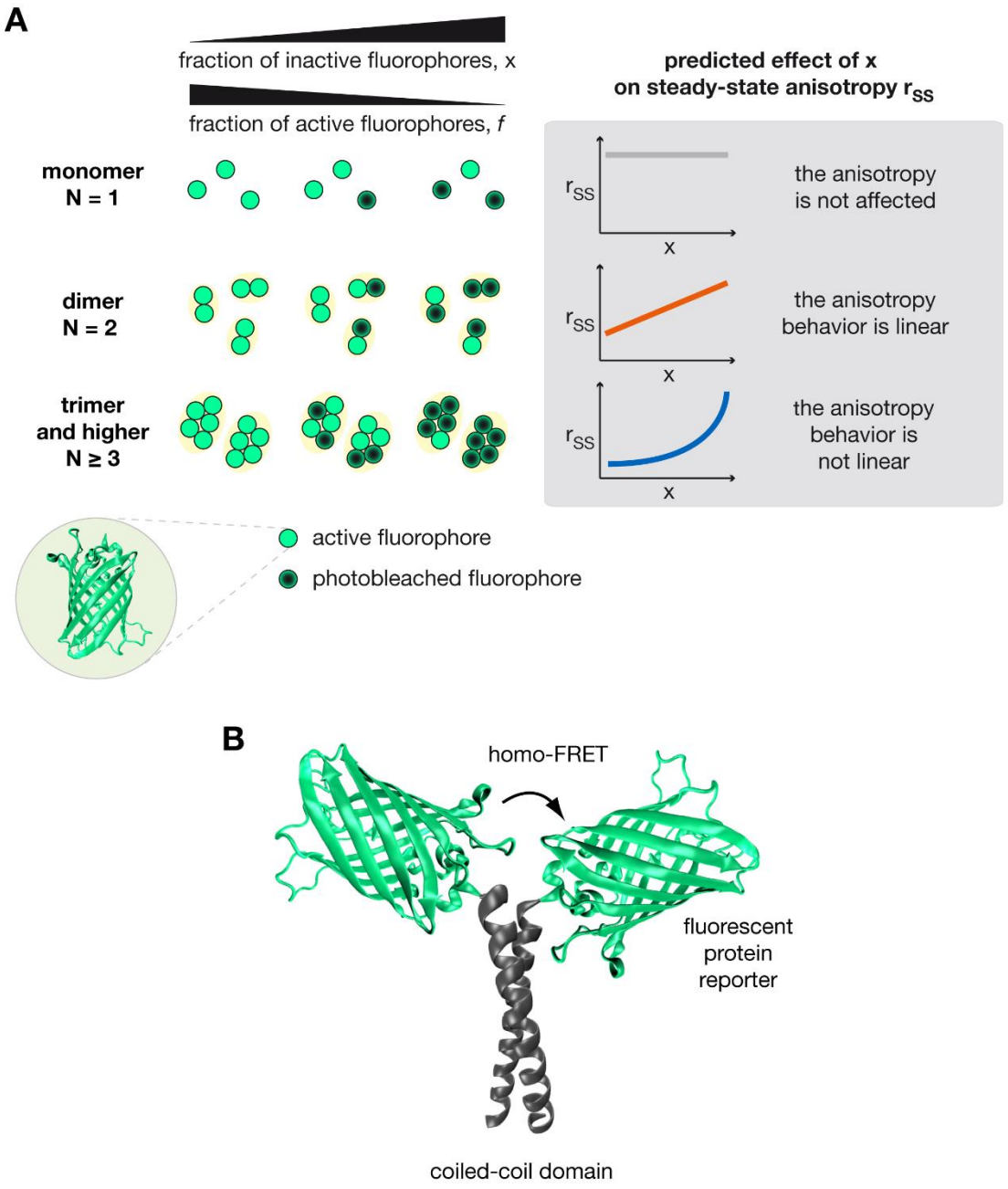


Figure 3: Characterizing the oligomeric state of fluorophore labeled protein clusters from steady-state anisotropy r_{SS} and fractional photobleaching. A) The steady-state anisotropy response for fractionally photobleached cluster of fluorophores provides an insight into protein stoichiometry without the extensive knowledge of photophysical parameters. Fractionally photobleaching a complex of fluorophores that interact via homo-FRET can be perceived as gradually removing participants from the complex. According to theoretical background by Yeow and Clayton [57], the steady-state anisotropy behavior is dependent on the stoichiometry of fluorophore-labeled proteins. For an increased fraction of inactive fluorophores x, the steady-state anisotropy r_{SS} behaves stationary for monomers, linear for dimers, and exponential for trimers and higher oligomers. **B)** Schematic concept of a fusion protein reporter system to characterize coiled-coil stoichiometry. The labeling of cluster-forming proteins with a fluorophore is achieved by linking the coiled-coil sequence covalently to a fluorescent protein domain in a chimeric protein. When coiled-coil domains assemble, fluorescent reporters are in proximity and interact via homo-FRET which alters the polarization of the emitted light. Thus, the measured steady-state anisotropy decreases with the number of interacting fluorophores N [46].

1.2 Green fluorescent protein – an appropriate fluorophore for a homo-FRET reporter system

Green fluorescent protein (GFP) is a very common fluorescent probe used not only in numerous biochemical assays but also in homo-FRET reporter systems [49, 51, 59].

The fluorophore of GFP is formed spontaneously by protein folding which is its most striking feature [47, 60]. Once being expressed in host cells, GFP matures in three distinct kinetic steps: (i) relatively slow protein folding ($t_{½}$ = 10 min), (ii) formation of imidazolinone (cyclization) by the nucleophilic attack of the amide of Gly67 on the carbonyl of residue 65 ($t_{½}$ = 3 min), (iii) and oxidation of the chromophore ($t_{½}$ > 19 min) (**Figure 4**) [61, 62].

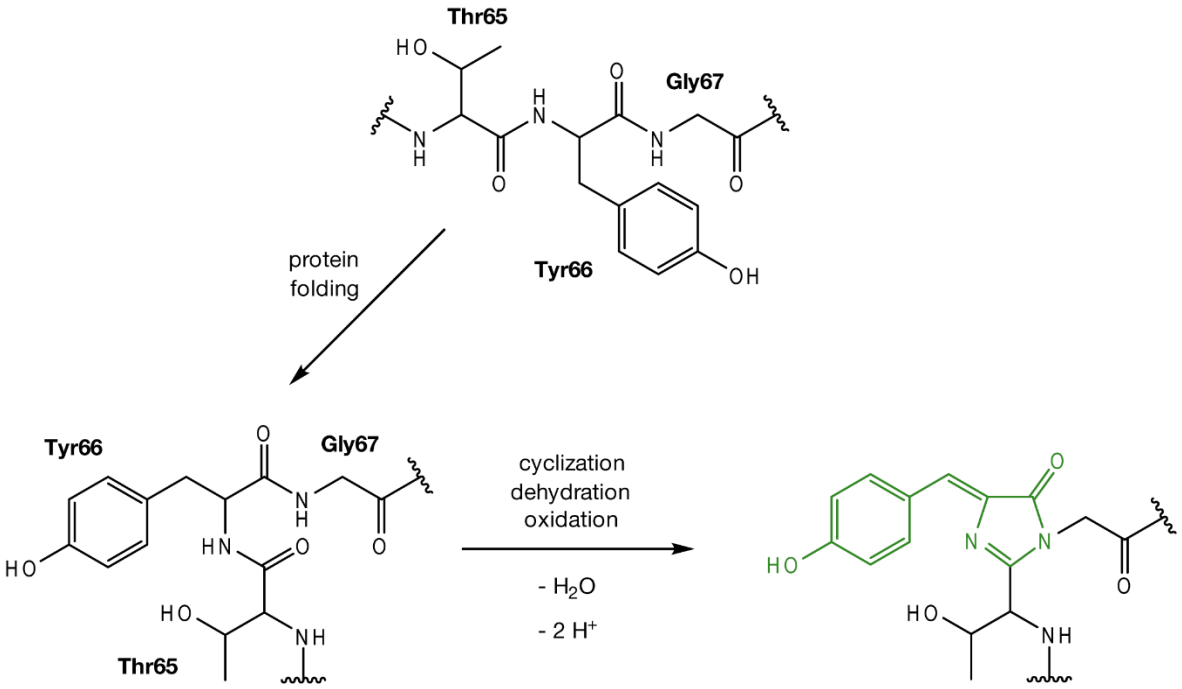


Figure 4: Maturation of GFP consists of protein folding, forming of imidazolinone, and oxidation. The maturation of superfolder GFP (sfGFP) is depicted as it has a threonine residue at position 65. The three neighboring residues Thr65, Tyr66, and Gly67 are rearranged during protein folding. Nucleophilic attack of the amide of Gly67 on the carbonyl of Thr65 leads to cyclization with a subsequent dehydration. Only in the presence of molecular oxygen, the fluorophore can finally mature by oxidation [63]. Thereby, the chromophore, a π-conjugated system (green), is formed between the phenol group of Tyr66, the linking methylene bridge, and the emerged imidazolinone.

1.2.1 The variety of fluorescent proteins

GFP was first found in extracts from the jellyfish *Aequorea victoria* in 1962 [64]. However, little is known about its biological function. Fluorescent proteins might have a photoprotective role for cnidarian organisms such as corals or jellyfishes [65]. More than 30 years after the first characterization on a protein level, Chalfie et al. could express functional GFP in *Escherichia coli* [60], enabling the application of GFP as a fluorescent probe to a broad scientific field [66-69].

Drawing enormous scientific interest in the past decades, GFP has been adjusted to experimental demands by mutating several residues. A brighter enhanced variant of GFP (EGFP) was developed changing Phe64 to Leucine and Ser65 to Threonine [70-72]. F64L and S65T mutations led to better protein solubility and an increased fluorescence intensity with an excitation peak at 488 nm and the emission maximum at 511 nm (**Figure 5**) [70, 73, 74]. Based on EGFP and wild-type GFP, numerous mutants have been established: A photoactivatable variant of GFP, pa-GFP [75, 76], a cysteine-free version of GFP, cfGFP, that is not affected by oxidative pressure [77], or red-shifted versions including the popular variant Venus [78, 79] are only selected examples among 110 descendants of wild-type GFP [80].

By altering the GFP sequence on several positions, Pédelacq et al. could develop a superfolder variant of GFP (sfGFP) that reduces the maturation half time by half compared to wtGFP [81, 82]. sfGFP is differing from EGFP in ten positions (S30R, Y39N, R80Q, F99S, N105T, Y145F, M153T, V163A, I171V, and A206V) (**Figure 5B**) and folds even in fusion with poorly folding and interfering peptides. Most prominently, the mutation of A206V disrupts the hydrophobic patch formed by Ala206, Leu221, and Phe223 [83], and thereby reduces the intrinsic dimerization of GFP [81].

The majority of the constructs in this thesis hold sfGFP as the fluorescent domain. As another fluorescent reporter, some of the fusion proteins hold an EGFP variant by Stauber et al. [84]. This EGFP version differs from the common variant by Cormack et al. in T65C and I167T and has its excitation and emission peaks at 479 nm and 507 nm, respectively [74].

```
  GFP (Chalfie)      1    MGKGEELFTG VVPILVELDG DVNGHKFSVS GEGEGDATYG KLTLKFICTT GKLPVPWPTL   60
 EGFP (Stauber)      1    MSKGEELFTG VVPILVELDG DVNGHKFSVS GEGEGDATYG KLTLKFICTT GKLPVPWPTL   60
sfGFP (Pedelacq)     1    MSKGEELFTG VVPILVELDG DVNGHKFSVR GEGEGDATNG KLTLKFICTT GKLPVPWPTL   60

  GFP (Chalfie)     61    VTTFSYGVQC FSRYPDHMKR HDFFKSAMPE GYVQERTIFF KDDGNYKTRA EVKFEGDTLV  120
 EGFP (Stauber)     61    VTTLCYGVQC FSRYPDHMKR HDFFKSAMPE GYVQERTIFF KDDGNYKTRA EVKFEGDTLV  120
sfGFP (Pedelacq)    61    VTTLTYGVQC FSRYPDHMKQ HDFFKSAMPE GYVQERTISF KDDGTYKTRA EVKFEGDTLV  120

  GFP (Chalfie)    121    NRIELKGIDF KEDGNILGHK LEYNYNSHNV YIMADKQKNG IKVNFKIRHN IEDGSVQLAD  180
 EGFP (Stauber)    121    NRIELKGIDF KEDGNILGHK LEYNYNSHNV YIMADKQKNG IKVNFKTRHN IEDGSVQLAD  180
sfGFP (Pedelacq)   121    NRIELKGIDF KEDGNILGHK LEYNFNSHNV YITADKQKNG IKANFKIRHN VEDGSVQLAD  180

  GFP (Chalfie)    181    HYQQNTPIGD GPVLLPDNHY LSTQSALSKD PNEKRDHMVL LEFVTAAGIT HGMDELYK    238
 EGFP (Stauber)    181    HYQQNTPIGD GPVLLPDNHY LSTQSALSKD PNEKRDHMVL LEFVTAAGIT HGMDELYK    238
sfGFP (Pedelacq)   181    HYQQNTPIGD GPVLLPDNHY LSTQSVLSKD PNEKRDHMVL LEFVTAAGIT HGMDELYK    238
```

Figure 5: Sequence alignment of GFP [60], EGFP [84], and sfGFP [81]. Differing residue positions among the aligned proteins are marked red. The underlined three residues are forming the chromophore. Grey areas represent sequences which fold into beta sheet structures.

After introducing GFP to a field outside the narrow niche of marine organisms, fluorescent proteins of different colors were derived from *Anthozoa* to be used in molecular biology, most prominently DsRed and "mFruit" variants [85-87]. To date, there are over 490 published fluorescent reporter proteins differing in structure, brightness, maturation, and decisively excitation and emission wavelength [80]. Color modification in fluorescent proteins is mainly achieved by changing the extent of the π-systems or by different protonation states of the chromophore [63, 69].

1.2.2 The application of GFP in a homo-FRET reporter system

Nevertheless, out of the diverse color palette of fluorescence proteins only a handful are commonly used in molecular biology and biochemistry, four presented and compared in **Table 1**.

In the past, other groups used concatenated Venus concatemers as standards for protein complexes labeled with fluorescent proteins [88, 89] or concatenated EGFP domains as standards for intracellular fluid dynamics [90]. Alternatively, a common reporter for gene expression mCherry could also be considered as a possible concatemer basis.

In this study, we decided to use sfGFP because it outcompetes Venus, EGFP, and mCherry as the appropriate fluorescent reporter. The yellowish GFP variant Venus has a small Stokes shift of 13 nm which is beneficial for an efficient homo-FRET resonation. However, it is outweighed by sfGFP in Förster distance and quantum yield [79, 80, 89, 91]. Compared to sfGFP, EGFP is in an inferior position particularly because of its weak brightness. mCherry

has a poor quantum yield (QY), an equally weak brightness, a larger Förster distance, and a long maturation time [80, 92]. It tends to be inappropriate for biophysical studies while its qualities may lie in the cell biological assays.

Table 1: Benchmark parameters for commonly used fluorescent proteins in molecular biology

	λ_{ex} (nm) [1]	λ_{em} (nm) [2]	Stokes shift (nm)	R_0 (nm) [3]	EC ($cm^{-1}M^{-1}$) [4]	QY [5]	Brightness	$mt_{½}$ (min) [6]
EGFP	488	507	19	4.6	55,900	0.6	33.54	14.5
sfGFP	485	510	25	4.6	83,300	0.65	54.15	13.6
Venus	515	528	13	5.3	92,200	0.57	52.55	17.6
mCherry	587	610	23	5.1	72,000	0.22	15.84	37

[1] Excitation wavelength. [2] Emission wavelength. [3] Förster radius / distance. [4] Molar attenuation / extinction coefficient. [5] Fluorescence quantum yield. [6] Maturation half-life.

1.2.3 The irreversible photobleaching of fluorescent proteins and its relation to photodegradation

Yeow and Clayton presented a method that is based on homo-FRET and requires the fractional and irreversible photobleaching of the fluorescent probe such as GFP (**Figure 3A**) [57]. In stark contrast to organic fluorophores, the photobleaching process of fluorescent proteins, particularly GFP, is still poorly studied [69]. While most of microscopy related techniques necessarily rely on fluorescent proteins that are robust against photobleaching, several techniques utilize the reversible ("blinking") and irreversible loss of fluorescence of fluorescent proteins [57, 93-96]. Reversible photobleaching of GFP is neither depending on the solution's oxygen content, nor on triplet state quenchers [97] but on cis-trans photoisomerization with altered protonation states [98-100].

The mechanistic and photophysical understanding of irreversible photobleaching of fluorescent proteins, however, is rudimentary. At present, irreversible photobleaching is presumed to be caused by the permanent structural change or deformation of the chromophore [69].

In several studies on the photoswitchable fluorescent protein IrisFP, Adam et al. found that photobleaching of fluorescent proteins and formation of radicals at the fluorophore are

associated [101-104]. Fluorophores that were irradiated with X-rays or high-intensity visible light exhibited "photodamage" at the C_α-C_5 π-bond of the methylene bridge which could be detected with Raman spectroscopy [101]. Strong illumination results in a radical intermediate state leading to an irreversible structural deformation, i.e. the decarboxylation of vicinal Glu212 and the distortion of the former phenol group, now sp^3-hybridized (**Figure 6**) [69, 103].

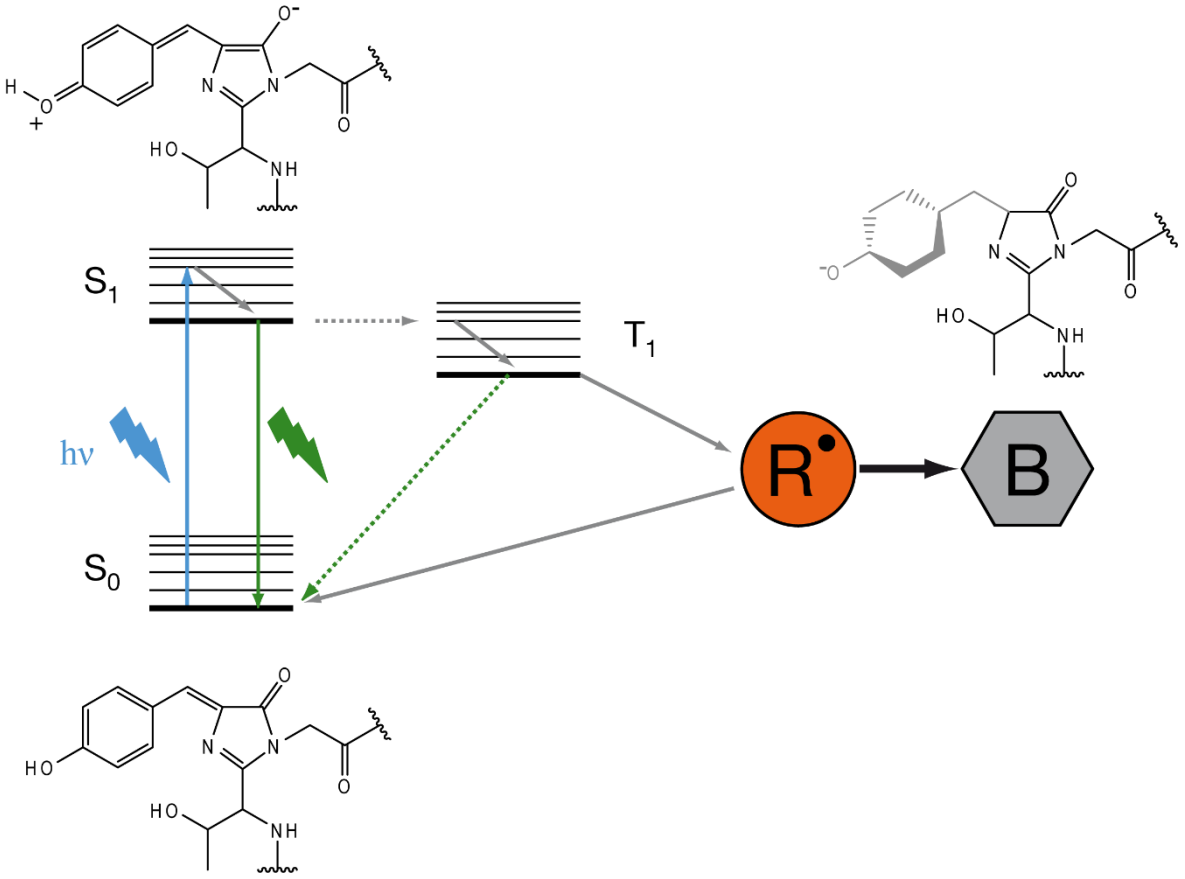


Figure 6: The proposed mechanism of the irreversible photobleaching of fluorescent proteins according to Adam et al. [101, 103]. By absorbing a high-energy photon, a fluorophore electron is brought from the ground (S_0) to an excited singlet state (S_1). As the electron relaxes to the ground state, a photon is emitted. After the excitation to S_1, a subsequent conversion to an excited triplet state T_1 and the relaxation from T_1 to S_0 is possible. Findings by Adam et al. show that fluorescent proteins might undergo radical formation ($R^\bullet$) when they are heavily illuminated over a certain time period. These fluorescent proteins partly exhibited "photodamage" at the C_α-C_5 π-bond of the methylene bridge and a distorted former phenol group. The protein cannot be excited anymore and irreversibly lost its fluorescence.

Massive structural deformation as a result of strong energetic irradiation is not exceptional for fluorescent proteins as there are several photoswitchable variants that undergo backbone breakage after near-UV irradiation. Most prominently, Kaede [105-107], EosFP [108] and IrisFP [109] exhibit fragmentation near the chromophore which leads to altered π-systems

and red-shifted emission spectra in their "photoswitched" form. For GFP in particular, backbone fragmentation has only been detected for a very exotic variant with fluorophore residues Ala65, Ser66, and Gly67 by Barondeau et al. [110] (**Figure 7**).

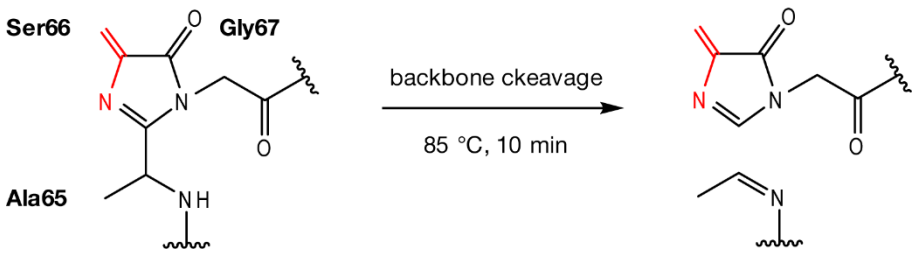


Figure 7: The proposed mechanism of GFP-Ala65-Ser66-Gly67 fragmentation according to Barondeau et al. [110]. The GFP variant GFP-Ala65-Ser66-Gly67 was incubated at 85 °C for 10 minutes. In denaturing polyacrylamide gel electrophoresis, fragments of 7 kDa and 20 kDa could be detected which indicated a fragmentation at the fluorophore. Barondeau et al. suggest an imine-enamine tautomerization (red marked region) and a final ejection of water as the reason for the backbone cleavage.

After heating the GFP variant, they could detect two different sized peptide fragments with approximately 7 kDa and 20 kDa in denaturing polyacrylamide gel electrophoresis. Their proposed mechanism comprises imine-enamine tautomerization (**Figure 7**). Barondeau et al. assumed that a similar tautomerization in (fluorescent) versions of GFP that hold an aromatic side-chain at the fluorophore center (sfGFP: Tyr66) would remove the connecting methylene bridge between the phenol and the imidazolinone. Due to the high energetic costs, they suggest that a backbone fragmentation, as they observed for the GFP-Ala65-Ser66-Gly67 variant, could not be possible for wild-type GFP or its conventional variants.

GFP can be used as a reporter protein in fluorescence anisotropy-based methods that involve the irreversible photobleaching of the fluorophore. Besides its undeniable advantages that were highlighted above, GFP and its relatives are apparently put in jeopardy of structural alteration when heavily irradiated or photobleached with high intensity. It is of interest whether fragmentation and backbone cleavage can be observed for ordinary GFP variants, such as EGFP and sfGFP, and how this affects the homo-FRET photobleaching approach.

2 Motivation

Anisotropy-based methods help to characterize the stoichiometry of fluorescently labeled protein assemblies. This could already be demonstrated for individual proteins in the past, mostly by using sophisticated equipment such as fluorescence microscopes and single-molecule resolving set-ups.

In this thesis, we wanted to test the applicability of anisotropy-based approaches to systematically study the oligomeric state of cluster-forming proteins. As a test system of choice, we intended to resolve the stoichiometry of oligomerizing coiled-coil proteins. To extend the application for these approaches, a novel methodic framework had to be developed, optimally requiring relatively simple equipment only.

As a fluorescent label, we intended to use the superfolder variant of GFP that is known to mature rapidly, to be bright, and to have a good quantum yield. Before testing non-covalently assembled protein complexes, a series of fusion proteins with concatenated sfGFP was to give first insights whether it is possible to differentiate between multiple sfGFP stoichiometries via fluorescence anisotropy. sfGFP concatemers could serve as model proteins to define important biophysical parameters and to elicit resolution limits for both, time-resolved and steady-state anisotropy. Moreover, it had to be tested whether existing theoretical models could interpret the observations made in our experiments correctly.

After clarifying these open questions, the stoichiometry of coiled-coil peptides was to be investigated via fluorescence anisotropy. Therefore, fusion proteins containing GFP and coiled-coil sequences had to be generated and analyzed. Besides classical biochemical approaches, this analysis had to comprise fluorescence anisotropy measurements with sophisticated and relatively simple instrumentation. At best, an oligomeric state N was to be determined from the fluorescence anisotropy. This could possibly be achieved with a combination of recording the steady-state anisotropy and fractionally photobleaching the GFP moiety.

In a parallel study, it had to be evaluated whether a fractionally photobleached GFP exhibits “photodamage” that was observed in the past for heavily irradiated fluorescent proteins. If GFP was affected by high-intensity illumination, the structural consequences had to be

specified. Furthermore, this could shed light on the poorly understood phenomenon of irreversible photobleaching of fluorescent proteins.

Conclusively, this study aimed to be a blue print of generating a test system for the investigation of small oligomerizing peptide sequences in fusion proteins, exploiting GFP reporters via homo-FRET. Beyond coiled-coil oligomerization, we intended this thesis to be the cornerstone for further projects where the stoichiometry characterization of – for instance – complex-forming transmembrane domains could be achieved.

3 Theory

3.1 Fluorescence anisotropy

Fluorescence anisotropy r is based on the principle of photoselective excitation of fluorophores by polarized light [47]. Fluorophores with an absorption transition dipole moment oriented parallel to the polarized exciting light are photoselected. The resulting fluorescence emission also occurs with a fixed transition dipole moment. The relative angle between these transition dipole moments determines the measured anisotropy.

Fluorescence anisotropy is maximal (r = 0.4) if the emission transition dipole moment is oriented parallel to the absorption dipole moment. With angles of 54.7° or 90° between absorption and emission dipole moment, anisotropy is decreased to r = 0 or to minimally r = -0.2, respectively [47]. The relative angle between the transition dipole moments is also a function of how fast a fluorophore rotates during its fluorescence lifetime. This relation can be expressed by the Perrin equation for anisotropy as a function of fluorescence lifetime [111]:

$$r(\tau) = \frac{r_0}{1 + \tau \cdot \phi^{-1}} \qquad (1)$$

where r_0 is the intrinsic anisotropy of the fluorophore, τ is the fluorescence lifetime, and ϕ is the rotational correlation time. The rotational correlation time is defined by the time the fluorophore needs to rotate 1 rad = 57.3°. The intrinsic anisotropy can maximally be 0.4. Small and fast tumbling fluorophores with short rotational correlation times experience a larger change in transition dipole moments and thus a stronger depolarization. Anisotropy is decreased. For bigger, slowly rotating fluorophores, such as the fluorescent proteins, ϕ is much longer than τ.

Experimentally, fluorescence anisotropy can be determined for fluorophores in the steady state. The steady-state anisotropy r_{SS} is defined by:

$$r_{SS} = \frac{I_{par} - G \cdot I_{per}}{I_{par} + 2 \cdot G \cdot I_{per}} \qquad (2)$$

Where I_{par} and I_{per} are the fluorescence intensities measured at parallel or perpendicular orientations relative to the polarized excitation. G is a calibration factor and describes the ratio of detection sensitivity for parallel and perpendicular polarized light.

3.2 Homo-FRET and time-resolved anisotropy

Besides rotational diffusion of the fluorophores, fluorescence anisotropy is affected by the extent of homo-FRET between interacting fluorophores. Homo-FRET is depending on the spatial orientation and the distance R between donor and acceptor fluorophore. If the distance between two resonating fluorophores equals the Förster distance R_0, homo-FRET is half-maximal. Förster distances can vary significantly. For GFP-GFP homotransfer, the Förster distance is R_0 = 4.6 nm [49, 112]. Homo-FRET can be quantitatively described by the homotransfer rate k_{FRET}:

$$k_{FRET} = \frac{1}{\tau} \cdot \left(\frac{R_0}{R}\right)^6 \tag{3}$$

where R_0 is the Förster distance and R is the distance between interacting fluorophores.

Additionally, k_{FRET} is a function of the homo-FRET efficiency E [113]. It is then defined by:

$$k_{FRET} = \frac{1}{\tau} \cdot \frac{E}{(1-E)} \tag{4}$$

The homotransfer rate k_{FRET} can be determined from time-resolved anisotropy analysis. Observing the anisotropy decay of a single molecule on a nanosecond time scale allows the discrimination between the effect of fluorophore rotation and the effect of homo-FRET. The anisotropy decay in the absence of homo-FRET is influenced by the molecule size and the viscosity of the environment. Then, the decay for a molecule without any rotational restrictions is defined as:

$$r(t) = r_0 \cdot e^{\left(\frac{-t}{\phi}\right)} \tag{5}$$

In this context, the intrinsic anisotropy r_0 is often referred to as the maximal initial, and thus limiting, anisotropy after excitation.

In most of the cases, $r(t)$ cannot be described as a simple single-exponential function, seen in eq. *(5)*, but as a multi-exponential decay [47]. Then, time-resolved anisotropy is a sum of multiple decays, each classified by a fractional amplitude g and the basic structure of eq. *(5)*. As practical approaches for GFP have shown [97], time-resolved data for monomeric GFP in the absence of homo-FRET is ideally understood with a two-rotator anisotropy model. Here, the anisotropy decay is affected by a slow component (with ϕ_{slow}) and a fast component (with ϕ_{fast} < 2 ns), both classified by the fractional amplitude g:

$$r(t) = r_0 \cdot (1 - g) \cdot e^{\left(\frac{-t}{\phi_{fast}}\right)} + g \cdot e^{\left(\frac{-t}{\phi_{slow}}\right)} \qquad (6)$$

For large and slowly rotating fluorescent proteins, such as GFP, rotational correlation time ϕ is much longer ($\phi_{GFP} \approx 20$ ns [97]) than its fluorescence lifetime τ ($\tau_{GFP} \approx 2.6$ ns). If multiple GFPs interact via homo-FRET, the rotational correlation time of the whole complex is even increased. In this particular case, the rotational effects can be disregarded for simplicity. Then, time-resolved anisotropy decay is solely dependent on the extent of homo-FRET ("hindered rotor") [47, 48, 51, 114] and $r(t)$ can be written as:

$$r(t) = (r_0 - r_\infty) \cdot e^{(-2 \cdot k_{FRET} \cdot t)} + r_0 \qquad (7)$$

where r_∞ is the anisotropy limit for t approaching infinity.

3.3 Homo-FRET and steady-state anisotropy

In a complex of proximal fluorophores that interact via homo-FRET, the steady-state anisotropy r_{SS} can be understood as a combination of both, the intrinsic anisotropy of the initially excited fluorophore r_1 and the anisotropy of those fluorophores that are indirectly excited via homo-FRET r_{ET} [46, 51].

In case of GFP as the fluorophore, one can assume that $r_1 \approx 0.3$, the steady-state anisotropy of monomeric GFP in aqueous solution [90, 115]. The anisotropy of the indirectly excited fluorophores, r_{ET}, is defined by their relative spatial orientation. The value of r_{ET} is close to zero [116] for small, randomly oriented fluorophores such as fluorescein [46]. However, it becomes significantly higher for less randomly organized assemblies such as complex-forming fluorescent proteins [51].

According to Runnels and Scarlata [46], the steady-state anisotropy of a complex of N fluorophores that interact via homo-FRET can be defined as the sum of two terms, one incorporating r_1, the anisotropy of the donor, and the other r_{ET}, the anisotropy of the acceptors. It is inversely proportional to the number N of interacting subunits:

$$r_{SS}(N) = r_1 \cdot \frac{(1 + k_{FRET} \cdot \tau)}{(1 + N \cdot k_{FRET} \cdot \tau)} + r_{ET} \cdot \frac{(N-1) \cdot k_{FRET} \cdot \tau}{(1 + N \cdot k_{FRET} \cdot \tau)} \qquad (8)$$

Practically, the number of interacting fluorophores N can be calculated by experimentally measuring r_{SS}, determining k_{FRET} from time-resolved anisotropy decay, and obtaining τ from the fluorescence intensity decay of a single fluorophore.

3.4 Fractional photobleaching and steady-state anisotropy

Yeow and Clayton [57] presented an approach to determine the oligomeric state N of interacting proteins by fractionally labeling the cluster with fluorophores that interact via homo-FRET. Experimentally, this requires the continuous determination of steady-state anisotropy while the fraction of labeled oligomers f is gradually increased. A similar result can be achieved by gradually removing labeled protein subunits from the complex via fractional photobleaching.

In this study, GFP reporters were fractionally photobleached. Therefore, Yeow and Clayton's formulas can be reformed by exchanging f with x, the fraction of photobleached, thus, inactive fluorophores:

$$x = 1 - f \qquad (9)$$

To determine the fraction of inactive fluorophores x experimentally, the fluorescence intensity after irradiation I_i at time point i is recorded and then related to the fluorescence intensity of a non-irradiated control I_0 at time point 0. The loss of active fluorophores due to irradiation x_i then is:

$$x_i = \frac{I_i}{I_0} \qquad (10)$$

When an assembly of fluorophores is fractionally photobleached, subunits that interact via homo-FRET are progressively removed and homotransfer is reduced. As a consequence, steady-state anisotropy r_{SS} increases to the point where most of the complexes exhibit only one active fluorophore and r_{SS} stagnates at monomer level (**Figure 3**).

In a first model, Yeow and Clayton described this behavior of r_{SS} upon fractional photobleaching as a polynomial expansion with coefficients derived from the $(N-1)$th row of the Pascal's triangle and with variables r_A, r_B, r_C, etc. yielded from an optimal fit:

$$N = 2 \qquad r_{SS}(x,2) = x \cdot r_A + (1-x) \cdot r_B$$

$$N = 3 \qquad r_{SS}(x,3) = x^2 \cdot r_A + (1-x) \cdot 2x \cdot r_B + (1-x)^2 \cdot r_C$$

$$N = 4 \qquad r_{SS}(x,4) = x^3 \cdot r_A + (1-x) \cdot 3x^2 \cdot r_B + (1-x)^2 \cdot 3x \cdot r_C + \\ +(1-x)^3 \cdot r_D \qquad (11)$$

$$N = 5 \qquad r_{SS}(x,5) = x^4 \cdot r_A + (1-x) \cdot 4x^3 \cdot r_B + (1-x)^2 \cdot 4x^2 \cdot r_C + \\ +(1-x)^3 \cdot 4x \cdot r_D + (1-x)^4 \cdot r_E$$

In another approach with the focus on heterogeneous oligomers and significant fraction of non-interacting fluorophores, r_{SS} can be described by a simple two-state bimodal model:

$$r_{SS}(x,N) = r_1 \cdot \left(f_{non} + (1 - f_{non}) \cdot x^{(N-1)}\right) = \\ = r_1 \cdot f_{non} + r_1 \cdot x^{(N-1)} - r_1 \cdot f_{non} \cdot x^{(N-1)} \qquad (12)$$

where r_1 is the steady-state anisotropy of the monomer and f_{non} is the fraction of non-interacting fluorophores. In the second line, the formula was rearranged for reasons of clarity: Here, the anterior part, $r_1 \cdot f_{non}$, describes the steady-state anisotropy for unbleached, thus, active fluorophores ($x = 0$). The succeeding part represents the change in anisotropy due to photobleaching ($0 < x < 1$).

This model is based mainly on geometrical considerations and becomes imprecise for fully assembled ($f_{non} = 0$) and non-bleached ($x = 0$) fluorophores. This inaccuracy is legitimate

as the formalism was created to describe the anisotropy behavior for fractional labeling (opposite effect, $x \rightarrow 0$) and heterogeneous oligomers with a significant monomer fraction. The most evident weakness of this formalism, however, is that it assumes that the steady-state anisotropy for unbleached samples ($x = 0$) is not depending on the oligomeric state N ($r_{SS}(0, N) = r_1 \cdot f_{non}$) which is contrary to physical reality [46].

Addressing this problem, I replaced $r_1 \cdot f_{non}$ in eq. *(12)* by r_{SS} from eq. *(8)*, which is Runnels and Scarlata's version to describe the steady-state anisotropy as a function of N subunits in a cluster, $r_{SS}(N)$. The anisotropy as a function of fractional photobleaching x and the oligomeric state N is then defined by:

$$\begin{aligned} r_{SS}(x, N) &= r_{SS}(N) + r_1 \cdot x^{(N-1)} - r_{SS}(N) \cdot x^{(N-1)}, \\ r_{SS}(N) &= r_1 \cdot \frac{(1 + k_{FRET} \cdot \tau)}{(1 + N \cdot k_{FRET} \cdot \tau)} + r_{ET} \cdot \frac{(N-1) \cdot k_{FRET} \cdot \tau}{(1 + N \cdot k_{FRET} \cdot \tau)} \end{aligned} \qquad (13)$$

This advanced model is able to maintain the inversely proportional character of r_{SS} for increasing N. It comprises the parameters k_{FRET}, τ, and r_{ET}, which can be obtained from time-resolved anisotropy.

In an alternative approach that is completely independent of these parameters, I exchanged $r_1 \cdot f_{non}$ in eq. *(12)* with $r_1 \cdot \frac{1 + a}{1 + N \cdot a}$, a simplified version of equation *(8)*, conceived for a better practicability. The resulting simplified model can be written as:

$$\begin{aligned} r_{SS}(x, N) &= r_1 \cdot \frac{1 + a}{1 + N \cdot a} + r_1 \cdot x^{(N-1)} - r_1 \cdot \frac{1 + a}{1 + N \cdot a} \cdot x^{(N-1)} = \\ &= r_1 \cdot \left(\frac{1 + a}{1 + N \cdot a} + x^{(N-1)} \cdot \frac{a \cdot (N-1)}{1 + N \cdot a} \right) \end{aligned} \qquad (14)$$

This term defines r_{SS} as inverse proportional and introduces parameter a, which can be determined empirically from r_{SS} for fractionally photobleached reference proteins. The parameters k_{FRET}, τ, and indirectly r_{ET} in eq. *(8)* are represented solely by the empirically determined parameter a which makes the simplified model completely independent from parameters gained from time-resolved anisotropy analysis.

3.5 Molecular diffusion

The diffusion coefficient of spherical particles, at very low Reynolds numbers, is a function of the temperature, the medium viscosity, and the particle size [117]:

$$D = \frac{k_B \cdot T}{6\pi \cdot \eta \cdot R_M} \qquad (15)$$

where k_B is Boltzmann's constant, T is absolute temperature, η is the dynamic viscosity of the medium, and R_M is the molecular radius which is the cube root of the molecule volume, $\sqrt[3]{V_M}$.

Assuming a globular shape for a concatemer of equally sized protein domains, such as GFP, V_M is directly proportional to the number N of concatenated domains. For comparing concatemer systems of different stoichiometries, N can be expressed as the molecular weight of a protein MW minus the weight of non-repeated elements MW_{non}, such as protein tags used for identification and purification, divided by the weight of one repeated protein domain MW_{rep}. According to that, $N = \frac{MW - MWnon}{MWrep}$.

For room temperature (T_∞ = 25 °C) and constant η_∞ (= 0.89 mPa·s; water at room temperature) and concatemers with gradually repetitive protein domains, the diffusion coefficient D in dependence of the protein's molecular weight MW can be expressed by:

$$D(MW) = D_1 \cdot \frac{1}{\sqrt[3]{N}} = D_1 \cdot \frac{1}{\sqrt[3]{\frac{MW - MW_{non}}{MW_{rep}}}} \qquad (16)$$

with D_1 as the diffusion coefficient of the concatemer with only one protein domain.

4 Methods

The buffers used for experiments in this dissertation were mostly prepared with deionized water (dH_2O). Their composition is listed in non-indexed tables next to the respective experiment. If a buffer was not prepared with dH_2O it is indicated in the table. All chemicals that were used for the buffers were purchased from Applichem (Darmstadt, Germany), Sigma-Aldrich (St. Louis, MO, USA), and Carl Roth (Karlsruhe, Germany) if no other distributor is specified. Restriction endonucleases were purchased from Thermo Fisher Scientific (Waltham, MA, USA) or New England Biolabs (Ipswich, MA, USA). The consumable lab equipment was acquired from Sarstedt (Nürnbrecht, Germany).

4.1 Plasmid design and construction

Constructs that contain a sfGFP domain are based on pET28a plasmids (Novagen, Merck, Darmstadt, Germany; Cat.-No: 69864-3; kind gift by Dr. Kathrin Castiglione, Technical University of Munich). In order to construct 1xGFP as the genetic basis for all other fusion proteins, a synthetic reading frame was used (BioCat, Heidelberg, Germany). The reading frame contained (i) a variant of the Nano-tag (DVEAWLGAR) for immunoblot detection [118]; (ii) sfGFP [81]; (iii) the 22 amino acid (aa) long glycine/serine-based flexible ‘linker A’; (iv) a Tobacco Etch Virus (TEV) protease recognition site (ENLYFQG) [119]; (v) a 26 aa long glycine/serine -based flexible ‘linker B’; (vi) a cysteine for potential labeling; (vii) a variant of the FLAG-tag (YKGDYKDHDG); (viii) and a polyhistidine-tag (xtHis, HNHHGHHNHHHHHH) (**Figure 8**). The sequences coding for all elements are all separated by unique restriction sites.

The reading frame, initially on a pUC57 vector, was cloned into the multiple cloning site of an empty pET28a via restriction based cloning with the enzymes *Nco*I and *Xho*I. The constructs 2xGFP, 3xGFP, 4xGFP, and 5xGFP were generated from 1xGFP by repeated restriction based tandem cloning with *Xba*I / *Pst*I and *Spe*I and *Pst*I. The constructs sfGFP-GCN4-p1, sfGFP-GCN4-pII, and sfGFP-ph3a were generated by exchanging the TEV protease cleavage site of 1xGFP with the respective coiled-coil sequences using cassette cloning with restriction enzymes *Nhe*I and *Bam*HI. Moreover, the N-terminal Nano-tag was replaced by an N-terminal xtHis sequence via restriction-free, PCR-based cloning. All

cloning procedures were monitored via restriction digestion or colony PCR control and a subsequent DNA sequencing control.

Plasmids with constructs EGFP, EGFP-CC-Di, EGFP-CC-Tet, and EGFP-CC-Pent [18] were gently provided by Ajitha Cristie-David and Neil Marsh, University of Michigan.

4.1.1 Transformation of competent Escherichia coli with plasmid DNA

Chemical competent *Escherichia coli* (*E. coli*) cells were transformed with the plasmid DNA via heat shock protocol. To generate competent *E. coli* chemically, the protocol by Chung et al. was used [120]. Therefore, lysogeny broth (LB) medium was prepared with ingredients as listed below and autoclaved before the usage. A volume of 100 mL prewarmed LB medium, containing 12.5 µg/mL tetracyclin (Serva, Heidelberg, Gernmany) for *E. coli* XL1-Blue and no antibiotic for *E. coli* BL21(DE3), was inoculated with a 1 mL overnight culture of the respective *E. coli* strain.

Table 2: *Escherichia coli* strains used in this thesis.

