MÉMOIRE PRÉSENTÉ POUR L'OBTENTION DU DIPLÔME DE

DOCTORAT EN CHIMIE ORGANIQUE

Université de Lille

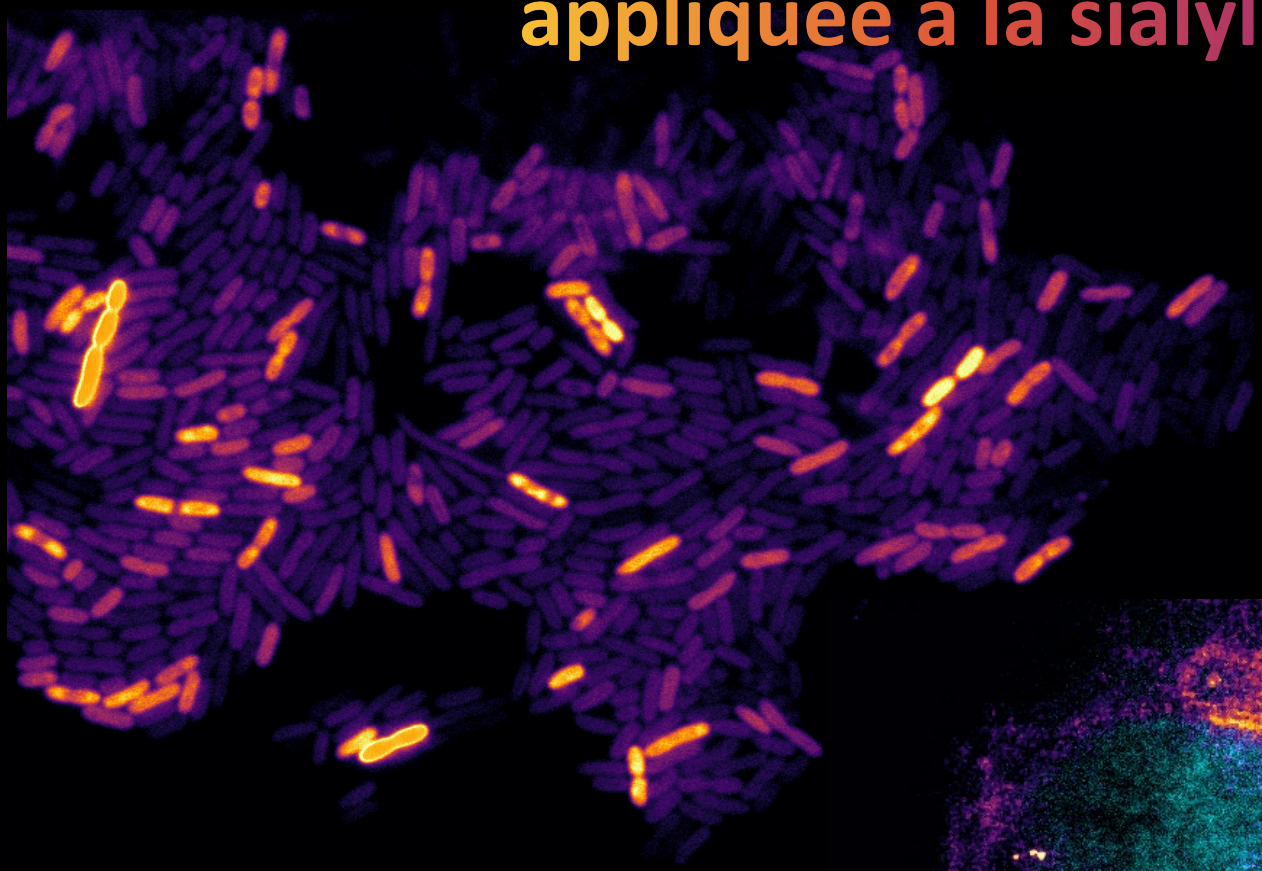
VINCENT RIGOLOT

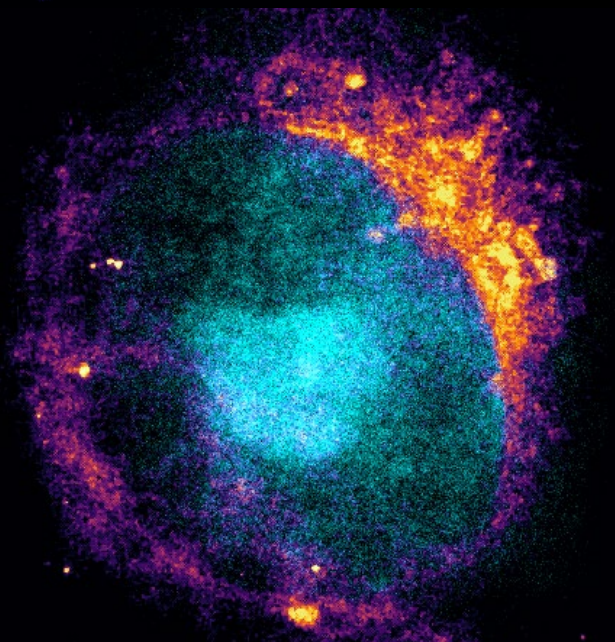
UMR 8576 CNRS

Unité de Glycobiologie Structurale et Fonctionnelle

École Doctorale Biologie Santé de Lille

Développements méthodologiques pour la chimie bioorthogonale appliquée à la sialylation





Thèse soutenue le 11 Décembre 2023 *devant le jury composé de :*

Président, Rapporteur	Dr. Florence Mahuteau	Directrice de Recherche	Institut Curie
Rapporteur	Pr. Eric Defrancq	Professeur des Universités	Université Grenoble Alpes
Co-directeur de thèse	Pr. Christophe Biot	Professeur des Universités	Université de Lille
Co-directeur de thèse	Dr. Cédric Lion	Maître de Conférences HDR	Université de Lille

Ces travaux ont été effectués au sein du laboratoire :
Unité de Glycobiologie Structurale et Fonctionnelle (UGSF, UMR CNRS 8576)
Université de Lille - Cité Scientifique, Bâtiment C9, Avenue Mendeleïev
59655, Villeneuve d'Ascq, France

Projet ANR-18-CE07-0042 financé par l'Agence Nationale de la Recherche

REMERCIEMENTS

Je tiens tout d'abord à adresser mes remerciements au Dr Florence Mahuteau et au Pr Eric Defrancq, pour avoir accepté de constituer le jury qui évaluera ces travaux de thèse. Je remercie également le Dr Sylvain Bohic et le Pr Mathieu Sollogoub pour avoir fait partie de mon comité de suivi, et le Dr Yann Guérardel pour m'avoir accueilli au sein de l'UGSF.

Merci à Christophe, pour m'avoir permis d'intégrer l'équipe il y a près de 5 ans, pour m'y avoir donné l'opportunité de participer à un projet passionnant, et pour avoir gardé *faith* durant toutes ces années.

Merci à Cédric, pour m'avoir guidé et épaulé sans relâche depuis le début, aussi bien en célébrant les instants glorieux, qu'en se tenant en véritable rempart lorsque je faisais face à des difficultés. *Aka'magosh, Lo'Gosh mog g'thazag cha. Lok'tar ogar!*

Merci à Angélina, pour avoir été une coéquipière incomparable en étant toujours à mes côtés et en m'ayant permis d'envisager ce qui me paraissait insurmontable.

Merci à Sylvain et Dima à l'ESRF pour leur accueil lors des différentes sessions au synchrotron, et à Aude et Michel au LASIRE pour avoir donné toute une nouvelle dimension à une partie de mon projet.

Merci à Corentin, pour m'avoir toujours permis de matérialiser mes idées ou besoins les plus insolites concernant les microscopes ou le traitement d'image, ainsi qu'à Yannick R., pour avoir résolu tant de mystères bactériologiques que j'avais fini par considérer insondables.

Merci à Clémence, qui m'a aidé dès mon arrivée au laboratoire, dès que je l'ai vue y revenir, et qui continue même quand je n'y suis plus.

Merci à Veronica pour avoir participé aux toutes premières manipulations de mon projet, et pour être revenue participer aux toutes dernières des années plus tard.

Merci à Anne-Sophie, Dorothée, Céline et Marlène pour leurs nombreux conseils en culture cellulaire, ainsi qu'à Olivier pour m'avoir formé aux manipulations en laboratoire de microbiologie.

Merci à Fred C., Fred K., Isabelle, Christophe M., et toutes les personnes qui m'ont accompagné pendant les années que j'ai passées au laboratoire. Merci à Audrey, Yannick, Louis-David, Jodie, Mathieu et Ryma pour tous les moments *hors* de celui-ci.

Merci à ma famille, et à mes parents, qui je l'espère seront aussi surpris qu'heureux de lire ces lignes, pour avoir attisé ma curiosité depuis tout petit, et pour m'avoir soutenu en toutes circonstances.

Merci à Lola, acolyte immémorial, pour sa confiance et son soutien indéfectible dans l'accomplissement des quêtes les plus périlleuses.

Merci aux illustres membres de la P2F, pour leur organisation.

This was a triumph.

I'm making a note here: "HUGE SUCCESS"

It's hard to overstate my satisfaction.

Aperture Science

We do what we must because we can.

For the good of all of us.

Except the ones who are dead.

But there's no sense crying over every mistake.

You just keep on trying till you run out of cake.

And the science gets done, and you make a neat gun

For the people who are still alive.

I'm not even angry.

I'm being so sincere right now.

Even though you broke my heart and killed me.

And tore me to pieces.

And threw every piece into a fire.

As they burned it hurt because I was so happy for you!

Now, these points of data make a beautiful line.

And we're out of beta; we're releasing on time.

So I'm glad I got burned.

Think of all the things we learned

For the people who are still alive.

"Still Alive", Jonathan Coulton

SOMMAIRE

LISTE DES FIGURES

Ac_4GalNAz : *N*-azidoacétylgalactosamine

Ac^4ManNAz : *N*-azidoacétylmannosamine

ADIBO : Azadibenzocyclooctyne

BARAC : Biarylazacyclooctynones

BCN : Bicyclo[6.1.0]nonyne

CCM : Chromatographie sur couche mince

CMP : Cytidine Monophosphate

CUP : Chaperone-Usher Pilus

CSS : CMP-Sialic acid synthetase

CuAAC : Copper-Catalyzed Azide-Alkyne Cycloaddition

DAPI : 4',6-diamidino-2-phénylindole

DCU : Dicyclohexylurée

DBCO : Dibenzocyclooctyne

DIC : *N*,*N*'-Diisopropylcarbodiimide

DMF : Diméthylformamide

DIBAC : Dibenzoazacyclooctyne

DIFO : Difluorocyclooctynes

EDC : 1-Éthyl-3-(-3-diméthylaminopropyl)carbodiimide

ESRF : European Synchrotron Radiation Facility

FRET : Transfert d'Énergie par Résonance de Förster

GFP : Green Fluorescent Protein

GNE : UDP-GlcNAc 2-épimérase

ICP-MS : Spectrométrie de masse couplée à plasma inductif

IEDDA : Réaction de Diels-Alder à demande électronique inverse

Kdn : Acide 2-ceto-3-désoxy-D-glycéro-D-galacto-nononique

LB : Lysogeny Broth

LPS : Lipopolysaccharide

ManNAc : *N*-acétyl-D-mannosamine

ManNAc-6-Az : 6-désoxy-6-azido-*N*-acétyl-D-mannosamine

MOE : Marquage Métabolique des Oligosaccharides

Neu5Ac : Acide *N*-acétylneuraminique

Neu5Gc : Acide *N*-glycolylneuraminique

NHS : *N*-Hydroxysuccinimide

O-GlcNAc : O-N-acétylglucosylamination

OGA : O-GlcNAcase

OGT : O-GlcNAc Transférase

PEG : Polyéthylène glycol

PEP : phosphoénolpyruvate

PSA : Acide polysialique

RMN : Résonance magnétique nucléaire

SiaNAl : Acide *N*-pentynoylneuraminique

SiaNAz : Acide *N*-azidoacétylneuraminique

TCO : Transcyclooctène

UV : Ultra-violet

I. MARQUAGE DES BIOMOLÉCULES

Au cours des vingt dernières années, l'avènement de la génomique et des techniques de biologie moléculaire telles que l'utilisation de gènes rapporteurs (*e.g.* GFP, luciférase, …) pour le suivi des protéines ont permis de grandes avancées dans notre compréhension des processus cellulaires. Néanmoins les lipides, les glucides et autres métabolites ne sont pas encodés génétiquement, rendant ces techniques inappropriées pour leur étude. Les voies de signalisation biologiques impliquent souvent des mécanismes de régulation complexes et des rétrocontrôles multipartites. De plus, de nombreuses modifications post-traductionnelles (MPT) telles que, par exemple, les glycosylations, la sulfatation, la phosphorylation, ou les lipidations modulent la fonction, la durée de vie ou encore la localisation des protéines.[1] Afin d'appréhender les processus cellulaires dans toute leur complexité, il est ainsi nécessaire de développer des technologies permettant de suivre toutes les classes de biomolécules dans le vivant.[2] De nombreux efforts ont été fournis par la communauté scientifique pour répondre à ces besoins, notamment grâce aux outils de la chimie, inscrivant ces recherches dans une mouvance relativement nouvelle : la chémobiologie. Dans ce contexte, de nouvelles techniques pour la visualisation et la détection de toutes les classes de biomolécules, y compris les glucides, lipides, mais également d'autres types de métabolites ou molécules exogènes (médicaments), ont ainsi été mises au point.

Historiquement, la principale approche permettant de détecter une biomolécule d'intérêt (BI) est d'utiliser la bioconjugaison, un ensemble de méthodes chimiques permettant de créer un lien covalent avec la BI, et ainsi de l'immobiliser sur une surface, de la vectoriser, ou de la détecter en la reliant à une sonde moléculaire observable (la principale catégorie de sondes moléculaires utilisée en laboratoires de recherche étant les fluorophores).

