

FINAL REPORT TO DESIGN AND SYNTHESIS OF PHOTOCAGED AMINO ACIDS WITH VISIBLE LIGHT ACTIVATION FOR EPIGENETIC STUDIES

Within this project we aimed at developing photocaged amino acids which can be activated by visible light. To achieve this aim we planned to functionalize amino acids with various photolabile protecting groups. By incorporating these non-canonical amino acids (ncAAs) into key positions of proteins via genetic code expansion, their functionality could be turned off and later rapidly restored by an external trigger (light). During the research project photolabile frames with certain properties were selected: they should have absorption in the visible range (>400 nm) with appreciable photochemical efficiency. The latter would require relatively large molar absorption coefficients (>15 000 M⁻¹cm⁻¹) and photochemical quantum yields (>0.1). Moreover, they should be reasonably small and compact to facilitate the incorporation of the non-canonical amino acids by means of amber-suppression technology. In accordance with these criteria, several photocages (mostly coumarin, quinoline, BODIPY derivatives) were chosen, synthesized and installed onto amino acids. The photocaged amino acids were then studied to assess their photoactivation characteristics under visible light irradiation (460 nm). The most active compounds were selected and applied to the genetic encoding experiments. In order to assess the efficiency of genetic incorporation, we have developed a flow cytometry method. However, no incorporation was observed for any of the new non-canonical amino acids by members of our existing tRNA synthetase library or by novel enzymes designed and constructed by us specifically for these given amino acids. Therefore, we turned our attention to alternative solutions, i.e., the “click-to-release” mechanism that offers reactivation of blocked amino acids by another external stimulus, more particularly a chemical trigger. As an alternative for tetrazines, we sought for new dienes and studied cyclopentadienone and pyrone derivatives in this respect. Meanwhile we also developed and genetically successfully encoded a novel, isonitrile-functionalized non-canonical amino acid, which could be selectively labeled in mutually orthogonal bioorthogonal labeling schemes with fluorophores equipped with a bulky, stable tetrazine.

RESULTS

1. Synthesis of the selected photocages

1.1. 4,5-dialkoxy-2-nitrobenzyl derivatives

First, we have synthesized the most commonly used 1-(4,5-dimethoxy-2-nitrophenyl)ethan-1-ol (**1**) compound and its improved derivative, 1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethan-1-ol (**2**) in an efficient and quick manner. In the subsequent step, they were conjugated to the epsilon amino group of lysine *via* a carbamate bond. These compounds were to be used in genetic encoding experiment as known UV-

activable photocaged Lys-derivatives. We envisioned, that this photocaging moiety could later be utilized as an orthogonal protecting group in combination with other photolabile protecting groups that can be cleaved at different wavelengths.

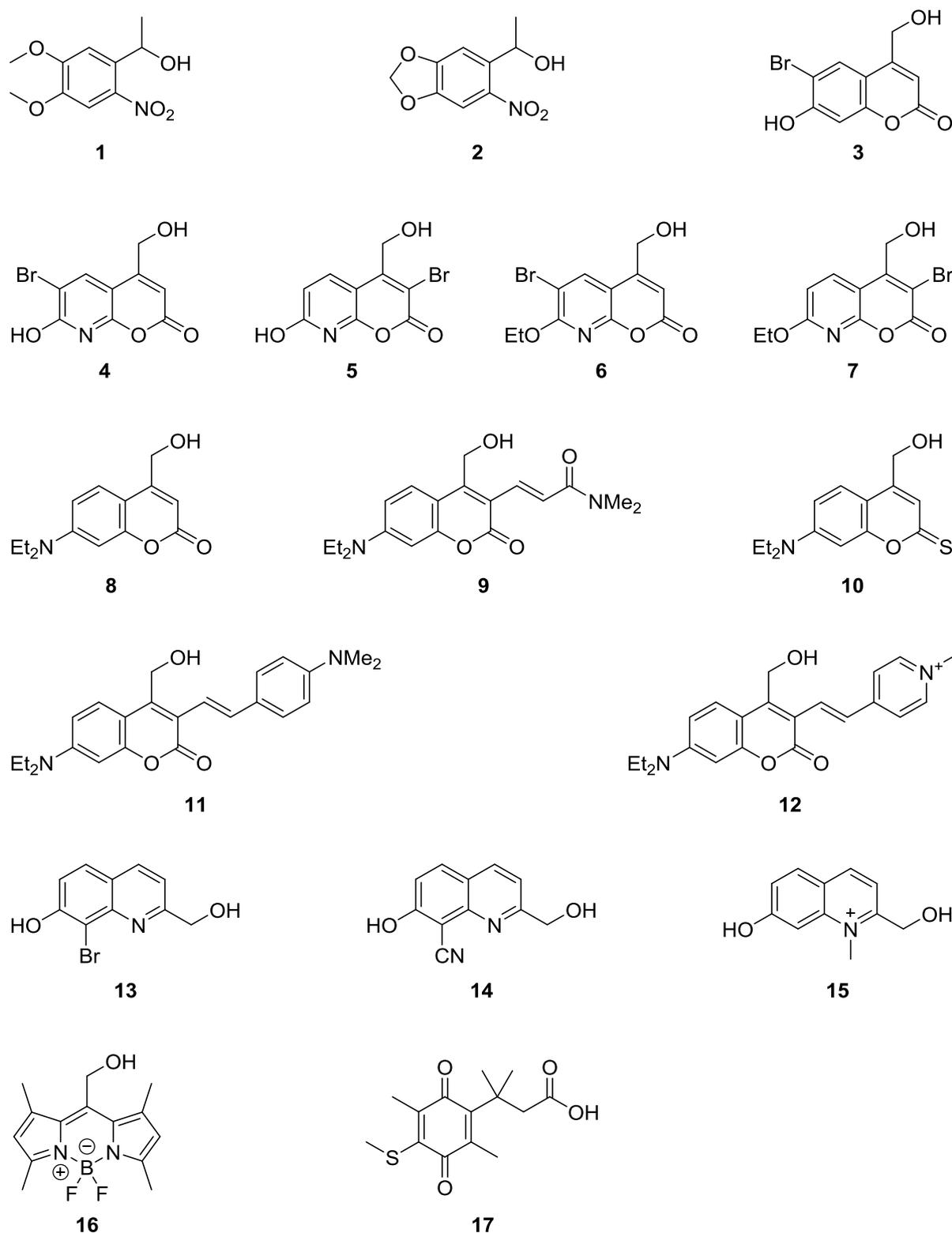


Figure 1. Selected and synthesized photocages.

1.2. Coumarin scaffolds

1) 6-Bromo-7-hydroxy-4-(hydroxymethyl)-coumarin (**3**) was prepared following literature procedures. This compound can be conjugated to amino acids via its 4-hydroxymethyl group. To this compound, commercially available tBu-protected glutamic acid (H-Glu(tBu)-OtBu) was conjugated *via* the photolabile site for in vitro release studies. However, upon irradiation with blue light (>400 nm), only negligible amount of glutamic-acid release was detected.

2) From the proposed 8-azacoumarin compounds, a set of four derivatives have been synthesized i.e. 6-bromo-7-hydroxy-4-(hydroxymethyl)-8-azacoumarin (**4**), 3-bromo-7-hydroxy-4-(hydroxymethyl)-8-azacoumarin (**5**), and their *O*-ethylated variants, 6-bromo-7-ethoxy-4-(hydroxymethyl)-8-azacoumarin (**6**) and 3-bromo-7-ethoxy-4-(hydroxymethyl)-8-azacoumarin (**7**), respectively. Each of these compounds was conjugated to H-Glu(tBu)-OtBu for photocleavage studies. The results have shown that the compounds with free 7-OH moiety (**4**, **5**) are more photolabile than their 7-OEt congeners (**6**, **7**), as these latter showed no amino acid release with blue light. Although literature examples suggest that **4** and **5** are efficient visible-light photocages, the photocleavage reaction was not quick enough under physiological conditions with a 462 nm LED light source for biological experiments.

3) The synthesis of the proposed 7-diethylamino-4-(hydroxymethyl)-coumarin (**8**) scaffold is reported in the literature. However, the synthetic route involves a SeO₂ mediated oxidation step with an unreliable and generally low yield. Therefore a new synthetic route was designed and the desired molecule was acquired in a faster and more efficient way. In the next step it was conjugated to H-Glu(tBu)-OtBu. Upon irradiation with blue light (>400 nm) release of the caged amino acid was observed in a moderate yield therefore a lysine derivative of this compound was also prepared for genetic encoding experiments. For this, a new synthetic route was established for the larger (multi-gram) scale, reliable and less expensive synthesis of the proposed and promising **8** coumarin scaffold, since it is directly necessary for the biological experiments as well as served as a starting material for the further coumarin-based photocages. The conjugation reaction to lysine was optimized and the photophysical and photochemical properties of the product (**8-Lys**) were quantified. The compound was found to be stable in the dark at room temperature, while the photolytic efficiency ($\epsilon \times \Phi$) was found to be 142 at the absorption maximum (385 nm) and was still over 100 at 410 nm, which is one of the frequently used laser sources in fluorescence microscopy. However, at 462 nm (at the emission maximum of the common blue LED lights) it drops to 0.14. This implies that the absorption wavelength should be further fine-tuned and shifted more towards the red region. It is important to note, that the photolytic efficiency ($\epsilon \times \Phi$) for commonly used, UV-light cleavable 1-(4,5-dimethoxy-2-nitrophenyl)ethan-1-ol derivatives is around 100 at 365 nm which is fast enough for most biological experiments.