Strain	Genotype
XL1-Blue	*rec*A1, *end*A1, *gyr*A96, *thi*, *hsd*R17(r_K^-, m_K^+), *sup*E44, *rel*A1, *lac*, [F′, *pro*AB^+, *lac*I^qZΔM15, ::Tn10(Tet^r)]
BL21(DE3)	F^-, *omp*T, *hsd*S_B, (r_B^-, m_B^-), *dcm*, *gal*, λ(DE3)
BL21(DE3) pLysS	F^-, *omp*T, *hsd*S_B, (r_B^-, m_B^-), *dcm*, *gal*, λ(DE3), pLysS CmR

While continuously measuring the optical density at λ = 600 nm (OD_{600}) via Ultrospec 3100pro photometer (Amersham Biosciences, Amersham, UK), the culture was incubated in a volume of 250 mL at 37 °C and 140 rpm until it reached OD_{600} = 0.3. The cells were cooled on ice for 10 minutes and centrifuged for 10 minutes at 1000 xg and 4 °C (Hermle Z513K centrifuge, Wehingen, Germany). The supernatant was discharged, suspended the cell pellet in 10 mL of fresh, cold, and sterile TSS buffer, generated aliquots of 100 µL, and subsequently froze the aliquots in liquid nitrogen to be stored at -80 °C.

For the transformation, the chemical competent cells were thawed on ice for 15 min. Heating steps were carefully avoided. Plasmid DNA was added to the cells, still remaining cooled on ice. Around 100 ng of plasmid DNA were sufficient to finally reach over 1000 positive clones. After an incubation of 30 minutes, the cells were heat shocked for 1 minute at 42 °C in the water bath, and again transferred to be incubated on ice for 2 minutes. Subsequently,

900 µL prewarmed LB medium with 20 mM glucose were given to the cells which were then incubated for 60 minutes at 37 °C and constant rotation. Afterwards, 100 µL cell suspension was plated out on LB-agar plates (LB medium with 1.5% (w/v) agar) with the respective antibiotic, 35 µg/mL kanamycin sulfate (Kan, Applichem) for plasmids coding for sfGFP-based constructs or 100 µg/mL ampicillin (Amp, Applichem) for plasmids coding for EGFP-based constructs.

LB medium

1.0% (w/v)	NaCl
1.0% (w/v)	Tryptone
0.5% (w/v)	Yeast extract

TSS buffer

10% (w/v)	PEG 3350
5% (v/v)	DMSO
0.48% (w/v)	$MgCl_2$
in LB medium	

4.1.2 Restriction-based cloning

In a restriction-based cloning approach, a linear insert DNA and a linear vector DNA are both separately generated with the help of restriction enzymes. Afterwards, insert and the vector DNA with compatible ends are mixed for reaction. To obtain insert and vector DNA, a donor and an acceptor plasmid were digested with 5 U of the specific restriction endonucleases in respective buffers for 1h at 37 °C. After the restriction digest, insert and vector DNA were separated from redundant DNA fragments via agarose gel electrophoresis. For DNA ligation, vector DNA and insert DNA in a ratio of 1:3 were mixed with T4 DNA ligase (Thermo Fisher Scientific). DNA mass estimation via band strength comparison with a lambda phage ladder proved to be very precise when setting up the right vector to insert ratio. The ligation was executed over night at 10 °C.

DNA ligation reaction

2µL	T4 DNA Ligase buffer (10x)
50 ng	Vector DNA
50 ng · 3 · length(insert) / length(vector)	Insert DNA
1 U	T4 DNA Ligase
to 20 µL	dH_2O

Sequences coding for coiled-coil peptides (**Table 3**) were integrated into the 1xGFP vector plasmid via cassette cloning. Therefore, the vector was digested using the two restriction enzymes *Nhe*I and *Bam*HI whose restrictions sites frame the sequence coding for a TEV cleavage site in the original plasmid (**Figure 8**). Differently to common cassette cloning approaches where cassettes were generated via oligonucleotide hybridization, the synthetic cassettes (BioCat) were already integrated on pUC57 vectors. To obtain the cassettes, these pUC57 vectors were digested using *Nhe*I and *Bam*HI. Subsequently the cassette was ligated with the linearized vector. Circumventing false positive clones, the control cassette 'You Shall Not Pass' (**Table 3**) with *Apa*I restriction sites was integrated beforehand. Plasmids without a cassette coding for a coiled-coil peptide could then be easily removed via *Apa*I digest (5 U *Apa*I, 60 min, 37 °C).

Table 3: DNA sequence of cassettes used in this thesis.

Cassette name	Sequence (5'-3')(bold: *Nhe*I / *Bam*HI sites)	ref.	PDB
GCN4-p1	**G|CTAGC**GCGGTGAGCCGTATGAAACAGCTGGAAGATAAAG TGGAAGAACTGCTGAGCAAAAACTATCATCTGGAAAACGAA GTGGCGCGTCTGAAAAAACTGGTGGGCGAACGTG**G|GATCC**	[12]	2ZTA
GCN4-pII	**G|CTAGC**GCCGTTTCTCGTATGAAACAGATTGAAGATAAAA TTGAAGAAATTCTGAGCAAAATTTATCATATTGAAAACGAA ATTGCGCGTATTAAAAAACTGATTGGCGAACGTG**G|GATCC**	[12]	1GCM
GCN4-pAA	**G|CTAGC**GCTGTCTCCCGTATGAAACAGGCGGAAGATAAAG CGGAAGAAGCGCTGAGCAAAGCGTATCATGCGGAAAACGAA GCGGCGCGTGCGAAAAAACTGGCGGGCGAACGC**G|GATCC**	this study	-
ph3a	**G|CTAGC**GCCGTATCCCGTATTTATAAAATTGAACAGAAAA TTTATCGTATTGAACAGAAAATTTATCGTATTGAACAGAAA ATTTATAAAATTGAACAGAAAATTTATG**G|GATCC**	this study	-
'You Shall Not Pass' control	**G|CTAGC**TAGGGCCCTGATTAACCGGGCCCTCACGGGCCCG TTAA**G|GATCC**	this study	-

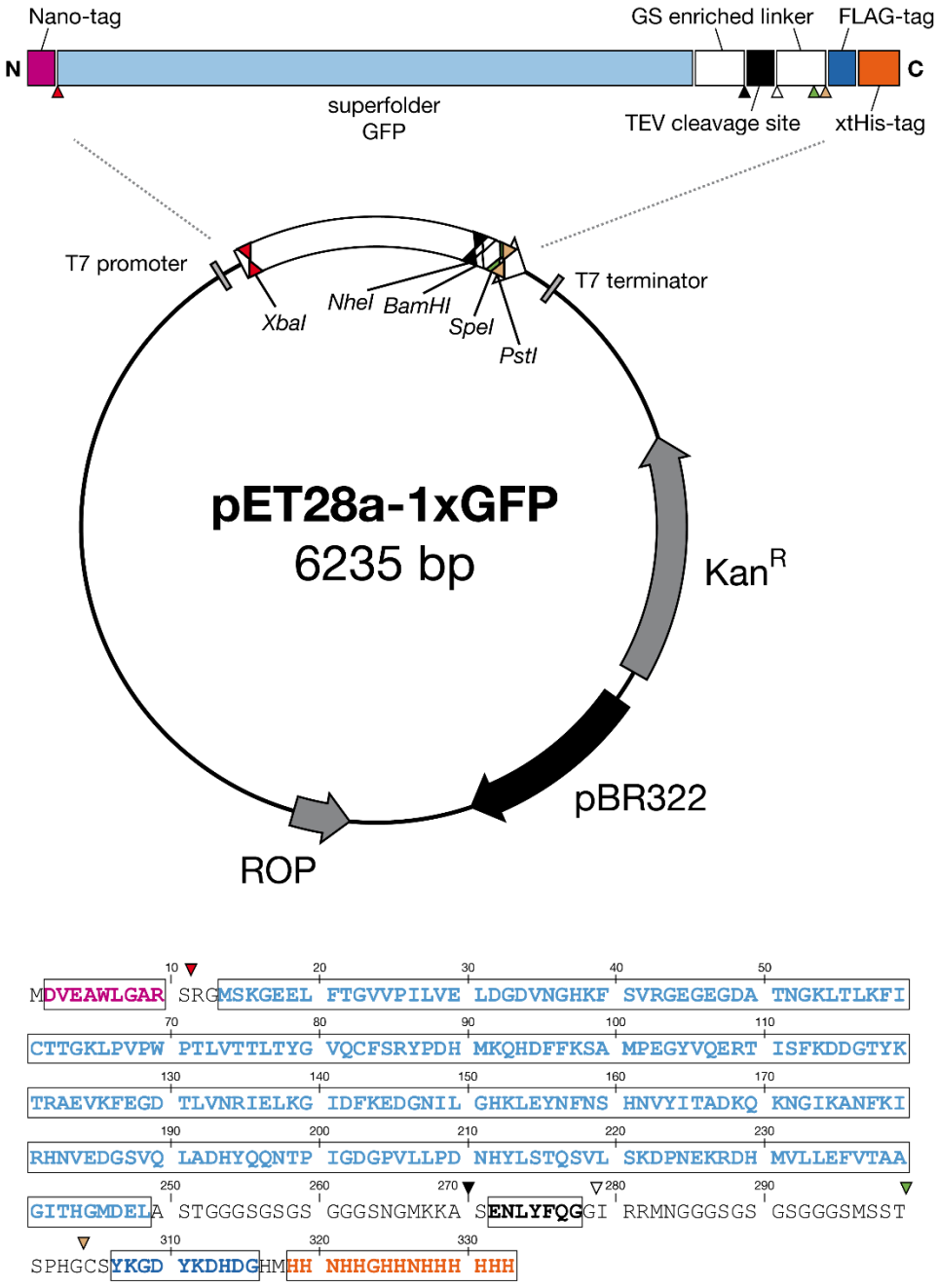


Figure 8: The schematic structure, the plasmid map, and the amino acid sequence of the 1xGFP construct. The 1xGFP vector is based on a pET28a vector. The pET28a-1xGFP plasmid contained a pBR322 origin of replication, a resistance gene against Kanamycin (Kan^R), a gene for the regulator protein Rop responsible for DNA replication (ROP), and the open reading frame with T7 promotor and terminator. The reading frame was coding for domains illustrated as boxes in this figure: superfolder GFP (sfGFP; light blue), Nano- (purple) / FLAG- (blue) / xtHis-Tag(orange), glycine/serine-rich flexible linkers (white), and a TEV cleavage site (black). In later cloning steps, coiled-coil protein sequences were integrated instead of the TEV cleavage site. The domains illustrated as boxes were bordered by distinctive restriction sites depicted as triangles: *Xba*I (red), *Nhe*I (black), *Bam*HI (white), *Spe*I (green), *Pst*I (ochre).

4.1.3 Tandem cloning with cohesive restriction ends

The plasmids for the GFP concatemers were generated via tandem cloning with cohesive restriction ends. With this technique it was possible to generate tandem repeats of sequences coding for sfGFP and linker regions, only by applying a repetitive restriction-based cloning approach. Therefore, a combination of the restriction enzymes *Pst*I, *Xba*I and *Spe*I was used. The enzymes *Xba*I and *Spe*I leave the same overhang sequences after digestion, yet, still differ in their recognition site. In two separate reactions, the plasmid for 1xGFP was cut with *Xba*I and *Pst*I to obtain an insert fragment of around 900 bp and with *Spe*I and *Pst*I to receive a vector backbone fragment (**Figure 8**). The reaction products were mixed and ligated, and *E. coli* XL1-Blue were transformed with the ligation product. The generated plasmid was sequenced and the procedure was repeated to produce plasmids coding proteins with up to five concatenated GFP domains. The resulting concatemers still hold one N-terminal Nano- and one C-terminal xtHis-tag. The linkers between the concatenated GFP domains are equal in length and sequence with one TEV cleavage site each.

4.1.4 Restriction-free, PCR-based cloning

Using a restriction-free, polymerase chain reaction (PCR)-based cloning approach [121], the N-terminal Nano-tag could be exchanged for an additional, N-terminal xtHis-tag. The restriction-free cloning approach comprises two separate PCR steps. The first step uses a pair of primers that contain complementary sequences to both, the target gene that is to be inserted and the vector backbone. These primers are used to amplify the insert sequence from a template vector that holds the gene of interest.

DNA amplification for megaprimer

10 µL	Phusion polymerase buffer (5x)
1 µL	dNTP mix (10 mM)
450 ng	Forward primer
450 ng	Reverse primer
100 ng	Template DNA (insert)
0.5 µL	Phusion polymerase (5 U)
to 50 µL	dH_2O

Amplification process

30 sec	98 °C	
10 sec	98 °C	
20 sec	70 °C	30x
30 sec	72 °C	
5 min	72 °C	

To guarantee high-fidelity PCR conditions, a Phusion High-fidelity DNA Polymerase (New England Biolabs) and an Eppendorf Mastercycler PCR cycler (Eppendorf, Hamburg, Germany) was used. For the given amplification process, one can assume an insert sequence length of 1000 bp which can be extended in 30 seconds by the Phusion polymerase.

In the a second PCR step, the product from the primary reaction is then used as a megaprimer. The terminal overhangs of the megaprimer are complementary to the sequences at the vector backbone which is amplified in both directions. The complementary sequences can be adjacent, which results in a gene insertion between the two complementary sequences, or distant, which leads to the deletion of the sequence in between.

Cloning PCR

4 µL	Phusion polymerase buffer (5x)
0.4 µL	dNTP mix (10 mM)
2 µL	Megaprimer / product 1st PCR
100 ng	Template DNA (vector)
0.2 µL	Phusion polymerase (2 U)
to 20 µL	dH_2O

Amplification process

30 sec	98 °C	
10 sec	98 °C	
20 sec	70 °C	16x
150 sec	72 °C	
5 min	72 °C	

After the secondary reaction, the parental, methylated DNA was removed via *Dpn*I restriction digest (5 U *Dpn*I, 90 min, 37 °C) to maintain only DNA amplified via PCR. Without any ligation, *E. coli* XL1-blue were transformed with nicked plasmid DNA.

4.1.5 Cloning controls

All generated plasmids were controlled for their integrity via control digest, colony PCR, and sequencing. To apply a control digest, a unique restriction site was integrated into the plasmid during the cloning procedure to differentiate parental from newly generated plasmids. If the inserted DNA sequence differed in length a restriction enzyme was used with two distal restriction sites on the plasmid. The reaction product was investigated via agarose gel electrophoresis. For a colony PCR control, primers were used that were exclusively complementary to the inserted sequence, or in the case of different sequence lengths, those primers that were complementary to regions vicinal to the potentially inserted sequence. For

colony PCR, a *Thermus aquaticus* polymerase (generated in-house) was used instead of the purchased Phusion polymerase. For template DNA preparation, the colony of interest was isolated and dissolved in 10 µL dH_2O. All generated plasmids were sequenced at Eurofins (Eurofins, Luxemburg, Luxemburg) and transferred into the plasmid bank at the institute.

4.1.6 DNA preparation, separation, and quantification

To isolate plasmid DNA from a single bacterial colony, cells from the colony of interest were transferred into a test tube with 5 mL LB medium including the respective antibiotic. The cultures were grown over night at 37 °C and constant rotation. For the successful isolation of plasmid DNA, prepacked columns and buffers of the NucleoSpin Plasmid Purification Kit (Macherey-Nagel, Düren, Germany) were used. The manufacture's protocol was followed with all optional preparation steps. Usually, the final concentration of plasmid DNA in dH_2O was around 150 ng/µL.

DNA separation was achieved via agarose gel electrophoresis with a Tris-acetate/EDTA (TAE) buffer. For a DNA fragment resolution between 200 bp and 1000 bp, 1.5% (w/v) agarose in TAE buffer was used. To resolve DNA fragments between 500 bp and 8000 bp, we used 1% (w/v) agarose in TAE buffer. In order to cast the agarose gel, a total volume of 50 mL with respective amount of agarose was heated it in the microwave. Afterwards, 0.03 µg/mL ethidium bromide was added to the solution that was casted into a combined casting and electrophoresis system, using a comb for sample pockets. After the gel had cooled down, the electrophoresis chamber was filled with TAE buffer and samples, mixed with DNA Loading Dye (6x) (Thermo Fisher Scientific), were loaded into the cavities. Parallel to the loaded samples, 7.5 µL of GeneRuler 1 kb DNA Ladder, GeneRuler 100 bp DNA Ladder (ladder depending on the length of the investigated sample), or Lambda DNA/*Eco*RI + *Hin*dIII Marker (all of them: Thermo Fisher Scientific) were loaded into a separate lane. DNA fragments were separated by running the electrophoresis at 80 V for 45 min. The ethidium bromide stained DNA was visualized via UV light at 312 nm.

TAE buffer (pH 8.0)

40 mM	TRIS
20 mM	Acetic Acid
1 mM	EDTA

In restriction based cloning, control experiments were included where DNA was separated via gel-electrophoresis and then extracted from the gel to be further used in subsequent working steps. Therefore, DNA bands were cut out thoroughly after the electrophoresis and purified via prepacked columns and buffers of the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel). The manufacture's protocol was followed with all optional purification steps. Commonly, the concentration of DNA fragments in dH_2O significantly decreased during this experiment.

The isolated and purified DNA was quantified by determining its extinction at 260 nm and 280 nm with an Ultrospec 3100pro photometer. Therefore, DNA was diluted 1:40 in dH_2O and given into a quartz cuvette. As a blank solution, dH2O was used. For pure and isolated DNA, the ratio between the extinction at 260 nm and at 280 nm should optimally be 1.8. For samples contaminated with RNA or remaining protein, one would expect a lower or a higher ratio, respectively.

4.2 Protein expression and purification

4.2.1 sfGFP concatemers

Competent *E. coli* BL21(DE3) pLysS were transformed with the plasmids coding for concatemer constructs. The transformed cells were cultivated overnight at 37 °C on LB-agar with 35 µg/mL Kan. Selected positives clones were used to inoculate 5 mL LB medium containing 35 µg/mL Kan. These cultures were grown for 18 h at 37 °C under constant rotation. Then, the culture was diluted to 1:100 in 500 mL LB medium with 35 µg/mL Kan and was incubated at 140 rpm and 37 °C until an OD_{600} = 0.5. Expression was induced by adding 250 µM isopropyl-β-D-thiogalactopyranosid (IPTG). After 1 hour, the color of the expressing cells had been visibly changed to green. The cells were harvested by gentle centrifugation at 4 °C and washed twice with PBS. The resulting cell pellets were stored at -20 °C.

Protein purification was completely executed at 4 °C or on ice. The concatemer proteins were purified by a two-step purification protocol, with a primary immobilized metal ion affinity chromatography (IMAC) that was succeeded by a secondary anion exchange chromatography (AEX). Cell pellets with expressed concatemer constructs were thawed and resuspended in concatemer lysis buffer. The cell lysis was performed via sonication at a Sonifier W-250D (Branson Ultrasonics, Danbury, CT, USA) for 5 minutes with short breaks after every 30 seconds at half maximal amplitude. The cell debris was separated from brightly green proteins in solution via centrifugation at 20,000 xg for 30 minutes at 4 °C. EDTA in the supernatant was removed by adding 20 mM $MgCl_2$ with an incubation of 1 h and mild shaking. The supernatant was then incubated with Ni-NTA resin (Macherey-Nagel) that had been equilibrated with concatemer IMAC buffer for 16 h under constant rotation.

The suspension was subsequently poured into a column to elute the proteins with a constant flow of 2 mL/minute (Äkta system, GE Bio-Sciences, Pittsburgh, PA, USA). The column was washed with 5 bed-volumes of concatemer IMAC buffer containing 20 mM imidazole and then with 1 bed volume of concatemer IMAC buffer without imidazole. For elution of the concatemer constructs, the imidazole concentration was gradually increased until the green fluorescent construct visibly dissociated from the Ni-NTA material. The imidazole concentration was kept constant to the point where the elution of the constructs seemed to be

complete (usually between 200-300 mM). The eluate was dialyzed against concatemer AEX buffer (1 mL against 100 mL, 3 changes, 24 h) in order to eliminate NaCl and imidazole. For a secondary purification step, the sample was incubated with Q sepharose Fast Flow (GE Bio-Sciences), equilibrated with concatemer AEX buffer, under constant rotation for 3 h. Again, the suspension was given into a column, washed with concatemer AEX buffer until the concentration of NaCl was stepwisely increased to 500 mM NaCl. Brightly green eluate containing the concatemer constructs was finally dialyzed against concatemer sample buffer (1 mL against 100 mL, 3 changes, 24 h). The protein concentration was determined via GFP absorption measurements. The purified concatemer proteins were frozen in liquid N_2 and stored at -80 °C for further experiments.

Concatemer lysis buffer (pH 8.0)

300 mM	NaCl
50 mM	NaH_2PO_4 / Na_2HPO_4
1 mM	EDTA
0.5 mM	β-Mercaptoethanol

Concatemer IMAC buffer (pH 8.0)

20 mM	NaCl
50 mM	NaH_2PO_4 / Na_2HPO_4
0.5 mM	β-Mercaptoethanol

Concatemer AEX buffer (pH 8.0)

50 mM	NaH_2PO_4 / Na_2HPO_4
0.5 mM	β-Mercaptoethanol

Concatemer sample buffer (pH 8.0)

1 mM	EDTA
0.5 mM	β-Mercaptoethanol
0.05% (v/v)	Sodium azide
in flow cytometry grade PBS (Thermo Fisher Scientific)	

4.2.2 Coiled-coil fusion proteins

Competent *E. coli* BL21(DE3) were transformed with the relevant plasmids and grown overnight at 37 °C on LB-agar with 35 µg/mL Kan (sfGFP-based fusion proteins) or 100 µg/mL Amp (EGFP-based fusion proteins). Selected colonies from the plate were used to inoculate overnight cultures in 5 mL LB medium with the respective antibiotic. The cultures were grown under constant rotation for 18 h at 37 °C. The overnight culture was diluted 1:100 in 500 mL 2xYT medium [122, 123] with 35 µg/mL Kan (sfGFP-based fusion proteins) or

100 µg/mL Amp (EGFP-based fusion proteins). The culture was incubated at 37 °C and 140 rpm to OD_{600} = 0.8. Then, 100 µM IPTG were added and the culture further incubated at 18 °C and 140 rpm for 18 h. After the incubation, cells appeared green in case of a successful expression. The cells were harvested by mild centrifugation at 1000 xg for 15 minutes at 4 °C, washed twice with PBS, and stored at -20 °C.

The protein purification of the coiled-coil fusion proteins was performed at 4 °C or on ice. The coiled-coil proteins were purified with a single-step IMAC purification protocol. Differently to the purification protocol of concatemer proteins, all buffers were based on 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and contain glycerol in order to prevent coiled-coil aggregation. The latter was based on the stabilizing effect of glycerol on coiled-coil protein folding as it has been demonstrated in the past [124, 125]. To purify the proteins, cells that expressed coiled-coil fusion protein were thawed and resuspended in coiled-coil lysis buffer that included the EDTA-free protease inhibitor cOmplete Mini (Roche Diagnostics, Basel, Switzerland) to prevent proteolysis and 70 mM lysozyme (Sigma-Aldrich) to enhance bacterial cell lysis. The cells were lysed via sonication and cell debris separated from soluble proteins via centrifugation at 20,000 xg. The supernatant was then incubated with Ni-NTA resin, equilibrated with coiled-coil lysis buffer beforehand, for at least 16 h under constant rotation. The slurry was washed with 10 bed volumes of coiled-coil lysis buffer. Subsequently, the coiled-coil fusion proteins were eluted using the coiled-coil elution buffer and dialyzed against the coiled-coil sample buffer (1 mL against 100 mL, 3 changes, 24 h). The purified proteins were frozen in liquid N_2 and stored at -80 °C for further experiments.

2xYT medium

1.6% (w/v)	Tryptone
1.0% (w/v)	Yeast extract
0.5% (w/v)	NaCl

Coiled-coil lysis buffer (pH 7.5)

300 mM	NaCl
50 mM	HEPES
25 mM	Imidazole
1 M	Urea
5% (v/v)	Glycerol
1 tablet / 20 mL	Protease inhibitor
70 mM	Lysozyme

Coiled-coil elution buffer (pH 7.5)

300 mM	NaCl
50 mM	HEPES
500 mM	Imidazole
5% (v/v)	Glycerol

Coiled-coil sample buffer (pH 7.5)

100 mM	NaCl
25 mM	HEPES
2 mM	EDTA
30% (v/v)	Glycerol

4.3 Biochemical protein analysis

4.3.1 Denaturing polyacrylamide gel electrophoresis

Denaturing polyacrylamide gel electrophoresis (PAGE) was used to separate proteins in solution by their molecular weight. Separating proteins by denaturing PAGE helped to understand the purity of a protein sample and the integrity of the target protein. In this study, gels with a polyacrylamide pore size gradient were used, particularly referred to as Schägger gradient gel [126], and without a pore size gradient [127]. In both cases, sodium dodecyl sulfate (SDS) was used to denature the protein samples and buffers based on tris(hydroxymethyl)aminomethane (TRIS). In order to cast and to run the gels, we used the PerfectBlue Wide Format Dual Gel System Twin ExW S (Peqlab, Erlangen, Germany) for gels sized 20 cm x 10 cm x 0.8 mm.

For gels without a pore size gradient, the separating gel contained 10% (v/v) acrylamide-bisacrylamid mixture and the covering stacking gel 5% (v/v) of the same mixture. The polymerization was catalyzed by the addition of ammonium persulfate (APS) and tetramethylethylenediamine (TEMED).

Separating gel	
375 mM	TRIS buffer (pH 8.8)
10% (v/v)	Acrylamide-Bisacrylamid Mixture (37.5:1)
0.1% (v/v)	SDS
0.01% (w/v)	APS
0.04% (v/v)	TEMED

Stacking gel	
62.5 mM	TRIS buffer (pH 6.8)
5% (v/v)	Acrylamide-Bisacrylamid Mixture (37.5:1)
0.1% (v/v)	SDS
0.01% (w/v)	APS
0.1% (v/v)	TEMED

TRIS buffer (pH 8.8)	
1.5 M	TRIS
with HCl to pH 8.8	

TRIS buffer (pH 6.8)	
0.5 M	TRIS
with HCl to pH 6.8	

Schägger gradient gels consisted of three different layers, the small-pore bottom layer gradually blending into a large-pore spacing layer, covered by the Schägger stacking gel layer. The small-pore gel contained 16.5% (v/v), the large-pore spacing gel 10% (v/v), and the Schägger stacking gel 4% (v/v) of the freshly prepared Schägger acrylamide-bisacrylamid mixture. As for the non-gradient gels, the polymerization was started with the addition of APS and TEMED with varying concentrations from layer to layer. Including glycerol in the small-pore bottom gel instead of water further helped to form the gradient.

Small-pore bottom Schägger gel

1 M	TRIS buffer (pH 8.45)
16.5% (v/v)	Schägger Acrylamide-Bisacrylamid Mixture (32:1)
0.1% (v/v)	SDS
0.006% (w/v)	APS
0.07% (v/v)	TEMED
10.5%	Glycerol

Large-pore spacing Schägger gel

1 M	TRIS buffer (pH 8.45)
10% (v/v)	Schägger Acrylamide-Bisacrylamid Mixture (32:1)
0.1% (v/v)	SDS
0.009% (w/v)	APS
0.11% (v/v)	TEMED

Schägger stacking gel

750 mM	TRIS buffer (pH 8.45)
16.5% (v/v)	Schägger Acrylamide-Bisacrylamid Mixture (32:1)
0.1% (v/v)	SDS
0.01% (w/v)	APS
0.13% (v/v)	TEMED

TRIS buffer (pH 8.45)

3 M	TRIS
with HCl to pH 8.45	

Schägger Acrylamide-Bisacrylamid Mix.

48% (w/v)	Acrylamid
1.5% (w/v)	Bisacrylamid

To resolve protein samples by their size, the samples were diluted to 1-2 µM in the Laemmli loading buffer and heated at 95 °C for 10 minutes to fully denature the proteins. Then the samples were pipetted into pockets in the stacking gel layer, formed by a comb beforehand. In this thesis, the unstained Pierce Protein Molecular Weight Marker and the prestained

PageRuler Protein Ladder (both: Thermo Fisher Scientific) were used to identify the molecular weight of the sample proteins. To run the electrophoresis, non-gradient gels required a TRIS-glycine running buffer. Electrophoreses with Schägger gradient gels, however, were run with specific Schägger anode and TRIS-tricine cathode buffers. The proteins were separated at 120 V for 1.5 h on the non-gradient and for 2.75 h on the Schägger gradient gel.

Laemmli loading buffer

62.5 mM	TRIS buffer (pH 6.8)
2% (v/v)	SDS
10% (v/v)	Glycerol
2.5% (v/v)	β-Mercaptoethanol
2 M	Urea
0.02% (w/v)	Bromphenol blue

TRIS-glycine buffer

20 mM	TRIS
192 mM	Glycine
0.1% (v/v)	SDS

Schägger anode buffer

200 mM	TRIS
with HCl to pH 8.9	

Schägger TRIS-tricine cathode buffer

100 mM	TRIS
100 mM	Tricine
0.1% (v/v)	SDS

If the proteins were not further transferred on a membrane via semi-dry blot, the gels were stained with Coomassie staining solution for 1 h, and destained in several steps with the destaining solution. Protein bands that remained dark-blue-stained were analyzed via densitometry with ImageStudio Lite (LI-COR, Lincoln, NE, USA). In experiments with bovine serum albumin (BSA) and GFP, the band strength determined from densitometry increases linearly for the loaded protein concentration (**Figure 9**).

Coomassie staining solution

50% (v/v)	Ethanol
30% (v/v)	Acetic acid
0.1% (w/v)	Coomassie Brilliant Blue G250

Destaining solution

30% (v/v)	Ethanol
10% (v/v)	Acetic acid

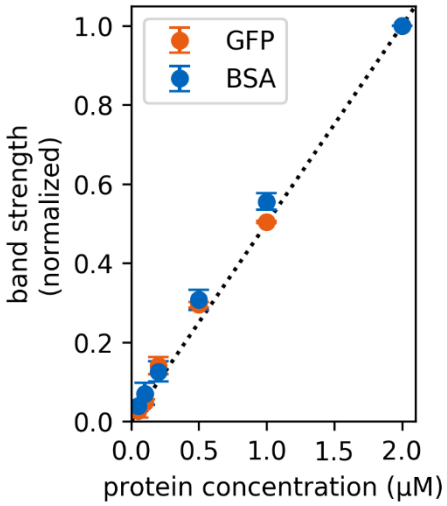


Figure 9: The band strength on denaturing PAGE is linearly proportional to the protein concentration. Bovine serum albumin (BSA) and the 1xGFP fusion protein were diluted to 2 µM in PBS. Dilution rows of both were separated via denaturing PAGE. The gel was stained with Coomassie Brilliant Blue G250. The strength of the main band was analyzed densitometrically and normalized to the band corresponding to 2 µM. Data points show means with standard deviation (n = 3). The dotted line shows a linear function (band strength = 0.5 · protein concentration).

4.3.2 Native polyacrylamide gel electrophoresis

In this thesis, color-less native PAGE was used to characterize the stoichiometry of coiled-coil fusion proteins. As the migration of folded protein oligomers is solely driven by their intrinsic charge, the isoelectric point pI of the proteins is decisive for an acceptable separation. An isoelectric point between 3 and 8 is desirable.

Except for the omission of the denaturant, the gel, sample and running buffers for the native PAGE did not differ from those for the denaturing PAGE with non-gradient 10% separation gels. The buffers were therefore prepared as described previously but without containing SDS. Furthermore, proteins were not heated before loading and the running time was increased to 2.5 h as the native proteins took longer to migrate through the gel. The staining of the gel was equal to its denaturing variant as well as the densitometric analysis of the stained gel.

4.3.3 Size exclusion chromatography

Size exclusion chromatography (SEC), often simply called gel filtration, was used to characterize coiled-coil fusion protein oligomers. In gel filtration, the separation of the proteins relies on the molecular size and is achieved by differently sized cavities in beads that are packed in a separation column. In this thesis, a Superdex 200 10/300 separation column (GE Bio-Sciences) was used whose resolution range lied between 10 and 300 kDa. For the equilibration of the column and the elution of the proteins, the same SEC buffer was used.

The proteins were loaded on the column and eluted at 0.3 mL/minute (Äkta system, GE Bio-Sciences). The peak fractions could be determined via UV_{280} and fluorescence intensity (λ_{ex} = 488 nm; λ_{em} = 520 nm) detection. The column was calibrated with standard protein solutions (Merck, Darmstadt, Germany) containing β-amylase, 200 kDa, alcohol dehydrogenase, 150 kDa, carbonic anhydrase, 29 kDa, and cytochrome C, 12.4 kDa. Blue dextran, 2000 kDa, was used to determine the exclusion volume (**Figure 10**). Using the calibration curves, the elution times of the protein samples were related to their molecular masses.

SEC buffer pH 7.5

25 mM	HEPES
100 mM	NaCl
2 mM	EDTA

4.3.4 Immunoblot

Proteins based on the 1xGFP fusion protein exhibit an N-terminal Nano-tag, a C-terminal FLAG-tag, and a C-terminal His-tag. They served as immunologically detectable epitopes to identify integrity of the investigated fusion proteins. In this study, this was achieved via a two-step immune staining of the proteins after the sample was transferred from a poly-acrylamide gel onto a membrane via semi-dry blotting.

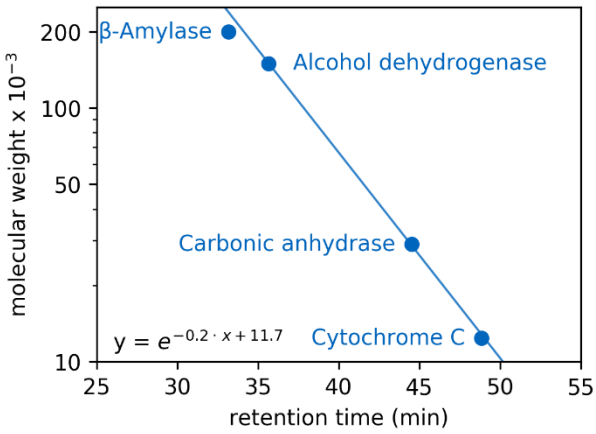


Figure 10: Calibration curve of the separation column for size exclusion chromatography. The calibration was executed with β-amylase, 4 mg/mL, alcohol dehydrogenase, 5 mg/mL, carbonic anhydrase, 3 mg/mL, and cytochrome C, 2 mg/mL.

In a semi-dry blotting step, proteins were transferred from the polyacrylamide gel onto a nitrocellulose (Berrytec, Grünwald, Germany) or a polyvinylidene difluoride (Immobilion, Merck) membrane. Therefore, filter papers were soaked with TRIS-glycine buffer containing 20% (v/v) methanol and placed on the anode of the blotting instrument (Modell SD 1, cti, Idstein, Germany). Then the membrane was placed on the filter papers. Nitrocellulose membranes were activated via soaking in 20% (v/v) methanol containing TRIS-glycine buffer, polyvinylidene difluoride membranes via soaking in isopropanol and dH_2O before usage. Subsequently, gel layers from denaturing PAGE containing separated protein samples were placed on the membrane, followed by a second stack of similarly soaked filter papers, and the cathode. Ballasting the blotting device, the proteins were transferred for 1.5 h at 1.5 mA/cm^2 transfer area. The transfer efficiency was controlled by staining the membrane with Ponceau S solution and destaining with dH_2O.

Ponceau S solution

3% (w/v)	Trichloroacetic acid
0.3% (w/v)	Ponceau S

In order to immunologically detect proteins on the membrane, epitope carrying proteins were marked by a specific primary antibody that could be detected in turn by a secondary antibody fused to an alkaline phosphatase (AP). The alkaline phosphatase hydrolyzes 5-bromo-4-

chloro-3-indolyl phosphate (BCIP). The reaction product can be further oxidized by nitro blue tetrazolium (NBT) which can be detected as visible, dark-blue areas on the membrane where proteins were immunologically marked.

Before the application of antibodies, the membrane was blocked overnight with 20 mL 3% (w/v) nonfat dried milk powder in Tris-buffered saline (TBS) at 4 °C. Afterwards, the membrane was washed for 5 minutes with TBS. Then, it was incubated for 1 h at room temperature with 20 mL blocking solution containing the primary antibody (**Table 4**). After this step, it was washed three times for 5 minutes with TBS containing polysorbate 20 (TBS-T). The membrane was incubated for 1 h at room temperature with 20 mL blocking solution containing 0.13 µg/mL anti-murine IgG AP Conjugate (Promega, Madison, WI, US) as the secondary antibody, followed by three 5-minute washing steps with TBS-T. Finally, the membrane was incubated with an BCIP/NBT staining solution to visualize the immunologically marked proteins and washed with water to stop the reaction.

TBS (pH 7.4)	
20 mM	TRIS
150 mM	NaCl

TBS-T (pH 7.4)	
	TBS (pH 7.4)
0.5% (v/v)	Polysorbate 20 “Tween”

BCIP/NBT staining solution (pH 9.5)	
100 mM	TRIS
100 mM	NaCl
5 mM	$MgCl_2$
0.3% (v/v)	BCIP solution
0.6% (v/v)	NBT solution
Preparation immediately before usage.	

BCIP solution (stored at -20 °C)	
	Dimethylformamide
5% (w/v)	BCIP

NBT solution (stored at -20 °C)	
	Dimethylformamide
5% (w/v)	NBT

Table 4: Primary antibodies used in this thesis

Antibody	**Origin**	**Distributor**	**Concentration**
Monoclonal Anti-His6	Murine	Roche, Basel, Switzerland	0.5 µg/mL
Monoclonal Anti-FLAG M2	Murine	Sigma-Aldrich, St. Louis, MO, US	0.1 µg/mL
Monoclonal Anti-Nano	Murine	MyBioSource, CA, US	0.1 µg/mL

4.4 Fluorescence measurements

All of the fusion proteins in this thesis held a green fluorescent protein domain. To detect GFP fluorescence, the protein sample was excited at λ_{ex} = 488 nm and the emitted light was recorded at λ_{em} = 520 nm.

In experiments where both fluorescence intensity and the steady-state fluorescence anisotropy were measured, purified protein samples were diluted to 2 µM in the respective sample buffer and pipetted the samples into 96-well plates (non-treated back plates, Nunc, Thermo Fischer Scientific). Fluorescence intensity measurements of samples in 96-well plates were executed with a FLUOstar microplate reader (BMG Labtech, Ortenberg, Germany).

To determine excitation and emission maxima, excitation and emission spectra of the monomeric sfGFP fusion protein were recorded in 10 mm precision quartz cuvettes (Hellma Analytics, Müllheim, Germany) with a fluorescence Cary Eclipse spectrophotometer (Varian, Darmstadt, Germany).

4.4.1 Fluorophore concentration measurements via Beer-Lambert law

The concentration c of a purified GFP fusion protein was determined via its absorbance A for an illumination at 488 nm described by the Beer-Lambert law:

$$c = \frac{A}{\varepsilon \cdot d} \qquad (17)$$

where the path length d was 1 cm. The molar attenuation coefficient ε for sfGFP is ε = 83,300 M^{-1} cm^{-1} and for EGFP ε = 56,000 M^{-1} cm^{-1}. The absorbance was measured with the Ultrospec 3100pro photometer.

4.4.2 Fluorescence quantum yield

Determining the fluorescence quantum yield QY of an unknown fluorescent sample, we applied a "relative" approach [128] with a standard fluorophore with a known QY as a reference to compare. In our experiments, fluorescein isothiocyanate (FITC) in H_2O was used as the standard fluorophore.

The QY of an unknown sample can be written as [128]:

$$QY_{sample} = \frac{\int FEm_{sample}}{\int FEm_{standard}} \cdot \frac{(1 - 10^{-A_{standard}})}{(1 - 10^{-A_{sample}})} \cdot \left(\frac{\zeta_{sample}}{\zeta_{standard}}\right)^2 \cdot QY_{standard} \qquad (18)$$

where $\int FEm$ is the area under the fluorescence emission spectrum, A is the absorbance at excitation wavelength λ_{ex} = 488 nm, and ζ the refractive index which is 1.331 for H_2O and 1.340 for PBS [129]. The fluorophore FITC has a QY = 0.84 in EtOH [130], where it was solved for storage, and a QY = 0.76 in H_2O or PBS [130].

Experimentally, 1xGFP and FITC were diluted to 2 µM in sample buffer and dH_2O, respectively, and measured the emission spectrum with the Cary Eclipse spectrophotometer and the absorbance with Ultrospec 3100pro photometer.

4.4.3 Steady-state fluorescence anisotropy measurements

In order to determine the steady-state fluorescence anisotropy, the protein samples were illuminated with polarized light (λ_{ex} = 488 nm). The parallel and perpendicular components of emitted fluorescence (λ_{em} = 520 nm) were recorded with a POLARstar microplate reader (BMG Labtech). To calculate the steady-state anisotropy, eq. *(2)*, the instrument dependent G-factor was required. It was determined with the help of FITC, diluted to 5 µM in dH_2O, as a reference. The calculation and the standardization of the measured fluorescence data was automatized with a Python script.

4.4.4 Steady-state anisotropy and fractional photobleaching

In this thesis, we tested the practicability to characterize the stoichiometry of coiled-coil fusion proteins via steady-state fluorescence anisotropy in combination with fractional photobleaching. As described in chapter 3.4, pp. 18-21, the steady-state anisotropy increases if neighboring, fluorophores that interact via homo-FRET are fractionally photobleached (**Figure 3**) [57].

To fractionally photobleach fluorophores, protein samples of 2 µM (in a volume of 1 mL) were irradiated with a 445 nm laser diode for fixed periods of time, usually 10-minute irradiation steps. A current of 450 mA was used, which resulted in a laser power of approximately 300 mW (**Figure 11**). The samples were kept in a tin-foil-wrapped cuvette at

a fixed position, cooled by a water circuit, and the laser diode sitting directly above. As 300 mW laser light can cause severe vision damage, protection glasses were worn constantly while working with the laser.

After each irradiation step, steady-state anisotropy values were determined with a microplate reader capable of detecting fluorescence polarization (POLARstar, BMG Labtech). To measure the irradiated sample in a 96-well plate, the sample was transferred from the cuvette and split into 6-9 technical replicates. On the 96-well plate, non-irradiated control replicates were measured in parallel. In order to reduce the day-to-day variability of anisotropy values, the data was standardized as described in the following subchapter. Additionally, the residual fluorescence intensity was measured (λ_{ex} = 488 nm; λ_{em} = 520 nm) to determine the fraction of inactive fluorophore x, eq. *(10)*. After the measurement procedure, the samples were pooled again and given into the cuvette to start the next irradiation step.

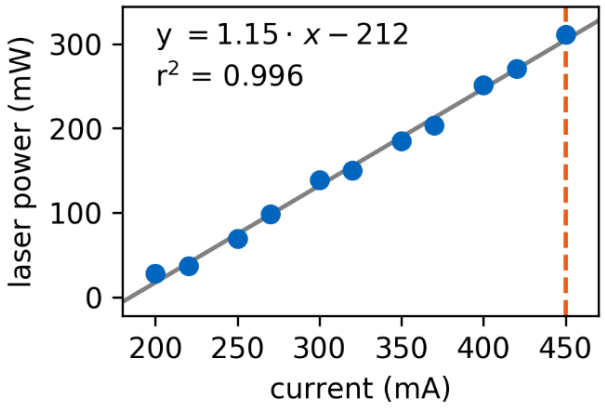


Figure 11: A blue laser (λ = 445 nm) with a power of 300 mW was used to photobleach the fluorophores in this study. The experimentally measured power of the laser is linear proportional to the supplying current.

4.4.5 Standardization of the steady-state fluorescence anisotropy data

The steady-state anisotropy in combination with fractional photobleaching was analyzed and standardized via Python scripts. In order to reduce the high day-to-day variability of anisotropy values for the same protein sample (**Figure 12A**), the data was corrected with the help of non-irradiated controls, always measured in parallel on the same 96-well plate.

In a first correction step, the anisotropy of irradiated samples was corrected by the mean anisotropy of non-irradiated controls (same experiment) (**Figure 12B**). That yielded the anisotropy differences Δr_{SS} between irradiated samples and non-irradiated controls as a function of the fraction of photobleached fluorophores x:

$$\Delta r_{SS}(x) = r_{SS}(x) - \overline{r_{SS\,non}}\ (\text{same experiment}) \qquad (19)$$

After determining the differences $\Delta r_{SS}(x)$ between irradiated samples and a non-irradiated control (measurement in parallel), the differences were added to the mean values of non-irradiated controls (**Figure 12C**):

$$r_{SS}(\text{converted}) = \Delta r_{SS}(x) + \overline{r_{SS\,non}}\ (all\ \text{experiments; same protein}) \qquad (20)$$

The mean anisotropy of all non-irradiated controls could differ from the values for single experiments (**Figure 12A**).

After this final step, the bleaching curves for one protein sample start all at the same r_{SS} (**Figure 12D**). The complete r_{SS} data in this thesis was corrected in this manner.

