Final report for NKFIH NN-116265: Development of new fluorogenic dyes for super-resolution microscopy of site-specifically engineered proteins

A. Genetically encodable non-canonical amino acids:

We have developed hydrophilic *trans*-cyclooctenes (TCO) as dienophile scaffolds for inverse electron demand Diels-Alder (IEDDA) reactions with tetrazines in bioconjugation schemes. Non-canonical amino acids (ncAAs) bearing TCOs with hydrophobic characteristics have a tendency to stick non-specifically onto non-polar surfaces inside cells. Even after vigorous washing cycles the remaining, non-specifically bound TCO-amino acids contribute to significant background labeling upon conjugation even with fluorogenic dyes. We have synthesized oxazolone-fused (OxTCO) and dioxo-TCO (DOTCO) as new TCO derivatives with low cLogP features and prepared their Lys conjugates. For comparison, we also included two regularly used dienophile modified ncAAs in these studies: TCO*-Lys and BCN-Lys together with a cyclooctyne (COMBO) scaffold that was formerly developed in our group. The TCO derivatives were conjugated to the e-amino group of Lys via carbamate linkage whereas COMBO was attached to the same amino acid function through an amide bond. Next, we have evaluated the genetic incorporation abilities of the two new Lys derivatives using an amber mutant GFP variant. Results showed that OxTCO-Lys and COMBO-Lys were not expressed in the model system, however, DOTCO-Lys was efficiently recognized by the enzyme and expressed efficiently. We also made a variant of OxTCO where the linkage between the TCO frame and Lys was shortened (sOxTCO), but it also failed. Thus we have concentrated onto DOTCO-Lys in further studies. Following satisfactory reaction rate and stability studies we have incorporated DOTCO-Lys into insulin receptor membrane proteins and labeled them with fluorescent tetrazine (Cy5-H-tet). Specific labeling confirmed the suitability of DOTCO in such biorthogonal tagging schemes. Most importantly we wanted to elaborate the clearance characteristics of the hydrophilic TCO modified Lys. To this end we performed a comparative washout assay with BCN-Lys, which was previously found to offer optimal performance in such assays and TCO*-Lys another often used non-canonical amino acid (ncAA). Cell culture medium containing excess ncAA was removed at different time points from nontransfected COS7 cells and residual cytoplasmic ncAA was detected in live cells after labeling for 10 min with 5 uM cell-permeable tetramethylrhodamine derivative (5-TAMRA-H-tet). This concentration was sufficient to observe the full amount of residual ncAA, but the reaction time was short enough to avoid excessive background staining by TAMRA (an often encountered problem in live-cell labeling). Efficient removal of excess TCO*-Lys from the cytoplasm was not possible, even after 6 h of washing. BCN-Lys gave almost background-free labeling after 6 h, in line with its almost one order of magnitude lower clogP value than TCO*-Lys. DOTCO-Lys on the other hand, was easily removed from the cytoplasm after 5 min washing; the fluorescence signal was identical to TAMRA background. Development of such hydrophilic-ncAAs is an important step that paves the way towards direct imaging after protein expression using genetic code expansion combined with rapid click labeling. While there is an obvious need for fluorogenic dyes for cell labeling methods based on reactions with small dyes, the described method and cell-permeable fluorogenic dyes in hand will greatly facilitate labeling schemes for fluorescence microscopy and even SRM, and will enable the study of numerous intracellular proteins, including those that turn over rapidly. (Kozma et al. *ChemBioChem*, **2016**, *17*, 1518 - This work was chosen to appear in the inside cover of the respective issue).

B. Fluorogenic probes

In line with our focus to *trans*-cyclooctene dienophiles we have started the development of fluorogenic scaffolds bearing tetrazine motifs. We demonstrated that near-infra red emitting phenoxazine scaffolds linked to tetrazines via a phenylene or vinylene linker are extremely efficient fluorogenic probes. The quenching in these case is the most efficient mode to diminish fluorescence as it works via through-bond-energy-transfer (TBET) mechanism. The developed phenoxazines turned out to be membrane permeable, NIR emitting and offered background-free fluorescence imaging of genetically altered (e.g. trans-cyclooctenylated) proteins (Knorr et al. *Chem. Eur. J*. **2016**, *22*, 8972).