4) Introducing an electron withdrawing moiety at the 3rd position of the coumarin ring is proven to shift absorption maximum towards longer wavelengths. Considering steric limitations in genetic encoding experiments, small *N,N*-dimethylacrylamide was chosen and 3-(4-hydroxymethyl-7-diethylamino-coumarin-3-yl)-*N,N*-dimethylacrylamide (**9**) was prepared from the above **8** coumarin in 4 steps and then conjugated to the ϵ -amino group of lysine under the previously optimized conditions. The compound (**9-Lys**) was found to be stable in the dark at room temperature, and the absorption maximum shifted to 445 nm and even the molar absorption coefficient (ϵ) increased, however, the photochemical quantum yield (Φ) and consequently the photolytic efficiency ($\epsilon \times \Phi$) decreased dramatically (i.e. 0.45 at 462 nm).

3) The synthesis of the proposed 7-diethylamino-4-(hydroxymethyl)-2-thiocoumarin scaffold (**10**) is reported in the literature and after a few optimization attempts was prepared from 7-diethylamino-4-(hydroxymethyl)-coumarin (**8**) in 3 steps in moderate yield. Conjugation to lysine was accomplished similarly to previous examples. The compound (**10-Lys**) was found to be stable in the dark at room temperature, with absorption maximum shifted to 470 nm and photolytic efficiency ($\epsilon \times \Phi$) was slightly lower (i.e. 40 at 462 nm), but still in the same order of magnitude as the *ortho*-nitrobenzyl derivatives'.

Compound	λ_{\max} / nm	ϵ_{\max} / $M^{-1}cm^{-1}$	ϵ_{462} / $M^{-1}cm^{-1}$	ϕ	$\epsilon_{\max} \times \phi$	$\epsilon_{462} \times \phi$
8-Lys	385	26463	25.16	$5,38 \cdot 10^{-3}$	142	0.14
9-Lys	445	37333	32195	$1,39 \cdot 10^{-5}$	0,52	0.45
10-Lys	470	25848	24522	$1,62 \cdot 10^{-3}$	41,8	39.7

Table 1. Photophysical and photochemical properties of the visible-light cleavable coumarin – lysine compounds.

4) During the preliminary testing of the previous photocaged amino acids we observed some incorporation of **8-Lys** and **9-Lys** (unfortunately false positives, as we later found out). Therefore, we have continued the synthesis of further caged lysine derivatives using new coumarin-based photolabile groups with superior photophysical characteristics, recently developed in our group (M. Bojtár et al. Org. Lett. 2019, 21, 9410) or by others (Q. Lin et al. Angew. Chem. Int. Ed. 2018, 57, 3722). Thus, we have installed blue light activatable 7-(diethylamino)-3-(4-(dimethylamino)styryl)-4-(hydroxymethyl)-coumarin (**11**) ($\lambda_{\max} = 446$ nm with extraordinary uncaging efficiency) and green-light activatable 4-(2-(7-(diethylamino)-4-(hydroxymethyl)-coumarin-3-yl)vinylyl)-1-methylpyridin-1-ium (**12**) ($\lambda_{\max} =$

493 nm) groups onto lysine. This latter derivative was readily soluble in water and also allows two-photon uncaging.

1.3. Quinoline-based photocages

The arising problems with the genetic encoding experiments of the 7-diethylaminocoumarin derivatives turned us towards smaller photocages. The concise structure and good aqueous solubility of the quinolone scaffold seemed promising, and although these are traditionally UV-activable molecules, some offer efficient two-photon uncaging with 740 nm fs-pulsed laser. According to protein docking analysis, these smaller photoactivable amino acids could be better recognized by the available tRNA transferases. Therefore two derivatives, 8-bromo-2-(hydroxymethyl)quinolin-7-ol (**13**) and 7-hydroxy-2-(hydroxymethyl)quinoline-8-carbonitrile (**14**) have been chosen, synthesized and conjugated to lysine. During the research project a 7-hydroxy-N-methylquinolinium chromophore (**15**) with an absorption maximum of 418 nm was published. To harness the visible-light absorption of this novel photolabile moiety, its lysine derivative was also prepared. After addressing the problems of the efficient and selective *N*-alkylation of the quinoline core, the preliminary results showed that the photocleavage reaction is fast at 460 nm and its water solubility is exceptionally good, which was very promising for the biological experiments.

1.4. BODIPY photocage

Several 8-(hydroxymethyl) BODIPY frames arose as visible-light cleavable photocages in recent years. To assess their usability in our project, we have chosen and synthesized the smallest stable derivative, 1,3,5,7-tetramethyl-9-(hydroxymethyl) difluoroborate (**16**). Preparation of the required lysine conjugate proved to elusive in the beginning as coupling to different lysine derivatives under a wide variety of conditions has failed or resulted in undesired products. After considerable efforts and optimization, the planned BODIPY-caged lysine derivative has been successfully synthesized, however, its very poor water solubility did not allow further biological experiments.

1.5. Quinone Trimethyl Lock

Although the original proposal did not include, the set of target compounds was extended with a moiety called quinone trimethyl lock because of the lack of success at encoding the coumarin scaffolds. First **17**, the most compact version of this compound family with the desired spectral properties was prepared by an improved reaction route. However, conjugation with H-Glu(tBu)-OtBu via activated ester coupling failed. Therefore, we attached the compound to 7-hydroxy-4-methyl coumarin, a commonly used molecule in release studies. Upon irradiation of this construct with blue light (462 nm) relatively fast release of the free coumarin was observed. Following optimization of the conjugation reaction, we prepared the photocaged lysine derivative in sufficient yields for the biological experiments.

2. Genetic encoding

2.1. Model systems

In order to elaborate the incorporation of photocaged-lysine derivatives into the proteins of interest we have planned to reproduce a test system established by Gautier et al. (*J. Am. Chem. Soc.* **2010**, *132*, 4086). In this system expression of a fusion protein construct (mCherry-linkerTAG-GFP) is monitored. While efficiency of transfection is followed by mCherry fluorescence, stop codon suppression i.e. the incorporation of the noncanonical amino acid (photocaged-lysine) is indicated by the appearance of the GFP signal. As a starting point, we planned to test three variants of pyrrolysine-tRNA transferase: the wild type pyrrolysine tRNA synthetase, and its two mutant congeners, the AF and U6 mutants. This work is currently in progress. At an early stage of the project, we have realized that the GFP-containing plasmid suitable for testing the incorporation of the visible-light reactive non-canonical amino acids should be reconstructed since excitation of GFP results in the cleavage of the photolabile groups. To this end we have made an mOrange-containing construct successfully.

To be able to follow the incorporation efficiency of our non-canonical amino acids and the subsequent photo-uncaging process, we designed a double fusion protein containing an mCherry fluorescent protein and a nuclear localization signal (NLS) tag. Efficiency of transfection can be followed by mCherry fluorescence intensity. The photocaged amino acid is to be incorporated at a key position of the NLS tag, resulting in disabled delivery of the construct into the nucleus. Upon photo-uncaging of the key Lys residue, the NLS tag is reactivated and delivers the construct into the nucleus. Thus, relocation of the mCherry fluorescence from the cytoplasm to the nucleus would clearly indicate successful uncaging.

2.2. Results with known tRNA synthetases

We started our experiments of genetic incorporation by means of amber suppression technology with **2-Lys** (as an established reference) together with **8-Lys**, **9-Lys**, **10-Lys**, selected from the *in vitro* studies. From the enzymatic side, during the project we have tried out 6 known tRNA synthetase systems: 3 pyrrolysine tRNA synthetase variant (the wild type PyIRS pyrrolysine tRNA synthetase, its AF-NES modification and the so-called PCK mutant which was developed for **2-Lys**) and 3 tyrosyl tRNA synthetase derivative (Azp (developed for azidophenylalanine), Acp (developed for acetylphenylalanine) and Bzp (developed for benzoylphenylalanine)).

Incorporating **2-Lys** with PCK was successful and served as a positive control (*Figure 2*, middle). **10-Lys** showed poor solubility in cell media and precipitated during the experiments, resulting in no production of mCherry-TAG-GFP proteins. **8-Lys** and **9-Lys** seemed to show very little, but at least some incorporation with Azp (*Figure 2*, bottom), and later with Bzp. However, when we tried the experiment with NLS-GFP or NLS-mCherry, no relocation of the proteins was observed after irradiation indicating unsuccessful photo-uncaging process.

