JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

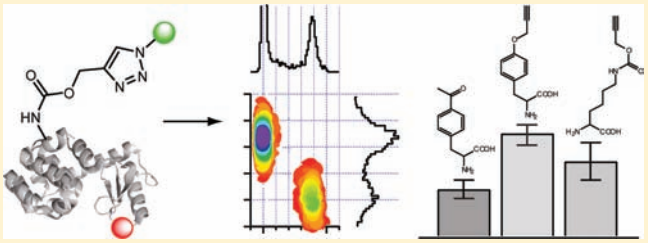
Click Strategies for Single-Molecule Protein Fluorescence

Sigrid Milles, Swati Tyagi, Niccolò Banterle, Christine Koehler, Virginia VanDelinder, Tilman Plass, Adrian P. Neal, and Edward A. Lemke*

EMBL, Structural and Computational Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany

Supporting Information

ABSTRACT: Single-molecule methods have matured into central tools for studies in biology. Foerster resonance energy transfer (FRET) techniques, in particular, have been widely applied to study biomolecular structure and dynamics. The major bottleneck for a facile and general application of these studies arises from the need to label biological samples site-specifically with suitable fluorescent dyes. In this work, we present an optimized strategy combining click chemistry and the genetic encoding of unnatural amino acids (UAAs) to overcome this limitation for proteins. We performed a systematic study with a variety of clickable UAAs and explored their potential for high-resolution single-molecule FRET (smFRET). We determined all parameters that are essential for successful single-molecule studies, such as accessibility of the probes, expression yield of proteins, and quantitative labeling. Our multiparameter fluorescence analysis allowed us to gain new insights into the effects and photophysical properties of fluorescent dyes linked to various UAAs for smFRET measurements. This led us to determine that, from the extended tool set that we now present, genetically encoding propargyllysine has major advantages for state-of-the-art measurements compared to other UAAs. Using this optimized system, we present a biocompatible one-step dual-labeling strategy of the regulatory protein RanBP3 with full labeling position freedom. Our technique allowed us then to determine that the region encompassing two FxFG repeat sequences adopts a disordered but collapsed state. RanBP3 serves here as a prototypical protein that, due to its multiple cysteines, size, and partially disordered structure, is not readily accessible to any of the typical structure determination techniques such as smFRET, NMR, and X-ray crystallography.



1. INTRODUCTION

Single-molecule (sm) methods are especially attractive and powerful for the study of complex biological and chemical systems and processes.[1] In contrast to conventional ensemble experiments, in single-molecule experiments, molecular properties are measured one molecule at a time and distributions in molecular behaviors are measured directly. This is particularly important when analyzing dynamic systems that cannot be easily synchronized. Single-molecule observation of such systems makes it possible to directly determine the connectivity between different states of complex biomolecular reactions rather than obtaining a view of the average behavior. Among single-molecule techniques, the ability to measure distances and dynamics using Foerster resonance energy transfer (FRET) has become a powerful tool in structural biology.[1−3] This impact is mainly due to the fact that FRET between a suitable donor dye and a suitable acceptor dye can be measured efficiently between 30 and 100 Å, and thus falls into size range of most molecular machines and their conformational changes. The application of single-molecule FRET (smFRET) technology has been boosted in recent decades by multiple technological advances that include optimized hardware, software, and data analysis.[1,2,4] However, site-specific labeling with a suitable dye pair, the key step for all smFRET studies, is still problematic for almost all proteins. Although typically not a major problem for DNA or RNA studies, for which many synthetic and semisynthetic strategies exist to economically generate labeled nucleic acid derivatives, site-specific labeling of proteins represents the major bottleneck to extend the applications of smFRET technology in biology and structural biology in particular.[5] In addition, smFRET necessitates exacting labeling methods compared to conventional ensemble fluorescence approaches, where low labeling yields can typically be compensated for by increasing the concentration of the sample. When possible, labeling of cysteine residues of proteins is a simple and efficient strategy, but is frequently not viable due to the natural abundance of this amino acid.

A plethora of chemical biology tools have been developed to site-specifically label proteins.[6−9] However, many of those are limited to the termini or require introduction of larger recognition elements, such as small peptide tags, into the protein. It was recently demonstrated that genetically encoding the unnatural amino acid (UAA) *p*-acetylphenylalanine (**1**) in *Escherichia coli* by means of an engineered AMBER suppressor tRNA/synthetase (tRNA/RS) system from *Methanocaldococcus jannaschii* can provide an efficient labeling site for hydroxylamine derivatives of single-molecule dyes.[5] However, this strategy has the major disadvantage that the oxime ligation optimally proceeds in acidic conditions, rendering it

Received: November 10, 2011
Published: February 22, 2012

   dx.doi.org/10.1021/ja210587q | *J. Am. Chem. Soc.* 2012, 134, 5187−5195

unattractive for use with many biological systems.[10] In contrast, click chemistry between an azide and a terminal alkyne is a widely used biocompatible reaction.[11] Its biggest advantages are that the reaction is fast, viable at physiological conditions, and independent of pH over a wide range. Here we present a strategy that utilizes genetically encoded clickable UAAs to make proteins accessible to single-molecule studies, with a focus on the specific demands necessary for smFRET, as a precise structural biology tool.

To this end, we performed a quantitative analysis of a range of clickable UAAs. We first analyzed the UAAs for highly important practical considerations that are critical to judge the potential for wide applicability for structural biology studies, including expression yield in different proteins using various expression systems, and commercial availability, and/or ease of synthesis. We systematically labeled a set of disordered and structured proteins with different incorporated UAAs and then performed smFRET multiparameter analysis of the samples, which differ from each other by only the choice of UAA. Intrinsically disordered proteins, which account for a large fraction of the eukaryotic proteome, were specifically included in this study, as they are typically more challenging proteins to express due to their lack of stable structure and high sensitivity toward proteases.[12] Furthermore, smFRET is one of the few techniques that can actually provide insights into IDP function,[3] which allows us to present a technology that goes clearly beyond demonstrating idealized model systems.

Highly precise measurements of FRET at the single-molecule level require that the dyes are attached via flexible, long linkers such that the dipole interaction, from which the FRET arises, averages isotropically. At the same time, long linkers result in a relatively large distance between the dye and the protein backbone. Correlating the measured smFRET distance with actual protein backbone distance is thus still a major problem,[3,13−15] and has been studied mainly on model systems (synthetic small DNA, RNA, or polyproline) that are also accessible to complementary computational methods. Therefore, even precise smFRET measurements are currently mainly applicable for relative comparison. Genetically encoding different UAAs and linking dyes to them also now generates an unexplored plethora of chemically and structurally diverse linkers. Using multiparameter smFRET technology, that is, recording fluorescence lifetimes, intensities, and anisotropies at the same time, we were able to access the individual consequences of the various UAA-dye linkages. Our results show that structure and size of the linker can affect the smFRET measurement in an unexpected way. Our data allow the individual effects to be distinguished and highlight the potential pitfalls when applying state-of-the-art protein engineering methods to smFRET experiments.

Overall, this work provides a resource on the advantages of different labeling systems for smFRET measurments, with a focus on biochemical considerations and effects on high-precision measurements in structural biology using smFRET. In fact, we were able to build upon the new strategy to show how a key regulatory protein of the nuclear transport pathway can now be studied with smFRET technology. Because of its amino acid composition, size, and complex structure, RanBP3 is typically not amenable to structural biology tools such as smFRET, NMR, or X-ray determination.

