To View Your Biomolecule, Click inside the Cell

Vincent Rigolot, Christophe Biot, Cedric Lion

▶ To cite this version:

Vincent Rigolot, Christophe Biot, Cedric Lion. To View Your Biomolecule, Click inside the Cell. Angewandte Chemie International Edition, Wiley-VCH Verlag, 2021, 60 (43), pp.23084-23105. 10.1002/anie.202101502 . hal-03386257

HAL Id: hal-03386257
https://hal.univ-lille.fr/hal-03386257

Submitted on 19 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TO VIEW YOUR BIOMOLECULE, CLICK INSIDE THE CELL

Vincent Rigolot, Christophe Biot* and Cédric Lion*

Univ. Lille, CNRS, UMR 8576 – UGSF – Unité de Glycobiologie Structurale et Fonctionnelle, 59000 Lille (France)
* Co-corresponding authors. E-mail: cedric.lion@univ-lille.fr ; christophe.biot@univ-lille.fr

ABSTRACT

The surging development of bioorthogonal chemistry has profoundly transformed chemical biology over the last two decades. Involving partners that specifically react together in highly complex biological fluids, this branch of chemistry now allows researchers to probe biomolecules in their natural habitat through metabolic labelling technologies. Chemical reporter strategies include metabolic glycan labelling, site-specific incorporation of unnatural amino acids in proteins, or nucleic acid postsynthetic labelling. While a majority of literature reports marks cell-surface exposed targets, implementing bioorthogonal ligations in the interior of cells constitutes a more challenging task. Owing to limiting factors such as reagent membrane permeability, fluorescence background due to hydrophobic interactions and off-target covalent binding, or perfectible kinetics/stability balances, these strategies need mindful planning to achieve success. In this review, we emphasize the hurdles encountered when targeting biomolecules localized in cell organelles, and wish to give an easily-accessible summary of the strategies at hand for imaging intracellular targets.

1. INTRODUCTION

For decades, first generation bioconjugation reactions such as amine-reactive NHS esters or isothiocyanates, thiol-reactive maleimides or pyridyldisulfides, or carbonyl-reactive hydrazides or oximes have been providing the scientific community with a wealth of valuable molecular tools for the visualization, detection, vectorization or immobilization of biomolecules of interest (BOI).[1,2]

However, for applications that require monitoring of a detectable bioconjugate in cell/tissue cultures or in whole organisms, these reactions present shortcomings that seriously narrow their scope of application. Because they readily interfere with chemical groups that are overwhelmingly present in living systems, the use of such reactions is indeed limited to the preparation of bioconjugates *in vitro* prior to their introduction in a complex biological sample. In addition, the ligation of bulky, often hydrophobic or charged amphiphilic probes such as organic fluorophores can significantly change the BOI's physico-chemical properties. It may negatively impact cell entry, localization, or affinity of the transmogrified BOI with receptors or enzymes thereby altering or even abolishing its biological behaviour. Shared with genetic approaches such as fluorescent proteins and other peptide tags, this pitfall long left in the dark various dynamic phenomena of interest including post-translational modifications of proteins and trafficking of small metabolites. Two decades ago, the emergence of metabolic labelling approaches involving small chemically engineered molecular spies started to fill that void,[3,4] while at the same time evidencing the need for new chemical transformations that would allow highly chemoselective covalent ligation of a reactive probe to a target BOI directly in the midst of its native biological habitat (*in vivo* or *ex vivo*). The hunt for bioorthogonal reactions had started.

In a period marked by major breakthroughs in the bioimaging field, the development of exogenous chemistry to interrogate biological processes *in situ* has been central in chemical biology. Historically, many early applications of the bioorthogonal chemical reporter strategy have addressed visualization at the cell surface, whereas detection inside the cell has proved to be a more challenging, yet increasingly achievable task (Figure 1). In this paper, we take a closer look at recent progress in the main bioorthogonal ligations with a special focus on technological advances that allow the detection of BOIs in intracellular compartments.

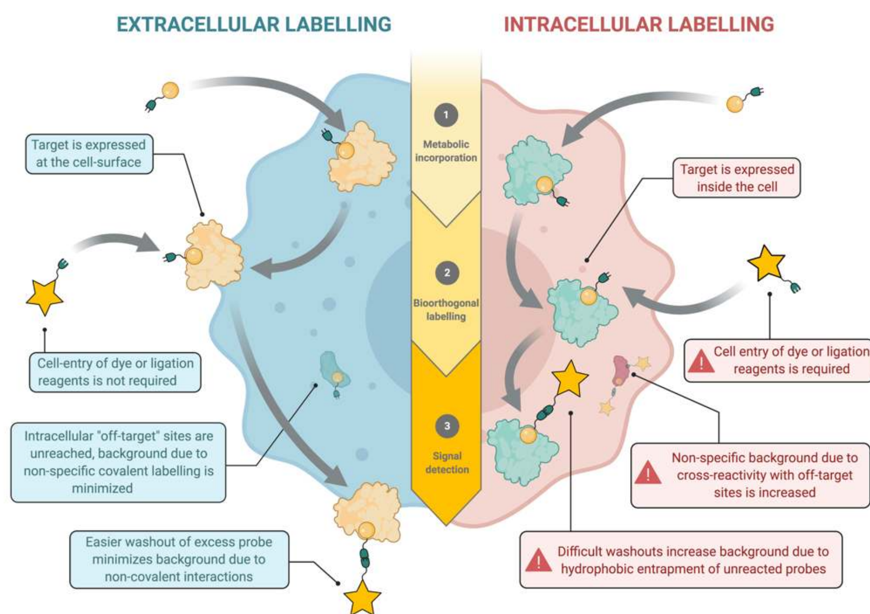


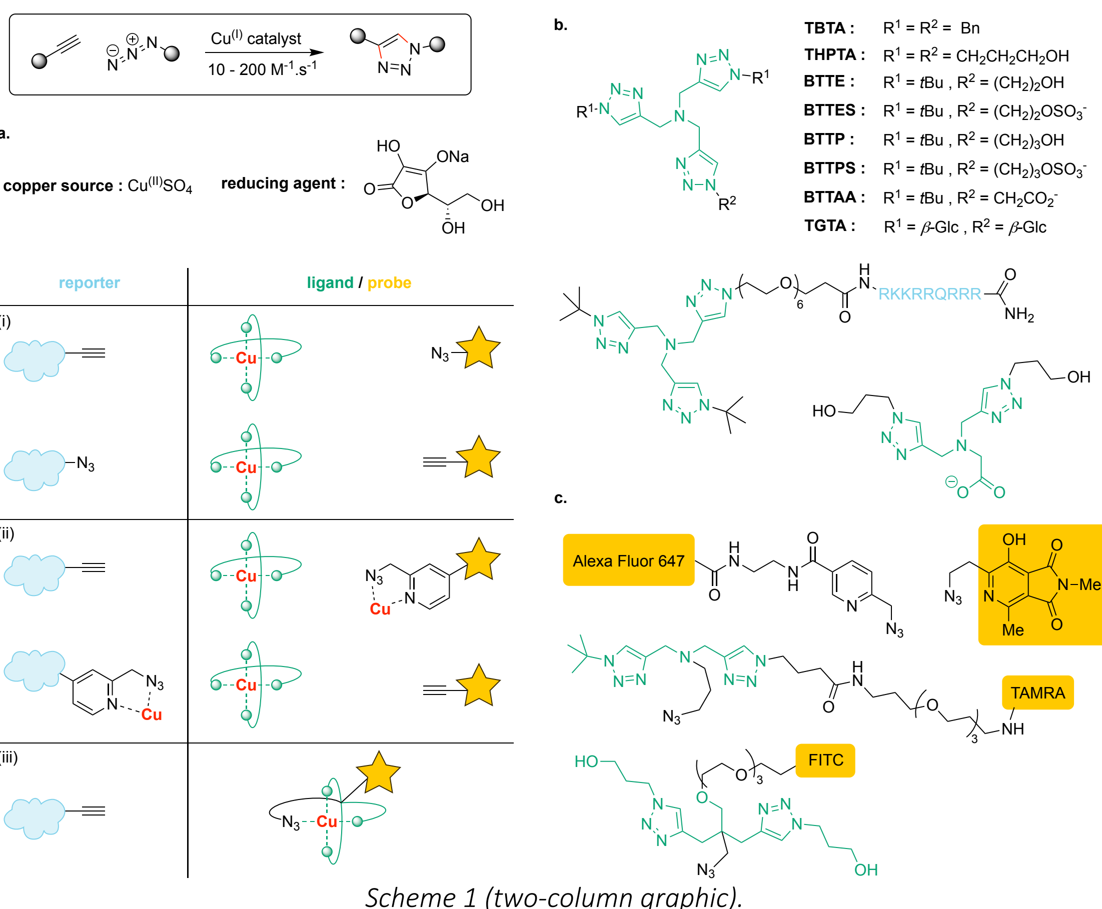
Figure 1 (two-column graphic).

Over the past 20 years, chemical biologists strived to develop a library of mutually selective pairs of chemical handles that react together efficiently to form a stable covalent linkage in the context of a complex biological environment, while minimizing interference with endogenous structures. Such reactions must proceed at high kinetic rates in physiological conditions of pH, pressure and temperature to allow satisfying signal-to-noise ratio with low concentrations of reactants and ensure specificity of the detected signal. Naturally, if the ligation is to take place on or in a living sample (as opposed to a fixed sample after prior metabolic incorporation of the reporter), reactants and by-products must not exert any perturbing or cytotoxic effect. Arguably, the first reaction that presented real bioorthogonal character was the Staudinger-Bertozzi ligation (SBL) between an azide reporter and a phosphine equipped with an electrophilic trap.[4,5] Although the SBL has found a great many applications over the years in bioconjugation and cell surface engineering, it quickly became apparent that such polar reactions were not the best candidates for intracellular detection given the often nucleophilic or electrophilic nature of biomolecules that, together with slow kinetics and phosphine sensitivity to oxidation, contribute to a disqualifying lack of specificity. The azide function, however, was revealed as a particularly suitable reporter group as it is virtually absent from most living organisms, is kinetically stable and does not induce cross-linking reactions with biomolecules. The focus shifted to pericyclic cycloadditions, in light of Sharpless' conceptual definition of ideal "*click chemistry*"[6] and of the ground-breaking discovery of the copper-catalysed azide-alkyne cycloaddition (CuAAC).[7,8]

2. THE MAJOR TRIAD OF BIOORTHOGONAL LIGATIONS

2.1. Copper-catalysed Azide-Alkyne Cycloaddition (CuAAC)

[3+2] Cycloadditions between azide dipoles and alkyne dipolarophiles proceed in the ground state according to a concerted, thermal $\pi 4s + \pi 2s$ mechanism, with perfect atom economy but high activation energy leading to poor kinetics. The limits of the original Huisgen reaction (requirement for organic solvents, high temperatures, long reaction times) were initially overcome for terminal alkynes by introducing Cu^{I} catalysts, boosting reaction rates up to a 10^{7} factor in all solvents including water, with the added benefit of 1,4-triazole regiospecificity (Scheme 1).



Scheme 1 (two-column graphic).

If all fields of chemistry rapidly embraced CuAAC as the quintessential click chemistry reaction in a myriad of applications, early efforts to use it for bioconjugation purposes *in vivo* were frustrated by adverse effects of the catalytic system. Due to Cu^{I}-induced cytotoxicity through generation of reactive oxygen species (ROS), cuprous ions have to be generated *in situ* from a cupric ion source (generally $CuSO_4$) with the help of excess reducing agent (typically, biocompatible ascorbate ions). Despite a few early reports, it was not until thorough optimization and standardization of protocols and the emergence of highly efficient water-soluble tetravalent copper-chelators of the *tris*(triazolylmethyl)amine family that CuAAC was truly considered a viable bioorthogonal reaction in eukaryotic models[9,10] (Scheme 1b). It probably remains to this day the most widely used reaction for metabolic labelling *ex vivo* as it presents many advantages while its drawbacks can usually be overcome with fine-tuning of reaction parameters. However, difficulties linked to copper toxicity, cell uptake and intracellular deactivation of the catalytic complex still often preclude applications that require long-term viability of labelled cells.

Firstly, both terminal alkynes and azides are absent from and non-reactive with living systems, stable, and very small. These moieties can both be used indifferently as reporter or as reactive group of the probe. Swapping their roles is indeed a facile affair[11,12] (Scheme 1a), making CuAAC very versatile and open to multiple labelling strategies.[13] Secondly, with an appropriate ligand its kinetics are further enhanced thus allowing low concentrations to be used. If the fluorochrome is chosen well, post-labelling removal of excess alkyne- or azide-functionalized probes is very efficient, leading to excellent signal-to-noise ratio. Thirdly, the ease of synthesis of azide- and alkyne-functionalized molecules is unsurpassed when compared to chemical partners in other bioorthogonal reactions, and many reagents are now commercially available and affordable. However, the complexity of the reacting system can be a deterring factor: not only must all parameters be optimized for each application to ensure success and reproducibility of labelling and minimization of copper-toxicity (namely, concentrations of Cu^{II} source, reducing agent and ligand, reaction time, washing protocols) making CuAAC less suited for one-shot experiments than for long-term methodology, but the need for great

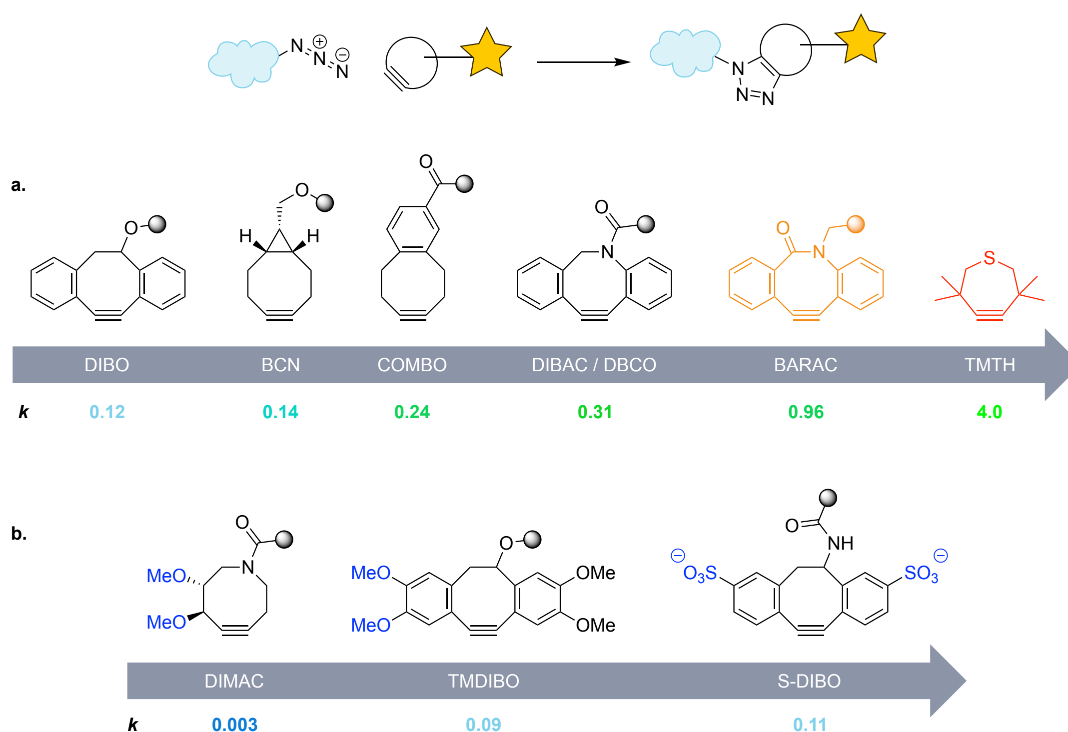
care in the preparation of the catalytic system can also be discouraging to chemistry-naïve users. For example, the copper source should be mixed with the ligand first so as to allow time for chelation to occur prior to adding the preformed complex and probe to the biological sample.[14] Although copper-induced cytotoxicity is well-known, the toxicity potentially caused by ascorbate itself should not be overlooked either, as its oxidation products have the potential to react with arginine, histidine or lysine residues in the sample. Ascorbate should thus be introduced last, to initiate the reaction, and its concentration optimized. Preparation and excessive storage of so-called "click buffers" containing all reagents is a frequent mistake that should be avoided under penalty of protein cross-linking and/or oxidation that may perturb cell function and lead to cell death,[14–16] and even in the case of CuAAC applied to fixed cells it is not advised to store and reuse solutions containing both copper and ascorbate. These considerations are important since they can be conducive to unsuccessful experiments and lesser reproducibility of results due to human error. Although this may seem trivial to the specialist, the ultimate aim of such reactions is easy technological transfer to routine users in biology and medicine, who might not always be aware of these aspects, and it constitutes a weak point. If the fact that CuAAC commercial kits are now widely available for the most frequent reporters is undoubtedly positive and highly contributes to the dissemination of the technique in various application fields, it must be stressed that the transparency of the standard protocols described could often be improved. The name and structure of the catalytic system components are usually undisclosed and the user is sometimes not even made aware of the nature of some reagents, which are simply referred to as numbered additives thus precluding full appreciation of the experiment and adaptation to potential issues. Last but not least, copper-toxicity cannot be entirely alleviated and remains the main issue when long-term viability of the sample is required.