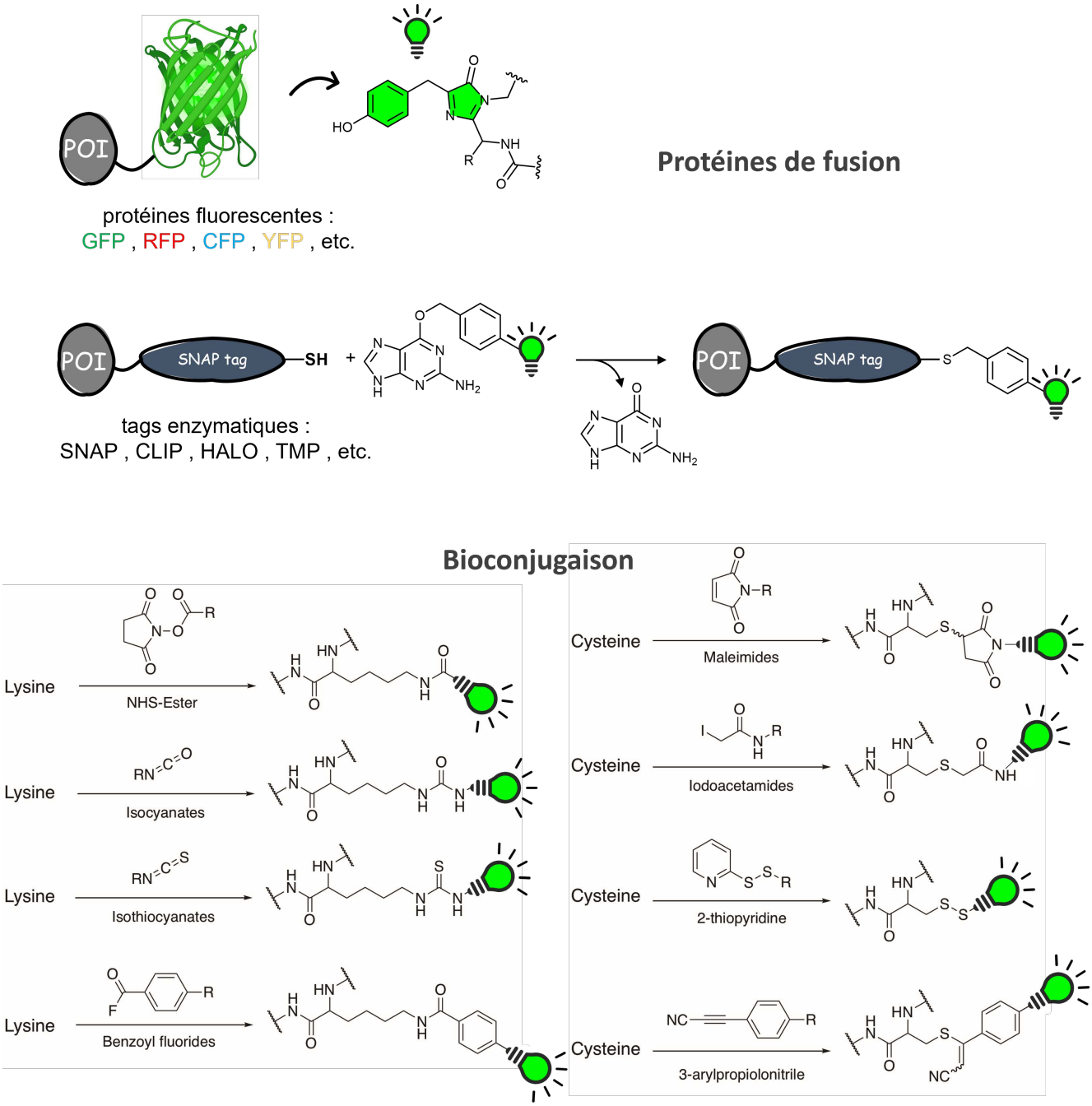


Figure 1 Principales approches pour le marquage des biomolécules. L'introduction de gènes rapporteurs encodant une protéine fluorescente ou un tag enzymatique dans la séquence génétique d'une protéine d'intérêt permet de l'observer in situ. Les composés chimériques obtenus par ce biais sont appelés protéines de fusion. Une autre approche, la bioconjugaison, consiste à lier une sonde observable par le biais de réactions chimiques spécifiques à certains groupements communément retrouvés au sein des composés biologiques. Pour des raisons de spécificité ces réactions sont souvent limitées à des applications in vitro.

De nombreuses réactions de bioconjugaison spécifiques à certains groupements fonctionnels ont ainsi été développées, notamment via l'utilisation d'isothiocyanates ou d'esters de *N*-hydroxysuccinimide réagissant avec les amines, de maléimides réagissant avec les thiols, ou d'hydrazides et d'oximes réagissant avec les carbonyles.[3,4] Cependant, ces réactions interfèrent avec les groupements chimiques naturellement présents dans les systèmes vivants, les rendant uniquement adaptées à la préparation de bioconjugués *in*

vitro avant leur introduction en milieu biologique. Les outils moléculaires générés par cette approche (appelés bioconjugués), tels que les anticorps fluorescents, les conjugués enzymes anticorps, ou les acides nucléiques marqués, trouvent aujourd'hui de nombreuses applications en biologie[5] (Figure 1).

Bien qu'adaptée aux biomolécules de taille conséquente (protéines, acides nucléiques), la bioconjugaison présente cependant ses limites dans le cas où la BI est de petite taille (monosaccharide, lipide, métabolite, etc.). En effet, les sondes utilisées (molécules organiques ou organométalliques, nanoparticules, etc.) présentent un encombrement stérique important comparativement à ce type de molécule. Créer un bioconjugué à partir d'un monosaccharide ou d'un lipide n'est donc pas pertinent, en particulier pour des applications *in vivo* visant à mettre en évidence le transport et/ou la métabolisation de la BI. En effet, les propriétés physico-chimiques du bioconjugué sont inévitablement altérées, et sa reconnaissance par des récepteurs, des protéines de transport ou des enzymes s'en trouvent fortement perturbée, voire annihilée.

En 2022, le prix Nobel de Chimie a été attribué à Morten Meldal, K. Barry Sharpless et Carolyn Bertozzi pour leur travail pionnier dans le développement de la chimie click et de la chimie bioorthogonale.[6] Ces concepts seront introduits dans la section suivante, les travaux de thèse que j'ai effectués s'inscrivant directement dans ce cadre.

II. LA STRATÉGIE DU RAPPORTEUR CHIMIQUE

Une alternative envisagée pour contourner les limites précédemment citées est la stratégie du rapporteur chimique. Cette stratégie, introduite par l'équipe de Carolyn R. Bertozzi en 1997 pour l'étude des glucides[7], repose sur la capacité des enzymes impliquées dans ces voies de métabolisation à tolérer de petites modifications structurales de leurs substrats.[8] Son principe se résume en deux étapes fondamentales (Figure 2) :

1 – l'*incorporation métabolique* d'un analogue de la BI (ou d'un de ses précurseurs métaboliques) portant un groupement fonctionnel synthétique choisi (le rapporteur chimique). Cette modification structurale doit être la plus anodine possible pour que la

molécule soit reconnue comme le serait le précurseur naturel (petite taille, charge neutre, stable et inerte vis-à-vis des molécules constituant les systèmes vivants), permettant ainsi de détourner la voie métabolique ciblée et d'incorporer le substrat portant le rapporteur chimique dans une structure biologique complexe (par exemple, une glycoprotéine).

2 – la *ligation* chimique d'une sonde moléculaire permettant de détecter le rapporteur chimique préalablement incorporé. La réaction utilisée doit permettre de créer une liaison covalente stable entre le rapporteur et la sonde au sein d'un échantillon biologique complexe (e.g. lysat cellulaire, cellules, tissus, organes, …) de manière rapide, spécifique et efficace. En comparaison de la bioconjugaison classique, dans ce cas la réaction chimique doit donc cibler certaines biomolécules, sans interférer avec l'ensemble des autres molécules et groupements fonctionnels endogènes constituant les systèmes vivants (*e.g.* amines, thiols, carboxylates, esters, ...). De plus, la réaction doit avoir lieu en conditions physiologiques (température, pH, milieu aqueux) et sans être toxique pour l'organisme au sein duquel elle se déroule. Les réactions chimiques répondant à ces critères exigeants appartiennent à un nouveau sous-domaine appelé *chimie bioorthogonale*, terme défini par Carolyn R. Bertozzi.

L'approche en plusieurs étapes apporte ainsi plusieurs avantages : la séparation nette entre l'incorporation du rapporteur et sa détection permet d'éviter l'utilisation de sondes contraignantes lors de l'incorporation et donc de minimiser les effets de la modification structurale sur les propriétés de la BI. Même dans les rares cas où la bioconjugaison classique pourrait convenir, le modèle obtenu par la stratégie du rapporteur chimique est de fait plus fidèle au système biologique étudié. De plus, cette stratégie permet d'effectuer des marquages longue durée sans avoir à se soucier des pertes en intensité de fluorescence des fluorophores (phénomène de photoblanchiment), ou du temps de demi-vie des radioisotopes par exemple, les sondes pouvant être liées juste avant l'acquisition.

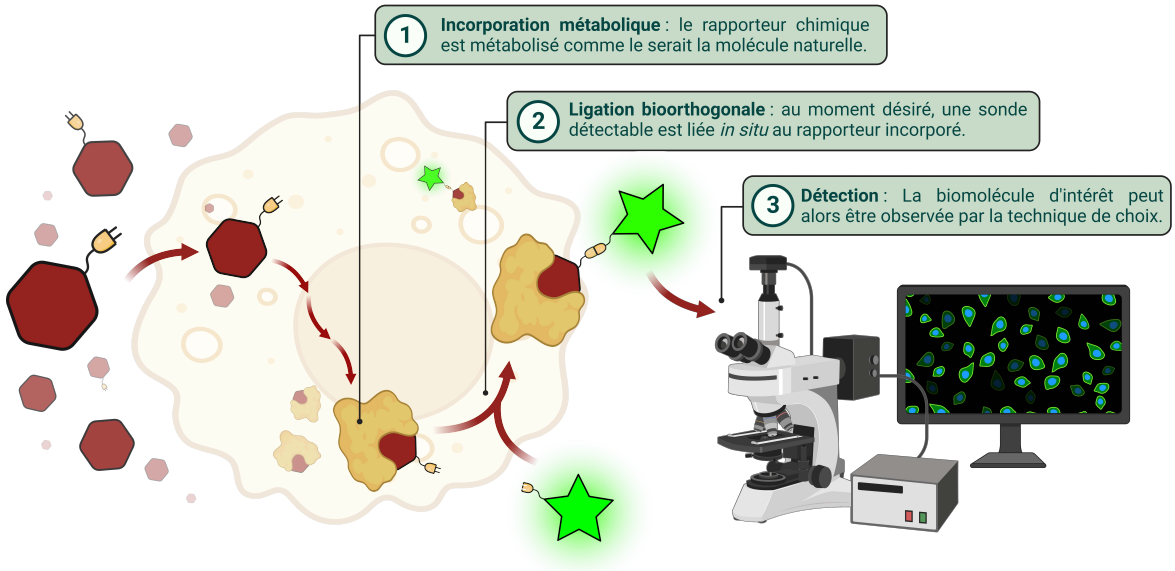


Figure 2 Stratégie du rapporteur chimique. Le rapporteur chimique (hexagone rouge) est incorporé au sein de cellules en culture ou d'un organisme. Au moment désiré, une sonde observable (étoile verte) est liée spécifiquement au rapporteur. L'utilisation de groupements à la fois discrets et absents des biomolécules naturelles, ainsi que la séparation nette entre l'étape d'incorporation et de ligation à une sonde permet une observation extrêmement fidèle de la biomolécule d'intérêt.

La réaction de ligation entre la sonde et le rapporteur doit être opérée de manière à minimiser le marquage aspécifique pour assurer un rapport signal sur bruit satisfaisant lors de la détection. Ainsi, la sonde utilisée doit réagir uniquement avec le rapporteur chimique et avec rien d'autre au sein de la cellule, du tissu ou de l'organisme vivant étudié. De plus, les propriétés physico-chimiques de la sonde telles que la solubilité ou l'hydrophobie doivent être optimales, de sorte à ce qu'elle ne précipite ou ne s'accumule pas de manière aspécifique dans certains compartiments de l'échantillon (*e.g.* membranes, partie hydrophobe des protéines, …), qu'elle puisse traverser librement les membranes, etc. Cet aspect a constitué, et constitue toujours aujourd'hui, un véritable challenge technologique : apparues au cours des vingt dernières années, les réactions dites bioorthogonales doivent répondre à une série de critères validant leur utilisation *in vivo*.

III. CHIMIE BIOORTHOGONALE

1. PRINCIPALES REACTIONS BIOORTHOGONALES

Le terme bioorthogonal, introduit par Carolyn R. Bertozzi, tire son origine du terme mathématique et désigne des réactions chimiques indépendantes d'un environnement biologique (i.e. qui n'interfèrent pas et ne sont pas influencées par ce dernier). Il englobe tous les prérequis précédemment cités que les réactions doivent respecter pour être utilisables lors de l'étape de ligation. Ainsi, la plupart des réactions polaires classiques en chimie organique (substitutions nucléophiles, additions, …) sont d'emblée éliminées, car les biomolécules constituant les systèmes vivants présentent une trop grande diversité de groupements sensibles à ces réactions (alcools, amines, thiols, acides carboxyliques, phosphates, …). De fait, les réactions de bioconjugaison classique comme par exemple les réactions hydrazide-carbonyles[7] ont d'abord été testées dans ce cadre, mais mènent à un niveau de background non-spécifique bien trop important. De plus, dans le cadre de la stratégie du rapporteur chimique, les groupements fonctionnels utilisés doivent également être les plus discrets possibles, pour ne pas perturber l'étape de métabolisation,[9] ce qui représente un défi de taille. La communauté de chimistes s'est donc mise en quête de la réaction bioorthogonale idéale. Les principales d'entre elles seront présentées brièvement ci-après.