2. MATERIALS AND METHODS

2.1. Protein Expression and Purification. For quantitative expression tests, a GFP with an AMBER mutation (i.e., mutation to a TAG stop codon) at position 39 and an Maltose Binding Protein (MBP) with an AMBER mutation at position 11, both under the control of a pBAD promoter, were used following the procedure described by Plass et al.[16] For labeling studies, a single-cysteine-containing T4Lysozyme (T4L) mutant gene with an AMBER mutation at position 38 was used.[5] A single cysteine mutant of the hNup153 domain (residues 875−1475) with an AMBER mutation at position 938 was used analogously to the procedure described in Milles et al.[17] The intrinsically disordered protein (IDP) domain of the *Saccharomyces cerevisiae* Nup49,[18,19] residues 1−260, was expressed and purified as an intein-6His fusion construct. Similarly to T4L and hNup153, the yNup49FG construct contained a unique cysteine at position 250 and an AMBER TAG mutation at site 191 (Figure 1a).

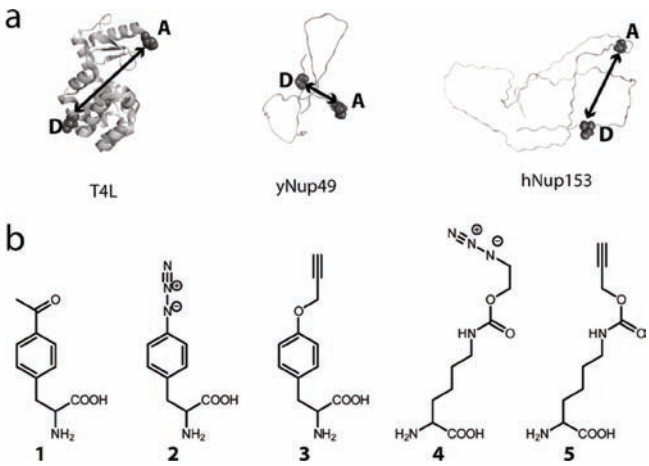


Figure 1. Proteins (a) and UAAs (b) used in this study for smFRET measurements.

T4L and yNup49 were cloned into pBad vectors and hNup153 into a pTXB3 vector. All vectors contained an ampicillin resistance gene and were cotransformed with plasmids harboring the tRNA/RS system for the respective UAA into *E. coli* Bl21(AI) (Invitrogen, Carlsbad, CA). For selective expression with **1**, **2** (*p*-azidophenylalanine), and **3** (*p*-(propargyloxy)-phenylanlanine) (Figure 1b), the recently reported optimized mutant variants of the pEvol system from the tyrosine synthetase of *M. jannaschii* were used, and termed pEvol-*pAcF* -*pAzF* and -*PrF*, for encoding compounds **1**, **2**, and **3**, respectively.[20]

For expression with **4** and **5**,[21,22] the recently introduced analogous pEvol system of the pyrrolysine tRNA/RS system from *Methanosarcina mazei* was used, termed pEvol PylRS.[16] Expression and purification of all proteins was performed as previously described for Nup153,[17] GFP,[16] and T4L, but more details and composition of standard purification buffers are given also in the Supporting Information Figures S1 and S2.[23,24] In short, expression was done in terrific broth (TB) medium at 37 °C. At OD_{600} = 0.2−0.5, 1 mM of the UAA was added. The cells were induced at OD_{600} = 0.4−1 with 0.02% arabinose and additional 1 mM IPTG in case of hNup153 and incubated until harvesting. Depending on the constructs, Ni and/or chitin beads were used for the purification in phosphate buffered saline (PBS, pH 7.4), and for the IDPs with 2 M urea following the manufacturer's protocols for Ni (Qiagen, Dusseldorf, Germany) and chitin (NEB, Ipswich, MA) beads, respectively. All purification buffers contained 0.2 mM *tris*(2-carboxyethyl)phosphine (TCEP) and 1 mM phenylmethanesulfonylfluoride (PMSF).

The gene encoding for RanBP3 was cloned into a pTXB1 vector as an intein-6His fusion construct and expressed in Bl21(AI) (Invitrogen, Carlsbad, CA) cotransformed with the pEvol PylRS to genetically encode **5** twice at the AMBER sites 181 and 294. Cultures were grown in TB medium at 37 °C and induced at OD_{600} = 0.4−0.6 with

arabinose and IPTG. Bacteria were harvested and purified under native conditions according to standard purification protocols for His-tagged proteins.

Expression efficiency using Amber suppression is not only UAA dependent, but also depends on the incorporation site as well as on the protein itself (see Figure 2 for comparison between relative GFP and MBP expressions). Typical expression yields for 1 L of expression cultures were >20 mg for **1**, ∼6 mg for **2**, ∼10 mg for **3**, ∼13 mg for **4**, and ∼12 mg **5** incorporated into GFP. For hNup153, which is very difficult to expresses due to its large size and intrinsically disordered structure, we typically obtained ∼1 mg (**1**), >0.5 mg (**3**), and >0.5 mg (**5**); for yNup49, 8 mg (**1**), >1 mg (**3**), and >2 mg (**5**); and for T4L, >20 mg (**1**), >2 mg (**3**), and >10 mg (**5**), respectively.

2.2. Site-Specific Dual Labeling. For quantitative comparison of UAA and dye linker effects on the quality of single-molecule measurements, all proteins (incorporating **1**, **3**, and **5**) were prepared in the same manner with Alexa 594 attached to equivalent cysteines using standard maleimide chemistry.[24] The ketone groups of proteins containing **1** were labeled under denaturing conditions (4 M guanidinium chloride) at pH 4 in sodiumacetate buffer with a hydroxylamine derivative of Alexa 488 as previously reported.[5,23,24]

Labeling of the alkyne functionality (proteins containing **3** and **5**, respectively) with commercially available azide derivatives of the Alexa 488 fluorophore was performed using copper-catalyzed alkyne−azide cycloaddition reaction (CuAAC). The exact labeling conditions, as well as kinetics and yields are given in detail in Supporting Information Figures S1 and S2.

RanBP3 was labeled using a 5 molar excess of Alexa 488 azide and 6 molar excess of Alexa 594 azide dyes in a single reaction pot. The reaction was quenched with 5 mM EDTA. All proteins were further purified on a Superdex-200. Concentrations were determined by UV−vis spectroscopy and/or BCA assay (Fisher Scientific, Rockford, IL) and labeling efficiencies were above 70% per labeling site (Supporting Information Figure S2).

