Protein Ligation

A Bioorthogonal Click Chemistry Toolbox for Targeted Synthesis of Branched and Well-Defined Protein–Protein Conjugates**

Mathis Baalmann[+], Laura Neises[+], Sebastian Bitsch, Hendrik Schneider, Lukas Deweid,
Philipp Werther, Nadja Ilkenhans, Martin Wolfring, Michael J. Ziegler, Jonas Wilhelm,
*Harald Kolmar, and Richard Wombacher**

Abstract: Bioorthogonal chemistry holds great potential to generate difficult-to-access protein–protein conjugate architectures. Current applications are hampered by challenging protein expression systems, slow conjugation chemistry, use of undesirable catalysts, or often do not result in quantitative product formation. Here we present a highly efficient technology for protein functionalization with commonly used bioorthogonal motifs for Diels–Alder cycloaddition with inverse electron demand (DA_{inv}). With the aim of precisely generating branched protein chimeras, we systematically assessed the reactivity, stability and side product formation of various bioorthogonal chemistries directly at the protein level. We demonstrate the efficiency and versatility of our conjugation platform using different functional proteins and the therapeutic antibody trastuzumab. This technology enables fast and routine access to tailored and hitherto inaccessible protein chimeras useful for a variety of scientific disciplines. We expect our work to substantially enhance antibody applications such as immunodetection and protein toxin-based targeted cancer therapies.

Introduction

Bioorthogonal chemistries provide ample opportunities to be applied in the manufacturing of therapeutic biologics due to their biocompatible properties. Further they possess power

[*] Dr. M. Baalmann,[+] L. Neises,[+] P. Werther, N. Ilkenhans,
M. Wolfring, M. J. Ziegler, J. Wilhelm, Dr. R. Wombacher
Institute of Pharmacy and Molecular Biotechnology, Heidelberg
University
Im Neuenheimer Feld 364, 69120 Heidelberg (Germany)
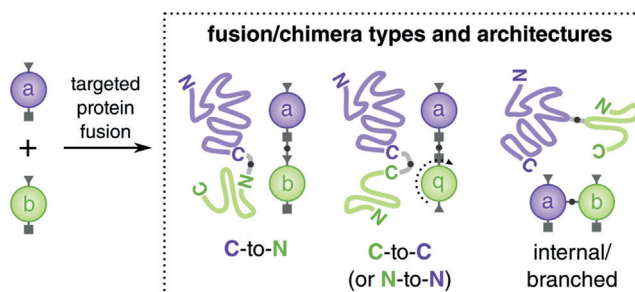E-mail: wombacher@uni-heidelberg.de

S. Bitsch, H. Schneider, L. Deweid, Prof. Dr. H. Kolmar
Institute for Organic Chemistry and Biochemistry, Technische
Universität Darmstadt
Alarich-Weiss-Straße 4, 64287 Darmstadt (Germany)

[+] These authors contributed equally to this work.

[**] A previous version of this manuscript has been deposited on
a preprint server (https://doi.org/10.26434/chemrxiv.10743344.v1).

📖 Supporting information and the ORCID identification number(s) for
ⓘD the author(s) of this article can be found under https://doi.org/10.
1002/anie.201915079.

to generate fusion types that are challenging to access using common protein expression or conjugation technologies (Scheme 1).[1] This is particularly the case for proteins that require special expression systems and thus limits the production as a fusion protein to a certain expression host, like for example, fusion proteins with toxic domains[2] or vaccines with non-proteinogenic features.[3]

To fully unleash the potential of bioorthogonal chemistry for the production of advanced biologicals, the method of conjugation needs to be highly efficient, quantitative and free of undesired side products to obtain homogeneous products—essential criteria to meet the high requirements of pharmaceutical manufacturing.[4]

Among all bioorthogonal reactions, the Diels–Alder cycloaddition with inverse electron demand (DA_{inv}) between alkenes or alkynes (dienophiles) and N-heteroaromatic compounds (dienes) has become the most promising chemoselective bond-forming chemistry in terms of catalyst-free reaction conditions and reaction rates.[5] It is thus not surprising that DA_{inv} has already been adapted for the generation of antibody–drug conjugates (ADCs),[6] investigated for diagnostic antibody radiolabeling[7] or related tumor pretargeting[8] approaches.

Advances in the past years gave rise to a variety of co-translational or posttranslational means to introduce DA_{inv} motifs site-specifically to proteins to prime them for subsequent site-specific derivatization.[9] A major challenge still remains in homogeneously labeling proteins with free choice of modification site and full independence of the expression system. In contrast to co-translational approaches, posttranslational labeling methods entail strict separation of installation of the bioorthogonal motif from protein translation and

Scheme 1. Protein fusion and conjugate architectures accessible through expression (only left), enzymatic ligation strategies or site-specific bioorthogonal chemistry.

the expression host. This also minimizes exposure time to prevent the reactive scaffold from losing integrity[10] with the consequence of more efficient conjugation.[11] Thus, a key consideration toward fast and quantitative protein labeling is that the installed bioorthogonal motif is (I) as reactive as possible and (II) entirely stable until the chemoselective derivatization is performed (stability–reactivity tradeoff). With the increasing number of bioorthogonal DA$_{inv}$ motifs reported in literature,[5,12] a profound comparison of the currently used dienophile/diene pairs would be highly beneficial for experimentalists interested in protein bioconjugation, also with regard to the production of biologicals, where well-defined conjugate populations with high purity and little variability are essential.

Here we present an extensive comparative study of various bioorthogonal chemistries for quantitative protein conjugation, aimed at identifying the ideal reaction partners for the production of defined and branched protein–protein conjugates by DA$_{inv}$. We report the synthesis of a large panel of the most popular dienophile and diene scaffolds as carboxylic acid derivatives and tested them as substrates for lipoic acid ligase (LplA). A methyltetrazine substrate and a stable bicyclononyne (BCN) substrate was found to be compatible with fast and quantitative protein derivatization. The combination of both substrates allows for site-specific and almost quantitative DA$_{inv}$-based protein–protein conjugation. We demonstrate the applicability and efficiency of our conjugation method with the full-length therapeutic antibody trastuzumab. All protein–protein conjugates were fully functional as proven in fluorescence, binding, and stability studies. The speed, robustness and efficiency of the method guarantees a simple workflow to generate protein–protein conjugates of choice within hours starting from recombinant proteins regardless of their expression host. With the technology presented here, we open up new avenues for advanced applications in therapeutics or diagnostics.

Results and Discussion

Synthesis and identification of well-accepted diene and dienophile substrates for site-specific ligation to peptide tags

To compare various bioorthogonal motifs for protein modification, a suitable assay platform is required that allows evaluation of modification efficiency and product integrity under conditions close to the desired application.

Building on innovative work by the Ting group,[13] we chose the engineered lipoic acid protein ligase A LplA^{W37V} from *Escherichia coli* that ligates unnatural lipoic acid analogues to the 13 amino acid recognition motif lipoate acceptor peptide (LAP). LAP can be attached terminally to as well as internally into recombinant proteins from any host organism.[13] Several substrates with bioorthogonal motifs for site-selective tag-based protein functionalization including azides,[14] norbornenes,[15] *trans*-cyclooctenes,[9b] methyltetrazines and triazines[16] have been described.

