JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Biosynthesis of a Fluorescent Protein with Extreme Pseudo-Stokes Shift by Introducing a Genetically Encoded Non-Natural Amino Acid outside the Fluorophore

Sebastian M. Kuhn, Marina Rubini, Michael A. Müller, and Arne Skerra*

Munich Center for Integrated Protein Science (CIPS-M) and Lehrstuhl für Biologische Chemie, Technische Universität München, Emil-Erlenmeyer-Forum 5, 85350 Freising-Weihenstephan, Germany

Supporting Information

ABSTRACT: A novel kind of fluorescent protein relying on the intramolecular interplay between two different fluorophores, one of chemical origin and one of biological origin, was developed. The fluorescent non-natural amino acid L-(7-hydroxycoumarin-4-yl)ethylglycine was site-specifically incorporated into the recombinant enhanced cyan fluorescent protein (eCFP) at a permissible surface position ~20 Å away from the protein fluorophore using amber suppression in *Escherichia coli* with an engineered cognate *Methanococcus jannaschii* tRNA synthetase. The resulting eCFP^{Cou} exhibited almost quantitative intramolecular Förster resonance energy transfer (FRET) between its two fluorophores, showing brilliant cyan emission at 476 nm upon excitation in the near-UV at 365 nm (a wavelength easily accessible via conventional laboratory UV sources), in contrast to its natural counterpart. Thus, this fluorescent protein with unprecedented spectroscopic properties reveals an extreme apparent Stokes shift of ~110 nm between the absorption wavelength of the coumaryl group and the emission wavelength of eCFP.

Since the first application of the recombinant green fluorescent protein (GFP) from *Aequorea victoria*,[1] the class of visibly fluorescent proteins (FPs) has become an invaluable molecular tool in cell and developmental biology, biophysics, and biotechnology. Numerous engineered versions of GFP itself as well as orthologues from other (mainly marine) organisms have been described.[2] From alteration of the substitution pattern or intraprotein environment of the 4-(*p*-hydroxybenzylidene)imidazolin-5-one core fluorophore, which autocatalytically forms from three amino acids as part of the polypeptide chain, a whole series of FPs with a broad spectrum of absorption and emission characteristics can result. Several attempts have been made to substitute the aromatic side chain at position 66 of GFP (originally Tyr), whose π-electron system is conjugated to the heterocyclic core group (see Figure 1) and strongly influences its spectroscopic properties.

However, the chemical space of the 20 natural amino acid building blocks is limited, and early on there was a desire to introduce other types of side chains. For example, 4-aminotryptophan has been incorporated at this position of the enhanced cyan fluorescent protein (eCFP) via metabolic pressure to yield a "golden FP".[3] Furthermore, *O*-methyl-L-tyrosine was introduced at the same position of the uvGFP variant via amber suppression employing an orthogonal aminoacyl-tRNA synthetase (aaRS)/tRNA pair, leading to another biosynthetic FP with unique spectral properties.[4−6]

With regard to the design of novel FPs, the latter approach in principle also allows insertion of spectroscopically active side chains at positions outside the central fluorophore, thus indirectly modulating its excitation and/or emission behavior. In this report, we describe the incorporation of a coumaryl residue at a permissible surface position of eCFP that forms a highly efficient Förster pair with the inner fluorophore, yielding an FP with an extreme apparent Stokes shift of ~110 nm between the absorption wavelength of the coumaryl group and the emission wavelength of eCFP.

The fluorescent amino acid L-(7-hydroxycoumarin-4-yl)ethylglycine (**1**) (Figure 1) was previously incorporated into proteins via amber suppression with an engineered aaRS.[7] In its phenolate state, **1** exhibits strong fluorescence at ~450 nm with a high quantum yield (QY) of 0.63 upon excitation at 360 nm.

Here we show that the spectral properties of **1** also make it a suitable intramolecular donor for eCFP by means of Förster resonance energy transfer (FRET).[8,9] FRET is a quantum-mechanical phenomenon that occurs when two fluorophores in sufficient spatial proximity (<100 Å) show spectral overlap between the emission band of the shorter-wavelength fluorophore (the "donor") and the absorption band of the longer-wavelength fluorophore (the "acceptor"). As the so-called Förster effect is strongly distance-dependent, it is often applied in biochemistry as a kind of macromolecular ruler.[8,10]