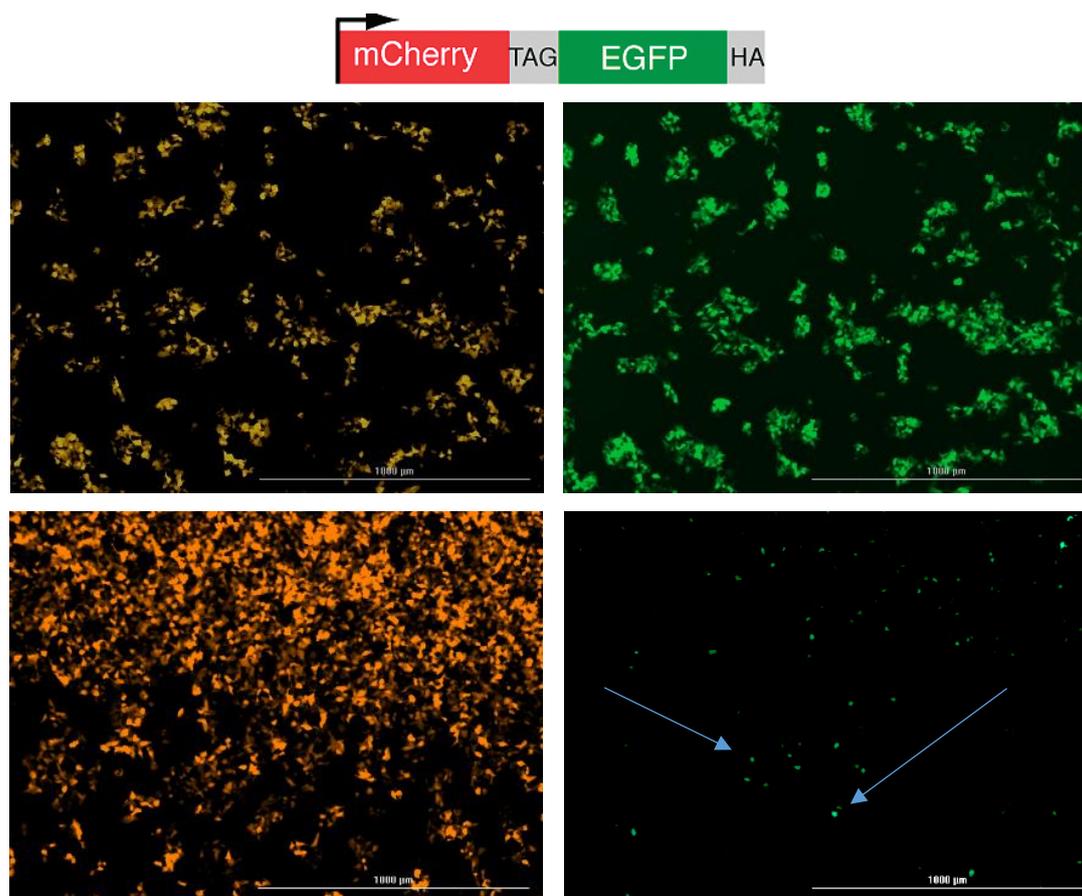


Figure 2. Top: the schematic representation of the ncAA containing fusion peptide; middle: mCherry and GFP fluorescence of cells using the **2-Lys** / PCK system; bottom: mCherry and GFP fluorescence of cells with **8-Lys** / Azp. mCherry signal (left) shows efficient transfection in both cases, while GFP signal (right) shows ncAA incorporation levels.

As the artificial tRNA synthetases are known to be “leaky” (sometimes they incorporate a natural amino acid at the position of the STOP codon, resulting in protein transcription and false positive signals), which are hard to detect in the microscope, we set out to develop a flow cytometry method to validate and quantify real ncAA incorporation. *Figure 3* shows example from these results. As positive control (*Figure 3*, left), cells expressing the mCherry-TAG-GFP protein were treated with azidophenylalanine, the natural substrate of the Azp enzyme. Strong and numerous signals in the upper right corner indicates efficient transfection and translation of the whole protein (high mCherry intensity) due to efficient incorporation of the ncAA (high GFP intensity). On the other hand, in case of **8-Lys** (*Figure 3*, middle) the GFP intensity was much weaker and comparable to the case where no ncAA was added to the cells (*Figure 3*, right), meaning no real incorporation and the enzyme is indeed “leaky”. Unfortunately, all other tRNA synthetases were found to be inefficient in terms of recognizing the caged amino acids and resulted in low level of incorporation into proteins under several conditions even after long optimization experiments with both **8-Lys** and **9-Lys** and later also for **13-Lys**, **14-Lys**, **15-Lys**, and **17-Lys**.

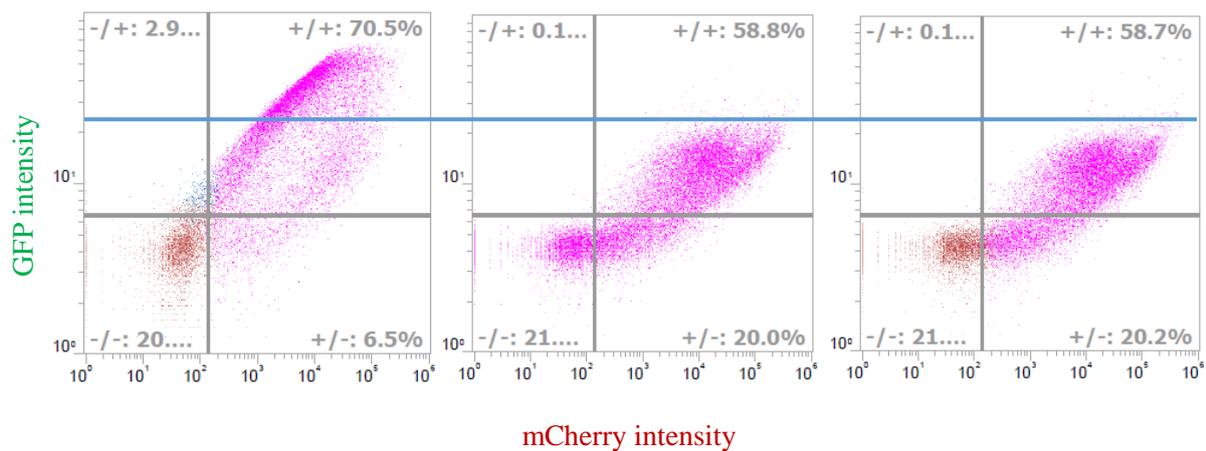
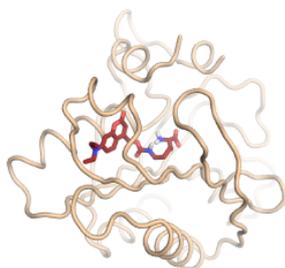


Figure 3. Flow cytometry examination of cell populations expressing mCherry-TAG-GFP proteins with Azp tRNA synthetase. Left: cells treated with azidophenylalanine (positive control); middle: cells treated with **8-Lys**; right: cells treated with no ncAA (negative control).

2.3. tRNA synthetase mutation experiments

The unsuccessful encoding experiments have prompted us to start *in silico* docking studies with the different ncAAs and tRNA synthetases. Although these suggested that Bzp could be a good choice for our smaller amino acids (which later turned out to be false), it also helped us to identify steric problems and revealed key amino acids which should be mutated in order to result in an optimal size of binding pocket suitable for accommodating the ncAA substrates. Based on these findings, we have chosen the AF-NES mutant and designed a number of modifications at different amino acid positions, summarized on Figure 4. Successful preparation of the new enzymes were verified by protein sequencing from both directions.

Vertical mutants



Coumarin ring replaces **Y306** (A)

Pocket-enlarging mutations:

L306 - to M or A or G

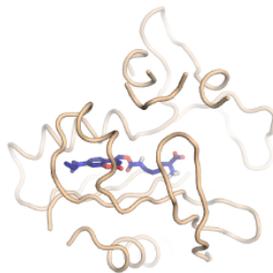
I413 - to A or G

Stabilizing mutations: none

All variants contain the obligatory pocket-enlarging mutation at **M350** (to G)

To enhance folding of proteins, **Y385** is always replaced by F

Horizontal mutants



Coumarin ring replaces **W417** (A)

Pocket-enlarging mutations:

L306 - to A or G

I413 - to A or G

Stabilizing mutations: **I405** - to S

#	Mutations	
	1	AG
2	AG	IAW
3	AA	IGW
4	AA	IAW
5	AM	IGW
6	YG	IGA
7	YG	IAA
8	YG	SGA
9	YG	SAA
10	YA	IGA

Figure 4. 8-Lys docked into the active site of AF-NAS and the proposed changes to its amino acid sequence with the table showing the chosen combinations and successfully prepared mutants.

With the new tRNA synthetases in hand we have repeated our previous experiments of the genetic encoding of **8-Lys**, **13-Lys**, **14-Lys** and **17-Lys** and the accompanying flow cytometry studies. However, the results were almost the same, i.e., efficient transfection with negligible increase of ncAA incorporation. Regardless, we have chosen the best apparent synthetases for **13-Lys** and **17-Lys**, and attempted the nuclear relocation experiment with NLS-GFP and NLS-mCherry proteins, using **2-Lys** and PCK tRNA synthetase as control (*Figure 5*). We have managed to reproduce the expected results with **2-Lys**, however **13-Lys** and **17-Lys** proteins showed no relocation (or any change for that matter) after photoactivation which once again can only be explained by the “leaky” nature of the tRNA synthetases.