1.1. LIGATION DE STAUDINGER-BERTOZZI

Historiquement, la première réaction bioorthogonale à avoir été utilisée *in vivo* pour la stratégie du rapporteur chimique est dérivée de la réaction de Staudinger.[10] Elle fait intervenir un azoture et une phosphine. L'azoture est petit, non chargé et globalement inerte vis-à vis des composants cellulaires, ce qui en fait un candidat de choix pour être utilisé comme rapporteur chimique. C'est Carolyn Bertozzi qui permit d'utiliser cette réaction au sein de systèmes biologiques en concevant une phosphine particulière munie d'un piège électrophile, un ester, permettant de lier la sonde au rapporteur azoture préalablement incorporé à la surface de cellules.[11] L'intermédiaire iminophosphorane

formé par la réaction de Staudinger réagit en effet avec l'ester en position ortho, intramoléculairement, produisant ainsi une liaison amide très stable plutôt que de libérer une amine comme c'est le cas avec une phosphine classique. Cette version de la réaction est communément appelée la ligation de Staudinger-Bertozzi (Figure 3).

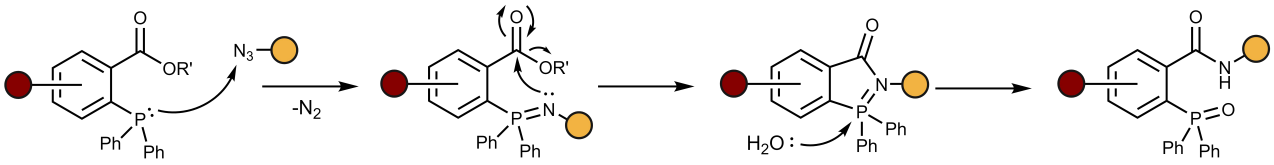


Figure 3 Ligation de Staudinger-Bertozzi. L'ester présent sur la phosphine agit comme un piège électrophile et permet de créer un lien amide stable plutôt que de libérer une amine comme dans la réaction de Staudinger originelle.

Cette réaction présente néanmoins des inconvénients majeurs : la phosphine n'est pas une fonction que l'on peut qualifier de biologiquement inerte car elle est sensible à l'oxydation, sa durée de vie est donc courte au sein de l'organisme. De plus, cette réaction présente une cinétique très lente, les temps de marquage sont longs, ce qui induit un bruit de fond important et la rend incompatible avec les applications *in vivo*. Après ce premier exemple prometteur (qui peut encore aujourd'hui être utilisé dans une stratégie de bioconjugaison classique, ou pour des marquages en milieu extracellulaire et à la surface des cellules),[12] la communauté scientifique s'est donc rapidement tournée vers un autre moyen d'utiliser les azotures comme rapporteurs chimiques : les cycloadditions 1,3-dipolaires.

1.2. CYCLOADDITIONS ALCYNES-AZOTURES

1.2.1. COPPER-CATALYZED AZIDE-ALKYNE CYCLOADDITION (CUAAC)

La première version de ces réactions, décrite par Huisgen en 1963, est la cycloaddition [3+2] thermique entre un alcyne et un azoture.[13] En faisant réagir un azoture comme dipôle avec un alcyne comme dipolarophile, elle permet l'obtention d'un cycle triazole stable, mais souffre d'une cinétique pauvre, et n'est utilisable qu'en solvants organiques. En 2001, K. Barry Sharpless définit le concept de *chimie click* comme un ensemble de critères auxquelles répondent des réactions chimiques dites idéales, permettant de

« cliquer » ensemble deux blocs moléculaires de manière efficace, propre, menant à un produit stable, dans des conditions faciles et sans libération de sous-produits gênants,[14] concept pour lequel il sera récompensé par le prix Nobel de Chimie en 2022[6] (Figure 4).

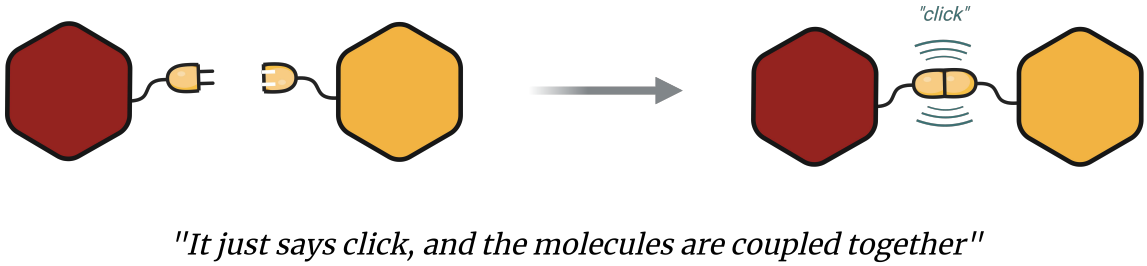


Figure 4 Illustration du concept de chimie click. Une réaction click permet de lier deux molécules ensemble de manière idéale : rapide, facile, totale, spontanée, spécifique, non-toxique.

Dans ce cadre, Sharpless adapta la cycloaddition décrite par Huisgen afin de l'effectuer dans l'eau et à température ambiante en utilisant un catalyseur Cu^{I}.[15] De manière concomitante, l'équipe de Morten Meldal publia aussi des travaux décrivant cette réaction en phase solide.[16] Les deux équipes découvrirent donc en même temps cette réaction, baptisée *Copper Catalyzed Azide-Alkyne Cycloaddition* (CuAAC), et qui devint rapidement la principale réaction de chimie click bien que limitée aux alcynes vrais. Devenue depuis une des réactions les plus célèbres tous domaines de la chimie confondus, la CuAAC fut rapidement transposée en milieu biologique. Outre une cinétique très favorable (de 10 à 200 $M^{-1}.s^{-1}$), l'avantage de cette réaction pour la stratégie du rapporteur chimique est que l'on peut aussi bien utiliser l'alcyne que l'azoture en tant que groupement rapporteur, les deux étant absents des organismes vivants, relativement inertes en milieu physiologique, peu encombrants et non chargés. On pourra ainsi envisager d'utiliser un rapporteur alcyne révélé par une sonde azoture, ou inversement (Figure 5).

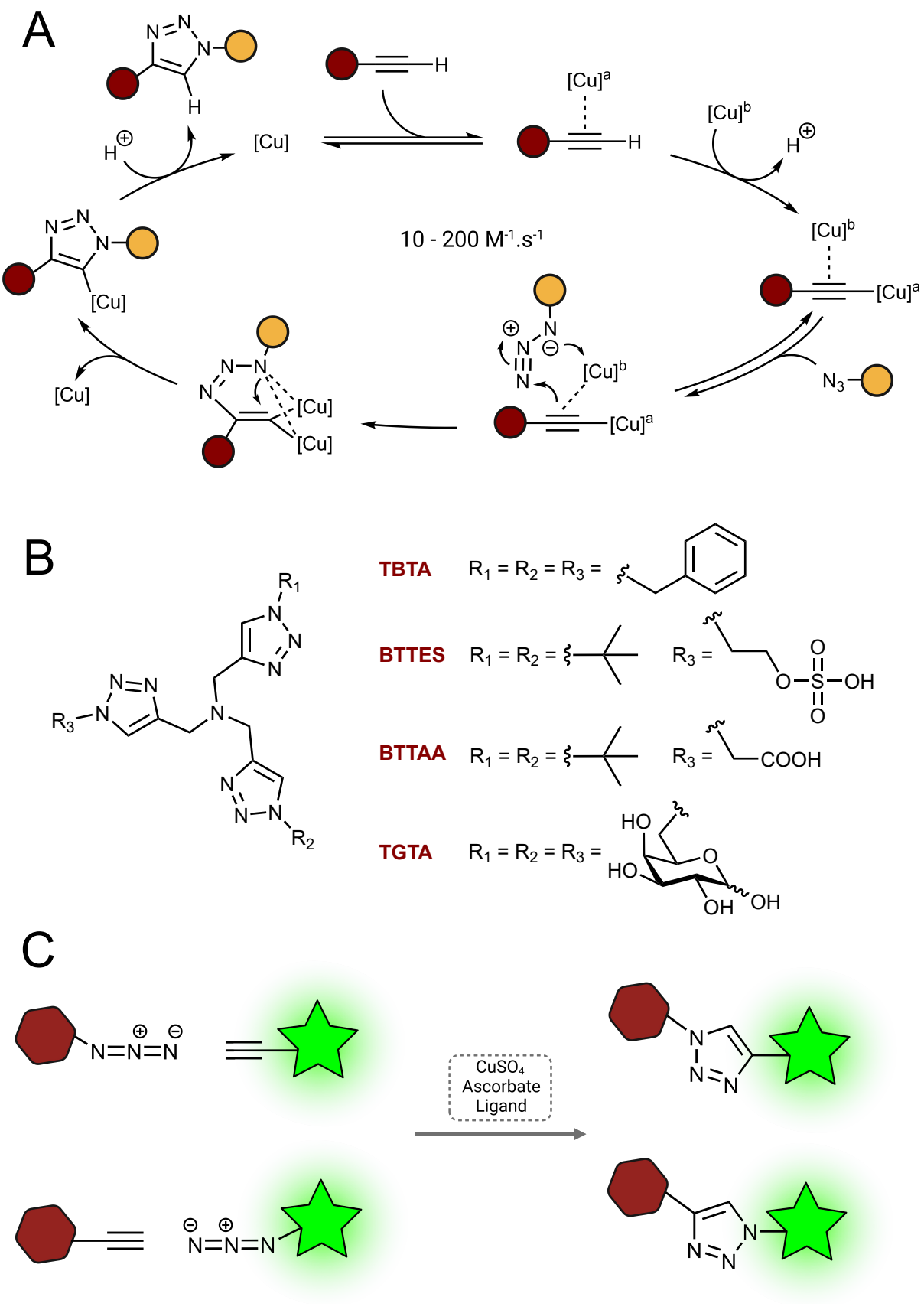


Figure 5 Copper Catalysed Azide Alkyne Cycloaddition (CuAAC). (A) Mécanisme réactionnel. (B) Ligands chélateurs couramment utilisés pour effectuer la réaction. (C) Les groupements azotures et alcynes, de par leur nature discrète et relativement inerte en milieu biologique, peuvent être utilisés aussi bien sur la sonde que sur le rapporteur, de façon interchangable.

Cependant, l'utilisation des ions Cu^{I} présente des problèmes de toxicité dans les systèmes biologiques. Les ions cuivreux génèrent en effet des espèces réactives de l'oxygène et de l'azote toxiques.[17] Ce problème est aujourd'hui minimisé par deux biais : on utilise pour la réaction une source de Cu^{II}, non toxique pour la cellule, ainsi qu'un réducteur biocompatible léger tel que les ions ascorbates, permettant de libérer *in situ* une faible quantité de Cu^{I}, assez pour catalyser la réaction tout en en limitant la concentration à l'instant *t*. De plus, on effectue la réaction en présence de ligands tétravalents, qui, en coordinant le Cu^{I}, limitent sa cytotoxicité et forment un catalyseur activé qui augmente la cinétique de la cycloaddition. Les premiers ligands utilisés, tels que le TBTA, souffraient d'une mauvaise solubilité en milieu aqueux, de nombreux développements ont donc été effectués pour les améliorer. C'est la mise au point du ligand BTTES qui permit en 2011 pour d'utiliser la CuAAC pour la première fois au sein d'embryons de *Danio rerio* (poisson zèbre),[18] ce qui était jusqu'alors impossible. En se basant sur sa structure, de nombreux ligands ont alors été développés, les plus utilisés aujourd'hui étant le BTTAA, qui permet d'obtenir des intensités de fluorescences trois à quatre fois plus importantes après marquage (comparé à un marquage au BTTES dans les mêmes conditions),[19] et le TGTA, dont la structure comporte 3 galactoses, assurant une très bonne solubilité aqueuse.[20]

Une autre limite de la CuAAC est le grand nombre de réactifs requis (sonde moléculaire, source de Cu^{II}, réducteur, ligand …), rendant son implémentation parfois difficile suivant le milieu biologique ou le tissu visé. Les chercheurs ont cherché de nombreux moyens de pallier ce problème, par exemple en développant des réactifs « tout-en-un » tels qu'un ligand fluorescent chélatant le Cu^{I},[21] ou en utilisant un peptide riche en Arginine qui chélate le Cu^{I} et pénètre les membranes passivement, améliorant l'incorporation du cuivre tout en maintenant une faible cytotoxicité.[22]

1.2.2. STRAIN-PROMOTED ALKYNE AZIDE CYCLOADDITION (SPAAC)

Cherchant à trouver une alternative à la CuAAC, l'équipe de Carolyn Bertozzi permit une autre grande avancée parallèle dans ce domaine : l'utilisation d'un alcyne cyclique contraint tel qu'un cyclooctyne permet de s'affranchir de l'utilisation d'un catalyseur de type Cu^{I} pour effectuer la réaction, tout en conservant sa compatibilité avec les conditions physiologiques et une cinétique satisfaisante.[23] La tension de cycle induite sur l'alcyne, normalement de géométrie linéaire, le déstabilise et augmente sa réactivité. L'énergie d'activation de la réaction est donc fortement abaissée, comparativement aux cycloadditions [3+2] avec un alcyne terminal, et le mécanisme péricyclique thermique procède spontanément à température ambiante sans exiger l'utilisation d'un catalyseur exogène, facilitant grandement l'implémentation de la réaction. C'est la *Strain Promoted Alkyne Azide Cycloaddition* (SPAAC). Cette réaction mena à la première application *in vivo* de la chimie click en 2008, en permettant de marquer des embryons de poisson zèbre.[24]