2.3. Multiparameter Single-Molecule Fluorescence Spectroscopy. For smFRET experiments, the doubly labeled proteins were diluted to a final concentration of 50−200 pM in PBS, 2 mM dithiothreitol (DTT), pH 7.4, and measured on a custom built multiparameter (MP) spectrometer centered around an Olympus IX81 microscope (Hamburg, Germany) equipped with a high numerical water objective (60×, 1.2 NA) as previously described.[17] A laser diode (LDH 485, Picoquant, Berlin, Germany) and a white light laser (SuperK Extreme, NKT Photonics, Denmark) filtered through a z578/10 excitation filter (Chroma, Olching, Germany), which were pulsed alternately at 80 MHz total, were used to excite freely diffusing labeled proteins with linearly polarized light. Fluorescence originating from single molecules was first spatially filtered by a 100 μm pinhole and then split into parallel and perpendicular polarization directions before separation into green donor (D) and red acceptor (A) fluorescence light. D and A fluorescence was detected with micro photon counting devices (MPDs for the green (Picoquant, Berlin, Germany) and avalanche photo diodes for the red (Perkin-Elmer, Waltham, MA). Acquired data were subjected to multiparameter fluorescence analysis.[25,26] After burst recognition, a threshold of 100 photons was applied over I_D and I_A. Only bursts below a maximum brightness of 600 photons and a maximum length of 8 ms were considered for data evaluation. Foerster resonance energy transfer,

$$E_{FRET} = \frac{I_A}{I_D + I_A}$$

and fluorescence anisotropies of D,

$$r = \frac{I_{||} - GI_{\perp}}{I_{||} + 2GI_{\perp}}$$

were determined for every data set as described in Milles et al.[17] I_A and I_D describe the A and D intensities, respectively. $I_{||}$ and $I_{\perp}$ are parallel and perpendicular polarized light intensities and G is the G-factor of the instrument. τ, r, and E_{FRET} were summed up burst-wise and plotted in 2D τ versus E_{FRET} and r versus τ histograms. The populations in these histograms of several hundred events were fit with 2D Gaussian functions, which determine E_{FRET}, r and τ for the respective population. The Perrin curve, described by the equation

$$r = r_0/\left(1 + \frac{\tau}{\rho}\right)$$

with the fundamental anisotropy $r_0 = 0.4$ and the rotational correlation coefficient ρ was plotted onto the r versus τ histogram to visualize the behavior of these two parameters in an ideal FRET system. 2D stoichiometry histograms were plotted from E_{FRET} and the stoichiometry

$$S = \frac{I_A + I_D}{I_D + I_A + I_A^{dir}}$$

with I_A^{dir} being the signal from the directly excited acceptor molecules using orange excitation light.[27,28]

All data sets were measured in triplicates and corrected for leakage of D fluorescence into the A channel. Single-molecule data were processed and analyzed with a custom written code in IgorPro (Wavemetrics, Lake Oswego, OR). Error bars represent standard deviations unless noted otherwise.

2.4. Distance Determination from Multiparameter Single-Molecule Fluorescence Spectroscopy. Supporting Information Table S2 lists the detailed protocol of how the distances were calculated from MP smFRET data. The FRET distances were calculated from the fluorescence lifetimes as lifetime-based FRET is less affected by measurement artifacts than intensity-based FRET,[4] with the equation

$$E_{FRET}^{\tau} = 1 - \frac{\tau_{DA}}{\tau_{D(0)}}$$

where $\tau_{D(0)}$ and τ_{DA} are the fluorescence lifetimes of the donor only population and of the FRET-species. Since yNup49 and hNup153 are intrinsically disordered proteins and thus highly flexible, all their FRET efficiencies underlie fast fluctuations and can therefore be described according to the equation

$$\langle E \rangle = \int_0^{\infty} E(r)P(r)\,dr$$

where

$$\langle E \rangle = E_{FRET}^{\tau}$$

and the radial probability distribution of a Gaussian chain (see Supporting Information).[17,29−33]

$$P(r) = 4\pi r^2 \left(\frac{3}{2\pi\langle r^2 \rangle}\right)^{3/2} \exp\left(-\frac{3r^2}{2\langle r^2 \rangle}\right)$$

End-to-end distance values $R_E = (\langle r^2 \rangle)^{1/2}$ satisfying this equation were determined numerically using IgorPro (Wavemetrics, Lake Oswego, OR).

3. RESULTS

3.1. Choice of System and Unnatural Amino Acids. To date, more than 80 UAA have been genetically encoded into *E. coli*.[34] A few of those offer unique chemical handles that can be reacted with commercially available derivatives of single-molecule-suitable dyes. Commercial availability or easy accessibility are essential criteria, since only very few fluorophores actually fulfill the high demands of single-molecule studies, specifically high photostability, low blinking, and ideally single exponential lifetime decay kinetics. If not commercially available, they are typically only accessible to particularly skilled chemistry laboratories. For copper-catalyzed alkyne−azide click (CuAAC) chemistries, genetically encoded

alkyne and azide functionalities are most suitable. The *M. jannaschii* system that was used for incorporation of **1** has additionally been used for the genetic encoding of compounds **2** and **3**.[35,36] The structure of all those amino acids contains an aromatic, rather stiff phenylalanine derivative side chain. Aliphatic alkyne residues have been added to the repertoire of clickable UAAs by exploring the promiscuity of the natural AMBER suppressor of the pyrrolysine $tRNA^{Pyl}$/PylRS pair from *M. mazei* or *Methanosarcina bakeri*.[21] This allows the site-specific encoding of **4** and **5**.[22] The four UAAs, **2**–**5**, can be reacted in CuAAC reactions with the respective alkyne and azide single-molecule dyes. Furthermore, all four UAAs are readily available on a large scale, either commercially (**2**) or by simple, few step, efficient synthesis (see also Discussion below and Supporting Information).

3.2. Protein Engineering and Yield. Arguably, the most important practical discriminator for the suitability of genetically encoding UAAs for single-molecule research is the yield of protein expression. Although sample consumption in smFRET experiments is small, proteins first need to be prepared to high purity, which is complicated with low-expressing proteins. Furthermore, for most proteins other than classical model systems, expression is the limiting factor, and a drop in expression due to inefficient AMBER suppression often jeopardizes the experiment. Thus, knowing the best expressing system is of utmost importance for the success of the technology in biology studies. Thus, we first aimed to quantitatively compare the AMBER suppression efficiency of the various systems.

Genetically encoding UAAs requires cotransformation of the expression host *E. coli* with a plasmid harboring an orthogonal tRNA/RS pair along with a plasmid carrying the protein of interest, which typically contains an AMBER stop codon at the site of UAA incorporation. A particular concern is if the available systems to encode clickable UAAs maintain high incorporation fidelity compared to the established systems using **1** and if labeling ultimately yields a protein suitable for single-molecule observation in a straightforward way. We thus included **1** as the current "gold standard" in all our analysis.[5,17,23,24]

For determination of the expression yield, we used two reporter systems in order to ensure a representative readout. A GFP containing a unique AMBER TAG site at amino acid position 39 is a convenient check of protein expression by means of the fluorescence from either a cell culture suspension or purified GFP via a C-terminal 6His handle using a Ni-affinity resin.[16] As AMBER suppression heterogeneity for different proteins is possible, we also expressed maltose binding protein (MBP) with a TAG site in an artificial N-terminal linker. Figure 2 shows that for the *M. jannaschii* system, **3** expresses $GFP^{TAG\rightarrow UAA}$ better than **2** and MBP only marginally worse. While **2** is commercially available and has also been used before for protein engineering,[37,38] **3** can be easily synthesized (within about a day) in gram quantities using basic laboratory procedures convenient also for labs not specialized in chemistry (see Supporting Information). Most importantly, azides in general, and in particular aromatic azides,[39] are often reduction-sensitive, and are less stable in culture and during biochemical purification than their alkyne counterparts, which potentially impacts labeling yields. To experimentally verify the potential advantages or disadvantages of both chemistries, we expressed T4L encoding **2**–**5** side-by-side and also purified the proteins under standard biochemical conditions. Labeling was then