We synthesized a panel of potential LplA^{W37V} substrates based on all commonly used dienes and dieonophiles for

DA$_{inv}$. This includes the axial 2E regioisomer of *trans*-cyclo-ctenol (**ax.-TCO***$_{a-d}$),[10b] *endo*-bicyclononynyl-methanol (**en-do-BCN**$_{a-c}$),[17] mono- and dimethylcyclopropenyl-methanol (**MMCy**$_{a-d}$ and **DMCy**$_{a-d}$)[18] derivatized with linear amino acids of varying length via a carbamate (Scheme 2, Table S1 in the Supporting Information). First, we used a HPLC-based assay to screen various putative substrate candidates for Mg^{2+}-ATP-dependent ligation onto LAP within 15 min in phosphate buffer (pH 7.0) at 37°C by both the wildtype-LplA and LplA^{W37V}. From each substrate class, an optimal candidate for maximum ligation efficiency by the LplA^{W37V} (Table S1) was identified. Carbamates with 5-aminopentanoic acid consistently gave the highest LplA^{W37V} acceptance which is why we additionally prepared derivatives with cycloct-2-yn-1-ol (**SCO**$_S$),[19] the equatorial 2E and axial 4E regioisomers of *trans*-cycloctenol (**eq.-TCO***$_S$ and **ax.-TCO**$_S$),[10b,20] as well as
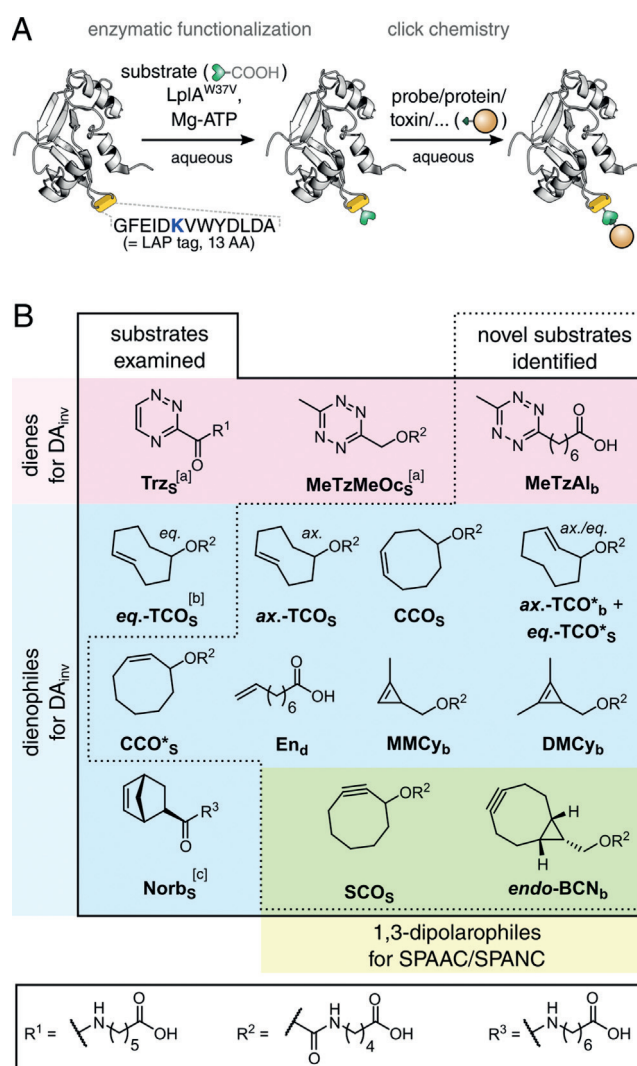
Scheme 2. A) Two-step chemoenzymatic procedure with LplA^{W37V} and click chemistry to generate protein–protein conjugates. B) Substrate scope for LplA^{W37V}-mediated site-specific protein functionalization and ensuing Diels–Alder cycloaddition with inverse electron demand (DA$_{inv}$), strain-promoted alkyne–azide cycloaddition (SPAAC) or strain-promoted alkyne–nitrone cycloaddition (SPANC). References: a: Baalmann et al.,[16] b: Liu et al.,[9b] c: Best et al.[21]

the 2Z and 4Z isomers of *cis*-cyclooctenol (**CCO*$_S$** and **CCO$_S$**), and validated them in the HPLC assay. We also included the equatorial 4E TCO-based substrate first reported by the Ting group[9b] (denoted as ***eq.*-TCO$_S$**) and the norbornene substrate **Norb$_S$** previously developed by our group[21] into our comparison. Complementary to our formerly reported diene-based methyltetrazinylmethoxycarbonyl (**MeTzMeOc$_S$**) and triazine amide (**Trz$_S$**) substrates,[16] we further synthesized methyltetrazinyl alkanoic acids (**MeTzAl$_{a-e}$**)[22] and identified the optimized substrate derivative **MeTzAl$_b$**.

This gives a total set of 15 different bioorthogonal motifs as substrates for LplA^{W37V}, covering the most commonly used diene and dienophile moieties from literature[12] to conduct DA$_{inv}$ in its entire reactivity range ($k_2 = 10^{-2}\,M^{-1}\,s^{-1}$–$10^5\,M^{-1}\,s^{-1}$)[23] (Scheme 2 and Table S1). Notably, the BCN and SCO moieties can further serve as 1,3-dipolarophiles for strain promoted alkyne–azide (SPAAC) or alkyne–nitrone (SPANC) cycloaddition, expanding the scope of the method to other bioorthogonal cycloaddition reactions.[17,23b,24] In the HPLC assay, all substrates resulted in uniform ligation products (Figures S2–S6 and S9–S13), except the TCO-based ones (Table S1 and Figures S7 and S8). Closely examining the retention times and m/z of the ***ax.*/*eq.*-TCO$_S$** ligation products in comparison with **CCO$_S$**, *trans*-to-*cis* isomerization of the TCO-modified peptides was apparent (Table S1 and Figure S7B,D). ***ax.*/*eq.*-TCO$_S$** stock solutions showed no sign of *cis*-isomerization prior to the LplA^{W37V} ligation reaction, indicating that the assay conditions lead to the observed *trans*-to-*cis* isomerization (Figures S7C). TCO* substrates (***ax.*-TCO*$_b$** and ***eq.*-TCO*$_S$**) with the previously reported more stable TCO* scaffold[10b] were prone to isomerization as well, albeit much less pronounced as ***ax.*/*eq.*-TCO$_S$**. For ***eq.*-TCO*$_S$**, isomerization was barely detectable (Table S1 and Figure S8).

Protein labeling, diene/dienophile performance and a tool to probe protein modification states

With the established substrate platform for peptide modification with all commonly used dienes and dienophiles, we moved on to evaluate the best performing LplA^{W37V} substrates for site-specific modification at the protein level.

During our work with a LAP tag internally inserted in a flexible loop of EGFP (EGFP$^{E172:LAP}$),[25] a change in electrophoretic mobility in semi-native SDS-PAGE (non-reducing and non-denaturing conditions) for both, substrate functionalized EGFP and the corresponding DA$_{inv}$ cycloadducts was observed (Figure 1). Comparable electrophoretic mobility changes due to lipoylation or octanoylation for E2 and H proteins have been reported.[26] Amino acid substitutions or posttranslational modifications can lead to an altered electrophoretic mobility and GFP's intrinsic properties seem to enhance this phenomenon.[27]

Inspired by the ability to probe the protein modification states directly at the protein level (Figure 1D), we started to investigate the ligation yields, DA$_{inv}$ conversion and side-reactions of all substrates (Figures 1, S14 and S15). Briefly, EGFP$^{E172:LAP}$ was quantitatively functionalized with each substrate (Figures S14 and S15) and subsequently treated with either **MeTzBnNH-TAMRA** or **TzBnNH-TAMRA** (Figure 1A–C, S14), respectively, or **TCO-Prop-TAMRA** and **BCN-Pip-TAMRA** (Figure S15). Reactions were quenched with a large excess of either ***eq.*-TCO-OH** or dimethyl tetrazine (**MeTzMe**), respectively. Additionally, separate reaction mixtures with LAP-tagged maltose-binding protein (MBP-LAP) were subjected to mass determination to further support the gel shifting behavior of EGFP (Figure S16 and S19).