Différentes générations de cyclooctynes ont été développées depuis leur apparition en 2004 dans le but d'améliorer leur solubilité ainsi que la cinétique de la réaction, en introduisant des atomes de fluor en position propargylique (difluorocyclooctynes, DIFO) puis en fusionnant le cyclooctyne avec des cycles aromatiques (dibenzocyclooctynes, DBCO, également appelés azadibenzocyclooctyne, ADIBO, ou dibenzoazacylooctyne, DIBAC) ou en introduisant des fonctions augmentant la tension de cycle (biarylazacyclooctynones, BARAC). Les dérivés BARAC sont douze fois plus réactifs que les dérivés DIFO et ne présentent aucune cytotoxicité,[25] mais sont plus difficiles à synthétiser (de manière générale, les alcynes cycliques sont plus fastidieux à obtenir que les alcynes terminaux). Certaines sondes pénètrent les membranes cellulaires tandis que d'autres, notamment les DIFO, sont trop polaires pour les traverser et peuvent donc être très facilement retirées du milieu par rinçage, après réaction. Ceci permit d'effectuer des multi-marquages résolus en temps chez *Caenorhabditis elegans*, en incubant avec une première sonde DIFO, suivi d'une étape de rinçage et d'une seconde incubation avec une sonde DIFO émettant à une autre longueur d'onde.[26]

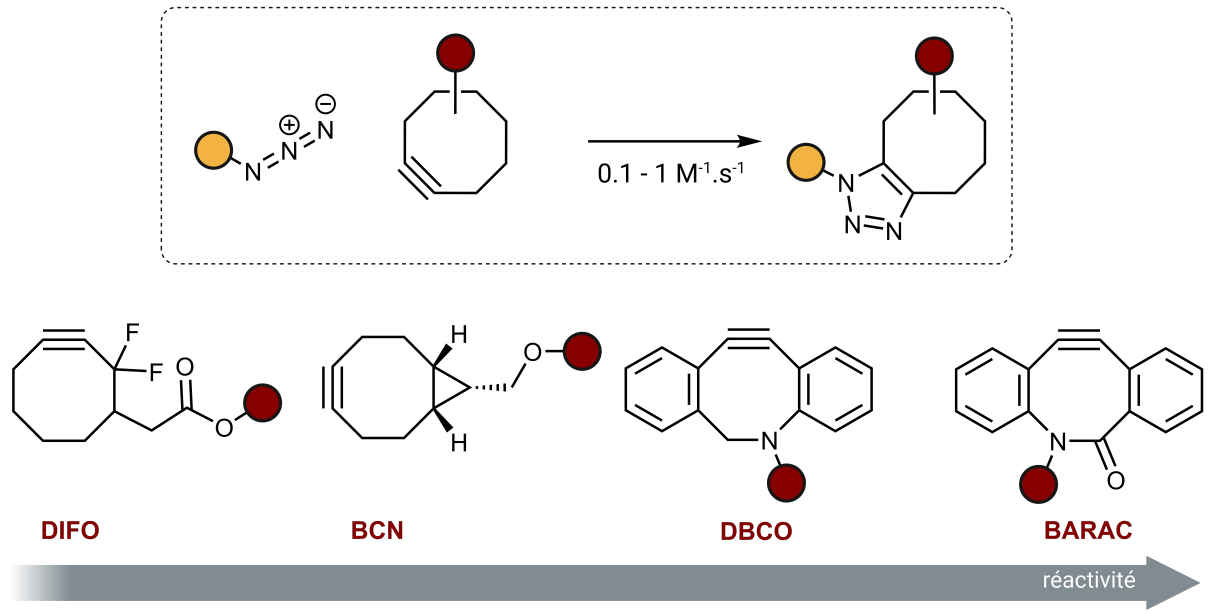


Figure 6 Strain-Promoted Alkyne Azide Cycloaddition (SPAAC). Les alcynes cycliques contraints de type cyclooctyne réagissent spontanément avec les azotures pour former un cycle triazole stable. Différentes générations de cyclooctynes ont été développées pour améliorer la vitesse de la réaction. Pour utiliser cette approche, le rapporteur chimique devra nécessairement porter le groupement azoture car les cyclooctynes sont trop encombrants et perturberaient l'étape d'incorporation métabolique.

Les alcynes cycliques – devant posséder au moins 8 chaînons, les cycles plus petits étant trop instables – présentent cependant un encombrement stérique important qui limite leur usage en tant que rapporteur chimique. Un marquage aspécifique plus important est également observé avec la SPAAC.[27] Ceci est principalement dû au caractère lipophile des cyclooctynes, qui facilite la mise en place d'interactions faibles hydrophobes avec les membranes cellulaires, ainsi qu'à l'existence d'une réaction secondaire possible avec les thiols libres des cystéines et la triple liaison C-C (addition thiol-yne), laquelle peut être évitée en pré-incubant les cellules avec des réactifs bloquant les thiols, tels que le iodoacétamide ou les maléimides.[28] La réactivité des cyclooctynes avec les thiols a d'ailleurs été mise à profit en développant une étiquette heptapeptide contenant des cystéines, permettant une fonctionnalisation sélective d'une protéine d'intérêt par SPAAC avec un fluorophore DBCO.[29] Les réactions de CuAAC et de SPAAC présentant chacune des avantages et inconvénients (facilité de mise en œuvre et de la SPAAC, meilleure spécificité de la CuAAC, …), elles doivent être perçues comme réactions complémentaires, pouvant être utilisées conjointement.

En 2008, de manière fortuite une fois encore, deux équipes de recherches ont simultanément fait la découverte d'une autre cycloaddition pour le marquage bioorthogonal, la réaction de [4+2] Diels-Alder entre une tetrazine et un alcène cyclique de type norbornène[30] ou transcyclooctène (TCO).[31] La réaction de Diels-Alder telle que couramment pratiquée implique l'utilisation de dienes appauvris électroniquement, qui sont souvent de bons accepteurs de Michael et risqueraient de réagir avec les nucléophiles naturellement présents dans la cellule (alcools, thiols, amines …). Pour pallier ce problème, les effets électroniques sur le diène et diénophile ont été inversés, on parle alors de réaction de Diels-Alder à demande électronique inverse (IEDDA) qui implique des dienes appauvris et de diénophiles enrichis électroniquement. La réaction effectuée avec les tetrazines présente l'avantage d'être irréversible contrairement aux réactions de Diels-Alder classiques. En effet la réaction s'effectue en deux étapes, une première réaction de Diels-Alder qui conduit à la formation d'un composés bicyclique contraint, lequel subit alors une réaction de rétro Diels-Alder irréversible qui forme un cycle 4,5-dihydropyridazine en libérant du diazote. La réaction est également possible avec des diénophiles de type alcyne, comme les bicyclononyne (BCN) également employés pour la SPAAC, qui forment directement le cycle pyridazine, avec une cinétique proche de la réaction entre une tétrazine et un TCO.[32]

La IEDDA possède une cinétique inégalée par les autres réactions bioorthogonales, elle est près de 10 000 fois plus rapide que la CuAAC, ce qui lui permet de réagir de manière efficace même à des concentrations très faibles, de l'ordre des concentrations auxquelles sont retrouvées les protéines en milieu physiologique.[33] De nombreuses optimisations ont été effectuées pour améliorer sans cesse la réaction, notamment en sélectionnant des partenaires réactionnels plus efficaces. La contrainte de cycle du diénophile joue notamment un rôle important. Les TCO souffrent par exemple d'une isomérisation possible en *cis*-cyclooctène, bien moins réactifs que l'isomère *trans* ; de nouveaux alcènes contraints ont de ce fait été étudiés, en fusionnant par exemple un cyclopropane ou *cis*-

dioxolane sur le TCO, menant à l'obtention des composés s-TCO et d-TCO respectivement, environ 160 fois plus réactifs que les TCO classique, mais souffrant d'une faible stabilité.[34,35] Les cyclopropènes ont également été identifiés comme de très bons réactifs pour effectuer la IEDDA.[36] Ceux-ci, bien que moins réactifs que les TCO, possèdent l'avantage d'être plus stable, ce qui peut être mis à profit selon l'application souhaitée.

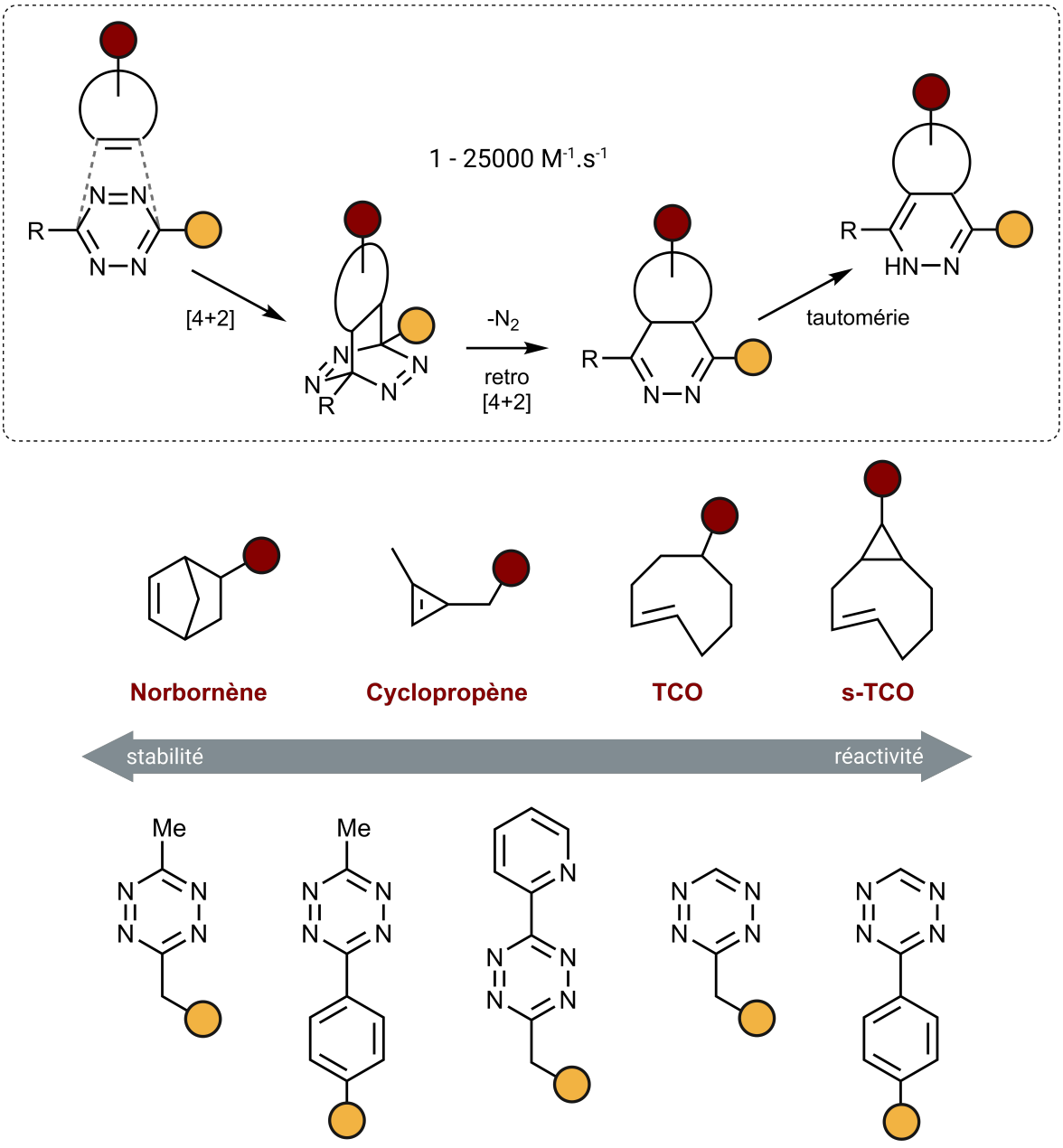


Figure 7 Inverse Electronic Demand Diels Alder (IEDDA). Cette réaction entre une tétrazine et un alcène permet de créer un cycle stable. Il s'agit de la réaction bioorthogonale la plus rapide à ce jour, mais les groupements impliqués souffrent d'un manque de stabilité et sont relativement encombrants.

IV. MARQUAGE INTRACELLULAIRE VS EXTRACELLULAIRE

Comme vu précédemment, la biosynthèse des acides sialiques et l'assemblage des glycoconjugués a majoritairement lieu à l'intérieur de la cellule. Ainsi, en parallèle des objectifs expérimentaux de ma thèse, dans un contexte sanitaire empêchant toute manipulation au laboratoire pendant parfois de longues périodes, nous nous sommes intéressés au cas particulier du marquage métabolique en milieu intracellulaire, conduisant à la publication d'une revue de littérature dans le prestigieux journal *Angewandte Chemie International Edition* en 2021.[37]

En effet, les réactions bioorthogonales doivent par essence respecter de nombreux critères, et le ciblage du milieu intracellulaire vient encore allonger la liste de ces prérequis. Ce milieu est densément compartimenté et peu accessible, plusieurs facteurs limitants tels que la perméabilité membranaire des réactifs ou la capacité à se débarrasser des signaux résiduels doivent donc être examinés encore plus scrupuleusement que lors de marquages extracellulaires où la cible est plus accessible et les rinçages plus aisés. Dans cette revue, nous présentons et examinons les différentes réactions bioorthogonales, puis insistons sur les prérequis particuliers pour le marquage intracellulaire des principaux types de biomolécules (glucides, protéines, lipides, acides nucléiques et autres). Celle-ci, et notamment la partie traitant du marquage des glycanes, introduit parfaitement le contexte dans lequel prend part mon projet de thèse ; j'ai donc choisi de l'insérer ci-après en tant que partie intégrante de l'introduction de mon manuscrit.