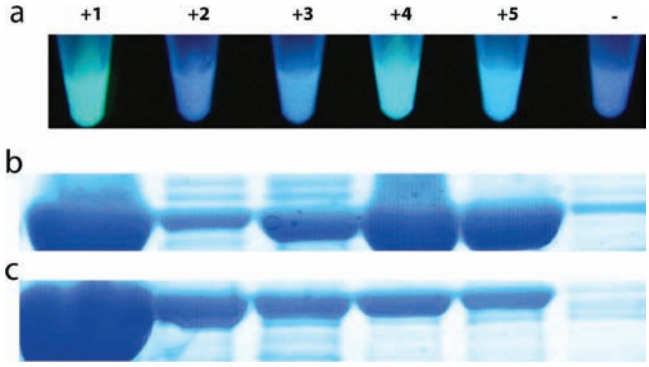


Figure 2. AMBER suppression expression test. (a) Fluorescence image of *E. coli* suspension expressing $GFP^{TAG\rightarrow UAA}$; (b) corresponding SDS–PAGE gel of purified $GFP^{TAG\rightarrow UAA}$; (c) SDS–PAGE gel of purified $MBP^{TAG\rightarrow UAA}$.

carried out at exactly the same protein concentrations using 5-fold excess of the complementary clickable dye derivative. Supporting Information Figure S1 confirms the major intuitive disadvantage of encoding azides compared to alkynes. Although all proteins behave well during labeling (no degradation), only **3** and **5** can be efficiently labeled under the tested standard conditions. It is therefore advisable to genetically encode the alkyne and perform the comparatively quick CuAAC labeling (see Supporting Information Figure S2 for labeling rates) in controlled clean buffer conditions with azido dyes rather than the other way around. Consequently, we discarded **2** and **4** from our further analysis and focused our comparative study on **1**, **3**, and **5**. Compound **5** is equally well synthetically accessible as compound **3** and can be synthesized in gram quantities within a day also by laboratories not specialized in chemistry (see Supporting Information).

3.3. Single-Molecule Multiparameter Spectroscopy. Single-molecule FRET is used in a variety of fields to study structure and dynamics of not only folded, but also disordered proteins.[1,3] In particular, the study of IDPs had recently gained much attention, since only very few techniques exist that can quantitatively provide insights into IDP plasticity. We selected three different proteins in order to extract general properties of site-specific labeling of various UAAs and to compare the different labeling strategies based on protein conformational studies. Therefore, in all experiments, the acceptor dye Alexa 594 (A) was attached via a single cysteine using standard maleimide chemistry serving as a reference in all measurements and the donor dye Alexa 488 (D) was reacted with the respective UAA. Since lifetime-based smFRET experiments provide considerably more information about the photophysical properties of D than A, this labeling strategy affords maximum insight into distinct UAA properties attached as a donor. Moreover, Alexa 488 is available in all the reactive forms required for this study and Alexa 488–Alexa 594 is a well-established (and probably most frequently used) dye pair for studying protein structure and dynamics.[17,32,40–43] Prior to labeling, the three proteins were reacted with maleimide A on their unique cysteines, and D was introduced under optimized experimental labeling conditions (see Materials and Methods). **1** was always reacted under denaturing conditions at pH 4 as in previously published procedures,[5,23,24] while the CuAAC for **3** and **5** was performed under physiological conditions (Figure 3a,b). See Supporting Information Figure S2 for more details on labeling conditions, kinetics, and yields.

dx.doi.org/10.1021/ja210587q | *J. Am. Chem. Soc.* 2012, 134, 5187–5195

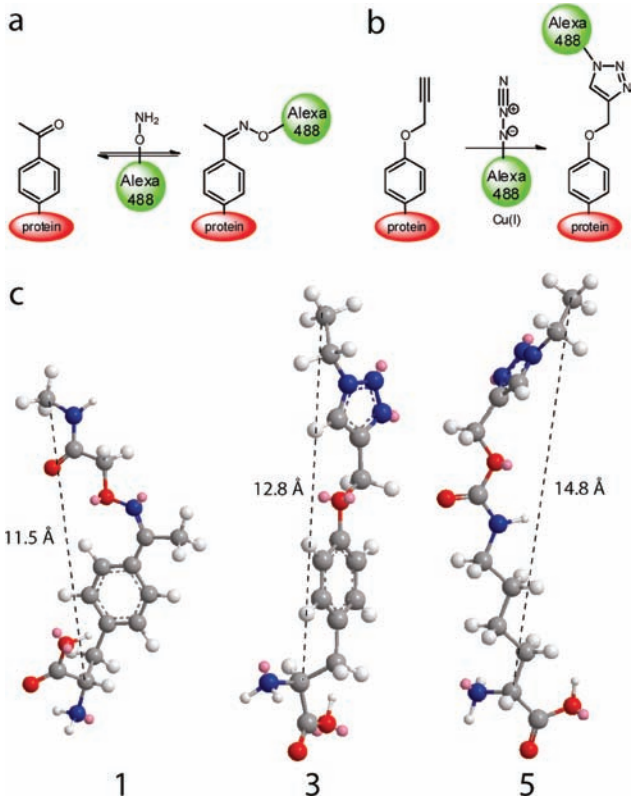


Figure 3. Scheme for labeling proteins via oxime ligation (a) and via CuAAC (b). (c) Energy minimized representation (using Chem-Bio3D, Cambridgesoft, MA) of chemical structures of UAA linked to dye derivative (truncated after linker C6 of Alexa 488). Distance drawn from Cα to linker-C of dye (see Supporting Information Figure S3 for more information).

Since commonly proteins are double-labeled using maleimide chemistry at two cysteines, little is known on how different dye−linker structure and compositions (Figure 3c, and Supporting Information Figure S5) can effect smFRET measurements. However, this is an important topic since FRET distance measurements can be affected by a perturbed local environment due to quenching or constrained mobility of the dyes.[2,4] Using MP smFRET, those effects can be measured. To assess these, the three proteins were subjected to multiparameter (MP) spectroscopy under identical conditions, which can detect lifetime anisotropy decays with high resolution.[17] The two-domain T4L serves as a model system for a folded protein.[5,23] However, amino acids in structured proteins are known to be exposed to a less homogeneous environment than those in rapidly fluctuating, polymer-like intrinsically disordered proteins (IDPs).[44] In IDPs, the dye is exposed to a more isotropic environment, so potential differences in a multiparameter single-molecule measurement due to changes in the UAA−dye architecture are more likely to be detected and interpreted correctly compared to a folded protein. We selected the yeast yNup49 intrinsically disordered protein[19] and the large human IDP hNup153.[45] These IDPs are well suited for our high-resolution study, as they lack tyrosine and tryptophan, which can dramatically alter dye photophysics and bias results.[46] For the IDPs, D and A were placed 40−70 amino acids away, so that FRET efficiencies were recorded under the most sensitive regime of the technique. Multi-parameter spectroscopy of freely diffusing single molecules was used to record fluorescence intensities I_A and I_D, fluorescence lifetimes τ, and anisotropies r of D for all three proteins.[4,17] Figure 4 shows exemplary experimental data for hNup153, and Supporting Information Figure S3 for yNup49 and Supporting Information Figure S4 for T4L.