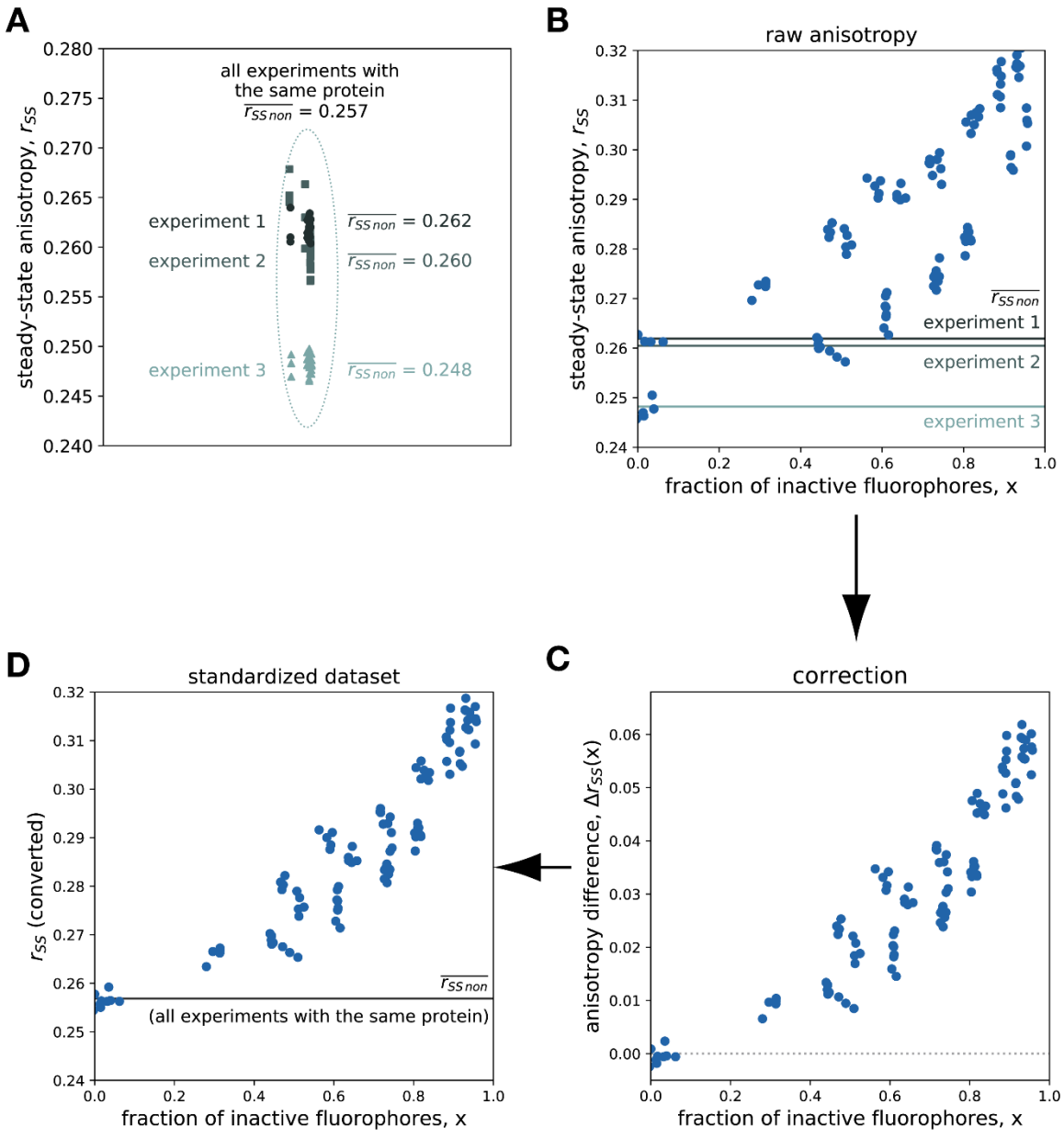


Figure 12: Standardization of steady-state fluorescence anisotropy increases the comparability of GFP concatemers when photobleached. Fusion proteins were diluted to 2 µM in sample buffer and were irradiated with a 445 nm laser (300 mW) for fixed periods of time. After bleaching, the residual fluorescence intensity was measured (λ_{ex} = 488 nm, λ_{em} = 520 nm). The fraction of inactive fluorophores was calculated as the residual fluorescence intensity divided by the initial intensity before photobleaching. Depolarization of emitted light was measured and steady-state anisotropy was calculated, see eq. *(2)*. **A)** The steady-state anisotropy for non-irradiated protein controls (r_{non}) was used for standardization. **B)** Without a correction, the comparison of raw steady-state anisotropy of a GFP fusion protein (exemplary: 5xGFP concatemer) proves to be difficult due to high sample-to-sample variability. **C)** Correction removes day-to-day variance of anisotropy. **D)** The datapoints were converted with the anisotropy mean of non-irradiated controls for all experiments with the same protein.

4.4.6 Time-resolved fluorescence anisotropy

Measuring time-resolved anisotropy for various GFP fusion proteins was executed in collaboration with Ganesh Agam at the Chair of Physical Chemistry I at the Ludwig Maximilans Universität (LMU) Munich. To determine the anisotropy and the fluorescence lifetime on a nanosecond scale, the samples were primarily diluted to a concentration of 2 µM in the corresponding sample buffer.

To excite the fluorophore, a polarized and pulsed 468 nm laser was used. The fluorescence emission was detected with a FLS1000 spectrometer (Edinburgh Instruments, UK) equipped with photomultipliers. The polarization parallel and perpendicular to the excitation laser was recorded. Resolving the anisotropy decay of the sample, we used time correlated single photon counting (TCSPC) detection (Time Harp 260, PicoQuant, Germany). For the calibration of the instruments, an Atto488 dye was used in the same concentration. The instrument dependent G-factor was 0.78. To fit the anisotropy decay in a time frame of 10 ns after the excitation, eq. *(6)* and eq. *(7)* were used.

The fluorescence lifetime was determined by Ganesh Agam via a sub-ensemble method. Here, the signals from both, the parallel and perpendicular channel, were combined together and the fluorescence intensity decay was analyzed. The instrument response function was obtained by exploiting a scattering solution. The combined intensity decay could be explained with a mono-exponential decay function, convoluted with the help of the instrument response function.

4.4.7 Fluorescence Correlation Spectroscopy

Supplementary analysis via fluorescence correlation spectroscopy (FCS) helped to determine the diffusion coefficients of concatenated sfGFP fusion proteins (**Figure 13C**). The experiments were executed by Ganesh Agam. Similar to the time-resolved anisotropy set-up, Ganesh Agam used a custom-build confocal microscope at the Chair of Physical Chemistry I at LMU Munich. The microscope was equipped with synchronized pulsed 482 nm excitation and time-correlated single photon counting capability [131]. The samples were diluted to 2 µM in their corresponding sample buffer. They were excited with approximately 100 µW laser power. The data was analyzed using a custom-written MATLAB script (The MathWorks) [132].

The resulting curves were fitted by Ganesh Agam with the model for free translational three-dimensional diffusion [133], determining the diffusion time τ_D and the diffusion coefficient D:

$$D = \frac{w_0^2}{4\tau_D} \qquad (21)$$

where w_0 is the radial diameter of the confocal volume.

4.5 Mass Spectrometry

4.5.1 Matrix-assisted laser desorption/ionization – time of flight mass spectrometry

Some fusion proteins generated this thesis were analyzed via matrix-assisted laser desorption/ionization – time of flight (MALDI-TOF) mass spectrometry. The experiments were executed by Dr. Martin Haslbeck at the Chair of Biotechnology at the Technical University of Munich (TUM) in Garching and determined the masse of samples in the range from 5-40 kDa. In MALDI-TOF mass spectrometry, proteins are embedded in a specific matrix and are ionized as the matrix absorbs laser energy. The proteins are ejected from the surface. They are analyzed by their mass-to-charge ratio (m/z) as the time to reach the detector is recorded.

The mixing process of MALDI matrix solution and the sample (ratio 1:1) was executed via C4 zipTips (Merck). The sample was simultaneously eluted with the matrix solution and spotted on an AnchorChip MALDI (Bruker, Billerca, MA, USA) target plate. The mixture was allowed to dry at room temperature, to form smooth matrix crystals. Afterwards, the spotted protein was analyzed with a Ultraflex MALDI TOF instrument (Bruker) at the Department of Chemistry at TUM by Dr. Martin Haslbeck.

MALDI matrix solution

50% (v/v)	Acetonitrile
saturated	α-Cyano-4-hydroxycinnamic acid

4.5.2 Electrospray ionization – time of flight mass spectrometry

As a second mass spectrometric method, electrospray ionization – time of flight (ESI-TOF) mass spectrometry helped to analyze fusion protein samples, alternatively to MALDI-TOF. ESI-TOF measurements were executed by Walter Stelzer at our group, Chair of Biopolymer Chemistry at the TUM in Freising.

Differently to MALDI-TOF, the ionization of sample proteins is achieved via a fine electrospray cone that ejects charged molecules into the flight tube. Therefore, fusion proteins of interest in sample buffer have to be prepared specifically for ESI-TOF. They were desalted and stepwise eluted from a tC18 Sep-Pak column (Waters, Milford, MA, USA) with 20%, 40%, 60%, 80%, or 100% acetonitrile, respectively, with 0.1% formic acid.

The eluates were subsequently measured with a Synapt G2 HDMS ESI-TOF setup (Waters) by Walter Stelzer. The data was analyzed with MassLynx 4.1 (Waters). For some of the samples, the raw data was deconvoluted with the maximum entropy algorithm MaxEnt 1 (Waters).

5 Results

5.1 Fluorescence anisotropy of GFP fusion proteins

5.1.1 Green fluorescent protein incorporated in the fusion protein 1xGFP

To date, GFP fusion proteins were used in many reporter assays to illuminate protein-protein interaction. In particular, protein stoichiometry was investigated via GFP fusion proteins in several studies. Apart from the usage of GFP as a fluorescent reporter, fusion proteins with a segmented and modular structure helped to classify protein-protein interaction in the past [134-136].

To systematically investigate the oligomerization of polypeptides (20-50 aa), 1xGFP was generated as a basic construct, a modular chimeric protein holding a superfolder GFP (sfGFP) domain [81], long flexible linkers, and an exchangeable region for the extension with the peptide of interest (**Figure 8**). We purposely decided for sfGFP among several other GFP reporters as it is very bright, has a short maturation time, and exhibits a good quantum yield (**Table 1**). Furthermore, sfGFP behaves as a monomer due to the mutation of A206V [81]. 1xGFP was expressed in *E. coli* BL21(DE3) pLysS and consecutively purified by immobilized metal affinity chromatography (IMAC) and anion exchange chromatography (AEX).

In order to guarantee that the fluorescent reporter still retained its functionality in the chimeric protein, 1xGFP was investigated adequately for its biochemical and photophysical properties, summed up in **Table 5**. The experimentally determined photophysical values show little difference compared to literature values for isolated sfGFP [80].

5.1.2 Generating sfGFP concatemers as model proteins

Fluorescence anisotropy is a powerful tool to characterize the oligomeric state N of a molecule that is labeled with the same kind of fluorophore. The extent of homo-FRET (**Figure 13A**) affects the steady-state anisotropy in a reciprocal manner: the steady-state anisotropy r_{SS} is inverse proportional to N, eq. *(8)* [46].

Table 5: Experimentally determined properties of 1xGFP

1xGFP	
Molecular weight, MW	36.4 kDa
Sequence length	333 aa
Diffusion coefficient, D	67.5 µm² s^{-1}
$\phi_{\text{aqueous buffer}}$ [1]	20.34 ns
$\phi_{\text{50\% glycerol}}$ [2]	104.38 ns
Excitation wavelength, λ_{ex}	488 nm
Emission wavelength, λ_{em}	511 nm
Stokes shift	23 nm
Fluorescence lifetime, τ	2.47 ns
Steady-state anisotropy, r_{SS}	0.297
Quantum yield, QY	0.70
$t_{1/2\ \text{bleaching}}$ [3]	13.62 min

[1] Rotational correlation time of 1xGFP in an aqueous buffer.
[2] Rotational correlation time of 1xGFP in 50% glycerol.
[3] Time in which half of the initial fluorescence is lost after irradiation with 300 mW strong laser light (445 nm).

However, the resolving power of fluorescence anisotropy and homo-FRET in determining oligomer stoichiometry strongly decreases with the spatial distance between fluorophore labels (**Figure 13B**). For short fluorophore distances the separation of oligomer anisotropy is maximal. With large fluorophores, such as fluorescent proteins, the inter-fluorophore distance can be larger than the Förster radius ($R/R_0 > 1$) (**Figure 13B,** the grey dotted line represents the case of GFP). Furthermore, it was demonstrated in the past that for GFP-labeled protein clusters the relative orientation of the fluorophores is not random [51]. In this case, the term incorporating r_{ET} becomes significant and the homo-FRET effect on anisotropy is limited, eq. *(8)* (**Figure 13B,** the grey surface). Therefore, the fluorophore species is heavily influencing the accuracy by which the stoichiometry of fluorophore-labeled proteins can be resolved.

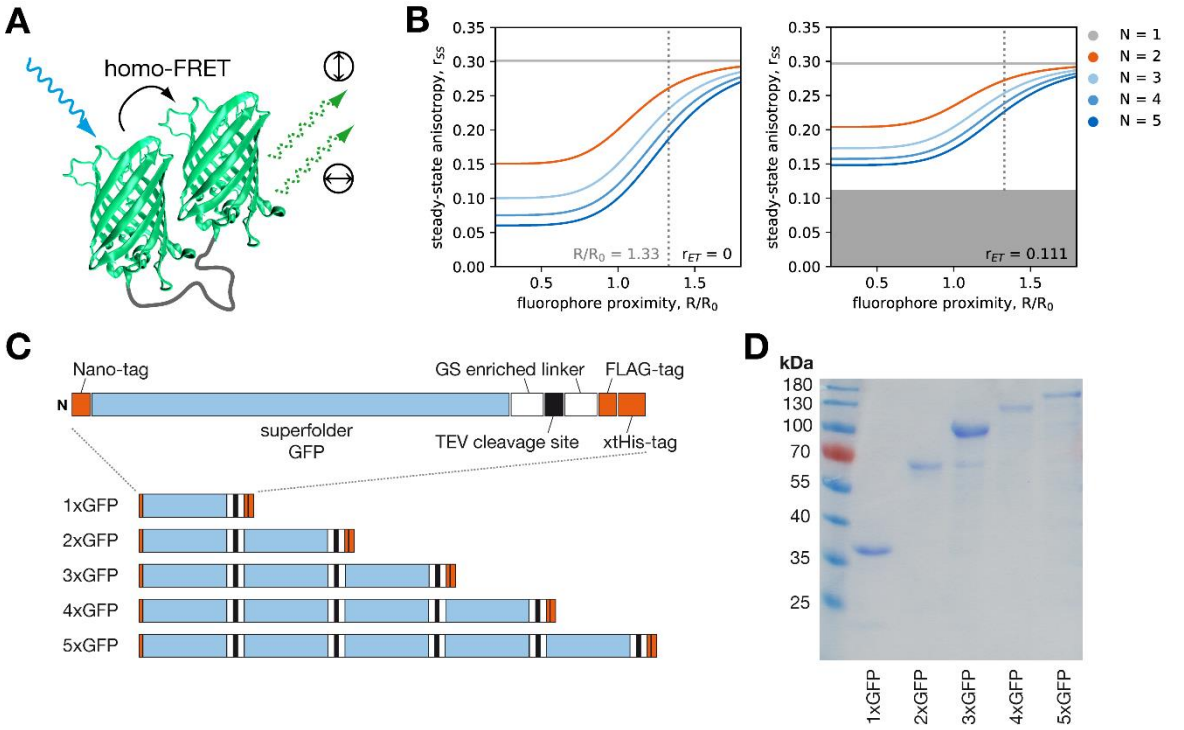


Figure 13: sfGFP concatemers as model proteins for sfGFP complexes interacting via homo-FRET. **A)** Schematic interpretation of homo-FRET between covalently linked sfGFP domains. The steady-state anisotropy can be experimentally determined via eq. *(2)*. **B)** In a complex of N proximal fluorophores interacting via homo-FRET, the steady-state anisotropy r_{SS} is inverse proportional to N [46]. The impact of homo-FRET on r_{SS} is depending on fluorophore proximity, expressed as the ratio between inter-fluorophore distance R and Förster Radius R_0, see eq. *(3)*. Further, r_{SS} is restricted when fluorophores are oriented non-randomly, depicted by the rear term of eq. *(8)*, incorporating r_{ET}. For randomly oriented fluorophores, the term that comprises r_{ET} can be neglected ($r_{ET} \approx 0$; left panel). In case of fluorescent proteins, however, r_{ET} becomes significant (grey box; r_{ET} = 0.111) [51]. The plotted curves represent values calculated with eq. *(8)* for N = 1-5 and variable R/R_0. The dotted line at R/R_0 = 1.33 was inserted exemplarily as it represents sfGFP fluorophores with a realistic distance of 6.1 nm. **C)** Schematic structure of sfGFP concatemers. For covalently linked GFP moieties, one can assume appropriate proximity to observe the homo-FRET effect on fluorescence anisotropy. **D)** Denaturing PAGE of sfGFP concatemers (c = 1 µM).

In a first and essential step, we wanted to ascertain whether homo-FRET between multiple sfGFP moieties would enable the determination of the exact number of interacting molecules. For this proof of concept, a series of fusion proteins with covalently linked, concatenated sfGFP was generated. There were similar approaches with GFP concatemers that represented protein oligomers, in the past [90, 137, 138].

Based on the plasmid of the monomeric fusion protein 1xGFP, the plasmids for concatenated sfGFP domains were generated. A restriction-based cloning approach with cohesive restriction ends helped to generated multiple repeats of the same sequence (see Methods: Tandem cloning with cohesive restriction ends). In this study, the concatemers 2xGFP,

3xGFP, 4xGFP, and 5xGFP represent dimeric (N = 2), trimeric (N = 3), tetrameric (N = 4), and pentameric (N = 5) protein oligomers, respectively (**Figure 13C**). Similarly to the monomeric 1xGFP, the sfGFP concatemers were expressed in *E. coli* BL21(DE3) pLysS, and purified by IMAC and AEX to reach near homogeneous samples, as shown by denaturing PAGE (**Figure 13D**).

As proven in a denaturing PAGE, the sfGFP concatemers 1xGFP (36.4 kDa), 2xGFP (46.6 kDa), 3xGFP (98.9 kDa), 4xGFP (130.1 kDa), and 5xGFP (161.3 kDa) did not show any major contamination (**Figure 13D**). Moreover, the molecular shape of 1xGFP-5xGFP was investigated.

Former studies tested whether concatemers with N GFP domains would be more elongated, ellipsoid, or spherical (**Figure 14A**) [90, 137, 138]. They determined the diffusion coefficients of the concatemers, as assessed by fluorescence correlation spectroscopy (FCS), and compared the experimental data with theoretical estimations based on the Einstein-Stokes equation for the diffusion of spherical particles in a classical fluid, see eq. *(15)* [117]. All of the recorded diffusion coefficients decreased for increasing N. However, the interpretation of the data was not mutually consistent among the studies. One group classified the concatemer shape as "rather rod-like" than spherical [138]. Others rated the concatemer shape as "rather elongated" than spherical although there was similarity between experimental data and theoretical estimations [90]. Lastly, another group assumed a globular shape [137].

In FCS measurements with the 1xGFP-5xGFP concatemers, conducted by Ganesh Agam, LMU Munich, it was possible to detect decreasing diffusion coefficients (D_{1xGFP} = 67.5 µm² s^{-1}, D_{2xGFP} = 56.2 µm² s^{-1}, D_{3xGFP} = 36.6 µm² s^{-1}, D_{4xGFP} = 38.2 µm² s^{-1}, D_{5xGFP} = 38.2 µm² s^{-1}) which is in line with the published reports (**Figure 14B**). Compared to estimations with the Einstein-Stokes equation, there was congruence for all concatemers except 3xGFP. This observation suggests that the shape of concatemers in this study is not perfectly spherical but putatively ellipsoid. It will be discussed later whether this gives further indications that covalently linked, concatenated sfGFP can structurally represent sfGFP complexes that are grouped together by protein-interaction.

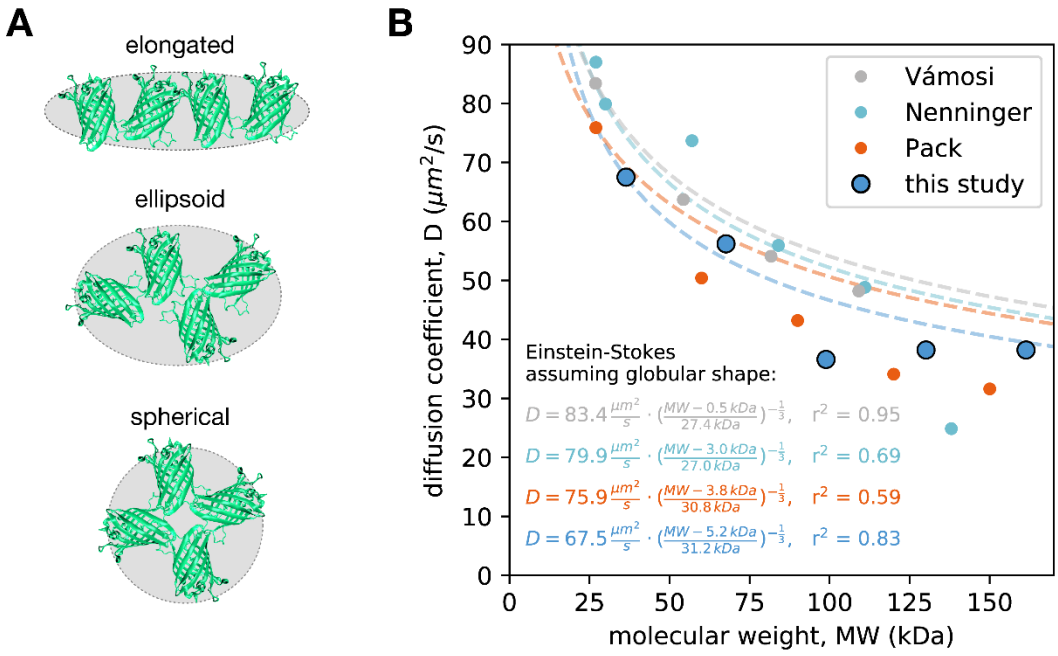


Figure 14: The diffusion coefficient of sfGFP concatemers, recorded by Ganesh Agam, LMU Munich, decreases with increasing GFP domains which is in line with other reports and indicates a ellipsoid macromolecular structure. A) Schematic approach of possible macromolecular structures for four concatenated GFP domains. **B)** The diffusion coefficient for sfGFP concatemers (1xGFP, 2xGFP, 3xGFP, 4xGFP, 5xGFP) is decreasing with increasing molecular weight, thus, with the number of concatenated GFP domains. A similar decrease in motility with an increased number of GFP domain for GFP concatemers was already described by others [90, 137, 138]. Purified superfolder GFP concatemers (1xGFP, 2xGFP, 3xGFP, 4xGFP, 5xGFP) were diluted to 10 nM in fluorescent-free PBS and the diffusion coefficient measured using fluorescence correlation spectroscopy (FCS; λ_{ex} = 488 nm, λ_{em} = 512 nm). The values of Nenninger et al. (2010) were measured in cells, and adjusted assuming an 8.9-fold increase in diffusion in water, as seen for their GFP *in vitro* control. The data was fitted (dashed lines) with a variation of the Einstein-Stokes model for spherical particles in a classical fluid, eq. *(16)*, assuming a globular molecular shape. Negative deviation of the measured and published data from the Einstein-Stokes model suggests that the concatemer shape is not perfectly globular.

5.1.3 Steady-state anisotropy of sfGFP concatemers

The steady-state anisotropy of 1xGFP and the sfGFP concatemers 2xGFP, 3xGFP, 4xGFP, and 5xGFP was determined with a microplate reader capable to measure fluorescence polarization. Therefore, we recorded the parallel and perpendicular components of the emitted light after excitation with polarized light (λ_{ex}: 488 nm; λ_{em}: 520 nm). Then the steady-state anisotropy could be calculated for all tested samples, according to eq. *(2)* (**Figure 15A**).

The steady-state anisotropy (r_{SS} ± standard deviation, n = 3) decreased from 1xGFP (0.297 ± 0.005) to 2xGFP (0.271 ± 0.007) and from 2xGFP to 3xGFP (0.251 ± 0.004). The differences were statistically significant (unpaired Student's *t*-test). For fusion proteins with N=1-3 GFP moieties, thus, r_{SS} behaves inversely proportional to N. However, r_{SS} did not

significantly differ between constructs with more than three concatenated GFP subunits: 3xGFP, 4xGFP (0.250 ± 0.014), and 5xGFP (0.254 ± 0.006). This observation is in line with anisotropy data for EGFP concatemers by Vámosi et al. (**Figure 15B**) [90].

Since 1xGFP, 2xGFP, and 3xGFP can be unequivocally differentiated, in chapter 5.2.3 these concatemers are used as standard proteins with reference steady-state anisotropies, valid for monomer, dimer, and trimer comparison.

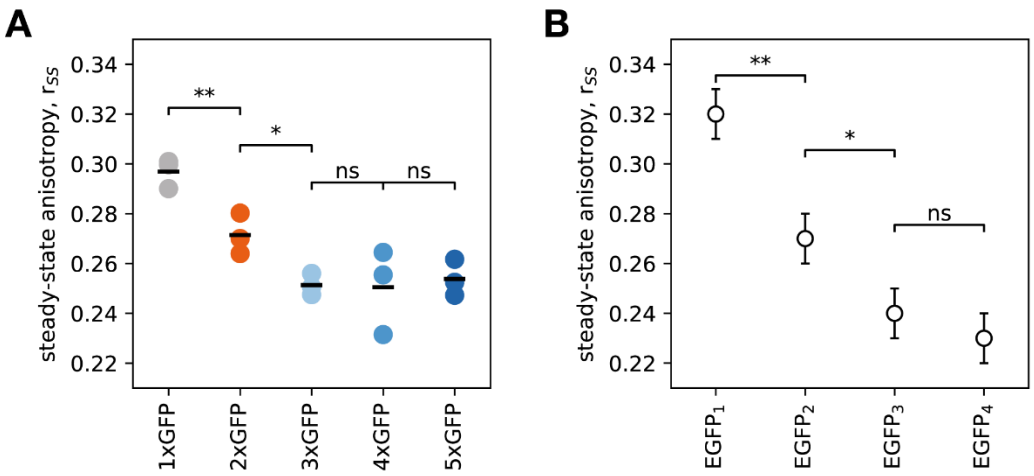


Figure 15: The steady-state anisotropy of sfGFP concatemers in comparison with similar constructs from Vámosi et al. [90]. A) The steady-state anisotropy of sfGFP concatemers. The determined values for r_{SS} are inversely proportional to N, the number of GFP domains in the concatemer, for $N \leq 3$. We could not observe significantly different steady-state anisotropies for concatemers with more than three sfGFP subunits. Purified sfGFP constructs were diluted to 2 µM in fluorescence-free PBS. Here displayed are the single values (6 technical replicates) and means (n = 3). Differences between fusion proteins were classified with an unpaired Student's *t*-test and ranked by their two-tailed p-values (ns = not significant, $p > 0.05$; * = significant, $p \leq 0.05$; ** = strongly significant, $p \leq 0.01$). **B)** Steady-state anisotropy of EGFP-based concatemers by Vámosi et al., directly taken from Table 2 of the original publication [90]. Anisotropy data and the representing standard deviations (±0.01; indicated by error bars) were not generated in that study and serve only for the purpose of data comparison.

5.1.4 Time-resolved anisotropy of sfGFP concatemers

In a next step, the sfGFP concatemers were investigated via time-resolved fluorescence anisotropy, measured by Ganesh Agam at the Chair of Physical Chemistry I, LMU Munich. Time-resolved anisotropy decays of fluorophores interacting via homo-FRET offer a detailed insight into the homotransfer rate k_{FRET} and consecutively on the distances between two interacting fluorophores. In case of large and slowly rotating fluorophores, such as GFP, the anisotropy decay can be best understood with the model for hindered rotors [47, 48, 51, 114]. Using this model, k_{FRET} could be easily determined from the time-resolved decay with the help of eq. *(7)*.

Recording the time-resolved anisotropy for every sfGFP concatemer, distinguishable responses could be detected for every fusion protein (**Appendix Fig. 4**). Furthermore, Ganesh Agam could determine the fluorescence lifetimes from the fluorescence intensity decay, generated from the combination of the perpendicular and the parallel emission signals. When analyzing the fluorescence anisotropy decay of sfGFP concatemers over a 10 ns time scale, the homotransfer rate k_{FRET}, the limiting anisotropy r_0, and the anisotropy for t approaching infinity r_∞ could be obtained. It must be noted that values for 1xGFP were calculated with help of eq. *(7)* which characterizes the anisotropy decay solely by homo-FRET dependency. Since there is no homo-FRET for the monomeric 1xGFP, values given in **Table 6** are not valid but listed to complement the ensemble.

The fluorescence lifetime is slightly but steadily increasing with a growing number of concatenated sfGFP domains (**Table 6**). Interestingly, a previous report on EGFP concatemers showed the opposite effect, with lifetimes decreasing for higher concatemers [90]. This suggests that the fusion protein arrangement affects the fluorescence lifetime although no explanation can be given for this observation.

Values of r_∞ decrease nearly linearly with increasing *N* (**Table 6**). Assuming the model of a hindered rotor, eq. *(7)*, r_∞ can be theoretically assumed to be near the steady-state anisotropy r_{SS} [113]. However, r_∞ in our case does not give the same value as r_{SS} acquired with a microplate reader. In fact, it is known that values for r_∞ generally show the strong tendency to be lower than measured r_{SS} values [51]. This can be explained by a stronger reduction of r_∞ based on cluster formation than seen for r_{SS} [51].

We found that the values for homotransfer rate k_{FRET} were increased for growing N which was originally not expected (**Table 6**). A similar effect can be seen for homo-FRET efficiency E which is a direct derivative of k_{FRET}, see eq. *(4)*. According to eq. *(3)*, k_{FRET} is a function of the spatial distance R between interacting fluorophores. In this relation, the fluorescent lifetime τ and the Förster radius R_0 are both fluorophore-specific constants. A changing k_{FRET} would imply the shortening of inter-fluorophore distances, as denoted in **Table 6**. As the linker length between single GFP domains remains unchanged for increased N, one could assume that the enhanced k_{FRET} is a result of condensed GFP packing in a spatially limited volume where homo-FRET is possible (R/R_0 < 1.8) [46].

Table 6: Analysis of sfGFP concatemer time-resolved anisotropy data

	τ (ns) [1]	k_{FRET} (ns^{-1}) [2]	R (nm) [3]	r_0 [4]	r_∞ [5]
1xGFP	2.47	(0.033)*	(7.00)*	(0.408)*	(0.162)*
2xGFP	2.54	0.055	6.38	0.365	0.201
3xGFP	2.57	0.069	6.14	0.36	0.168
4xGFP	2.66	0.071	6.07	0.321	0.102
5xGFP	2.89	0.075	5.94	0.264	0.062

Raw data provided by Ganesh Agam, LMU Munich. [1] Fluorescence lifetime. [2] Homotransfer rate. [3] Distance between interacting fluorophores. [4] Initial / limiting / fundamental anisotropy. [5] Anisotropy for t approaching infinity. ( )* Data fitted with the model for hindered rotor, eq. *(7)*, although no homo-FRET is expected; to receive the rotational correlation time ϕ, a fit with eq. *(6)* can be applied.

5.1.5 Combining steady-state and time-resolved fluorescence anisotropy to determine the number of concatenated sfGFP subunits

Being confronted with GFP fusion proteins of unknown stoichiometry, the r_{SS} values of GFP concatemers 1xGFP, 2xGFP, and 3xGFP can be used as reference points for monomer, dimer, and trimer, respectively.

Combining data from steady-state and time-resolved anisotropy measurements in the model of Runnels and Scarlata, eq. *(8)* defines a gold standard to obtain the number of interacting fluorophores N [46]. The calculations require the anisotropy of the monomer r_1(sfGFP concatemers: r_1 = 0.297) and the measured r_{SS} for the protein of interest. Furthermore, the homotransfer rate k_{FRET} and fluorescence lifetime τ (values for concatemers 1xGFP-5xGFP

are given in **Table 6**) have to be known. Some studies assume that the rear part of eq. *(8)* can be neglected [46, 112, 113, 139] as the decisive parameter r_{ET}, the average emission anisotropy following energy transfer, was determined to be very low at r_{ET} = 0.016 for fluorophores that are randomly oriented [116]. N was calculated with both, r_{ET} = 0 and r_{ET} = 0.016, for the sfGFP concatemers (**Table 7**). The calculations yield acceptable values for 1xGFP and 2xGFP but failed to correctly determine N for 3xGFP, as well as, 4xGFP and 5xGFP.

With GFP fusion proteins, by contrast, the orientation of fluorescent protein subdomains is not random. In this case, r_{ET} substantially contributes to eq. *(8)* (**Figure 13B**) [51]. The parameter r_{ET}, however, has to be determined empirically [51]. With the premise that N = 2 for 2xGFP and N = 3 for 3xGFP, we could solve eq. *(8)* for the parameter r_{ET}. This yielded r_{ET}(2xGFP) = 0.102, r_{ET}(3xGFP) = 0.108, and an average r_{ET} = 0.105.

Changing r_{ET} to 0.105, the subunit stoichiometry of 1xGFP, 2xGFP, and – most importantly – 3xGFP could be calculated more precisely (**Table 7**). The characterization of concatemers with N > 3 still proved to be difficult.

Table 7: Calculating the number of concatenated sfGFP from fluorescence anisotropy, using the model by Runnels & Scarlata [46]

	$N_{SS}(r_{ET} = 0)$ [1]	N_{SS} $(r_{ET} = 0.016)$ [2]	N_{SS} $(r_{ET} = 0.105)$ [3]	$N_{SS,standard}$ [4]
1xGFP	1.0	1.0	1.0	1.0
2xGFP	1.8	1.8	2.2	2.0
3xGFP	2.2	2.3	3.1	3.1
4xGFP	2.2	2.2	3.0	3.1
5xGFP	2.0	2.0	2.6	2.9

[1] Number of subunits N that interact via homo-FRET. N was calculated with Runnels and Scarlata's model, eq. *(8)*, and excluding the rear part by r_{ET} = 0 [46].

[2] N calculated with eq. *(8)*, by assuming r_{ET} = 0.016 [46, 116].

[3] N calculated with eq. *(8)* assuming a value for r_{ET} = 0.105, determined from concatemer standards. For the calculations, individual k_{FRET} values from time-resolved anisotropy data were used.

[4] N calculated with eq. *(8)* and r_{ET} = 0.105. Instead of individual values, a mean value for k_{FRET} (= 0.067 ns^{-1}) and τ (= 2.62 ns^{-1}) was used that was determined from standard concatemers.

To this point, the calculations were based on individual k_{FRET} and τ values that were obtained from time-resolved measurements. Since these values showed an unexpected but subtle variability, it was tested whether the stoichiometry calculations with mean k_{FRET} (= 0.067 ± 0.004 ns^{-1}) and mean τ (= 2.62 ns^{-1}) could deliver an improved determination of N.

The usage of a uniform set of parameters would facilitate the determination of N for any new, structurally similar sfGFP fusion protein. Calculations with this standard parameter set (r_1 = 0.297, k_{FRET} = 0.067 ns^{-1}, τ = 2.62 ns^{-1}, r_{ET} = 0.105) yielded the most accurate values, corresponding well to the known monomeric (1xGFP), dimeric (2xGFP), or trimeric (3xGFP) structures.

5.1.6 Fractional photobleaching and steady-state anisotropy to differentiate between GFP concatemers.

It had previously been proposed that fractional photobleaching in combination with steady-state anisotropy can resolve the stoichiometry of fluorophore labeled molecules, independently from time-resolved anisotropy analysis [57]. Here, the sfGFP concatemers can be used as model proteins, too, representing non-covalently assembled oligomers.

By irradiation with a strong 445 nm laser, GFP subunits can be irreversibly "switched off". This fractional photobleaching of GFP was thoroughly investigated in the chapter 5.3, pp. 81-91. The irrevocable deletion of fluorescence of FRET acceptors blocks the depolarization caused by energy transfer [140]. The anisotropy of a fluorophore labeled complex with N = 3, for instance, is increased when the fluorophores are fractionally photobleached, gradually exhibiting the anisotropy of complexes with N < 3. The anisotropy behavior for fractionally photobleached fluorophores with a cluster size N was theoretically described by Yeow and Clayton [57]. According to their formalism, the continuous fractional photobleaching does not affect the r_{SS} of monomers, increases the r_{SS} of dimers quasi-linearly, and increases the r_{SS} of trimers and higher oligomers in an exponential fashion.

When we determined the r_{SS} of sfGFP concatemers upon fractional photobleaching, these characteristic anisotropy patterns could be observed (**Figure 16A**). It was possible to differentiate between 1xGFP, 2xGFP and 3xGFP. The behavior of the dimer r_{SS} is linear and the behavior of trimers and higher oligomers is exponential. Nevertheless, the differentiation

between oligomers with $N > 3$ was as challenging as already described previously for unbleached samples.

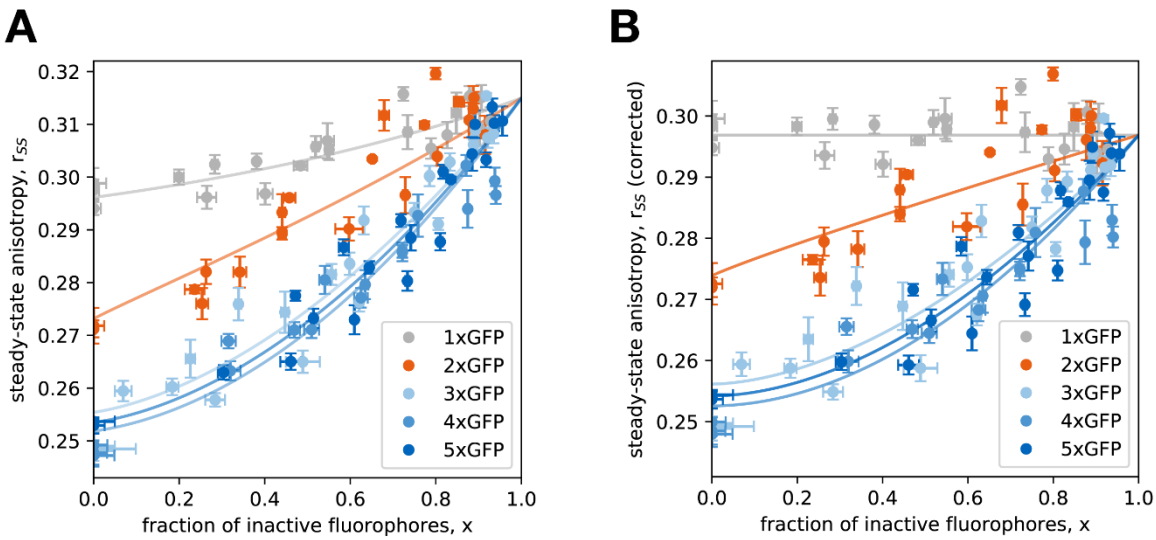


Figure 16: Steady-state anisotropy in combination with fractional photobleaching shows distinguishable, protein-specific behavior for monomeric, dimeric, and trimeric sfGFP concatemers. A) Steady-state anisotropy of the sfGFP concatemers (1xGFP, 2xGFP, 3xGFP, 4xGFP, 5xGFP) for an increased fraction of inactive / photobleached fluorophores x. The anisotropy data was fitted with eq. *(14)* and a = 0.10. Purified fusion proteins were diluted to 2 µM in concatemer sample buffer before they were irradiated with 445 nm laser light (300 mW) for fixed periods of time. After every bleaching step, the steady-state anisotropy and the residual fluorescence intensity was recorded (λ_{ex} = 488 nm; λ_{em} = 520 nm). The fraction of inactive fluorophores was calculated, eq. *(10)*. The raw steady-state anisotropy data was standardized to remove day-to-day variance, see Figure 12. **B)** Photobleaching data after correction for monomeric anisotropy increase, generally in eq. *(22)* and specified in eq. *(23)* (o = 0.011, p = 0.974, q = -0.012). The data was fitted with eq. *(14)* and a = 0.10.

However, we detected a slightly exponential increase of anisotropy for the monomeric 1xGFP. This observation was not sfGFP-specific since another monomeric fusion protein holding EGFP [18] also showed an anisotropy increase for fractional photobleaching, even with a stronger curvature (**Figure 17**).

This exponential response can be described with the following equation:

$$\Delta r_{SS}(x) = o \cdot e^{(p \cdot x)} + q \qquad (22)$$

where o, p, and q are the variables to exactly describe the curvature (o and p) and the y-intercept (q) of the graph.

A non-linear least square fit was used to determine o, p, and q:

$$\text{sfGFP monomer: } \Delta r_{SS}(x) = 0.011 \cdot e^{(0.974 \cdot x)} - 0.012$$
$$\text{EGFP monomer: } \Delta r_{SS}(x) = 0.001 \cdot e^{(4.056 \cdot x)} - 0.003 \qquad (23)$$

The exponential fits outcompete linear fits that yielded worse r^2 values, for both, the EGFP and sfGFP-containing monomer. As 1xGFP is monomeric, this increase in anisotropy is independent of any homo-FRET effect. It also occurs similarly for all fusion proteins holding sfGFP or EGFP. For this reason, the photobleaching responses in **Figure 16A** were corrected with the representative function of the sfGFP monomer, eq. *(23)*, yielding an even clearer picture for the differentiation between monomer, dimer and trimer (**Figure 16B**).

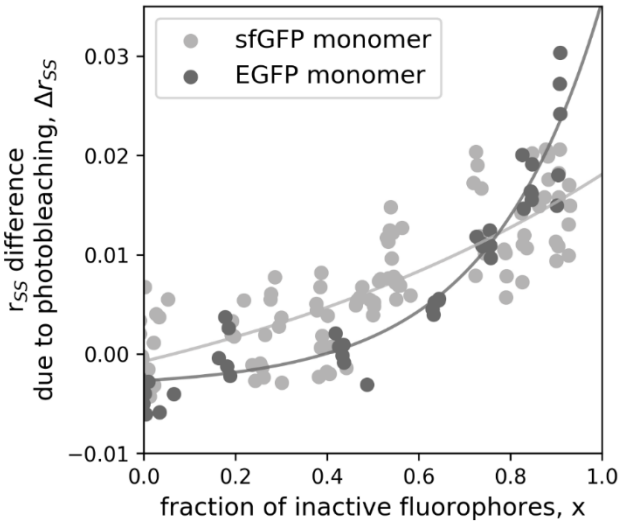


Figure 17: The steady-state anisotropy of the monomeric GFP fusion proteins does not stay stationary but increases for a growing fraction of inactive fluorophores. The change in steady-state anisotropy of sfGFP- and EGFP-based monomers was recorded for a growing fraction of photobleached fluorophores x, calculated with eq. *(10)*. Therefore, raw anisotropy was recorded as described in the Methods section. The anisotropy values of irradiated samples were subtracted by the values of non-irradiated controls to yield Δr_{SS}. The data was fitted with eq. *(23)*. sfGFP monomer: $\Delta r_{SS} = 0.011 \cdot \exp(0.974 \cdot x) - 0.012$. EGFP monomer: $\Delta r_{SS} = 0.001 \cdot \exp(4.056 \cdot x) - 0.003$. Values for Δr_{SS} exceeded a significant threshold (standard deviation of r_{SS} between two measurements ≈ 0.005) when around 50% of the fluorophores had been inactivated.

5.1.7 Determination of subunit stoichiometry via fractional photobleaching and steady-state anisotropy

Yeow and Clayton introduced several ways to use the steady-state anisotropy after fractional fluorophore labeling f to determine the number of fluorophores N that interact via homo-FRET [57]. Fractional photobleaching x is the reversed process of fractional labeling and therefore can be expressed by $x = 1 - f$.

In a simplistic approach, Yeow and Clayton described the steady-state anisotropy after photobleaching with a polynomial function [57]. The order of the polynomial increases with N with coefficients derived from the Pascal's triangle and individual parameters r_A, r_B, …, r_x, see eq. *(11)*. The functions do not include homotransfer or molecular rotation.