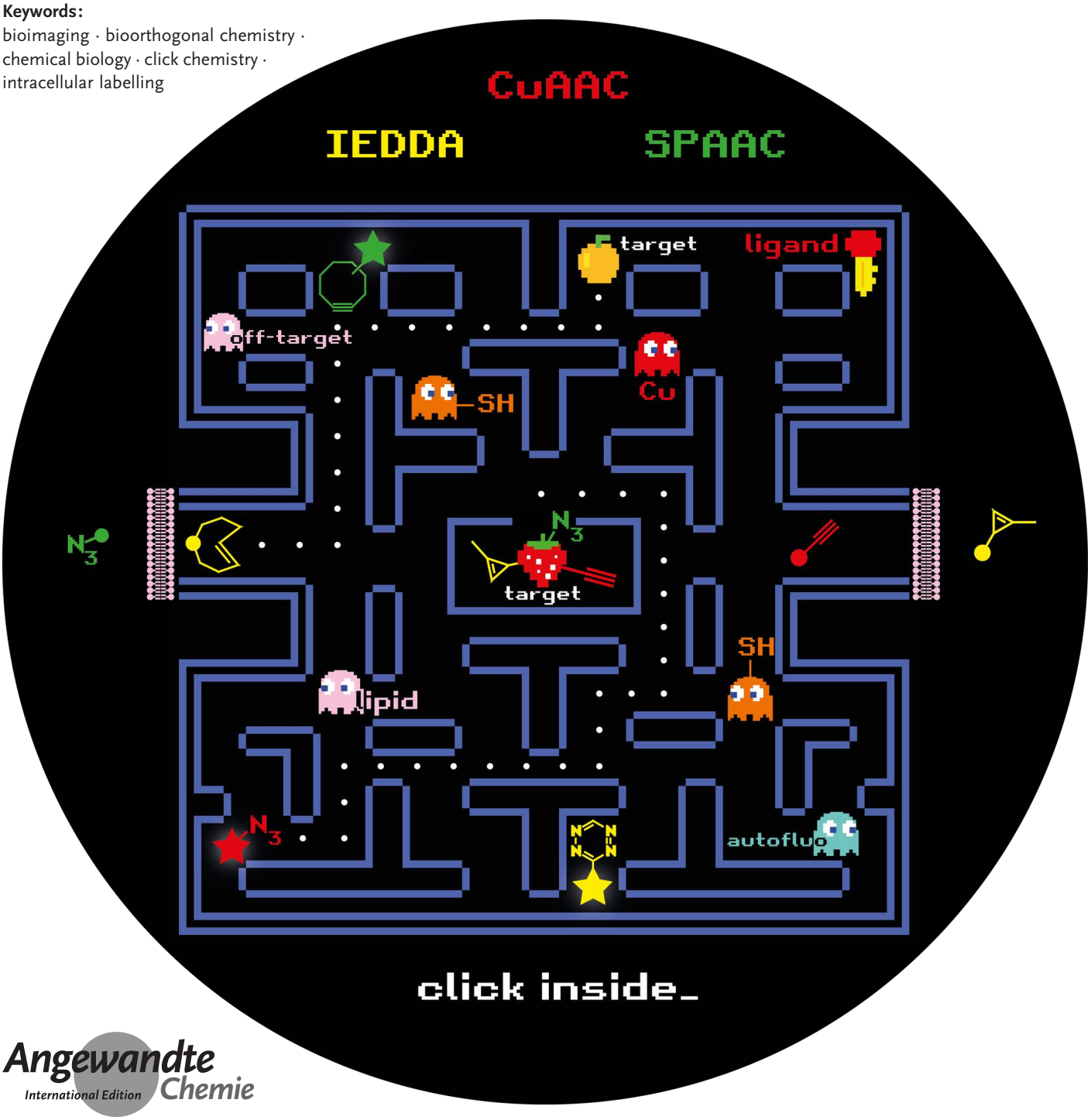
How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 23084–23105
International Edition: doi.org/10.1002/anie.202101502
German Edition: doi.org/10.1002/ange.202101502

Bioimaging

To View Your Biomolecule, Click inside the Cell

Vincent Rigolot, Christophe Biot, and Cedric Lion**

Keywords:
bioimaging · bioorthogonal chemistry · chemical biology · click chemistry · intracellular labelling



The surging development of bioorthogonal chemistry has profoundly transformed chemical biology over the last two decades. Involving chemical 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, and post-synthetic labelling of nucleic acids. While a majority of literature reports mark cell-surface exposed targets, implementing bioorthogonal ligations in the interior of cells constitutes a more challenging task. Owing to limiting factors such as membrane permeability of reagents, fluorescence background due to hydrophobic interactions and off-target covalent binding, and suboptimal balance between reactivity and stability of the designed molecular reporters and probes, these strategies need mindful planning to achieve success. In this review, we discuss the hurdles encountered when targeting biomolecules localized in cell organelles and give an easily accessible summary of the strategies at hand for imaging intracellular targets.

From the Contents

1. Introduction

For decades, first-generation bioconjugation reactions such as those involving amine-reactive *N*-hydroxysuccinimidyl esters or isothiocyanates, thiol-reactive maleimides or pyridyldisulfides, and carbonyl-reactive hydrazides or oximes have been providing the scientific community with a wealth of valuable molecular tools for the visualization, detection, vectorization, and immobilization of biomolecules of interest (BOIs).[1,2] However, for applications that require monitoring of a detectable bioconjugate in cell/tissue cultures or in whole organisms, these reactions have 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 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 behavior. Owing to this pitfall, which is also found with genetic approaches relying on fluorescent proteins and other peptide tags, various dynamic phenomena of interest, including post-translational modifications of proteins and trafficking of small metabolites, could not be addressed. 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 review, 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.

Over the past 20 years, chemical biologists have strived to develop 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 with low concentrations of reactants at physiological pH, pressure, and temperature to obtain a satisfactory signal-to-noise ratio and ensure the specificity of the detected signal. Naturally, if the ligation is to take place on or in a living sample (as opposed to a fixed biological 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 displayed real bioorthogonal character

[*] V. Rigolot, Prof. C. Biot, Dr. C. Lion
UMR 8576 CNRS, Unité de Glycobiologie Structurale et Fonctionnelle, Université de Lille, Faculté des Sciences et Technologies
Bât. C9, 59655 Villeneuve d'Ascq (France)
E-mail: christophe.biot@univ-lille.fr
cedric.lion@univ-lille.fr
The ORCID identification number(s) for the author(s) of this article can be found under:
https://doi.org/10.1002/anie.202101502.

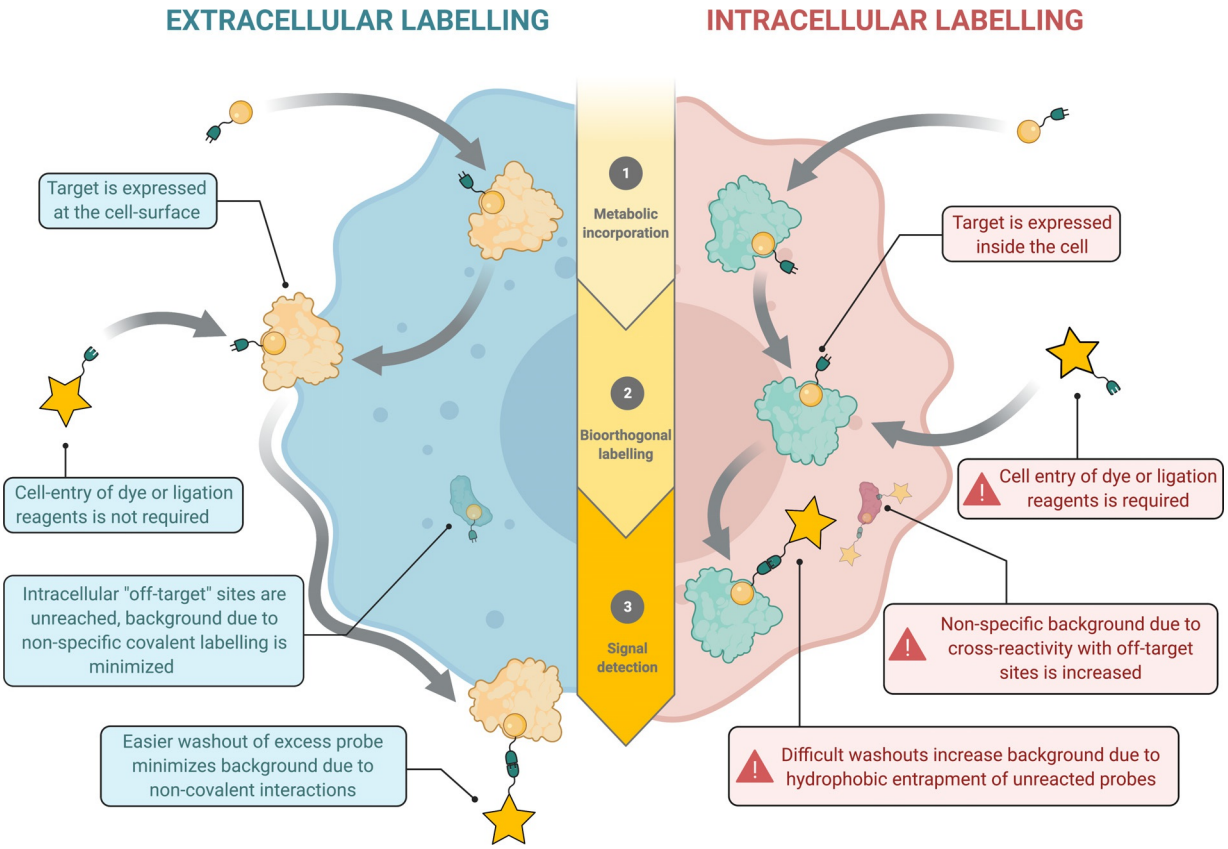


Figure 1. Inside or outside. Key differences between strategies for biorthogonal labelling with chemical reporters within the cell or at the cell surface.

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 (CheGB) group at the Institute of Structural and Functional Glycobiology (UMR8576–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.

Christophe Biot obtained his Ph.D. in synthetic chemistry (1995–1998) under the supervision of Prof. J. Brocard at the University of Lille and in collaboration with Dr. D. Dive at the Institut Pasteur de Lille. He conducted postdoctoral research in the team of Dr. E. Davioud-Charvet at the Institute of Biology of Lille (2000–2001) and with Pr. M. Rooman at the Ecole Polytechnique in Bruxelles (Belgium) (2001–2004). He was then recruited by the University of Lille in 2004 as assistant professor. He was appointed Professor in 2012. His research topics cover several aspects of high significance in chemical biology.

Cedric Lion obtained his Ph.D. in pharmaceutical 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 stays 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 Chemical Glycobiology team in the UGSF to develop bioorthogonal chemistry for bioimaging. His research interests include the development of multiple click methods to decipher glycosylation and more sialylation in human and bacterial models, as well as cell wall biopolymers in plants, in particular lignins.

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 sensitivity of phosphines 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 rapidly shifted to pericyclic cycloadditions, in light of Sharpless' conceptual definition of ideal "click chemistry"[6] and the ground-breaking discovery of the copper-catalysed azide–alkyne cycloaddition (CuAAC).[7,8]

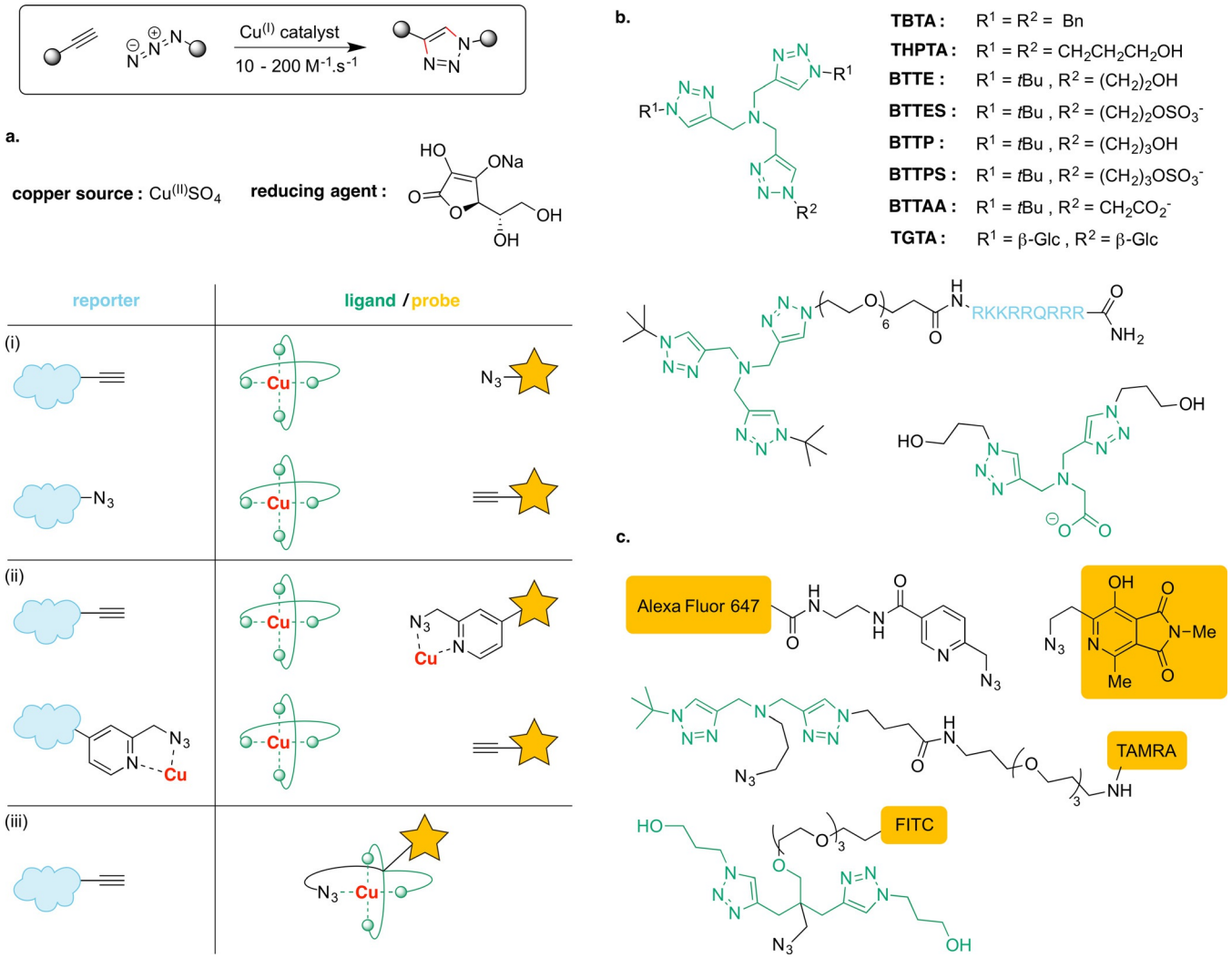
2. The Major Triad of Bioorthogonal Ligations

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, which boosted reaction rates by up to a factor of 10^7 in all solvents including water, with the added benefit of 1,4-triazole regiospecificity (Scheme 1). Although 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 the 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 as either 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 CuAAC 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, washout 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 before the preformed complex and probe are added 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 capability 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 advisable to store and reuse solutions containing both copper and ascorbate. These considerations are important since they may lead to unsuccessful experiments and lower reproducibility of results due to human error. Although this may seem trivial to the specialist, the ultimate aim of such reactions is the 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. Even though CuAAC commercial kits are now widely available for the most frequent reporters, which 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", and "in vitro" can be somewhat blurred and subjective when it comes to metabolic labelling with the two-



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 picolyl azides used in conjunction with ligands improve kinetics and labelling performance when compared to (i). For most applications, a picolyl azide 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, who introduced a picolyl azide 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.