All types of TCOs have remarkably fast kinetics in DA$_{inv}$ that are particularly useful for intracellular labeling.[28] However, previous works reported that TCO tends to isomerize to the almost non-reactive *cis*-isomer and TCO* shows elimination after DA$_{inv}$-product formation. Thus, we wanted to evaluate the extent of these unwanted side-reactions during LplA^{W37V} ligation and DA$_{inv}$. Although, ***eq.*-TCO$_S$**-modified proteins undergo rapid conjugation with both tetrazine fluorophore probes, we observed this conjugation to be incomplete limited to a maximum of about 53% (Figure 1B). This supports the notion that the previously observed *trans*-to-*cis* isomerization tendency in the peptide assay also takes place during protein labeling (Figure S7). Comparably, ***ax.*-TCO$_S$**-modified EGFP showed higher isomerization propensity. Interestingly, proteins directly modified with **CCO$_S$** were not fully unreactive toward **TzBnNH-TAMRA** resulting in non-negligible cycloadduct formation (Figure 1B). For ***ax.*-TCO*$_b$**-modified EGFP treated with **TzBnNH-TAMRA**, an unreactive EGFP population substantiates isomerization of ***ax.*-TCO***, albeit less pronounced as for the related 4E TCO substrates (Figure 1B). Especially for **MeTzBnH-TAMRA**-treated proteins, a protein species with the apparent MW of non-modified EGFP$^{E172:LAP}$ emerged (Figure 1B), which we attribute to the elimination of the carbamate function of the cycloadduct. This was further verified by data from mass spectrometry (Figure S16). Less elimination was observed when using **TzBnNH-TAMRA**, which is in agreement with the previously reported effect of the tetrazine scaffold for elimination of the TCO*-tetrazine cycloadduct.[29] High cycloaddition yield, minimal elimination and low isomerization were achieved solely with the ***eq.*-TCO*$_S$**/**TzBnNH-TAMRA** combination. However, we were unable to validate the reported exclusive orthogonality for MeTzBnNH-substituted probes under the chosen reaction conditions.[23b] Under the conditions applied, both TCO and TCO* are unsuitable for quantitative conjugation in DA$_{inv}$ reaction.

We then evaluated the reactivity of our novel BCN- and SCO-LplA^{W37V} substrates. Strained cyclic alkynes do not have an isomerization-susceptible configuration, are regarded to be stable and their cycloaddition products are not prone to elimination.[10b] To our delight, ***endo*-BCN$_b$**-functionalized proteins could be transformed nearly quantitatively to the cycloadduct with both tetrazine-TAMRA conjugates (Figure 1B and S17, S19). **SCO$_S$**-modified EGFP underwent almost full conversion to the cycloadduct with **TzBnNH-TAMRA**, but was only slightly reactive toward **MeTzBnNH-TAMRA**. This is in agreement with a previous work,[23b]
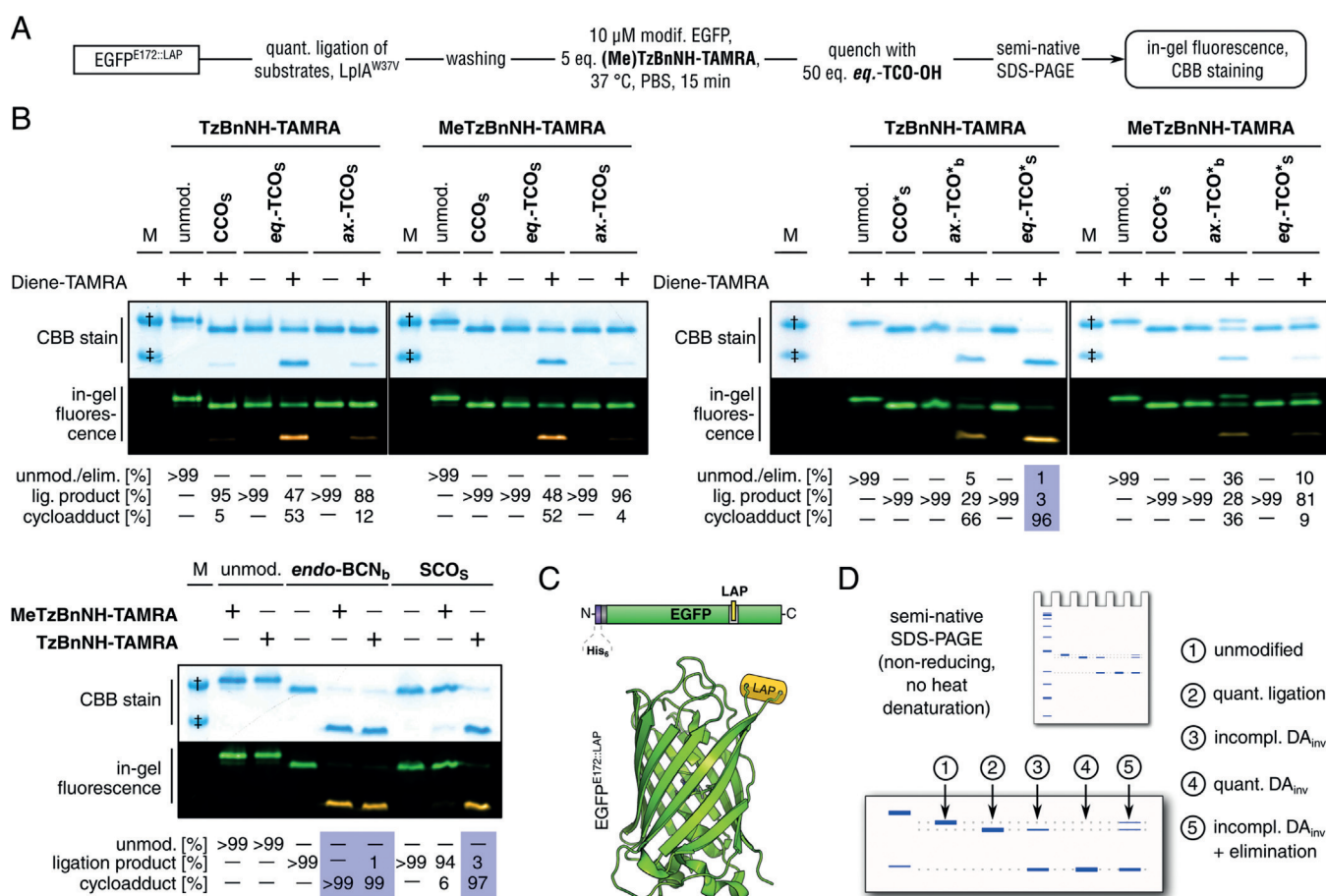
Figure 1. Individual EGFP modification states during two-step labeling can be traced by semi-native SDS-PAGE. A) Procedure of two-step EGFP modification and subsequent analysis. B) SDS gels assaying the reactivity of different dienophile LplA constructs using EGFP$^{E172:LAP}$ highlights high modification yields utilizing combinations of **endo-BCN$_b$** and **SCO$_s$** with **MeTzBnNH-TAMRA** and **TzBnNH-TAMRA** and **eq.-TCO*$_s$** together with **TzBnNH-TAMRA** (blue boxes). CBB = Coomassie Brilliant Blue. M = Marker. †: ≈ 35 kDa. ‡: ≈ 25 kDa. C) Protein construct used. D) Scheme illustrating EGFP shifting in the gel during modification steps. For the exhaustive side-by-side comparison of all DA$_{inv}$ substrates, see Figures S14 and S15. Note: Yields have been corrected for an EGFP impurity of 6% that was neither accessible for ligation nor cycloaddition reactions. Crystal structure PDB: 2Y0G.[32]

although complete orthogonality cannot be proven as reported.