The photobleaching data, (**Figure 16B**), was fitted with polynomial functions for N = 1-5 and tested via r^2 value comparison whether 1xGFP is represented best by the polynomial N = 1, 2xGFP is represented best by the polynomial N = 2, etc. Unexpectedly, for all of the concatemers, the polynomial function with the highest N (= 5) has the highest r^2 value. Thus, every photobleaching response is characterized as pentameric (**Figure 18**; **Table 8**).

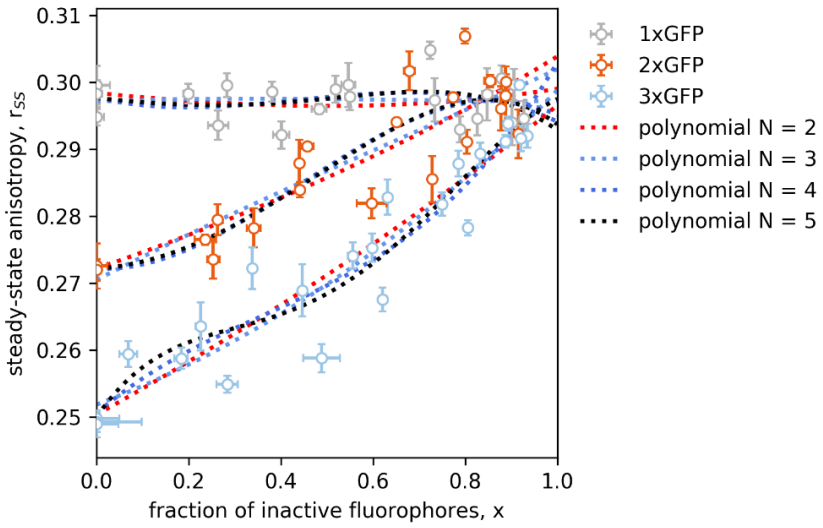


Figure 18: Modelling the steady-state anisotropy behavior for fractionally photobleached sfGFP concatemers with Yeow and Clayton's polynomial approach. The anisotropy behavior after fractional photobleaching of sfGFP concatemers was fitted with eq. *(11)*. Datapoints for 4xGFP and 5xGFP were not plotted for reasons of clarity as they are mostly overlapping with datapoints for 3xGFP. The function with the highest tested polynomial gained the best fit for all of the fusion proteins, becoming visible at highest r^2 values, listed in Table 8.

This observation can be explained by overfitting which usually occurs for models with too many parameters and datasets with not enough data points. Both could be responsible for the misleading predictions of N in Yeow and Clayton's polynomial approach.

Table 8: Determination of *N* via Yeow and Clayton's polynomial approach, eq. *(11)* [57]

	r^2 values from fitting the data with polynomial functions				
	f (N = 2)	**f (N= 3)**	**f (N= 4)**	**f (N= 5)**	$N_{predicted}$
1xGFP	-0.0868	-0.0058	0.0585	0.0626	5
2xGFP	0.7554	0.7618	0.7844	0.7847	5
3xGFP	0.8967	0.9014	0.9079	0.9114	5
4xGFP	0.9300	0.9357	0.9358	0.9493	5
5xGFP	0.8369	0.8808	0.8852	0.8861	5

In another approach, they described the photobleaching behavior of oligomers with a two-state bimodal model, see eq. *(12)* [57]. This model comprised two parameters, the anisotropy of the monomer r_1 and the parameter f_{non} that is the fractional fluorescence of non-interacting and thus monomeric fluorophores. We assessed f_{non} and N with a least square fit of the photobleaching data (**Figure 16B**) which yielded very high values for f_{non}, near its maximum f_{non} = 1 (**Table 9**). For increasing N, the values of f_{non} slightly decrease, yet, still remain close to 1. According to its definition, high values for f_{non} imply that the majority of the investigated fluorescent proteins are arranged as monomers. However, for the sfGFP concatemers, which are covalently linked assemblies of N concatenated GFP moieties, no predominant monomer fraction was detected in denaturing PAGE (**Figure 13D**). In addition, the calculated values for N were significantly different from the inherent N of sfGFP concatemers.

Therefore, it was not possible to describe the photobleaching response of the sfGFP concatemers and the correct N with the two-state bimodal model by Yeow and Clayton. However, we recognized that values of f_{non} were decreasing with the stoichiometry of the complexes (**Table 9**), similarly as already found for the steady-state anisotropy (**Figure 15A**). When plotting f_{non} in dependence of the number of sfGFP moieties in the concatemer N, a nearly linear behavior became apparent for monomer, dimer, and trimer. In

contrast, f_{non} is stagnating for concatemers with $N \geq 3$ which is also in line with observations with steady-state anisotropies (**Figure 15**). Evidently, f_{non} itself is a function of N (**Figure 19**) which is strongly supported by parallels between the steady-state anisotropy in Runnels and Scarlata's model – inversely proportional for N – and f_{non} from Yeow and Clayton, see chapter 3.4.

Table 9: Determination of *N* via Yeow and Clayton, eq. *(12)* [57], and the new theoretical model introduced in this study, eq. *(14)*

	Yeow and Clayton, eq. *(12)*		**this study, eq. *(14)***	
	f_{non} [1]	$N_{x,fnon}$ [2]	$N_{x,a}$ [3]	$N_{x,R\&S}$ [4]
1xGFP	1.002	8.2	1.0	1.0
2xGFP	0.912	1.7	1.9	1.9
3xGFP	0.848	2.3	2.7	2.8
4xGFP	0.846	2.8	2.9	3.0
5xGFP	0.857	2.9	2.8	2.9

[1] Fraction of non-interacting, monomeric fluorophores, according to Yeow and Clayton [57].
[2] Number of subunits that interact via homo-FRET, based on photobleaching data and eq. *(12)*.
[3] Number of subunits that interact via homo-FRET, based on photobleaching data and eq. *(14)*.
[4] Number of subunits that interact via homo-FRET, based on photobleaching data and eq. *(13)*.

As a consequence, the dependence of f_{non} on N is reciprocal and could be best explained with $f_{non} = \frac{1+a}{1+N \cdot a}$ and parameter a_{sfGFP} (= 0.10). The value of a_{sfGFP} was determined from the sfGFP concatemers with an optimal fit for N = 1-3 in a least squares approach. Values for 4xGFP and 5xGFP were excluded. Subsequently, we replaced f_{non} in eq. *(12)* with the reciprocal function and calculated N_x again from the photobleaching data. The obtained values resemble those values yielded for $N_{SS,standard}$, the declared "gold standard" requiring time-resolved anisotropy analysis and photophysical parameters such as k_{FRET}, τ, and r_{ET}.

When the whole Runnels and Scarlata equation, eq. *(8)*, was integrated into the Yeow and Clayton formalism, eq. *(12)*, and the photophysical parameters k_{FRET}, τ, r_{ET} were inserted, the predictive power could only slightly be improved. The resulting N_x values did not significantly differ for the calculations with a_{sfGFP}. This confirms that the combination of

steady-state anisotropy, photobleaching, and a simple model could deliver results comparable to those obtained using sophisticated equipment.

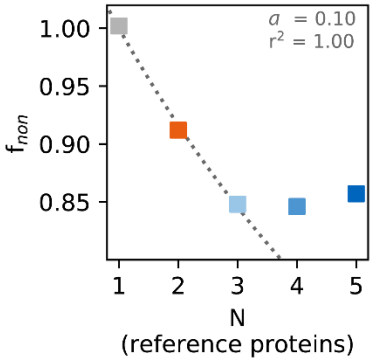


Figure 19: Parameter f_{non} in dependence of N GFP moieties in sfGFP concatemers that interact via homo-FRET. Values for f_{non} were calculated via eq. *(12)* from steady-state anisotropy data upon fractional photobleaching of sfGFP concatemers. A non-linear least squares approach for $f_{non} = \frac{1+a}{1+N \cdot a}$ was used to determine parameter a (valid for N=1-3).

Using our sfGFP concatemers, we could then show that it was feasible to distinguish between a complex of one, two, and three GFP reporters interacting via homo-FRET. This could be achieved both, with steady-state and with time-resolved anisotropy. It was further possible to determine biophysical parameters with the anisotropy decay in timescales of 10 ns by Ganesh Agam. Apart from directly comparing absolute numbers of r_{SS}, we confirmed the number of subunits N that interact via homo-FRET with the well-established model by Runnels and Scarlata. This approach however was only possible with parameters assessed by time-resolved anisotropy. A crucial new result obtained here is that recording the steady-state anisotropy of fractionally photobleached sfGFP subunits gave good approximations even without relying on time-resolved measurements. We now proceed to analyze the oligomerization of non-covalently linked GFP fusion proteins.

5.2 Determining the oligomeric state of non-covalently linked GFP fusion proteins via fluorescence anisotropy

5.2.1 Generating GFP fusion proteins that form non-covalently bound oligomers

As models of non-covalently linked protein assemblies, chimeric proteins were generated containing GFP as the fluorophore label and oligomerizing coiled-coil domains. Coiled-coil domains are well studied peptides that exist in various stoichiometries, as specified in the introductive sections.

In fusion proteins derived from parental 1xGFP protein, a coiled-coil element with known stoichiometry (GCN4-p1: dimer; GCN4-pII: trimer) or with unclear or unknown stoichiometry (ph3a and GCN4-pAA) was linked to a sfGFP domain (**Figure 20A**). Similar to the sfGFP concatemers, Gly-Ser enriched linker regions were used to connect the GFP moiety with the oligomerizing coiled-coil element. Additionally, tags were included for protein purification (xtHis) and immunochemical detection (FLAG). The primary structure of these sfGFP-based constructs only varied in the integrated coiled-coil peptide sequences (**Figure 20A**).

In parallel, fusion proteins with a different structural composition were investigated. These proteins were introduced by Cristie-David et al. [18] and hold *de novo* generated coiled-coil domains [14, 15] and EGFP instead of the sfGFP variant (**Figure 20B**). The stoichiometry of these EGFP-based fusion proteins had already been examined in the past by Cristie-David et al. [18], and they were tested again in this study. As a monomeric control for EGFP-based construct, an EGFP fusion protein with minor N-terminal extensions (polyHis tag and linker region) was used [18].

The GFP-coiled-coil fusion proteins were expressed in *E. coli* BL21(DE3) and purified with an altered purification protocol compared to the sfGFP concatemers. Applying the same purification protocol as for the sfGFP concatemers, we were confronted with several challenges including massive precipitation and minimal protein yield. The addition of 30% glycerol during the purification process proved to be beneficial.

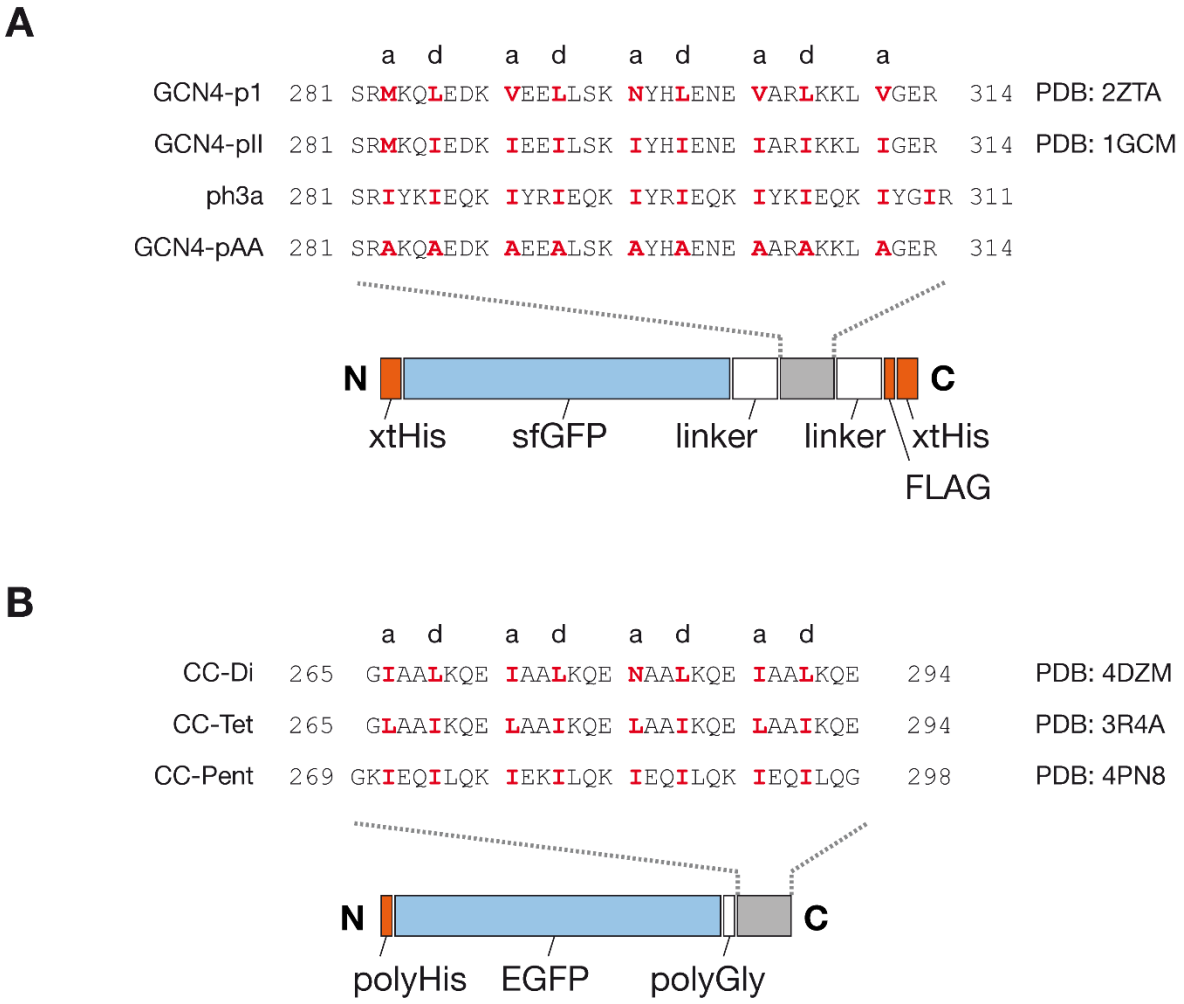


Figure 20: The amino acid sequences of all coiled-coil elements that were integrated in GFP fusion proteins in this study. A) The amino acid sequences of GCN4-p1 [35], GCN4-pII [12], the de novo generated peptide ph3a, and GCN4-pAA were inserted into a modular sfGFP-based fusion protein. **B)** The amino acid sequences of the de novo coiled-coil peptides CC-Di [14], CC-Tet [14], and CC-Pent [15] were inserted in an EGFP-based fusion protein. This protein set-up was introduced by Cristie-David et al. [18] who supplied us with the original plasmids generated for their study.

5.2.2 Biochemical analysis of the stoichiometry of oligomerizing GFP fusion proteins

The purified constructs were tested for integrity via denaturing PAGE (**Figure 21A&B**). Most of the constructs exhibited a clear and distinctive band corresponding to an intact fusion protein, particularly the EGFP-based chimeras (**Figure 21B**). With some of the sfGFP-based contaminations and fragments of minor molecular weight could be detected (**Figure 21A**). All GFP-coiled coil fusion proteins were analyzed for their stoichiometry via native PAGE and size exclusion chromatography (SEC) (**Figure 21C-F**).

In native PAGE experiments, nearly all of the proteins migrated into the 10% polyacrylamide gel with sfGFP-ph3a as the only exception. However, the migration behavior varied fundamentally for the tested samples, also between oligomers of the same stoichiometry and molecular weight but different structural composition. Comparing the predicted isoelectric points [141] for the tested samples, a strong difference between the more acidic EGFP- and the less acidic sfGFP-based proteins became visible, which might be a reason for the migration discrepancy. As only acidic protein species (pI < 7) are able to intrinsically migrate to the anode [25, 27], the only protein whose predicted isoelectric point is clearly above this threshold, sfGFP-ph3a, did not migrate (**Figure 21E**). As the number of charged groups is foremostly responsible for the migration into the gel, the use of a general protein standard in order to relate gel migration of any protein to its stoichiometry can be therefore regarded as difficult. Therefore, native PAGE had only a supportive role when characterizing the samples biochemically.

In SEC analysis (**Figure 21F**), a calibration curve with protein standards was recorded beforehand in order to derive the molecular weight and protein stoichiometry from sample retention time (**Figure 10**). This enables the comparison between differently structured protein assemblies. First, we analyzed fusion proteins that held GCN4-p1 and its variant GCN4-pII. In the past, it had been demonstrated by X-ray structures [35] and nuclear magnetic resonance spectroscopy [142, 143] that the isolated GCN4-p1 peptide is a dimer. In isolation, the GCN4-pII peptide forms a trimer according to its X-ray structure [144]. Integrated in the chimeric set-up, fusion proteins with GCN4-p1 and GCN4-pII eluate at retention times corresponding to the two- and threefold monomer molecular weight, respectively. Thus, their dimeric and trimeric stoichiometry could be confirmed. Gradually slower migration on the native PAGE in the order monomeric 1xGFP, dimeric sfGFP-GCN4-p1, and trimeric sfGFP-GCN4-pII further supports this interpretation. By mutating all interfacial residues in GCN4-p1 to alanine, the construct sfGFP-GCN4-pAA was assumed no longer to form oligomers. Surprisingly, two distinct peaks could be observed in SEC analysis with two peaks lying in between dimer and trimer and dimer and monomer (**Figure 21F**, **Appendix Fig. 3**).

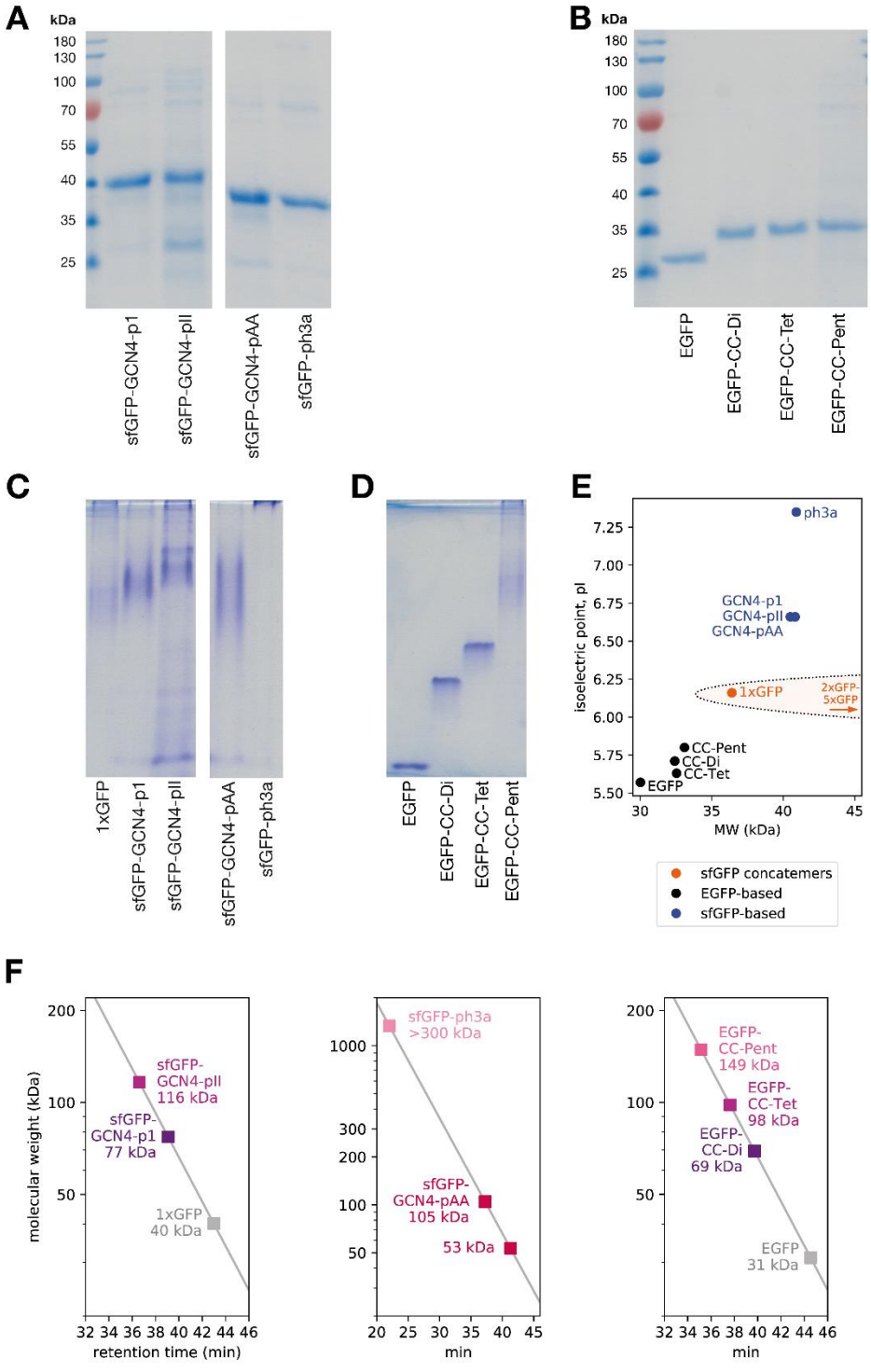


Figure 21: Biochemical analysis of non-covalently associated GFP fusion proteins. A) Denaturing PAGE of sfGFP-based fusion proteins. The concentration of all loaded proteins on denaturing as well as on native PAGE was 2 µM. **B)** Denaturing PAGE of EGFP-based fusion proteins. **C)** Native PAGE of sfGFP-based fusion proteins. **D)** Native PAGE of EGFP-based fusion proteins. **E)** Predicted isoelectric point [141] for all GFP fusion proteins. **F)** Size exclusion chromatography with GFP-coiled coil fusion proteins. sfGFP-based fusion proteins of unclear / unknown stoichiometry are displayed in the central panel. The data points represent UV and fluorescence intensity maxima (**Appendix Fig. 3**).

The initial assumption that mutating interfacial residues to Ala would monomerize the peptide could not be substantiated. Hence, sfGFP-GCN4-pAA was excluded from further anisotropy experiments.

Lastly, we tested the *de novo* coiled-coil peptide ph3a integrated into the sfGFP fusion protein. The *de novo* peptide ph3a was designed from scratch with the aid of CCBuilder, a tool to test coiled-coil architecture *in silico* (**Figure 22**) [16]. With trimer helix properties, the coiled-coil sequence of ph3a was generated starting from the amino acid sequence of the trimeric GCN4-pII (RMKQIEDK IEEILSK IYHIENE IARIKKL IGER). Gradually altering the sequence on specific positions, the changes were validated for their theoretical stability via Rosetta [145] and BUDE Score [146, 147] by the CCBuilder tool.

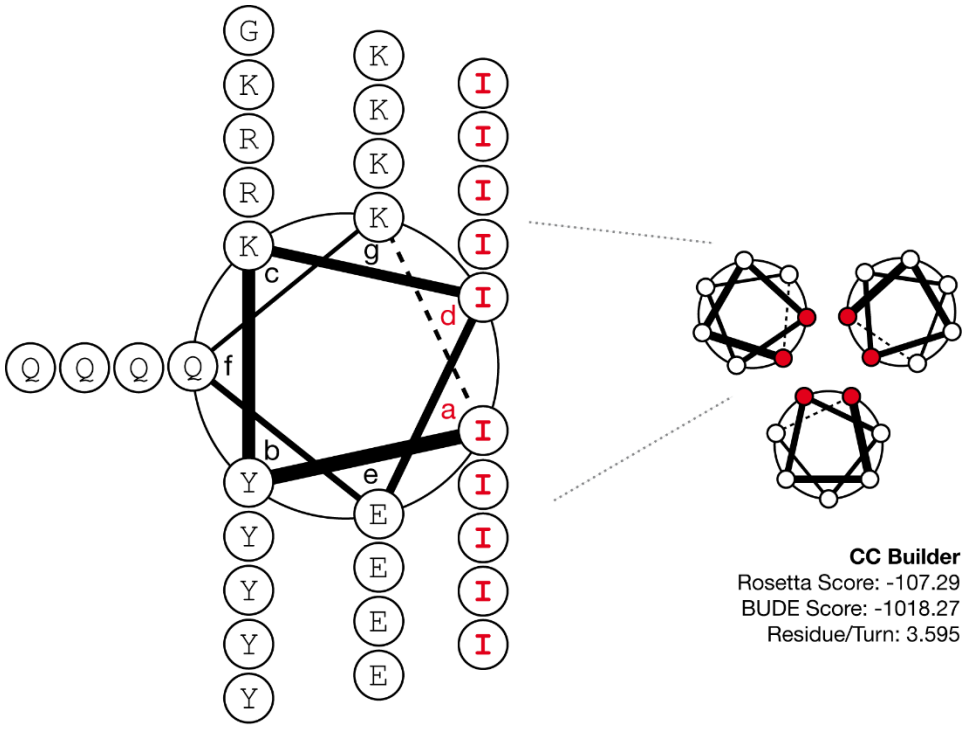


Figure 22:The generation of the *de novo* coiled-coil ph3a. Helix properties were chosen from CC-Tri, a *de novo* trimer (radius = 6.34 Å, pitch = 194 Å, interface angel = 19.98 °) [16]. The sequence for the ph3a peptide was based on the heptad pattern of GCN4-pII. Coiled-coils with isoleucine residues at the hydrophobic interface (*a* and *d* position) were classified as trimers in the past [12]. Intending a trimeric fusion protein, isoleucine was selected for heptad position *a* and *d*. As repeatedly seen for the natural coiled-coil element in GCN4, glutamic acid was selected for the *e* position and lysine for the *g* position, mutually stabilizing the protein via electrostatic interactions. For position *c* and *f*, polar and hydrophilic residues arginine, lysine, and glutamine were chosen. Position *b* was filled with tyrosine. Using the CC Builder tool introduced by Wood et al. [16], the resulting coiled-coil **p**eptide was **h**ypothesized as a strong parallel **tri**mer (eponymously: **ph3a**; Rosetta Score: -107.29; BUDE Score: -1018.27; residues per turn: 3.595).

Besides its hydrophobic Ile-based core motif, the sequence for ph3a exhibits further repetitive patterns, i.e. Tyr on heptade position *b*, Glu on *e*, Lys on *g*, Arg, Lys, and Gln on positions *c* and *f* (RIYKIEQK IYRIEQK IYRIEQK IYKIEQK IYGIR). The ph3a peptide was classified as a strong trimer *in silico*. However, there was no trimeric behavior observable in SEC experiments. sfGFP-ph3a eluted with the exclusion volume indicating a higher-order complex formation. In order to test whether anisotropy-based methods could resolve the stoichiometry of the higher-order oligomer, we included sfGFP-ph3a in the subsequent anisotropy analysis.

As an alternative fusion protein subset, EGFP-based fusion proteins were analyzed for their stoichiometry. All of the EGFP-based fusion proteins, initially published by Cristie-David et al. [18], contain *de novo* coiled-coil domains designed by the Woolfson group [14, 15]. These *de novo* peptides CC-Di, CC-Tet, and CC-Pent were characterized in their isolated form as dimer, tetramer, and pentamer, respectively, by crystallography. However, Cristie-David et al. found stoichiometry inconsistencies when integrating these elements in the protein chimeras (**Figure 20B**). Experiments with SEC, analytical ultracentrifugation, and native mass spectrometry characterized CC-Di as a dimer, as expected, but classified the putative tetramer CC-Tet as a trimer instead. Furthermore, CC-Pent could be only detected as a pentamer and prevented from aggregating after changing the linker length [18]. In our SEC experiments, we could confirm the dimeric, trimeric, and pentameric character of CC-Di, CC-Tet, and CC-Pent, supported by native PAGE, (**Figure 21D&F**) and prove the findings by Cristie-David et al. [18].

5.2.3 Distinguishing the oligomeric state non-covalently associated GFP-coiled-coil fusion proteins via fluorescence anisotropy

The steady-state and the time-resolved anisotropies of the GFP-coiled-coil fusion proteins were measured similar to the sfGFP concatemers (**Figure 23**). The samples were again separated in two subsets, sfGFP- and EGFP-based (**Table 10**). For the sfGFP-based subset, the steady-state anisotropies of the dimeric sfGFP-GCN4-p1 (r_{SS} = 0.271) and the trimeric sfGFP-GCN4-pII (r_{SS} = 0.246) are significantly decreased in comparison to the anisotropy of the monomeric 1xGFP. Both steady-state anisotropies resemble the values of their concatemer equivalent 2xGFP (r_{SS} = 0.271) and 3xGFP (r_{SS} = 0.251). In time-resolved

anisotropy measurements, executed by Ganesh Agam (**Appendix Fig. 5**), the homotransfer rates for dimer ($k_{FRET,\,\mathrm{sfGFP-GCN4-p1}}$ = 0.063 ns^{-1}) and trimer ($k_{FRET,\,\mathrm{sfGFP-GCN4-pII}}$ = 0.099 ns^{-1}) differed from their concatemer analogues ($k_{FRET,2xGFP}$ = 0.055 ns^{-1}; $k_{FRET,3xGFP}$ = 0.069 ns^{-1}) which indicates a stronger proximity between GFP reporters. Indeed, this could be affirmed when calculating the fluorophore distance R from k_{FRET} and τ, see eq. *(3)*. Here, the dimeric sfGFP-GCN4-p1 (R = 6.26 nm) and the trimeric sfGFP-GCN4-pII (R = 5.83 nm) show both shorter inter-fluorophore distances compared to their equivalents 2xGFP (R = 6.38 nm) and 3xGFP (R = 6.14 nm), respectively.

SEC analysis indicated that sfGFP-ph3a forms higher-order protein complexes. This result was confirmed by an extraordinarily low steady-state anisotropy (r_{SS} = 0.172) and a much stronger time-resolved decay than detected for any other GFP-coiled-coil fusion protein ($k_{FRET,\ sfGFP-ph3a}$ = 0.230 ns^{-1}). The inter-fluorophore distance R could be calculated and equals 5.03 nm, suggesting a very tightly packed macromolecular structure.

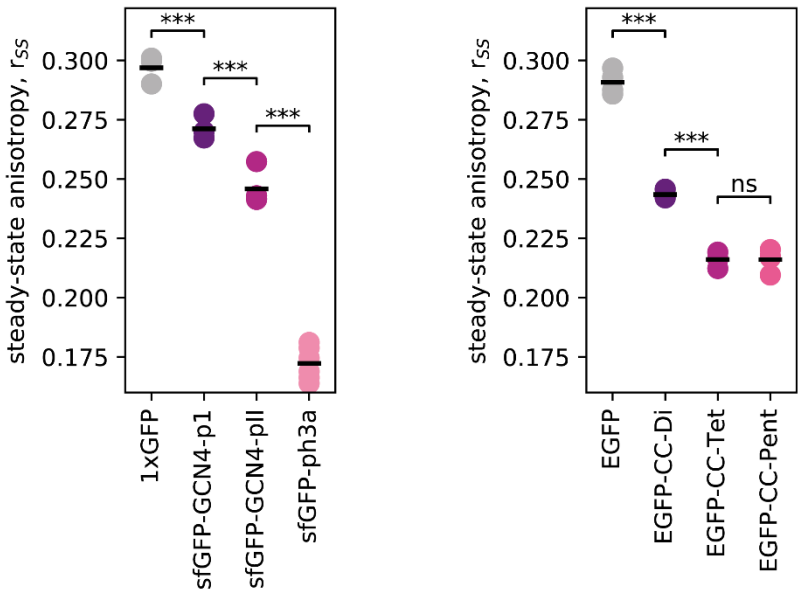


Figure 23: The steady-state anisotropy of GFP-based coiled-coil fusion proteins. Purified GFP-based coiled-coil fusion proteins were diluted to a concentration of 2 µM and were excited with polarized light (λ_{ex} = 488 nm). The fluorescence intensity (λ_{em} = 520 nm) parallel and perpendicular to the excitation light was recorded and the steady-state anisotropy was calculated using eq. *(2)*. Here, the single values and means (n = 4-6) are displayed. Differences between fusion proteins were classified with an unpaired Student's *t*-test and ranked by their two-tailed p-values (ns = not significant, $p > 0.05$; * = significant, $p \leq 0.05$; ** = strongly significant, $p \leq 0.01$; *** = extremely significant, $p \leq 0.001$).

The other fusion protein subset, EGFP-based coiled-coil fusion proteins, exhibited a greater decrease in fluorescence anisotropy than their sfGFP-based analogues (**Table 10**). The anisotropy of the monomeric EGFP (r_{SS} = 0.291) is similar to that of 1xGFP (r_{SS} = 0.297) (difference is not statistically significant). However, the anisotropy of the dimeric EGFP-CC-Di (r_{SS} = 0.243) and trimeric EGFP-CC-Tet (r_{SS} = 0.216) drops distinctly stronger than for the sfGFP-based fusion proteins including sfGFP concatemers. The steady-state anisotropy of the pentameric EGFP-CC-Pent does not show any difference from the trimeric EGFP-CC-Tet which confirms the previous observations for complexes with $N \geq 3$ GFP subunits. From time-resolved data, monomer, dimer, and trimer could be clearly distinguished ($k_{FRET,EGFP-CC-Di}$ = 0.105 ns^{-1}; $k_{FRET,EGFP-CC-Tet}$ = 0.139 ns^{-1}), but no difference was detected between trimer and pentamer ($k_{FRET,EGFP-CC-Pent}$ = 0.137 ns^{-1}). For the EGFP-based protein subset, the distance between interacting fluorophores was lower than for the sfGFP-based fusion proteins (**Table 10**) which could be explained with the de-facto shorter linkers.

Table 10: Analysis of anisotropy data for GFP-coiled-coil fusion proteins

	r_{SS} [1]	τ_{FRET} (ns)	k_{FRET} (ns^{-1}) [3]	R (nm) [4]	r_0 [5]	r_∞ [6]
1xGFP	0.297	2.47	(0.033)*	(6.99)*	(0.408)*	(0.162)*
sfGFP-GCN4-p1	0.271	2.50	0.063	6.26	0.384	0.220
sfGFP-GCN4-pII	0.246	2.43	0.099	5.83	0.359	0.179
sfGFP-GCN4-ph3a	0.172	2.53	0.230	5.03	0.389	0.137
EGFP	0.291	2.84	(0.063)*	(6.13)*	(0.435)*	(0.250)*
EGFP-CC-Di	0.243	2.92	0.105	5.59	0.393	0.229
EGFP-CC-Tet	0.216	2.75	0.139	5.40	0.376	0.205
EGFP-CC-Pent	0.216	2.89	0.137	5.37	0.365	0.183

[1] Steady-state anisotropy. [2] Fluorescence lifetime. [3] Homotransfer rate. [4] Distance between all interacting fluorophores. [5] Initial / limiting / fundamental anisotropy. [6] Anisotropy for $t \rightarrow$ infinity.
( )* Data fitted with eq. (7), although no homo-FRET is expected; a fit with eq. (6) is preferred in order to receive the rotational correlation time ϕ.

5.2.4 Determining the stoichiometry of non-covalently associated GFP fusion proteins from fluorescence anisotropy

The combination of measured steady-state anisotropy and parameters from time-resolved anisotropy analysis enables the determination of protein stoichiometry. As previously shown for sfGFP concatemers, we could determine the number of interacting fluorophores N using several variations of Runnels and Scarlata's model (**Table 7**).

Similar to the sfGFP concatemers, the oligomeric state N was calculated with measured r_{SS} values in combination with individual or standard parameters (**Table 11**). In a first step, N was calculated with eq. *(8)* and individual values for the homotransfer rate k_{FRET} and τ. The rear part was neglected (r_{ET} = 0). The results for all of the proteins varied systematically from the stoichiometry determined in SEC (N_{SEC}). All of the calculated values of N were lower. As already shown in the past [51] and with sfGFP concatemers in this study, GFP are not oriented randomly and the simplified formula (r_{ET} = 0) cannot be applied for clusters of GFP.

In a second step, we calculated N from eq. *(8)* including the rear part. This required r_{ET} which was determined for sfGFP concatemers (r_{ET} = 0.105). Again, individual values for k_{FRET} and τ were used. As a consequence, all of the N were higher than for r_{ET} = 0. The values of N for sfGFP-based chimeras came close to N observed with biochemical methods while N for EGFP-based fusion proteins were higher than expected. The value for r_{ET} = 0.105 was calculated for sfGFP concatemers where more than 50 linker amino acids separate the GFP domains. EGFP-based fusion proteins exhibit shorter linkers and less flexibility. Apparently, r_{ET} has to be adjusted for EGFP-based fusion proteins. We determined the parameter r_{ET} for EGFP-based fusion proteins empirically with eq. *(8)* and r_{SS} values, assuming that N = 2 for EGFP-CC-Di and N = 3 for EGFP-CC-Tet. We could determine r_{ET}(EGFP-CC-Di) = 0.073 and r_{ET}(EGFP-CC-Tri) = 0.085, and an average r_{ET} = 0.079.

By adjusting r_{ET} to 0.079, the determination of N_{SS} could be improved for EGFP-based samples, with N_{SS}(EGFP) = 1.0, N_{SS}(EGFP-CC-Di) = 2.2, N_{SS}(EGFP-CC-Tet) = 2.9, and N_{SS}(EGFP-CC-Pent) = 2.9. Calculating N_{SS} for the sfGFP-based proteins with the EGFP-specified parameter r_{ET} = 0.079, in turn, led to underestimations of N. This demonstrates that the choice of r_{ET} can be significant for a precise determination of N_{SS}. Comparing EGFP-

based and sfGFP-based samples, r_{ET} differences can be traced back to varying linker lengths and altered construct structures (**Figure 20**) which both might decisively change the arrangement of GFP domains. From these results, it became evident that a varied structural composition requires a different parameter r_{ET}.

In a final step, N_{SS} was determined via eq. *(8)* in combination with the standard parameter set that was introduced for the sfGFP concatemers (r_1 = 0.297, k_{FRET} = 0.067 ns^{-1}, τ = 2.62 ns^{-1}, r_{ET} = 0.105). In case of sfGFP-based proteins, the calculated N_{SS} (last column, **Table 11**) hardly differed from values calculated with individual parameters.

Table 11: Determining the oligomeric state *N* of assembled GFP-coiled-coil fusion proteins from fluorescence anisotropy, using the model by Runnels & Scarlata [46]

	N_{SEC} [1]	N_{SS} (r_{ET} = 0) [2]	N_{SS} (r_{ET} = 0.105) [3]	N_{SS} (r_{ET} = 0.079) [4]	$N_{SS,standard}$ [5]
1xGFP	1	1.0	1.0	(1.0)*	1.0
sfGFP-GCN4-p1	2	1.6	2.0	(1.9)*	2.0
sfGFP-GCN4-pII	3	1.9	2.6	(2.4)*	3.4
sfGFP-GCN4-ph3a	(7+)•	2.8	5.7	(4.4)*	13.5
EGFP	1	1.0	(1.0)*	1.0	1.0
EGFP-CC-Di	2	1.8	(2.5)*	2.2	2.1
EGFP-CC-Tet	3	2.2	(3.4)*	2.9	3.1
EGFP-CC-Pent	5	2.2	(3.4)*	2.9	3.1

[1] Oligomeric state estimated from analysis with size exclusion chromatography.

[2] Number of subunits *N* in a complex of fluorophores that interact via homo-FRET. *N* was calculated with Runnels and Scarlata's model, eq. (8), and excluding the rear part by r_{ET} = 0 [46].

[3] *N* calculated with eq. (8) assuming that the rear part must not be neglected [51]. The value for r_{ET} = 0.105 was determined from concatemer standards. For the calculations, individual k_{FRET} values from time-resolved anisotropy data were used.

[4] *N* calculated with eq. (8) and r_{ET} = 0.079. For the calculations, individual k_{FRET} values from time-resolved anisotropy data were used.

[5] *N* calculated with eq. (8) and a standard parameter set. For sfGFP-based fusion proteins (row 1-4): r_1 = 0.297, k_{FRET} = 0.067 ns^{-1}, τ = 2.62 ns, r_{ET} = 0.105. For EGFP-based fusion proteins (row 5-8): r_1 = 0.291, k_{FRET} = 0.127 ns^{-1}, τ = 2.85 ns, r_{ET} = 0.079.

( )• Parallel total internal reflection fluorescence microscope experiments indicated a higher-order complex.

( )* Calculations with an inappropriate r_{ET}.

As both sfGFP-based coiled-coil fusion proteins and sfGFP concatemer standards exhibit a similar structural composition, it was possible to confirm the biochemically observed stoichiometries with the standard parameter set from sfGFP concatemers. In stark contrast, EGFP-based proteins could not be determined correctly from parameters gauged with sfGFP-based fusion proteins. Here, the estimations were 1.2 for the monomer, 3.6 for the dimer, 5.9 for the trimer, and 5.9 for the pentamer (not listed in **Table 11**). As a consequence, a new standard parameter set, specifically for EGFP-based proteins, was generated. This was done in an equal way as for the sfGFP concatemers, with mean values for k_{FRET} and τ. The calculations with the new and adjusted parameter set (r_1 = 0.291, k_{FRET} = 0.127 ns^{-1}, τ = 2.85 ns^{-1}, r_{ET} = 0.079) were listed in **Table 11**. For the dimer and trimer, the determined values came very close to observations in biochemical analysis. The determination for $N \geq 3$, i.e. for EGFP-CC-pent, was not possible as already shown for sfGFP concatemers.

In conclusion, the determination of the subunit stoichiometry for assembled GFP-coiled-coil fusion proteins is possible via steady-state anisotropy and the appropriate model, eq. *(8)*. Nevertheless, it requires time-resolving instrumentation to obtain parameters k_{FRET}, τ, and r_{ET}.

5.2.5 Fractional photobleaching of non-covalently associated GFP fusion proteins

With sfGFP concatemers, we could determine the number of interacting GFP moieties N without requiring time-resolved anisotropy data, simply by fractionally photobleaching the molecules and recording the steady-state anisotropy.

In the same way, we fractionally photobleached GFP-coiled-coil fusion proteins and analyzed their anisotropy behavior upon photobleaching (**Figure 24**). For the dimeric sfGFP-GCN4-p1 and the trimeric sfGFP-GCN4-pII (**Figure 24A**), the recorded anisotropy behaviors were similar to their concatemers 2xGFP and 3xGFP, respectively (**Figure 24B**). In striking contrast, the photobleaching response of sfGFP-ph3a started at a very low steady-state anisotropy and grew much steeper than observed for fusion proteins with GCN4-1 variants or any sfGFP concatemer (**Figure 24B**).