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 ligation of the probe onto incorporated reporters. The latter constitutes a separate, second experimental step which can be applied to a variety of biological sample types such as living organisms, tissues, and cells, fixed and/or permeated samples, biopsies, lysate extracts, SDS-PAGE gels, etc. depending on the application and the desired 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.

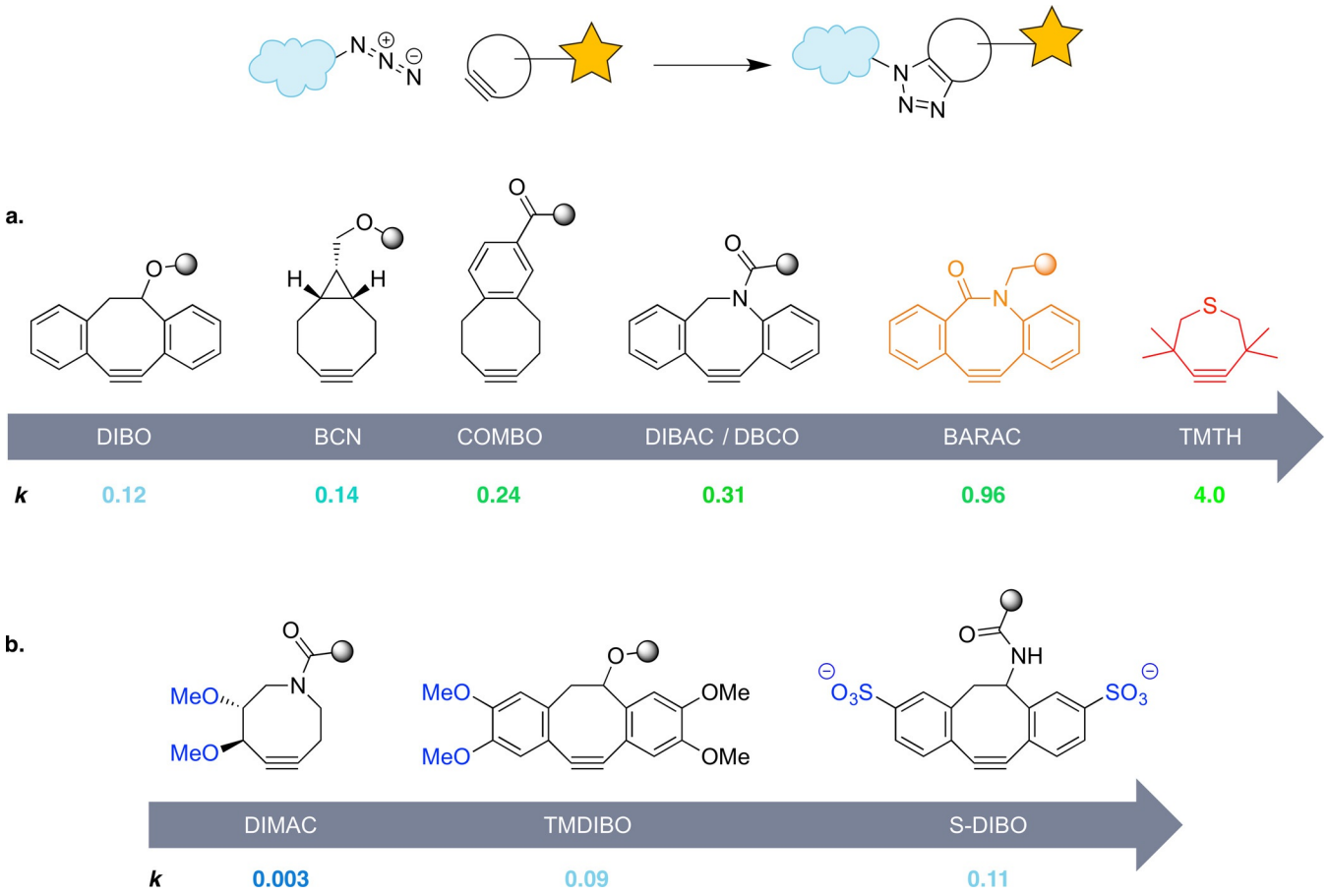
For applications in which the ligation step must be done in living samples, the biocompatibility of CuAAC has been greatly improved in the last few years, driven by advances in copper-chelating tools. Stable water-soluble Cu^{I} complexes have been shown to cross the cell membrane and promote intracellular cycloaddition in the absence of ascorbate,[20] and

the design of multidentate ligands continues with the recent description of an easily accessible ligand imparting CuAAC with improved kinetics and lower cytotoxicity compared to that of BTTPS[21] (Scheme 1b). More elaborate 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 in two articles on 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]

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

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 a forgotten cycloaddition variant involving cyclooctyne as the dipolarophile.[31] The distortion of its dihedral angle (163° in cyclooctyne instead of 180° in acyclic counterparts) confers important ring strain to the alkyne (ca. 18 kcal mol^{-1}). As a result, the purely thermal $\pi4s + \pi2s$ 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 $HOMO_{AZIDE}$–$LUMO_{ALKYNE}$ gap dominating the SPAAC reaction and to decrease the 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 many 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 react more slowly with alkyl azides. Interestingly though, this kinetic trend is reversed when cycloalkynes are reacted 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 nonspecific background signal due to hydrophobic trapping was also partially addressed with the development of hydrophilic cyclooctynes such as S-DIBO, TMDIBO, and DIMAC derivatives, which display better bioavailability and may prove more relevant than DIBAC/DBCO in some contexts despite slower kinetics[35–37] (Scheme 2b).

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 lower specificity, as cyclooctynes may react with endogenous nucleophiles (cysteine residues and other thiols) thus generating much higher non-specific background signal 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

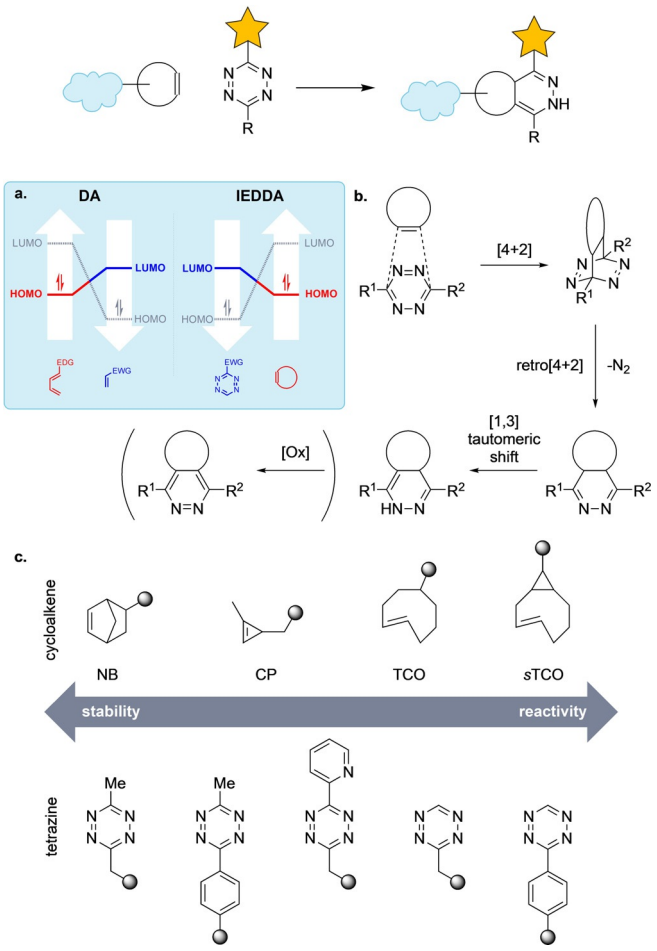


Scheme 2. Strain-promoted alkyne–azide cycloaddition (SPAAC). a) Kinetics of the main reagent families. DIBO: dibenzocyclooctyne; BCN: bicyclononyne; COMBO: carboxymethylmonobenzocyclooctyne; DIBAC: dibenzoazacyclooctyne (also known as DBCO: dibenzocyclooctyne); BARAC: biarylazacyclooctynone; TMTH: tetramethylthioheptyne. 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 $M^{-1}s^{-1}$.

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 frequently used reactants for thiol bioconjugation). In contrast, 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 efficiently create stable (dihydro)pyridazine linkages with the BOI without the need for metal catalysis (Scheme 3c).



Scheme 3. Inverse electronic-demand Diels–Alder reaction (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 strain, 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.

IEDDA kinetics are unequalled when TCO reporters are used—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, and steric effects, but also electronic effects on the tetrazine diene.[33] Intriguingly, IEDDA reactions are dramatically faster in aqueous solutions because of the strong effect of solvent on hydrogen bonding and hydrophobic interactions. Rate constants up to 10000-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 the 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) and 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 with still respectable 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. CP-functionnalized reporters 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 unsubstituted

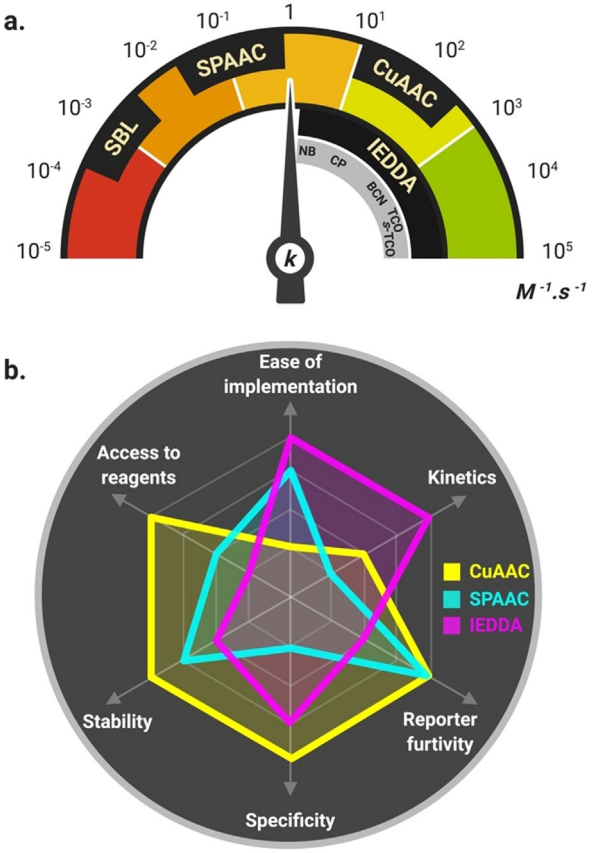


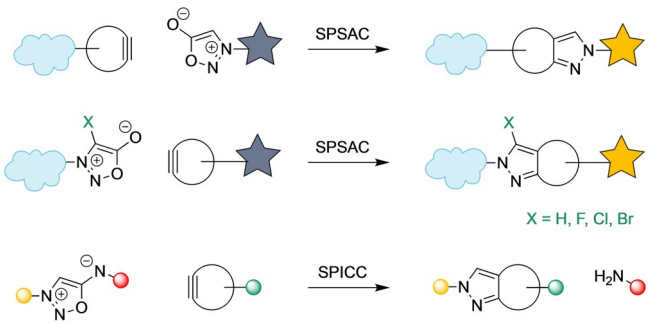
Figure 2. Differences and complementarity. a) Comparison of the estimated kinetic range of SBL, CuAAC, SPAAC and IEDDA ligations. b) Qualitative evaluation of important criteria influencing method design.

tetrazines, at the expense of slower kinetics, see Scheme 3c), and using the fastest reacting tetrazine derivative might not always be optimal.[50] Importantly, some tetrazine reactants also come with the added benefit of the "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, and coumarin derivatives (see Section 6). 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 washout steps. While a small number of profluorophores are available for CuAAC and SPAAC ligations, the unique photophysical properties of tetrazines gives 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 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 regarding the relevance of their biological behaviour. In addition to their increased steric hindrance over that of the azide/alkyne pair of CuAAC, strained alkenes and tetrazines can also be less stable. For a given case, the practical choice of reacting IEDDA partners might thus not prove to be faster than the CuAAC alternatives. 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 the experimental context (Figure 2).

2.4. Other Cycloadditions

Among a multitude of candidate reactions, other cycloadditions have also emerged that show promise for bioorthogonality, including the photoinduced alkene–tetrazole conjugation[51–53] and 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



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.

(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.

2.5. Cellular Thiols Standing in the Way

Biological thiols are the main villain in 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 non-specific labelling and high background levels for SPAAC.[39] Combined with its slow kinetics, this significantly lowers the sensitivity of the SPAAC reaction, as the labelled BOI must 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, the choice of a type of reporter and a bioorthogonal reaction to link it to a probe is very dependent on 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, and peptide tags.[66] Despite their undeniable usefulness, the requirement for genetic fusion of an appended sequence may perturb the 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. The expansion of the genetic code to designer building blocks exploiting 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.