In order to minimize autofluorescence we have chosen fluorescent frameworks that are excitable in the red-near infrared (NIR) range of the electromagnetic spectrum. In these studies different cyanine-type frames were selected (Cy3, Cy5; and Dy630, Dy650). These frames were rendered fluorogenic and bioorthogonally applicable at the same time by installing tetrazine motifs via conjugated but electronically decoupled (twisted) linkers (phenyl or vinyl).
All derivatives showed moderate fluorogenicities (4-10 times increase) in contrast to formerly tested dyes with more compact structures. These studies suggest that fluorescence of more conjugated systems are less modulated by the presence of the quencher moieties.
Regardless of the above results Cy3 and Cy5 derivatives were tested in live cell labeling studies using cells that expressed membrane proteins (i.e. insulin receptors, IR) genetically modified with GFP and a non-canonical amino acid bearing *trans*-cyclooctene. Results showed that vinyl-linked Cy-derivatives are suitable for such labeling schemes. Phenyl-linked Cy3 and Cy5 derivatives on the other hand passed through the membrane thus no specific membrane-protein labeling could be seen. However, these latter species can be used in intracellular protein tagging in the future. (manuscript is under preparation).
Dy630 and Dy650 labels were also tested for fluorogenicities and likewise moderate (5-6 × increase in fluorescence intensity) fluorogenicities were detected. Isolated human serum albumin (HSA) was modified with cyclooctyne (BCN) and reacted with Dy630- and Dy650-tetrazines. Gel electrophoretic studies showed that only cyclooctyne bearing HAS was labeled with the dyes.
These studies were the subject of a BSc diploma work and a manuscript is under preparation).

In the last year of the funded period we have accomplished the synthesis, spectroscopic and biological evaluation of tetrazine quenched, bioorthogonally applicable fluorogenic probes. Cyanine frames (Cy3 and Cy5) were rendered bioorthogonal and fluorogenic by installation of a tetrazine moiety via phenylene or vinylene linkers. Spectroscopic evaluation revealed that one of the four derivatives can reach up to 14 fold fluorescence increase upon reaction with cyclooctynes. We have tested the specific labeling potentials of the dyes using phalloidin-cyclooctyne tagged actin filaments. While Cy5 dyes showed no specific staining, both Cy3 derivatives labeled specifically the actin filaments. Furthermore, the probe with the highest fluorogenicity showed only negligible background, which prompted us to further test it upon no-wash conditions. To our satisfaction, no substantial difference was observed between the images acquired under wash or no-wash conditions. This suggests that the over 10-fold fluorescence enhancement enables efficient distinction between reacted and unreacted forms with a good signal-to-noise ratio. Although indocyanines are not preferred in STED-based super-resolution techniques, we have tested the applicability of the best performing dye in STED imaging of actin filaments, using a 660 nm continuous wave laser for depletion. To our delight, it was found to be suitable in STED and resulted in subdiffraction imaging of the labeled actin meshwork at resolutions similar to reported values. Results were published in Bioconjugate Chemistry (Knorr et al. *Bioconjugate Chem*. **2018**, *29*, 1312).

C. Double fluorogenic systems

As it was observed fluorogenicities decrease remarkably towards the red range of the excitation spectrum, we addressed this problem by introducing the concept of multiple fluorogenicity. To this end we have elaborated the effect of more quenchers (azide or tetrazine) or the combination of more quenching mechanisms.

We have introduced a new, bisazide, thus bis-quenched fluorogenic scaffold, which enables more specific two-point binding anchoring to bis-cyclooctynylated motifs with much higher fluorogenicity (i.e. 200 fold fluorescence increase upon binding). In these studies a suitable peptide motif containing two cyclooctyne functions was also identified. Such short peptide motifs in combination with bisazide fluorogenic probes can be good surrogates for tetracysteine-containing short peptide tags in combination with biarsenical probes as bisazides and biscyclooctynylated motifs do not require reducing environment, are not toxic and free of any competing side reactions. (Demeter et al. *Chem. Eur. J*. **2016**, *22*, 6382).