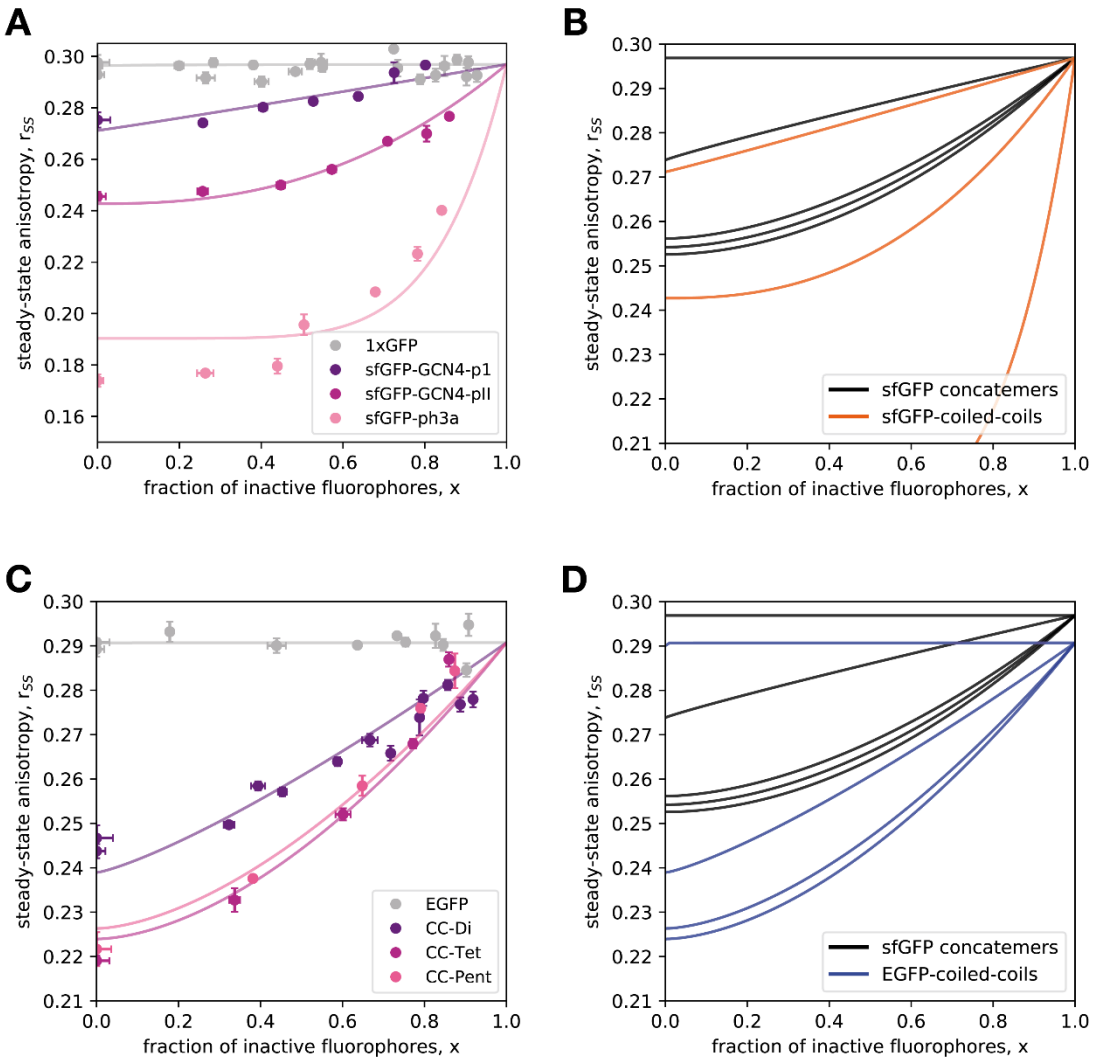


Figure 24: The steady-state anisotropy of fractionally photobleached GFP-coiled-coil fusion proteins (left panels), compared to the anisotropy behavior of concatemer reference proteins (right panels). Purified fusion proteins were diluted to 2 µM in coiled-coil sample buffer before they were irradiated with 445 nm laser light (300 mW) for fixed periods of time. After every bleaching step, the steady-state anisotropy and the residual fluorescence intensity was recorded (λ_{ex} = 488 nm; λ_{em} = 520 nm). The fraction of inactive fluorophores was calculated, eq. *(10)*. **A)** Steady-state anisotropy of the sfGFP-based coiled-coil fusion proteins for a growing fraction of inactive fluorophores x. The anisotropy data was fitted with eq. *(14)* and a = 0.10. The raw steady-state anisotropy data was standardized due to massive day-to-day variance and corrected for the monomeric anisotropy increase, eq. *(23)* (o = 0.011, p = 0.974, q = -0.012). **B)** Comparison between the anisotropy behavior of sfGFP-based concatemers and coiled-coil fusion proteins upon photobleaching. The steady-state anisotropy behavior is similar for dimers and trimers. The curves were generated with eq. *(14)* and a = 0.10. **C)** Steady-state anisotropy of the EGFP-based coiled-coil fusion proteins for an increased fraction of inactive fluorophores x. The anisotropy data was fitted with eq. *(14)* and a = 0.21. The raw steady-state anisotropy data was standardized and corrected for the for monomeric anisotropy increase, eq. *(23)* (o = 0.001, p = 4.056, q = -0.003). **D)** Comparison between the anisotropy behavior of sfGFP-based concatemers and EGFP-based coiled-coil fusion proteins upon photobleaching. The steady-state anisotropy behavior is not congruent. The curves were generated with eq. *(14)* and a = 0.10 for sfGFP-based concatemers, see Figure 16, and a = 0.21 for EGFP-coiled-coils.

When EGFP-based fusion proteins were photobleached (**Figure 24C**), the detected steady-state anisotropy patterns were not congruent to those from sfGFP-based fusion proteins including the sfGFP concatemers (**Figure 24D**). The photobleaching responses appeared to be "shifted" towards smaller r_{SS}. As previously described, the steady-state anisotropy of unbleached ($x = 0$) EGFP-based fusion proteins is lower than that of their sfGFP-based analogues. This underlines that the direct comparison between sfGFP- and the EGFP-based system is difficult.

We could determine the number of assembled subunits *N* from the photobleaching responses recorded with non-covalently associated GFP fusion proteins (**Table 12**). Therefore, the bimodal model by Yeow and Clayton [57], eq. *(12)*, and the model introduced and specified for sfGFP concatemers, eq. *(14)*, were compared.

Table 12: Determining the oligomeric state *N* of assembled GFP-coiled-coil fusion proteins from fluorescence anisotropy in combination with fractional photobleaching.

		Yeow and Clayton, eq. ***(12)***		**this study, eq.** ***(14)***	
	N_{SEC} [1]	f_{non} [2]	$N_{x,fnon}$ [3]	$N_{x,a}$ **(a = 0.10)** [4]	$N_{x,a}$ **(a = 0.21)** [5]
1xGFP	1	1.000	8.2	1.0	(1.0)*
sfGFP-GCN4-p1	2	0.921	2.3	2.0	(1.7)*
sfGFP-GCN4-pII	3	0.829	4.1	3.4	(2.6)*
sfGFP-ph3a	(7+)•	0.594	4.5	7.1	(4.9)*
EGFP	1	1.000	3.7	(1.0)*	1.0
EGFP-CC-Di	2	0.846	2.9	(2.9)*	2.3
EGFP-CC-Tet	3	0.750	2.4	(3.7)*	2.7
EGFP-CC-Pent	5	0.759	2.3	(3.5)*	2.7

[1] Oligomeric state estimated from analysis with size exclusion chromatography.
[2] Fraction of non-interacting, monomeric fluorophores according to Yeow and Clayton [57].
[3] Number of subunits that interact via homo-FRET, based on photobleaching data and eq. *(12)*.
[4] Number of subunits that interact via homo-FRET, based on photobleaching data and eq. *(14)* and a = 0.10.
[5] Number of subunits that interact via homo-FRET, based on photobleaching data and eq. *(14)* and a = 0.21.
( )• Parallel total internal reflection fluorescence microscope experiments indicated a higher-order complex.
( )* Calculations with an inappropriate a.

For Yeow and Clayton's model, the parameter f_{non} decreases for an increasing number of fluorophores interacting via homo-FRET, as already seen for concatenated GFP domains. However, the resulting determination of $N_{x,fnon}$ does not correspond to the results hemical analysis.

Yet, with the help of the newly introduced model, eq. *(14)*, and a = 0.10, determined from sfGFP concatemer standards, we could improve the calculations for N_x. Now, GCN4-p1 could indeed be classified as a dimer and GCN4-pII as a trimer. Furthermore, we predicted ph3a to be a heptamer. The oligomeric state of EGFP-based proteins, though, is systematically overestimated. Again, this could be explained by differences in the structural composition between sfGFP- and EGFP-based fusion proteins. Similar to calculations in the previous subsection – where r_{ET} was adjusted – the parameter a had to be adapted (a = 0.21) in order to cope with the altered design of EGFP-based fusion proteins (**Figure 25**). Thus, relying on steady-state anisotropy upon photobleaching, we could characterize the stoichiometry of oligomerizing GFP-coiled-coil fusion proteins. This was feasible even without fluorescence data recorded on a nanosecond scale. Besides the characterization of monomer, dimer, and trimer stoichiometry, it was possible to differentiate between higher-order and lower-order oligomers.

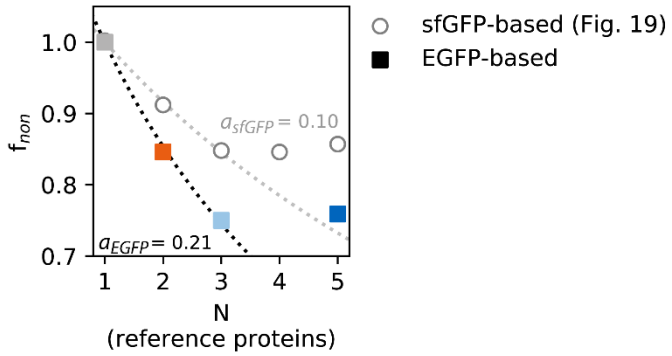


Figure 25: Adjusting parameter a, eq. *(14)*, for EGFP-based fusion proteins. To obtain parameter a_{EGFP}, values for f_{non} were calculated via eq. *(12)* from steady-state anisotropy data upon fractional photobleaching. We assumed that N = 1 for EGFP, N = 2 for EGFP-CC-Di, N = 3 for EGFP-CC-Tet, and N = 5 for EGFP-CC-Pent (treating the EGFP-based samples as reference proteins). A non-linear least squares approach for $f_{non} = \frac{1+a}{1+N \cdot a}$ was used to determine parameter a (valid for N = 1-3). For comparison, the data points and the fit from Figure 19 were indicated.

5.3 GFP photobleaching and photodegradation

5.3.1 Photobleaching of green fluorescent protein

When determining the stoichiometry of GFP fusion proteins via fractional photobleaching, the protein samples were irradiated for a significant amount of time (> 30 min) with strong blue laser light. To understand the process of GFP photobleaching, the effect of high-intensity laser light on GFP fusion proteins was classified hereafter.

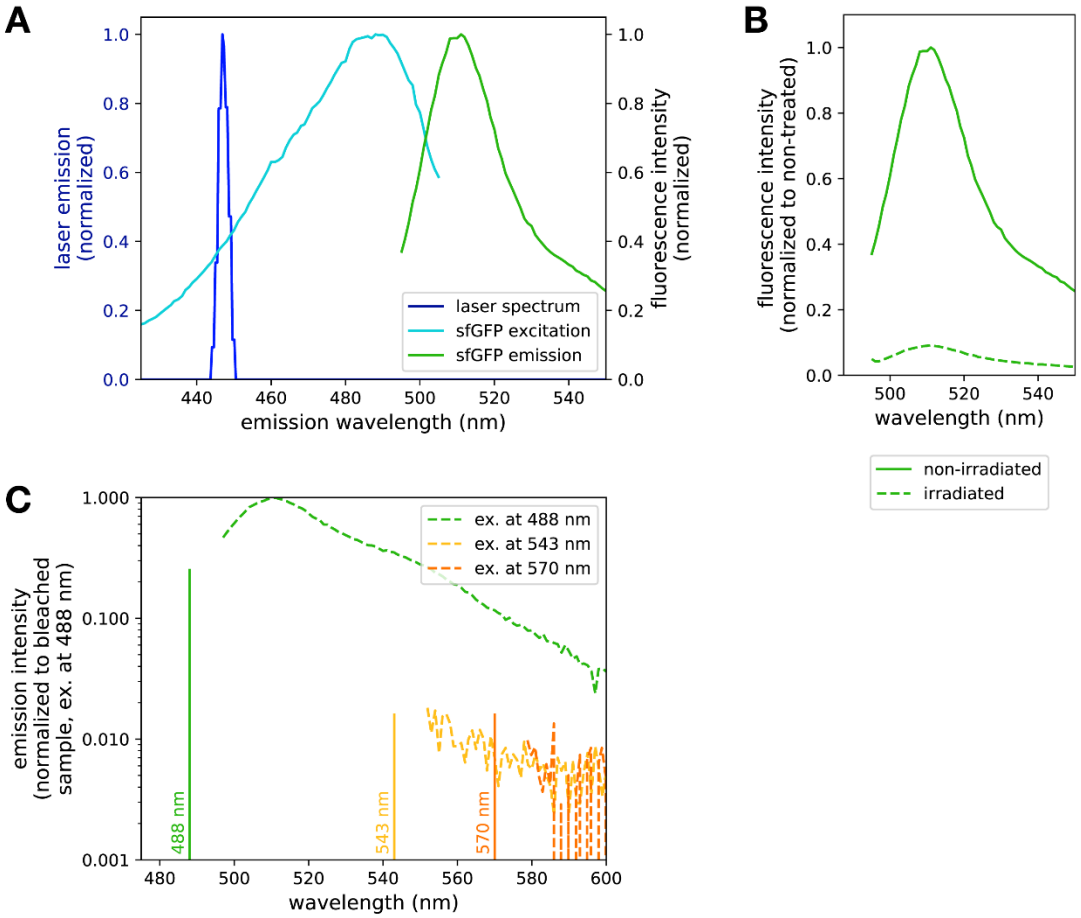


Figure 26: Spectroscopic analysis of sfGFP after photobleaching reveals loss in emission intensity at 511 nm, but no increase in emission at longer wavelengths as seen for Kaede [105] or EosFP [108]. A) Excitation and emission spectrum of sfGFP as part of the 1xGFP fusion protein. The spectrum of the 445 nm laser (± 5 nm peak) used for photobleaching is also indicated. Excitation and emission spectra were recorded with a fluorescence spectrophotometer (Cary Eclipse, wavelength scan from 400 to 600 nm). **B)** When excited with the 445 nm laser diode, photobleaching greatly reduces the height of the 511 nm emission peak. The protein sample was irradiated until fluorescence intensity was below 10% of the original. **C)** Photobleaching did not lead to red shift of emission, as shown for photoswitchable proteins Kaede (λ_{ex1} = 543 nm; λ_{ex2} = 572 nm) and EosFP (λ_{ex} = 571 nm) where irradiation leads to backbone breakage.

The superfolder variant of GFP, that was predominantly used in this thesis, has an excitation maximum at 485 nm and an emission maximum at 510 nm (**Figure 26A**). Unlike wild-type GFP with a Stokes shift of 114 nm, sfGFP has a relatively small Stokes shift of 25 nm. The small Stokes shift of sfGFP facilitates homo-FRET.

In the photobleaching experiments, sfGFP fusion proteins were irradiated with a 445 nm laser diode (± 5 nm). Here, the excitation is half-maximal (**Figure 26A**). Constant irradiation at 445 nm led to a decline of > 90% in emission intensity (**Figure 26B**). In the past, a similar behavior has been published for irreversibly photobleached EGFP [148]. With the irradiation, neither the wavelength of the maximal absorbance, nor the wavelength of the maximal emission changed which would be characteristic if sfGFP was photoswitchable such as proteins Kaede or EosFP (**Figure 26C**). Irreversible photobleaching of sfGFP depended on the irradiation power and time. The sfGFP containing construct 1xGFP was continuously irradiated with 18 mW (200 mA at the laser diode), 133 mW (300 mA), and 306 mW (450 mA) (**Figure 27**).

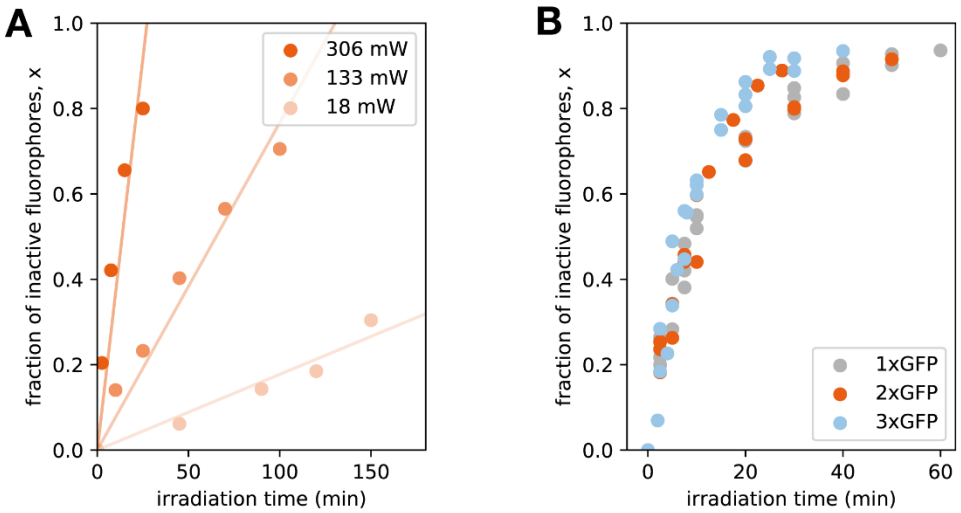


Figure 27: The effect of irradiation power and bleaching time on the fraction of inactive fluorophores. A) Decreasing the laser power increases the irradiation time necessary for photobleaching. **B)** Comparing GFP concatemers of growing repeat lengths, the increase in fraction of inactive fluorophores was similar for 1xGFP, 2xGFP, and 3xGFP constructs. Purified GFP constructs were diluted in fluorescent free PBS, 1 mM EDTA, 0.5 mM β-mercaptoethanol, and 0.05% (v/v) sodium azide. The fusion proteins were irradiated with a 445 nm laser diode (306 mW). After the photobleaching for a period of time, the residual fluorescence intensity was measured (λ_{ex} = 488 nm, λ_{em} = 520 nm). The fraction of inactive fluorophores was calculated by dividing the residual fluorescence intensity by the initial intensity before photobleaching.

By dividing the remaining fluorescence intensity of photobleached samples by the initial intensity of non-irradiated controls, the fraction of inactive fluorophores x could be determined which served as an intuitive scale to characterize the extent of photobleaching. GFP photobleaching is increased with the power of the laser light and the irradiation time (**Figure 27A**). The photobleaching of sfGFP proved to be similar for 1xGFP, 2xGFP, and 3xGFP (**Figure 27B**).

5.3.2 GFP degradation caused by irradiation with blue laser light

FCS measurements of GFP fusion proteins, executed by Ganesh Agam (**Appendix Fig. 6**), revealed that the structure of the fusion proteins is altered when irradiated with highly-energetic laser light.

To investigate the quantity of damaged molecules resulting from irradiation, the band intensity of irradiated 1xGFP fusion proteins on denaturing PAGE was examined (**Figure 28A**). The band corresponding for the intact 36.4 kDa 1xGFP fusion protein (running at 40 kDa) weakened for an increasing fraction of inactive fluorophores. As shown for protein concentrations from ~0.1 µM to 2 µM in PBS (**Figure 9**), the band strength of Coomassie stained proteins is linear proportional to the protein concentration.

Specifically, the concentration of intact 1xGFP decreased linearly with continuous photobleaching (**Figure 28B**). Starting with an initial concentration of 2 µM of intact GFP, the concentration of the 36.4 kDa main fraction was strongly reduced to 0.75 µM after 80% of GFP had been bleached. Assuming a steadily linear decrease, a final concentration of 0.6 µM of non-fluorescent 36.4 kDa proteins can be extrapolated. Yet, it is questionable whether these proteins are structurally intact.

Experiments with a sample consisting of 1xGFP (2 µM) and one additive protein (2 µM), i.e. bovine serum albumin, carbonic anhydrase, or equine cytochrome C, demonstrated that blue laser light degrades exclusively 1xGFP but no other proteins in a mixed sample (**Figure 28C**).

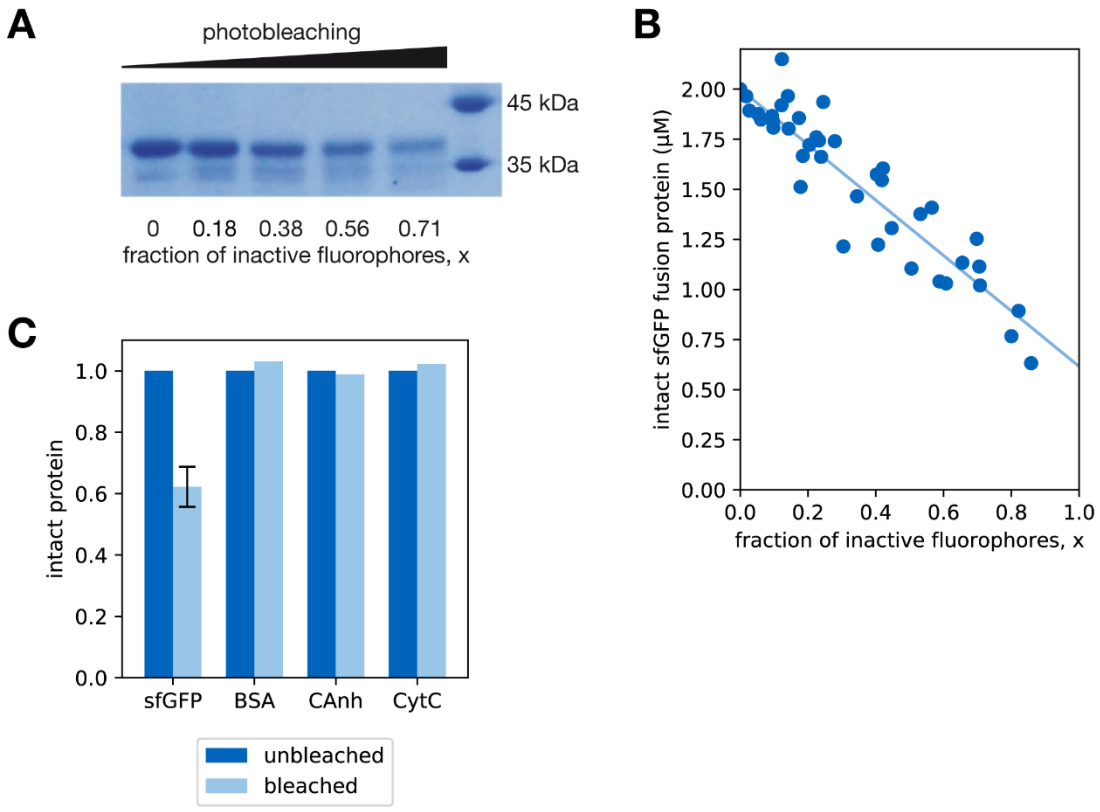


Figure 28: Blue laser light degrades the sfGFP containing fusion proteins. Purified sfGFP fusion proteins were diluted in fluorescence free PBS, 1 mM EDTA, 0.5 mM β-mercaptoethanol, and 0.05% (v/v) sodium azide. They were irradiated with a 445 nm laser (450 mA) for fixed periods of time. After bleaching, the residual fluorescence intensity was measured (λ_{ex}: 488 nm; λ_{em}: 520 nm) and the fraction of inactive fluorophores *x* was calculated, eq. *(10)*. **A)** Photobleaching decreases the abundance of the intact sfGFP fusion protein. **B)** The densitometric evaluation of sfGFP bands suggests that the decrease is linear in relation to the decrease in fluorescence. Samples were separated by denaturing PAGE and stained with Coomassie Brilliant Blue G250, band strength was analyzed densitometrically ($y = -0.70x + 1.01$; $r^2 = 0.89$). **C)** Blue laser light degrades sfGFP but no other proteins in a mixed sample. Mixtures of sfGFP (2 µM) and bovine serum albumin (BSA, 2 µM), sfGFP and bovine carbonic anhydrase (CAnh, 2 µM), and sfGFP and equine cytochrome C (CytC, 2 µM) were irradiated to 60% intensity and separated by gradient denaturing PAGE. The band strengths were determined via densitometry and normalized to unbleached samples. The bars for sfGFP show the mean value for sfGFP in all three mixtures, the error bar represents the respective standard deviation (n = 3).

5.3.3 Characterization of GFP degradation by a great abundance of different and specific fragments

The irradiated / photobleached 1xGFP was analyzed via gradient Schägger gels (**Figure 29A**). Besides the intact construct at 40 kDa, there were 7 lighter molecules detectable. The molecules are labeled with the prefix “sf-” for superfolder GFP and a letter. Here, the molecules were further specified with their molecular weights calculated by interpolation (using protein standards).

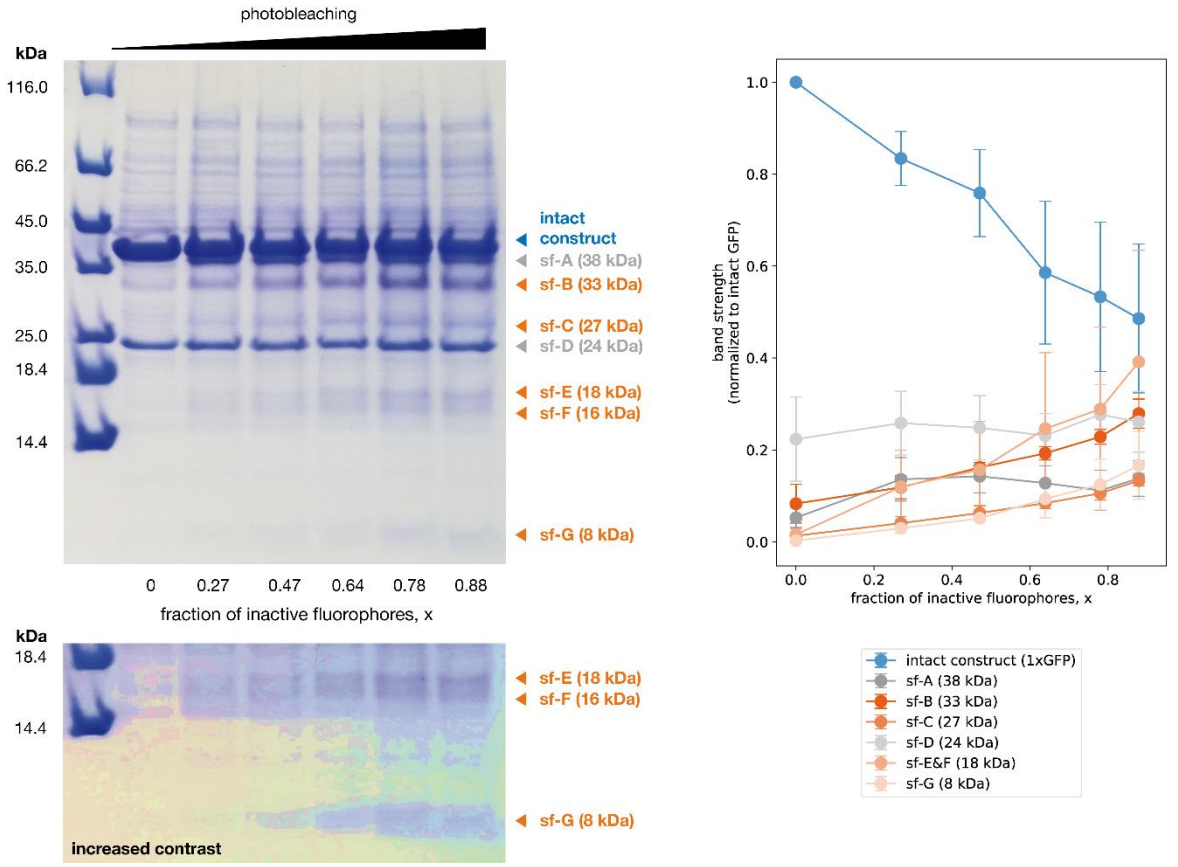


Figure 29: The irradiation of sfGFP increases abundance of a number of protein fragments. Purified 1xGFP, a sfGFP containing fusion protein, diluted to 14 µM in fluorescence free PBS, 1 mM EDTA, 0.5 mM β-mercaptoethanol, and 0.05% (v/v) sodium azide was irradiated with a 445 nm laser (450 mA) for fixed periods of time. After bleaching, the residual fluorescence intensity was measured (λ_{ex}: 488 nm; λ_{em}: 520 nm) and the fraction of inactive fluorophores was calculated, eq. *(10)*. Samples were separated by gradient denaturing PAGE [126] and stained with Coomassie Brilliant Blue G250 (left panels). For protein detection below 18 kDa, the contrast was increased for a better identification of the protein bands. The band strength was analyzed densitometrically (right panel). Corresponding molecular weights were estimated from a protein standard. The irradiation of 1xGFP gradually changes the sample composition by fragmenting the fusion protein (intact construct) into smaller molecules sf-A (38 kDa), sf-B (33 kDa), sf-C (27 kDa), sf-E (18 kDa), sf-F (16 kDa), and sf-G (8 kDa). The densitometric analysis shows the increase of smaller molecules (sf-B, sf-C, sf-E, sf-F, sf-G; orange) and the decrease in band of the intact GFP fusion proteins (1xGFP, blue), gradually when photobleached (n = 3). For practical reasons, the densitometric analysis for sf-E (18 kDa) and sf-F (16 kDa) was combined.

From the seven detected molecules, sf-A (38 kDa) and sf-D (24 kDa) bands did not steadily get stronger during irradiation. Since molecule sf-A was only formed after the first bleaching step and stagnated in band strength afterwards, it can be assumed that it might be a version of sfGFP that was conformationally changed due to irradiation. Molecule sf-D was already present in the non-irradiated sample, was not affected by irradiation, and is presumably not a fragment of 1xGFP. Bands sf-B (33 kDa), sf-C (27 kDa), sf-E (18 kDa), sf-F (16 kDa;

hardly being separated from sf-E), and sf-G (8 kDa) visibly gained in strength over increasing photobleaching. A densitometric analysis of these five bands confirmed this observation. Due to significantly higher loading concentrations for gradient Schägger gels than for gels in **Figure 28**, the detection of very weak bands was possible. However, the densitometric analysis of the intact GFP became more difficult due to overstaining. It has to be noted that the band of the intact construct was estimated to ~40 kDa, around 10% higher than the actual mass of the protein. We suggest that this might be due to an inaccuracy for masses higher than 30 kDa in Schägger tricine-denaturing PAGE gels [149] and protein marker differences in glycine- and tricine-based gels. Consequently, also the masses of the sf-A and sf-B can be estimated 10% lower yielding 34 and 30 kDa, respectively.

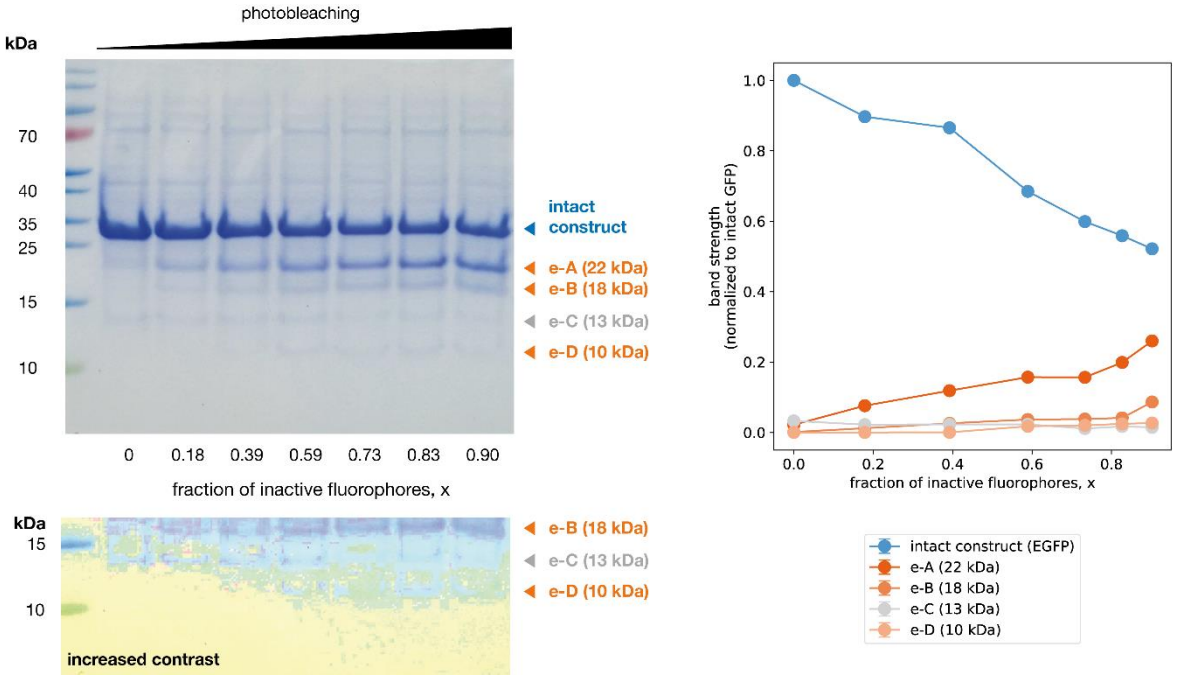


Figure 30: The irradiation of EGFP results in a defined number of fragments. The EGFP sample protein was diluted to 14 µM in the corresponding sample buffer. The sample was irradiated with a 445 nm laser (450 mA) for fixed periods of time. After bleaching, the residual fluorescence intensity was measured (λ_{ex}: 488 nm; λ_{em}: 520 nm) and the fraction of inactive fluorophores was calculated, eq. *(10)*. Samples were separated by gradient denaturing PAGE [126] and stained with Coomassie Brilliant Blue G250 (left panels). For a better identification of the protein bands, the contrast was increased for molecular weigths < 18 kDa. The band strength was analyzed densitometrically (right panel). Corresponding molecular weights were estimated from a protein standard. When irradiating EGFP with blue laser light, the band strength of smaller molecules e-A (22 kDa), e-B (18 kDa), e-D (10 kDa) is increased. The densitometry shows the increase of smaller fragments (e-A, e-B, e-D, orange) and the decrease in band of the intact fusion protein (EGFP, blue) upon fractional photobleaching.

In a second approach, a fusion protein based on EGFP (30.0 kDa) [18] was irradiated under equal conditions and analyzed via gradient Schägger gel. EGFP was similarly affected by blue laser light, being degraded gradually for continuous photobleaching (**Figure 30**). As seen for the sfGFP based fusion proteins, the amount of smaller molecules was increased for the photobleached samples. The faster-migrating molecules were labeled with the prefix "e-" for enhanced GFP and a letter, similar as for 1xGFP. Except from the main band, corresponding for the structurally intact 30 kDa fusion protein, four smaller bands were detected. Among these, the band strength of e-C (13 kDa) was not altered while the abundance of e-A (22 kDa), e-B (18 kDa), and e-D (10 kDa) grew during the photobleaching process. Molecules whose band strength increased by continuous photobleaching were classified as fragments of the original GFP fusion proteins.

5.3.4 Immunoblot characterization of GFP fragments

To further characterize the fragments resulting from irradiation, unbleached and bleached GFP fusion proteins were analyzed via protein immunoblot analysis. The intact, sfGFP-based construct 1xGFP has an N-terminal Nano-tag epitope and a C-terminal FLAG-tag epitope (**Figure 31A**, **Appendix Fig. 1**). A non-fragmented protein should comprise both epitopes. Similar to Coomassie stained gels, Ponceau S stained membranes showed that the abundance of fragments sf-B, sf-C, sf-E, sf-F, and sf-G grew during photobleaching. All of the fragments could be detected by either an anti-Nano antibody or by anti-FLAG and anti-His. The band representing the intact protein was positive for both antibodies. The fragments sf-C (27 kDa), sf-E (18 kDa), and sf-G (8 kDa) could be detected with the anti-Nano antibody. These fragments still hold the original N-terminus. The bands representing sf-E and sf-G were very weak but detectable.

Using anti-FLAG and anti-His antibodies, the fragments sf-B (33 kDa) and sf-F (16 kDa) could be detected still holding the original C-terminus. Furthermore, it was possible to detect a C-terminal 25 kDa band with anti-His and anti-FLAG antibodies. This band can be neither assigned to the heavier sf-C nor the lighter sf-D band, that were detected previously with Coomassie or Ponceau S staining. From now on, this C-terminal 25 kDa fragment will be referred to as sf-Ct25. Interestingly, we detected stronger bands for larger molecules and higher background signal in irradiated samples, particularly for the C-terminal anti-FLAG-tag antibody. In contrast, no growing bands corresponding to large molecules were detected

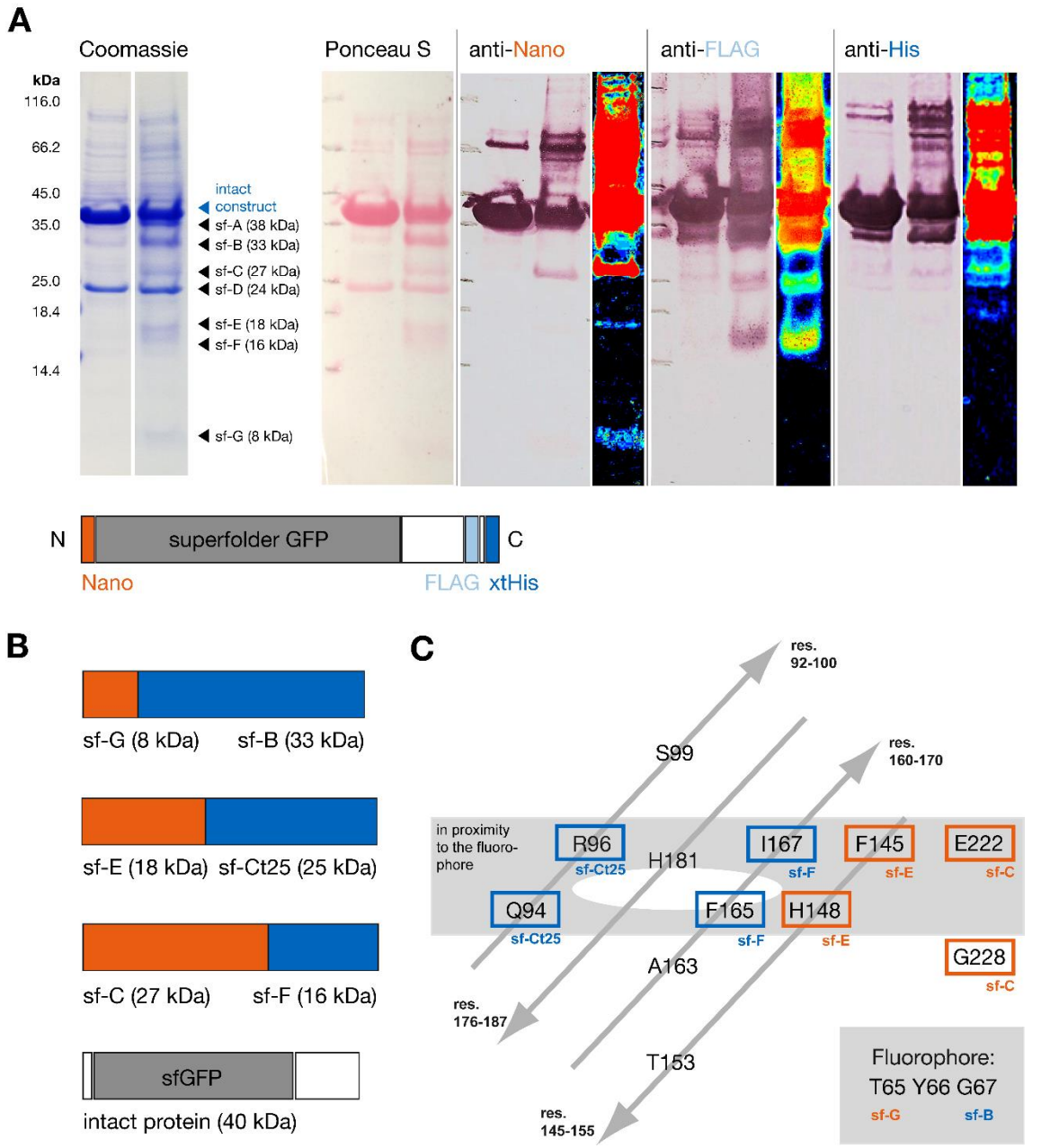


Figure 31: Immunoblot analysis of 1xGFP fragmentation to identify possible cleavage sites. A) Immunochemical characterization of non-irradiated (left lane) and irradiated (central and right lane) 1xGFP to identify N- and C-terminal epitopes**.** The fragments sf-C (25 kDa), sf-E (18 kDa) and sf-G (8 kDa) contain the N-terminal Nano-tag. The fragments sf-B (33 kDa) and sf-F (16 kDa) contain the C-terminal FLAG-tag and His-tag, respectively. The band sf-Ct25 (25 kDa) which was detectable in immunoblot analysis exhibits C-terminal epitopes. 1xGFP was irradiated with a 445 nm laser (450 mA) and separated by gradient denaturing PAGE [126]. For the analysis of weaker bands, the image contrast was increased (right lane). **B)** Complementary N- (orange) and C-terminal (blue) fragments. Combined masses are comparable to the intact 40 kDa 1xGFP (mass estimation via PAGE and standard proteins). **C)** Schematic overview over candidate residues for potential fragmentation sites. The residue numbers refer to the residue sequence position in the isolated sfGFP [81]. The fluorophore (Thr65/Tyr66/Gly67) is represented by a white ellipse in the background, beta-strands by arrows. Candidate residues are marked according to their corresponding fragments. Phe145 and Phe165 form major contacts to the fluorophore [150]. His148, Gln94, and Arg96 are related to the fluorophore maturation [78, 150]. A proton wire is formed by His181 and Arg96 [151]. Glu222 and residues 99, 153, 163 are related to the fluorescence intensity and the excitation maximum [69, 150, 152].

in Coomassie stained gels or in Ponceau S stained membranes. This indicates that the irradiation affected the epitope accessibility for subtly aggregating fractions.

The immunochemical analysis of sfGFP fragmentation demonstrated that there are three N- and three C-terminal fragments. These fragments can be pairwise combined to molecules of 41 kDa (sf-G and sf-B) or 43 kDa (sf-E and sf-Ct25; sf-C and sf-F) (**Figure 31B**). The combined masses congruently approach a weight of 40 kDa for the intact protein, estimated from gel electrophoresis and protein standards. The fragment combinations also indicate that two of the cleavages (sf-G and sf-B; sf-E and sf-Ct25) are taking place inside the GFP domain whereas the C-terminal cleavage (sf-C and sf-F) cannot be unequivocally characterized as such. The estimation of potential fragmentation sites from gel electrophoresis alone is very speculative as the fragment masses are inaccurate. Based on the approximate masses from PAGE analysis and the terminal epitopes, we could determine potential candidate residues (**Figure 31C**). Most of these candidate residues are located on one side of the protein's beta-barrel and proximal to the fluorophore. They occupy key roles in fluorophore maturation and stabilization [150]. In chapter 6.2, the hypothetical cleavage sites for sfGFP will be scrutinized in detail.

In a consecutive approach, the fragmentation of the EGFP fusion protein upon photobleaching was analyzed via immunoblot (**Figure 32**). Again, Coomassie stained gradient gels and membranes after Ponceau S staining demonstrated that the fragments e-A, e-B, and e-D grew stronger for heavily irradiated samples (**Figure 32A**, left and middle panel). Unfortunately, the EGFP fusion protein only holds one epitope for immunological identification, an N-terminal His_6 tag, detectable with an anti-His antibody (**Appendix Fig. 2**). The bands e-A (22 kDa) and e-B (18 kDa) were both negative for the anti-His antibody. Thus, one must conclude that they are lacking the N-terminal epitope. On the contrary, e-D (10 kDa) was detectable with the anti-His antibody as it is presumably an N-terminal fragment. Furthermore, a weak band could be detected at around 13-14 kDa, above the e-C (13 kDa) band. The N-terminal 14 kDa fragment will be referred to as e-Nt14 from now on. Similar to the sfGFP fusion protein, the N- and C-terminal fragments can be complemented (**Figure 32B)**. The combined masses of the fragment pairs resemble that of the intact protein.

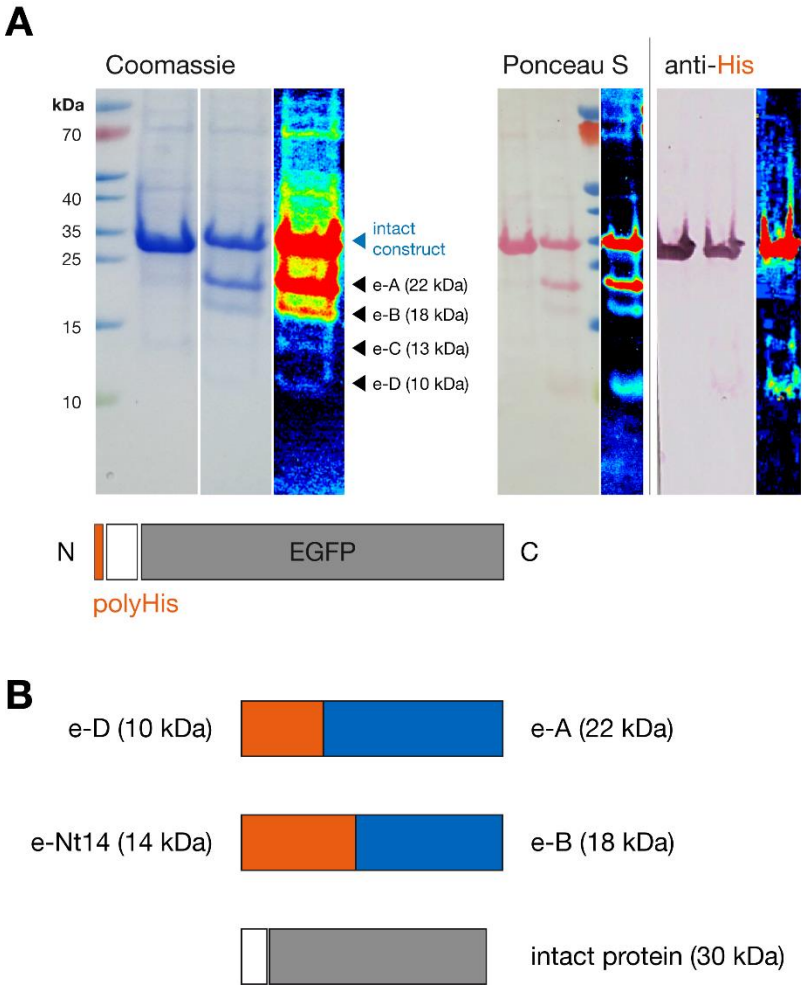


Figure 32: Immunoblot analysis of non-irradiated (left lane) and irradiated (right lane) EGFP fusion in order to evaluate the N-terminal integrity of protein fragments. A) In immunochemical analysis, e-D (10 kDa) and a fragment at 14 kDa (e-Nt14) were found to hold an N-terminal His_6 epitope. Purified EGFP (10 µM in fluorescence free PBS, 1 mM EDTA, 0.5 mM β-mercaptoethanol, and 0.05% (v/v) sodium azide) was irradiated with a 445 nm laser (450 mA) for fixed periods of time. Unbleached and bleached samples were separated by gradient denaturing PAGE [126], transferred to a nitrocellulose membrane via a semi-dry blotting step (with a subsequent Ponceau S staining step), marked with an anti-His antibody, and individually visualized with an antibody coupled alkaline phosphatase. With the help of denaturing PAGE and a protein ladder, the molecular weight of the bands was estimated. **B)** Complementary N- (orange) and C-terminal (blue) fragments. The masses of the combined fragments equal approximately 30 kDa, the molecular weight of the intact EGFP fusion protein (mass estimation via PAGE and standard proteins).