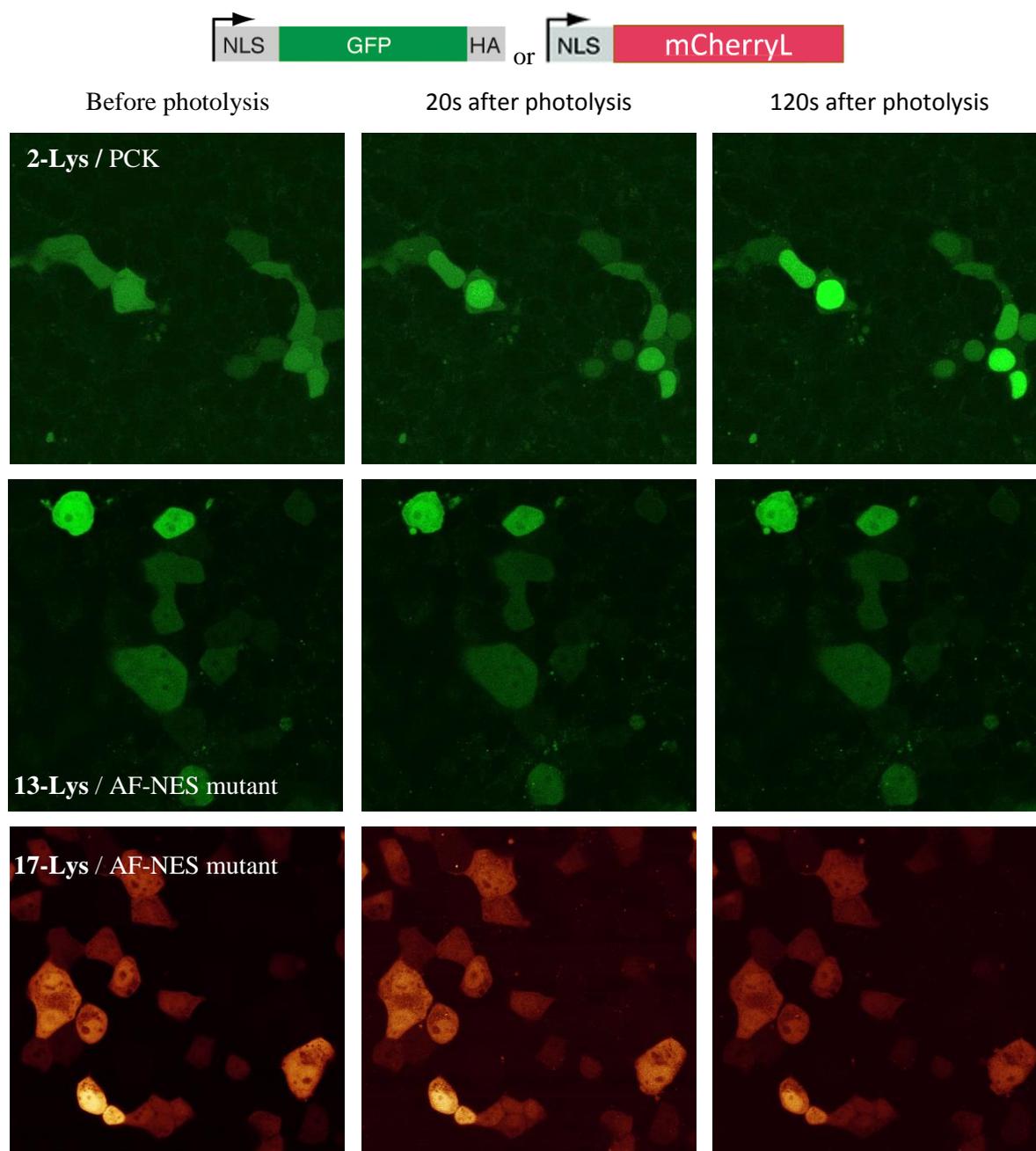


Figure 5. Fluorescence of NLS-GFP or NLS-mCherry proteins and their distributions before and after photolysis. Top: **2-Lys**; middle: **13-Lys**; bottom: **17-Lys**.

3. Bioorthogonal systems

3.1. Cyclopentadienone derivatives

Because of the lack of success of the genetic encoding of our photolabile amino acids, in the second part of the project we have turned towards another activation mechanism, and decided to involve such systems in our caged-Lys repertoire. In recent years considerable attention was focused on the fast and reliable in vivo release of caged compounds by so-called click-and-release mechanism (Carlson et al. *J.*

Am. Chem. Soc. **2018**, *140*, 3603). Among such chemically cleavable cages, tetrazine-triggered bioorthogonal elimination holds greatest promise, however the most reactive tetrazines are also the most unstable. Last year Ji et al. (*J. Org. Chem.* **2017**, *82*, 1471) reported a set of cyclopentadienone derivatives (**18**) which behave similarly to tetrazines, but are more stable and have the added bonus of inherent fluorogenicity. However, some features – i.e. UV excitation, poor water solubility, reaction kinetics – limit their efficient use in living systems. To improve these molecules and to provide an alternative to photocaged systems, we designed and synthesized new derivatives of such cyclopentadienones and studied their reaction with BCN (a strained cyclooctyne) and TCO (a strained *trans*-cyclooctene used as caging a group).

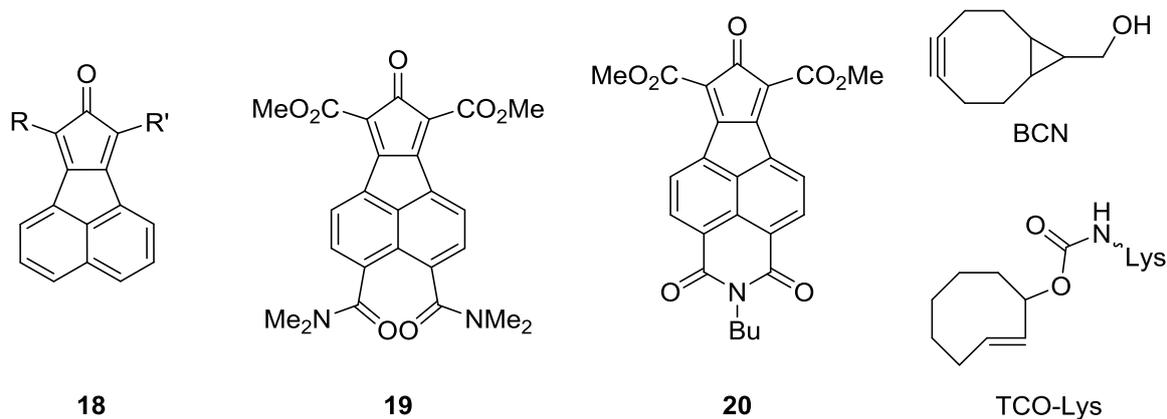


Figure 6. The original (**18**) and the cyclopentadienone derivatives (**19**, **20**) prepared by us and their strained reaction partners (BCN and TCO).

One of these new compounds had increased reaction speed with BCN and appreciably increased aqueous solubility (**19**). At the same time it has retained a virtually infinite fluorogenicity, although with excitability in the UV region (**19-BCN**). Another derivative (**20**) showed remarkable reaction kinetics with BCN and red-shifted excitation/emission of the fluorescent product (**20-BCN**) (*Table 2*). Unfortunately these outstanding properties were accompanied with low water solubility. As expected, reaction of **19** with TCO was even faster, however no significant click-and-release process was observed under physiological conditions.

Compound	k_2 with BCN / $M^{-1}s^{-1}$	Compound	$\lambda_{max}(abs)$ / nm	$\lambda_{max}(em)$ / nm
18	406	18-BCN	365	465
19	636	19-BCN	372	482
20	spontaneous	20-BCN	410	560

Table 2. Reaction rate and spectral properties of the cyclopentadienone derivatives under physiological conditions (10% DMSO in PBS + 0,1% SDS) at 25 °C.

Probe **19** was chosen for its favorable reactivity and solubility and was used in a bioorthogonal labeling scheme and compared to the original molecule **18**. First, BCN-modified phalloidin was conjugated to the actin filaments of COS-7 cells, then a 10 μM solution of the probe was added. The fluorescent images were taken without washing out excess dyes. In the control experiments unmodified phalloidin was used. As can be seen *Figure 7*, the improved properties of the new dye enable more efficient labeling compared to the original molecule.

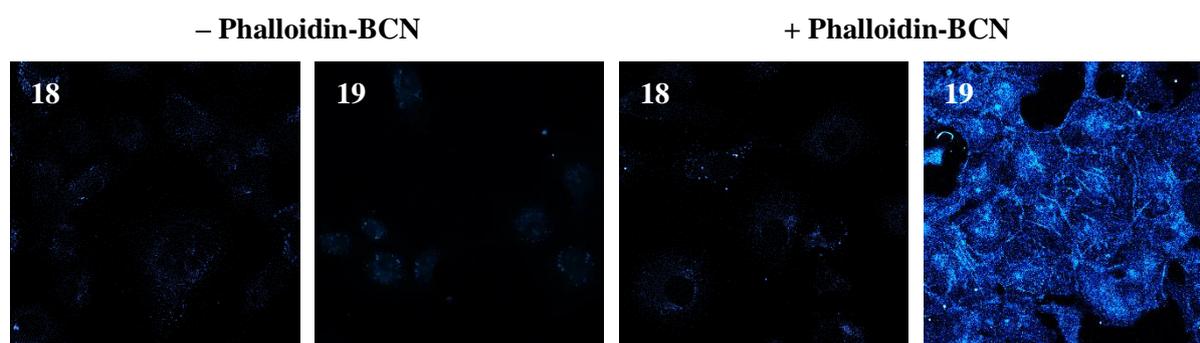


Figure 7. Actin labeling experiment with **18** and **19**.