Although the strained cyclooctyne substrates already had demonstrated great potential for quantitative cycloadditions, we were curious to investigate means to prevent the observed isomerization of the TCO. Isomerization of TCOs has mainly been attributed to the influence of thiols via a radical-based mechanism.[9b] We selected the two radical scavengers Trolox[30] and ascorbic acid as possible isomerization suppressors in the ligation mixture with **ax.-TCO$_s$** without effect (Figure S18A). Next, we suspected the cysteine residue (Cys85) located in the binding pocket of the substrate-bound form of the LplA^{W37V}[31] to be responsible for the observed TCO isomerization. The double mutant LplA$^{W37V/C85A}$ was prepared and displayed ligase activity for **endo-BCN$_b$** and **eq.-TCO$_s$** but did not alleviate or abolish isomerization of TCO (Figure S18B). This suggests that protein environments during the ligation reaction are sufficient to trigger TCO isomerization. Using BCN as a dienophile for DA$_{inv}$ based posttranslational protein modification takes advantage of the reactivity–stability tradeoff and outperforms TCO and TCO*.

While maintaining a high reaction rate, BCN provides quantitative conjugation yields. With **endo-BCN$_b$**, we identified the most efficient substrate for LplA^{W37V} to prime proteins for efficient DA$_{inv}$ conjugation. **En$_d$-**, **Norb$_s$-**, **MMCy$_b$-** and **DMCy$_b$-**functionalized EGFPs were expectedly much less reactive in DA$_{inv}$, but their side-by-side comparison might be interesting for some readers (Figure S14). We could also confirm almost quantitative cycloaddition of tetrazine-modified EGFP for **MeTzMeOc$_c$** with TCO and BCN probes and for the novel **MeTzAl$_b$** with TCO under the chosen reaction conditions (Figure S15). Even the triazine substrate **Trz$_s$** provided reasonable reactivity with both TCO and BCN probes.

Targeted protein-protein conjugation using DA$_{inv}$

Ideally, a method for protein–protein conjugation is directly applicable to the recombinant protein from the expression host of choice using enzymatic or chemical strategies. Sortase or split-intein strategies have been used

for post-translational covalent bond formation between two or more native proteins to a continuous polypeptide chain.[2,33] Especially for larger multidomain protein fusions, it might be advantageous to precisely attach proteins at defined internal positions within a single polypeptide chain. The generation of these internal or branched protein fusions from individual native proteins purely by enzymatic means is a challenge. To our knowledge, only the approaches based on CnaB2 and D4 Ig-like domains (Spy-/Snoop- × Tag/Catcher) from the Howarth lab provide enzyme-catalyzed covalent internal protein connectivity in a building block principle.[34] Amine and thiol reactive chemistry has enabled internal protein–protein conjugation much long before the advent of genetic fusions,[35] but is severely limited in site-specificity and thus not suited for precise construction of branched protein architectures.[36] Protein–protein conjugation utilizing bioorthogonal reactions has attracted attention over the last years due to its potential to overcome terminal attachment restrictions or non-selective reactions. It is surprising that targeted and conjugation between proteins with internal connectivity via DA$_{inv}$ has not been investigated yet, given that all potential tools in genetic code expansion are available for a couple of years.[38] Apart from conjugations between TCO- and tetrazine-modified proteins pre-modified via non-site-specific NHS chemistry,[39] there are only a few examples for targeted protein–protein conjugation using DA$_{inv}$.[11b,40]

Stimulated by the modularity of the stable LplA^{W37V} substrates and the quick and near-quantitative ligation reaction, we envisioned functionalizing proteins with a methyltetrazine substrate (**MeTzMeOc$_s$** or **MeTzAl$_b$**) and an *endo-*

BCN$_b$ dienophile substrate separately. Subsequently, the primed proteins can be combined for conjugation by the DA$_{inv}$ cycloaddition, benefitting from the fast reaction rate of tetrazines with strained dienophiles and the benign catalyst-free[5] conditions (Figure 2).[25,41] Introduction of LAP in permissive internal positions should provide the basis to construct branched protein chimeras.[25,42] We thus chose the conjugation between the two model proteins MBP-LAP and an EGFP construct with an internal LAP tag (EGFP$^{Q157:LAP}$) as a proof of concept. For 1:2 protein ratios at 10 μM and 2–4 h incubation time, the conjugation gave nearly quantitative yield related to the protein reaction partner which is not present in excess (fluorescent protein species at ≈85 kDa, Figure 2A and Figure S21). Non-purified reaction mixtures of the protein cycloaddition reactions were further analyzed by intact protein mass determination showing full conversion to the cycloaddition product (Figure 2C and S20). Notably, the N- and C-termini of both proteins remain in their native unmodified form providing additional branching options for even more complex protein chimeras. To analyze the kinetics of this protein-protein conjugation method, we incubated *endo-*BCN$_b$-functionalized EGFP$^{E172:LAP}$ with varying concentrations of methyltetrazine-functionalized mRuby3-LAP and monitored the Förster resonance energy transfer (FRET) acceptor fluorescence of mRuby3 over time. Under pseudo-first order kinetics, remarkably high second order rate constants of ≈70 M^{-1}s^{-1} and ≈50 M^{-1}s^{-1} for the *endo-*BCN$_b$/MeTzMeOc$_s$ and *endo-*BCN$_b$/MeTzAl$_b$ combinations, respectively, were determined at 37 °C in PBS (Figure 3 and S22). In