Upon continuation of this topic, we have extended the concept to biologically more relevant frames, i.e. to cyanine frameworks. We designed bisazide-Cy3 probes and meanwhile identified another peptide sequence that was modified with a cyclooctyne dubbed BCN at the terminal Lys residues (Ac-K(BCN)AEAADAEAAK(BCN)-NH2) as target platform. Upon reaction of the bisazides and the peptide, all dyes were found to be fluorogenic and gave 12-, 28-, 39- and 11-fold increase relative to the original fluorescence of the dark forms. We also investigated the stability of these bisazide probes in the presence of various thiols as possible reducing agents under physiological conditions. Azide-quenched dyes are often prone to thiol mediated reduction that can lead to fluorescent products. We found that the weak fluorescence intensity of the bisazides did not change substantially in reducing media such as excessive amounts of ethanethiol, cysteine or glutathione especially in comparison to the intensive signal evolution upon azide-alkyne reaction with strained alkyne reaction partners.
To further demonstrate the potential of our bisazides, we sought to explore crosslinking of site-specifically modified proteins. To this end, we first expressed GFP-6His harboring a single TAG (Amber mutation) at position 39 in E.coli. In each case an intensely fluorescent band appeared in SDS PAGE gel analysis at a size double of GFP, whereas control experiment with TAMRA-azide showed only the formation of the monomeric labeled species. This suggests that the bisazides can indeed be used as fluorogenic crosslinkers to follow protein-protein interactions.

Our further aim was to use these bisazide-cyanine probes in protein labeling schemes on proteins, genetically tagged with our biscyclooctynylated peptide tag. We have therefore expressed GFP that was fused with K(BCN)AEAADAEAAK(BCN) at its C-terminus. However, the azide-alkyne cycloaddition turned out to be too slow to achieve appreciable labeling within a reasonable timeframe (Demeter et al. *Bioconjugate Chem.*, **2017**, *28,* 1552). Therefore we have designed and synthesized cyanines that were modified with two tetrazines as quenchers and bioorthogonal moieties.
We have tested the fluorogenicities of these compounds against Ac-K(BCN)AEAADAEAAK(BCN)-NH2. To our delight 10-55 × fluorescence increase was observed and the reactions were indeed 3 orders of magnitude faster than with bisazides. These double-quenched, bioorthogonally applicable fluorogenic bistetrazine-cyanine probes possessed emission maxima between 600–620 nm and excitation wavelengths matching well with sources of fluorescent microscopes. We have studied the fluorogenic potential of the probes upon reaction with a biscyclooctynylated peptide and observed 10-fold enhancements in fluorescence. This increase is considered quite remarkable among fluorogenic Cy3 dyes and might be sufficient to allow application in live cell imaging. Unlike the formerly tested bisazide cyanine probes these bistetrazines were not that easily prone to degradation by daylight. We explored the labeling potential of these probes on a double-tagged protein, where two bioorthogonalized non-

canonical amino acids (ncAAs) were implemented via Amber suppression technology. Experiments indicated formation of the right cyclic probe–protein conjugate. Since, Amber-tagging with bioorthogonalized ncAAs is not restricted to N- or C-terminal modifications, these platforms enable mutations virtually at any optimized location. Thus, we envision that this concept, upon further optimization, enables the exploration of enzyme activation/inactivation, conformational changes using minimally perturbed engineered proteins etc. We also studied the cross-linking potential of these fluorogenic, bifunctional platforms in biorthogonal labeling schemes on a monocyclooctynylated model protein and concluded that they can be useful tools for such purposes. We believe that these fluorogenic covalent cross-linking bistetrazine probes are ideal choices to directly visualize protein–protein interactions even by super-resolution microscopy. In comparison to bimolecular fluorescence complementation (BiFC) technology, which is one of the most popular means to follow protein communications the presented bistetrazine probes in combination with Amber suppression technology are ideal surrogates. Unlike BiFC, where the original function is considerably perturbed by a comparable fusion tag and signal evolution is limited by the time required for assembly of the signaling fusion protein, the presented scaffolds require minimal perturbation in the structure of the protein studied and the fast kinetics of the inverse electron demand Diels-Alder reaction ensures real-time monitoring of protein–protein interactions. Results were published in Chemistry a European Journal (Kormos et al*. Chem. Eur. J*. **2018**, *24*, 8841).

A different approach for multiply quenched systems was also explored, where reactivity based fluorogenicity was combined with polarity based fluorogenicity. Along these lines we have designed and synthesized azido-dyes. These frames are known for their structural-change based fluorogenicity especially in the context of DNA binding, while azide moiety is responsible for the reactivity based fluorogenicity. Comparative studies with the azides and their non-azide congeners in the absence/presence of DNA and cyclooctyne reaction partners revealed that both the presence of DNA and cyclooctyne is necessary to trigger the appearance of fluorescence. One of the dyes had a remarkable 720 × fluorescence intensity increase in the presence of both. This topic was the base of a BSc diploma work and an account on these results is under preparation.