An important feature when analyzing protein conformation is the brightness of the labeling dye. Alexa fluorophores, in particular Alexa 488, have high quantum yields and relatively long fluorescence lifetimes. These two connected properties might, however, change when the fluorophore is conjugated to a protein and are likely to also depend on the linker properties. In addition, fluorescence lifetimes and E_{FRET} are also directly related to each other. The longer the initial fluorescence lifetime of the donor (τ_0), the more sensitive the lifetime-based FRET efficiency is because small distance changes result in larger lifetime changes. The correlation between these two parameters is visible in the 2D τ versus E_{FRET} histograms, shown in Figure 4 for hNup153 (Supporting Information Figure S3 for yNup49 and Supporting Information Figure S4 for T4L). The two populations discernible in the 2D τ vs E_{FRET} histograms in Figure 4 correspond to the 0-peak population (inactive or absent acceptor, high lifetime) and to the FRET population (lower lifetime). The 0-peak and FRET populations are fitted with Gaussian functions from which E_{FRET} and τ can be obtained. Each protein displays a consistent trend with respect to UAA in both E_{FRET} (based on fluorescence intensity recordings) and τ (i.e., τ_0, τ_{DA}), as can be seen in the summary Figure 5. Alexa 488 has a shorter τ and higher E_{FRET} when attached to **3** than when attached to **1** or **5**, which behave rather similarly.

While both **1** and **3** are based on a rather stiff phenylalanine backbone, **5** is derived from the much more flexible lysine. On the other hand, the triazole ring originating from the labeling of **3** and **5** might render additional stiffness to the donor molecule conjugated via this reaction compared to the oxime linker between **1** and the dye. Fluorescence anisotropies report on differences in rotational freedom of dyes and might therefore reflect such variability of linker properties. Figure 4 shows the r versus τ plots (for the FRET population) of the **1**, **3**, and **5** conjugated D fluorescence of hNup153 (Supporting Information Figure S3 for yNup49 and Supporting Information Figure S4 for T4L). Since r is an observable that depends on the observation time of a single molecule after a laser excitation pulse, it is directly related to the fluorescence lifetime through the Perrin equation. ρ, on the contrary, is an intrinsic property of the fluorophore.[4] To draw reliable conclusions about ρ, both the 0-peak population and the FRET population have to lie on the same Perrin curve. This is true for hNup153 (Figure 4) and for yNup49 (Supporting Information Figure S3). T4L (Supporting Information Figure S4) shows slight deviations from the Perrin curve under all labeling conditions (with all UAAs). This observation indicates possible interference of the folded structure with dye flexibility, and not necessarily UAA properties. In general, care should be taken when evaluating FRET data from proteins that do not follow the prediction of the Perrin curve: elevated rotational correlation times can also affect the assumption that the rotational averaging factor κ^2 = 2/3, which can affect the accuracy of distance estimates.[14] To avoid any bias in this study, Table 1 and Supporting Information Table S2 list the actual distances that were extracted only for the IDPs. Figure 5d summarizes the rotational correlation coefficients of D labeled to proteins. IDPs labeled with UAA **3** show highest ρ (i.e., lowest rotational

dx.doi.org/10.1021/ja210587q | J. Am. Chem. Soc. 2012, 134, 5187−5195

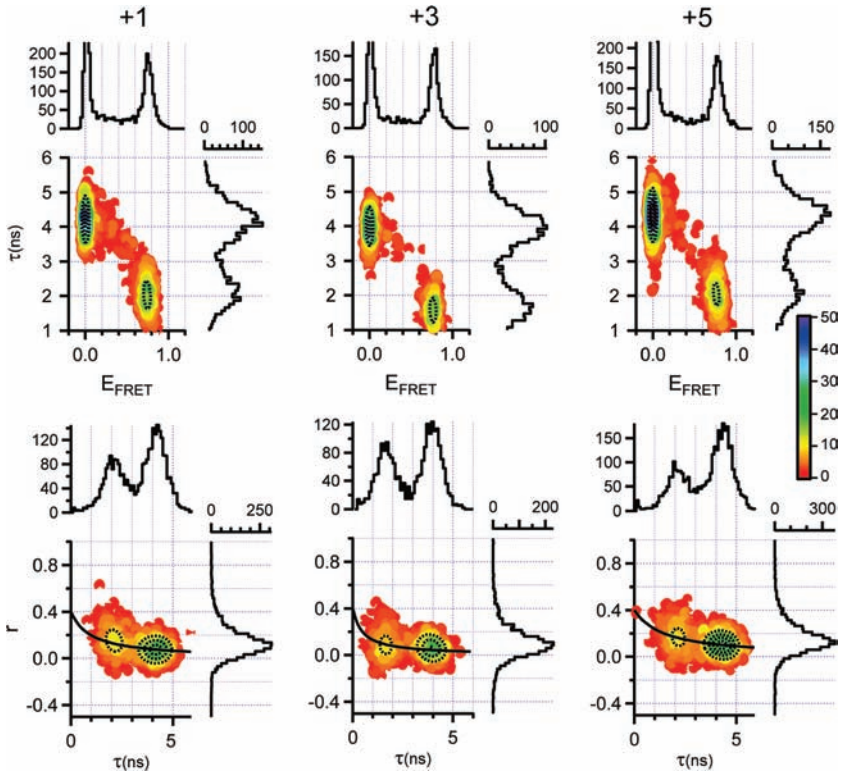


Figure 4. MP smFRET data for hNup153 with D labeled via **1**, **3** and **5**, respectively. In each column, burst integrated fluorescence lifetime (BIFL) analysis is in the upper panel and corresponding anisotropy r vs τ in the lower panel. The 2D plots are gray scale coded for frequency of occurrence, while the top and right histograms are maximum projections of the data along the vertical τ axis and along the horizontal E_{FRET} axis. In the lower row, the black line shows the expected trend according to the Perrin equation. Dotted lines outline fit results.

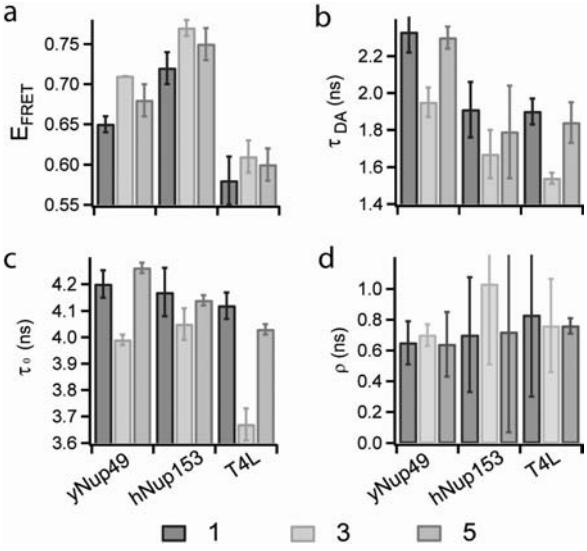


Figure 5. Summary of data measured for all UAA containing proteins. (a) E_{FRET} (from intensity measurement); (b) τ_{DA}; (c) τ_0; and (d) ρ.

freedom), while dyes attached to **1** and **5** indicate faster rotation (low ρ), following the same pattern as for the other two parameters (E_{FRET} and τ_{DA}) reported in Figure 5, with **3** behaving different than the other two.

Since both, high fluorescence lifetimes of D and rotational freedom of the D and A are important for high-resolution single-molecule FRET, UAAs **1** and **5** seem the ideal handles on which to attach fluorophores for multiparameter fluorescence studies. In addition, compared to **1**, **5** can also be labeled under physiological conditions and is therefore the choice for general single-molecule applications.