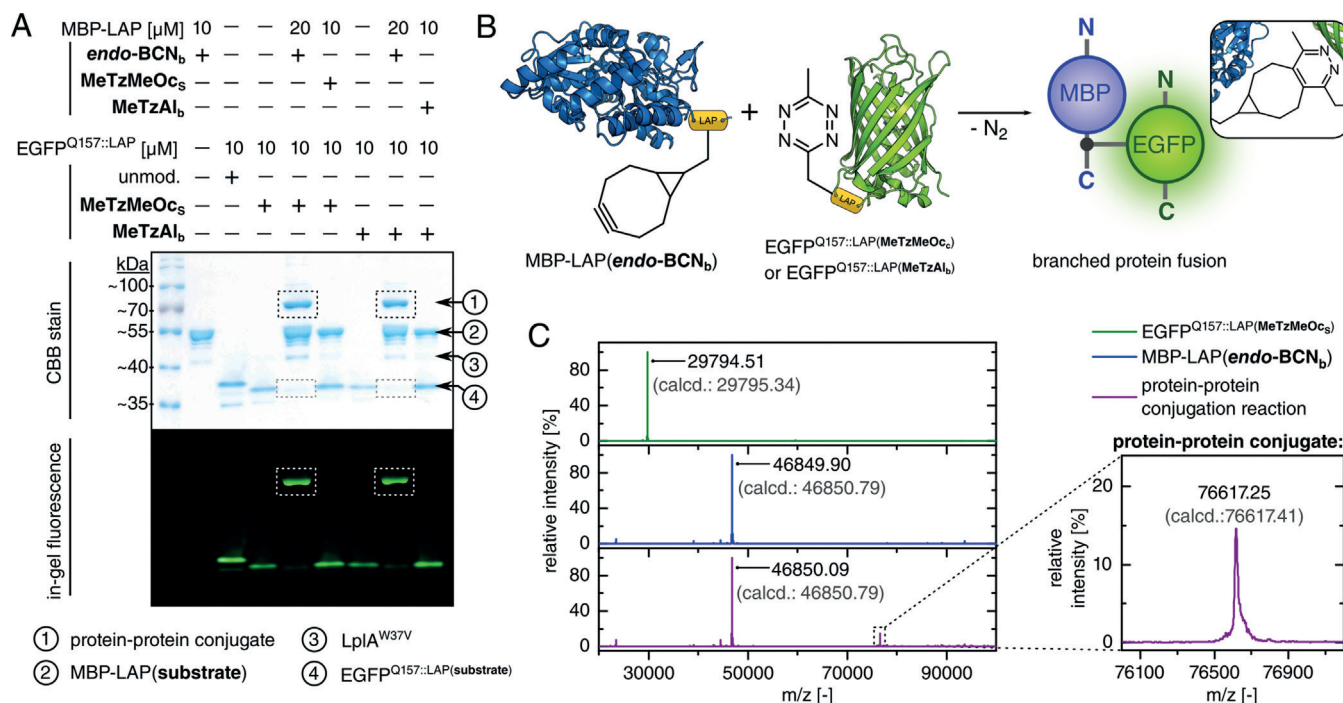**Figure 2.** Protein–protein conjugation using model proteins MBP-LAP and EGFP$^{Q157:LAP}$. Combination of *endo-*BCN$_b$ with MeTzMeOc$_s$ or MeTzAl$_b$ was efficiently used for targeted protein–protein conjugation by DA$_{inv}$. A) SDS-PAGE analysis highlights almost quantitative formation of MBP-EGFP conjugate (≈70 kDa) within 2 h. CBB = Coomassie Brilliant Blue. B) Schematic view of the model proteins primed for protein–protein conjugation by DA$_{inv}$. C) Mass spectrometry analysis of conjugation reaction. Crystal structure PDBs: 1ANF[37] and 2Y0G.
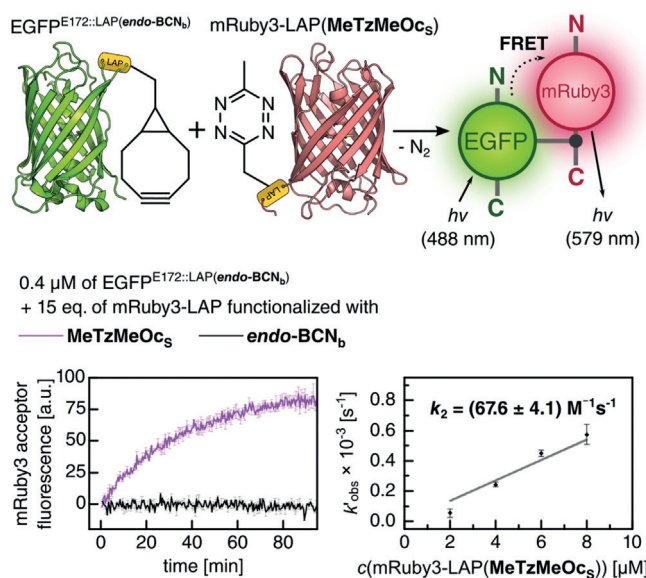
Figure 3. Reaction kinetics for targeted protein–protein conjugation between EGFP and mRuby3 monitored by FRET. Left graph shows mRuby3 FRET acceptor response. Right graph shows determination of second order rate constant. Error bars in left graph: ±1 standard deviation; in right graph: ±1 highest error estimate. Crystal structure PDBs: 2Y0G and 3U0L.[43]

line with our expectations, FRET partners both equipped with ***endo*-BCN_b** did not show any FRET response (Figure 3).

Functional protein–antibody conjugates

Next, we wanted to demonstrate the power of our conjugation procedure for the generation of therapeutic protein chimeras. We genetically equipped the monoclonal antibody trastuzumab with the LAP motif on the C-termini of each heavy chain (trastuzumab-hc^{LAP}/lc^{WT}, further referred to as trastuzumab-(LAP)$_2$). This immunoglobulin targets HER2-positive cancer cells and is applied as parental antibody in the FDA approved drug Herceptin.[44] Very recently, LplA^{W37V}-mediated TCO ligation was used in combination with transglutaminase for one-pot dual labeling of trastuzumab.[6a] The data presented in the report indirectly point toward an incomplete conjugation of TCO modified trastuzumab-(LAP)$_2$ which is in agreement with our observation of an impaired stability of TCO for quantitative in vitro bioconjugation. We reasoned that antibody functionalization with ***endo*-BCN_b** might serve as a conjugation hub enabling both DA$_{inv}$/SPAAC with small molecules (fluorophores or cytotoxic drugs) and even whole proteins. Ligation of ***endo*-**
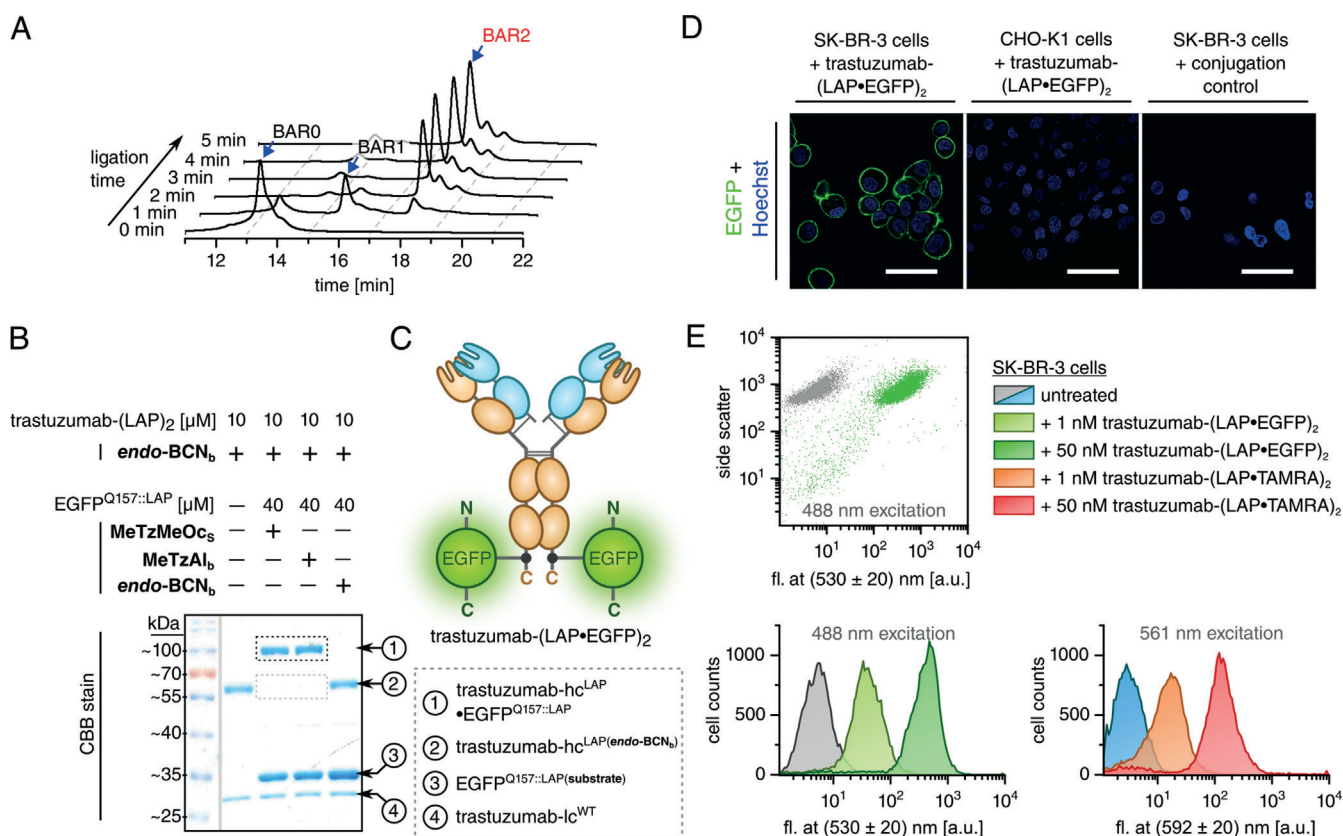

Figure 4. Protein–antibody conjugation using DA$_{inv}$ with ***endo*-BCN_b** and **MeTzMeOc_s/MeTzAl_b** yields a functional protein conjugate with the full-length antibody trastuzumab. A) Ligation of ***endo*-BCN_b** to the doubly LAP-tagged trastuzumab, BAR = BCN-to-antibody ratio. See Figure S23 for expanded ligation time. B) SDS-PAGE analysis shows quantitative conjugation EGFP to trastuzumab. hc = heavy chain, lc = light chain. C) Scheme of trastuzumab-(LAP·EGFP)$_2$ conjugate architecture. D) Confocal microscopy with the trastuzumab-(LAP·EGFP)$_2$ conjugate shows specific binding to SK-BR-3 and no binding to CHO-K1 cells. Scale bar = 20 μm. Conjugation control: trastuzumab-(LAP)$_2$ and EGFP$^{Q157:LAP}$ both functionalized with ***endo*-BCN_b**. E) Flow cytometry analysis shows concentration-dependent binding of trastuzumab-(LAP·EGFP)$_2$ and trastuzumab-(LAP·TAMRA)$_2$ conjugates to SK-BR-3 cells. fl. = fluorescence.