We also examined the effect of the leaving group in the retro-Diels–Alder step by changing the carbonyl group to a sulfonyl function and identified new thiophene dioxide derivatives (**21**, **22**), which release sulfur dioxide during their fluorogenic reaction with cyclooctynes. We envisioned, that this later could later be used to release SO_2 in a controlled and quantitative manner inside cells to study its effects, e.g. its role in DNA methylation and epigenetics. Examination of the reaction between **14** and BCN proved the possibility of the controlled release of sulfur-dioxide, though at slower reaction rates ($k_2 \sim 2\text{-}20 \text{ M}^{-1}\text{s}^{-1}$). Sadly, during the project similar results were published in the literature by Wang et al. (*Chem. Commun.* **2017**, 53, 1370) / Ji et al. (*Org. Lett.* **2017**, 19, 818).

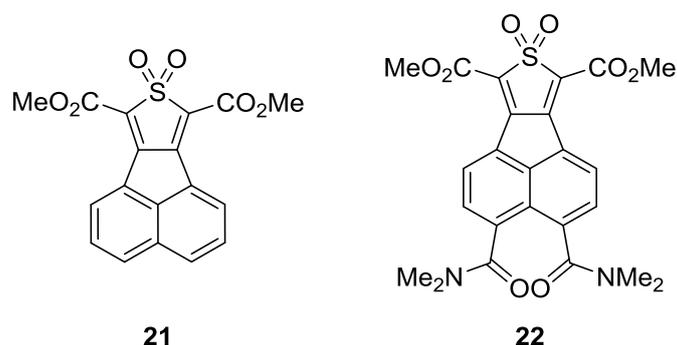


Figure 8. Structure of fluorogenic dyes capable of releasing SO_2 after click reaction with BCN.

3.2. Isonitrile

While we were looking for alternative bioorthogonal click-and-release mechanisms, Tu et al. (*Angew. Chem. Int. Ed.* **2019**, 58, 9043) disclosed a study on bulky isonitriles, which had preferential reactivity towards sterically demanding (and thus more stable) tetrazines and offered a new type of click-and-release reaction. This has prompted us to develop a non-canonical amino acid bearing a suitable isonitrile (NC) function and use our existing tRNA synthetase library to genetically encode the new ncAA. To prove successful incorporation and the possibility for mutually orthogonal bioorthogonal modifications we planned selective fluorescent labeling studies against traditional tetrazines and BCN. Thus we designed and prepared **23** lysine derivative and a sterically demanding nicotinic acid-derived tetrazine, which was coupled to rhodamine-piperazine and sulfo-Cy3 dyes to furnish a membrane-permeable probe **24** and the membrane-impermeant **25**, respectively (*Figure 9*).

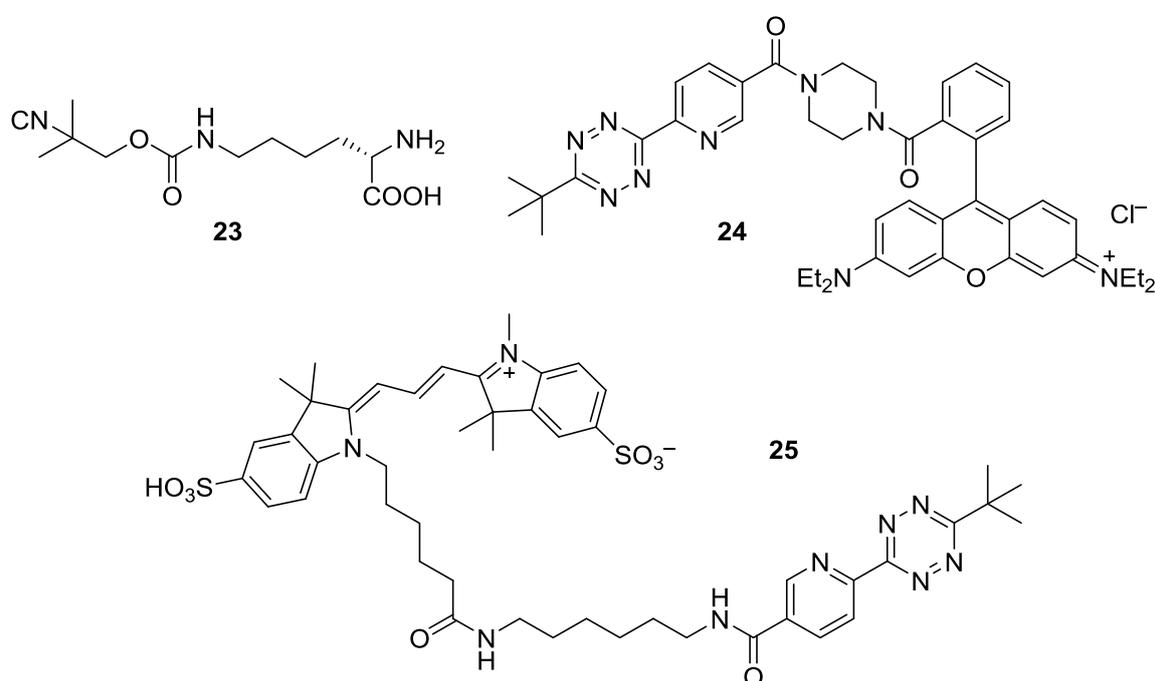


Figure 9. The isonitrile-lysine ncAA (**23**) and the two fluorescent dyes (**24**, **25**) equipped with our bulky tetrazine moiety.

To test mutual orthogonality of the cycloaddition of the bulky isonitrile group and the sterically demanding *tert*-butyl-tetrazine versus the inverse electron demand Diels–Alder reaction (IEDDA) of BCN and conventional tetrazines to label proteins, we selected Transferrin (TF, 76 kDa) and Bovine Serum Albumin (BSA, (66 kDa) because they are comparable in size, but still can be fully resolved on SDS-PAGE gels. We conjugated each protein with an isonitrile group or BCN and after removing the unreacted reagents we added a solution of a tetrazine functionalized silicon rhodamine probe (Tet-SiR) and **24** either separately or together to different combinations of bioorthogonalized, i.e., isonitrile- or

BCN-modified, proteins. Following 2 h incubation, the proteins were worked up using Sephadex G-25, and were immediately run on an 8% SDS-polyacrylamide gel (*Figure 10*).

The results indicated that Tet-SiR reacted specifically with BCN in either combination. Reaction of **24** with isonitrile, on the other hand, was partially specific, i.e., no cross reaction was observed with BSA-BCN, while a faint but visible band appeared with TF-BCN (*Figure 10b*, lane 4). Cross labeling was also present when a mixture of TF-BCN and BSA-NC was treated with **24** alone. However, when both bioorthogonalized proteins and both dyes were present, the two reactions were specific in each combination (*Figure 10b* lanes 5 and 8), indicating that cross reactivity is only present when the BCN moiety is not consumed in the fast IEDDA reaction. Gratifyingly, no unspecific binding of any of the dyes to any of the proteins was detected in the absence of NC and BCN (*Figure 10a,b*, lanes 12 and 13).

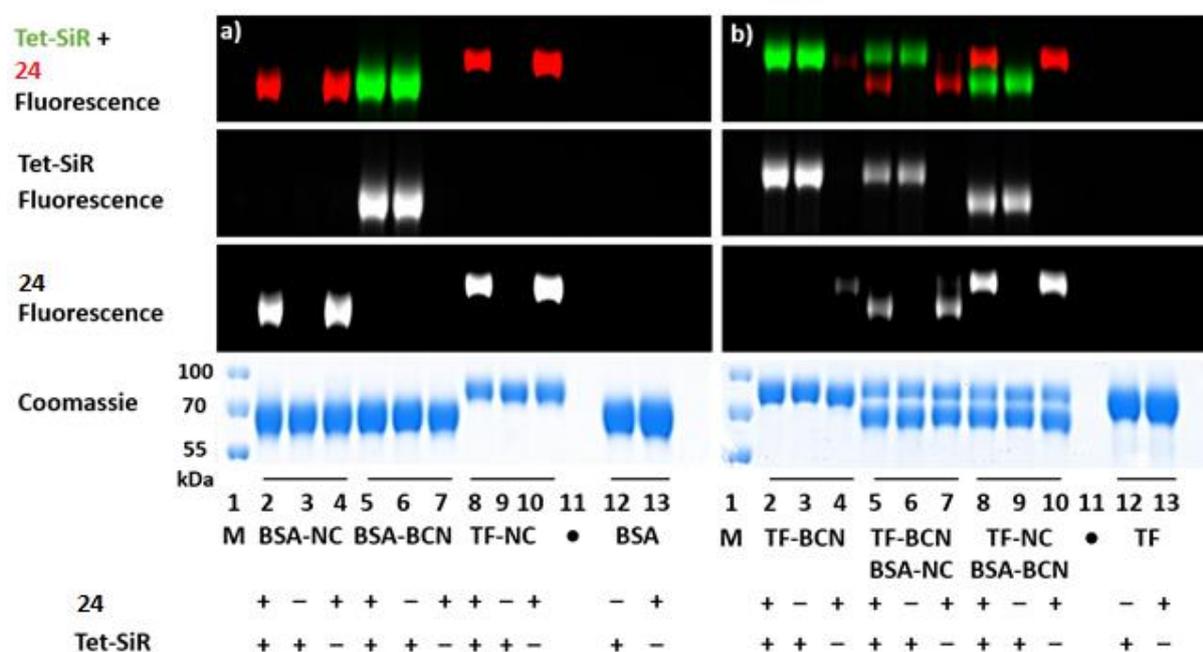


Figure 10. Bioorthogonal labeling of serum proteins Transferrin (TF) and Bovine Serum Albumin (BSA) to test mutual orthogonality between IEDDA and [4+1] cycloaddition represented by in-gel fluorescence imaging of SDS-PAGE gels.