The terminology used in research articles to describe biological experiments has to be considered here, as the frontiers between such terms as "living cells/tissue", "*in vivo*", "*ex vivo*" or "*in vitro*" can be somewhat blurred and subjective when it comes to metabolic labelling with the two-step chemical reporter strategy. Whereas the metabolic incorporation of chemical reporters must be carried out in living samples, it is not necessarily the case for the subsequent bioorthogonal probe ligation onto incorporated reporters. The latter constitutes a separate, second experimental step which can be applied indifferently to living samples (organisms, tissues, cells), fixed and/or permeated samples, biopsies, lysate extracts, SDS-PAGE gels, *etc*. depending on application and output. Naturally, the ultimate aim of such methodologies is to be fully achievable *in vivo*. For intracellular bioimaging purposes, however, cell fixation prior to bioorthogonal labelling is regularly the practical choice for technical reasons (*e.g.*, improved signal-to-noise ratio and specificity due to more efficient removal of excess probe, alleviation of toxicity issues, of safety issues for the operator in the case of highly pathogenic samples, mid- or long-term storage of labelled samples for re-observation, required transport, etc.).[17,18] It can also be advantageous for strategic reasons in the experiment design (*e.g.*, time-freeze of a given biological state for kinetic studies).[19] Typically, articles refer to a method using one of the above-mentioned terms only, but it is not unusual for the metabolic incorporation to be done *in vivo* and the bioorthogonal ligation *ex vivo* for example.

For applications in which the ligation step must be done in living samples, the biocompatibility of CuAAC has been greatly improved in the last few years, driven by advances in copper-chelating tools. Stable water-soluble Cu^{I} complexes have been shown to cross the cell membrane and promote intracellular cycloaddition in the absence of ascorbate[20] and multidentate ligand design continues with the recent description of an easily accessible ligand imparting CuAAC with improved kinetics and lesser cytotoxicity compared to BTTPS[21] (Scheme 1b). More elaborated chelation-assisted approaches have also emerged in recent years. Building on important work by Ting and co-workers, who reported the use of pyridinyl-based copper-chelating azides together with THPTA or BTTAA leading to a dramatic 10- to 25-fold decrease in Cu^{I} concentrations needed in cell-based experiments,[22] in 2014 Taran and co-workers proposed a new generation of integrated probes equipped with a reactive azide and a stronger *bis*-(triazolylmethyl)amino- moiety (Scheme 1c). They showed that using such "all-in-one" probes further enhances kinetics in a dramatic fashion and alleviates the need for external tetravalent ligands altogether, and demonstrated the intracellular capability of CuAAC by detecting paclitaxel-alkyne conjugates inside human hepatoma cells.[23] Although synthetic access to such "all-in-one" fluorophores is rather more complex and low-yielding compared to simple ligands, the same group later

suggested a scalable synthetic approach for these invaluable tools.[24] Similar probes based on a central sp^3 carbon rather than nitrogen were also described, providing an extra arm to the structure for increased versatility.[25] Most recently, Flon et al. used the same strategy and developed a small-sized all-in-one copper-chelating probe in which the fluorochrome itself also plays the role of ligand,[26] allowing intracellular imaging of alkyne-tagged biomolecules generated with an alkyne-maleimide heterobifunctional linker. In alternative heterogeneous approaches, Qu and co-workers recently reported two articles for intracellular CuAAC detection by fluorescence imaging and flow cytometry in which they used metal-organic framework Cu-complexes targeted at mitochondria[27] and Cu-nanoparticles grown on mesoporous carbon nanospheres activated by near-infrared light for Cu^I generation *in situ*,[28] respectively.

2.2. Strain-Promoted Alkyne-Azide Cycloaddition (SPAAC)



Scheme 2 (two-column graphic).

Bertozzi and co-workers took a different path. In yet another stroke of genius, they rapidly identified a clever way to avoid copper catalysis altogether[29,30] by unearthing an old forgotten variant involving cyclooctyne as the dipolarophile.[31] The distortion of its dihedral angle (163° in cyclooctyne instead of 180° in acyclic counterparts) confers an important ring strain to the alkyne (*ca*. 18 kcal.mol^{-1}). As a result, the purely thermal π4s + π2s pericyclic process of the strain-promoted alkyne-azide cycloaddition (SPAAC) is greatly accelerated and proceeds efficiently in aqueous medium without catalysis, at reaction rates that were originally on par with the SBL. Various groups then embarked on a quest to develop novel cycloalkyne derivatives with a view to further improve water solubility and kinetics, in a delicate exercise to reach a good balance between high reactivity and stability[32] (Scheme 2). The main approaches took advantage of substituent electronic effect or of increased sp^2 character in the ring to modulate the HOMO-LUMO gap dominating the SPAAC reaction and to decrease its activation barrier. A flurry of synthetic efforts and mechanistic studies which are outside the scope of this review led to great advances, and we refer the reader to the captivating review by Deb et al.[33] for more information on the theoretical grounds of strain-promoted cycloadditions. Among numerous examples, two families have emerged as the most popular reactants for SPAAC in recent years: DIBAC/DBCO dibenzoannulated derivatives, and cyclopropane-fused BCN derivatives, both of which have found numerous applications for metabolic labelling *in vivo* and for bioconjugation *in vitro*. Probes comprising DIBAC/DBCO

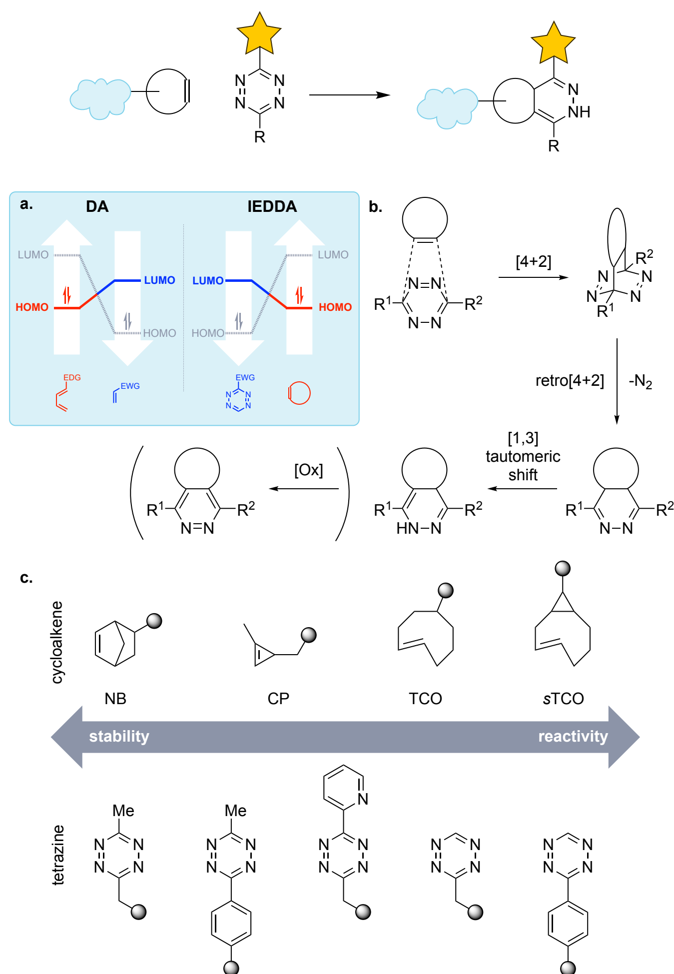
moieties have been the reactants of choice so far owing to their better kinetics with alkyl azides. Indeed, the presence of an endocyclic sp^2 nitrogen in the ring makes them faster than the first generation DIBO derivatives (Scheme 2a). However, dibenzoannulated compounds also have high lipophilicity and significant sizes, which lead to potential solubility and specificity issues. In contrast, BCN derivatives are smaller and less lipophilic in addition to being more easily synthesized, but are slower reactants with alkyl azides. Interestingly though, this kinetic trend is reversed when reacting cycloalkynes with aryl azides as this process is dominated by the reverse $LUMO_{AZIDE}$-$HOMO_{ALKYNE}$ interaction, therefore BCN reactants should be preferred then. More recently, a monobenzoannulated cyclooctyne (COMBO) was added to the toolbox as an interesting compromise.[34] Alternatively, high non-specific background due to hydrophobic trapping was also partially addressed with the development of hydrophilic cyclooctynes such as S-DIBO, TMDIBO or DIMAC derivatives, which display better bioavailability and may prove more relevant than DIBAC/DBCO in some contexts despite slower kinetics[35–37] (Scheme2b).

Strain-promoted cycloadditions, not requiring the use of metal catalysts, are much more biocompatible than CuAAC as they do not exert significant toxicity. In addition, the implementation of SPAAC for metabolic labelling is far simpler, as it does not require careful control of a complex catalytic system (which translates to better reproducibility of results for routine users) and cell uptake of multiple reactants. Yet, the SPAAC ligation does not only have advantages. Reduced toxicity comes at the cost of lesser specificity, as cyclooctynes may react with endogenous nucleophiles (cysteine residues and other thiols) thus generating much higher non-specific background than CuAAC, which is particularly hampering in the case of intracellular visualization.[38,39] This effect is not aided by the greater lipophilicity of dibenzoannulated cyclooctynes compared to terminal alkynes, which remains the major contributor to non-specificity in most cases. Additionally, the cyclooctyne moiety generates too much steric hindrance to allow its use as a reporter handle in most metabolic labelling approaches. Synthetic access to DIBAC/DBCO derivatives also remains much more difficult and costly, despite the commercial availability of stable cyclooctyne building blocks in recent years. Most importantly, despite tremendous synthetic efforts SPAAC still trails far behind other bioorthogonal reactions in terms of rate constants. Sadly, attempts to improve SPAAC kinetics seem to have stalled recently, as faster reactants such as BARAC derivatives or thiocycloheptyne TMTH suffer from poor stability, which has precluded their application to living cells so far. Although the impact of this copper-free click reaction can hardly be overpraised in cell-surface engineering *in vivo*, labelling with SPAAC has thus remained more challenging for intracellular purposes.

2.3. Inverse Electronic-Demand Diels-Alder Reaction (IEDDA)

Considering the very high potential of [4+2] cycloadditions, their rate acceleration promoted by ring strain was also soon explored and the Inverse Electronic-Demand Diels-Alder (IEDDA) reaction emerged as the fastest *in vivo* example to date. Historically, as one of the most widely applied transformations in chemistry in the twentieth century normal-demand Diels-Alder reactions (DA) involving electron-rich dienes and electron-poor dienophiles were always under scrutiny in the bioconjugation field, but with unmet expectations even *in vitro*. Additionally, the capacity of electron-poor dienophiles to act as Michael acceptors in addition reactions involving endogenous nucleophiles did not give great hope for specificity *in vivo* (for instance, maleimides are the most used reactants for thiol bioconjugation). On the opposite, the use of non-electrophilic electron-rich alkenes as dienophiles is allowed in reactions controlled by an inverted HOMO-LUMO interaction, making IEDDA a perfect candidate for bioorthogonality (Scheme 3a). In important contributions that are probably on equal footing with the discovery of CuAAC with regards to biological applications, in 2008 the groups of Fox and Hildebrand concomitantly reported the use of 1,2,4,5-tetrazine probes as dienes to label proteins equipped with *trans*-cyclooctene (TCO) dienophile reporters *in vitro*[40] and norbornene (NB) dienophile reporters *in vivo*,[41] respectively. Since then, the alkene-tetrazine IEDDA has undoubtedly been the most rapidly developing bioorthogonal reaction for several reasons. Unlike the normal-demand Diels-Alder reaction, it is an ultrafast irreversible process compelled to operate as a [4+2] / retro-[4+2] cascade by strong driving forces such as ring distortion release and nitrogen gas production, oft followed by 1,3-prototropic rearrangement or aromatization (Scheme 3b). Various strained cycloalkenes can thus be

used to create stable (dihydro)pyridazine linkages with the BOI without the need for metal-catalysis (Scheme 3c).



Scheme 3 (one-column graphic).