BCN_b to both LAP tags of intact trastuzumab was monitored by hydrophobic interaction chromatography (HIC, Figure 4A and S23) and quantitative conversion of trastuzumab to the doubly BCN-functionalized species within 3 min was observed, demonstrating rapid and clean antibody functionalization.[45]

With fast, targeted and efficient posttranslational protein conjugation to antibodies being a hitherto unmet challenge, we sought to directly assess this concept with internally tetrazine-functionalized EGFP. Protein conjugation to trastuzumab smoothly went to near completion at 40 μM of tetrazine-modified EGFP$^{Q157:LAP}$ and 10 μM of doubly BCN-functionalized trastuzumab-(LAP)$_2$ at 37 °C for 4 h in PBS (Figure 4B). Encouraged by these promising results, the biological activity of doubly EGFP-modified trastuzumab (trastuzumab-(LAP·EGFP)$_2$, Figure 4C) was assessed via immunofluorescence staining of living mammalian cells using confocal microscopy (Figure 4D and S24A). As expected, the fluorescent signal of EGFP-conjugated trastuzumab accumulated at the outer cell membrane of the HER2-positive breast cancer cell line SK-BR-3, but not at the HER2-negative control cell line CHO-K1. No fluorescent signal was detected for the negative control conjugation (Figure 4D), supporting the lack of mutual cross reactivity between BCN-functionalized proteins (Figure 2A and 4B). Incubation of both cell lines with a 1:1 mixture of trastuzumab-(LAP·EGFP)$_2$ and trastuzumab-(LAP·TAMRA)$_2$ resulted in colocalizing fluorescent signals of only on the HER2-positive cell line (Figure S24B).

To examine how the LAP tag and subsequent conjugation with a probe influence biological properties of the antibody, subsequent binding studies were performed. A comparative binding study of trastuzumab-(LAP·EGFP)$_2$, trastuzumab-(LAP·TAMRA)$_2$ and trastuzumab-(LAP)$_2$ on HER2-positive cells revealed potent binding of all constructs (Figure 4E and S25). To further substantiate these findings, we determined the binding constants on cells applying flow cytometry in vitro (Figure S26). Trastuzumab with the C-terminal LAP motif revealed single-digit nanomolar binding properties ($K_D = 5.6$ nM) equal to unmodified wildtype trastuzumab ($K_D = 6.3$ nM).[46] Notably, trastuzumab equipped with EGFP exhibited equipotential binding ($K_D = 7.9$ nM). Additionally, the further modified trastuzumab-(LAP·TAMRA)$_2$ construct revealed only a negligible loss of binding potential ($K_D = 12.9$ nM).

Encouraged by these results, we further investigated the impact of the LAP motif itself and the conjugation to EGFP on the stability of the antibody. Thermal shift assays revealed equal melting points for the trastuzumab-(LAP)$_2$ compared to the unmodified wildtype antibody. Additionally, no loss in stability was found for trastuzumab-(LAP·EGFP)$_2$, as three melting points equal to the Fc- and the Fab-part of the unmodified trastuzumab and to solitaire EGFP were obtained (Figure S27 and Table S28).

Antibody–drug conjugates (ADCs) by SPAAC click chemistry

To investigate, whether our strategy of site-specific protein conjugation allows for the efficient generation of antibodies equipped with highly potent cytotoxins, we generated an ADC in a two-step procedure. Trastuzumab-(LAP)$_2$ was modified with **endo-BCN_b** by LplA^{W37V}-catalysis followed by conjugation of highly toxic monomethyl auristatin E (MMAE) by SPAAC applying **N$_3$-PEG$_3$-vc-PABC-MMAE**, resulting in trastuzumab-(LAP·MMAE)$_2$. Next, we assessed the in vitro potency of the trastuzumab-(LAP·MMAE)$_2$. For this, a cell proliferation assay on SK-BR-3 cells and CHO-K1 cells was performed (Figure 5). Potent, subnanomolar (IC$_{50}$ = 0.31 nM) cell-killing properties were shown for trastuzumab-(LAP·MMAE)$_2$ on HER2-postive cells, while the construct was found innocuous to HER2-negative cells. As expected, trastuzumab-(LAP)$_2$ lacking a toxin warhead was found non-toxic for both HER2-postive and negative cells. The introduced peptide may have antigenic potential upon in vivo
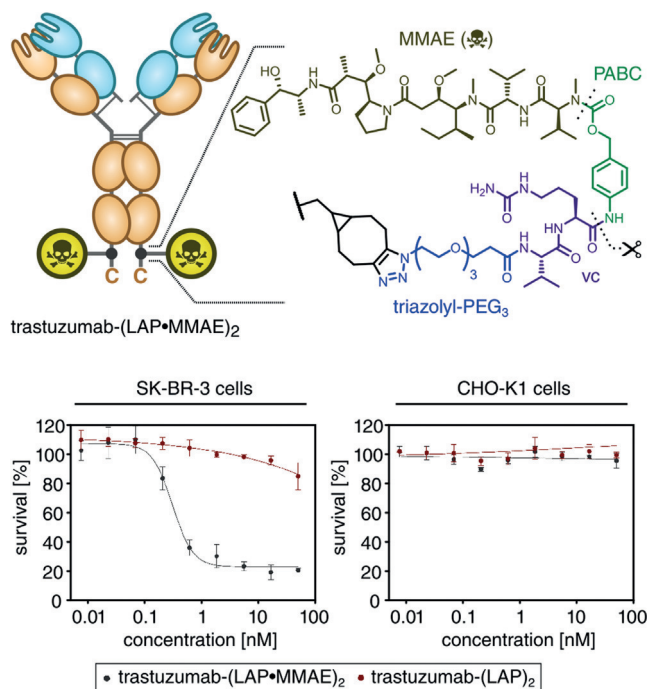

Figure 5. Generation of the potent antibody–drug conjugate trastuzumab-(LAP·MMAE)$_2$ by **endo-BCN_b** ligation followed by SPAAC with **N$_3$-PEG$_3$-vc-PABC-MMAE**. Top: Conjugate architecture. vc = valine-citrulline, PABC = *para*-aminobenzyl carbamate, MMAE = monomethyl auristatin E. Dotted lines indicate cleavable bonds for drug release. Scissors indicate cleavage site for lysosomal proteases. Bottom: Cell proliferation assays with SK-BR-3 and CHO-K1 cells treated with trastuzumab-(LAP)$_2$ and trastuzumab-(LAP·MMAE)$_2$ (DAR2). Error bars in graphs: ±1 standard deviation.

administration as for several other antibody tags that are currently used for orthogonal modification of antibodies. Consequently, careful evaluation of the antigenicity of the LAP sequence that may not only depend on the peptide sequence, but also on other factors such as the mode of

administration or the antibody concentration will be necessary in frame of potential medical applications.