To genetically incorporate the new **23** ncAA into mammalian proteins we used our previously established mCherry-TAG-GFP system containing an Amber STOP codon between a red (mCherry) and a green (EGFP) fluorescent reporter protein. Docking studies indicated that the AF-NES mutation of the pyrrolysine tRNA synthetase would be suitable. To our delight, microscopy analysis indicated the appearance of green fluorescence in a significant number of cells alongside red fluorescence, suggesting that **23** is efficiently incorporated into proteins (*Figure 11*, top). Flow cytometry analysis of cells was utilized to quantify the efficiency and specificity of ncAA incorporation (*Figure 11*, bottom). We

observed that the incorporation of **23** was similarly efficient to the known substrate BCN-Lys and significantly higher than negative control that contains no ncAA.

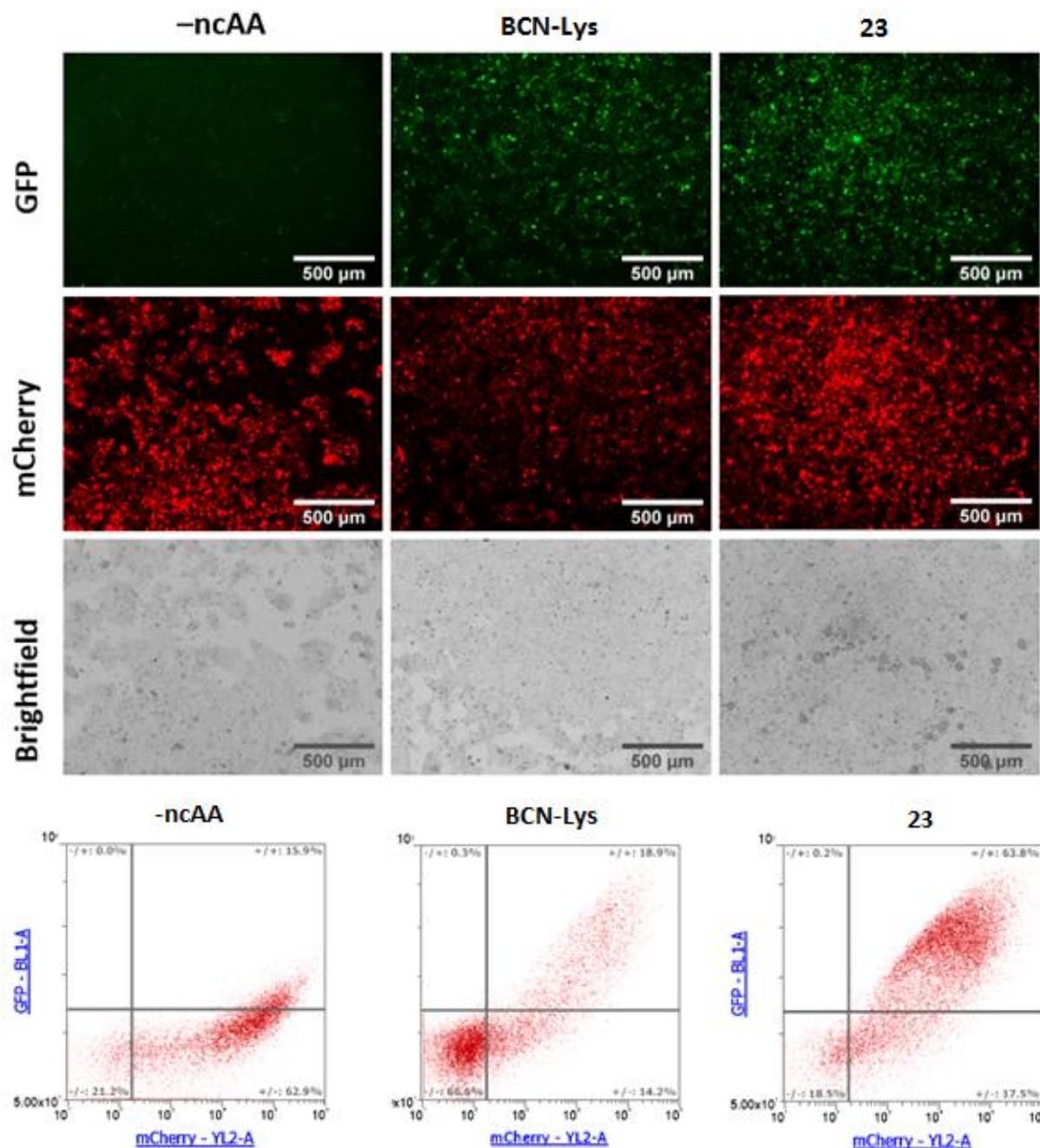


Figure 11. Top: Live cell fluorescent widefield imaging of specific incorporation of BCN-Lys (middle column) and **23** (right column) into pmCherry-TAG-EGFP in HEK293T cells. Bottom: Flow cytometry analysis to assess the efficiency of genetic incorporation of **23** (right) as compared to BCN-Lys (middle), or no ncAA added (left) by AF-NES.

To achieve double labeling of ultrastructures in fixed cells we recruited the help of the HaloTag system, i.e., an intracellular POI fused to the HaloTag self-labeling enzyme. This allowed easy functionalization of various structures such as vimentin and LaminA without with a HaloTag-BCN substrate. The

isonitrile group, on the contrary, was used to modify a secondary antibody (IgGNC) together with anti-TOMM20 primary antibody to label mitochondria. At this point each of the two bioorthogonal reactive groups (BCN and isonitrile) were ready to receive their reaction partners. First, we stained the specimens with Tet-SiR, then after extensive washing, we proceeded with staining with the sterically demanding, *tert*-butyl-tetrazine-rhodamine (**24**) probe. Confocal microscopy revealed successful dual-colour labeling of mitochondria and vimentin or LaminA using the orthogonal-bioorthogonal chemistry (Figure 12).

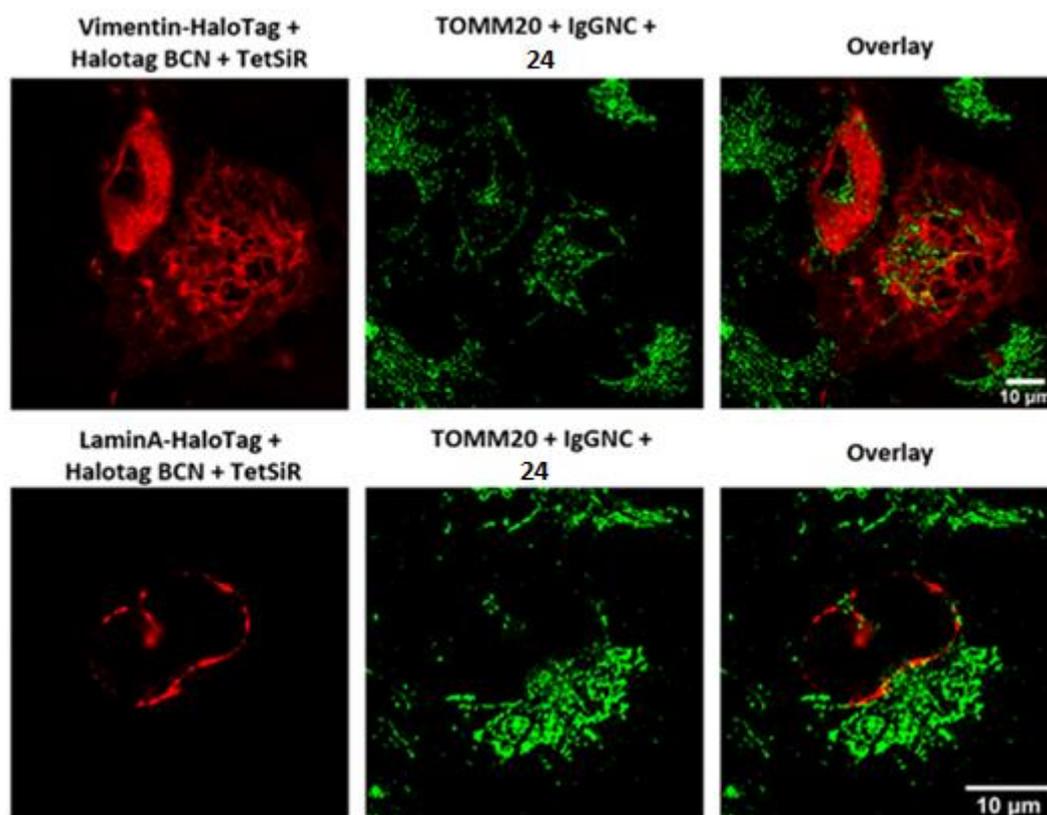


Figure 12. Confocal images of dual colour fixed cell labeling. Top: COS-7 cells expressing Vimentin-HaloTag construct; bottom: COS-7 cells expressing LaminA-HaloTag. False colours: Red: Tet-SiR labeling of HaloTag + HaloTag ligand-BCN tagged nuclear proteins; Green: Labeling of TOMM20 with **24** by cycloaddition between isonitrile/*tert*butyl-tetrazine groups.