IEDDA kinetics are unequalled when using TCO reporters – even more so with ring-fused TCOs that are locked in highly reactive half-chair conformations such as *s*-TCO derivatives, however the intracellular stability of the latter is compromised as they are more susceptible to deactivation by thiol-based isomerization to the *cis* isomer. Factors affecting kinetic trends include the type of strained dienophile, the electronic effect of its substituents, as well as steric effects, but also electronic effects on the tetrazine diene.[33] Intriguingly, IEDDA reactions are dramatically faster in aqueous solutions because of strong solvent effects on hydrogen bonding and hydrophobic interactions. Rate constants up to 10.000-fold higher than CuAAC have been estimated (Figure 2) that permit the use of very low concentrations. These tremendous reaction rates opened new doors beyond fluorescence imaging, as they overcome poor pharmacokinetics of pretargeting agents and enable bioorthogonal labelling to be used with a number of tomographic techniques *in vivo* for nuclear medicine.[42,43] It is also a highly tunable and versatile reaction, as other strained alkenes such as NBs (commercially available as a wide range of building blocks) or cyclopropene derivatives (CP)[44] can advantageously replace TCOs in applications that do not require ultrafast reaction rates. The relatively small size of CP tags combined to still honorable kinetics that remain on par with most routine CuAAC protocols (1-27 $M^{-1}.s^{-1}$ in aqueous media)[45] makes them particularly useful for experiments in which the reporter must undergo multiple metabolic transformations by specific enzymes that do not incorporate large hydrophobic TCOs. They have been successfully used in various contexts for extracellular labelling[46–49] and have high potential for intracellular applications. Cycloalkynes such as BCN can also be used as dienophiles in IEDDA. Interestingly, the properties of tetrazine dienes can be tuned to obtain the most desirable balance between reactivity and stability in a given setting (*e.g.*, methyltetrazine derivatives exhibit less degradation over time than hydrogen substituted

tetrazines, at the expense of slower kinetics), and using the fastest reacting tetrazine derivative might not always be optimal.[50]

Importantly, some tetrazine reactants also come with the added benefit of "turn-on" fluorescence effect. Tetrazines can indeed act as quenchers toward a number of fluorochromes that are useful in UV-Vis fluorescence bioimaging, including BODIPY, rhodamine or coumarin derivatives (see summary and outlook). Such fluorogenic probes are particularly sought after in bioorthogonal chemistry, as they provide minimal fluorescence background thus reducing (in some cases alleviating) the need for arduous washing steps. While a small number of profluorophores are available for CuAAC and SPAAC ligations, the unique photophysical properties of tetrazines give the IEDDA reaction a clear advantage in that regard. For more comprehensive details on tetrazine ligations, we refer the reader to the review by Bernardes and co-workers.[45] Despite all the advantages that make IEDDA the most bioorthogonal of bioorthogonal reactions to this date, it is nonetheless not devoid of drawbacks. Its main strength over its competitors is indeed its tremendous kinetics, but these reactions involve large hydrophobic TCO reporters that are not be suitable for all purposes because of steric issues potentially perturbing enzymes or preventing metabolic incorporation. Even CP reporters, although smaller, are not as inconspicuous as azides or terminal alkynes and thus do not offer the same level of confidence on the relevance of their biological behaviour. In addition to their increased steric hindrance over the azide/alkyne pair of CuAAC, strained alkenes and tetrazines can also be less stable. For a given case, the practical choice of reacting partners for IEDDA might thus not prove to be faster than CuAAC. Finally, although some building blocks have become commercially available (usually at prohibitive costs), tetrazines, TCOs and CPs are not readily accessible and their synthesis is rather demanding.

Currently, CuAAC, SPAAC and IEDDA form the major triad of bioorthogonal reactions that cover the vast majority of applications. Each reaction has its *pros* and *cons*. Fortunately, they are rather complementary thus providing the chemical biology community with an effective toolbox: reaction, reporter and probe can be carefully selected depending on the limiting criteria of each experimental context (Figure 2).

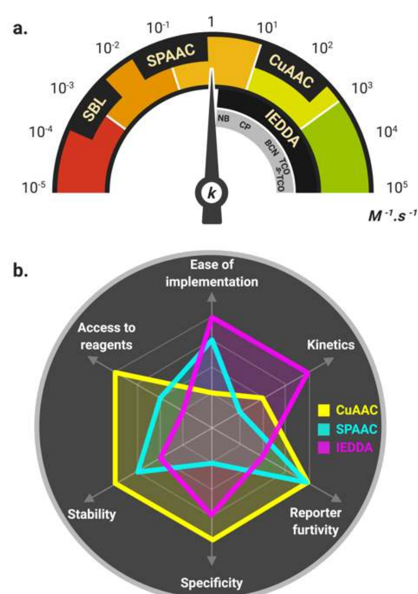
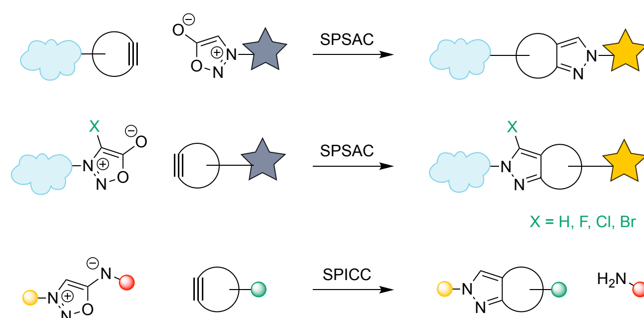


Figure 2 (one-column graphic).

2.4. Other cycloadditions

Among a multitude of candidate reactions, other cycloadditions have also emerged that show potential for bioorthogonality, including the photoinduced alkene-tetrazole conjugation[51–53] or the strain-promoted

alkyne-nitrone cycloaddition.[54] However, these reactions have not yet fulfilled their potential and have struggled so far to find their way to intracellular *in vivo* applications. Most recently, sydnone-based cycloadditions raised high interest owing to the stability of these mesoionic dipoles, and were revealed as promising contenders for future bioorthogonal applications[55] (Scheme 4). Strain-promoted sydnone-alkyne cycloaddition (SPSAC) was indeed shown to be suitable for biochemical applications *in vitro* by labelling BCN-tagged proteins with sydnone-functionalized fluorophores[56] and profluorophores.[57] Friscourt's group also successfully used sydnones as reporters to label BSA *in vitro* with a fluorogenic DIBO probe,[58] and SPSAC was shown by Taran's group to be compatible with intracellular labelling in *ex vivo* experiments[59] as well as with pretargeting PET imaging *in vivo*.[60] Interestingly, Taran and co-workers reported a variant reaction involving iminosydnones and strained cycloalkynes (SPICC),[61] which they applied inside cells with fluorogenic double turn-on effect.[62] Their ingenuous *"click-and-release"* strategy could well lead to new advances in drug delivery and theranostics as well as bioimaging in the coming years.



Scheme 4 (1-column graphic).

2.5. Cellular thiols standing in the way

Biological thiols are the main villain of this story as they impact most bioorthogonal reactions, a problem that is emphasized in the intracellular milieu. Not only are thiols responsible for copper ligand inactivation thus dramatically reducing the rate of CuAAC,[63] but they can also react with cyclooctynes via thiol-yne addition thus leading to nonspecific labelling and high background levels for SPAAC.[39] Combined to its slow kinetics, this significantly lowers the sensitivity of the SPAAC reaction as it demands that the labelled BOI be present at rather high concentration to be detectable. As an illustration of the reactivity between thiols and cyclooctyne, it is interesting to note that bioconjugation methods relying on the thiol-yne addition have been developed,[64] for example with a DBCO-reactive cysteine-containing heptapeptide tag. Some TCO reporters used in IEDDA reactions can undergo isomerization to the much less reactive *cis* isomer, a process that is facilitated in the presence of thiols.[65] IEDDA dienes are not spared, as the fastest tetrazines are also susceptible to thiol attack, leading to higher background and causing researchers to fall back on less reactive but more stable tetrazines. These problems can be partially overcome by deactivating endogenous thiols with reagents such as iodoacetamide or maleimides prior to the detection step, which has been shown to significantly improve intracellular SPAAC[39] and CuAAC[63] labelling for example, or by fixing cells so as to allow the efflux of highly concentrated thiolated species from the cell, but it remains a practical limitation. Among many factors influencing bioorthogonality, preventing competitive thiol reactivity has probably been the main reason behind the tremendous synthetic efforts to fine-tune reporter and probe handle pairs in the past ten years. Besides, thiol-based issues are not only linked to cross-reactivity with bioorthogonal handles. Indeed, cysteines can also react with monosaccharide reporters when used in the per-*O*-acetylated form, and generate off-target non-specific labelling in glycosylation studies. Overall, choosing a type of reporter and a bioorthogonal reaction to link it to a probe is very dependent of the application and type of BOI.

3. TARGETING PROTEINS WITH BIOORTHOGONAL CHEMISTRY

3.1. Site-specific vs. residue-specific labelling with unnatural amino acids

A variety of methods have been developed over the years to visualize protein function and trafficking inside living cells, including genetic fusion with autofluorescent proteins, self-labelling enzymes or peptide tags.[66] Despite their undeniable usefulness, the requirement for genetic fusion of an appended sequence may lead to perturbation of structure and activity of the protein of interest (POI). In addition, installing a probe at a chosen position of the POI remains a challenge with these methods, an increasingly important aspect given that super-resolution imaging is currently making great strides. The expansion of the genetic code to designer building blocks using the cellular machinery provides a complementary approach to tag a POI in a more furtive manner. Indeed, the inclusion of a synthesized unnatural amino acid (UAA) can be carried out in a site-specific manner,[12] an elegant method that requires three components in addition to the UAA: a reassigned codon inserted in the gene of interest (generally, the rare amber codon), a cognate transfer ribonucleic acid (tRNA) capable of decoding it, and an aminoacyl-tRNA synthetase (aaRS) capable of charging this orthogonal tRNA with the UAA. These exogenous elements are transfectionally introduced, leading to the incorporation of the UAA at the desired position of the biosynthesized POI (Figure 3a). This approach can also be more simply applied without genetic manipulation to target the entire proteome of a cell in a residue-specific manner (as opposed to site-specific) by hijacking endogenous aaRS/tRNA pairs, a much less selective method leading to global labelling that is nonetheless complementary and useful, for example in proteomic studies (Figure 3b). In the latter approach, the UAA is provided to the cells with a medium deprived of a relevant amino acid (typically methionine). As a result of selective pressure, the host chosen aaRS/tRNA pair partially or completely replaces the natural amino acid with the UAA in response to the corresponding codon.

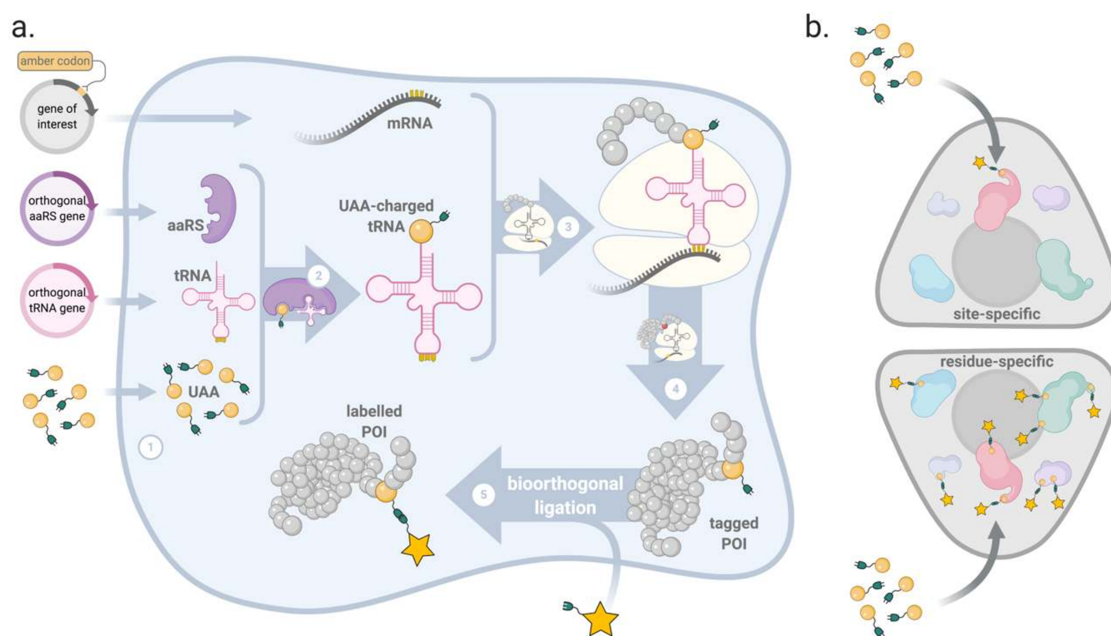


Figure 3.

A plethora of UAAs have been successfully introduced into proteins in living cells, including UAAs bearing a fluorochromic side chain.[67] The latter, however useful, are not always optimal for intracellular bioimaging in the current state of the art for sensitivity reasons (high-energy wavelengths in regions of the spectrum with marked autofluorescence in most samples) and for specificity reasons (high non-specific background due to difficult post-labelling removal of the excess fluorescent UAA from the cell, usually incubated at millimolar concentrations) that are both detrimental to signal-to-noise ratio. The versatility and flexibility introduced by bioorthogonal chemistry thus led chemists and biologists to work synergistically and develop a library of UAA equipped with bioorthogonal handles.

3.2. Copper-dependent labelling of intracellular proteins

Although CuAAC has been used numerous times to tag POIs in a residue-specific or site-specific manner at the cell surface or on fixed samples,[22] probe-ligation onto proteins localized in the challenging interior of living cells long remained out of reach due to Cu^{I} toxicity. Chen and co-workers demonstrated the efficiency of intracellular CuAAC in both the periplasm and cytoplasm of *E. coli* bacteria by site-specific labelling of an acid chaperone protein with ACPK (Scheme 5) and a solvatochromic fluorophore, using *tris*(triazolylmethyl)amine ligands.[68,69] In these important articles, BTTAA and BTTP were identified as the most suitable Cu-chelators, as they passively crossed membranes and significantly reduced the production of ROS thus allowing intracellular CuAAC in living bacteria. On the opposite, BTTPS was not fit for intracellular reaction due to lack of cell uptake, while TBTA complexes remained highly toxic to the bacteria. Recently, Cai and co-workers reported the first example of CuAAC protein labelling inside living human cells, improving uptake of the copper-chelator by generating a peptide-ligand hybrid.[63] To this aim, they equipped the *tris*(triazolylmethyl)amino scaffold with a cell penetrating sequence referred to as Tat peptide (RKKRRQRRR) and used it to label homopropargylglycine reporters (HpG) in a residue-specific manner. Importantly, the authors devised a method to quantify the uptake of the catalytic system components and estimate membrane and intracellular CuAAC yields by LC/MS-MS. They confirmed that tethering the ligand with a positively charged Tat peptide significantly increases the intracellular concentration of both ligand and copper when compared to BTTAA as the reference, and that cycloaddition does take place inside the cell. However, the CuAAC yields for cytosolic proteins were decreased by a 20-fold factor when compared to cell-surface proteins in untreated adhering cells, which was not the case for scraped cells with compromised membranes. As they identified an inverse correlation between glutathione concentration and catalytic efficacy, this difference was mostly attributed to competition of the ligand with endogenous copper-chelating moieties such as thiols, and the authors re-established intracellular CuAAC yields on par with that of membrane proteins by implementing *N*-ethylmaleimide capping of cellular thiols prior to bioorthogonal ligation. Although CuAAC has been somewhat shunned when it comes to *in vivo* applications because of the toxicity issues, these recent advances illustrate that it might still be a very serious candidate for protein labelling inside cells, the key factors being further development of optimized chelating structures. The development of more elaborated "all-in-one" probe-ligand-azide structures such as that developed by Taran and co-workers[23,24] might provide new avenues in this regard by allowing lower concentrations to be used.