Table 1. Linker Distances and Measured FRET Distances[a]

	R_E (Å)	ΔR_E (Å)	linker (Å)	Δlinker (Å)
yNup49$^{TAG→1}$	63.8	0.0	11.5	0.0
yNup49$^{TAG→3}$	57.8	−6.0	12.8	1.3
yNup49$^{TAG→5}$	62.4	−1.4	14.8	3.3
hNup153$^{TAG→1}$	57.6	0.0	11.5	0.0
hNup153$^{TAG→3}$	53.2	−4.4	12.8	1.3
hNup153$^{TAG→5}$	55.6	−2.0	14.8	3.3

[a] ΔR_E and Δlinker correspond to measured value of UAA **1**.

3.3. FRET-Labeling of the Large and Cysteine-Rich RanBP3. Having now an optimized system in hand, we aimed to apply this technology to a common problem for smFRET in biology: the site-specific dual labeling with full choice of labeling sites (i.e., not just the termini) of large and cysteine-rich proteins that require mild, physiological labeling conditions, ideally in one step. RanBP3, a key player in regulation of nucleocytoplasmic transport, is such a protein; since partially disordered, no structural information can be obtained from X-ray crystallography, and due to its size (>50 kDa), it is essentially inaccessible by NMR.[47,48] While these limitations are not a concern for smFRET measurements in general, the multiple cysteines (>5), that could be of functional relevance, render RanBP3 also inaccessible to standard smFRET approaches. As a consequence, only a small region of RanBP3 has been amenable to crystallographic analysis but the overall architecture of RanBP3 is still elusive.[48] RanBP3, however, can undergo a variety of tasks, and regulate nuclear

dx.doi.org/10.1021/ja210587q | J. Am. Chem. Soc. 2012, 134, 5187−5195

export of various cargos, including mRNA, in the form of a tetrameric complex with RanGTP, the cargo and major export factor Crm1.[47−49] The part of RanBP3 that is not accessible to X-ray crystallography is heterogeneous in sequence and contains an unusual nuclear localization sequence (NLS) (Figure 6). Particularly interesting is also the existence of two

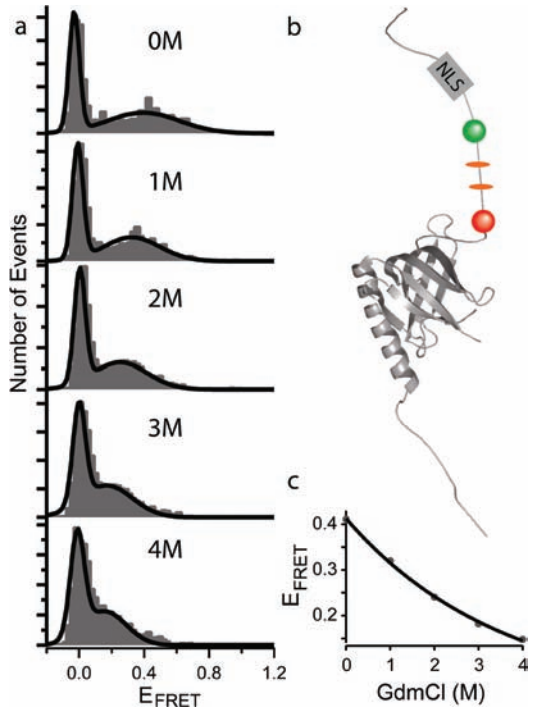


Figure 6. (a) GdmCl titration of RanBP3 clearly shows a gradual decrease in E_{FRET} when transferred to denaturing condition. (b) Cartoon of overall RanBP3 architecture showing disordered N- and C-terminal domains, as well as the structured Ran binding domain (RBD) and labeling sites. The position of the nuclear localization sequence (NLS) is depicted in gray and FxFG-repeats in orange. (c) E_{FRET} obtained for RanBP3 under different GdmCl concentrations can be described by an exponential fit (black line) and decreases from ∼0.4 in PBS to ∼0.15 in 4 M GdmCl.

FxFG repeats within this region. These FG repeats are usually highly enriched in nucleoporins that line the inner channel of the nuclear pore complex and constitute the permeability barrier. The function of these FG repeats in a protein that itself can shuttle the pore is yet completely unknown, as well as the structure of this region. To demonstrate the utility of our technology, we introduced TAG mutations at AA site 181 and 294 and expressed RanBP3 in the appropriate *E. coli* strain to twice genetically encode **5**. A single reaction with a mixture of Alexa 488 and Alexa 594 azide derivative dyes yielded efficient and site-specific labeling under physiological conditions. The doubly labeled RanBP3 gives a FRET around E_{FRET} = 0.4 under physiological conditions (Figure 6a). The FRET peak moves to E_{FRET} = 0.15 in the presence of increasing concentrations of guanidinium chloride (GdmCl). Thus, despite predicted to be an intrinsically disordered region, the region can be further "unfolded" through denaturation. Notably, the corresponding E_{FRET} and GdmCl-concentrations can be fit with an exponential decay, which indicates the absence of stable secondary structure, but points toward the existence of transient secondary structure.[30] Taken together, the occurrence of a single E_{FRET} population that can be gradually denatured is evidence of the existence of a native collapsed state, which is distinct from a fully unfolded or folded state.[30,32] In such a state, intramolecular interactions are favored versus a fully unfolded population and could function in presenting the FxFG motives for complex formation as part of the still unknown molecular transport mechanism. This result demonstrates that click labeling of two site-specifically encoded **5** is a reliable and efficient tool to label even complex proteins in order to study protein folding and conformational changes. Further mutagenic screening following the same procedure will now allow us to fully map the disordered domains and the conformational motions of RanBP3 in its various active states and during active transport. This will make the disorder-mediated regulatory mechanism of this proteins and how it can facilitate nuclear transport by binding to various proteins such as Importin-α, Importin-β, Crm1, and RanGTP finally amenable to structural biology.

4. DISCUSSION

Encoding UAAs for site-specific protein labeling is one of the rare tools for accessing structurally complex and demanding proteins in single-molecule FRET studies. However, not all potentially available UAAs fulfill the requirements for future high-resolution conformational studies. As also summarized in Supporting Information Table S3, in this study we started out from five known UAAs that can in principle be used for biochemical labeling. Because of the reduction sensitivity of azides during standard protein expression and purification of proteins, we quickly discarded the further use of azides **2** and **4** (Supporting Information Figure S1). This is a particularly noteworthy result, since **2**, likely owing to its commercial availability, has been used frequently to engineer proteins in applications where, for example, inefficient labeling was not detrimental.[38]