For dual color live cell labeling we incorporated **23** into GFP-fused insulin receptor (IR) using genetic code expansion in HEK293T cells and introduced the second bioorthogonal moiety (BCN) through a plasmid encoding an intracellular protein of interest (POI) fused to the HaloTag self-labeling enzyme, similarly to the fixed-cell protocol. Our workflow is summarized in Figure 6a. As POIs we chose nuclear proteins, LaminA and H2B. Tet-SiR was added to label the first subcellular structure. This labeling reaction is rapid, resulting in specific and background-free labeling, and consumption of the BCN moieties. After extensive washing of the cells, the second probe, **25**, was added. Confocal microscopy imaging was without fixation. This way, we could carry out efficient double bioorthogonal labeling of

insulin receptors and histone H2B or LaminA in live HEK293T cells (*Figure 13*). We detected no unspecific labeling, which meant that the BCN moieties were fully consumed in the first step, resulting in specific labeling of H2B or LaminA at the inner lining of cell nuclei. Insulin receptors bearing the **23** ncAA were also specifically labeled in the following step. These results were published in a special issue of *Molecules*.

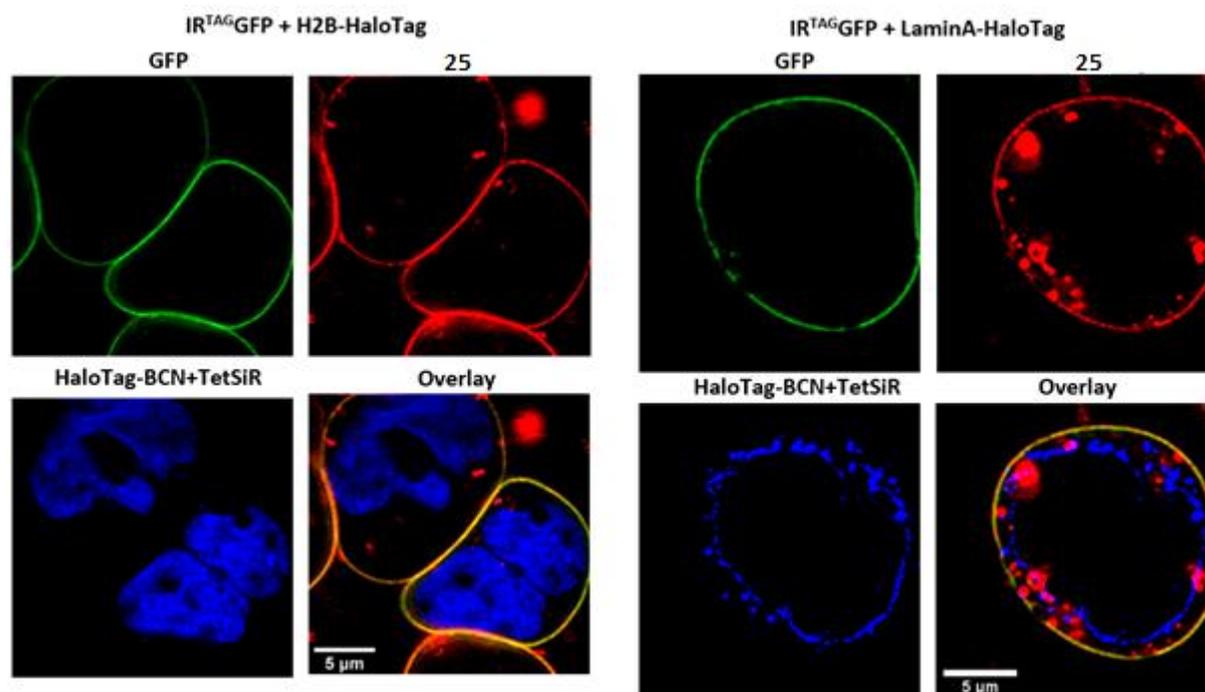


Figure 13. Confocal images of dual-color live cell labeling with the orthogonal-bioorthogonal reactions between the ncAA **23** genetically incorporated into IR-^{TAG}-GFP and **25**, and nuclear proteins expressed in fusion with the HaloTag self-labeling enzyme labeled with Tet-SiR through HaloTag-BCN. Left: H2B-HaloTag construct; right: LaminA-HaloTag construct was used. False color codes: Green: the reporter fluorophore GFP shows location of efficient IR expression; Red: labeling of IR with **25**; Magenta: labeling of cell nucleus related proteins with Tet-SiR; Yellow in overlays: green GFP and red **25** labeling of insulin receptors result in yellow color.

3.3. Pyrone derivatives

During our work with cyclopentadienone derivatives and search for new bioorthogonal reactions and click-and-release systems we have made an interesting observation that pyrones might be able to react with strained cyclooctenes and cyclooctynes in an inverse electron demand Diels–Alder reaction (IEDDA) and release carbon dioxide under physiological conditions. The literature only describes similar reactions with electron deficient alkynes at elevated temperatures or long reaction times and even higher temperatures for CO₂ elimination. To test our hypothesis we reacted methyl coumalate (**26**) with BCN

which indeed gave the desired **26-BCN** product. Although the reaction was slower than with tetrazines, pyrones later proved to be stable under physiological conditions.

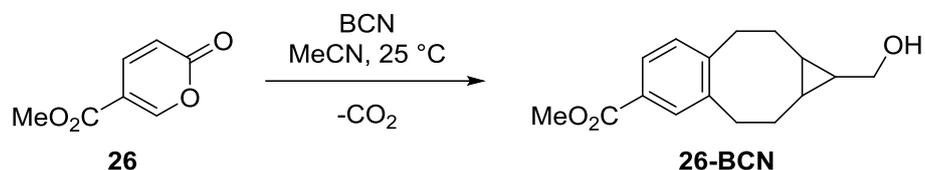


Figure 14. The preliminary reaction between **26** pyrone and BCN cyclooctyne.

To harness and test the feasibility of this new bioorthogonal reaction, we set out to create a set of fluorescent dyes furnished with a pyrone moiety. We also anticipated that the different nature of the pyrone and the benzene ring (which is formed in the IEDDA reaction) will dramatically change the spectrophysical properties of the probes and hopefully result in fluorogenic dyes. We have chosen the coumarin ring as the fluorophore core and attached the pyrone function to its 3- or 7-position directly or through vinyl or acetylene linker to synthesize compounds **27-30** (Figure 15).

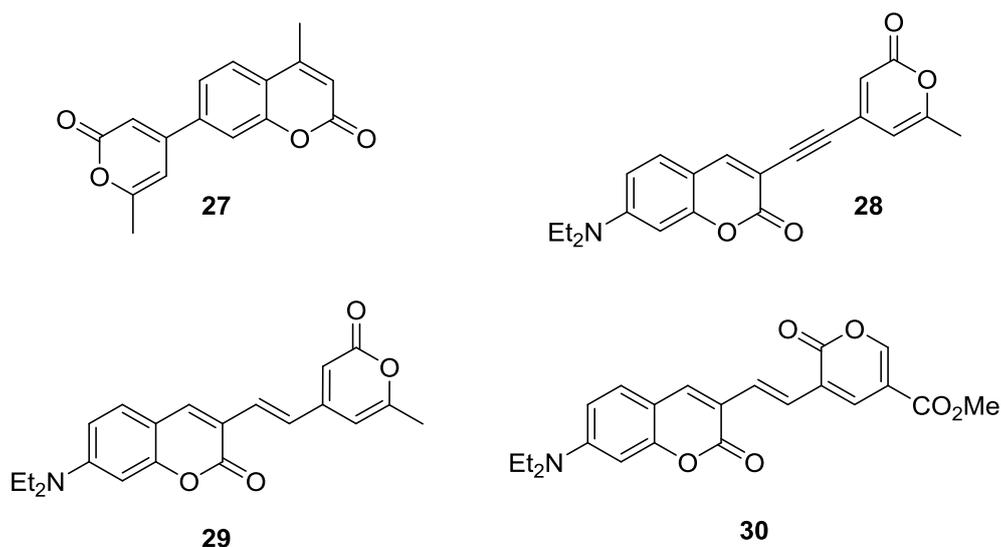


Figure 15. Pyrone-functionalized bioorthogonal coumarin probes.