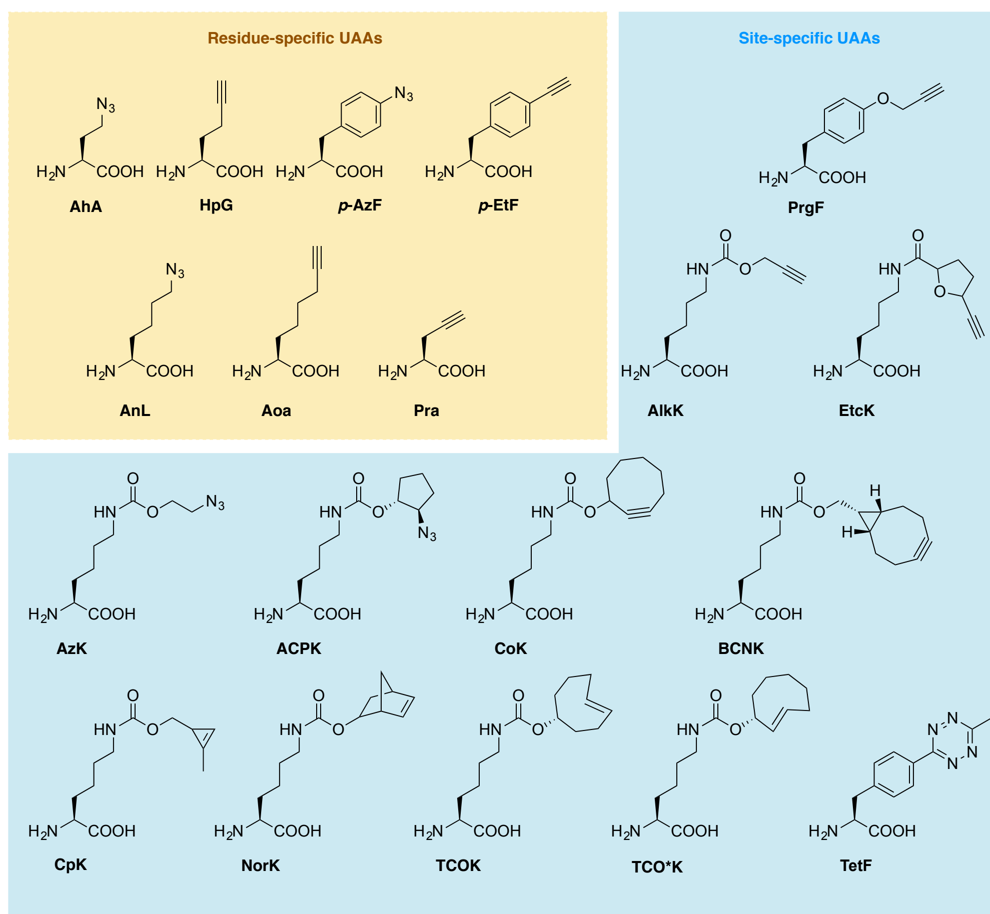
3.3. Copper-free labelling of intracellular proteins

Copper-free reactions have attracted much more attention for protein applications. Successful expansion of the genetic code to a number of IEDDA- and SPAAC-ready UAAs now allows site-specific incorporation of various unnatural residues such as cyclopropene (CP), norbornene (NB), *trans*-cyclooctene (TCO), cyclooctyne or bicyclononyne (BCN) moieties[70–74] (Scheme 5). A vast majority of such UAAs exploit pyrrolysyl-tRNA synthetase (PylRS)/tRNA pairs, which have the advantage of being orthogonal in both animal and bacterial cells, but Tyrosyl-tRNA (TyrRS)/tRNA systems can also be designed to accommodate aromatic UAAs. Reporter size does not appear to be a burden for metabolic incorporation in this approach as orthogonal aaRS/tRNA pairs that accommodate large groups can be engineered.[12] For the labelling step itself, reporter size is not an inconvenience either as, in fact, the modification of a single amino acid is markedly less invasive than in any other protein labelling method, even with strained alkene or alkyne reporters. Allied to the development of dyes that are increasingly tuned for intracellular detection of low biomolecule concentrations,[75,76] this has led to bioorthogonal labelling of various types of POIs inside living prokaryotic and eukaryotic cells alike. The ultrafast kinetics of TCO/tetrazine and BCN/tetrazine IEDDA ligations are particularly attractive, and the approach is even amenable to super-resolution microscopy.[73,74] Demonstrated on POIs of rather high expression levels, fluorescence background can however remain problematic for efficient labelling of less abundant proteins, an issue that might not be solely dependent on the choice of dye because of competitive labelling of TCO-functionalized POI versus TCO-charged tRNAs.[77] Some TCOs display a rather short half-life in the presence of intracellular thiols (TCOK), which is a limiting factor *in vivo*, but other isomers such as TCO*K[78] can exhibit higher stability, which is presumably due to shielding effect of the carbamate.[45] By inverting probe and reporter handle stable tetrazine-UAAs such as TetK, suitable for use with TyrRS/tRNA pairs, were developed to circumvent this issue while still benefiting from the rapid kinetics, and slower dienophiles with

higher resilience to thiols such as CPs can also be resorted to. The versatility of BCN, which can be used for both IEDDA or SPAAC ligations, is also of particular interest.

Sakin et al. successfully incorporated TCO*K and BCNK UAAs in HIV-1 envelop glycoproteins expressed in mammalian cells, then performed live-cell IEDDA to ligate a tetrazine-Cy5 probe in a trafficking study.[78] Although this work mostly focused on cell surface labelling, when using cell-permeable dyes they observed intracellular pools of the endocytosed engineered protein that were otherwise inaccessible to immunostaining alternatives. This interesting side result illustrates future possibilities to differentially stain extracellular and intracellular pools of a given POI using a unique bioorthogonal handle, by simply using a combination of cell-permeable and non-cell-permeable dyes.

The choice of tetrazine component must also be carefully guided, as the fastest kinetics do not necessarily translate to the best imaging data in the context of living cells. For example, methyl-substituted tetrazines have been preferred to their faster reacting hydrogen-substituted counterparts for some imaging studies in intracellular compartments, as their improved stability may largely compensate for the kinetic differential and offer more specific staining. Very recently, tetrazine-amino acids were also introduced as a reactive handle in a POI and intracellularly labelled with *s*-TCO probes at an even more impressive rate (80.000 $M^{-1}.s^{-1}$), showing great promise for future applications of IEDDA.[79]



Scheme 5 (2-column graphic).

Although SPAAC performs less efficiently than IEDDA, it has nevertheless been shown to be applicable for intracellular purposes as well despite sluggish kinetics.[80,81] For SPAAC labelling of azido-UAAs, lipophilicity is a crucial criterion in the choice of cyclooctyne for solubility, cell-permeability and aspecificity issues. The use of dibenzoannulated probes such as DIBAC/DBCO is challenging for these reasons while the less hydrophobic BCN derivatives react more slowly with alkyl azides than with aryl azides.[32] BCN can however be a valuable alternative even for aliphatic azides, as recently illustrated by Tirrell and Ho in bacteria.[82] Given the advances and advantages of IEDDA site-specific labelling, the interest in applying SPAAC for single-marking experiments

via genetic expansion of intracellular proteins has somewhat dwindled in the very recent past. However, the latter reaction is still of very high significance when it comes to multichannel bioimaging as both reactions are compatible and can be combined to track two independent reporters.[83,84]

3.4. Combining bioorthogonal chemistry and genetic tags

Interestingly, bioorthogonal chemistry is also suited for combinations with self-labelling enzymes such as SNAP-, CLIP- or Halo-tags. In 2015 Murrey et al. published an impressive article combining the use of genetically encoded HaloTags and strain-promoted bioorthogonal chemistry and comparing the reactive moieties of strain-promoted cycloadditions in various organelles of living cells,[85] and proposed silver-stabilized sTCO reagents with an improved stability/reactivity balance. The SNAP/CLIP-tag technology was also used in conjunction with IEDDA and SPAAC for dual "orthogonal/bioorthogonal" intracellular labelling, further enhancing the versatility and complementarity of this approach.[86] Admittedly, the use of bioorthogonal chemistry directed at self-labelling enzymes does not address the same concerns as genetic code expansion strategies. Indeed, these fusion proteins are as large as Tsien, Shimomura and Chalfie's intrinsically fluorescent proteins derived from *Aequorea victoria* and have the same potential to perturb the POI's function. However, these approaches still overcome some of their weaknesses (such as low brightness, limited spectrum coverage, need for long oxygen-dependent fluorochrome maturation or propensity for photobleaching), while avoiding the need for directed evolution of exogenous aaRS/tRNA pairs (specific to each UAA but orthogonal to endogenous pairs) and circumventing possible low efficiency of amber suppression. Beyond amino acid sequence, the localization and function of proteins is also largely dependent on post-translational modifications (PTMs), among which glycosylation plays a prominent role. However, PTMs are not encoded in a protein's gene and are only introduced at the maturation stage after its biosynthesis, leaving genetic fusion tags unable to shed light on their biological impact.

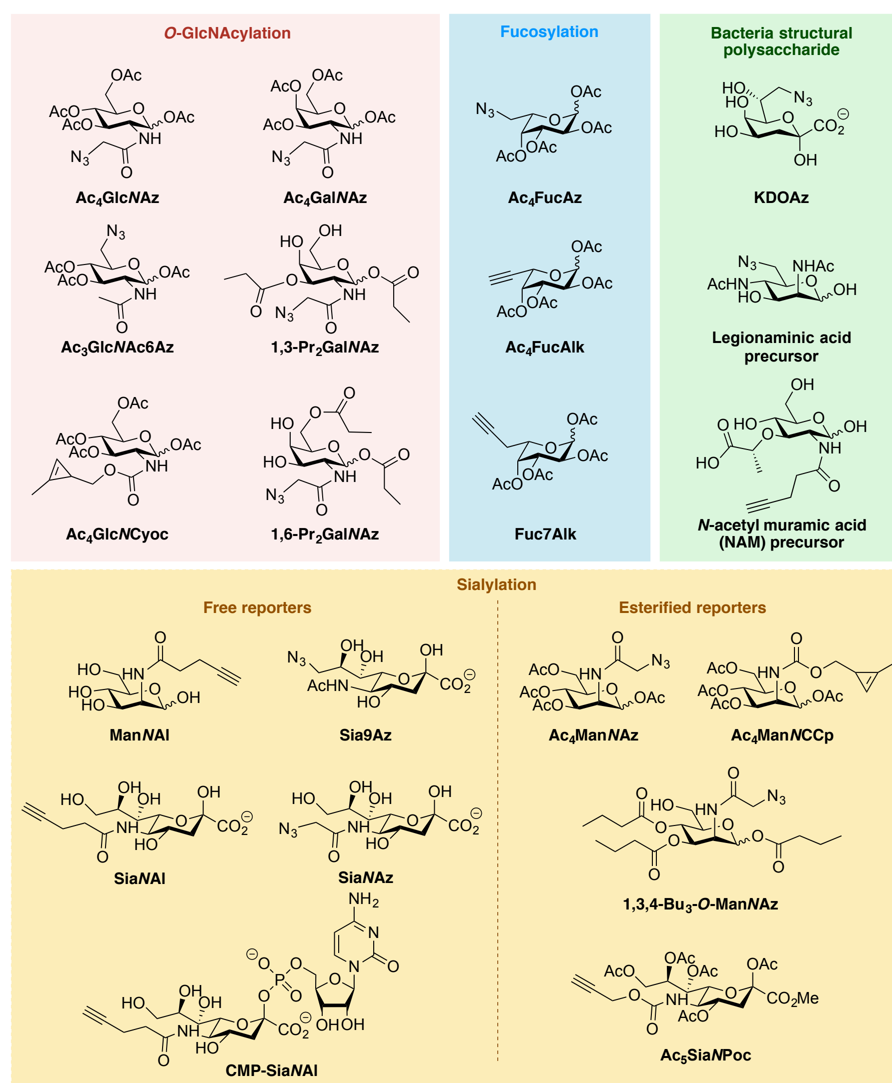
4. TARGETING GLYCANS WITH BIOORTHOGONAL CHEMISTRY

4.1. A plurality of glycosylations

Around 50% of eukaryotic proteins are modified with glycosylation PTMs (*i.e.*, covalent attachment of a mono-, oligo- or polysaccharide), which are crucial for protein folding and activity and mediate molecular recognition and interactions in various phenomena such as cell adhesion, infection, immunity, or cancer. In addition to glycoproteins, polysaccharides are also essential energy storage and structural components of cells, and are present in other structures such as glycolipids. Since carbohydrate biosynthesis is non-templated and depends on multiple metabolic enzymes that use monosaccharides or other small molecules as substrates, monitoring such pathways within cells has always been a difficult task. It is thus not surprising that, from the early days, bioorthogonal chemistry attracted much attention from the glycochemistry and glycobiology communities.

Pioneered by Bertozzi et al.,[3] the very first attempt at developing any bioorthogonal chemical reporter strategy involved the metabolization of keto-functionalized mannosamine derivatives into cell surface sialic acids before ligation of a hydrazide probe. Nine years later, the development of this concept dubbed metabolic glycan labelling (MGL) or metabolic oligosaccharide engineering (MOE) combined to newer bioorthogonal reactions had led to the first *in vivo* glycan labelling in zebrafish embryos by the same team.[87] Since then, MGL has been applied in myriad studies taking advantage of the main click reactions presented above: CuAAC, SPAAC and IEDDA. The available toolbox of monosaccharide reporters is rather comprehensive and now enables the application of MGL to a vast array of molecular targets in animal (including human),[88–90] bacterial[84,91–93] or plant biological models,[13,94,95] in various experiments ranging from cell culture to *ex vivo* tissues to whole organisms (Scheme 6). This includes detection of essential glycosylation types such as sialylation, fucosylation or *O*-linked β-*N*-acetylglucosaminylation (*O*-Glc*N*Acylation), as well as structural polysaccharides in plants (*e.g.*, pectins) and bacteria (*e.g.*, lipopolysaccharides or peptidoglycans). Enumerating every example of MGL application would be beyond the scope of this review and this topic has

been extensively covered. Interestingly however, most MGL reports address cell-surface expression of the engineered glycoproteins, for which probe ligation takes place in the extracellular medium thus partly avoiding the difficulties linked to cell entry, toxicity or specificity of reagents. Here, we focus on the hurdles to overcome when aiming at detecting glycans in the interior of living or fixed cells.



Scheme 6 (2-column graphic).

4.2. Monosaccharide reporter uptake and metabolic incorporation in eukaryotes

To label and image glycans in intracellular compartments, several aspects need to be considered. The reporter must be able to enter the cell, then undergo significant metabolic incorporation while in competition with its natural counterpart (which, admittedly, is also required in most chemical reporter strategies). In addition to controlling the lipophilicity-hydrophilicity balance of the reporter, this also entails that the introduced chemical handles have a size compatible with the substrate specificity of all enzymes involved (which must display a certain degree of promiscuity) and a stability compatible with the time necessary for metabolic transformation and incorporation into glycoconjugates (typically, several hours to several days of incubation). Understandably, terminal alkyne and azide reporters for CuAAC and SPAAC ligations have thus been largely favoured in the context of glycan labelling. IEDDA has been scarcely used for MGL so far, mostly for extracellular labelling and only with cyclopropene (CP) mini-tags[44,46,96–99] and norbornene (NB) reporters[100] as TCO reporters appear too bulky to be accommodated by most enzymes constituting glycosylation pathways. Although only one of these reports concerns intracellular cytosolic glycoproteins so far,[96] this state of affairs might perhaps change in the coming years, driven by fast advances in IEDDA methodologies.

Firstly, the reporter itself must display sufficient cellular uptake at physiologically relevant concentrations. Unfortunately, this is not always true for sugars in the free form because of high polarity, and only those that enter via efficient active transport in the appropriate biological model may be used as such. Limitations linked to cell entry of the reporter are usually addressed by esterification (typically acetylation) of the carbohydrate's free hydroxyl groups, generating a more apolar molecular spy able to cross the cell membrane by passive diffusion. Upon reaching the cytosolic compartment, the free monosaccharide is released after cleavage of the ester bonds by nonspecific esterases and intercepts the metabolic pathway under scrutiny (Figure 4). Used in numerous cell-surface applications, these tools are also exploited to track and observe intracellular glycosylation.