Apart from a good expression yield and the possibility of labeling under physiological conditions, the way dyes are attached to proteins can have a major impact on the interpretation of sm data. A possible explanation for the different ρ measured with the various UAAs is that **1** and **5** contain only one rigid aromatic ring (benzene ring from the phenylalanine and triazole from the click reaction, respectively), while **3** contains two (both, benzene and triazole rings). smFRET requires the isotropic averaging of dye orientation during the lifetime, so that the orientational averaging factor κ^2 can be assumed to be 2/3. Any deviation from this would jeopardize accurate distance information.[50] This is particularly important, since no easy way exists to directly measure κ^2 and thus the assumption needs to be valid in order not to compromise the experiment. In principle, all the tested UAAs fulfill these requirements. Since the κ^2 criterion requires high flexibility and thus long linkers, a direct correlation of measured FRET distance with actual residue distance within a protein labeled via very short (no) linkers is not possible. A conversion between these two distances is thus necessary. For simple model systems that are also accessible to computational complementation using accessible volume or molecular dynamics calculations, such as labeled polyproline, DNA, and RNA, it has been shown that very high, sometimes even effectively atomic resolution can be achieved.[13−15,29] As these procedures cannot be easily extended to large proteins with unknown structure, an approximation of dye−linker contributions in protein structure determination by smFRET is nontrivial. We thus calculated for the highly flexible intrinsically

dx.doi.org/10.1021/ja210587q | J. Am. Chem. Soc. 2012, 134, 5187−5195

disordered proteins yNup49 and hNup153 the dye-to-dye distance, also termed the end-to-distance (R_E). To accurately take the flexibility of the molecule into account, the basic Foerster equation cannot be used but needs to be expanded to take the fast fluctuating character of the protein into account (see Materials and Methods and Supporting Information Table S2 for more information).[17,29−33] In Table 1, we summarize the obtained R_E distances and also plot the relative difference to the measurement with **1**, which serves as a reference in this study. Our data (Figure 5a, Table 1, and Supporting Information Table S2) already show that, while the energy minimized structure of for example **5** is >3 Å longer than for **1**, proteins that were labeled via **1** show consistently longer distances compared to when they were labeled via **5**. As another example, a measurement of yNup49$^{TAG\rightarrow1}$ yields a distance of 63.8 Å. It would be common to assume that due to the approximately 1.3 Å-longer linker of **3** compared to **1**, a measurment of yNup49$^{TAG\rightarrow3}$ should yield a distance of 64.1 Å. In fact, however, a distance of 57.8 Å is measured and thus the total error could be more than 7 Å. Although we can measure very small but consistent differences between the proteins labeled on the different UAAs (Figure 5), accurate absolute protein backbone distance measurements still need to be developed in general. However, our data permit relative, but precise, comparison between different experiments and different UAAs, and show that **1**, **3**, and **5** are suitable for FRET distance measurements. Furthermore, our data show how the different linkers that "counterintuitively" affect the FRET distance measurement can be used as a reference to relate measurements with different UAAs to another. Compared to **3**, **1** and **5** obey higher lifetimes and lower anisotropy and thus clearly depict advantages with respect to higher dynamic range and higher measurement precision for smFRET measurements (Figure 5, Supporting Information Table S1).

5. CONCLUSION

We have shown in this work that genetically encoded clickable UAAs can be used for an efficient and general strategy to finally make proteins accessible to fluorescence and, in particular, to high-resolution single-molecule FRET measurements, independent of size or if the protein is folded or disordered. Taking the accumulated information together from this systematic study, for high-resolution smFRET measurements the tRNA^{Pyl}/PylRS system from *M. mazei* genetically encoding **5** is the best system for several reasons: (i) most importantly, expression using our optimized pEvol system guarantees efficient and selective AMBER suppression and high-yielding protein production even in ultrarich media such as terrific broth. Good expression thus also allowed double AMBER suppression, which can then, at the cost of random labeling, be used to label a protein with full labeling freedom at any site in only one step. (ii) **5** can easily be synthesized even with basic chemical tools in gram quantities and reacted with commercially available compounds; (iii) the alkyne moiety is less prone to reduction compared to the azide (particularly important during long-term expression in living cultures); (iv) the D dye maintains a long fluorescence lifetime, supporting a high dynamic range for FRET measurements; (v) CuAAC proceeds readily under physiological conditions; (vi) the initial anisotropy of labeled proteins is very low, encouraging the application of the key assumption for all high-resolution smFRET measurements of κ^2 being 2/3.

Furthermore, the availability to use chemically distinct UAAs with phenylalanine- (**3**, aromatic, hydrophobic) and lysine-like (**5**, aliphatic, hydrophilic) backbones can be useful when the protein backbone structure requires minimal perturbance, such as folding sensitive proteins, where a canonical Phe could be replaced with either **1** or **3**, or a canonical Lys with **5**. **3** might also have advantages when large changes in anisotropy need to be recorded with high precision, such as in drug screening.[51] Whenever the protein is not denaturation-sensitive and can be labeled under acidic conditions, **1** seems to be the ideal choice, as it offers similar photophysical properties as **5** but a higher expression yield. We note that sometimes proteins cannot be labeled using copper-based click chemistry typically due to intolerance against copper. In those cases, our recently demonstrated use of genetically encoding strained alkynes that allow for catalyst-free click chemistry can in principle present a solution.[16] However, strained alkynes are difficult to synthesize and it would not be easy to express large cultures as necessary for the preparation of high purity samples of very challenging to express IDPs such as hNup153.

Finally, our optimized system allowed us to make the challenging protein RanBP3 accessible to smFRET studies by encoding two UAAs at any position we wanted, followed by a simple single-step reaction. We showed a particular region in the IDP-domain adopts collapsed conformation. This now affords the possibility to study also the regulatory mechanism of this multifunctional, partially unfolded protein at single-molecule resolution and to map conformational changes and changes of the protein dynamics induced through protein−protein interactions that are relevant in nuclear transport. Combining our approach in the future with recent developments that facilitate multifunctional and multisite incorporation of UAAs will likely influence the power of smFRET in biology studies dramatically.[52−55] In more general applications, our strategy will also facilitate the labeling with other reporters such as spin markers for NMR, EPR, and MRI as well as post-translational modifications such as glycosylation and ubiquitinylation.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Supporting Information including additional methods and results. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

lemke@embl.de

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by the Emmy Noether program of the DFG. S.M. acknowledges a fellowship from the BIF; S.T. from the Darwin Trust foundation; and T.P. from the VCI. We thank P.G. Schultz (Scripps) for the pEvol system to incorporate **1**, **2**, and **3**.

■ REFERENCES

(1) Deniz, A. A.; Mukhopadhyay, S.; Lemke, E. A. *J. R. Soc. Interface* **2008**, *5*, 15.
(2) Roy, R.; Hohng, S.; Ha, T. *Nat. Methods* **2008**, *5*, 507.