While initially applied in conjunction with chemical labeling of biomolecules, FRET between different GFPs—as biological gene products—has also been exploited, thus opening its use in cell biology.[11−13] However, a problem of FPs for many practical applications is their typically inefficient excitation in the near-UV spectral range. The most commonly used eGFP and, in particular, eCFP show poor brightness when illuminated with a typical laboratory UV light source or a laser at 366 nm. To address this general limitation, we sought to evoke an intramolecular FRET effect by combining the fluorescence activity of a natural FP with the spectrally matching fluorescence of a non-natural amino acid incorporated during its biosynthesis. As an initial attempt, we chose eCFP as the FRET acceptor for **1** because its absorption maximum in the range of

Received: November 6, 2010
Published: February 22, 2011

dx.doi.org/10.1021/ja1099787 | *J. Am. Chem. Soc.* 2011, 133, 3708–3711

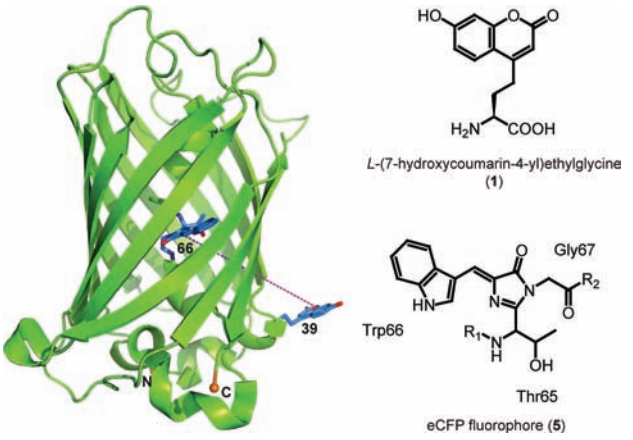

Figure 1. Molecular structures of the fluorophores used in this study. (left) Three-dimensional model of eCFP^{Cou}, which carries the fluorescent amino acid **1** at the surface-exposed position 39, based on the crystal structure of eCFP (PDB entry 2WSN). (right) Chemical structures of (bottom) the eCFP fluorophore 4-[(1*H*-indol-3-yl)methylidene]imidazolin-5-one (**5**; λ_{ex} = 434 nm, λ_{em} = 476 nm) and (top) the fluorescent non-natural amino acid L-(7-hydroxycoumarin-4-yl)ethylglycine (**1**; λ_{ex} = 360 nm, λ_{em} = 450 nm), which together can form a Förster pair. In the structure of **5**, R_1 and R_2 denote the N- and C-terminal polypeptide segments, respectively.

430−450 nm almost perfectly coincides with the 7-hydroxycoumarin emission maximum (Figure 2). Although eCFP has a moderate QY of 0.4 for 476 nm emission,[14] its high excitation coefficient of 32 500 M^{-1} cm^{-1} at 434 nm is ideal for efficient (radiationless) absorption of donor fluorescence energy (i.e., FRET).

1 was chemically synthesized according to a published procedure[7,15] with an improved purification by crystallization (see the Supporting Information). To achieve cotranslational incorporation of **1** into eCFP, its coding region was expressed in *Escherichia coli* using a recently developed one-plasmid expression system[6] that also carried the genes for an aaRS specific to **1**[7] and for a cognate amber suppressor $tRNA^{Tyr}_{CUA}$. For site-directed incorporation of the foreign amino acid, an amber stop codon (TAG) was introduced at the position of Tyr39 in the eCFP reading frame, and the *Strep*-tag II[16] was appended at the C-terminus to allow affinity purification. In the crystal structure of eCFP[17] (PDB entry 2WSN), the side chain of Tyr39 is located in a loop on the surface of the β-can motif at a distance of ~20 Å from the central eCFP fluorophore, which is in sufficient proximity to ensure efficient FRET (Figure 1).

From a 1 L culture of *E. coli* BL21 harboring this system in rich medium supplemented with **1**, ~0.1 mg of the recombinant protein (dubbed eCFP^{Cou}) was obtained after purification via *Strep*-Tactin affinity chromatography and polishing by anion-exchange and size-exclusion chromatography. High protein purity was confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and electrospray ionization mass spectrometry (ESI-MS) (Figure 3). In particular, the mass spectrum revealed a unique prominent peak at 27959.8 Da, which closely matches the calculated value of 27959.4 Da. This demonstrated essentially quantitative incorporation of **1** instead of the original Tyr residue, apart from full maturation of the central eCFP fluorophore, which is accompanied by a loss of H_2O and H_2, as well as complete processing of the N-terminal Met residue.