The same concept was applied in the design of a series of siliconrhodamine (SiR) derivatives. SiRs are widely used membrane-permeable NIR dyes for SRM applications. Besides their high photostability and brightness, the unique environment-dependent fluorescence of carboxy-SiRs due to a polarity dependent lactone-formation offers an appealing opportunity to distinguish between specific and non-specific labelling, (polarity-based fluorogenicity). When bound to polar protein surfaces, carboxyl-SiRs exist in a fluorescent zwitterionic form, while upon non-specific adhesion to hydrophobic surfaces, it closes to non-fluorescent spirolactone. We aimed at combining this polarity based fluorogenicity with reactivity based turn-on features by installment of bioorthogonally applicable tetrazine unit. Electron-withdrawing groups are known to markedly increase the reactivities of tetrazines in inverse electron demand Diels-Alder (IEDDA) reactions, thus we also included trifluoromethyl-phenyl derivatized tetrazines in the study. We also synthesized -NEt2 substituted derivatives. To our delight, all derivatives showed much larger fluorescence enhancements (FE) than formerly reported fluorogenic tetrazine NIR dyes (cf. 5.6×). We propose that this is attributed to the double fluorogenic nature of our SiR dyes. As expected, ethyl substitution on the N-atom improved turn-on rates of the dyes. The second order rate constants were also determined and were in good agreement with k2 values of previously reported methyl-tetrazine derivatives with a similarly reactive trans-cyclooctene (TCO*). As expected, trifluoromethylphenyl derivatives reacted much faster than methyl derivatized congeners with TCO. Due to the lower electron density of the tetrazines in these latter species

reactions went to completion within minutes compared to methyl substituted probes where reaction required ca. 1.5 h. We also investigated the polarity dependent absorbance of the Diels-Alder product of alkene-SiR. In agreement with previous findings, in solvents with lower dielectric constants, absorption maximum, characteristic of the spirolactone form at 290 nm was dominant. In polar solvents, however, the zwitterionic form with a characteristic absorption at 645 nm dominated the spectrum. As a consequence, in solvents with dielectric constants <50 the conjugate was non-fluorescent while fluorescence gradually reinstated with increasing polarity.
With the SiR-dyes in hand, showing exceptional fluorogenicity in the NIR region, we were eager to evaluate their performance in cellular environment. We used the previously reported skeletal protein vimentin116TAG-mOrange to test the intracellular labelling ability of the dyes. Using orthogonal tRNAPyl/PylRSAF pair we genetically encoded commercially available cyclooctynylated-lysine into vimentin116TAG-mOrange and performed live cell labelling experiments with the dyes in mammalian cells. Clear live cell labelling was observed with two of the dyes. The labelling was specific with no strong background fluorescence. Although CF3-dyesreacted faster in vitro, they did not give any specific labelling in vivo. We attribute this to the low membrane permeability of these dyes.
To date, there are only a few reports on bioorthogonally applicable, fluorogenic tetrazine-mediated intracellular labelling inside living cells and subsequent superresolution microscopy (SRM). Among them, to the best of our knowledge, there are no reports of site-specific labelling with double-fluorogens. We set out to test the suitability of two of our dyes in SRM of vimentin. One of the dyes was suitable for SRM and an enhanced resolution (28 nm) was observed.
Results of this latter study was published in Chemical Communications (Kozma et al. *Chem. Commun.* **2017**, *53*, 6696) and was selected as a cover picture. Due to remarkable feedbacks we were invited to include these results in a protocol paper: Kozma et al. *Methods Mol. Biol*. **2018**, *1728*, 337, as well.