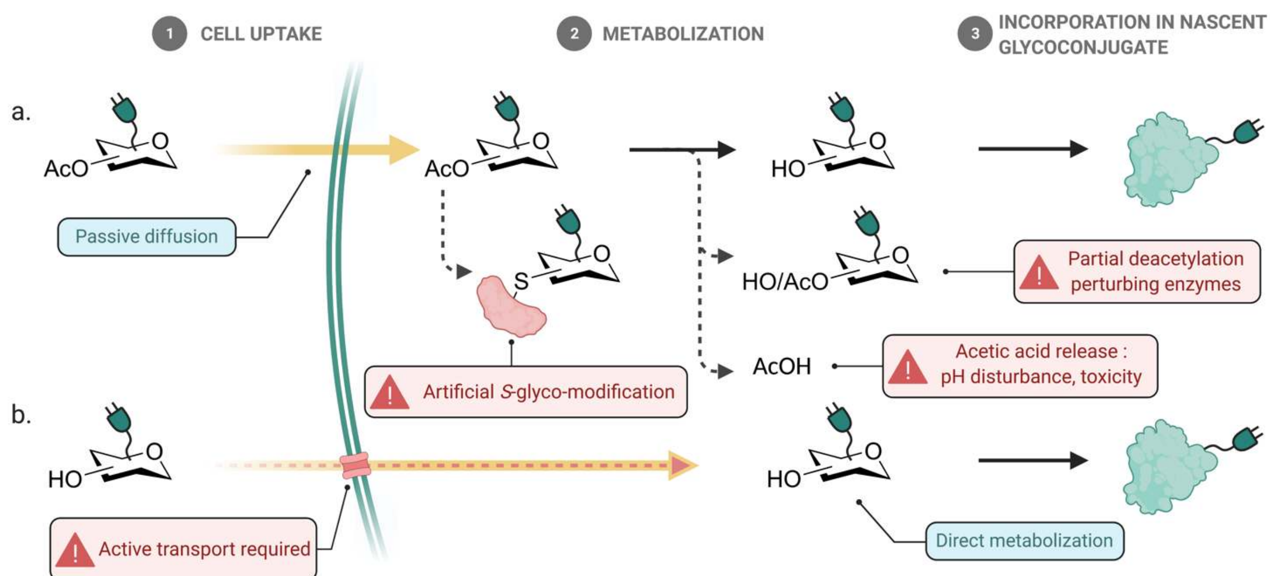


Figure 4 (2-column graphic).

For example, *per-O*-acetylated fucose analogues equipped with azide[101] or alkyne[102] handles were incorporated into human cancer cells via the fucose salvage pathway, and subsequent fluorescent labelling by CuAAC then allowed detection of fucosylated glycoconjugates in the *Golgi* compartments of the fixed cells. This illustrates the interchangeable nature of azide and alkyne tags when using CuAAC, although care should be taken when interpreting fluorescence data as toxicity as well as enzyme kinetics and incorporation rates might differ. In these studies, 6-azidofucose derivatives were converted to GDP-fucose with greater metabolic efficiency than their 6-alkynyl counterpart. Unfortunately, they also showed higher toxicity. In 2016, Kizuka et al. designed several alkynyl fucose analogues with various chain lengths, among which 7-alkynylfucose led to CuAAC labelling of *N*-glycan cores in mouse embryonic fibroblasts with increased efficiency and low associated toxicity.[103]

Another widespread form of PTM is *O*-Glc*N*Acylation, *i.e*. the covalent attachment of a *N*-acetylglucosamine (Glc*N*Ac) onto serine or threonine residues via a glycosidic linkage. Given that *O*-Glc*N*Acylated proteins are often cytosolic and that this PTM has an inherently more dynamic and temporary character than other glycosylation types, its bioimaging is rather difficult to address. To that aim, per-*O*-acetylated derivatives of *N*-acetylgalactosamine (Gal*N*Ac) and Glc*N*Ac have usually been employed, although Glc*N*Ac reporters might also lead to labelling of *N*-linked glycosylation to a certain extent.[104] *O*-Glc*N*Acylation was recently imaged using Förster resonance energy transfer (FRET) experiments between a fluorescent protein donor and a fluorochrome acceptor introduced selectively by incorporation of per-*O*-acetylated Gal*N*Ac or Glc*N*Ac reporters followed by SPAAC or IEDDA labelling.[96,105,106] This approach overcomes one of the main drawbacks of MGL, as it allows protein-specific imaging (in these examples, Tau, *O*-GlcNAc transferase and β-catenin) as opposed to detecting the whole pool of *O*-Glc*N*Acylated glycoconjugates. The use of Ac_4Glc*N*Cyoc for that purpose by Wittmann and co-workers[96] is promising as it demonstrates the feasibility of IEDDA for

intracellular MGL. Nevertheless, further investigation is needed to determine whether CP reporters are incorporated with the same efficiency as smaller tags or induce perturbation of the enzymes involved, and for more straightforward studies not involving FRET further investigation will be needed to confirm whether off-target labelling might be a hindrance.

If *N*-acetylated hexose reporters are often modified on the *N*-acyl chain for synthetic convenience, the position on which the reporter handle is introduced has indeed been shown to affect metabolic incorporation. In a detailed study reported by Pratt and colleagues,[107] Western-blot and flow cytometry were combined to analyse the incorporation of three peracetylated azide reporters. After CuAAC, both intracellular and extracellular glycoproteins were labelled using Ac_4Gal*N*Az and Ac_4Glc*N*Az, whereas only intracellular glycoproteins were labelled using Ac_3Glc*N*Ac6Az, leading the authors to identify the latter as being more specific of *O*-Glc*N*Acylation and highlighting the underestimated flexibility of this biosynthetic pathway. The reporter esterification pattern also plays an important role. In addition, Chen and colleagues showed that free cysteines can react with per-*O*-acetylated monosaccharides thus generating false positives in the identification of protein *O*-Glc*N*Acylation sites.[108] This phenomenon, originally named artificial *S*-glycosylation and since redubbed artificial *S*-glyco-modification (ASG), is not enzyme-mediated and was very recently suggested to produce 3-thiolated sugars thus generating off-target labelling (Figure 4). The proposed mechanism involves E1cb acetate β-elimination in basic microenvironments of proteins leading to the formation of α,β-unsaturated aldehydes, which then undergo Michael addition of cysteine residues.[109] The use of partially acetylated or propionylated derivatives (*e.g.*, 1,3-Pr_2Gal*N*Az or better still, 1,6-Pr_2Gal*N*Az)[109,110] was suggested to prevent this effect leading to a new generation of reporters with increased specificity for *O*-Glc*N*Acylation.

Sialylation has perhaps been the most studied glycosylation type with MGL,[88] although the overwhelmingly extracellular nature of sialoconjugate expression and function has made reports focusing on intracellular trafficking less frequent than cell-surface analyses. Sialylated glycan engineering can be envisaged according to two approaches: using modified *N*-acetylmannosamine (Man*N*Ac) reporters for an early interception of the metabolic pathway, or sialic acid (Sia) reporters for incorporation at a later stage of the biosynthesis, respectively.[11] Sia analogues have been suggested to be more specific of the sialic acid pathway and more versatile than Man*N*Ac analogues as they bypass three enzymatic conversions, but their synthesis is more demanding despite chemoenzymatic methods being available.[111] They are different tools for different purposes, but overall per-*O*-acetylated Man*N*Ac analogues have been the most popular choice in cell-surface studies. The technology enables intracellular detection, for example with the use of alkynyl derivative Ac_4Man*N*Al and CuAAC.[102] Swapping alkyne for azide, Steet and co-workers later used Ac_4Man*N*Az with copper-free SPAAC ligation inside living cells to evidence the abnormal accumulation of sialylated glycoconjugates in intracellular vesicles of Niemann–Pick Type C fibroblasts.[112]

Reporter per-*O*-acetylation is not a panacea because of potential ASG-linked non-specificity and because conversion of the per-*O*-acetylated reporter to its free form by intracellular esterases leads to liberation of acetic acid, which can be conducive to acidification of the intracellular milieu, metabolic perturbation and cytotoxicity. In addition, the production of partially acetylated forms also has the potential to deregulate sialic acid metabolism,[113] suggesting that other metabolic pathways could be affected too. When problematic, this can be addressed by modifying the length of the esterified chain and using partial acylation patterns, just as in *O*-Glc*N*Acylation studies. For instance,1,3,4-*O*-Bu_3Man*N*Az led to a labelling profile equivalent to those obtained with Ac_4Man*N*Az at concentrations 3- to 5-fold lower, with decreased toxicity.[114] More simply, free unprotected monosaccharides can also be used efficiently, which has the benefit of completely overcoming ASG at the cost of using higher concentrations. For example, unprotected Man*N*Al and Sia*N*Al incubated at 200-500 µM were shown to permit significant CuAAC MGL in human fibroblasts in studies that suggested the existence of a yet unidentified ManNAc membrane transporter, confirmed cellular entry of exogenous Sia*N*Al by endocytosis and sialin-mediated export to the cytosol, or evidenced lysosomal trapping of sialic acids due to sialin deficiency in the context of Salla disease.[19,111] Co-localization of the obtained signal is typically achieved by immunostaining with fluorescent antibodies targeted at an organelle-specific protein or by genetically engineering an organelle specific fusion-tag (*e.g.*, lysosome-associated membrane proteins

LAMP1/2, early endosome antigen EEA1, golgi marker TGN46). Azide derivatives such as Sia*N*Az can also be employed for intracellular SPAAC ligation, in which case using cyclooctyne-functionalized ratiometric fluorophores can prove useful to enhance the specificity (*e.g.*, pH-dependent for lysosomes, redox-dependent for mitochondria).[115] Encapsulation and vectorization of monosaccharide reporters can also be successfully implemented with the added benefit of cell-type selectivity when using ligand-targeted liposomes,[116] a method that allows crossing plasma membranes but also the blood brain barrier to target brain sialylation in living mice. Using liposome-encapsulated unprotected Sia9Az, Sun *et al*. detected labelled sialoglycans localized in endosomes and lysosomes and showed that this reporter is also exported to the cytosol by the sialin transporter.[117] Interestingly, exclusively extracellular methods such as selective exoenzymatic labelling (SEEL) can also give access to sialoglycan recycling inside cells after glycoconjugate internalization and accumulation in intracellular vesicles.[118] With this method, cytidine monophosphate activated sialic acid reporters (CMP-Sia) do not cross the membrane but are directly incorporated onto glycans expressed at the cell surface with an exogenous recombinant sialyltransferase.

4.3. Monosaccharide reporter uptake and metabolic incorporation in prokaryotes

Moreover, not every organism possesses an esterase activity sufficient to release free reporters inside the cells. This is especially the case in prokaryotes including *Escherichia coli* and *Legionella pneumophila*. Lipopolysaccharide targeted labelling was indeed achieved with unprotected KDO and legionaminic acid precursors, while per-*O*-acetylated reporters did not lead to significant signal.[91,92] The peptidoglycan of *E. coli* was also recently detected using a free azide-functionalized *N*-acetyl muramic acid.[93]

Few examples successfully applied MGL in bacteria with per-*O*-acetylated monosaccharides, an approach that is highly dependent on the species under scrutiny. For example, it was successful in *Helicobacter pylori,*[119,120] and more recently in commensal bacteria of the *Bacteroides* genus.[84] In this impressive work, Hudak et al achieved triple labelling of cell wall structures, employing free KDOAz to label the lipopolysaccharide via SPAAC ligation while simultaneously using a per-*O*-acetylated *N*-cyclopropenyl galactose to label capsular glycans by IEDDA (as well as a fluorescent D-amino acid specifically incorporated in the bacteria's peptidoglycan). The methodology was used to track host-bacteria interactions in the intestine of mice by *ex vivo* co-culture experiments. Although the targeted glycoconjugates are not intracellular, it is worth mentioning these studies here given the structural complexity of gram-negative cell walls. Besides, the labelled bacteria were also detected intracellularly inside invaded mouse macrophages.[84,93]

4.4. To fix or not to fix

In addition to the reporter, the probe (as well as the catalytic complex in the case of CuAAC) also needs to cross membranes at the bioorthogonal detection step and reach the location of the engineered glycoconjugate (*Golgi* apparatus, lysosomal or endosomal vesicles, cytosol, *etc*.). The reaction must occur with fast kinetics to allow low concentrations to be used, but not at the cost of reagent stability or specificity. Finally, unreacted fluorophore molecules need to be removed during washouts to avoid fluorescence background due to non-covalent trapping in hydrophobic regions. These washout processes can be tedious, particularly in the case of fragile samples that cannot withstand thorough cleaning (*e.g.*, living cells or whole organisms). In that regard, there is a high interest in developing fluorogenic click-ready probes.[121] When possible, cell-fixation with paraformaldehyde or other reagents can prove useful. Despite the methodological drive to develop selective chemistry in living cells, we would also like to remind the reader that fixing cells prior to bioorthogonal labelling may also be an effective, voluntary, useful strategy in order to study trafficking and metabolization kinetics,[19,111] rather than a necessary evil. In addition, it seems superfluous that the click chemistry step be carried out on living cells in the context of multicolour experiments involving other co-staining reagents that require cell fixation or permeabilization anyway. In a large majority of the examples cited here above that involve mammalian cells, samples were fixed prior to the bioorthogonal labelling step, after metabolic incorporation *in vivo* or *ex vivo*. This aspect comes in contrast to cell-surface engineering experiments in which cell entry of the probe is usually not desired, in order to avoid detection of intracellular species for specificity and sensitivity reasons.