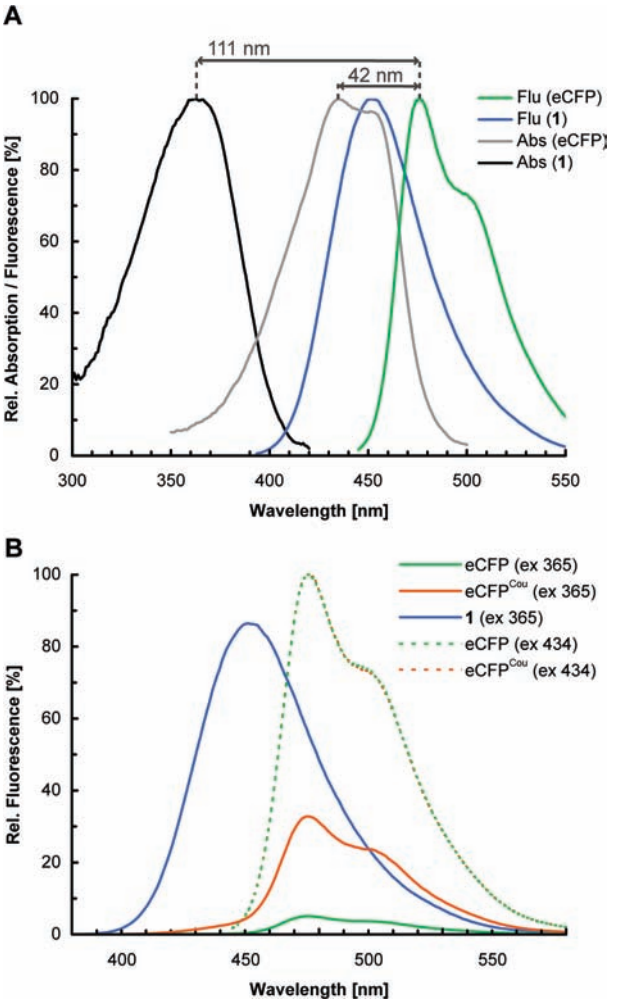

Figure 2. Spectroscopic analysis of the novel FP. (A) Normalized absorption and fluorescence spectra of **1** [Abs(**1**), Flu(**1**)] and eCFP [Abs(eCFP), Flu(eCFP)]. The major absorption band of eCFP shows considerable overlap with the fluorescence spectrum of **1**, thus fulfilling a prerequisite for efficient FRET. (B) Fluorescence spectra of **1**, eCFP, and eCFP^{Cou} (2.2 μM each) upon excitation at 365 and 434 nm. Upon excitation at 365 nm, the coumaryl donor emission in eCFP^{Cou} is almost completely quenched (comparison of blue and orange lines, respectively, at 450 nm) whereas its acceptor-group emission is dramatically increased (orange line at 476 nm) in comparison with that of the unmodified eCFP (green). Notably, upon excitation at 434 nm, both eCFP and eCFP^{Cou} show identical high emission (dashed green and orange lines, respectively).

The biosynthetic eCFP^{Cou} was investigated by collecting its absorption and fluorescence spectra at pH 9.0 (Figure 2 and Figure S4 in the Supporting Information) and comparing them with those of the synthetic fluorescence donor compound **1** and the natural bacterially produced fluorescence acceptor eCFP. The isolated donor and acceptor fluorescence spectra showed the expected maximal absorption and fluorescence wavelengths, respectively, around 360 and 450 nm for **1** and 434 and 476 nm for eCFP. Notably, eCFP^{Cou} exhibited an absorption spectrum corresponding closely to the theoretical absorption of an equimolar mixture of **1** and eCFP (Figure 2 and Figure S4).

However, upon excitation at 365 nm, the anticipated donor fluorescence in eCFP^{Cou} was efficiently quenched (Figure 2B, orange line at 450 nm) whereas a bright acceptor fluorescence

dx.doi.org/10.1021/ja1099787 |*J. Am. Chem. Soc.* 2011, 133, 3708–3711

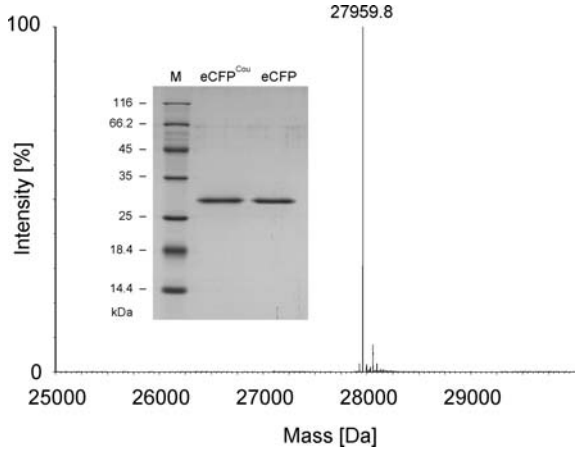

Figure 3. Analysis of the bacterially produced and purified fluorescent proteins eCFP and eCFP^{Cou} by ESI-MS and Coomassie-stained SDS-PAGE (inset). For unlabeled eCFP (calcd mass 27877.3 Da), only a minute signal (amounting to less than 2.6% of the intensity for eCFP^{Cou}) was detectable in the mass spectrum, thus confirming essentially quantitative incorporation of **1** instead of the natural Tyr residue at position 39.