(3) Schuler, B.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2008**, *18*, 16.
(4) Sisamakis, E.; Valeri, A.; Kalinin, S.; Rothwell, P. J.; Seidel, C. A. *Methods Enzymol.* **2010**, *475*, 455.
(5) Brustad, E. M.; Lemke, E. A.; Schultz, P. G.; Deniz, A. A. *J. Am. Chem. Soc.* **2008**, *130*, 17664.
(6) Blaschke, U. K.; Silberstein, J.; Muir, T. W. *Methods Enzymol.* **2000**, *328*, 478.
(7) Scheck, R. A.; Dedeo, M. T.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2008**, *130*, 11762.
(8) Sun, X.; Zhang, A.; Baker, B.; Sun, L.; Howard, A.; Buswell, J.; Maurel, D.; Masharina, A.; Johnsson, K.; Noren, C. J.; Xu, M. Q.; Correa, I. R. Jr. *ChemBioChem* **2011**, *12*, 2217.
(9) Yi, L.; Sun, H.; Itzen, A.; Triola, G.; Waldmann, H.; Goody, R. S.; Wu, Y. W. *Angew. Chem., Int. Ed.* **2011**, *50*, 8287.
(10) Dirksen, A.; Hackeng, T. M.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 7581.
(11) Hackenberger, C. P.; Schwarzer, D. *Angew. Chem., Int. Ed.* **2008**, *47*, 10030.
(12) Tompa, P. *Structure and Function of Intrinsiclaly Disordered Proteins*; CRC PRESS: Boca Raton, FL, 2010.
(13) Hoefling, M.; Lima, N.; Haenni, D.; Seidel, C. A.; Schuler, B.; Grubmuller, H. *PLoS One* **2011**, *6*, e19791.
(14) Sindbert, S.; Kalinin, S.; Nguyen, H.; Kienzler, A.; Clima, L.; Bannwarth, W.; Appel, B.; Muller, S.; Seidel, C. A. *J. Am. Chem. Soc.* **2011**, *133*, 2463.
(15) Muschielok, A.; Andrecka, J.; Jawhari, A.; Bruckner, F.; Cramer, P.; Michaelis, J. *Nat. Methods* **2008**, *5*, 965.
(16) Plass, T.; Milles, S.; Koehler, C.; Schultz, C.; Lemke, E. A. *Angew. Chem., Int. Ed.* **2011**, *50*, 3878.
(17) Milles, S.; Lemke, E. A. *Biophys. J.* **2011**, *101*, 1710.
(18) Denning, D. P.; Patel, S. S.; Uversky, V.; Fink, A. L.; Rexach, M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2450.
(19) Strawn, L. A.; Shen, T.; Shulga, N.; Goldfarb, D. S.; Wente, S. R. *Nat. Cell Biol.* **2004**, *6*, 197.
(20) Young, T. S.; Ahmad, I.; Yin, J. A.; Schultz, P. G. *J. Mol. Biol.* **2010**, *395*, 361.
(21) Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. *Nat. Chem. Biol.* **2008**, *4*, 232.
(22) Nguyen, D. P.; Lusic, H.; Neumann, H.; Kapadnis, P. B.; Deiters, A.; Chin, J. W. *J. Am. Chem. Soc.* **2009**, *131*, 8720.
(23) Lemke, E. A.; Gambin, Y.; Vandelinder, V.; Brustad, E. M.; Liu, H. W.; Schultz, P. G.; Groisman, A.; Deniz, A. A. *J. Am. Chem. Soc.* **2009**, *131*, 13610.
(24) Lemke, E. A. *Methods Mol. Biol.* **2011**, *751*, 3.
(25) Koshioka, M.; Sasaki, K.; Masuhara, H. *Appl. Spectrosc.* **1995**, *49*, 224.
(26) Schaffer, J.; Volkmer, A.; Eggeling, C.; Subramaniam, V.; Striker, G.; Seidel, C. A. M. *J. Phys. Chem. A* **1999**, *103*, 331.
(27) Kapanidis, A. N.; Lee, N. K.; Laurence, T. A.; Doose, S.; Margeat, E.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8936.
(28) Muller, B. K.; Zaychikov, E.; Brauchle, C.; Lamb, D. C. *Biophys. J.* **2005**, *89*, 3508.
(29) Best, R. B.; Merchant, K. A.; Gopich, I. V.; Schuler, B.; Bax, A.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18964.
(30) Hoffmann, A.; Kane, A.; Nettels, D.; Hertzog, D. E.; Baumgartel, P.; Lengefeld, J.; Reichardt, G.; Horsley, D. A.; Seckler, R.; Bakajin, O.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 105.
(31) Laurence, T. A.; Kong, X.; Jager, M.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17348.
(32) Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2649.
(33) Gopich, I. V.; Szabo, A. *J. Phys. Chem. B* **2007**, *111*, 12925.
(34) Liu, C. C.; Schultz, P. G. *Annu. Rev. Biochem.* **2010**, *79*, 413.
(35) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124*, 9026.
(36) Deiters, A.; Schultz, P. G. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1521.
(37) Seo, M. H.; Lee, T. S.; Kim, E.; Cho, Y. L.; Park, H. S.; Yoon, T. Y.; Kim, H. S. *Anal. Chem.* **2011**, *83*, 8849.
(38) Sommer, S.; Weikart, N. D.; Brockmeyer, A.; Janning, P.; Mootz, H. D. *Angew. Chem., Int. Ed.* **2011**, *50*, 9888.
(39) Staros, J. V.; Bayley, H.; Standring, D. N.; Knowles, J. R. *Biochem. Biophys. Res. Commun.* **1978**, *80*, 568.
(40) Gambin, Y.; VanDelinder, V.; Ferreon, A. C.; Lemke, E. A.; Groisman, A.; Deniz, A. A. *Nat. Methods* **2011**, *8*, 239.
(41) Ferreon, A. C.; Gambin, Y.; Lemke, E. A.; Deniz, A. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 5645.
(42) Muller-Spath, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Ruegger, S.; Reymond, L.; Nettels, D.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 14609.
(43) Nettels, D.; Muller-Spath, S.; Kuster, F.; Hofmann, H.; Haenni, D.; Ruegger, S.; Reymond, L.; Hoffmann, A.; Kubelka, J.; Heinz, B.; Gast, K.; Best, R. B.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20740.
(44) Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1999**, *293*, 321.
(45) Lim, R. Y.; Huang, N. P.; Koser, J.; Deng, J.; Lau, K. H.; Schwarz-Herion, K.; Fahrenkrog, B.; Aebi, U. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9512.
(46) Chen, H.; Ahsan, S. S.; Santiago-Berrios, M. B.; Abruna, H. D.; Webb, W. W. *J. Am. Chem. Soc.* **2010**, *132*, 7244.
(47) Lindsay, M. E.; Holaska, J. M.; Welch, K.; Paschal, B. M.; Macara, I. G. *J. Cell Biol.* **2001**, *153*, 1391.
(48) Langer, K.; Dian, C.; Rybin, V.; Muller, C. W.; Petosa, C. *PLoS One* **2011**, *6*, e17011.
(49) Mueller, L.; Cordes, V. C.; Bischoff, F. R.; Ponstingl, H. *FEBS Lett.* **1998**, *427*, 330.
(50) Iqbal, A.; Arslan, S.; Okumus, B.; Wilson, T. J.; Giraud, G.; Norman, D. G.; Ha, T.; Lilley, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11176.
(51) Owicki, J. C. *J. Biomol. Screening* **2000**, *5*, 297.
(52) Johnson, D. B.; Xu, J.; Shen, Z.; Takimoto, J. K.; Schultz, M. D.; Schmitz, R. J.; Xiang, Z.; Ecker, J. R.; Briggs, S. P.; Wang, L. *Nat. Chem. Biol.* **2011**, *7*, 779.
(53) Mukai, T.; Hayashi, A.; Iraha, F.; Sato, A.; Ohtake, K.; Yokoyama, S.; Sakamoto, K. *Nucleic Acids Res.* **2010**, *38*, 8188.
(54) Wan, W.; Huang, Y.; Wang, Z. Y.; Russell, W. K.; Pai, P. J.; Russell, D. H.; Liu, W. R. *Angew. Chem., Int. Ed.* **2010**, *49*, 3211.
(55) Neumann, H.; Wang, K. H.; Davis, L.; Garcia-Alai, M.; Chin, J. W. *Nature* **2010**, *464*, 441.