D. Heterobifunctional linkers

Part of this project, bifunctional linkers (called chemical reporters) were also developed. These chemical reporters enabled specific targeting of rare amino acid side chains (cysteine) with specific warheads on the one hand, and carried a biorthogonal function on the other (cyclooctyne, azide, terminal alkyne). In various studies we have demonstrated the use of such linkers in two step, sequential protein labeling schemes using fluorogenic labels (Söveges et al. *Org. Biomol. Chem*. **2016**, *14*, 6071). We demonstrated the synthesis and use of a heterobifunctional linker that enables highly specific and efficient modification of proteins via free-sulfhydryl functions of cysteine residues through a methylsulfonyl-oxadiazole warhead. The cyclooctyne functionalized bifunctional methylsulfonyl linker readily creates a bioorthogonal platform that offers easy and high-yielding modification with azide or tetrazine modified fluorescent or fluorogenic probes. We have used this heterobifunctional linker to screen for possible FRET pairs from our existing bioorthogonalized fluorescent pool. Following identification of a promising FRET pair, we have used this Cys-specific linker to modify a protein (p38C162S) and its interacting partner MK2 or its docking motif pepMK2 with a FRET donor and acceptor, respectively. Tandem-MS measurements verified the selective and specific labeling of both protein components. FRET studies involving donor and acceptor labeled p38C162S–MK2 or p38C162S–pepMK2 partners confirmed the specific interaction of the proteins. Such heterobifunctional linkers are suitable for the selective modification of proteins possessing accessible free sulfhydryl functions. On the other hand, the high FRET efficiency between the identified dye pairs (ca. 80%) is suitable for the detection of protein-protein interactions and allowed estimation of the strength of affinity. Results were published in Organic and Biomolecular Chemistry and the paper was also selected to appear on the cover of the respective issue. (Söveges et al. *Org. Biomol. Chem*. **2018**, *16*, 5756).
Part of this project we have established an oligonucleotide based system suitable for the fast screening

for FRET pairs. We designed bioorthogonalized oligonucleotide strands based on reported sequences as platforms suitable for fast and efficient labeling by azide or tetrazine modified dyes. Not only are the selected sequences suitable for elaboration of FRET efficiencies, the length of the oligonucleotides also allows for estimation of Förster radii. We have used these platforms to test selected, bioorthogonally applicable fluorescent probes from our existing dye-set. Results indicated that the selected dyes exhibit strong FRET signals (70-90%) when the dyes are situated at proximal termini. As expected, decreased FRET efficiencies were detected when the distal termini were labeled. We have also examined the changes in FRET efficiencies as function of temperature and concluded that in case of the double stranded oligos labeled at the proximal ends the Tm shifted towards higher temperatures presumeably due to attractive interactions between the dyes. Specificity of the FRET signals was affirmed by competitive studies using a non-labeled counterstrand. Regarding the pairs tested, the use of Cy3Tetrazine-Cy5Ttetrazine pairs is proposed when decrease in fluorescent intensity of the donor dye is desirable upon interactions between macromolecules, while Cy1Azide-CBRD1Azide or Cy1Azide-Cy3Tetrazine pairs are suggested as primary choices when evolution of acceptor signal is to be detected. The Cy1Azide-Cy3Tetrazine pairs were further tested in live cells. Confocal microscopy images showed no considerable cross-talk between the donor and the acceptor, while a remarkable FRET signal was observed in HEK293T cells transfected with complementary DNA oligomers labeled fluorescently with the donor and the acceptor dyes, respectively. The evolution of FRET effect was further proved by the donor bleaching kinetics in the presence of the acceptor. Furthermore, the Cy3Tetrazine labeled oligomers were applied in FISH detection of H. pylori in gastric biopsy samples showing further use of the two-step –bioorthogonalization and bioorthogonal labeling– modification scheme. Results were published in Organic and Biomolecular Chemistry (Petrovics et al. *Org. Biomol. Chem*. **2018**, *16*, 2997).

E. Synthesis of key precursors for fluorogenic dyes

As a side project we have reported a synthetic route that allows efficient synthesis of 7-azido-3-formylcoumarin a key compound in the synthesis of fluorogenic dyes. The synthesis applies a two-step sequence of a Wittig-reaction and photoisomerization process, which can even be carried out in on pot. Results were published in Journal of Heterocyclic Chemistry (Pünkösti et al. *J. Het. Chem*. **2018**, *55*, 1183).

F. Other results:

Within the grant period two PhD degrees were obtained and two other were submitted. Furthermore, and 5 MSc, 5 BSc works were prepared. Works were also presented at international and national conferences and workshops.

All in all, during the funded 3-year-period we have published 13 papers (12 reserach and 1 invited review – Kozma et al. *ChemBioChem* **2017**, *18*, 486) in prestigious journals (ΣIF=48,788). Three of these publications were also selected to appear as covers (ChemComm, OrgBiomolChem, ChemBioChem). 2 graduate students successfully defended their theses and two others submitted their works. We have attended in several conferences with oral (3 invited lectures) and poster presentations.