5. TARGETING NUCLEIC ACIDS WITH BIOORTHOGONAL CHEMISTRY

Another important category of BOI for which bioorthogonal chemistry has aroused interest is nucleic acids. There are three major ways of introducing bioorthogonal handles on DNA or RNA strands. Firstly, the oligonucleotide sequence can be prepared by solid-phase synthesis, affording perfect control of the number and position of introduced bioorthogonal handles. However, the click-reactive moieties must be stable to the harsh conditions required during the synthesis, such as strong oxidants, acids or bases. Since stability of the reporter groups is usually inversely correlated with their reactivity, this is a limitation. The second method is the use of primer extension (PEX) to generate the sequence using DNA or RNA polymerases *in vitro*, with the major drawback of essentially having no control of the number and position of introduced tags. In these two methods, the synthesized oligonucleotide can be introduced into the cell by transfection to be later bioorthogonally labelled with the corresponding dye.[122] The third way uses small nucleoside reporters for metabolic labelling (or pronucleotides), similar to technologies used for proteins with amino acids or glycans with monosaccharides (Figure 5). This strategy avoids complex synthesis of the nucleic acid sequence by exploiting the metabolic machinery of the cell, but here again the number and position of labelled sites is not controlled. Intracellular by essence, the use of chemical reporter strategies to label and track DNA or RNA in the cellular environment has been extensively reviewed, and we refer the reader to two recent reviews by the groups of Wagenknecht and Rentmeister for more comprehensive details.[123,124]

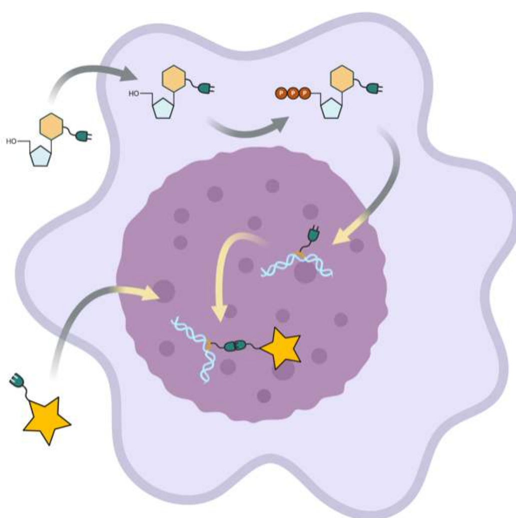
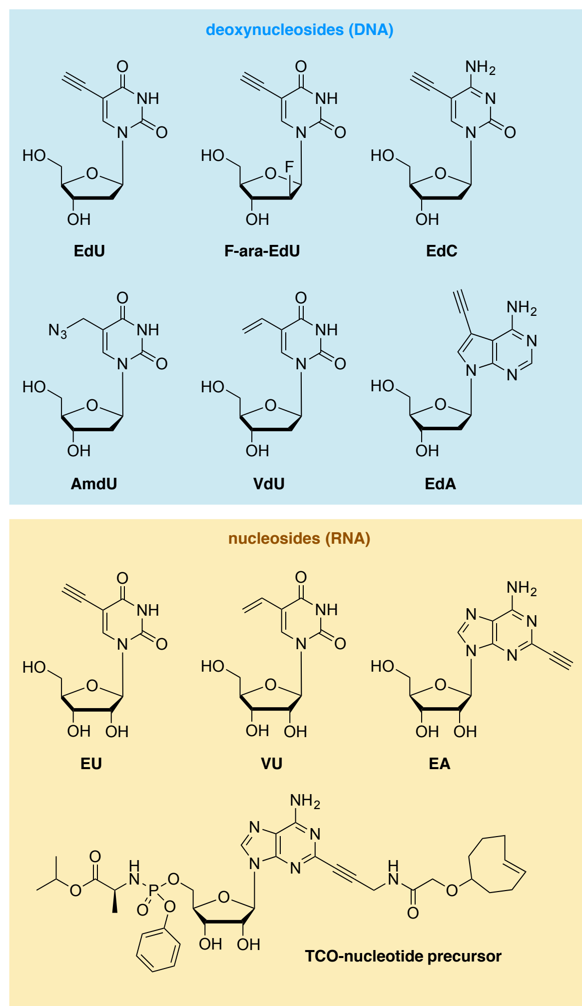


Figure 5 (1-column graphic).

Focusing here on metabolic labelling approaches, desoxyuridine and uridine reporters containing a propargyl group for *in vivo* visualization of both DNA and RNA were originally described by Salic and co-workers.[125,126] The metabolic deoxythymidine analogue bearing an ethynyl group at position 5 named EdU (Scheme 7) has since become the reporter of choice for DNA metabolic marking and can for example be used in combination with profluorescent azide probes.[127] Nevertheless, this reporter is a toxic antimetabolite, and a fluorinated version (F-ara-EdU) with minimal impact on genome function has been proposed, which can be used *in vivo* as demonstrated in zebrafish.[128] Deoxynucleosides functionalized for copper-free chemistry are also available. For example, azide derivative AmdU can be employed with CuAAC, but also with SPAAC thus allowing visualization in samples without fixation, and vinyl derivatives have also been reported (*e.g.*, VdU). Regarding purines, an alkynyl deoxyadenosine derivative (EdA) also allowed efficient DNA labelling, while its guanosine counterpart was not satisfactory.[129] Similarly, alkynyl nucleosides such as EU or EA are being used for RNA. Publications describing the detection of nucleic acids in living cells and organisms by metabolic labelling with modified nucleosides remain sparse, but is currently a very active topic of interest. For example, the alkene-tetrazine IEDDA ligation was recently used to detect RNA incorporated vinyl-functionalized

reporters (not using strain-promoted alkenes),[130,131] and a TCO-equipped adenosine derivative protected as an aryl phosphoramidate was also described.[132] Alkene-tetrazole photoclick chemistry was also successfully applied in zebrafish embryos.[133]



Scheme 7 (1-column graphic).

6. SUMMARY AND OUTLOOK

In summary, the field of bioorthogonal chemistry has forged ahead in two decades of existence. The main triad of reactions has found its way beyond methodological work and is rapidly providing new biological and clinical insight, which constitutes a testimony to the synergistic power of the chemical biology community. Although it is still early days, intracellular applications of CuAAC, SPAAC and IEDDA have allowed the detection of the main types of biomolecules in various organelles of eukaryotic cells and in bacteria (Figure 6). This type of chemistry has been best used in conjunction with metabolic chemical reporter strategies that allow inclusion of the chemical tag in a larger biomacromolecule during biosynthesis, although methods that allow incorporation within the cell by other means have also emerged, such as combining bioorthogonal chemistry with self-labelling enzymes, or simple cellular staining with small metabolites or drugs. A majority of the scientific production has concerned cell-surface marking as intracellular targets are considerably more challenging, but this review highlights that the latter have been increasingly reachable in recent years. Click-ready molecular probes with improved solubility now allow more efficient washouts that decrease non-covalent background, covering the full UV-visible spectrum and near-IR region. In particular, great expectations come from the continuing development of cell-permeable fluorogenic dyes. The nature of pericyclic mechanisms, which by essence deeply impacts π-systems, is indeed particularly suited to release fluorescence quenching effects (*e.g.*, photoinduced electron transfer) or modify fluorescence properties (wavelength shift) when the reactive moiety is conjugated with the fluorochromic core. Such profluorophores

appeared quite early on for CuAAC, like coumarins[134,135] and naphthalimides,[101,102] and more recently a range of azido-functionalized probes of the fluorescein, rhodamine and silicon rhodamine families designed by Bertozzi's group, termed the CalFluors,[121,136] became commercially available. Profluorescent cyclooctynes have also been reported for SPAAC ligations,[137–139] and there are numerous examples of profluorophores for IEDDA owing to the intrinsic fluorogenicity of cycloadditions involving aryltetrazines.[140–144]

The difficulties are linked to the higher complexity of the intracellular milieu that negatively impacts specificity, and nucleophilic biothiols in particular have been suggested as the main cause of interference with both reporters and probes. The challenges ahead for intracellular bioimaging thus also include further investigations aimed at quantifying and studying the mechanisms responsible for off-target covalent labelling. Such studies, specific to each type of reporter, would give crucial insight toward improving signal-to-noise ratio for the detection of intracellular targets that are expressed in low abundance.

Outside the three click chemistries under the spotlight here, there is also a number of new reactions knocking at the door, such as sydnone-based cycloadditions. Methodological development of bioorthogonal reactions still has a bright future ahead, for example to advance multicomponent marking. Although this aspect has not been emphasized here, it is noteworthy that a number of the cited examples implemented multiple labelling by combining one click reaction with other bioconjugation methods, or by combining two or more compatible bioorthogonal reactions. Optimization of reporters and reagents for the known reactions must continue, and new chemistries are being explored (not limited to cycloadditions). However, the field is now reaching a maturity that also allows routine use outside the chemistry or chemical biology laboratory. In that regard, the technological transfer would benefit from a more systematic production of standardized protocols, and such advances ought to further encourage the promotion of interdisciplinarity at the interface of chemistry, biology and medicine.

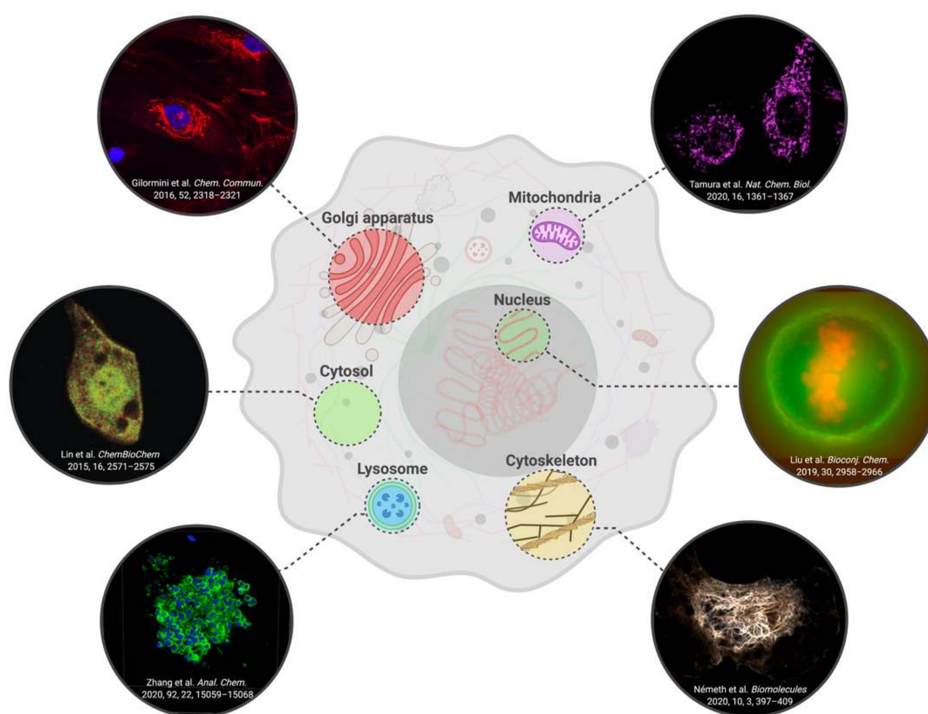


Figure 6.

REFERENCES

[1] O. Boutureira, G. J. L. Bernardes, *Chem. Rev.* **2015**, *115*, 2174–2195.
[2] I. Dovgan, O. Koniev, S. Kolodych, A. Wagner, *Bioconjugate Chem.* **2019**, *30*, 2483–2501.
[3] L. K. Mahal, K. J. Yarema, C. R. Bertozzi, *Science* **1997**, *276*, 1125–1128.
[4] E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007–2010.
[5] E. Saxon, J. I. Armstrong, C. R. Bertozzi, *Org. Lett.* **2000**, *2*, 2141–2143.
[6] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.
[7] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *114*, 2708–2711.
[8] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064.
[9] D. Soriano del Amo, W. Wang, H. Jiang, C. Besanceney, A. C. Yan, M. Levy, Y. Liu, F. L. Marlow, P. Wu, *J. Am. Chem. Soc.* **2010**, *132*, 16893–16899.
[10] C. Besanceney-Webler, H. Jiang, T. Zheng, L. Feng, D. Soriano del Amo, W. Wang, L. M. Klivansky, F. L. Marlow, Y. Liu, P. Wu, *Angew. Chem. Int. Ed.* **2011**, *50*, 8051–8056.
[11] S. J. Moons, G. J. Adema, M. T. Derks, T. J. Boltje, C. Büll, *Glycobiology* **2019**, *29*, 433–445.
[12] K. Lang, J. W. Chin, *Chem. Rev.* **2014**, *114*, 4764–4806.
[13] C. Simon, C. Lion, C. Spriet, F. Baldacci-Cresp, S. Hawkins, C. Biot, *Angew. Chem. Int. Ed.* **2018**, *57*, 16665–16671.
[14] V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angew. Chem. Int. Ed.* **2009**, *48*, 9879–9883.
[15] R. H. Nagaraj, D. R. Sell, M. Prabhakaram, B. J. Ortwerth, V. M. Monnier, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 10257–10261.
[16] S. C. Fry, *Biochem. J.* **1998**, *332*, 507–515.
[17] J. T. Ngo, E. M. Schuman, D. A. Tirrell, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 4992–4997.
[18] A. Shajahan, S. Parashar, S. Goswami, S. M. Ahmed, P. Nagarajan, S.-G. Sampathkumar, *J. Am. Chem. Soc.* **2017**, *139*, 693–700.
[19] P. A. Gilormini, C. Lion, D. Vicogne, Y. Guérardel, F. Foulquier, C. Biot, *J Inherit Metab Dis* **2018**, *41*, 515–523.
[20] J. Miguel-Ávila, M. Tomás-Gamasa, A. Olmos, P. J. Pérez, J. L. Mascareñas, *Chem. Sci.* **2018**, *9*, 1947–1952.
[21] Y.-Y. Guo, B. Zhang, L. Wang, S. Huang, S. Wang, Y. You, G. Zhu, A. Zhu, M. Geng, L. Li, *Chem. Commun.* **2020**, *56*, 14401–14403.
[22] 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.
[23] V. Bevilacqua, M. King, M. Chaumontet, M. Nothisen, S. Gabillet, D. Buisson, Cø. Puente, A. Wagner, F. Taran, *Angew. Chem. Int. Ed.* **2014**, *53*, 5872–5876.
[24] A. Sallustrau, S. Bregant, C. Chollet, D. Audisio, F. Taran, *Chem. Commun.* **2017**, *53*, 7890–7893.
[25] Y. Su, L. Li, H. Wang, X. Wang, Z. Zhang, *Chem. Commun.* **2016**, *52*, 2185–2188.
[26] V. Flon, M. Bénard, D. Schapman, L. Galas, P.-Y. Renard, C. Sabot, *Biomolecules* **2020**, *10*, 619–631.
[27] X. Qu, F. Wang, Y. Zhang, Z. Du, Z. Liu, L. Zhang, *Angew. Chem. Int. Ed.* **2019**, *58*, 6987–6992.
[28] Y. You, F. Cao, Y. Zhao, Q. Deng, Y. Sang, Y. Li, K. Dong, J. Ren, X. Qu, *ACS Nano* **2020**, *14*, 4178–4187.
[29] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.
[30] J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16793–16797.
[31] A. T. Blomquist, L. H. Liu, *J. Am. Chem. Soc.* **1953**, *75*, 2153–2154.
[32] J. Dommerholt, F. P. J. T. Rutjes, F. L. van Delft, *Top Curr Chem (Z)* **2016**, *374*, 16.