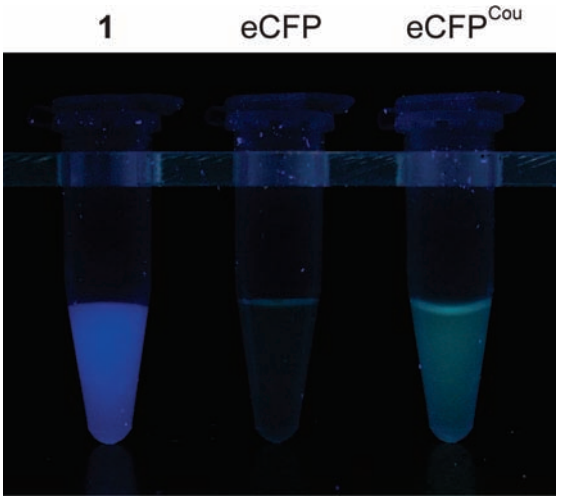

Figure 4. Eppendorf tubes filled with 1 μM solutions of **1**, eCFP or eCFP^{Cou}, all excited with a conventional UV hand lamp at 366 nm. The strongly enhanced fluorescence of eCFP^{Cou} vs eCFP is visible, and the fluorescence is clearly red-shifted in comparison with that of **1**.

at 476 nm appeared (Figure 2B, orange line). In fact, the coumaryl fluorescence was suppressed by a factor of 17, as estimated from a comparison of the stationary peak fluorescence at 450 nm with that of an equimolar mixture of **1** with eCFP, while the concomitant eCFP^{Cou} emission had a peak intensity more than 6-fold higher than the one of eCFP upon excitation at 365 nm (Figure 2B, comparison of orange and green lines). Still, it was also possible to excite eCFP^{Cou} at 434 nm, again giving rise to the typical eCFP fluorescence at 476 nm with unchanged QY.

The observed strong quenching of the hydroxycoumaryl donor group in eCFP^{Cou} indicated potent intramolecular FRET. The FRET efficiency in biochemical systems is usually quantified according to the change in the ratio of acceptor to donor stationary peak fluorescence upon excitation at the donor's excitation wavelength (here 365 nm). In the case of eCFP^{Cou}, this ratio showed a remarkable 9.3-fold increase, from a value of 0.70 measured for a mixture of the two isolated fluorophores to 6.5 in the biosynthetic protein (using donor and acceptor emission wavelengths of 452 and 476 nm, respectively). Since eCFP exhibits a double-peak fluorescence spectrum,[14] the ratio of acceptor to donor emission even increased to 16.3 when the acceptor emission at 505 nm was probed, because the spectral overlap with **1** in the equimolar mixture serving as the reference decreased at this higher fluorescence wavelength. On the other hand, the fluorescence of eCFP was only ~30% lower at this longer wavelength than at its first emission maximum at 476 nm.

The highly efficient Förster transfer in our biosynthetic hybrid protein is remarkable and deserves further discussion. The 7-hydroxycoumarylethyl amino acid employed here as the donor has three intra-side-chain rotational degrees of freedom ($C_\alpha-C_\beta$, $C_\beta-C_\gamma$, $C_\gamma-C_\delta$). With its solvent- exposed position at the protein surface, it is unlikely to be trapped in an orientation unsuitable for FRET with the central eCFP fluorophore upon consideration of the roles of fluorescence anisotropy and polarization. The fluorescence lifetime of 7-hydroxy-4-methylcoumarin is ~15 ns for the anionic form.[18] On the other hand, the rotational relaxation times of amino acid side chains in unfolded proteins lie in the significantly shorter range of 200 ps to 1.6 ns,[19] thus preventing an unfavorable orientation factor. Furthermore, a calculation of the critical Förster transfer distance R_0 (using an orientation factor $\kappa^2 = {}^2/_3$)[20] based on the excitation and emission spectra collected for the eCFP fluorophore and **1**, respectively, yielded a value of 45 Å, which is considerably larger than their spatial distance in the biosynthetic hybrid protein (Figure 1) and hence explains the high efficiency of this process.