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 state-of-the-art intracellular bioimaging for sensitivity reasons (high-energy wavelengths in regions of the spectrum with marked autofluorescence in most samples) and for specificity reasons (high nonspecific 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 UAAs 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. In contrast, 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 (Scheme 1b).[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 20-fold 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

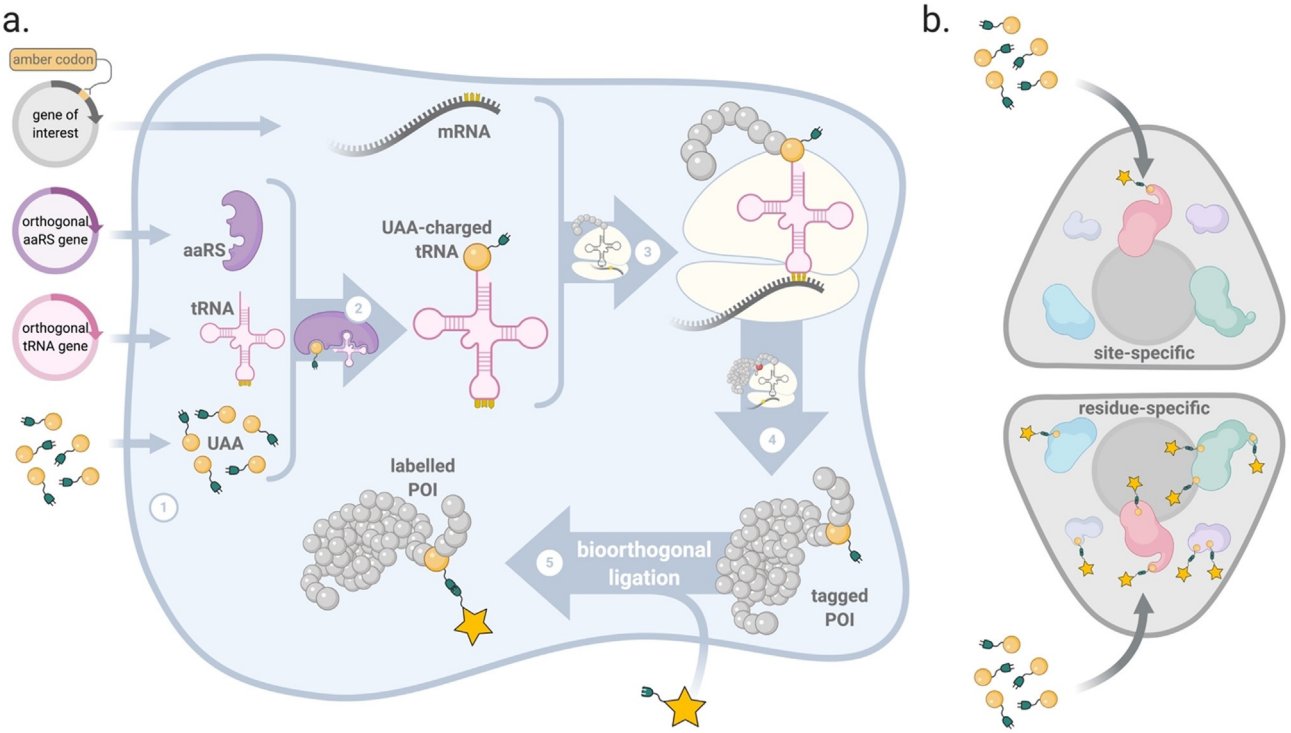
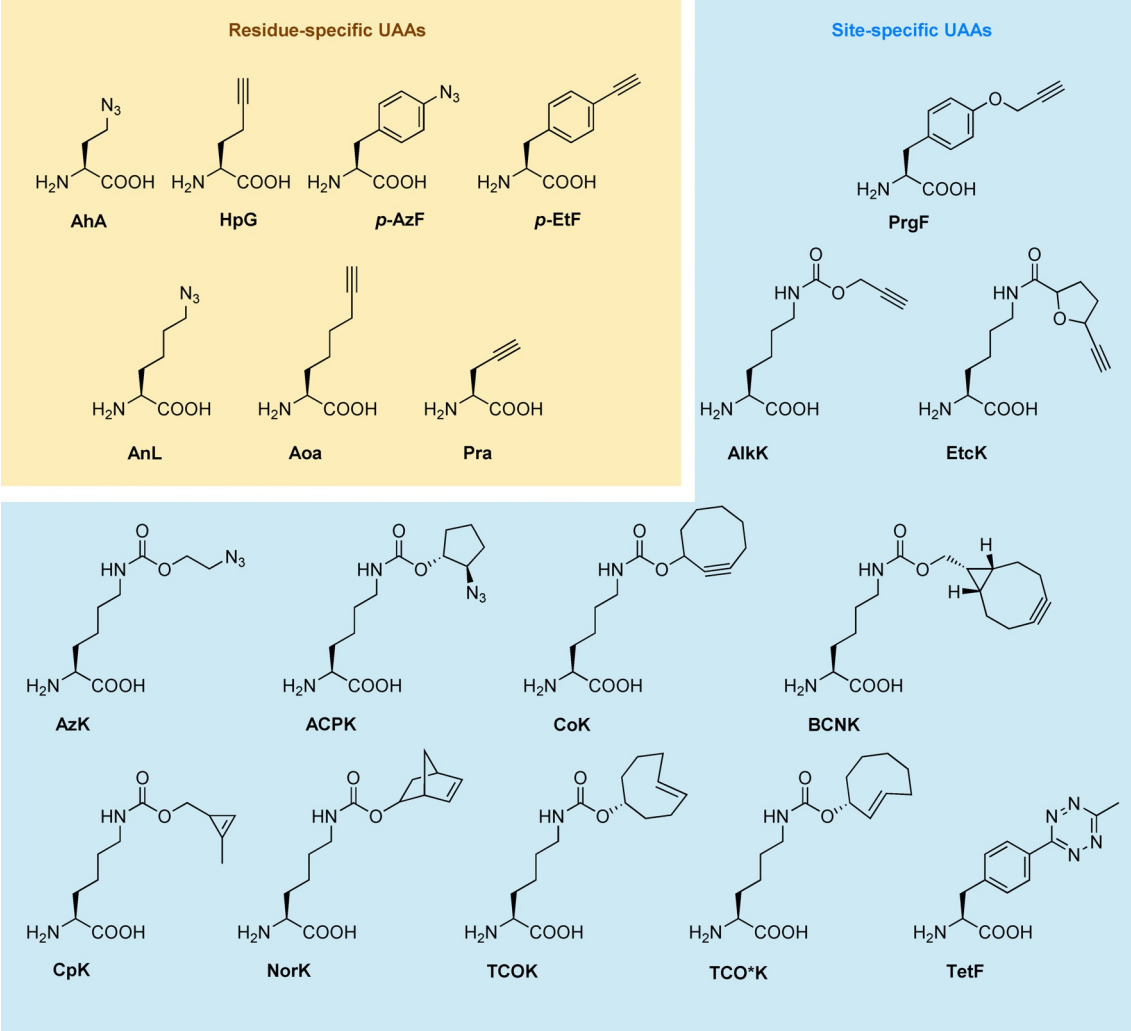


Figure 3. Metabolic labelling of proteins. a) Site-specific incorporation of UAAs in a protein of interest (POI). The orthogonal aaRS/tRNA pair is introduced by transfection as well as the gene coding for the POI, modified with a reassigned amber codon (1). Once inside the cell, UAAs must be recognized 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.

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, and 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 with 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 with 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 the 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 a shielding effect by the carbamate.[45] By inverting probe and reporter handle, stable tetrazine-UAAs suitable for use with TyrRS/tRNA pairs, such as TetK, 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



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), and 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).

both IEDDA and 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 cell-impermeable 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 over their faster reacting unsubstituted counterparts for some imaging studies in intracellular compartments, as their improved stability may largely compensate for the kinetic difference 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]

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 systems based on self-labelling enzymes, such as SNAP, CLIP, and Halo tags. In 2015 Murrey et al. published an impressive article combining the use of genetically encoded Halo tags and strain-promoted bioorthogonal chemistry; they compared the reactive moieties of strain-promoted cycloadditions in various organelles of living cells[85] and proposed silver-stabilized *s*-TCO reagents with an improved stability/reactivity balance. The SNAP/CLIP-tag technology was also used in conjunction with IEDDA and SPAAC for dual "orthogonal/biorthogonal" 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, and 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

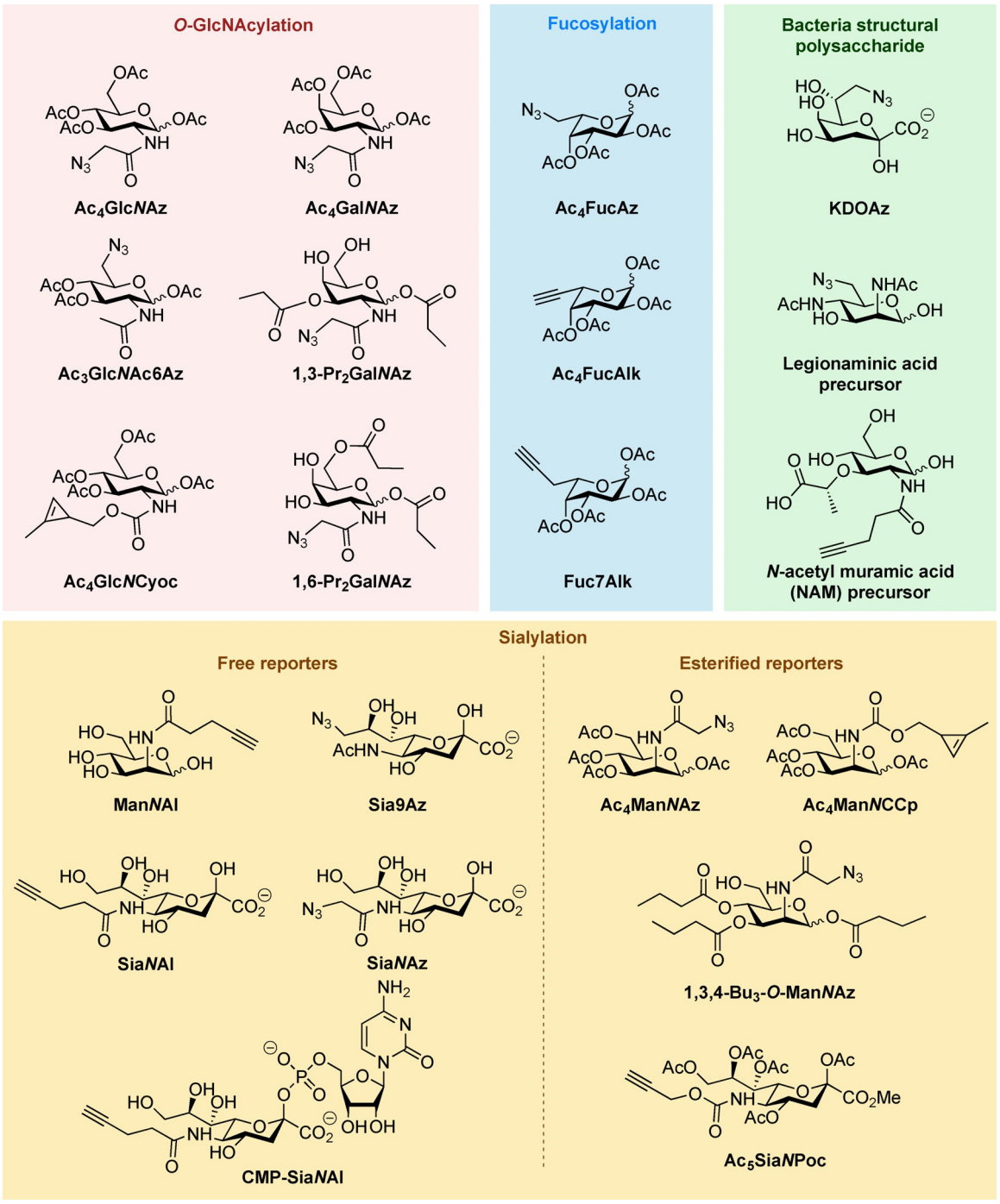
4.1. A Plurality of Glycosylations

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, and 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 with 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] and plant biological models,[13,94,95] in various experiments ranging from cell cultures to ex vivo tissues to whole organisms (Scheme 6). This includes detection of essential glycosylation types such as sialylation, fucosylation, and *O*-linked β-*N*-acetylglucosaminylation (*O*-Glc*N*Acylation), as well as structural polysaccharides in plants (e.g., pectins) and bacteria (e.g., lipopolysaccharides and 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, and specificity of reagents. Here, we focus on the hurdles to overcome when aiming at detecting glycans in the interior of living or fixed cells.

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



Scheme 6. Selection of monosaccharide reporters that target glycosylation.

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.