[33] T. Deb, J. Tu, R. M. Franzini, *Chem. Rev.* **2021**, DOI: 10.1021/acs.chemrev.0c01013.
[34] B. R. Varga, M. Kμllay, K. Hegyi, S. BØni, Pø. Kele, *Chem. Eur. J.* **2012**, *18*, 822–828.
[35] F. Friscourt, P. A. Ledin, N. E. Mbua, H. R. Flanagan-Steet, M. A. Wolfert, R. Steet, G.-J. Boons, *J. Am. Chem. Soc.* **2012**, *134*, 5381–5389.
[36] H. Stöckmann, A. A. Neves, S. Stairs, H. Ireland-Zecchini, K. M. Brindle, F. J. Leeper, *Chem. Sci.* **2011**, *2*, 932–936.
[37] E. M. Sletten, C. R. Bertozzi, *Org. Lett.* **2008**, *10*, 3097–3099.
[38] B. Amgarten, R. Rajan, N. Martınez-Sàez, B. L. Oliveira, I. S. Albuquerque, R. A. Brooks, D. G. Reid, M. J. Duer, G. J. L. Bernardes, *Chem. Commun.* **2015**, *51*, 5250–5252.
[39] R. van Geel, G. J. M. Pruijn, F. L. van Delft, W. C. Boelens, *Bioconjugate Chem.* **2012**, *23*, 392–398.
[40] M. L. Blackman, M. Royzen, J. M. Fox, *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519.
[41] N. K. Devaraj, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2008**, *19*, 2297–2299.
[42] R. Rossin, P. Renart Verkerk, S. M. van den Bosch, R. C. M. Vulders, I. Verel, J. Lub, M. S. Robillard, *Angew. Chem. Int. Ed.* **2010**, *49*, 3375–3378.
[43] R. Rossin, M. S. Robillard, *Curr Opin Chem Biol* **2014**, *21*, 161–169.
[44] D. M. Patterson, L. A. Nazarova, B. Xie, D. N. Kamber, J. A. Prescher, *J. Am. Chem. Soc.* **2012**, *134*, 18638–18643.
[45] B. L. Oliveira, Z. Guo, G. J. L. Bernardes, *Chem. Soc. Rev.* **2017**, *46*, 4895–4950.
[46] A.-K. Späte, H. Bußkamp, A. Niederwieser, V. F. Schart, A. Marx, V. Wittmann, *Bioconjugate Chem.* **2014**, *25*, 147–154.
[47] C. Lion, C. Simon, B. Huss, A.-S. Blevarcq, L. Tirot, D. Toybou, C. Spriet, C. Slomianny, Y. Guérardel, S. Hawkins, C. Biot, *Cell Chem. Biol.* **2017**, *24*, 326–338.
[48] M. Merkel, S. Arndt, D. Ploschik, G. B. Cserép, U. Wenge, P. Kele, H.-A. Wagenknecht, *J. Org. Chem.* **2016**, *81*, 7527–7538.
[49] T. S. Elliott, F. M. Townsley, A. Bianco, R. J. Ernst, A. Sachdeva, S. J. Elsässer, L. Davis, K. Lang, R. Pisa, S. Greiss, K. S. Lilley, J. W. Chin, *Nat. Biotechnol.* **2014**, *32*, 465–472.
[50] M. R. Karver, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2011**, *22*, 2263–2270.
[51] T. Carell, M. Vrabel, *Top Curr Chem (Z)* **2016**, *374*, 9. doi: 10.1007/s41061-016-0010-x.
[52] B. Lehmann, H.-A. Wagenknecht, *Org. Biomol. Chem.* **2018**, *16*, 7579–7582.
[53] R. K. V. Lim, Q. Lin, *Acc. Chem. Res.* **2011**, *44*, 828–839.
[54] D. A. MacKenzie, A. R. Sherratt, M. Chigrinova, L. L. Cheung, J. P. Pezacki, *Curr Opin Chem Biol* **2014**, *21*, 81–88.
[55] E. Decuypere, L. Plougastel, D. Audisio, F. Taran, *Chem. Commun.* **2017**, *53*, 11515–11527.
[56] S. Wallace, J. W. Chin, *Chem. Sci.* **2014**, *5*, 1742–1744.
[57] C. Favre, F. Friscourt, *Org. Lett.* **2018**, *20*, 4213–4217.
[58] C. Favre, L. de Cremoux, J. Badaut, F. Friscourt, *J. Org. Chem.* **2018**, *83*, 2058–2066.
[59] L. Plougastel, M. R. Pattanayak, M. Riomet, S. Bregant, A. Sallustrau, M. Nothisen, A. Wagner, D. Audisio, F. Taran, *Chem. Commun.* **2019**, *55*, 4582–4585.
[60] M. Richard, C. Truillet, V. L. Tran, H. Liu, K. Porte, D. Audisio, M. Roche, B. Jego, S. Cholet, F. Fenaille, B. Kuhnast, F. Taran, S. Specklin, *Chem. Commun.* **2019**, *55*, 10400–10403.
[61] S. Bernard, D. Audisio, M. Riomet, S. Bregant, A. Sallustrau, L. Plougastel, E. Decuypere, S. Gabillet, R. A. Kumar, J. Elyian, M. N. Trinh, O. Koniev, A. Wagner, S. Kolodych, F. Taran, *Angew. Chem. Int. Ed.* **2017**, *56*, 15612–15616.
[62] M. Riomet, K. Porte, A. Wijkhuisen, D. Audisio, F. Taran, *Chem. Commun.* **2020**, *56*, 7183–7186.
[63] S. Li, L. Wang, F. Yu, Z. Zhu, D. Shobaki, H. Chen, M. Wang, J. Wang, G. Qin, U. J. Erasquin, L. Ren, Y. Wang, C. Cai, *Chem. Sci.* **2017**, *8*, 2107–2114.
[64] C. Zhang, P. Dai, A. A. Vinogradov, Z. P. Gates, B. L. Pentelute, *Angew. Chem. Int. Ed.* **2018**, *57*, 6459–6463.
[65] N. K. Devaraj, *ACS Cent. Sci.* **2018**, *4*, 952–959.

[66] J. Liu, Z. Cui, *Bioconjugate Chem.* **2020**, *31*, 1587–1595.
[67] Z. Cheng, E. Kuru, A. Sachdeva, M. Vendrell, *Nat. Rev. Chem.* **2020**, *4*, 275–290.
[68] M. Yang, Y. Song, M. Zhang, S. Lin, Z. Hao, Y. Liang, D. Zhang, P. R. Chen, *Angew. Chem. Int. Ed.* **2012**, *51*, 7674–7679.
[69] M. Yang, A. S. Jalloh, W. Wei, J. Zhao, P. Wu, P. R. Chen, *Nat. Commun.* **2014**, *5*, 4981.
[70] T. Plass, S. Milles, C. Koehler, J. Szymański, R. Mueller, M. Wießler, C. Schultz, E. A. Lemke, *Angew. Chem. Int. Ed.* **2012**, *51*, 4166–4170.
[71] 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.
[72] A. Borrmann, S. Milles, T. Plass, J. Dommerholt, J. M. M. Verkade, M. Wießler, C. Schultz, J. C. M. van Hest, F. L. van Delft, E. A. Lemke, *ChemBioChem* **2012**, *13*, 2094–2099.
[73] C. Uttamapinant, J. D. Howe, K. Lang, V. Beránek, L. Davis, M. Mahesh, N. P. Barry, J. W. Chin, *J. Am. Chem. Soc.* **2015**, *137*, 4602–4605.
[74] T. Peng, H. C. Hang, *J. Am. Chem. Soc.* **2016**, *138*, 14423–14433.
[75] G. Lukinavičius, K. Umezawa, N. Olivier, A. Honigmann, G. Yang, T. Plass, V. Mueller, L. Reymond, I. R. Corrêa Jr, Z.-G. Luo, C. Schultz, E. A. Lemke, P. Heppenstall, C. Eggeling, S. Manley, K. Johnsson, *Nat. Chem.* **2013**, *5*, 132–139.
[76] A. Egyed, A. Kormos, B. Söveges, K. Németh, P. Kele, *Bioorg. Med. Chem.* **2020**, *28*, 115218.
[77] R. Serfling, L. Seidel, A. Bock, M. J. Lohse, P. Annibale, I. Coin, *ACS Chem. Biol.* **2019**, *14*, 1141–1149.
[78] V. Sakin, J. Hanne, J. Dunder, M. Anders-Össwein, V. Laketa, I. Nikić, H.-G. Kräusslich, E. A. Lemke, B. Müller, *Cell Chem. Biol.* **2017**, *24*, 635–645.
[79] H. S. Jang, S. Jana, R. J. Blizzard, J. C. Meeuwsen, R. A. Mehl, *J. Am. Chem. Soc.* **2020**, *142*, 7245–7249.
[80] K. E. Beatty, J. D. Fisk, B. P. Smart, Y. Y. Lu, J. Szychowski, M. J. Hangauer, J. M. Baskin, C. R. Bertozzi, D. A. Tirrell, *ChemBioChem* **2010**, *11*, 2092–2095.
[81] S. E. Geissinger, A. Schreiber, M. C. Huber, L. G. Stühn, S. M. Schiller, *ACS Synth. Biol.* **2020**, *9*, 827–842.
[82] S. H. Ho, D. A. Tirrell, *ACS Cent. Sci.* **2019**, *5*, 1911–1919.
[83] L.-L. Huang, K. Liu, Q. Zhang, J. Xu, D. Zhao, H. Zhu, H.-Y. Xie, *Anal. Chem.* **2017**, *89*, 11620–11627.
[84] J. E. Hudak, D. Alvarez, A. Skelly, U. H. von Andrian, D. L. Kasper, *Nat. Microbiol.* **2017**, *2*, 17099.
[85] 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.
[86] M. Macias-Contreras, H. He, K. N. Little, J. P. Lee, R. P. Campbell, M. Royzen, L. Zhu, *Bioconjugate Chem.* **2020**, *31*, 1370–1381.
[87] S. T. Laughlin, J. M. Baskin, S. L. Amacher, C. R. Bertozzi, *Science* **2008**, *320*, 664–667.
[88] P. R. Wratil, R. Horstkorte, W. Reutter, *Angew. Chem. Int. Ed.* **2016**, *55*, 9482–9512.
[89] C. Agatemor, M. J. Buettner, R. Ariss, K. Muthiah, C. T. Saeui, K. J. Yarema, *Nat. Rev. Chem.* **2019**, *3*, 605–620.
[90] K. K. Palaniappan, C. R. Bertozzi, *Chem. Rev.* **2016**, *116*, 14277–14306.
[91] J. Mas Pons, A. Dumont, G. Sautejeau, E. Fugier, A. Baron, S. Dukan, B. Vauzeilles, *Angew. Chem. Int. Ed.* **2014**, *53*, 1275–1278.
[92] A. Dumont, A. Malleron, M. Awwad, S. Dukan, B. Vauzeilles, *Angew. Chem. Int. Ed.* **2012**, *51*, 3143–3146.
[93] H. Liang, K. E. DeMeester, C.-W. Hou, M. A. Parent, J. L. Caplan, C. L. Grimes, *Nat. Commun.* **2017**, *8*, 15015.

[94] M. Dumont, A. Lehner, B. Vauzeilles, J. Malassis, A. Marchant, K. Smyth, B. Linclau, A. Baron, J. Mas Pons, C. T. Anderson, D. Schapman, L. Galas, J.-C. Mollet, P. Lerouge, *Plant J* **2016**, *85*, 437–447.
[95] Y. Zhu, X. Chen, *ChemBioChem* **2017**, *18*, 1286–1296.
[96] F. Doll, A. Buntz, A.-K. Späte, V. F. Schart, A. Timper, W. Schrimpf, C. R. Hauck, A. Zumbusch, V. Wittmann, *Angew. Chem. Int. Ed.* **2016**, *55*, 2262–2266.
[97] D.-C. Xiong, J. Zhu, M.-J. Han, H.-X. Luo, C. Wang, Y. Yu, Y. Ye, G. Tai, X.-S. Ye, *Org. Biomol. Chem.* **2015**, *13*, 3911–3917.
[98] C. M. Cole, J. Yang, J. Šečkutė, N. K. Devaraj, *ChemBioChem* **2013**, *14*, 205–208.
[99] D. M. Patterson, K. A. Jones, J. A. Prescher, *Mol. BioSyst.* **2014**, *10*, 1693–1697.
[100] A.-K. Späte, J. E. G. A. Dold, E. Batroff, V. F. Schart, D. E. Wieland, O. R. Baudendistel, V. Wittmann, *ChemBioChem* **2016**, *17*, 1374–1383.
[101] M. Sawa, T.-L. Hsu, T. Itoh, M. Sugiyama, S. R. Hanson, P. K. Vogt, C.-H. Wong, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12371–12376.
[102] T.-L. Hsu, S. R. Hanson, K. Kishikawa, S.-K. Wang, M. Sawa, C.-H. Wong, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 2614–2619.
[103] Y. Kizuka, S. Funayama, H. Shogomori, M. Nakano, K. Nakajima, R. Oka, S. Kitazume, Y. Yamaguchi, M. Sano, H. Korekane, T.L. Hsu, H.Y. Lee, C.H. Wong, N. Taniguchi, Cell Chem. Biol. 2016, 23, 782–792.
[104] B. W. Zaro, Y.-Y. Yang, H. C. Hang, M. R. Pratt, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8146–8151.
[105] W. Lin, L. Gao, X. Chen, *ChemBioChem* **2015**, *16*, 2571–2575.
[106] A. Kasprowicz, C. Spriet, C. Terryn, V. Rigolot, S. Hardiville, M. G. Alteen, T. Lefebvre, C. Biot, *Molecules* **2020**, *25*, 4501.
[107] K. N. Chuh, B. W. Zaro, F. Piller, V. Piller, M. R. Pratt, *J. Am. Chem. Soc.* **2014**, *136*, 12283–12295.
[108] W. Qin, K. Qin, X. Fan, L. Peng, W. Hong, Y. Zhu, P. Lv, Y. Du, R. Huang, M. Han, B. Cheng, Y. Liu, W. Zhou, C. Wang, X. Chen, *Angew. Chem. Int. Ed.* **2018**, *57*, 1817–1820.
[109] K. Qin, H. Zhang, Z. Zhao, X. Chen, *J. Am. Chem. Soc.* **2020**, *142*, 9382–9388.
[110] Y. Hao, X. Fan, Y. Shi, C. Zhang, D. Sun, K. Qin, W. Qin, W. Zhou, X. Chen, *Nat. Commun.* **2019**, *10*, 4065.
[111] P.-A. Gilormini, C. Lion, D. Vicogne, T. Levade, S. Potelle, C. Mariller, Y. Guérardel, C. Biot, F. Foulquier, *Chem. Commun.* **2016**, *52*, 2318–2321.
[112] N. E. Mbua, H. Flanagan-Steet, S. Johnson, M. A. Wolfert, G.-J. Boons, R. Steet, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 10207–10212.
[113] R. Schauer, G. Srinivasan, D. Wipfler, B. Kniep, R. Schwartz-Albiez (2011) O-Acetylated Sialic Acids and Their Role in Immune Defense. In: Wu A. (eds) The Molecular Immunology of Complex Carbohydrates-3. Advances in Experimental Medicine and Biology, **2011**, *705*, 525-548. Springer, Boston, MA.
[114] R. T. Almaraz, U. Aich, H. S. Khanna, E. Tan, R. Bhattacharya, S. Shah, K. J. Yarema, *Biotechnol. Bioeng.* **2012**, *109*, 992–1006.
[115] E. Zhang, Y. Shi, J. Han, S. Han, *Anal. Chem.* **2020**, *92*, 15059–15068.
[116] R. Xie, L. Dong, Y. Du, Y. Zhu, R. Hua, C. Zhang, X. Chen, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 5173–5178.
[117] Y. Sun, S. Hong, R. Xie, R. Huang, R. Lei, B. Cheng, D. Sun, Y. Du, C. M. Nycholat, J. C. Paulson, X. Chen, *J. Am. Chem. Soc.* **2018**, *140*, 3592–3602.
[118] T. Sun, S.-H. Yu, P. Zhao, L. Meng, K. W. Moremen, L. Wells, R. Steet, G.-J. Boons, *J. Am. Chem. Soc.* **2016**, *138*, 11575–11582.
[119] P. Kaewsapsak, O. Esonu, D. H. Dube, *ChemBioChem* **2013**, *14*, 721–726.
[120] E. L. Clark, M. Emmadi, K. L. Krupp, A. R. Podilapu, J. D. Helble, S. S. Kulkarni, D. H. Dube, *ACS Chem. Biol.* **2016**, *11*, 3365–3373.