Conclusion

In conclusion, we have developed a highly efficient method to synthesize well-defined, homogeneous protein(–protein) conjugates utilizing enzyme-mediated protein modification and bioorthogonal DA_{inv} chemistry. Various diene and dienophile substrates were analyzed for ligation to proteins using $LplA^{W3V}$ as well as for their ability to yield homogeneous conjugates in subsequent DA_{inv}. For this we utilized the unique gel shifting behavior of EGFP to evaluate stability and reactivity of bioorthogonal motifs at the protein level. We identified novel tetrazine- and BCN substrates to yield quantitative homogeneous protein–protein conjugates without any detectable by-products or toxic remnants, within a single working day starting from purified protein. By efficiently labeling the therapeutic antibody trastuzumab with EGFP or the antimitotic agent MMAE, respectively, we have shown that the novel methodology offers great potential for advances in the production of ADCs and antibody protein toxin conjugates. Our approach provides great flexibility to conjugate a probe, drug or even a whole protein to various internal positions, resulting in branched protein architectures that are difficult to accomplish with other technologies. We assume that the versatility of this conjugation strategy bears potential for various applications, for example, for facile drug development,[45b] generation of biomaterials[47] or nanofabrication.[47,48] In particular, with biologics being the fastest growing class of drugs, we see the herein reported method as an attractive technology that can speed up development timelines and manufacturing, important aspects to enhance the biologics drug pipeline.

Conflict of interest

The authors declare no conflict of interest.

[1] M. D. Witte, J. J. Cragnolini, S. K. Dougan, N. C. Yoder, M. W. Popp, H. L. Ploegh, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11993–11998.

[2] T. Pirzer, K. S. Becher, M. Rieker, T. Meckel, H. D. Mootz, H. Kolmar, *ACS Chem. Biol.* **2018**, *13*, 2058–2066.

[3] P. Costantino, S. Viti, A. Podda, M. A. Velmonte, L. Nencioni, R. Rappuoli, *Vaccine* **1992**, *10*, 691–698.

[4] P. Agarwal, C. R. Bertozzi, *Bioconjugate Chem.* **2015**, *26*, 176–192.

[5] B. L. Oliveira, Z. Guo, G. J. L. Bernardes, *Chem. Soc. Rev.* **2017**, *46*, 4895–4950.

[6] a) D. N. Thornlow, E. C. Cox, J. A. Walker, M. Sorkin, J. B. Plesset, M. P. DeLisa, C. A. Alabi, *Bioconjugate Chem.* **2019**, *30*, 1702–1710; b) B. Oller-Salvia, G. Kym, J. W. Chin, *Angew. Chem. Int. Ed.* **2018**, *57*, 2831–2834; *Angew. Chem.* **2018**, *130*, 2881–2884.

[7] B. M. Zeglis, P. Mohindra, G. I. Weissmann, V. Divilov, S. A. Hilderbrand, R. Weissleder, J. S. Lewis, *Bioconjugate Chem.* **2011**, *22*, 2048–2059.

[8] a) R. Rossin, P. R. Verkerk, S. M. van den Bosch, R. C. Vulders, I. Verel, J. Lub, M. S. Robillard, *Angew. Chem. Int. Ed.* **2010**, *49*, 3375–3378; *Angew. Chem.* **2010**, *122*, 3447–3450; b) N. K. Devaraj, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2008**, *19*, 2297–2299.

[9] a) T. Plass, S. Milles, C. Koehler, J. Szymanski, R. Mueller, M. Wiessler, C. Schultz, E. A. Lemke, *Angew. Chem. Int. Ed.* **2012**, *51*, 4166–4170; *Angew. Chem.* **2012**, *124*, 4242–4246; b) D. S. Liu, A. Tangpeerachaikul, R. Selvaraj, M. T. Taylor, J. M. Fox, A. Y. Ting, *J. Am. Chem. Soc.* **2012**, *134*, 792–795; c) J. K. Dozier, S. L. Khatwani, J. W. Wollack, Y. C. Wang, C. Schmidt-Dannert, M. D. Distefano, *Bioconjugate Chem.* **2014**, *25*, 1203–1212; d) R. J. Blizzard, D. R. Backus, W. Brown, C. G. Bazewicz, Y. Li, R. A. Mehl, *J. Am. Chem. Soc.* **2015**, *137*, 10044–10047; e) M. Best, I. Porth, S. Hauke, F. Braun, D. P. Herten, R. Wombacher, *Org. Biomol. Chem.* **2016**, *14*, 5606–5611.

[10] a) K. Lang, L. Davis, S. Wallace, M. Mahesh, D. J. Cox, M. L. Blackman, J. M. Fox, J. W. Chin, *J. Am. Chem. Soc.* **2012**, *134*, 10317–10320; b) J. E. Hoffmann, T. Plass, I. Nikic, I. V. Aramburu, C. Koehler, H. Gillandt, E. A. Lemke, C. Schultz, *Chem. Eur. J.* **2015**, *21*, 12266–12270.

[11] a) E. J. Sherman, D. A. Lorenz, A. L. Garner, *ACS Comb. Sci.* **2019**, *21*, 522–527; b) A. T. Krueger, C. Kroll, E. Sanchez, L. G. Griffith, B. Imperiali, *Angew. Chem. Int. Ed.* **2014**, *53*, 2662–2666; *Angew. Chem.* **2014**, *126*, 2700–2704.

[12] K. Lang, J. W. Chin, *ACS Chem. Biol.* **2014**, *9*, 16–20.

[13] C. Uttamapinant, K. A. White, H. Baruah, S. Thompson, M. Fernandez-Suarez, S. Puthenveetil, A. Y. Ting, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10914–10919.

[14] a) H. Baruah, S. Puthenveetil, Y. A. Choi, S. Shah, A. Y. Ting, *Angew. Chem. Int. Ed.* **2008**, *47*, 7018–7021; *Angew. Chem.* **2008**, *120*, 7126–7129; b) C. Uttamapinant, A. Tangpeerachaikul, S. Grecian, S. Clarke, U. Singh, P. Slade, K. R. Gee, A. Y. Ting, *Angew. Chem. Int. Ed.* **2012**, *51*, 5852–5856; *Angew. Chem.* **2012**, *124*, 5954–5958.

[15] M. Best, A. Degen, M. Baalmann, T. T. Schmidt, R. Wombacher, *ChemBioChem* **2015**, *16*, 1158–1162.

[16] M. Baalmann, M. J. Ziegler, P. Werther, J. Wilhelm, R. Wombacher, *Bioconjugate Chem.* **2019**, *30*, 1405–1414.

[17] J. Dommerholt, S. Schmidt, R. Temming, L. J. Hendriks, F. P. Rutjes, J. C. van Hest, D. J. Lefeber, P. Friedl, F. L. van Delft, *Angew. Chem. Int. Ed.* **2010**, *49*, 9422–9425; *Angew. Chem.* **2010**, *122*, 9612–9615.

[18] a) F. Doll, J. Hassenruck, V. Wittmann, A. Zumbusch, *Methods Enzymol.* **2018**, *598*, 283–319; b) J. Yang, J. Seckute, C. M. Cole, N. K. Devaraj, *Angew. Chem. Int. Ed.* **2012**, *51*, 7476–7479; *Angew. Chem.* **2012**, *124*, 7594–7597.

[19] T. Plass, S. Milles, C. Koehler, C. Schultz, E. A. Lemke, *Angew. Chem. Int. Ed.* **2011**, *50*, 3878–3881; *Angew. Chem.* **2011**, *123*, 3964–3967.

[20] M. L. Blackman, M. Royzen, J. M. Fox, *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519.