As a result of the almost complete intramolecular energy transfer in the folded protein, our eCFP^{Cou} can be regarded as a novel type of FP with a large pseudo-Stokes shift of ~110 nm, which contrasts with the much lower shift of ~40 nm for the natural eCFP (Figure 2A). Until now, such a pseudo-Stokes shift above 100 nm could be achieved almost only with tandem dyes or quantum dots.[21]

Our new FP should be useful not only for biophysical and spectroscopic studies. As one advantage, the fluorescence of eCFP^{Cou} upon irradiation with a standard 366 nm UV lamp available in most laboratories is easily detectable by the naked eye, whereas the natural eCFP shows only very faint fluorescence under such conditions (Figure 4). More generally, our approach should be applicable to other combinations of fluorescent proteins with non-natural fluorescent amino acids, providing novel reagents with spectroscopic properties tailored for various purposes.

■ ASSOCIATED CONTENT

Supporting Information. Details of plasmid design, recombinant protein production and purification, synthesis and characterization of the non-natural amino acid, absorption and fluorescence measurements, MS analyses, and pH titration data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
skerra@tum.de

REFERENCES

(1) Shimomura, O. *Angew. Chem., Int. Ed.* **2009**, *48*, 5590.
(2) Tsien, R. Y. *Angew. Chem., Int. Ed.* **2009**, *48*, 5612.
(3) Bae, J. H.; Rubini, M.; Jung, G.; Wiegand, G.; Seifert, M. H.; Azim, M. K.; Kim, J. S.; Zumbusch, A.; Holak, T. A.; Moroder, L.; Huber, R.; Budisa, N. *J. Mol. Biol.* **2003**, *328*, 1071.
(4) Wang, L.; Xie, J.; Deniz, A. A.; Schultz, P. G. *J. Org. Chem.* **2003**, *68*, 174.
(5) Kajihara, D.; Hohsaka, T.; Sisido, M. *Protein Eng. Des. Sel.* **2005**, *18*, 273.
(6) Kuhn, S. M.; Rubini, M.; Fuhrmann, M.; Theobald, I.; Skerra, A. *J. Mol. Biol.* **2010**, *404*, 70.
(7) Wang, J.; Xie, J.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 8738.
(8) Stryer, L. *Annu. Rev. Biochem.* **1978**, *47*, 819.
(9) Wu, P.; Brand, L. *Anal. Biochem.* **1994**, *218*, 1.
(10) Hillisch, A.; Lorenz, M.; Diekmann, S. *Curr. Opin. Struct. Biol.* **2001**, *11*, 201.
(11) Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. *Nat. Rev. Mol. Cell. Biol.* **2002**, *3*, 906.
(12) Frommer, W. B.; Davidson, M. W.; Campbell, R. E. *Chem. Soc. Rev.* **2009**, *38*, 2833.
(13) Day, R. N.; Davidson, M. W. *Chem. Soc. Rev.* **2009**, *38*, 2887.
(14) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509.
(15) Brun, M. P.; Bischoff, L.; Garbay, C. *Angew. Chem., Int. Ed.* **2004**, *43*, 3432.
(16) Schmidt, T. G.; Skerra, A. *Nat. Protoc.* **2007**, *2*, 1528.
(17) Lelimousin, M.; Noirclerc-Savoye, M.; Lazareno-Saez, C.; Paetzold, B.; Le Vot, S.; Chazal, R.; Macheboeuf, P.; Field, M. J.; Bourgeois, D.; Royant, A. *Biochemistry* **2009**, *48*, 10038.
(18) Schulman, S. G.; Rosenberg, L. S. *J. Phys. Chem.* **1979**, *83*, 447.
(19) Creighton, T. E. *Proteins: Structures and Molecular Properties*, 2nd ed.; W. H. Freeman: New York, 1992.
(20) Hink, M. A.; Visser, N. V.; Borst, J. W.; van Hoek, A.; Visser, A. J. W. G. *J. Fluoresc.* **2003**, *13*, 185.
(21) Rantanen, T.; Pakkila, H.; Jamsen, L.; Kuningas, K.; Ukonaho, T.; Lovgren, T.; Soukka, T. *Anal. Chem.* **2007**, *79*, 6312.

NOTE ADDED AFTER ASAP PUBLICATION

In the version of this Communication published ASAP February 22, 2011, $C_\beta - C_\gamma$ was omitted from the list of three intra-side-chain rotational degrees of freedom for the 7-hydroxycoumarylethyl amino acid. The corrected version was published February 25, 2011.