The findings in this section imply a similarity between photodegradation of sfGFP and EGFP. Both variants are fragmented into a small N-terminal and a larger C-terminal part upon irradiation with highly-energetic laser light. Besides this analogy, it is unclear whether the residual fragments can be attributed to identical cleavage sites. With the support of mass spectrometric analysis, a detailed picture will be drawn at the end of the Discussion section.

6 Discussion

6.1 Characterization of protein stoichiometry via fluorescence anisotropy

In this thesis, we examined the applicability of several approaches to study the stoichiometry of protein assemblies via fluorescence anisotropy, homo-FRET, and fractional photobleaching. Therefore, coiled-coil peptides, oligomerizing α-helical peptides (~30 amino acids in length) most prominently represented by the leucine zipper dimerization domain GCN4-p1 (**Figure 1**), were tested with different anisotropy-based approaches. As coiled-coil peptides have no intrinsic fluorescence, fluorescent reporter tags help to resolve their stoichiometry. To guarantee that the protein of interest is labeled effectively with a fluorescent probe, our strategy is based on expressing the oligomerizing coiled-coil peptide in a fusion protein, C-terminally to a green fluorescent protein (GFP) (**Figure 8**, **Figure 20**). Furthermore, Nano-, FLAG-, and polyHis-tags proved to be beneficial for protein purification (**Figure 13D**, **Figure 21A**) and identification of N- and C-terminal epitopes (**Figure 31A**).

Compared to isolated GFP, the photophysical properties of GFP used in chimeric proteins are not altered significantly.

As a fluorescent reporter of choice, we integrated the superfolder variant of GFP (sfGFP) into fusion proteins. In the past, sfGFP was used as a fluorescent protein tag in many biochemical and cell biological studies [81, 136, 153-156]. The monomeric construct 1xGFP, the basis for all constructs generated in this work, comprises 333 aa in full length (**Figure 8**), and is significantly larger than the original 238 aa long sfGFP introduced by Pédelacq et al. [81]. Based on these structural differences, the question arises whether the altered structural composition and the increased molecular size of the 1xGFP construct has a significant influence on fluorescence properties of GFP, especially fluorescence anisotropy.

In experiments with 1xGFP, the excitation maximum (λ_{ex} = 488 nm) and the emission maximum (λ_{em} = 511 nm) were both slightly red-shifted with a Stokes shift of 23 nm. This is similar to the literature values (λ_{ex} = 485 nm, λ_{em} = 510 nm, Stokes shift = 25 nm)

(**Figure 26A**). The fluorescent lifetime is τ = 2.47 ns for 1xGFP specifically. For all sfGFP-based samples in this thesis, the lifetime was τ = 2.56 ± 0.13 ns (standard deviation) on average. This resembles published lifetimes of other GFP variants [90, 157], reported to be in the range τ = 2.4-2.9 ns. The measured quantum yield for 1xGFP was QY = 0.70 and resembles the literature value of QY = 0.65.

To our knowledge, the steady-state anisotropy of sfGFP has never been determined in a scientific publication so far. Experimentally determined anisotropy values for monomeric EGFP, though, usually lie around $r_{SS} \approx 0.29$ [90, 115, 158]. We recorded a similar value for monomeric sfGFP (r_{SS} = 0.297). In a crucial test, this experimental value can be scrutinized by comparing it to a theoretical anisotropy value, calculated with the help of the Perrin equation, eq. *(1)*. Therefore, we solved the Perrin equation with the empirically determined initial anisotropy r_0 = 0.408 (**Table 6**), the fluorescence lifetime τ = 2.47 ns (**Table 6**), and the rotational correlation time ϕ = 20.34 ns (**Table 5**). The calculations yield an anisotropy value of r = 0.364 that differs significantly from the experimental reality. How can the difference between the theoretical value and the recorded r_{SS} be explained? In an enlightening work by Volkmer et al., the anisotropy decays of several GFP variants were characterized, including S65T-GFP with the same fluorophore forming residues as sfGFP [159]. According to this study, the initial anisotropy is r_0 = 0.34 for S65T-GFP upon single-photon excitation and thus significantly lower than the hypothetical r_0 = 0.4. Volkmer et al. further found ϕ = 16 ns and τ = 3.01 ns for S65T-GFP. Inserting these parameters into the Perrin equation yields a theoretical anisotropy of r = 0.286, a value that reflects the experimental reality in this study very well.

Overall, we found that the inclusion of the sfGFP domain in a protein chimera did not substantially change the photophysical properties of the fluorescent protein, as strong similarities between values from the literature and our own results indicate.

Covalently bound, concatenated GFP domains serve well as model proteins to study protein stoichiometry via fluorescence anisotropy.

In this thesis, we introduced concatemer model proteins (**Figure 13C**) to evaluate whether their numbers of GFP moieties could be studied via anisotropy. As seen in former studies, fluorescent protein concatemers served as a reference for complexes labeled with fluorescent proteins [88, 89]. The concatemer structure ensures a spatial proximity of fluorophores which is a basic prerequisite for homo-FRET [44]. Additionally, homo-FRET interaction can be studied as a function of fixed numbers of fluorophores.

It was demonstrated that the sfGFP concatemers in this study form macromolecular structures which are neither perfectly globular, nor "rod-like", but rather ellipsoid (**Figure 14**). For similar GFP concatemers in the literature, the question of whether the structure is more extended or globular is still open [90, 137, 138, 160]. Hence, the experimental findings in this thesis are within the range of previous data. We suggest that oligomerizing GFP-coiled-coil fusion proteins form a bouquet-like macrostructure, indicated in **Figure 3B**. In those macromolecular structures, the GFP tags are assumed to be assembled in an ellipsoid shape, too. Therefore, the anisotropy values determined for covalently linked, concatenated sfGFP domains can serve as reference points for non-covalently assembled oligomers labeled with sfGFP reporters.

We found that the steady-state anisotropy of GFP concatemers could well distinguish between monomeric, dimeric, or trimeric assemblies (**Figure 15A**). The difference between complexes higher than trimers were not statistically significant (unpaired Student's *t*-test). This is in line with a study on EGFP concatemers linked with short interspacing residues (Gly-Pro-Val-Ala-Thr) where the steady-state anisotropy could also differentiate between monomeric, dimeric and trimeric concatemers (**Figure 15B**) [90]. The anisotropy spread between monomer and trimer was decisively higher for the EGFP concatemers with short linkers. This leads to the conclusion that a smaller inter-fluorophore distance results in a higher homo-FRET efficiency and thus in a stronger decrease in anisotropy. Interestingly, the difference between EGFP trimer and EGFP tetramer was mistakenly rated as significant in the original publication [90] which was not supported by statistical analysis in our calculations (unpaired Student's *t*-test).

Observations with sfGFP concatemers showed that maximally three assembled GFP domains can be resolved via steady-state anisotropy. This resolution limit was also visible for time-resolved anisotropy, measured by Ganesh Agam in a cooperation with the Chair of Physical Chemistry I at LMU Munich. The time-dependent anisotropy decay was recorded with a multiparameter fluorescence detection (MFD) pulsed interleaved excitation (PIE) equipment [161] that is usually used to detect the fluorescence anisotropy of small organic fluorophores. Measurements with the MFD PIE setup demonstrated exactly the same difficulty to distinguish trimer from tetramer or pentamer behavior (**Appendix Fig. 4A**). However, changing the emission detection setup to an FLS1000 spectrometer equipped with photomultipliers enabled the discrimination between concatemers with more than $N = 3$ GFP moieties (**Appendix Fig. 4B**). From this data, the homotransfer rates k_{FRET} were determined precisely for all concatemers (**Table 6**). This observation suggests that a resolution limit can be prevented by fluorescence detection methods optimized for the specific fluorophore species. When detecting the steady-state anisotropy with a microplate reader, the observed resolution limit for the GFP concatemers is – in consequence – not caused by restricted homo-FRET. It is likely to be prevented by increasing the detection sensitivity for $N \geq 3$. Since the microplate reader used for steady-state anisotropy determination in this thesis is not a custom-built instrument, changing only a part of the setup was not possible. Overcoming these technical limitations in future applications, however, would broaden the field of application and increase the validity of steady-state analysis for trimers and higher oligomers.

The number of subunits interacting via homo-FRET can be determined via fluorescence anisotropy, exemplarily shown for sfGFP concatemers.

Making use of eq. *(8)* and a combination of measured steady-state anisotropy values and time-resolved anisotropy data, the number of interacting subunits N could be calculated for sfGFP concatemers (**Table 7**). We show that Runnels and Scarlata's model compulsorily requires knowledge of the restricting parameter r_{ET}. This was expected as r_{ET} was hypothesized as negligibly low for randomly oriented small fluorophores but substantial for more uniformly oriented fluorophores, such as fluorescent proteins [51]. Calculations with $r_{ET} = 0.105$ yielded the most accurate predictions for 1xGFP, 2xGFP, and 3xGFP (**Table 7**).

When comparing calculations of N that used individual values for homotransfer rate k_{FRET} and fluorescent lifetime τ with those that used a standard parameter set with average values (k_{FRET} = 0.067 ns^{-1}; τ = 2.62 ns), there was no substantial difference detectable (**Table 7**). This implies that similarly-structured proteins can be analyzed with averaged parameters obtained by time-resolved measurements. Time-resolved measurements, in turn, need sophisticated instrumentation and a significant amount of expertise [161]. In an attempt to circumvent time-resolved anisotropy measurements, we could demonstrate that it is possible to determine N from steady-state anisotropy of fractionally photobleached proteins alone. Similar to the steady-state anisotropy of unbleached samples, the anisotropy behavior during fractional photobleaching differs for concatemers with one, two, and three concatenated GFP domains (**Figure 16**). These responses agree with the theoretical predictions for fractionally inactivated, clusters with N subunits that interact via homo-FRET (**Figure 3B**) [57].

Unexpectedly, the steady-state anisotropy of monomeric GFP fusion proteins is growing upon fractional photobleaching (**Figure 17**). This effect can be observed for both, sfGFP- and EGFP fusion proteins, and contradicts the expectation that only complexes interacting via homo-FRET exhibit an increase in steady-state anisotropy when being fractionally photobleached. Since the effect is systematic, the entire data could be corrected for the increase of the monomeric fusion proteins, eq. *(23)*. The phenomenon is caused by an increase in parallel signal strength, on average 6% for photobleached (x > 0.9) monomeric GFP. It can be explained by a growing noise to signal ratio due to the reduction of fluorescent and thus detected samples.

Alternatively, the effect can be explained by assuming that photobleaching happens asymmetrically, preferably affecting fluorophores with lower individual anisotropies (**Figure 33**). Here, two basic premises must be made:

First, the fluorescence lifetime of GFP fluorophores in solution is normally distributed, where the majority of fluorophores have individual lifetimes around an average value μ (= 3 ns for GFP) (**Figure 33A**). Since the fluorescence anisotropy is a function of fluorescence lifetime (**Figure 33B**), eq. *(1)*, also the fluorescence anisotropy itself exhibits a Gaussian distribution.

And second, molecules with longer fluorescence lifetimes are more likely to face photobleaching than those with shorter lifetimes. If a fluorophore remains longer in the

excited state ($\tau > \mu$), an intersystem crossing from the excited singlet state (S_1) to the triplet state (T_1) is more probable. According to studies by Adam et al., the triplet state transition is directly related to the photobleaching of the fluorophore [101, 103].

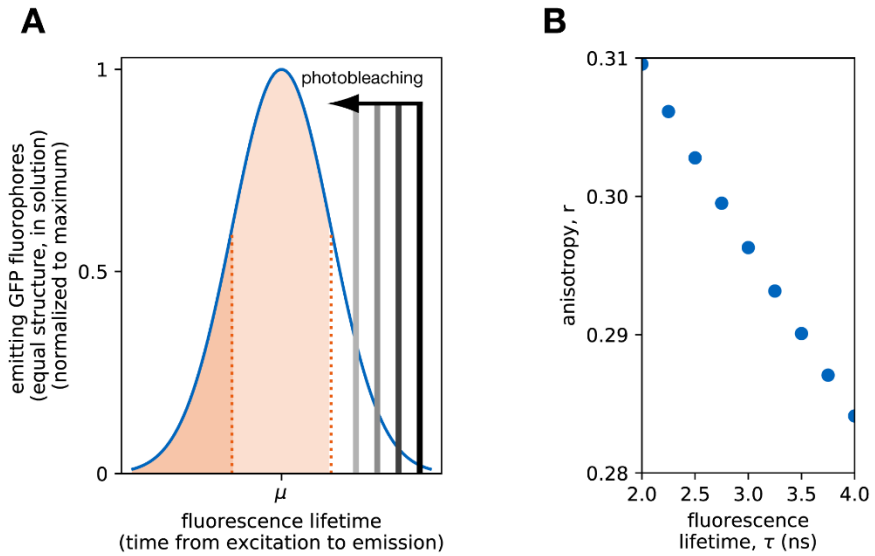


Figure 33: The asymmetric photobleaching preference for fluorophores with longer individual lifetimes leads to an increase in steady-state anisotropy of GFP upon photobleaching. A) Schematic representation of normally distributed fluorescence lifetimes for equal GFP fluorophores in solution (average lifetime μ). For GFP, the average lifetime is usually around 3 ns. Dark orange: fluorophores with individual lifetimes that are shorter than μ + standard deviation. Light orange: fluorophores within standard deviation range. No color: fluorophores with individual lifetimes that are longer than μ + standard deviation. According to the theoretical considerations in section 6.2, fluorophores with shorter lifetimes are less prone to the photobleaching process than those with longer lifetime. This asymmetry might be the reason for the steady-state anisotropy increase of monomeric GFP fusion proteins upon photobleaching. **B)** The fluorescence anisotropy is a function of fluorescence lifetime, see Perrin equation, eq. *(1)*. Here, the anisotropies were exemplarily calculated with and eq. *(1)*, variable lifetimes, and constant r_0 = 0.34 and ϕ = 20.34, experimentally determined for 1xGFP. The observed increase of more than Δr = 0.010 can be realistically traced back to the photobleaching asymmetry as a lifetime difference of ±0.5 ns already results in an anisotropy shift of ±0.006.

Thus, it can be assumed that GFP fluorophores with lifetimes shorter than average are more resistant against photobleaching (**Figure 33A**). In consequence, GFP fluorophores with low individual anisotropies are more likely to be photobleached first than fluorophores with high individual anisotropies. Due to this asymmetry, the steady-state anisotropy is slightly raised for the growing fraction of inactive fluorophores.

When calculating N from concatemer anisotropy data, both models by Yeow and Clayton [57], a polynomial approach, eq. *(11)*, and a two-state bimodal model, eq. *(12)*, could not determine N satisfactorily (**Figure 18**, **Table 8**; 2nd column in **Table 9**). We found that the parameter f_{non} in the two-state bimodal model, which was proposed by Yeow and

Clayton to describe the fraction of non-interacting fluorophores, is close to its maximum f_{non} = 1 for all of the investigated assemblies. Although, this suggests a large fraction of non-interacting – and thus – monomeric fluorophores, the biochemical analysis of sfGFP concatemers shows no major monomer fraction (**Figure 13D**). This evident divergence could be based on the circumstances that the model was conceived for the fractional labeling of heterogeneous complexes instead of the fractional photobleaching of homogeneous fluorophore assemblies, as presented in our case. Indeed, for a fully assembled (f_{non} = 0) and non-bleached (x = 0) cluster of fluorophores, we calculated that the steady-state anisotropy equals zero for all oligomers, according to Yeow and Clayton's bimodal model. Certainly, this does not reflect biophysical reality since clusters of homogeneous fluorophores differ in steady-state anisotropy with inverse proportionality to the number of subunits N [46]. As a consequence, the bimodal model was complemented with an extension replacing f_{non} with a term that incorporates N, eq. *(14)*. Using this new model, the calculated values are comparable to those based on time-resolved data and the relevant parameters k_{FRET} and τ (**Table 9**). This strikingly demonstrates that it is feasible to systematically determine the stoichiometry of monomers, dimers, and trimers with a relatively simple setup.

The stoichiometry of coiled-coil peptides can be determined with the steady-state fluorescence anisotropy.

Similar to the concatemer model proteins, a selection of oligomerizing, α-helical coiled-coil peptides was tested (**Figure 20**). Differently to sfGFP concatemers where the number of GFP reporter proteins is defined, working with non-covalently assembled GFP-coiled-coil fusion proteins required adequate controls by biochemical methods (**Figure 21**). Analysis with native PAGE and SEC, however, revealed several difficulties to resolve the oligomeric state of some of the GFP-coiled-coil fusion proteins. As some could not be separated properly in native PAGE due to their too basic isoelectric point (sfGFP-based proteins), others exhibited ambiguous elution profiles in SEC.

For GFP-coiled-coil proteins, the steady-state as well as time-resolved anisotropy can clearly distinguish between a monomeric, dimeric, or trimeric assembly, with statistical significance (**Figure 23**). The anisotropy analysis allows the determination of coiled-coil stoichiometry integrated in GFP fusion proteins (**Table 11, Table 12**). The results in this thesis showed that

the stoichiometry of sfGFP-coiled-coil can be determined with the same parameters as for sfGFP concatemers, the average values for r_{ET} and k_{FRET}, eq. *(8)*. Since linker lengths and the structural composition vary between the EGFP- and sfGFP-based constructs, the parameters to calculate oligomeric state *N* had to be adjusted. When determining *N* of EGFP-coiled-coil fusion protein via fractional photobleaching, eq. *(14)*, the parameter a was adapted as well (**Figure 25**).

The calculated oligomeric states for dimeric sfGFP-GCN4-p1 and trimeric sfGFP-GCN4-pII are in line with X-ray structures published for the isolated peptides in the past [35, 144]. The results for the dimeric EGFP-CC-Di, trimeric EGFP-CC-Tet, and pentameric EGFP-CC-Pent coincide with published protein stoichiometries, too [14, 18, 162]. Determining *N* for sfGFP-ph3a (**Figure 22**) demonstrated that the *de novo* peptide is likely to form a higher-order protein complex when covalently linked to sfGFP. Results from total internal reflection fluorescence microscopy by Ganesh Agam confirmed a super structure with more than seven subunits (data not shown). Hereby, it becomes apparent that the anisotropy-based approach differs well between higher-order oligomers with very condensed GFP reporter domains and lower-order oligomers. Interestingly, this was possible even with the technical limitation for complexes with $N > 3$ fluorophores, explained in a previous section.

With a view on the future anisotropy analyses of other GFP-coiled-coil oligomers, the results in this thesis emphasize that an altered structural composition potentially requires arranged parameters to calculate the oligomeric state N. When systematically studying the stoichiometry of various oligomerizing proteins via fluorescence anisotropy, one uniform composition should be used.

6.2 Photodegradation of green fluorescent protein

The molecular mechanism behind the irreversible photobleaching of fluorescent proteins is still poorly understood. Findings by Adam et al. [101] suggest that photobleaching directly follows radical-formation at the fluorophore (**Figure 6**). This assumption is based on observations for the photoswitchable fluorescent protein irisFP that was irradiated with X-rays [101, 102].

In our photobleaching experiments, we recognized that the irradiation of GFP fusion proteins with visible blue laser light (**Figure 26A**, **Figure 11**) not only led to the irreversible reduction of fluorescence intensity (**Figure 26B**, **Figure 27**), but also to a loss in structural integrity of the proteins. The photodegradation upon high-intensity illumination is observed solely for GFP. Non-fluorescent proteins, jointly in solution and simultaneously irradiated, are not affected by the irradiation (**Figure 28C**).

The effect was initially detected by determining the molecular diffusion coefficients of sfGFP concatemers 1xGFP, 2xGFP, 3xGFP before and after photobleaching using fluorescence correlation spectroscopy (FCS), conducted by Ganesh Agam (**Appendix Fig. 6**). In FCS analysis, the fusion proteins moved 1.64 times faster through the focal plane after having been heavily irradiated. This suggests that the irradiation with highly-energetic laser light alters the molecular size of the GFP fusion proteins and potentially their structure or conformation. The extent of GFP damage due to irradiation, however, remained unclear in FCS measurements. It must be also noted that the recorded diffusion coefficients came from molecules which were apparently still fluorescent.

Surprisingly, a similar phenomenon has never been described for fluorescent proteins irradiated with visible light. Though, it is known that the irradiation of photoswitchable proteins such as Kaede, EosFP, and IrisFP with near-UV light leads to backbone breakage accompanied by the formation of altered chromophore π-systems [105, 108, 109]. Their structural deformation leads to a decrease in molecular weight [108] and a new emission peak at 580 nm after the “switched-on” Kaede is excited at 543 nm or 572 nm, EosFP at 571 nm, and IrisFP at 551 nm. However, the irradiated GFP fusion proteins in this study do not show any red-shifted emission maxima (**Figure 26C**).

This phenomenon raises several questions, most prominently whether photodegradation observed in our experiments is related to the backbone cleavage [105, 108-110], whether it can be traced back to radical formation during the photobleaching process [101, 103], and whether photobleaching necessarily leads to the photodegradation.

Irreversible photobleaching of GFP does not cause complete photo-degradation.

The irradiation of GFP fusion proteins with visible, high-intensity laser light (λ = 445 nm; 300 mW) led to a decrease of fluorophores. It also reduced the amount of intact protein (**Figure 28A**). Thereby, not every photobleached fluorophore was degraded: according to the densitometric analysis of photobleached samples (**Figure 28B**), the percentage of intact proteins is still around 35% when only 20% of the initial fluorescence is left. Practically, it was never possible to completely reduce the fluorescence as the decrease in active fluorophores upon photobleaching describes an exponential decay. Consequently, the photobleaching does not necessarily cause photodegradation. This also brings up the question whether photodegraded proteins have to be necessarily photobleached. Fluorescence correlation spectroscopy, a method that relies on the fluorescence of investigated particles, demonstrates that the diffusion of photobleached GFP fusion proteins is increased which is a direct indication for photodegradation or conformational changes of the molecule (**Appendix Fig. 6**). It implies that the measured fluorescent proteins are both, still fluorescent but structurally affected by the irradiation. However, studies on split or truncated forms of GFP [163-165] and studies on its minimally required structure for fluorescence [166] demonstrated that fragments of GFPs are not functional. Based on these findings, the idea of a still fluorescent but severely fragmented GFP has to be excluded. Nevertheless, this does not except conformational changes of the heavily irradiated proteins that may be the reason for the raised diffusion coefficients. Indeed, a conformationally altered version of the sfGFP fusion protein 1xGFP could be detected in PAGE and immunoblot analysis (sf-A, **Figure 31A**). The corresponding band was only detectable upon photobleaching and migrated faster than the sfGFP main band.

The photodegradation of GFP is not arbitrary and occurs at a number of preferred sites.

The irradiation of sfGFP and EGFP leads to a characteristic pattern of fragments (**Table 13**) that can be detected via PAGE (**Figure 29**), immunoblot (**Figure 31**, **Figure 32**), and mass spectrometry (**Appendix Fig. 7-10**).

For fusion proteins holding sfGFP [81] and EGFP [84], a small N-terminal (sf-G, e-D) and a larger C-terminal fragment (sf-B, e-A) are found, whose approximate molecular masses can be complemented to the intact protein size. Besides this prominent fragmentation site, three other sites were reconstructed by combining further N-terminal with C-terminal fragments, sf-E with sf-Ct25 and sf-C with sf-F for sfGFP (**Figure 31B**), and e-Nt14 with e-B for EGFP (**Figure 32B**).

In a MALDI-TOF analysis of 1xGFP (**Appendix Fig. 7**), executed by Dr. Martin Haslbeck, Chair of Biotechnology at TUM, Garching, the masses of the small sf-G (8341.0 ± 0.6 Da) and the large complementary sf-B fragment (27997 ± 124 Da) were determined. Interestingly, peaks that were strongly increased after irradiation were already weakly present in non-irradiated samples. This could imply a weak fragmentation already caused by ambient light. MALDI-TOF measurements with non-irradiated and irradiated EGFP samples identified the mass of fragment e-A (20016 ± 72.3 Da), e-Nt14 (14137.7 Da), e-B (15993.3 Da), and e-D (10044 Da) (**Appendix Fig. 8**). The mass of fragment e-D could only be determined when splitting the broad maximum at m/z = 10017 (grey in **Appendix Fig. 8**) into a sharp peak at m/z = 10008, corresponding to the double-charged e-A fragment, and a maximum at m/z = 10045 (**Appendix Fig. 9**).

For all MALDI-TOF measurements, the standard deviations of the calculated masses of differently charged but identical molecules were approximately 100-150 Da. When detecting the same peak, we also registered standard deviations of 35 Da on average. Contrasting our expectations, this shows that the MALDI-TOF data did not resolve the mass of the protein and the protein fragments with atomic precision.

Table 13: Detailed analysis of GFP fragmentation patterns upon high-intensity irradiation

Fusion protein (kDa)	Band	~*MW*, PAGE (kDa)	Response upon irrad. [1]	Terminus [2]	Compl. fragment pair [3]	MALDI-TOF (Da) [4]	ESI-TOF (Da)
1xGFP (36.4 kDa)	intact construct	(40)*	↓	N, C	-	36449	-
	sf-A	(38)*	→	N, C	-	-	-
	sf-B	(33)*	↑	C	**I**	27997	-
	sf-C	27	↑	N	**III**	-	-
	sf-Ct25	25	↑	C	**II**	-	-
	sf-D	24	→	-	-	-	-
	sf-E	18	↑	N	**II**	-	-
	sf-F	16	↑	C	**III**	-	-
	sf-G	8	↑	N	**I**	8340	8344
EGFP (30.0 kDa)	intact construct	30	↓	N	-	30084	30025
	e-A	22	↑	~~N~~	**I**	20016	20030
	e-B	18	↑	~~N~~	**IV**	15992	-
	e-Nt14	14	↑	N	**IV**	14138	
	e-C	13	→	~~N~~	-	12456	-
	e-D	10	↑	N	**I**	10044	-

[1] Band strength response upon irradiation with a 445 nm high-intensity laser.
[2] N- and C-terminal location, verified via immunoblot. For EGFP, only one terminus held an epitope; thus, molecules could be classified as N-terminal (N) and non-N-terminal (~~N~~).
[3] Complementary fragment pair (I, II, III, or IV) according to results from denaturing PAGE, immunoblot, MALDI-TOF, and ESI-TOF mass spectrometry.
[4] ± 100-150 Da (standard deviation).
( )* Schägger gradient gels are optimal to resolve molecules < 30 kDa [149]. For molecules of higher *MW*, inaccuracies in *MW* estimation might be a consequence.

In order to narrow down the precise masses of some fragments, the irradiated 1xGFP and EGFP fusion proteins were measured with ESI-TOF mass spectrometry by Walter Stelzer, at our group, Chair of Biopolymer Chemistry at TUM (**Appendix Fig. 10**). The noisy mass spectrum for irradiated 1xGFP (**Appendix Fig. 10A**) shows three prominent peaks which correspond to multiple-charged molecules with a mass of 8344.36 ± 0.03 Da. No other maximum could be detected.

When analyzing irradiated EGFP with ESI-TOF mass spectrometry, two peaks could be detected only when deconvoluting the data (**Appendix Fig. 10B**), a maximum at 20030 Da and at 30025 Da, which coincides with the proteins actual mass 30025.67 Da. A putative maximum near 10 kDa cannot be clearly determined due to high noise for masses below a 10-12 kDa threshold. Remarkably, all three masses determined with ESI-TOF mass spectrometry are congruent to results with MALDI-TOF equipment and observations from PAGE. It should also be noted that masses from PAGE and masses from MALDI-TOF analysis match better for fragments < 15-20 kDa, possibly due to a bad resolution for higher molecular weights by Schägger gel analysis (**Table 13**). Reconstructing possible cleavage sites, PAGE results for lower masses are prioritized against those with higher masses as a consequence.

In sum, we found that the light-induced fragmentation of GFP is not an arbitrary process. It involves the cleavage at specific areas of the GFP domain in our proteins. This results in complementary fragment pairs I, II, III, and IV (**Table 13**), assuming that the cleavage site represented by fragment pair I is identical for both GFP variants. From the combination of PAGE, immunoblot, and mass spectrometry results, it was possible to reconstruct cleavage sites at residue resolution.

The photodegradation of GFP occurs at the C_α atom of the former Thr65 within the mature fluorophore.

The analysis of GFP upon high-intensity laser irradiation (λ = 445 nm, 300 mW) repeatedly demonstrated a prominent pattern: a small N-terminal and a large C-terminal fragment, complementary to the full-sized protein (fragment pair I) (**Table 13**). The combination of exact fragment masses and the sequences that coincide with terminal parts of the intact protein can be used in order to determine a first potential cleavage site (**Figure 34**).

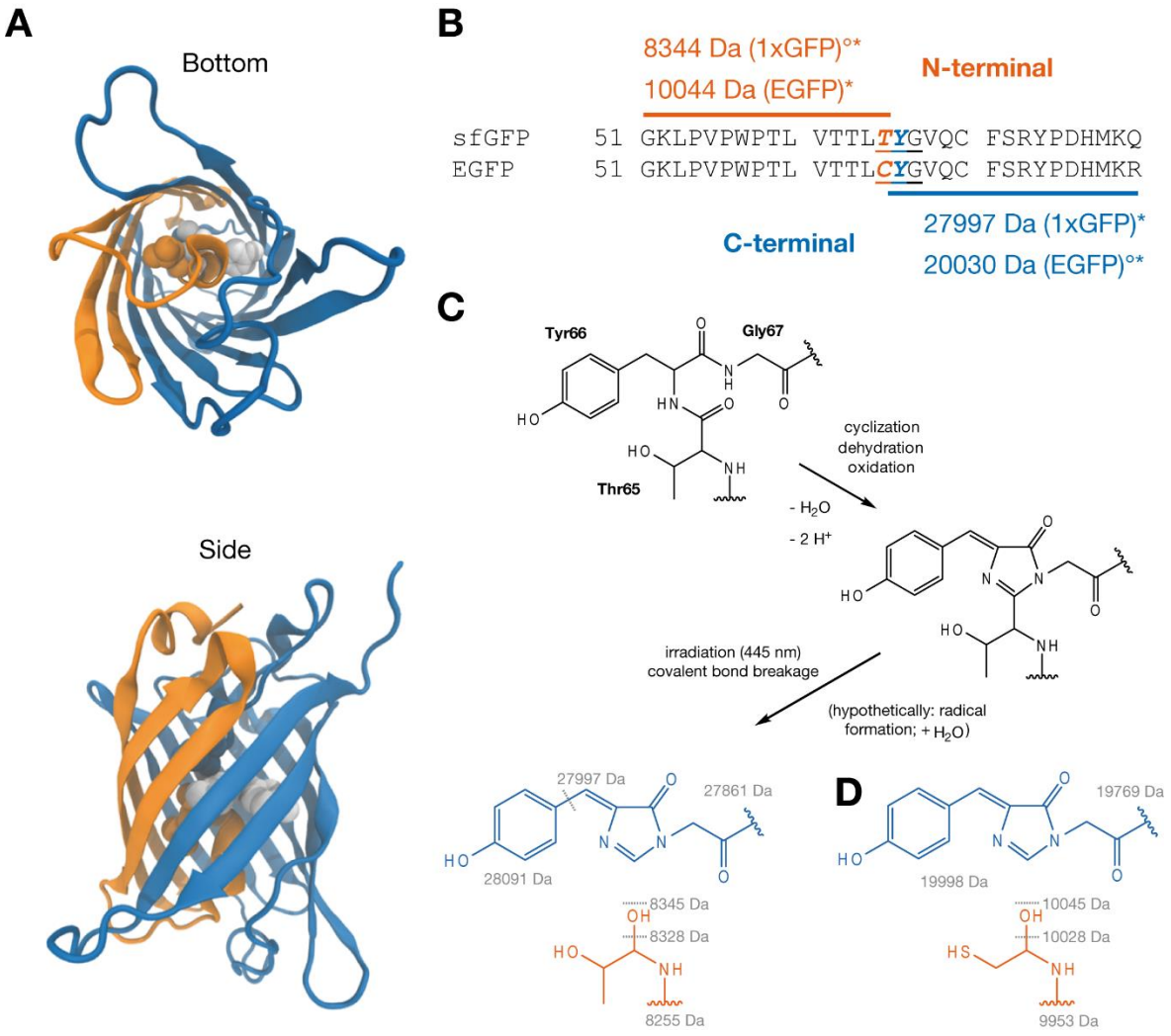


Figure 34: GFP is fragmented at the fluorophore upon high-intensity irradiation. A) Bottom and side view of GFP with the highlighted fluorophore. The fluorophore (sGFP: Thr65, Tyr66, Gly67; EGFP: Cys65, Tyr66, Gly67; white) divides the GFP molecule into a small N-terminal (orange) and a large C-terminal part (blue). It is located in the center of the beta-barrel tertiary structure. **B)** The irradiation of 1xGFP, a sfGFP fusion protein, and an EGFP fusion protein [18] with high-intensity laser light (λ = 445 nm) causes a covalent bond breakage. The protein cleavage can be located at the fluorophore by relating the size of the emerged fragments (*: fragment size determined by MALDI-TOF spectrometry; °: fragment size determined by ESI-TOF spectrometry). **C)** Hypothetical product after the covalent bond breakage of 1xGFP (36418 Da). The comparison between calculated masses (grey) and the measured masses of the fragments (8344 Da and 27997 Da) indicate that the protein might be fragmented between the C_α atom of the former residue 65 and the imidazolinone, dividing the whole fusion protein into a small N-terminal (orange) and a large C-terminal fragment (blue). The mass of the small fragment suggests an addition of a hydroxyl group, hypothetically due to radical formation and the presence of H_2O. For a better comparison, the residue numbers refer to the residue sequence position in the isolated sfGFP [81]. **D)** Hypothetical product after the covalent bond breakage of EGFP (30026 Da).

For 1xGFP, a monomeric sfGFP fusion protein (36418 kDa, matured fluorophore and formylmethionine as the starting residue), the masses of the N-terminal fragment (8344 Da) and C-terminal fragment (27997 ± 124 Da) helped to reconstruct the approximate position of the cleavage site (**Figure 34B**). They revealed that the fragmentation takes place directly after the C_α atom of the former Thr65 (small fragment) and before the imidazolinone of the transformed Tyr66 and Gly67 (large fragment). When we compare the mass of the N-terminal fragment (8344 Da) with the mass of a possible fragment cleaved at the C_α atom (8328 Da) (**Figure 34C**), it becomes apparent that the C_α atom is likely to be modified, potentially with a hydroxyl group (8345 Da). Hypothetically, this could be caused by radical formation as a result of the irradiation.

The fragment masses from the investigated EGFP fusion protein (30025.67 Da) [18] can be assigned to a cleavage site exactly between the former fluorophore-forming residues Cys65 and Tyr66, the same position as detected for 1xGFP (**Figure 34B**). The mass of N-terminal fragment (10044 Da) supports the hypothesis of the hydroxyl group modification at the C_α atom since the calculated mass for a suchlike fragment would be 10045 Da (**Figure 34D**). The measured mass for the C-terminal fragment (20016 ± 72.3 Da) coincides with the calculated fragment size as well.

This demonstrates that the cleavage is located at the fluorophore (**Figure 4**) in the spatial center of the beta-barrel structure of GFP (**Figure 34A**). In the past, the position N-terminally to the fluorophore (residue 65) was classified as the fragmentation site for photoswitchable fluorescent proteins Kaede, EosFP, and IrisFP upon near-UV irradiation [105-109] or for GFP-Ala65-Ser66-Gly67 under very high temperatures [110]. Seemingly, this position can be – very figuratively – understood as the weakest bond of an intensively irradiated GFP.

The chemical mechanism of the fragmentation, however, is not understood. Therefore, the exact fragment sizes have to be known. Since the mass spectrometric analysis could not resolve the exact fragment masses in order to pinpoint the cleavage site on an atomic resolution, there is no basis for a mechanistic reconstruction.

EGFP is cleaved at Phe99, a residue relevant to protein folding that is not in direct proximity to the fluorophore.

Besides fragment pair I, the masses for fragment pair IV (exclusively for EGFP, 30025.67 Da) could be determined via mass spectrometry (**Table 13**). The N-terminal fragment (14138 Da) and the C-terminal fragment (15992 Da) strongly indicate that EGFP is cleaved at Phe99 (**Figure 35A&B**).

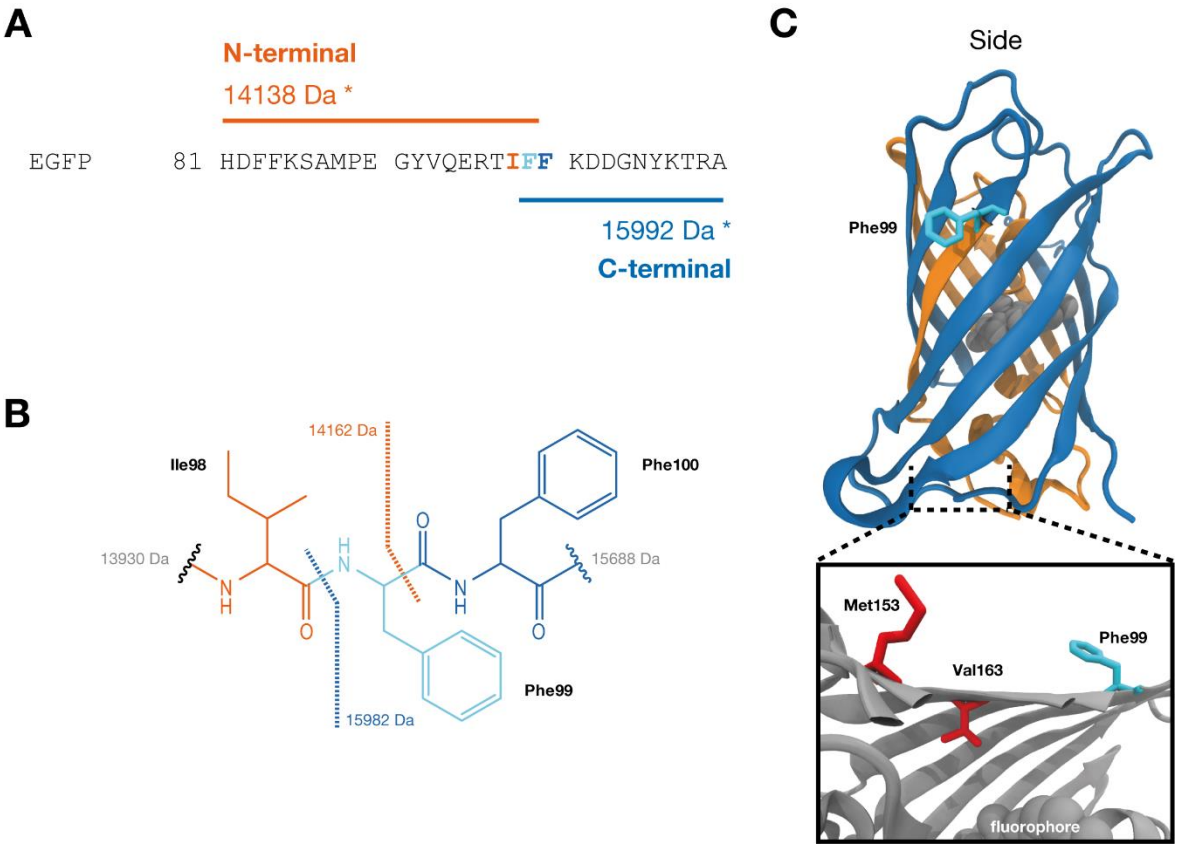


Figure 35: EGFP is cleaved at Phe99 upon high-intensity irradiation. A) The irradiation of an EGFP fusion protein [18] with high-intensity laser light (λ = 445 nm) causes a covalent bond breakage at Phe99. The cleavage site was reconstructed from the fragments masses determined by MALDI-TOF spectrometry, n = 1 (*). **B)** The calculated fragments (14162 Da and 15982 Da) and the measured (14138 Da and 15962 Da) coincide well. However, the precise cleavage site, N- or C-terminal of the Phe99 C_α atom, could not be determined. The residue numbers refer to the residue sequence position in the isolated EGFP [84]. **C)** Side view of EGFP with the highlighted Phe99 (cyan) and the fluorophore (grey). The cleavages at Phe99 divides the GFP into a N-terminal (orange) and a C-terminal part (blue). Bottom panel: Cutting EGFP along its vertical axis reveals that Phe99 (cyan), Met153 (red), and Val163 (red), are located in line but in three different beta-strands. Studies in the past demonstrated that these three residues are essential for GFP folding [150, 167, 168].

Surprisingly, Phe99 is not proximal to the fluorophore and its side chain is located outside the beta-barrel (**Figure 35C**). Together with Met153 and Val163, it defines three essential residues responsible for GFP folding [150, 167, 168]. In the past, the mutant F99S/M153T/V163A ("Cycle3") was found to have a highly increased fluorescence intensity and a blue-shifted excitation maximum (λ = 395 nm) [80, 152, 167-169]. When cutting the GFP beta-barrel along its vertical axis (**Figure 35C**), it becomes visible that residue 99, 153, and 163 are located in three different beta-strands and in one straight line parallel to the vertical axis [152]. The sfGFP variant, in contrast, contains the "Cycle3" mutations (**Figure 31C**) and does not break at position Phe99 which was supported by results from PAGE and mass spectrometry.

Other candidate cleavage sites include Phe145, His148, Phe165, and Ile167.

The precise characterization of all cleavage sites is not feasible from the current data, due to weak resolution of fragment masses in gradient gel analysis. Nevertheless, identifying the fragmentation of EGFP at Phe99 helped to gauge two further cleavage areas (fragment pair II and III) in the sfGFP fusion protein. Previously, we could determine several candidate residues (**Figure 31C**). Four of them, Phe145, His148, Phe165, and Ile167 are in proximity to the fluorophore and next to the residues 99, 181, 163, and 153 ("Cycle 3" mutant) (**Figure 36**). Possibly, the region around residue 99 and the highlighted residues of the adjacent beta-strands is destabilized upon high-intensity irradiation.

In theory, all of the previously described fragmentations would entail the loss of fluorescence [163-165]. FCS results however imply that – at least – some of the fusion proteins are cut C-terminally to the beta-barrel, thereby remaining fluorescent and exhibiting a faster molecular diffusion (**Appendix Fig. 6**). In fragment analysis, only one molecule, the N-terminal sf-C (27 kDa), fits to this specification. Thus, a possible fragmentation at Gly228 is conceivable (**Figure 31C**) as GFP molecules with a truncated C-terminus up to position 227 are reported to remain fluorescent [163]. It should be noted that the approximate mass of 27 kDa of the fragment must be treated with caution since gradient PAGE analysis in this thesis systematically overestimates the masses for fragments > 20 kDa. Further supporting mass spectrometric measurements could finally clarify this theory.