[21] M. Best, A. Degen, M. Baalmann, T. T. Schmidt, R. Wombacher, *ChemBioChem* **2015**, *16*, 1158–1162.

[22] J. Yang, M. R. Karver, W. Li, S. Sahu, N. K. Devaraj, *Angew. Chem. Int. Ed.* **2012**, *51*, 5222–5225; *Angew. Chem.* **2012**, *124*, 5312–5315.

[23] a) A. Niederwieser, A. K. Späte, L. D. Nguyen, C. Jungst, W. Reutter, V. Wittmann, *Angew. Chem. Int. Ed.* **2013**, *52*, 4265–4268; *Angew. Chem.* **2013**, *125*, 4359–4363; b) J. A. Wagner, D. Mercadante, I. Nikic, E. A. Lemke, F. Grater, *Chem. Eur. J.* **2015**, *21*, 12431–12435.

[24] a) N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047; b) X. Ning, R. P. Temming, J. Dommerholt, J. Guo, D. B. Ania, M. F. Debets, M. A. Wolfert, G.-J. Boons, F. L. van Delft, *Angew. Chem. Int. Ed.* **2010**, *49*, 3065–3068; *Angew. Chem.* **2010**, *122*, 3129–3132; c) C. S. McKay, J. Moran, J. P. Pezacki, *Chem. Commun.* **2010**, *46*, 931–933.

[25] J. G. Plaks, R. Falatach, M. Kastantin, J. A. Berberich, J. L. Kaar, *Bioconjugate Chem.* **2015**, *26*, 1104–1112.

[26] a) K. Fujiwara, K. Okamura-Ikeda, Y. Motokawa, *J. Biol. Chem.* **1992**, *267*, 20011–20016; b) J. Quinn, A. G. Diamond, A. K. Masters, D. E. Brookfield, N. G. Wallis, S. J. Yeaman, *Biochem. J.* **1993**, *289*, 81–85; c) S. W. Jordan, J. E. Cronan, Jr., *J. Bacteriol.* **2003**, *185*, 1582–1589.

[27] T. Aoki, T. Tahara, K. Satoh, H. Fujino, H. Watabe, *Anal. Biochem.* **2003**, *317*, 107–115.

[28] H. E. Murrey, J. C. Judkins, C. W. Am Ende, T. E. Ballard, Y. Fang, K. Riccardi, L. Di, E. R. Guilmette, J. W. Schwartz, J. M. Fox, D. S. Johnson, *J. Am. Chem. Soc.* **2015**, *137*, 11461–11475.

[29] a) R. M. Versteegen, R. Rossin, W. ten Hoeve, H. M. Janssen, M. S. Robillard, *Angew. Chem. Int. Ed.* **2013**, *52*, 14112–14116; *Angew. Chem.* **2013**, *125*, 14362–14366; b) X. Fan, Y. Ge, F. Lin, Y. Yang, G. Zhang, W. S. Ngai, Z. Lin, S. Zheng, J. Wang, J. Zhao, J. Li, P. R. Chen, *Angew. Chem. Int. Ed.* **2016**, *55*, 14046–14050; *Angew. Chem.* **2016**, *128*, 14252–14256.

[30] Y. Fang, J. C. Judkins, S. J. Boyd, C. W. am Ende, K. Rohlfing, Z. Huang, Y. Xie, D. S. Johnson, J. M. Fox, *Tetrahedron* **2019**, *75*, 4307–4317.

[31] K. Fujiwara, N. Maita, H. Hosaka, K. Okamura-Ikeda, A. Nakagawa, H. Taniguchi, *J. Biol. Chem.* **2010**, *285*, 9971–9980.

[32] A. Royant, M. Noirclerc-Savoye, *J. Struct. Biol.* **2011**, *174*, 385–390.

[33] a) H. Mao, S. A. Hart, A. Schink, B. A. Pollok, *J. Am. Chem. Soc.* **2004**, *126*, 2670–2671; b) J. M. Antos, M. C. Truttmann, H. L. Ploegh, *Curr. Opin. Struct. Biol.* **2016**, *38*, 111–118.

[34] a) B. Zakeri, M. Howarth, *J. Am. Chem. Soc.* **2010**, *132*, 4526–4527; b) J. O. Fierer, G. Veggiani, M. Howarth, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E1176–E1181; c) G. Veggiani, T. Nakamura, M. D. Brenner, R. V. Gayet, J. Yan, C. V. Robinson, M. Howarth, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 1202–1207.

[35] a) R. A. F. Reinhart, P. D. Bragg, *Biochim. Biophys. Acta Biomembr.* **1977**, *466*, 245–256; b) J. Carlsson, H. Drevin, R. Axen, *Biochem. J.* **1978**, *173*, 723–737; c) S. Hashida, E. Ishikawa, *Anal. Lett.* **1985**, *18*, 1143–1155.

[36] L. Wang, G. Amphlett, W. A. Blattler, J. M. Lambert, W. Zhang, *Protein Sci.* **2005**, *14*, 2436–2446.

[37] F. A. Quiocho, J. C. Spurlino, L. E. Rodseth, *Structure* **1997**, *5*, 997–1015.

[38] a) J. L. Seitchik, J. C. Peeler, M. T. Taylor, M. L. Blackman, T. W. Rhoads, R. B. Cooley, C. Refakis, J. M. Fox, R. A. Mehl, *J. Am. Chem. Soc.* **2012**, *134*, 2898–2901; b) K. Lang, L. Davis, J. Torres-Kolbus, C. Chou, A. Deiters, J. W. Chin, *Nat. Chem.* **2012**, *4*, 298–304.

[39] S. Chaudhuri, T. Korten, S. Diez, *Bioconjugate Chem.* **2017**, *28*, 918–922.

[40] A. Rutkowska, T. Plass, J. E. Hoffmann, D. A. Yushchenko, S. Feng, C. Schultz, *ChemBioChem* **2014**, *15*, 1765–1768.

[41] A. Hagen, M. Sutter, N. Sloan, C. A. Kerfeld, *Nat. Commun.* **2018**, *9*, 2881.

[42] a) Y. Zheng, M. Kielian, *Virology* **2015**, *484*, 412–420; b) Y. Xu, L. Peng, S. Wang, A. Wang, R. Ma, Y. Zhou, J. Yang, D. E. Sun, W. Lin, X. Chen, P. Zou, *Angew. Chem. Int. Ed.* **2018**, *57*, 3949–3953; *Angew. Chem.* **2018**, *130*, 4013–4017.

[43] J. Akerboom, N. Carreras Calderon, L. Tian, S. Wabnig, M. Prigge, J. Tolo, A. Gordus, M. B. Orger, K. E. Severi, J. J. Macklin, R. Patel, S. R. Pulver, T. J. Wardill, E. Fischer, C. Schuler, T. W. Chen, K. S. Sarkisyan, J. S. Marvin, C. I. Bargmann, D. S. Kim, S. Kugler, L. Lagnado, P. Hegemann, A. Gottschalk, E. R. Schreiter, L. L. Looger, *Front. Mol. Neurosci.* **2013**, *6*, 2.

[44] J. M. Lambert, A. Berkenblit, *Annu. Rev. Med.* **2018**, *69*, 191–207.

[45] a) G. Falck, K. M. Müller, *Antibodies* **2018**, *7*, 4; b) N. Diamantis, U. Banerji, *Br. J. Cancer* **2016**, *114*, 362–367.

[46] H. Schneider, L. Deweid, T. Pirzer, D. Yanakieva, S. Englert, B. Becker, O. Avrutina, H. Kolmar, *ChemistryOpen* **2019**, *8*, 354–357.

[47] C. A. DeForest, K. S. Anseth, *Nat. Chem.* **2011**, *3*, 925–931.

[48] M.-E. Aubin, D. G. Morales, K. Hamad-Schifferli, *Nano Lett.* **2005**, *5*, 519–522.