[121] P. Shieh, V. T. Dien, B. J. Beahm, J. M. Castellano, T. Wyss, C. R. Bertozzi, *J. Am. Chem. Soc.* **2015**, *137*, 7145–7151.
[122] I. Ivancová, D.-L. Leone, M. Hocek, *Curr. Opin. Chem. Biol.* **2019**, *52*, 136–144.
[123] D. Ganz, D. Harijan, H.-A. Wagenknecht, *RSC Chem. Biol.* **2020**, *1*, 86–97.
[124] N. Klöcker, F. P. Weissenboeck, A. Rentmeister, *Chem Soc Rev* **2020**, *49*, 8749–8773.
[125] A. Salic, T. J. Mitchison, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2415–2420.
[126] C. Y. Jao, M. Roth, R. Welti, A. Salic, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15332–15337.
[127] T. Ishizuka, H. S. Liu, K. Ito, Y. Xu, *Sci. Rep.* **2016**, *6*, 33217.
[128] A. B. Neef, N. W. Luedtke, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20404–20409.
[129] A. B. Neef, F. Samain, N. W. Luedtke, *ChemBioChem* **2012**, *13*, 1750–1753.
[130] H. Liu, T. Ishizuka, M. Kawaguchi, R. Nishii, H. Kataoka, Y. Xu, *Bioconjugate Chem.* **2019**, *30*, 2958–2966.
[131] M. Kubota, S. Nainar, S. M. Parker, W. England, F. Furche, R. C. Spitale, *ACS Chem. Biol.* **2019**, *14*, 1698–1707.
[132] K. Wu, M. He, I. Khan, P. N. A. Okai, Q. Lin, G. Fuchs, M. Royzen, *Chem. Commun.* **2019**, *55*, 10456–10459.
[133] Y. Wu, G. Guo, J. Zheng, D. Xing, T. Zhang, *ACS Sens.* **2019**, *4*, 44–51.
[134] K. Sivakumar, F. Xie, B. M. Cash, S. Long, H. N. Barnhill, Q. Wang, *Org. Lett.* **2004**, *6*, 4603–4606.
[135] Z. Zhou, C. J. Fahrni, *J. Am. Chem. Soc.* **2004**, *126*, 8862–8863.
[136] P. Shieh, M. J. Hangauer, C. R. Bertozzi, *J. Am. Chem. Soc.* **2012**, *134*, 17428–17431.
[137] J. C. Jewett, C. R. Bertozzi, *Org. Lett.* **2011**, *13*, 5937–5939.
[138] F. Friscourt, C. J. Fahrni, G.-J. Boons, *J. Am. Chem. Soc.* **2012**, *134*, 18809–18815.
[139] J.-J. Shie, Y.-C. Liu, J.-C. Hsiao, J.-M. Fang, C.-H. Wong, *Chem. Commun.* **2017**, *53*, 1490–1493.
[140] N. K. Devaraj, S. Hilderbrand, R. Upadhyay, R. Mazitschek, R. Weissleder, *Angew. Chem. Int. Ed.* **2010**, *49*, 2869–2872.
[141] L. G. Meimetis, J. C. T. Carlson, R. J. Giedt, R. H. Kohler, R. Weissleder, *Angew. Chem. Int. Ed.* **2014**, *53*, 7531–7534.
[142] A. Wieczorek, P. Werther, J. Euchner, R. Wombacher, *Chem. Sci.* **2017**, *8*, 1506–1510.
[143] H. Wu, J. Yang, J. Šečkutė, N. K. Devaraj, *Angew. Chem. Int. Ed.* **2014**, *53*, 5805–5809.
[144] A. Vázquez, R. Dzijak, M. Dračínský, R. Rampmaier, S. J. Siegl, M. Vrabel, *Angew. Chem. Int. Ed.* **2017**, *56*, 1334–1337.

FIGURE LEGENDS

Figure 1. Inside or outside. Key differences to consider between intracellular and cell-surface bioorthogonal labelling chemical reporter strategies.

Scheme 1. Copper-catalysed Alkyne-Azide Cycloaddition (CuAAC). a) strategies. (i) the most commonly employed catalytic system uses separate tetravalent ligands to chelate copper. Cu^{+} ions are generated in situ by reducing Cu^{2+} with sodium ascorbate. Tag swap is facile.(ii) copper-chelating picolylazides used in conjunction with ligands improve kinetics and labelling performance when compared to (i). For most applications, a picolylazide probe may be reacted with the alkyne reporter. Tag swap is not facile, but is possible for proteins when using self-labelling fusion-tags, as shown by Ting and co-workers by introducing a picolylazide reporter on the BOI with a lipoic acid ligase (LplA) / LAP-peptide fusion-tag system.(iii) All-in-one azide probes comprising strong chelating moieties alleviate the use of separate ligands altogether thus further enhancing kinetics and performance of the labelling. Tag swap, i.e. all-in-one alkyne probes, has not been described yet. **b)** Frequently used tris(triazolylmethyl)amine ligands. **c)** New generation copper-chelating azides.

Scheme 2. Strain-Promoted Alkyne-Azide Cycloaddition (SPAAC). a) kinetics of the main reagent families. DIBO: dibenzocyclooctyne; BCN: bicyclononyne; COMBO: carboxymethylmonobenzocyclooctyne; DIBAC: dibenzoazacyclooctyne (also known as DBCO: dibenzocyclooctyne); BARAC:biarylazacyclooctynone; TMTH:tetramethylthioheptyne. Typically, intracellular experiments employ DIBAC/DBCO or BCN. BARAC and TMTH have better kinetics, but are too unstable. **b)** examples of hydrophilic cyclooctyne reagents. DIMAC: dimethoxyazacyclooctyne; TMDIBO: tetramethoxydibenzocyclooctyne; S-DIBO: sulfonylated dibenzocyclooctyne. Kinetic rates, measured in reaction with benzyl azide in organic (co)solvents, are given in $M^{-1}.s^{-1}$.

Scheme 3. Inverse Electronic-Demand Diels-Alder (IEDDA). a) Frontier molecular orbitals (FMO) analysis. Normal-demand Diels-Alder involving the $HOMO_{DIENE}$-$LUMO_{DIENOPHILE}$ interaction require electron-poor dienophiles bearing electron-withdrawing groups (EWG) that are also Michael acceptors susceptible to nucleophilic attack. In contrast, IEDDA allows the use of non-electrophilic dienophiles as they are controlled by the inverse $LUMO_{DIENE}$-$HOMO_{DIENOPHILE}$ transfer. The energy of the dienophile FMOs is raised by ring distortion, while FMOs of tetrazine dienes are lowered by electronic perturbation. **b)** IEDDA mechanism. The [4+2] / retro[4+2] cascade is usually followed by isomerization of the 4,5-dihydropyridazine cycloadduct to the 1,4-dihydropyridazine. Subsequent aromatization may occur depending on the reporter. **c)** structures of the main strained *E*-cycloalkene reporters and tetrazine probes ranked by reactivity/stability balance.

Scheme 4. Strain-Promoted Sydnone-Alkyne Cycloaddition (SPSAC) and Strain-Promoted Iminosydnone-Cycloalkyne Cycloaddition (SPICC). Sydnone-based pericyclic reactions have a high potential for future bioorthogonal applications. In particular, the click-and-release SPICC may provide new avenues in drug delivery and theranostics.

Figure 2. Differences and complementarity. a) comparison of the estimated kinetic range of SBL, CuAAC, SPAAC and IEDDA ligations. **b)** qualitative evaluation of important criteria influencing method design.

Figure 3. Metabolic labelling of proteins. a) site-specific incorporation of UAAs in a protein of interest (POI). The orthogonal aaRS/tRNA pair is introduced by transfection as well as the gene coding for the POI, modified with a reassigned amber codon (1). Once inside the cell, UAAs must be recognized by the exogenous aaRS but not by host aaRS to be selectively transferred onto the exogenous tRNA (2). The produced UAA-charged tRNA is used by ribosomes to insert the UAA at the desired position of a growing POI in response to the amber codon (3). After translation, folding, maturation and relocalization (4), the POI can be detected with biorthogonal labelling (5).**b)**In simpler residue-specific approaches, the UAA acts as a competitive substrate for an endogenous aaRS/tRNA pair leading to replacement of the natural amino acid. UAA-aaRS recognition can be enhanced by introducing mutations. The whole proteome is marked.

Scheme 5. Examples of UAAs that allow marking of POIs with CuAAC, SPAAC or IEDDA. ***Residue specific:*** Azidohomoalanine (Aha) and homopropargylglycine (Hpg) are the main methionine surrogates. Azidonorleucine (Anl), 2-aminooctynoic acid (Aoa) or propargylglycine (Pra) can also be used with mutant MetRS. *para*-azidophenylalanine (*p*-AzF) and *para*-ethynylphenylalanine (*p*-EtF) are phenylalanine surrogates using PheRS. ***Site-specific:*** PylRS/tRNA pairs can be engineered to allow site-specific incorporation of aliphatic azido-UAAs (*e.g.*, AzK, ACPK), alkynyl-UAAs (*e.g.*, AlkK, EtcK, CoK, BCNK), alkenyl-UAAs (*e.g.*, CpK, NorK, TCOK, TCO*K) analogues of pyrrolysine. Orthogonal TyrRS/tRNA pairs can be designed to accommodate tetrazine-UAAs (TetF) or alkynyl-UAAs (PrgF).

Scheme 6. Selection of monosaccharide reporters that target glycosylation.

Figure 4. Per-*O*-acetylated monosaccharides versus free monosaccharides. Apolar per-*O*-acetylated reporter uptake is facilitated by passive diffusion thus allowing lower concentrations to be used for metabolic incorporation, but can lead to increased background because of off-target ASG. Enzymatic activity in the pathway under scrutiny as well as others might also be perturbed by pH variations or interactions with

partially deacetylated intermediates. In contrast, unprotected monosaccharides directly enter the metabolic pathway with minimum perturbation but require active transport and typically higher concentrations impacting the metabolic flux.

Figure 5. DNA and RNA metabolic labelling. A cell-permeable nucleoside or pronucleotide equipped with a bioorthogonal reactive handle is activated as a nucleotide triphosphate by the cellular machinery then incorporated into nascent DNA or RNA strands.

Scheme 7. Examples of reporter nucleosides and deoxynucleosides.

Figure 6. Exploring the intracellular map with bioorthogonal chemistry. The three main click reactions have allowed detection of specific targets in various organelles of eukaryotic cells as well as in bacteria. (top right) Mitochondrial phosphatidylcholine labelling achieved in HeLa cells by incorporation of azido-choline followed by SPAAC with a DBCO-tetraethylrhodamine probe (magenta) [Tamura et al. Nat. Chem. Biol. 2020, 16, 1361–1367]. (middle right) Dual-labelling of DNA and RNA achieved by incorporation of EdU and VrU (incorporated in DNA and RNA, respectively) followed by CuAAC with Alexa594-azide (red) and IEDDA with tetrazine-biotine then streptavidin-Alexa Fluor 488 (green) [Liu et al. Bioconj. Chem. 2019, 30, 2958−2966]. (bottom right) Labelling of cytoskeletal vimentin (Vim) by incorporation of VimLys-BCN-mOrange in COS-7 cells followed by IEDDA with a fluorogenic tetrazine probe (blue) along with detection of the mOrange signal (orange) [Németh et al. Biomolecules 2020, 10, 3, 397–409]. (top left) Labelling of sialylated glycoconjugates achieved in human fibroblasts by incorporation of SiaNAl followed by CuAAC with AzidoFluor 545 (red) [Gilormini et al. Chem. Commun. 2016, 52, 2318–2321]. (middle left) Labelling of glycosylated EGFP-B-catenin achieved in MDCK II cells by incorporation of Ac4GalNAz followed by CuAAC with TAMRA-alkyne and FLIM-FRET intensity-weighted lifetime detection of the EGFP signal [Lin et al. ChemBioChem 2015, 16, 2571–2575]. (bottom left) Labelling of lysosomal sialylated glycoproteins achieved in HeLa cells by incorporation of Sia9Az and Lamp1-GFP followed by SPAAC with DBCO-Lyso-Blue probe (blue) along with detection of the GFP signal (green) [Zhang et al. Anal. Chem. 2020, 92, 22, 15059–15068].

BIOGRAPHIES

Vincent Rigolot obtained his M.Sc. in bioorganic chemistry and chemical biology in 2019 at the University of Lille (France). He then joined the Chemical GlycoBiology (ChemGB) group at the Unité de Glycobiologie Structurale et Fonctionnelle (UMR 8576 – UGSF) where he is currently working on his Ph.D. under the guidance of Prof. Christophe Biot and Dr. Cedric Lion. His research focuses on the design, synthesis and application of novel bioorthogonal chemical tools for bioimaging, aimed at studying mammal and bacterial sialylation phenomena.

Christophe Biot has a scientific background characterized by a very strong thematic mobility reflecting his desire to explore the fields of chemistry, biochemistry and biophysics. He obtained his Ph.D. in synthetic chemistry (1995-1998) under the supervision of Prof. J. Brocard at the University of Lille. His collaboration with Dr. D. Dive at the Institut Pasteur de Lille allowed him to acquire solid knowledge at the interface between chemistry and biology. In order to extend his training in biochemistry, he then did a first post-doctoral fellowship (2000-2001) in the team of Dr. E. Davioud-Charvet at the Institute of Biology of Lille. He did a second post-doctoral fellowship (2001-2004) in molecular modelling at the Ecole Polytechnique in Bruxelles (Belgium) with Pr. M. Rooman. He was then recruited at the University of Lille in 2004 as assistant professor where he developed new areas of research, including the study of mechanisms of action of anti-infective agents. After one year (2010-2011) hosted by the CNRS, he then integrated the Institute of Structural and Functional Glycobiology. He was appointed Professor in 2012. Currently, Christophe Biot coordinates the Chemical GlycoBiology (CheGB) team with scientists displaying complementary know-how from organic synthesis to cell biology. His research topics cover several aspects of high significance in chemical biology.

Cedric Lion obtained is Ph.D. in medicinal chemistry from the School of Pharmacy at the University of Nottingham (UK) in 2005, under the guidance of Prof. Malcolm F. G. Stevens and Prof. Andrew D. Westwell. After post-doctoral fellowships at the Centre for Biomolecular Sciences (Nottingham, UK) and at the Institut National des Sciences Appliquées (Rouen, France), he was appointed associate professor at the University of Lille (France) in 2008. In 2014, he joined the Institute of Structural and Functional Glycobiology (UMR8576 – UGSF) in the Chemical Glycobiology team to develop the bioorthogonal chemistry axis. Focused on developing synergistic collaborations at the interface of chemistry and biology, his research interests include the development of multiple click bioimaging methods to decipher glycosylation (and more particularly sialylation) in human and bacterial models, as well as cell wall biopolymers in plants (in particular, lignins). He is currently the director of the Erasmus Mundus Joint Master Degree Advanced Spectroscopy in Chemistry funded by the European Union.

KEYWORDS

Bioorthogonal chemistry ; Metabolic Labelling ; Bioimaging ; In vivo chemistry ; Click chemistry ; Intracellular

FRONTISPIECE