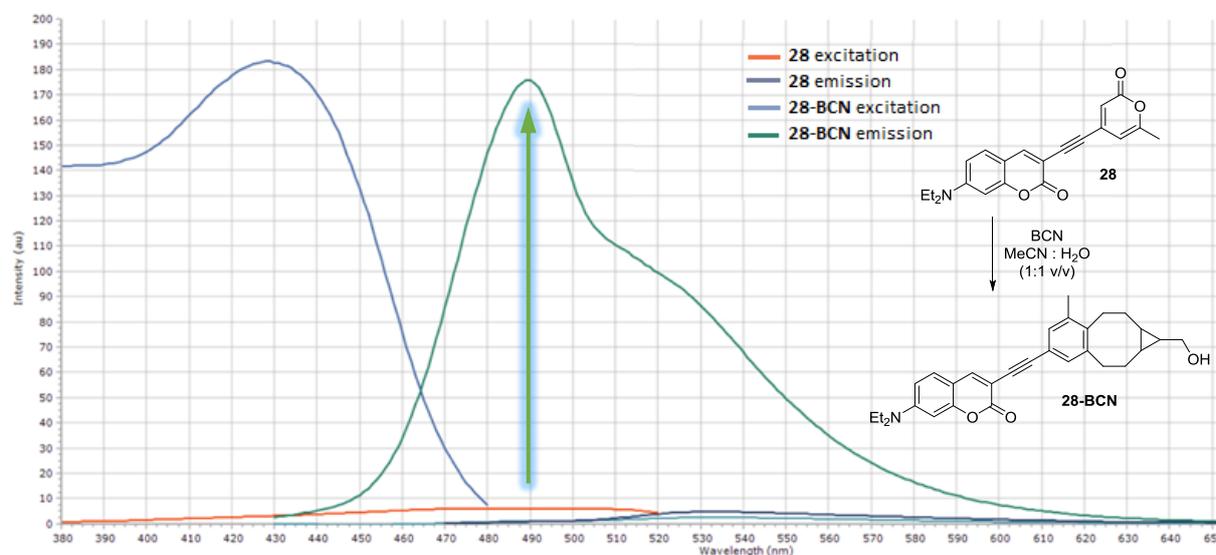
With the desired molecules in hand, we tested their reaction with BCN and monitored the changes in their fluorescent properties. The results (Table 3) showed that all the compounds reacted readily with BCN, however the change in the fluorescent properties varied considerably. Probe **27** and **27-BCN** showed similar absorption and emission spectra with significant fluorescence enhancement, however, required UV excitation, which is not ideal for later biological experiments. **28** showed blue-shifted excitation and emission upon reaction, accompanied by a huge increase in fluorescence. On the other

hand, reaction of **29** with BCN decreased the fluorescent intensity and the formerly large Stokes-shift disappeared. Changing the connection point and introducing an electron-withdrawing group onto the pyrone ring (**30**) resulted in a hypsochromic shift compared to **29**, but only a slight increase in fluorescence could be observed upon reaction.

	27	27-BCN	28	28-BCN	29	29-BCN	30	30-BCN
$\lambda_{\text{max}}(\text{ex}) / \text{nm}$	335	334	465	426	469	538	407	422
$\lambda_{\text{max}}(\text{em}) / \text{nm}$	429	424	536	487	593	565	490	490
<i>Change in fluorescence</i>	Increase (12-13 \times)		Huge increase (>100 \times)		Decrease		Slight increase (1.4-1.6 \times)	

Table 3. Excitation and emission maxima of the new fluorescent dyes and the change in fluorescence intensity upon reaction of BCN. Conditions: 50% MeCN in H₂O, 25 °C.

Based on these properties, we chose **28** for further biological experiments. We quantified its reaction rate with BCN ($k_2 = 0.095 \text{ M}^{-1}\text{s}^{-1}$) and also the molar absorption coefficients and fluorescence quantum yields in 50% MeCN in H₂O at 25 °C (*Figure 16*).



$\epsilon_{465\text{nm}}$ of 28 / $\text{M}^{-1}\text{cm}^{-1}$	ϕ of 28	$\epsilon_{426\text{nm}}$ of 28-BCN / $\text{M}^{-1}\text{cm}^{-1}$	28 / $\text{M}^{-1}\text{cm}^{-1}$	Fluorescence enhancement
52730	0.009	33150	1.0	111 \times

Figure 16. Spectral properties with ϵ and ϕ values for **28** and **28-BCN**.

To test the possibility of fluorogenic protein labelling with **28**, we conjugated Transferrin (TF, 76 kDa) with BCN. After removing the unreacted reagents, we added the solution of **24** in different

concentrations. Following overnight incubation, the proteins were worked up using Sephadex G-25, and were immediately run on an 8% SDS-polyacrylamide gel (*Figure 17*). As control experiments we used native Transferrin without BCN.

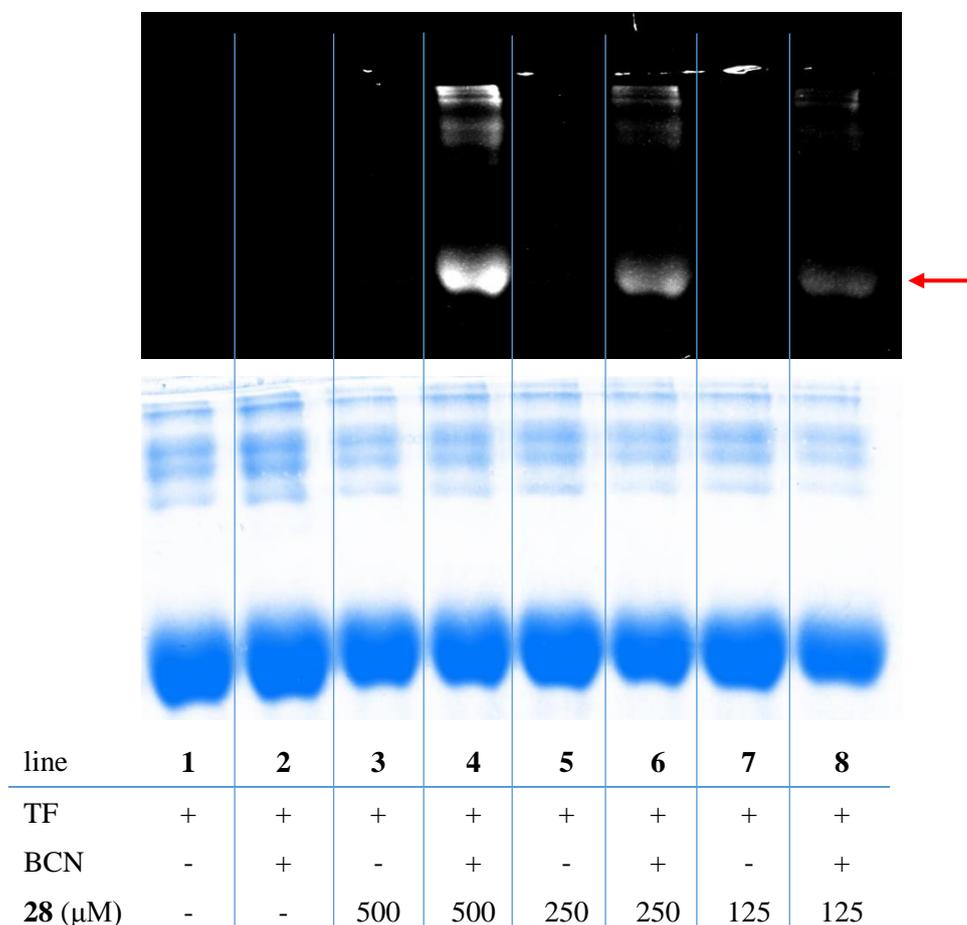


Figure 17. Bioorthogonal labeling of serum protein Transferrin (TF) with **28**. Line 1 contains only the protein; lines 2, 4 and 6 contain TF + **28**; lines 3, 5 and 7 contain BCN-functionalized TF + **28** in varying concentrations. Excitation: 488 nm.

As can be seen on the fluorescent imaging of the gel, the labelling was successful and only minimal aspecific signal was observed.

PRESENTATION OF RESULTS

The results of the isonitrile labelling were published in *Molecules* **2021**, *26*, 4988.

Our findings on pyrones are summarized in a manuscript which is about being submitted to *Synthesis*.

Results have been presented at the following conferences:

EMBO Workshop: Chemical Biology 2018, Heidelberg, Germany (poster presentation);

Annual Heterocyclic and Bioorganic Workshop in Balatonszemes (oral presentation);

6th ECBS/LS-EuCheMS Symposium (2019) in Madrid, Spain (poster presentation);

EMBO | EMBL Symposia: Seeing is Believing - Imaging the Molecular Processes of Life (2019) in Heidelberg, Germany (poster presentation).

During the funding period of the project 1 MSc and 3 BSc degrees were obtained and 2 students participated in the National Scientific Students' Associations Conference (TDK/OTDK).