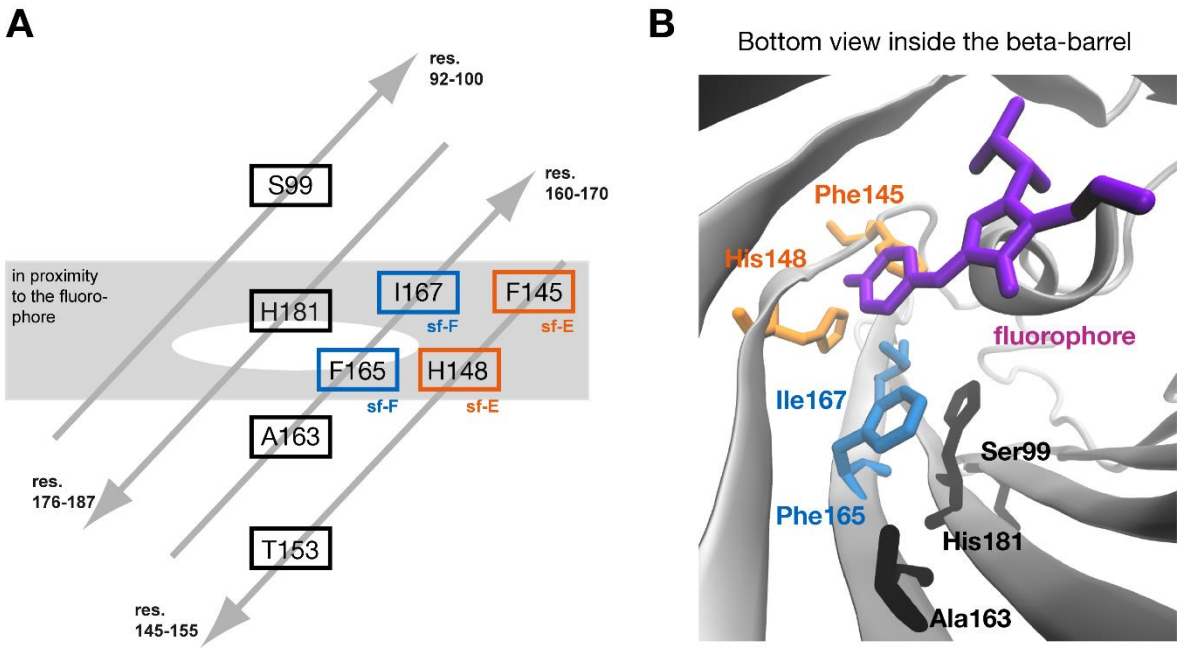


Figure 36: Further fragmentation of sfGFP upon high-intensity irradiation may take place proximally to Ser99/His181/Ala163/Thr153. A) Four positions Phe145, His148, Phe165, and Ile167 were speculated to be candidate residues for GFP fragmentation. The residue numbers refer to the residue sequence position in the isolated sfGFP [81]. The highlighted residues are proximal to the fluorophore (A: white ellipse; B: purple group) and close to "Cycle 3" mutant-related residues 99, 181, 163, and 153 (black). A schematic overview (**A**) and a corresponding protein structure (**B**) illustrate the position of the relevant residues. The arrows represent beta-strands (with residues). Residues at possible fragmentation sites are marked according to their corresponding fragments (orange = N-terminal, blue = C-terminal).

The effect of GFP photodegradation on homo-FRET-based approaches is negligible but might be considerable for other methods using photobleaching.

The new insights into GFP photodegradation lead to a substantial question: How strong is the impact of light-induced degradation on the design of experiments that rely on photobleaching as performed in chapters 5.1 and 5.2 of this thesis?

According to the previous considerations, it can be assumed that the effect of GFP photodegradation on techniques that exploit fluorescence anisotropy in combination with photobleaching is negligible. GFP photodegradation was shown to be accompanied with the partition of GFP in two fragments (e.g., at C_α atom of residue 65 or at Phe99). This necessarily leads to the loss of fluorescence in the majority of the cases, with no detectable effect on anisotropy as a consequence. The putative fragmentation at Gly228 is the only possibility that GFP photodegradation would directly affect the fluorescence anisotropy of the molecule. The densitometric analysis of gradient gels shows that the fragment that

corresponds to the fragmentation at Gly228, sf-C, exhibits a proportion of around 13% of total fragment band intensity. This indicates only a minor presence of the potential Gly228 fragments in the irradiated sample. Thus, solely little influence on the steady-state and time-resolved anisotropy can be assumed. Indeed, the oligomeric state of interacting coiled-coil sequences as well as the number of concatenated sfGFP subunits were determined without introducing a respective correction factor (**Table 9**, **Table 12**). Although the overall bias is expected to be minor, a better understanding of GFP photodegradation and the development of photodegradation-resistant GFP variants would help to eliminate this source of error.

The photobleaching of GFP is not only used in homo-FRET-based methods, but also in a number of different applications such as fluorescence recovery after photobleaching (FRAP) to measure protein diffusion [94, 170, 171] or fluorescence loss in photobleaching (FLIP) to determine the dynamics of membrane proteins [172, 173]. Until now, the implicit assumption in these methods is that photobleaching reduces the fluorescence intensity of GFP and leaves the protein intact. Our findings however suggest that the protein integrity upon photobleaching is not guaranteed. The extent to which photodegradation has an influence on these techniques is not clearly evident and should be elucidated in future studies.

6.3 Outlook

In future approaches, the applicability of these techniques could be extended in order to determine the stoichiometry of membrane-bound proteins. Most of the membrane-spanning proteins are assumed to form non-covalent assemblies [174]. There are already techniques, such as the ToxR, TOXCAT, BLaTM, or the GALLEX system, that can resolve homo- and hetero-dimerization *in vivo*. These tools are able to define the strength of transmembrane domain interaction very well, yet, they cannot distinguish dimers from trimers or higher oligomers [136, 175-177]. Moreover, when investigating integral membrane proteins *in vitro*, methods that are classically used to determine the stoichiometry of soluble protein assemblies are stretched to their limits. If one wants to study transmembrane proteins integrated into a lipid bilayer, techniques that are based on hydrodynamic measurements – most prominently, analytical ultracentrifugation or size exclusion chromatography – are not suitable. The here presented method to investigate protein stoichiometry via fluorescence anisotropy may thus inspire the future investigation of membrane-bound oligomers.

In parallel GFP photobleaching experiments, GFP was demonstrated to undergo covalent bond breakage when irradiated with high-intensity visible light. This was never shown before. The cleavage site could be localized vicinal to the fluorophore. Further experiments could specify the exact position of this site as well as the second cleavage area that could only be vaguely indicated in this thesis. This could substantiate the understanding of limits in fluorescent protein application and could help to improve future GFP variants. Beyond that, a technical application where proteins that are linked to the N- and C-terminal to GFP are gradually separated via irradiation-based cleavage might be promising.

7 References

1. Goodsell, D. S. and Olson, A. J. (2000) Structural Symmetry and Protein Function. *Annual Review of Biophysics and Biomolecular Structure* 29(1), 105-153.
2. Brinda, K. V. and Vishveshwara, S. (2005) Oligomeric protein structure networks: insights into protein-protein interactions. *BMC bioinformatics* 6, 296-296.
3. Nooren, I. M. A. and Thornton, J. M. (2003) Structural Characterisation and Functional Significance of Transient Protein–Protein Interactions. *Journal of Molecular Biology* 325(5), 991-1018.
4. Ali, M. H. and Imperiali, B. (2005) Protein oligomerization: How and why. *Bioorganic & Medicinal Chemistry* 13(17), 5013-5020.
5. Lupas, A., Van Dyke, M., and Stock, J. (1991) Predicting coiled coils from protein sequences. *Science* 252(5009), 1162.
6. O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1989) Evidence that the Leucine Zipper is a Coiled Coil. *Science* 243(4890), 538-542.
7. Hodges, R. S., Sodek, J., Smillie, L. B., and Jurasek, L. (1973) Tropomyosin: Amino Acid Sequence and Coiled-Coil Structure. *Cold Spring Harbor Symposia on Quantitative Biology* 37, 299-310.
8. Crick, F. H. C. (1953) The packing of α-helices: simple coiled-coils. *Acta Crystallographica* 6(8-9), 689-697.
9. Scholtz, J. M., Qian, H., Robbins, V. H., and Baldwin, R. L. (1993) The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry* 32(37), 9668-9676.
10. Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E., and Matthews, C. R. (1995) Probing the Folding Mechanism of a Leucine Zipper Peptide by Stopped-Flow Circular Dichroism Spectroscopy. *Biochemistry* 34(39), 12812-12819.
11. Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240(4860), 1759.
12. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) A Switch Between Two- , Three-, and Four-Stranded Coiled Coils in GCN4 Leucine Zipper Mutants. *Science* 262(5138), 1401-1407.
13. Lupas, A. N. and Gruber, M. (2005) The structure of alpha-helical coiled coils. *Advances in Protein Chemistry* 70, 37-78.
14. Fletcher, J. M., Boyle, A. L., Bruning, M., Bartlett, G. J., Vincent, T. L., Zaccai, N. R., Armstrong, C. T., Bromley, E. H. C., Booth, P. J., Brady, R. L., Thomson, A. R., and Woolfson, D. N. (2012) A Basis Set of de Novo Coiled-Coil Peptide Oligomers for Rational Protein Design and Synthetic Biology. *ACS Synthetic Biology* 1(6), 240-250.
15. Thomson, A. R., Wood, C. W., Burton, A. J., Bartlett, G. J., Sessions, R. B., Brady, R. L., and Woolfson, D. N. (2014) Computational design of water-soluble α-helical barrels. *Science* 346(6208), 485-488.
16. Wood, C. W., Bruning, M., Ibarra, A. Á., Bartlett, G. J., Thomson, A. R., Sessions, R. B., Brady, R. L., and Woolfson, D. N. (2014) CCBuilder: an interactive web-based tool for building, designing and assessing coiled-coil protein assemblies. *Bioinformatics* 30(21), 3029-3035.

17. Wood, C. W. and Woolfson, D. N. (2018) CCBuilder 2.0: Powerful and accessible coiled-coil modeling. *Protein Science* 27(1), 103-111.
18. Cristie-David, A. S., Sciore, A., Badieyan, S., Escheweiler, J. D., Koldewey, P., Bardwell, J. C. A., Ruotolo, B. T., and Marsh, E. N. G. (2017) Evaluation of de novo-designed coiled coils as off-the-shelf components for protein assembly. *Molecular Systems Design & Engineering* 2(2), 140-148.
19. Svedberg, T. and Fåhraeus, R. (1926) A new method for the determination of the molecular weight of the proteins. *Journal of the American Chemical Society* 48(2), 430-438.
20. Freundlich, H. (1940) The Ultracentrifuge. By The Svedberg and Kai O. Pedersen. *The Journal of Physical Chemistry* 44(7), 952-953.
21. Schachman, H. K. (1957) Ultracentrifugation, diffusion, and viscometry. in *Methods in Enzymology*, 32-103.
22. Schuck, P. (2013) Analytical Ultracentrifugation as a Tool for Studying Protein Interactions. *Biophysical reviews* 5(2), 159-171.
23. Gell, D. A., Grant, R. P., and Mackay, J. P. (2012) The Detection and Quantitation of Protein Oligomerization. *Protein Dimerization and Oligomerization in Biology*, ed Matthews J. M. (Springer New York, New York, NY), 19-41.
24. Schmid, J. A., Scholze, P., Kudlacek, O., Freissmuth, M., Singer, E. A., and Sitte, H. H. (2001) Oligomerization of the Human Serotonin Transporter and of the Rat GABA Transporter 1 Visualized by Fluorescence Resonance Energy Transfer Microscopy in Living Cells. *Journal of Biological Chemistry* 276(6), 3805-3810.
25. Schägger, H., Cramer, W. A., and Vonjagow, G. (1994) Analysis of Molecular Masses and Oligomeric States of Protein Complexes by Blue Native Electrophoresis and Isolation of Membrane Protein Complexes by Two-Dimensional Native Electrophoresis. *Analytical Biochemistry* 217(2), 220-230.
26. Wittig, I., Braun, H. P., and Schägger, H. (2006) Blue native PAGE. *Nature protocols* 1(1), 418-428.
27. Krause, F. (2006) Detection and analysis of protein–protein interactions in organellar and prokaryotic proteomes by native gel electrophoresis: (Membrane) protein complexes and supercomplexes. *Electrophoresis* 27(13), 2759-2781.
28. Hillar, A., Culham, D. E., Vernikovska, Y. I., Wood, J. M., and Boggs, J. M. (2005) Formation of an Antiparallel, Intermolecular Coiled Coil Is Associated with in Vivo Dimerization of Osmosensor and Osmoprotectant Transporter ProP in Escherichia coli. *Biochemistry* 44(30), 10170-10180.
29. Solis, Gonzalo P., Hoegg, M., Munderloh, C., Schrock, Y., Malaga-Trillo, E., Rivera-Milla, E., and Stuermer, Claudia A. O. (2007) Reggie/flotillin proteins are organized into stable tetramers in membrane microdomains. *Biochemical Journal* 403(2), 313.
30. Bada, M., Walther, D., Arcangioli, B. t., Doniach, S., and Delarue, M. (2000) Solution structural studies and low-resolution model of the Schizosaccharomyces pombe sap1 protein11Edited by M. F. Moody. *Journal of Molecular Biology* 300(3), 563-574.
31. Wood, Z. A., Poole, L. B., Hantgan, R. R., and Karplus, P. A. (2002) Dimers to doughnuts: Redox-sensitive oligomerization of 2-cysteine peroxiredoxins. *Biochemistry* 41(17), 5493-5504.
32. Putnam, C. D., Hammel, M., Hura, G. L., and Tainer, J. A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics* 40(3), 191-285.

33. Clore, G. M., Omichinski, J. G., Sakaguchi, K., Zambrano, N., Sakamoto, H., Appella, E., and Gronenborn, A. M. (1994) High-resolution structure of the oligomerization domain of p53 by multidimensional NMR. *Science* 265(5170), 386-391.
34. Bocharov, E. V., Volynsky, P. E., Pavlov, K. V., Efremov, R. G., and Arseniev, A. S. (2010) Structure elucidation of dimeric transmembrane domains of bitopic proteins. *Cell Adhesion and Migration* 4(2), 284-298.
35. O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) X-ray Structure of the GCN4 Leucine Zipper, a Two-Stranded, Parallel Coiled Coil. *Science* 254(5031), 539-544.
36. Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) Core Structure of gp41 from the HIV Envelope Glycoprotein. *Cell* 89(2), 263-273.
37. Woolfson, D. N., Bartlett, G. J., Bruning, M., and Thomson, A. R. (2012) New currency for old rope: From coiled-coil assemblies to α-helical barrels. *Current Opinion in Structural Biology* 22(4), 432-441.
38. Gadella Jr, T. W. J. and Jovin, T. M. (1995) Oligomerization of epidermal growth factor receptors on A431 cells studied by time-resolved fluorescence imaging microscopy. A stereochemical model for tyrosine kinase receptor activation. *Journal of Cell Biology* 129(6), 1543-1558.
39. Overton, M. C. and Blumer, K. J. (2000) G-protein-coupled receptors function as oligomers in vivo. *Current Biology* 10(6), 341-344.
40. Truong, K. and Ikura, M. (2001) The use of FRET imaging microscopy to detect protein–protein interactions and protein conformational changes in vivo. *Current Opinion in Structural Biology* 11(5), 573-578.
41. DeGrado, W. F., Gratkowski, H., and Lear, J. D. (2003) How do helix-helix interactions help determine the folds of membrane proteins? Perspectives from the study of homo-oligomeric helical bundles. *Protein Science* 12(4), 647-665.
42. Meyer, B. H., Segura, J. M., Martinez, K. L., Hovius, R., George, N., Johnsson, K., and Vogel, H. (2006) FRET imaging reveals that functional neurokinin-1 receptors are monomeric and reside in membrane microdomains of live cells. *Proceedings of the National Academy of Sciences of the United States of America* 103(7), 2138-2143.
43. Margineanu, A., Chan, J. J., Kelly, D. J., Warren, S. C., Flatters, D., Kumar, S., Katan, M., Dunsby, C. W., and French, P. M. W. (2016) Screening for protein-protein interactions using Förster resonance energy transfer (FRET) and fluorescence lifetime imaging microscopy (FLIM). *Scientific Reports* 6, 28186.
44. Förster, T. (1948) Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik* 437(1- 2), 55-75.
45. Wawilow, S. J. and Lewschin, W. L. (1923) Beiträge zur Frage über polarisiertes Fluoreszenzlicht von Farbstofflösungen. II - Die Bestätigung des Weigertschen Fluoreszenzeffekts und neue Beobachtungen. *Zeitschrift für Physik* 16(1), 135-154.
46. Runnels, L. W. and Scarlata, S. F. (1995) Theory and application of fluorescence homotransfer to melittin oligomerization. *Biophysical Journal* 69(4), 1569-1583.
47. Lakowicz, J. R. (2006) *Principles of fluorescence spectroscopy* (Springer, 233 Spring Street, New York, NY 10013, USA) 3rd Ed.
48. Gautier, I., Tramier, M., Durieux, C., Coppey, J., Pansu, R. B., Nicolas, J. C., Kemnitz, K., and Coppey-Moisan, M. (2001) Homo-FRET microscopy in living cells to measure monomer-dimer transition of GFP-tagged proteins. *Biophysical Journal* 80(6), 3000-3008.

49. Sharma, P., Varma, R., Sarasij, R. C., Ira, Gousset, K., Krishnamoorthy, G., Rao, M., and Mayor, S. (2004) Nanoscale Organization of Multiple GPI-Anchored Proteins in Living Cell Membranes. *Cell* 116(4), 577-589.
50. Goswami, D., Gowrishankar, K., Bilgrami, S., Ghosh, S., Raghupathy, R., Chadda, R., Vishwakarma, R., Rao, M., and Mayor, S. (2008) Nanoclusters of GPI-Anchored Proteins Are Formed by Cortical Actin-Driven Activity. *Cell* 135(6), 1085-1097.
51. Bader, A. N., Hofman, E. G., Voortman, J., van Bergen en Henegouwen, P. M. P., and Gerritsen, H. C. (2009) Homo-FRET Imaging Enables Quantification of Protein Cluster Sizes with Subcellular Resolution. *Biophysical Journal* 97(9), 2613-2622.
52. Szabó, A., Horváth, G., Szöllosi, J., and Nagy, P. (2008) Quantitative characterization of the large-scale association of ErbB1 and ErbB2 by flow cytometric homo-FRET measurements. *Biophysical Journal* 95(4), 2086-2096.
53. Gholami, Z., Brunsveld, L., and Hanley, Q. (2013) PNA-Induced Assembly of Fluorescent Proteins Using DNA as a Framework. *Bioconjugate Chemistry* 24(8), 1378-1386.
54. Zolmajd-Haghighi, Z. and Hanley, Q. S. (2014) When one plus one does not equal two: Fluorescence anisotropy in aggregates and multiply labeled proteins. *Biophysical Journal* 106(7), 1457-1466.
55. Ganguly, S., Clayton, A. H. A., and Chattopadhyay, A. (2011) Organization of higher-order oligomers of the serotonin 1A receptor explored utilizing homo-FRET in live cells. *Biophysical Journal* 100(2), 361-368.
56. Melo, A. M., Fedorov, A., Prieto, M., and Coutinho, A. (2014) Exploring homo-FRET to quantify the oligomer stoichiometry of membrane-bound proteins involved in a cooperative partition equilibrium. *Physical Chemistry Chemical Physics* 16(34), 18105-18117.
57. Yeow, E. K. L. and Clayton, A. H. A. (2007) Enumeration of oligomerization states of membrane proteins in living cells by homo-FRET spectroscopy and microscopy: Theory and application. *Biophysical Journal* 92(9), 3098-3104.
58. Kozer, N., Kelly, M. P., Orchard, S., Burgess, A. W., Scott, A. M., and Clayton, A. H. A. (2011) Differential and Synergistic Effects of Epidermal Growth Factor Receptor Antibodies on Unliganded ErbB Dimers and Oligomers. *Biochemistry* 50(18), 3581-3590.
59. Rocheleau, J. V., Edidin, M., and Piston, D. W. (2003) Intrasequence GFP in class I MHC molecules, a rigid probe for fluorescence anisotropy measurements of the membrane environment. *Biophysical Journal* 84(6), 4078-4086.
60. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994) Green fluorescent protein as a marker for gene expression. *Science* 263(5148), 802-805.
61. Reid, B. G. and Flynn, G. C. (1997) Chromophore Formation in Green Fluorescent Protein. *Biochemistry* 36(22), 6786-6791.
62. Tsien, R. Y. (1998) The Green Fluorescent Protein. *Annual Review of Biochemistry* 67(1), 509-544.
63. Heim, R., Prasher, D. C., and Tsien, R. Y. (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences* 91(26), 12501.
64. Johnson, F. H., Shimomura, O., Saiga, Y., Gershman, L. C., Reynolds, G. T., and Waters, J. R. (1962) Quantum efficiency of Cypridina luminescence, with a note on that of Aequorea. *Journal of Cellular and Comparative Physiology* 60(1), 85-103.

65. Salih, A., Larkum, A., Cox, G., Kuhl, M., and Hoegh-Guldberg, O. (2000) Fluorescent pigments in corals are photoprotective. *Nature* 408(6814), 850-853.
66. Chudakov, D. M., Lukyanov, S., and Lukyanov, K. A. (2005) Fluorescent proteins as a toolkit for in vivo imaging. *Trends in Biotechnology* 23(12), 605-613.
67. Jan Willem, B. and Antonie, J. W. G. V. (2010) Fluorescence lifetime imaging microscopy in life sciences. *Measurement Science and Technology* 21(10), 102002.
68. Donner, J. S., Thompson, S. A., Kreuzer, M. P., Baffou, G., and Quidant, R. (2012) Mapping intracellular temperature using green fluorescent protein. *Nano Letters* 12(4), 2107-2111.
69. Acharya, A., Bogdanov, A. M., Grigorenko, B. L., Bravaya, K. B., Nemukhin, A. V., Lukyanov, K. A., and Krylov, A. I. (2017) Photoinduced Chemistry in Fluorescent Proteins: Curse or Blessing? *Chemical Reviews* 117(2), 758-795.
70. Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173(1), 33-38.
71. Yang, T. T., Cheng, L., and Kain, S. R. (1996) Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucleic Acids Research* 24(22), 4592-4593.
72. Patterson, G. H., Knobel, S. M., Sharif, W. D., Kain, S. R., and Piston, D. W. (1997) Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophysical Journal* 73(5), 2782-2790.
73. Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995) Improved green fluorescence. *Nature* 373, 663.
74. Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A., and Tsien, R. Y. (1995) Understanding, improving and using green fluorescent proteins. *Trends in Biochemical Sciences* 20(11), 448-455.
75. van Thor, J. J., Gensch, T., Hellingwerf, K. J., and Johnson, L. N. (2001) Phototransformation of green fluorescent protein with UV and visible light leads to decarboxylation of glutamate 222. *Nature Structural Biology* 9, 37.
76. Patterson, G. H. and Lippincott-Schwartz, J. (2002) A Photoactivatable GFP for Selective Photolabeling of Proteins and Cells. *Science* 297(5588), 1873.
77. Suzuki, T., Arai, S., Takeuchi, M., Sakurai, C., Ebana, H., Higashi, T., Hashimoto, H., Hatsuzawa, K., and Wada, I. (2012) Development of cysteine-free fluorescent proteins for the oxidative environment. *PLoS ONE* 7(5).
78. Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., and Remington, S. J. (1996) Crystal Structure of the Aequorea victoria Green Fluorescent Protein. *Science* 273(5280), 1392.
79. Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nature Biotechnology* 20, 87.
80. Lambert, T. (2018) Fluorescent Protein Base (tlambert03/FPbase: v1.1.0). 10.5281/zenodo.1316956. https://www.fpbase.org/. Accessed on July 19, 2018.
81. Pedelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology* 24(1), 79-88.
82. Balleza, E., Kim, J. M., and Cluzel, P. (2017) Systematic characterization of maturation time of fluorescent proteins in living cells. *Nature Methods* 15, 47.

83. Zacharias, D. A., Violin, J. D., Newton, A. C., and Tsien, R. Y. (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* 296(5569), 913-916.
84. Stauber, R. H., Horie, K., Carney, P., Hudson, E. A., Tarasova, N. I., Gaitanaris, G. A., and Pavlakis, G. N. (1998) Development and applications of enhanced green fluorescent protein mutants. *Biotechniques* 24(3), 462-466, 468-471.
85. Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L., and Lukyanov, S. A. (1999) Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature Biotechnology* 17, 969.
86. Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E., and Tsien, R. Y. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology* 22, 1567.
87. Day, R. N. and Davidson, M. W. (2009) The fluorescent protein palette: tools for cellular imaging. *Chemical Society Reviews* 38(10), 2887-2921.
88. Thaler, C., Koushik, S. V., Puhl Iii, H. L., Blank, P. S., and Vogel, S. S. (2009) Structural rearrangement of CaMKIIα catalytic domains encodes activation. *Proceedings of the National Academy of Sciences of the United States of America* 106(15), 6369-6374.
89. Nguyen, T. A., Sarkar, P., Veetil, J. V., Koushik, S. V., and Vogel, S. S. (2012) Fluorescence Polarization and Fluctuation Analysis Monitors Subunit Proximity, Stoichiometry, and Protein Complex Hydrodynamics. *PLoS ONE* 7(5), e38209.
90. Vámosi, G., Mücke, N., Müller, G., Krieger, J. W., Curth, U., Langowski, J., and Tóth, K. (2016) EGFP oligomers as natural fluorescence and hydrodynamic standards. *Scientific Reports* 6, 33022.
91. Kremers, G. J. and Goedhart, J. (2009) Chapter 5 Visible fluorescent proteins for FRET. *Laboratory Techniques in Biochemistry and Molecular Biology* 33, 171-223.
92. Akrap, N., Seidel, T., and Barisas, B. G. (2010) Förster distances for fluorescence resonant energy transfer between mCherry and other visible fluorescent proteins. *Analytical Biochemistry* 402(1), 105-106.
93. Kimura, H. and Cook, P. R. (2001) Kinetics of core histones in living human cells: Little exchange of H3 and H4 and some rapid exchange of H2B. *Journal of Cell Biology* 153(7), 1341-1353.
94. Reits, E. A. J. and Neefjes, J. J. (2001) From fixed to FRAP: Measuring protein mobility and activity in living cells. *Nature Cell Biology* 3(6), E145-E147.
95. Ulbrich, M. H. and Isacoff, E. Y. (2007) Subunit counting in membrane-bound proteins. *Nature Methods* 4(4), 319-321.
96. Burnette, D. T., Sengupta, P., Dai, Y., Lippincott-Schwartz, J., and Kachar, B. (2011) Bleaching/blinking assisted localization microscopy for superresolution imaging using standard fluorescent molecules. *Proceedings of the National Academy of Sciences* 108(52), 21081.
97. Swaminathan, R., Hoang, C. P., and Verkman, A. S. (1997) Photobleaching recovery and anisotropy decay of green fluorescent protein GFP-S65T in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. *Biophysical Journal* 72(4), 1900-1907.
98. Andresen, M., Wahl, M. C., Stiel, A. C., Gräter, F., Schäfer, L. V., Trowitzsch, S., Weber, G., Eggeling, C., Grubmüller, H., Hell, S. W., and Jakobs, S. (2005) Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proceedings*

of the National Academy of Sciences of the United States of America 102(37), 13070-13074.

99. Henderson, J. N., Ai, H.-w., Campbell, R. E., and Remington, S. J. (2007) Structural basis for reversible photobleaching of a green fluorescent protein homologue. *Proceedings of the National Academy of Sciences* 104(16), 6672.
100. Nifosì, R. and Tozzini, V. (2006) Cis–trans photoisomerization of the chromophore in the green fluorescent protein variant E2GFP: A molecular dynamics study. *Chemical Physics* 323(2), 358-368.
101. Adam, V., Carpentier, P., Violot, S., Lelimousin, M., Darnault, C., Nienhaus, G. U., and Bourgeois, D. (2009) Structural basis of X-ray-induced transient photobleaching in a photoactivatable green fluorescent protein. *Journal of the American Chemical Society* 131(50), 18063-18065.
102. Adam, V. (2009) Mechanistic studies of photoactivatable fluorescent proteins: a combined approach by crystallography and spectroscopy. Doctoral dissertation (Université Joseph-Fourier Gernoble).
103. Duan, C., Adam, V., Byrdin, M., Ridard, J., Kieffer-Jaquinod, S., Morlot, C., Arcizet, D., Demachy, I., and Bourgeois, D. (2013) Structural Evidence for a Two-Regime Photobleaching Mechanism in a Reversibly Switchable Fluorescent Protein. *Journal of the American Chemical Society* 135(42), 15841-15850.
104. Adam, V., Berardozzi, R., Byrdin, M., and Bourgeois, D. (2014) Phototransformable fluorescent proteins: Future challenges. *Current Opinion in Chemical Biology* 20, 92-102.
105. Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proceedings of the National Academy of Sciences* 99(20), 12651.
106. Mizuno, H., Mal, T. K., Tong, K. I., Ando, R., Furuta, T., Ikura, M., and Miyawaki, A. (2003) Photo-Induced Peptide Cleavage in the Green-to-Red Conversion of a Fluorescent Protein. *Molecular Cell* 12(4), 1051-1058.
107. Hayashi, I., Mizuno, H., Tong, K. I., Furuta, T., Tanaka, F., Yoshimura, M., Miyawaki, A., and Ikura, M. (2007) Crystallographic Evidence for Water-assisted Photo-induced Peptide Cleavage in the Stony Coral Fluorescent Protein Kaede. *Journal of Molecular Biology* 372(4), 918-926.
108. Wiedenmann, J., Ivanchenko, S., Oswald, F., Schmitt, F., Röcker, C., Salih, A., Spindler, K.-D., and Nienhaus, G. U. (2004) EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. *Proceedings of the National Academy of Sciences of the United States of America* 101(45), 15905.
109. Adam, V., Lelimousin, M., Boehme, S., Desfonds, G., Nienhaus, K., Field, M. J., Wiedenmann, J., McSweeney, S., Nienhaus, G. U., and Bourgeois, D. (2008) Structural characterization of IrisFP, an optical highlighter undergoing multiple photo-induced transformations. *Proceedings of the National Academy of Sciences of the United States of America* 105(47), 18343-18348.
110. Barondeau, D. P., Kassmann, C. J., Tainer, J. A., and Getzoff, E. D. (2006) Understanding GFP Posttranslational Chemistry: Structures of Designed Variants that Achieve Backbone Fragmentation, Hydrolysis, and Decarboxylation. *Journal of the American Chemical Society* 128(14), 4685-4693.
111. Perrin, F. (1926) Polarisation de la lumière de fluorescence. Vie moyenne des molécules dans l'etat excité. *Journal de Physique et le Radium* 7(12), 390-401.

112. Warren, S. C., Margineanu, A., Katan, M., Dunsby, C., and French, P. M. W. (2015) Homo-FRET Based Biosensors and Their Application to Multiplexed Imaging of Signalling Events in Live Cells. *International journal of molecular sciences* 16(7), 14695-14716.
113. Bader, A. N., Hofman, E. G., van Bergen En Henegouwen, P. M., and Gerritsen, H. C. (2007) Imaging of protein cluster sizes by means of confocal time-gated fluorescence anisotropy microscopy. *Optics Express* 15(11), 6934-6945.
114. Tanaka, F. and Mataga, N. (1979) Theory of time-dependent photo-selection in interacting fixed systems. *Photochemistry and Photobiology* 29(6), 1091-1097.
115. Squire, A., Verveer, P. J., Rocks, O., and Bastiaens, P. I. H. (2004) Red-edge anisotropy microscopy enables dynamic imaging of homo-FRET between green fluorescent proteins in cells. *Journal of Structural Biology* 147(1), 62-69.
116. Agranovich, V. M. G., M.D. (1978) Electronic Excitation Energy Transfer in Condensed Matter. . *Izdatel'stvo Nauka*, 384.
117. Einstein, A. (1905) Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik* 322(8), 549-560.
118. Lamla, T. and Erdmann, V. A. (2004) The Nano-tag, a streptavidin-binding peptide for the purification and detection of recombinant proteins. *Protein Expression and Purification* 33(1), 39-47.
119. Parks, T. D., Leuther, K. K., Howard, E. D., Johnston, S. A., and Dougherty, W. G. (1994) Release of Proteins and Peptides from Fusion Proteins Using a Recombinant Plant Virus Proteinase. *Analytical Biochemistry* 216(2), 413-417.
120. Chung, C. T., Niemela, S. L., and Miller, R. H. (1989) One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Sciences of the United States of America* 86(7), 2172-2175.
121. van den Ent, F. and Löwe, J. (2006) RF cloning: A restriction-free method for inserting target genes into plasmids. *Journal of Biochemical and Biophysical Methods* 67(1), 67-74.
122. Miller, H. (1987) Practical aspects of preparing phage and plasmid DNA: growth, maintenance, and storage of bacteria and bacteriophage. *Methods in Enzymology* 152, 145-170.
123. Miller, J. H. (1974) Experiments in Molecular Genetics. *The Quarterly Review of Biology* 49(2), 151-151.
124. Corrêa, F. and Farah, C. S. (2007) Different effects of trifluoroethanol and glycerol on the stability of tropomyosin helices and the head-to-tail complex. *Biophysical Journal* 92(7), 2463-2475.
125. Eberhard, D. and Ilian, J. (2000) *Thermodynamic Analysis of Cavity Creating Mutations in an Engineered Leucine Zipper and Energetics of Glycerol-Induced Coiled Coil Stabilization*.
126. Schägger, H. and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry* 166(2), 368-379.
127. Laemmli, U. K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227(5259), 680-685.
128. Jameson, D. (2014) *Introduction to Fluorescence* (Taylor & Francis, Boca Raton, FL, USA).

129. Dieguez, L., Darwish, N., Mir, M., Martinez, E., Moreno, M., and Samitier, J. (2009) Effect of the Refractive Index of Buffer Solutions in Evanescent Optical Biosensors. *Sensor Letters* 7, 851-855.
130. Zhang, X.-F., Zhang, J., and Liu, L. (2014) Fluorescence Properties of Twenty Fluorescein Derivatives: Lifetime, Quantum Yield, Absorption and Emission Spectra. *Journal of Fluorescence* 24, 819-826.
131. Nicoli, F., Barth, A., Bae, W., Neukirchinger, F., Crevenna, A. H., Lamb, D. C., and Liedl, T. (2017) Directional Photonic Wire Mediated by Homo-Forster Resonance Energy Transfer on a DNA Origami Platform. *ACS Nano* 11(11), 11264-11272.
132. Schrimpf, W., Barth, A., Hendrix, J., and Lamb, D. C. (2018) PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophysical Journal* 114(7), 1518-1528.
133. Rigler, R., Mets, Ü., Widengren, J., and Kask, P. (1993) Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. *European Biophysics Journal* 22(3), 169-175.
134. Galarneau, A., Primeau, M., Trudeau, L. E., and Michnick, S. W. (2002) Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. *Nature Biotechnology* 20(6), 619-622.
135. Hong, H., Blois, T. M., Cao, Z., and Bowie, J. U. (2010) Method to measure strong protein-protein interactions in lipid bilayers using a steric trap. *Proceedings of the National Academy of Sciences of the United States of America* 107(46), 19802-19807.
136. Schanzenbach, C., Schmidt, F. C., Breckner, P., Teese, M. G., and Langosch, D. (2017) Identifying ionic interactions within a membrane using BLaTM, a genetic tool to measure homo-and heterotypic transmembrane helix-helix interactions. *Scientific Reports* 7(43476).
137. Nenninger, A., Mastroianni, G., and Mullineaux, C. W. (2010) Size dependence of protein diffusion in the cytoplasm of Escherichia coli. *Journal of Bacteriology* 192(18), 4535-4540.
138. Pack, C., Saito, K., Tamura, M., and Kinjo, M. (2006) Microenvironment and effect of energy depletion in the nucleus analyzed by mobility of multiple oligomeric EGFPs. *Biophysical Journal* 91(10), 3921-3936.
139. Robia, S. L., Flohr, N. C., and Thomas, D. D. (2005) Phospholamban Pentamer Quaternary Conformation Determined by In-Gel Fluorescence Anisotropy. *Biochemistry* 44(11), 4302-4311.
140. Varma, R. and Mayor, S. (1998) GPI-anchored proteins are organized in submicron domains at the cell surface. *Nature* 394, 798.
141. Kozlowski, L. P. (2016) IPC – Isoelectric Point Calculator. *Biology Direct* 11(1), 55.
142. Oas, T. G., McIntosh, L. P., O'Shea, E. K., Dahlquist, F. W., and Kim, P. S. (1990) Secondary structure of a leucine zipper determined by nuclear magnetic resonance spectroscopy. *Biochemistry* 29(12), 2891-2894.
143. Goodman, E. M. and Kim, P. S. (1991) Periodicity of amide proton exchange rates in a coiled-coil leucine zipper peptide. *Biochemistry* 30(50), 11615-11620.
144. Harbury, P. B., Kim, P. S., and Alber, T. (1994) Crystal structure of an isoleucine-zipper trimer. *Nature* 371, 80.
145. Bradley, P., Misura, K. M. S., and Baker, D. (2005) Toward High-Resolution de Novo Structure Prediction for Small Proteins. *Science* 309(5742), 1868.

146. McIntosh-Smith, S., Price, J., Sessions, R. B., and Ibarra, A. A. (2015) High performance in silico virtual drug screening on many-core processors. *The international journal of high performance computing applications* 29(2), 119-134.
147. McIntosh-Smith, S., Wilson, T., Ibarra, A. Á., Crisp, J., and Sessions, R. B. (2011) Benchmarking Energy Efficiency, Power Costs and Carbon Emissions on Heterogeneous Systems. *The Computer Journal* 55(2), 192-205.
148. Cinelli, R. A., Ferrari, A., Pellegrini, V., Tyagi, M., Giacca, M., and Beltram, F. (2000) The enhanced green fluorescent protein as a tool for the analysis of protein dynamics and localization: local fluorescence study at the single-molecule level. *Photochemistry and Photobiology* 71(6), 771-776.
149. Schägger, H. (2006) Tricine-SDS-PAGE. *Nature protocols* 1(1), 16-22.
150. Banerjee, S., Schenkelberg, C. D., Jordan, T. B., Reimertz, J. M., Crone, E. E., Crone, D. E., and Bystroff, C. (2017) Mispacking and the Fitness Landscape of the Green Fluorescent Protein Chromophore Milieu. *Biochemistry* 56(5), 736-747.
151. Shinobu, A. and Agmon, N. (2009) Mapping Proton Wires in Proteins: Carbonic Anhydrase and GFP Chromophore Biosynthesis. *The Journal of Physical Chemistry A* 113(26), 7253-7266.
152. Battistutta, R., Negro, A., and Zanotti, G. (2000) Crystal structure and refolding properties of the mutant F99S/M153T/V163A of the green fluorescent protein. *Proteins* 41(4), 429-437.
153. Qi, Lei S., Larson, Matthew H., Gilbert, Luke A., Doudna, Jennifer A., Weissman, Jonathan S., Arkin, Adam P., and Lim, Wendell A. (2013) Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* 152(5), 1173-1183.
154. Tanenbaum, Marvin E., Gilbert, Luke A., Qi, Lei S., Weissman, Jonathan S., and Vale, Ronald D. (2014) A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell* 159(3), 635-646.
155. Tian, L., Hires, S. A., Mao, T., Huber, D., Chiappe, M. E., Chalasani, S. H., Petreanu, L., Akerboom, J., McKinney, S. A., Schreiter, E. R., Bargmann, C. I., Jayaraman, V., Svoboda, K., and Looger, L. L. (2009) Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* 6, 875.
156. Julius, A., Laur, L., Schanzenbach, C., and Langosch, D. (2017) BLaTM 2.0, a Genetic Tool Revealing Preferred Antiparallel Interaction of Transmembrane Helix 4 of the Dual-Topology Protein EmrE. *Journal of Molecular Biology* 429(11), 1630-1637.
157. Subramaniam, V., Hanley, Q. S., Clayton, A. H. A., and Jovin, T. M. (2003) [6] Photophysics of green and red fluorescent proteins: Implications for quantitative microscopy. *Methods in Enzymology*, (Academic Press), Vol 360, 178-201.
158. Bader, A. N., Hoetzl, S., Hofman, E. G., Voortman, J., van Bergen en Henegouwen, P. M. P., van Meer, G., and Gerritsen, H. C. (2011) Homo-FRET Imaging as a Tool to Quantify Protein and Lipid Clustering. *ChemPhysChem* 12(3), 475-483.
159. Volkmer, A., Subramaniam, V., Birch, D. J., and Jovin, T. M. (2000) One- and two-photon excited fluorescence lifetimes and anisotropy decays of green fluorescent proteins. *Biophysical Journal* 78(3), 1589-1598.
160. Eun Kim, Y., Kim, Y.-N., A Kim, J., Kim, H., and Jung, Y. (2015) Green fluorescent protein nanopolygons as monodisperse supramolecular assemblies of functional proteins with defined valency. *Nature Communications* 6, 7134.

161. Kudryavtsev, V., Sikor, M., Kalinin, S., Mokranjac, D., Seidel, C. A. M., and Lamb, D. C. (2012) Combining MFD and PIE for accurate single-pair förster resonance energy transfer measurements. *ChemPhysChem* 13(4), 1060-1078.
162. Thomson, A. R., Wood, C. W., Burton, A. J., Bartlett, G. J., Sessions, R. B., Brady, R. L., and Woolfson, D. N. (2014) Computational design of water-soluble α-helical barrels. *Science* 346(6208), 485.
163. Kim, H. K. and Kaang, B. K. (1998) Truncated green fluorescent protein mutants and their expression in Aplysia neurons. *Brain Research Bulletin* 47(1), 35-41.
164. Cabantous, S., Terwilliger, T. C., and Waldo, G. S. (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nature Biotechnology* 23(1), 102-107.
165. Foglieni, C., Papin, S., Salvadè, A., Afroz, T., Pinton, S., Pedrioli, G., Ulrich, G., Polymenidou, M., and Paganetti, P. (2017) Split GFP technologies to structurally characterize and quantify functional biomolecular interactions of FTD-related proteins. *Scientific Reports* 7(1), 14013-14013.
166. Li, X., Zhang, G., Ngo, N., Zhao, X., Kain, S. R., and Huang, C.-C. (1997) Deletions of the Aequorea victoria Green Fluorescent Protein Define the Minimal Domain Required for Fluorescence. *Journal of Biological Chemistry* 272(45), 28545-28549.
167. Fukuda, H., Arai, M., and Kuwajima, K. (2000) Folding of Green Fluorescent Protein and the Cycle3 Mutant. *Biochemistry* 39(39), 12025-12032.
168. Crameri, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. C. (1996) Improved Green Fluorescent Protein by Molecular Evolution Using DNA Shuffling. *Nature Biotechnology* 14(3), 315-319.
169. Ito, Y., Suzuki, M., and Husimi, Y. (1999) A Novel Mutant of Green Fluorescent Protein with Enhanced Sensitivity for Microanalysis at 488 nm Excitation. *Biochemical and Biophysical Research Communications* 264(2), 556-560.
170. Misteli, T., Gunjan, A., Hock, R., Bustin, M., and Brown, D. T. (2000) Dynamic binding of histone H1 to chromatin in living cells. *Nature* 408(6814), 877-881.
171. Snapp, E. L., Hegde, R. S., Francolini, M., Lombardo, F., Colombo, S., Pedrazzini, E., Borgese, N., and Lippincott-Schwartz, J. (2003) Formation of stacked ER cisternae by low affinity protein interactions. *The Journal of Cell Biology* 163(2), 257.
172. Leake, M. C., Chandler, J. H., Wadhams, G. H., Bai, F., Berry, R. M., and Armitage, J. P. (2006) Stoichiometry and turnover in single, functioning membrane protein complexes. *Nature* 443(7109), 355-358.
173. Shimi, T., Koujin, T., Segura-Totten, M., Wilson, K. L., Haraguchi, T., and Hiraoka, Y. (2004) Dynamic interaction between BAF and emerin revealed by FRAP, FLIP, and FRET analyses in living HeLa cells. *Journal of Structural Biology* 147(1), 31-41.
174. Neumann, J., Klein, N., Otzen, D. E., and Schneider, D. (2014) Folding energetics and oligomerization of polytopic α-helical transmembrane proteins. *Archives of Biochemistry and Biophysics* 564, 281-296.
175. Langosch, D., Brosig, B., Kolmar, H., and Fritz, H. J. (1996) Dimerisation of the glycophorin A transmembrane segment in membranes probed with the ToxR transcription activator. *Journal of Molecular Biology* 263(4), 525-530.
176. Russ, W. P. and Engelman, D. M. (2000) The GxxxG motif: A framework for transmembrane helix-helix association. *Journal of Molecular Biology* 296(3), 911-919.

177. Schneider, D. and Engelman, D. M. (2003) GALLEX, a measurement of heterologous association of transmembrane helices in a biological membrane. *Journal of Biological Chemistry* 278(5), 3105-3111.

8 List of Figures

Figure 1: Schematic structure of GCN4-p1, the most thoroughly studied coiled-coil peptide. 2
Figure 2: Schematic representation of homo-FRET between two fluorescent proteins and its effect on steady-state anisotropy.....4
Figure 3: Characterizing the oligomeric state of fluorophore labeled protein clusters from steady-state anisotropy rSS and fractional photobleaching. 6
Figure 4: Maturation of GFP consists of protein folding, forming of imidazolinone, and oxidation. 7
Figure 5: Sequence alignment of GFP [60], EGFP [84], and sfGFP [81]. 9
Figure 6: The proposed mechanism of the irreversible photobleaching of fluorescent proteins according to Adam et al. [101, 103]. 11
Figure 7: The proposed mechanism of GFP-Ala65-Ser66-Gly67 fragmentation according to Barondeau et al. [110].....12
Figure 8: The schematic structure, the plasmid map, and the amino acid sequence of the 1xGFP construct. 26
Figure 9: The band strength on denaturing PAGE is linearly proportional to the protein concentration.....38
Figure 10: Calibration curve of the separation column for size exclusion chromatography. 40
Figure 11: A blue laser (λ = 445 nm) with a power of 300 mW was used to photobleach the fluorophores in this study. 45
Figure 12: Standardization of steady-state fluorescence anisotropy increases the comparability of GFP concatemers when photobleached.....47
Figure 13: sfGFP concatemers as model proteins for sfGFP complexes interacting via homo-FRET.....53
Figure 14: The diffusion coefficient of sfGFP concatemers, recorded by Ganesh Agam, LMU Munich, decreases with increasing GFP domains which is in line with other reports and indicates a ellipsoid macromolecular structure.....55
Figure 15: The steady-state anisotropy of sfGFP concatemers in comparison with similar constructs from Vámosi et al. [90].....56

Figure 16: Steady-state anisotropy in combination with fractional photobleaching shows distinguishable, protein-specific behavior for monomeric, dimeric, and trimeric sfGFP concatemers. 61
Figure 17: The steady-state anisotropy of the monomeric GFP fusion proteins does not stay stationary but increases for a growing fraction of inactive fluorophores. 62
Figure 18: Modelling the steady-state anisotropy behavior for fractionally photobleached sfGFP concatemers with Yeow and Clayton's polynomial approach. 63
Figure 19: Parameter $fnon$ in dependence of N GFP moieties in sfGFP concatemers that interact via homo-FRET. 66
Figure 20: The amino acid sequences of all coiled-coil elements that were integrated in GFP fusion proteins in this study. 68
Figure 21: Biochemical analysis of non-covalently associated GFP fusion proteins. 70
Figure 22:The generation of the *de novo* coiled-coil ph3a. 71
Figure 23: The steady-state anisotropy of GFP-based coiled-coil fusion proteins. 73
Figure 24: The steady-state anisotropy of fractionally photobleached GFP-coiled-coil fusion proteins (left panels), compared to the anisotropy behavior of concatemer reference proteins (right panels). 78
Figure 25: Adjusting parameter a, eq. *(14)*, for EGFP-based fusion proteins. 80
Figure 26: Spectroscopic analysis of sfGFP after photobleaching reveals loss in emission intensity at 511 nm, but no increase in emission at longer wavelengths as seen for Kaede [105] or EosFP [108]. 81
Figure 27: The effect of irradiation power and bleaching time on the fraction of inactive fluorophores. 82
Figure 28: Blue laser light degrades the sfGFP containing fusion proteins. 84
Figure 29: The irradiation of sfGFP increases abundance of a number of protein fragments. 85
Figure 30: The irradiation of EGFP results in a defined number of fragments. 86
Figure 31: Immunoblot analysis of 1xGFP fragmentation to identify possible cleavage sites. 88
Figure 32: Immunoblot analysis of non-irradiated (left lane) and irradiated (right lane) EGFP fusion in order to evaluate the N-terminal integrity of protein fragments. 90

Figure 33: The asymmetric photobleaching preference for fluorophores with longer individual lifetimes leads to an increase in steady-state anisotropy of GFP upon photobleaching. 96
Figure 34: GFP is fragmented at the fluorophore upon high-intensity irradiation. 104
Figure 35: EGFP is cleaved at Phe99 upon high-intensity irradiation. 106
Figure 36: Further fragmentation of sfGFP upon high-intensity irradiation may take place proximally to Ser99/His181/Ala163/Thr153. 108

9 List of Tables

Table 1: Benchmark parameters for commonly used fluorescent proteins in molecular biology..... 10
Table 2: *Escherichia coli* strains used in this thesis..... 23
Table 3: DNA sequence of cassettes used in this thesis..... 25
Table 4: Primary antibodies used in this thesis..... 42
Table 5: Experimentally determined properties of 1xGFP..... 52
Table 6: Analysis of sfGFP concatemer time-resolved anisotropy data..... 58
Table 7: Calculating the number of concatenated sfGFP from fluorescence anisotropy, using the model by Runnels & Scarlata [46]..... 59
Table 8: Determination of *N* via Yeow and Clayton's polynomial approach, eq. *(11)* [57] 64
Table 9: Determination of *N* via Yeow and Clayton, eq. *(12)* [57], and the new theoretical model introduced in this study, eq. *(14)*..... 65
Table 10: Analysis of anisotropy data for GFP-coiled-coil fusion proteins..... 74
Table 11: Determining the oligomeric state *N* of assembled GFP-coiled-coil fusion proteins from fluorescence anisotropy, using the model by Runnels & Scarlata [46]..... 76
Table 12: Determining the oligomeric state *N* of assembled GFP-coiled-coil fusion proteins from fluorescence anisotropy in combination with fractional photobleaching..... 79
Table 13: Detailed analysis of GFP fragmentation patterns upon high-intensity irradiation..... 102

10 List of Abbreviations

*	Statistically significant (unpaired Student's *t*-test); $p < 0.05$
**	Statistically strongly significant (unpaired Student's *t*-test); $p < 0.01$
***	Statistically extremely significant (unpaired Student's *t*-test); $p < 0.001$
$\Delta r_{SS}(x)$	Steady-state anisotropy difference between irradiated samples and non-irradiated controls for an increased fraction of photobleached fluorophores x
$\int FEm$	Area under the fluorescence emission spectrum
μ	Mean in a normally distributed dataset
a	Empirically determined parameter to correctly describe the steady-state anisotropy data for photobleached fluorophores
A	Absorbance
aa	Amino acids
AEX	Anion exchange chromatography
Ala	Alanine (A)
Amp	Ampicillin
AP	Alkaline phosphatase
APS	Ammonium persulfate
Arg	Arginine (R)
Asn	Asparagine (N)
Asp	Aspartic acid (D)
BCIP	5-bromo-4-chloro-3-indolyl phosphate
bp	Base pairs
BSA	Bovine serum albumin
c	Concentration
CAnh	Bovine carbonic anhydrase
cfGFP	Cysteine-free version of GFP
Cys	Cysteine (C)
CytC	Equine cytochrome C
D	Diffusion coefficient
d	Path length
ddH_2O	Double-distilled water
dH_2O	Deionized water
DNA	Deoxyribonucleic acid
dNTP	Deoxyribose nucleoside triphosphate
E	Homo-FRET efficiency
EC	Molar attenuation / extinction coefficient
EDTA	Ethylenediaminetetraacetic acid

EGFP	Enhanced variant of GFP
ESI-TOF	Electrospray ionization – time of flight
f	Fraction of active fluorophores / fraction of labeled oligomers
FCS	Fluorescence correlation spectroscopy
FITC	Fluorescein isothiocyanate
FLIP	Fluorescence loss in photobleaching
f_{non}	Fraction of non-interacting fluorophores
FRAP	Fluorescence recovery after photobleaching
FRET	Förster resonance energy transfer
G	Calibration factor for anisotropy measurements
g	Fractional amplitude for eponential time-resolved anisotropy decays
GFP	Green fluorescent protein
Gln	Gltuamine (Q)
Glu	Glutamic acid (E)
Gly	Glycine (G)
GS enriched	Glyser enriched
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine (H)
Homo-FRET	FRET between identical fluorophores
I_0	Fluorescence intensity after irradiation at time point 0
I_i	Fluorescence intensity after irradiation at time point i
Ile	Isoleucine (I)
IMAC	Immobilized metal ion affinity chromatography
I_{par} or I_{per}	Fluorescence intensities measured at parallel or perpendicular orientations relative to the polarized excitation
IPTG	Isopropyl-β-D-thiogalactopyranosid
Kan	Kanamycin
Kan^R	Kanamycin resistance gene
k_B	Boltzmann's constant
k_{FRET}	Homotransfer rate
Laser	Light amplification by stimulated emission of radiation
LB	Lysogeny broth
Leu	Leucine (L)
LMU	Ludwig Maximilans Universität, Munich
Lys	Lysine (K)
m/z	Mass-to-charge ratio
MALDI-TOF	Matrix-assisted laser desorption/ionization – time of flight
Met	Methionine (M)

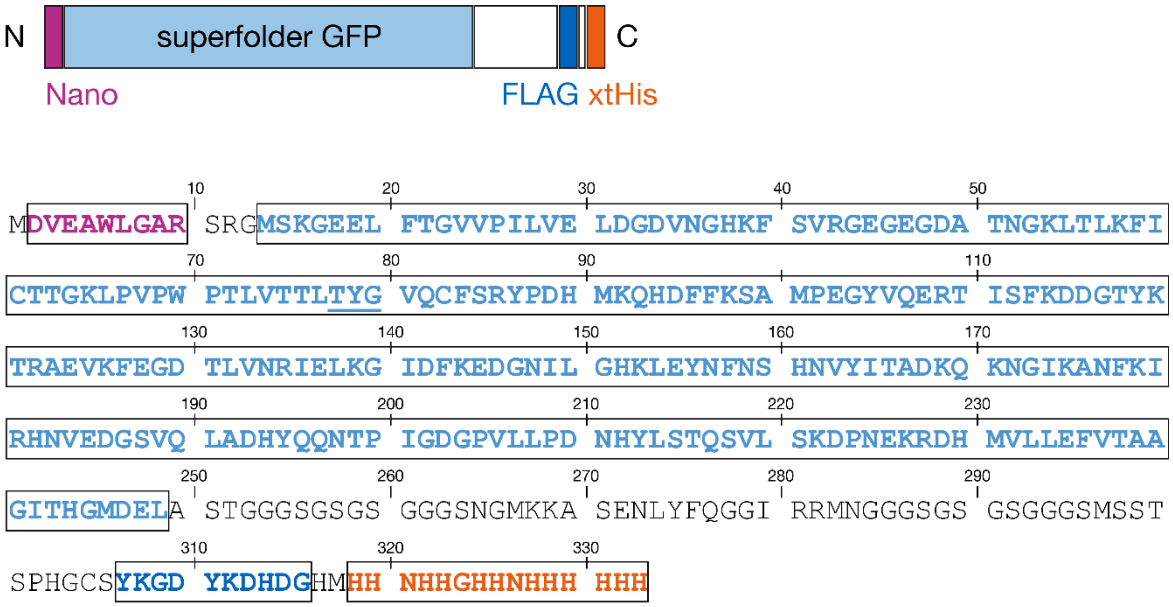
MFD	Multiparameter fluorescence detection
$mt_{½}$	Maturation halftime
MW	Molecular weight
MW_{non}	Molecular weight of non-repeated elements
MW_{rep}	Molecular weight of one repeated protein domain
N	Number of fluorophores in a cluster that interact via homo-FRET
n	Sample size
$Na_XH_XPO_4$	NaH_2PO_4 / Na_2HPO_4 buffer
NBT	Nitro blue tetrazolium
NMR	Nuclear magnetic resonance
ns	Statistically not significant (unpaired Student's *t*-test)
N_{SS}	N, calculated from steady-state anisotropy with parameters from time-resolved anisotropy
N_x	N, calculated from steady-state anisotropy for fractionally photobleached (x) samples
o	1st variable to describe the curvature of an exponential response
OD_{600}	Optical density at λ = 600 nm
p	Two-tailed probability value (unpaired Student's *t*-test)
p	2nd variable to describe the curvature of an exponential response
PAGE	Polyacrylamide gel electrophoresis
pa-GFP	Photoactivatable variant of GFP
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDB	Protein database catalogue number
Phe	Phenylalanine (F)
pI	Isoelectric point
PIE	Pulsed interleaved excitation
Pro	Proline (P)
q	Variable to describe the y-intercept of an exponential response
QY	Fluorescence quantum yield
r	Anisotropy
R/R_0	Fluorophore proximity: absolute distance divided by Förster radius
R_0	Förster radius / distance
r_0	Intrinsic / initial / limiting / fundamental anisotropy of a fluorophore
r_1	Steady-state anisotropy
r^2	Coefficient of determination
r_∞	Anisotropy limit for t approaching infinity
r_{ET}	Anisotropy of the indirectly excited fluorophores
R_M	Molecular radius

ROP	Regulator protein Rop responsible for DNA replication
rpm	Revolutions per minute
r_{SS}	Steady-state anisotropy
$r_{SS\ non}$	Steady-state anisotropy of non-irradiated controls
S_0	Ground state (Jablonski diagram)
S_1	Excited singlet state (Jablonski diagram)
SDS	Sodium dodecyl sulfate
SEC	Size-exclusion chromatography
Ser	Serine (S)
sfGFP	Superfolder variant of GFP
t	Time
T	Absolute temperature
$t_{½}$	Half-life
T_1	Excited triplet state (Jablonski diagram)
T_∞	Room temperature
TAE	Tris-acetate/EDTA
TBS-T	TBS containing polysorbate 20 “Tween”
TEMED	Tetramethylethylenediamine
TEV	Tobacco etch virus
Thr	Threonine (T)
TRIS	Tris(hydroxymethyl)aminomethane
Trp	Tryptophan (W)
TUM	Technical University of Munich
Tyr	Tyrosine (Y)
UV	Ultraviolet light
UV_{280}	Ultraviolet light at λ = 280 nm
v/v	Volume per volume
Val	Valine (V)
V_M	Molecular volume
w/v	Weight per volume
w_0	Radial diameter of the confocal volume
wt	Wild-type
x	Fraction of inactive fluorophores
xg	Times gravity
z	Charge
ε	Molar attenuation coefficient, “extinction coefficient”
ζ	Refractive index
η	Dynamic viscosity of the medium

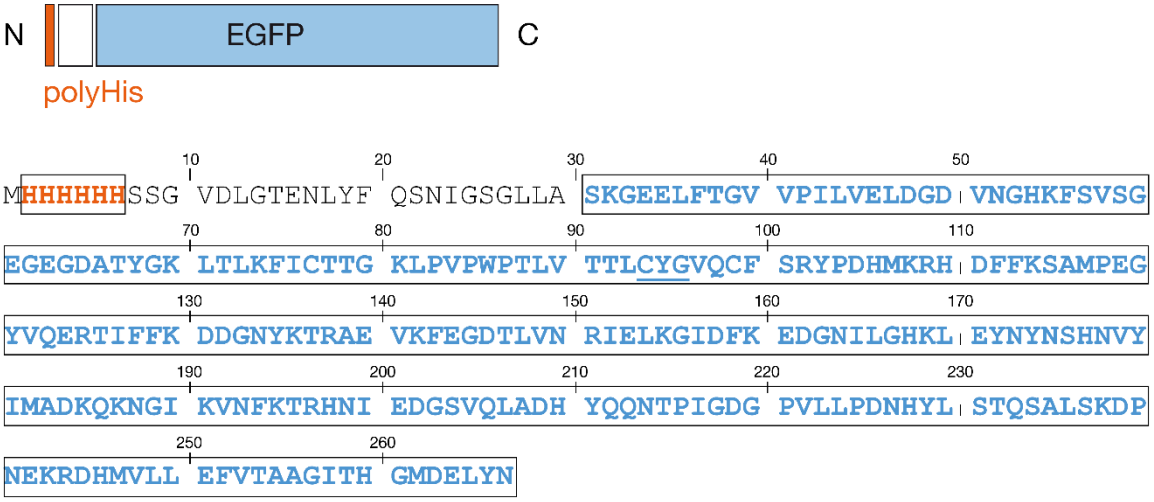
η_∞	Viscosity at room temperature
λ	Wavelength
$\lambda_{em.}$	Wavelength of fluorophore emission
$\lambda_{ex.}$	Wavelength of fluorophore excitation
τ	Fluorescence lifetime
τ_D	Diffusion time
ϕ	Rotational correlation time

11 Appendix

11.1 Amino acid sequences of proteins in this thesis



Appendix Fig. 1: Sequence of 1xGFP.



Appendix Fig. 2: Sequence of EGFP.

2xGFP – 5xGFP

MDVEAWLGARS**(**RGMSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKL
PVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLV
NRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFKIRHNVEDGSVQLADHYQQNTP
IGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMVLLEFVTAAGITHGMDELASTGGGSGSGSGGGSNGM
KKASENLYFQGGIRRMNGGGSGSGSGGGSMSSTRGMSKGEELFTGVVPILVELDGDVNGHKFSVRGE
GEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIS
FKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFK
IRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMVLLEFVTAAGITHGM
DELASTGGGSGSGSGGGSNGMKKASENLYFQGGIRRMNGGGSGSGSGGGSMSST**)** **repeated element**
SPHGCSYKGDYKDHDGHMHHNHHGHHNHHHHH

sfGFP-GCN4-p1

MGHHNHHGHHNHHHHHHPRGMSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFKIRHNVEDGSVQLADH
YQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMVLLEFVTAAGITHGMDELASTGGGSGSGSG
GGSNGMKKASAVSRMKQLEDKVEELLSKNYHLENEVARLKKLVGERGIRRMNGGGSGSGSGGGSMSS
TSPHGCSYKGDYKDHDGHMHHNHHGHHNHHHHH

sfGFP-GCN4-pII

MGHHNHHGHHNHHHHHHPRGMSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFKIRHNVEDGSVQLADH
YQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMVLLEFVTAAGITHGMDELASTGGGSGSGSG
GGSNGMKKASAVSRMKQIEDKIEEILSKIYHIENEIARIKKLIGERGIRRMNGGGSGSGSGGGSMSS
TSPHGCSYKGDYKDHDGHMHHNHHGHHNHHHHH

sfGFP-ph3a

MGHHNHHGHHNHHHHHHPRGMSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKF
EGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKANFKIRHNVEDGSVQLADH
YQQNTPIGDGPVLLPDNHYLSTQSVLSKDPNEKRDHMVLLEFVTAAGITHGMDELASTGGGSGSGSG
GGSNGMKKASAVSRIYKIEQKIYRIEQKIYRIEQKIYKIEQKIYGIRRMNGGGSGSGSGGGSMSSTS
PHGCSYKGDYKDHDGHMHHNHHGHHNHHHHH

EGFP-CC-Di

MHHHHHHSSGVDLGTENLYFQSNIGSGLLASKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDAT
YGKLTLKFICTTGKLPVPWPTLVTTLCYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGN
YKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKTRHNIE
DGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITGGGGGGIAA
LKQEIAALKQENAALKQEIAALKQE

EGFP-CC-Tet

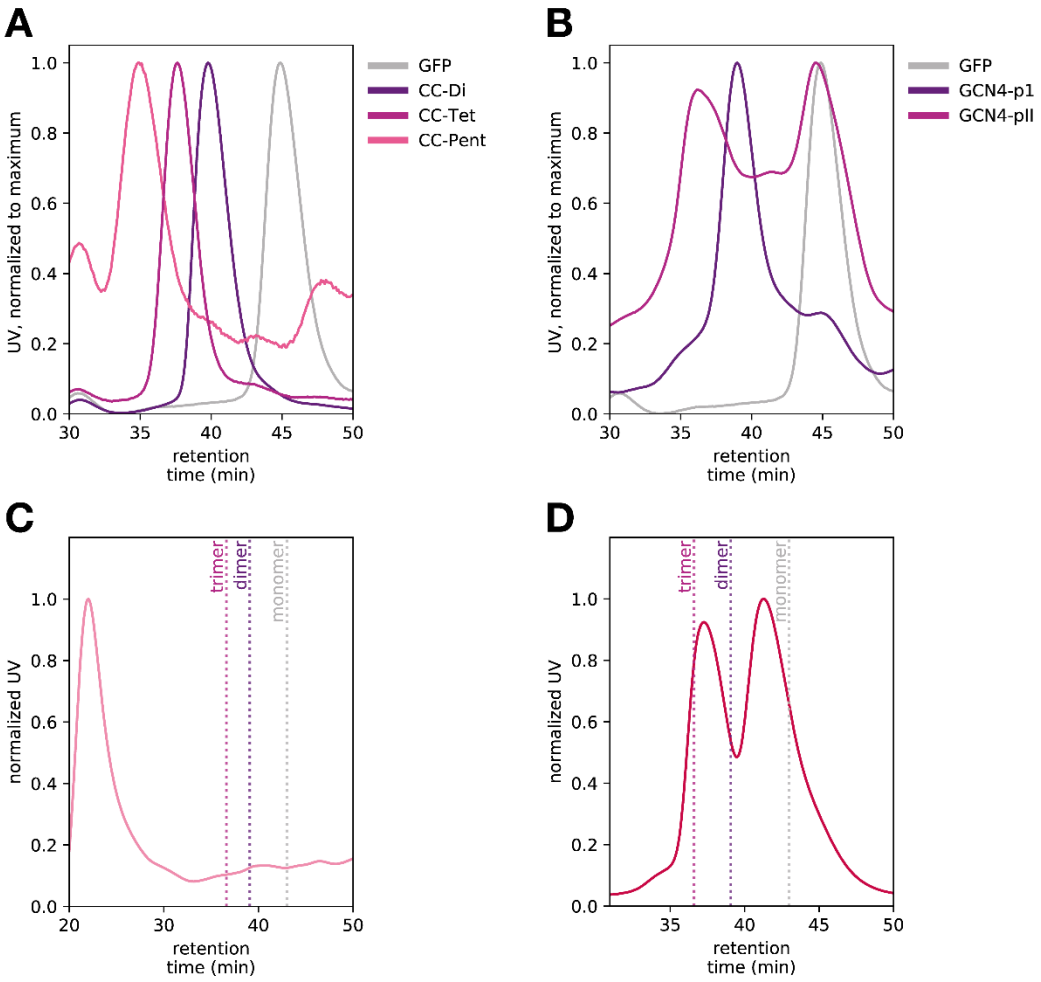
MHHHHHHSSGVDLGTENLYFQSNIGSGLLASKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDAT
YGKLTLKFICTTGKLPVPWPTLVTTLCYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGN
YKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKTRHNIE
DGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITGGGGGGELA
AIKQELAAIKQELAAIKQELAAIKQE

EGFP-CC-Pent

MHHHHHHSSGVDLGTENLYFQSNIGSGLLASKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDAT
YGKLTLKFICTTGKLPVPWPTLVTTLCYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGN
YKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKTRHNIE
DGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITGGGGGGGGG
GKIEQILQKIEKILQKIEQILQKIEQILQG

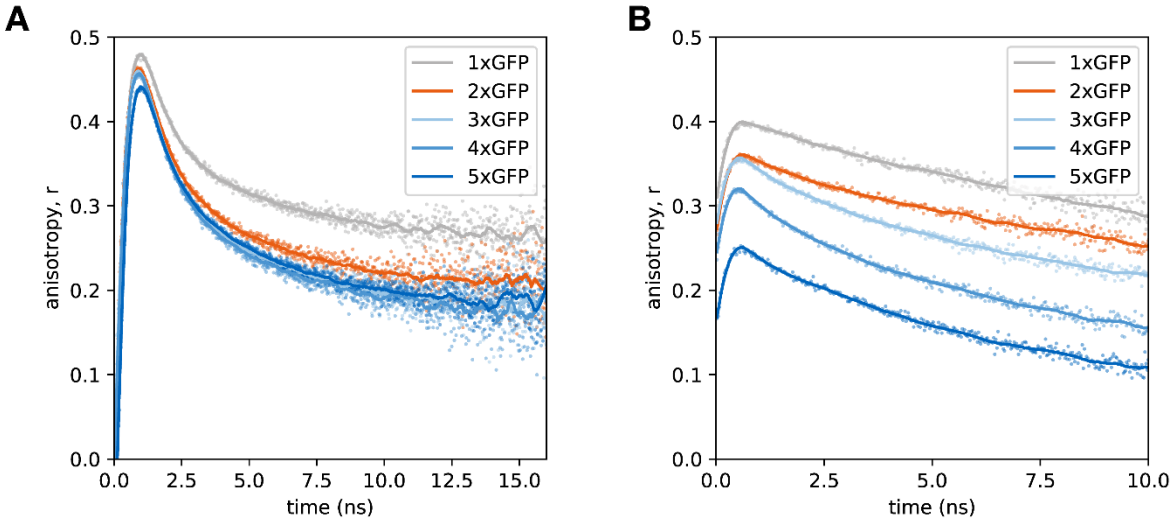
11.2 Raw data and data from collaboration partners

11.2.1 Size exclusion chromatography (SEC)

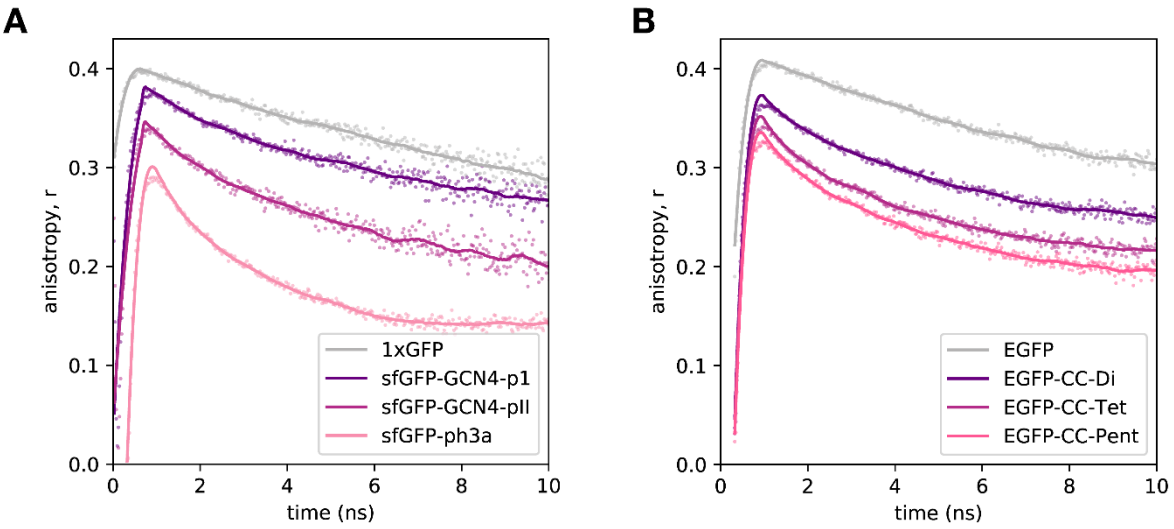


Appendix Fig. 3 Elution profiles of GFP-coiled-coil fusion proteins, analyzed via size exclusion chromatography (SEC). A) SEC analysis of EGFP-coiled-coils, originally introduced by Cristie-David et al. (2017) (loading concentrations: GFP, 106 µM; CC-Di, 142 µM; CC-Tet, 101 µM; CC-Pent, 21 µM) [18]. **B)** SEC analysis of GFP (106 µM) and sfGFP-coiled-coil fusion proteins GCN4-p1 (40 µM) and GCN4-pII (20 µM). For sfGFP-based fusion proteins, there was a strong monomer peak detectable. This indicates partial dissociation of the complexes when the sample is separated on the column. **C)** SEC analysis of sfGFP-ph3a (46 µM). **D)** SEC analysis of sfGFP-pAA (50 µM). The samples were loaded in a 25 mM HEPES (pH 7.5) buffer, containing 100 mM NaCl, 2 mM EDTA, and 30 % glycerol. The chromatography buffer was identical to the sample buffer but without any glycerol.

11.2.2 Time-resolved anisotropy

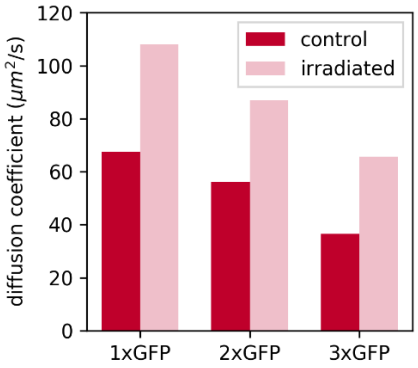


Appendix Fig. 4: Time-resolved fluorescence anisotropy of sfGFP concatemers detected with two different custom-built setups. The data was recorded by Ganesh Agam, LMU Munich. Purified sfGFP concatemers (1xGFP, 2xGFP, 3xGFP, 4xGFP, 5xGFP) were diluted to 2 µM in 30% (v/v) glycerol in fluorescent-free PBS. Their time-resolved anisotropy decays were analyzed with two custom-built setups. **A**) The first setup comprised a multiparameter fluorescence detection (MFD) pulsed interleaved excitation (PIE) equipment to monitor time-dependent fluorescence anisotropy decay [161]. It is usually used to detect small organic fluorophores. **B**) The second custom-built setup comprised a polarized and pulsed 468 nm laser for the excitation and a FLS1000 spectrometer equipped with photomultipliers for the detection of the emission. Raw data was smoothened using a Savitzky-Golay filter with a window size of 51 and a polynomial order of 3. The homotransfer rate k_{FRET} was calculated from the anisotropy decay from B), assuming the model of a hindered rotor, eq. *(7)*. [47].



Appendix Fig. 5: Time-resolved fluorescence anisotropy of GFP-based coiled-coil fusion proteins. The data was recorded Ganesh Agam, LMU Munich. Time-resolved anisotropy decay was detected via custom-built setup comprising a polarized and pulsed 468 nm laser for the excitation and a FLS1000 spectrometer equipped with photomultipliers for the detection of the emission. Raw data was smoothened using a Savitzky-Golay filter with a window size of 51 and a polynomial order of 3. The homotransfer rate k_{FRET} was calculated from the anisotropy decay using eq. *(7)*. In order to reduce molecular rotation, the proteins were kept in sample buffer containing 30% glycerol.

11.2.3 Fluorescence correlation spectroscopy (FCS)



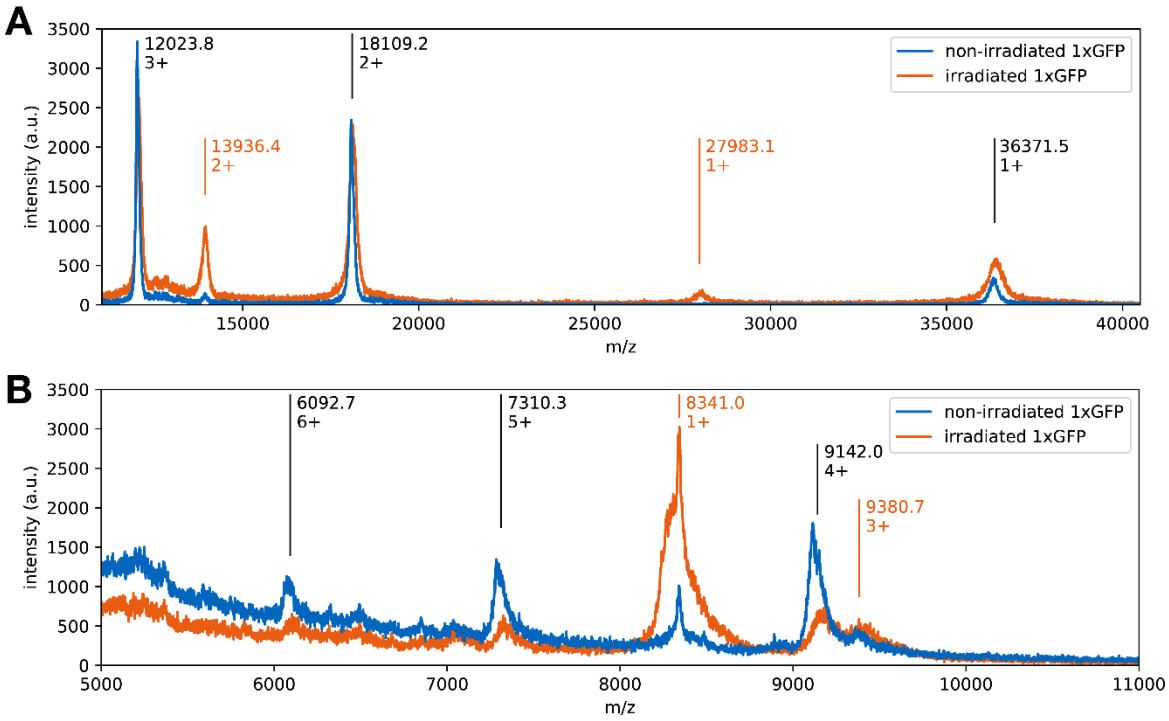
Appendix Fig. 6: The molecular diffusion of GFP fusion proteins is increased after irradiation with a high-intensity 445 nm laser. The data was recorded via fluorescence correlation spectroscopy (FCS) by Ganesh Agam, LMU Munich. According to the Einstein-Stokes model for spherical particles, eq. *(15)*, the molecular diffusion is a function of molecular size. Initially, no change in motility and molecular size was expected due to irradiation, as photobleaching should not structurally affect the molecule in theory. However, the diffusion coefficient of 1xGFP (36.4 kDa) increased from 67.5 µm^2/s for the unbleached sample to 108.0 µm^2/s after photobleaching (x > 0.9), indicating higher motility in aqueous solution. A similar acceleration of molecular movement was detected for 2xGFP (67.6 kDa) and 3xGFP (98.9 kDa). For GFP alone (26.9 kDa), diffusion coefficients of ~80 µm^2/s in aqueous solution have been published [90, 138]. Samples were diluted to 10 nM in fluorescence free PBS. Note that diffusion could only be measured for proteins that retained fluorescence (λ_{ex}: 488 nm; λ_{em}: 512 nm).

11.2.4 Mass Spectrometry

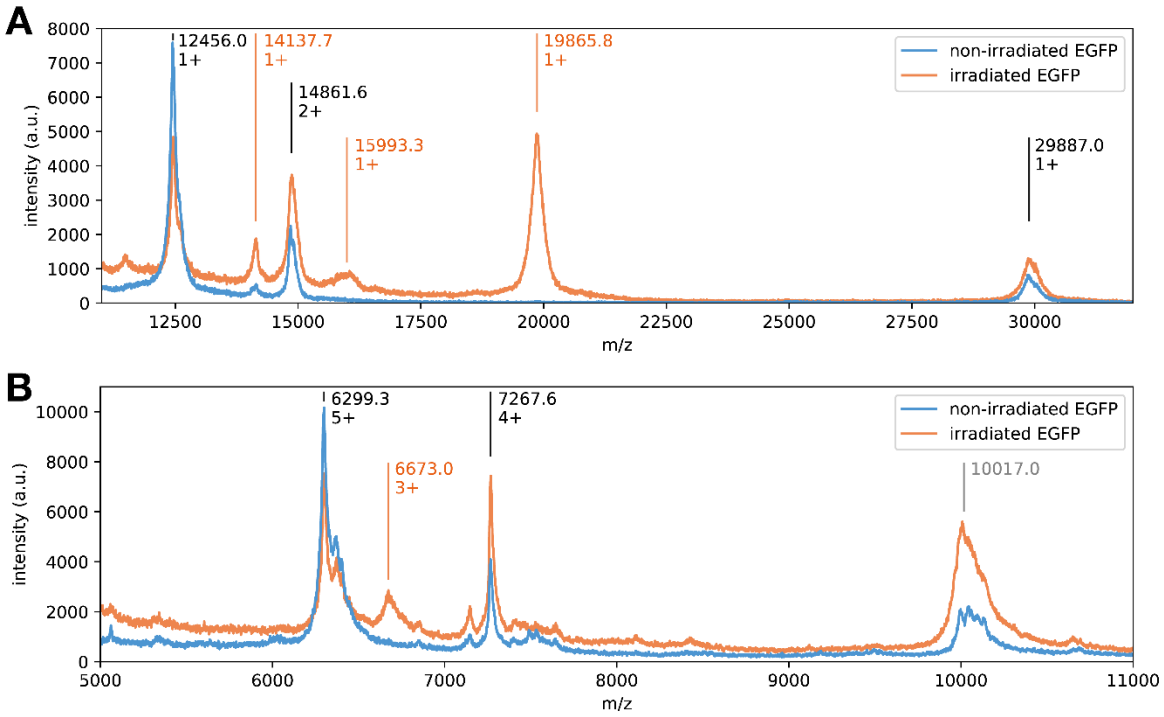
In order to characterize the light-induced fragmentation of GFP by finding residue specific or even higher resolved fragmentation sites, the exact masses of the fragments had to be determined. Therefore, MALDI-TOF and ESI-TOF mass spectrometry were used and the fragments of non-irradiated and irradiated 1xGFP (**Appendix Fig. 7**, **Appendix Fig. 10**) and EGFP (**Appendix Fig. 8**, **Appendix Fig. 10**) were compared.

MALDI-TOF measurements were executed by Dr. Martin Haslbeck at the Chair of Biotechnology at TUM, Garching.

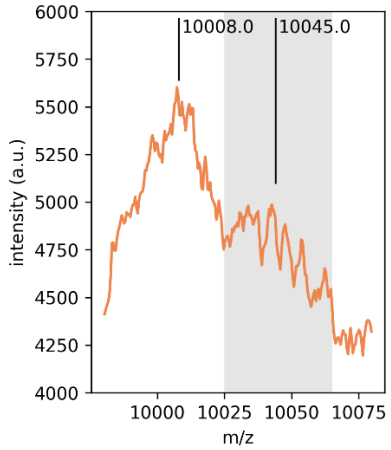
ESI-TOF measurements were executed by Walter Stelzer at our group, Chair of Biopolymer Chemistry at TUM, Freising.



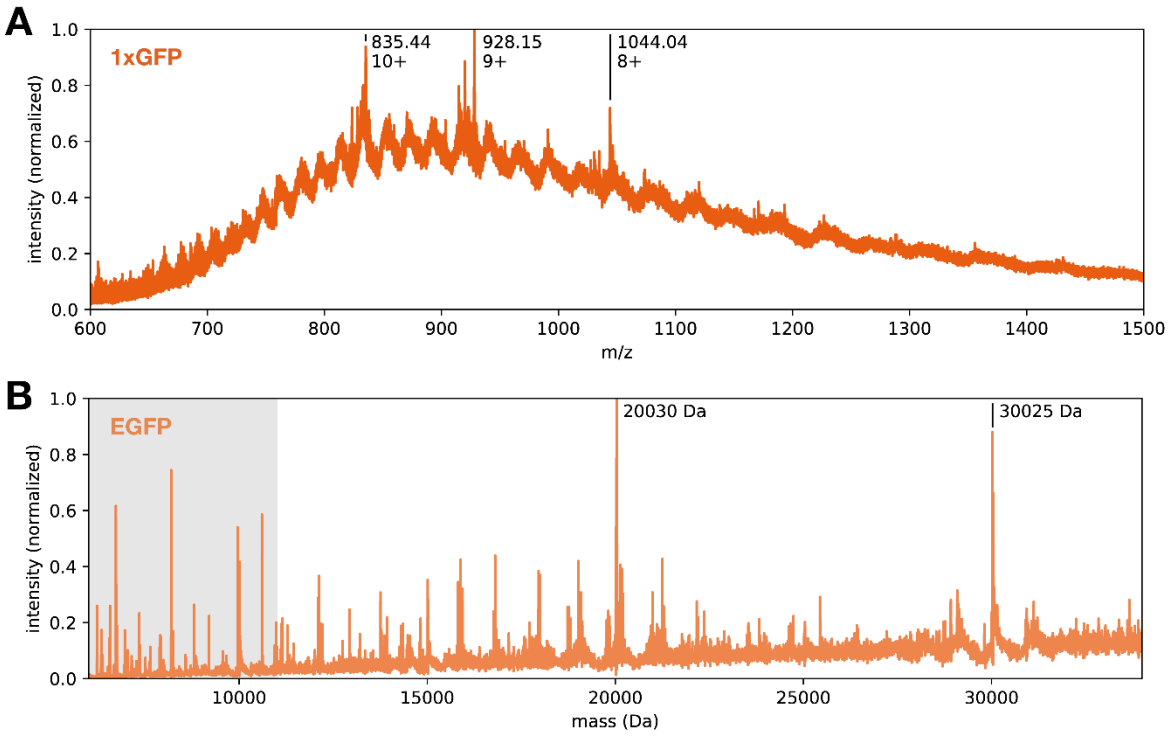
Appendix Fig. 7: MALDI-TOF spectrum of non-irradiated and irradiated the sfGFP fusion protein 1xGFP. The samples were prepared, spotted and measured in a collaboration, see Method section. There were two detector settings used, one for m/z above 11000 (**A**) and another for m/z below 11000 (**B**). Six peaks can be determined that correspond to the intact construct at differently charged states. This can be observed for the unbleached as well as for the bleached sample. With these peaks, an average mass of 36449 ± 109 Da (standard deviation) could be calculated for the intact protein which is similar to the mass calculated from the primary structure (36410 Da). Besides the peaks corresponding to multiply-charged 1xGFP, there are four maxima that significantly grew during the irradiation: peaks at m/z = 27983.1, 13936.4, 9380.7, and 8341.0.



Appendix Fig. 8: MALDI-TOF spectrum of non-irradiated and irradiated the EGFP fusion proteins. The samples were prepared, spotted and measured in a collaboration, see Method section. There were two detector settings used, one for m/z above 11000 (**A**) and another for m/z below 11000 (**B**). The average mass for the intact EGFP fusion protein was calculated from all identified peaks to be 30084 Da (marked black). The huge standard deviation of 879 Da, though, indicates that some of the peaks might be falsely assigned to the EGFP chimera. We assume that the strong peak is the equivalent of the 13 kDa contaminant, visible in the PAGE analysis (band e-C). Five maxima can only be detected for bleached EGFP: peaks at m/z = 19865.8, 15993.3, 14137.7, and 6673.0. The broad maximum at m/z = 10017.0 (grey) is very special as it might be a compound of (i) the 3+ charged original intact fusion protein, (ii) the 2+ charged 20 kDa fragment, or (iii) a complementary piece to 20 kDa fragment. In all three cases, one could expect a strong maximum at 10,000. However, only the latter two possibilities would result in a significantly increased peak for the irradiated sample.



Appendix Fig. 9: MALDI-TOF spectrum of irradiated EGFP at 10 kDa. A peak at 10008 Da and a clearly distinguishable shoulder with 10045 Da as a maximum is visible. The peak at 10008 Da putatively corresponds to the 3+ charged intact EGFP molecule and/or the 2+ charged 20 kDa fragment. The maximum at 10045 Da possibly represents the 1+ charged N-terminal fragment e-D.



Appendix Fig. 10: ESI-TOF spectrum of photobleached GFP fusion proteins. The sample was desalted and stepwise eluted from a tC18 Sep-Pak (Waters) column with 60% acetonitrile, 0.1% formic acid. It was measured with a Synapt G2 HDMS (Waters). **A)** Raw data spectrum of the photobleached sfGFP construct. A pattern of three prominent peaks can be distinguished from the relatively strong noise. From the three peaks, a mass of 8344.36 ± 0.03 Da could be calculated. **B)** Deconvoluted spectrum of the photobleached EGFP construct. Two major peaks can be distinguished, one at 200030 Da, and the other at 30025 Da. Below 10-12 kDa (grey area), the noise is increased and thus a potential peak at around 10 kDa cannot be clearly distinguished.

About the author

Philipp Johann Heckmeier

Born on March 17, 1989 in Ingolstadt, I grew up in Vohburg an der Donau, a small town on the banks of the river Danube. After my studies in Molecular Biotechnology at the Technical University of Munich (TUM), I worked as a research associate at the Chair of Biopolymer Chemistry at TUM.

Contact: philipp@heckmeier.net

Homepage: www.heckmeier.net

Publications, scientific presentations, and conferences

2017/02	61th Annual Meeting of the Biophysical Society in New Orleans, Louisiana, USA. Poster presentation.
2017	**Heckmeier PJ**, Teese MG, Langosch D. A Novel FRET Technique to Characterize the Oligomerization State of Protein-Protein Interactions. Biophysical Journal, Volume 112, Issue 3, Supplement 1, 452A-453A, February 03, 2017
2013	Smejkal B, Agrawal NJ, Helk B, Schulz H, Giffard M, Mechelke M, Ortner F, **Heckmeier P**, Trout BL, Hekmat D. Fast and scalable purification of a therapeutic full-length antibody based on process crystallization. Biotechnology and Bioengineering, Volume 110, Issue 9, 2452–2461, September 2013