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 CuAAC is used, 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, that is, the covalent attachment of a *N*-acetylglucosamine (Glc*N*Ac) to 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*-Glc*N*Ac 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 at 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

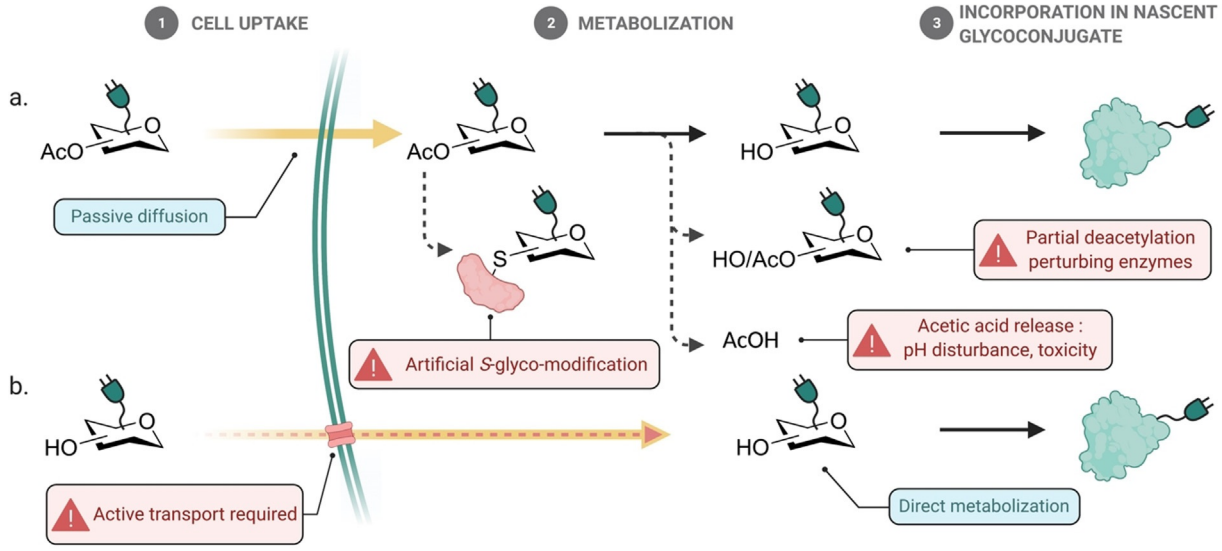


Figure 4. Per-*O*-acetylated monosaccharides versus free monosaccharides. The uptake of an apolar per-*O*-acetylated reporter is facilitated by passive diffusion, thus allowing lower concentrations to be used for metabolic incorporation, but this can lead to increased background because of off-target 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.

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.[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 nonspecificity 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 three- to fivefold 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 Man*N*Ac membrane transporter, confirmed cellular entry of exogenous Sia*N*Al by endocytosis and sialin-mediated export to the cytosol, and 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 for organelle-directed MGL 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 ligand-targeted liposomes are used,[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

Not every organism possesses an esterase activity sufficient to release free reporters from per-*O*-acetylated precursors 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 by washing to avoid fluorescence background due to noncovalent 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 in this review 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 modified oligonucleotide 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 (or pronucleotides) reporters as monomeric building blocks for metabolic labelling, similar to technologies used for proteins with amino acids or glycans with monosaccharides (Figure 5). This strategy avoids the complex synthesis of a 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]

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 deoxythymidine analogue EdU, bearing an ethynyl group at position 5 (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 instance, azide derivative AmdU can be employed with CuAAC as well as with SPAAC, thus allowing visualization in samples without fixation, and vinyl derivatives have also been reported (e.g., VdU). Regarding purines, an alkynyl deoxy-

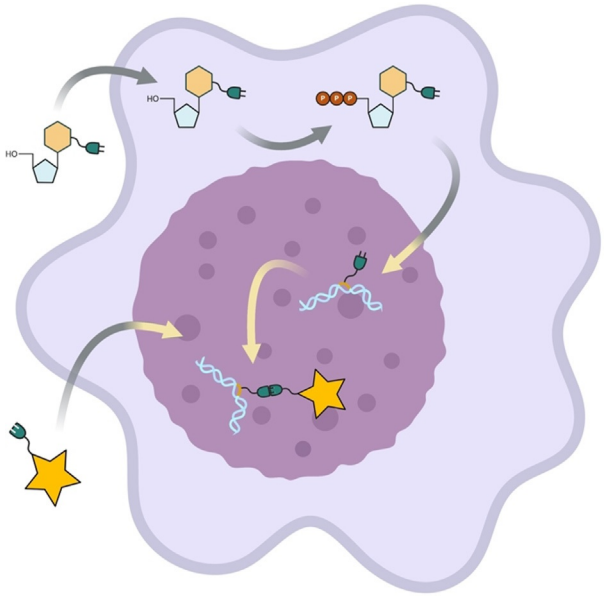
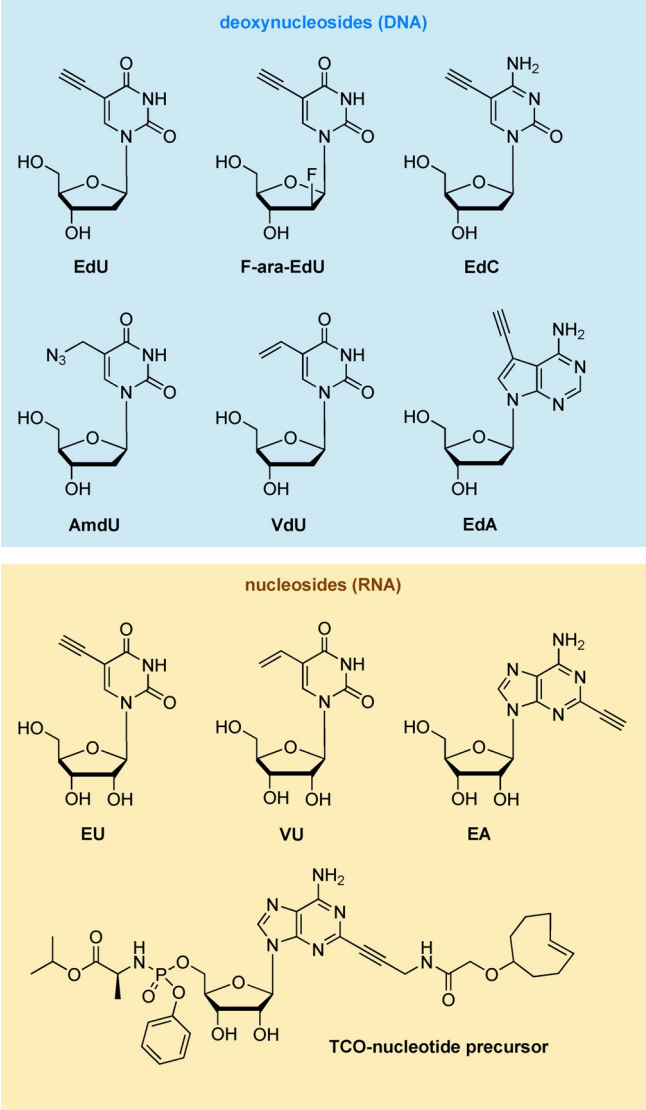


Figure 5. DNA and RNA metabolic labelling. A cell-permeable nucleoside or pronucleotide equipped with a bioorthogonal reactive handle is activated as a nucleotide triphosphate by the endogenous cellular machinery, then incorporated into nascent DNA or RNA strands.



Scheme 7. Examples of reporter nucleosides and deoxynucleosides.

adenosine 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 this 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]

6. Summary and Outlook

In summary, the field of bioorthogonal chemistry has forged ahead in its 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, and simple cellular staining with small metabolites or drugs. A majority of the scientific studies 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 noncovalent background, covering the full UV/Vis 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 impact 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.

In addition to the three click chemistries highlighted here, there are 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 examples cited here implemented multiple labelling by combining one bioorthogonal reaction with other bioconjugation methods, or by combining two or more compatible bioorthogonal reactions.

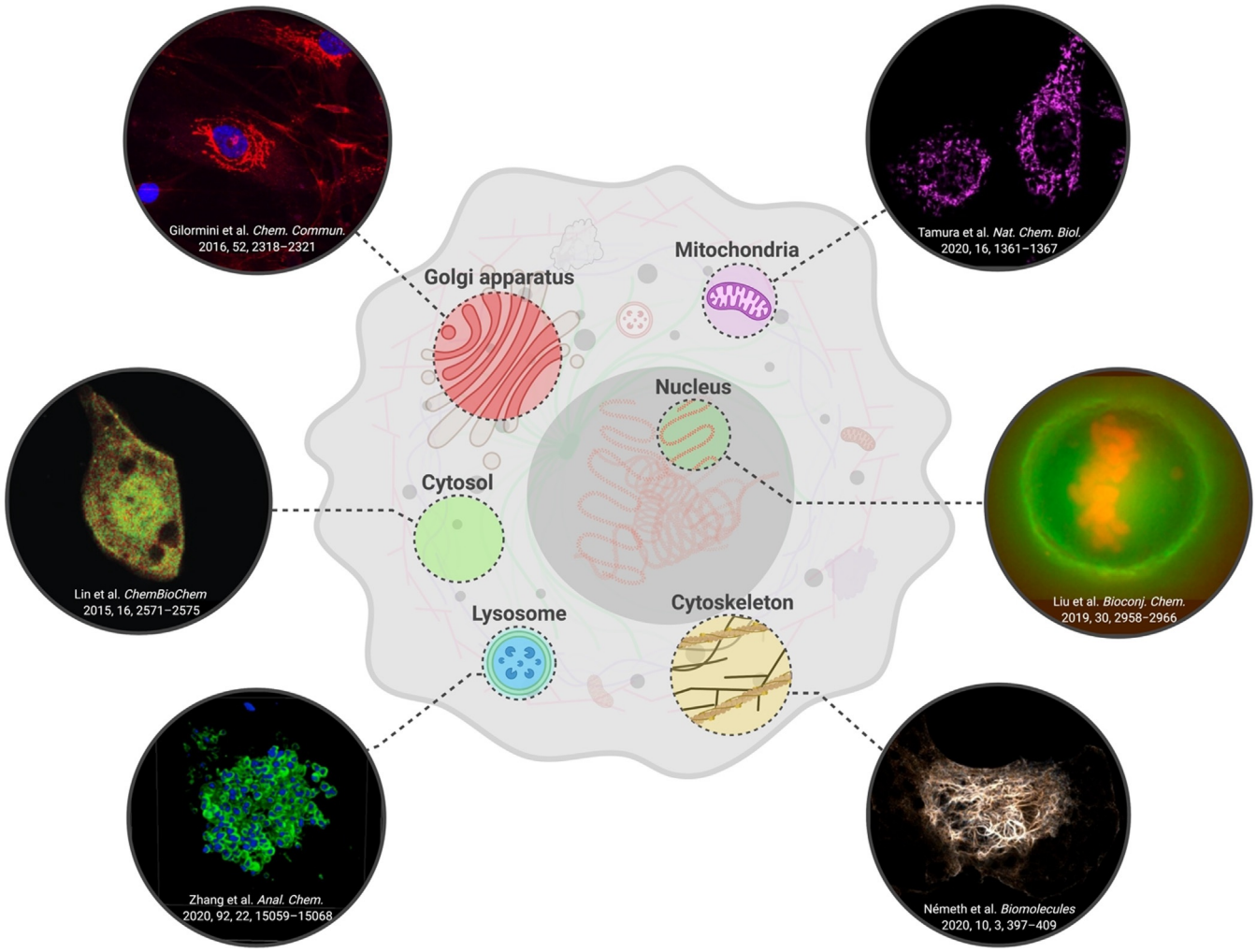


Figure 6. Exploring the intracellular map with bioorthogonal chemistry. The three main click reactions have allowed detection of specific targets in various organelles of eukaryotic cells as well as in bacteria. Top right: Mitochondrial phosphatidylcholine labelling achieved in HeLa cells by incorporation of azido-choline followed by SPAAC with a DBCO-tetraethylrhodamine probe (magenta).[145] 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-biotin then streptavidin-Alexa Fluor 488 (green).[130] 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 along with detection of the mOrange signal. Image shows the merged colocalization signal.[146] Top left: Labelling of sialylated glycoconjugates achieved in human fibroblasts by incorporation of SiaNAl followed by CuAAC with AzidoFluor 545 (red).[111] Middle left: Labelling of glycosylated EGFP-β-catenin achieved in MDCK II cells by incorporation of Ac_4GalNAz followed by CuAAC with TAMRA-alkyne and FLIM-FRET intensity-weighted lifetime detection of the EGFP signal.[105] 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).[115]

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 level of maturity that also allows routine use outside the chemistry or chemical biology laboratory. In that regard, the technology 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.

Acknowledgements

The authors thank the French Agence Nationale de la Recherche (ANR) for their financial support under reference NEURAPROBE – ANR-18-CE07-0042.

Conflict of Interest

The authors declare no conflict of interest.

[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; *Angew. Chem.* **2001**, *113*, 2056 – 2075.
[7] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596 – 2599; *Angew. Chem.* **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; *Angew. Chem.* **2011**, *123*, 8201 – 8206.
[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; *Angew. Chem.* **2018**, *130*, 16907 – 16913.
[14] V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angew. Chem. Int. Ed.* **2009**, *48*, 9879 – 9883; *Angew. Chem.* **2009**, *121*, 10063 – 10067.
[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. Inherited 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; *Angew. Chem.* **2012**, *124*, 5954 – 5958.
[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; *Angew. Chem.* **2014**, *126*, 5982 – 5986.
[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] F. Wang, Y. Zhang, Z. Du, Z. Liu, L. Zhang, X. Qu, *Angew. Chem. Int. Ed.* **2019**, *58*, 6987 – 6992; *Angew. Chem.* **2019**, *131*, 7061 – 7066.
[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.* **2016**, *374*, 16.
[33] T. Deb, J. Tu, R. M. Franzini, *Chem. Rev.* **2021**, https://doi.org/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; *Angew. Chem.* **2010**, *122*, 3447 – 3450.
[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.* **2016**, *374*, 9.
[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. Decuypère, 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; *Angew. Chem.* **2017**, *129*, 15818 – 15822.
[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; *Angew. Chem.* **2018**, *130*, 6569 – 6573.
[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; *Angew. Chem.* **2012**, *124*, 7794 – 7799.
[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; *Angew. Chem.* **2012**, *124*, 4242 – 4246.
[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; *Angew. Chem.* **2016**, *128*, 9632 – 9665.
[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; *Angew. Chem.* **2014**, *126*, 1299 – 1302.
[92] A. Dumont, A. Malleron, M. Awwad, S. Dukan, B. Vauzeilles, *Angew. Chem. Int. Ed.* **2012**, *51*, 3143 – 3146; *Angew. Chem.* **2012**, *124*, 3197 – 3200.
[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; *Angew. Chem.* **2016**, *128*, 2303 – 2308.
[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; *Angew. Chem.* **2018**, *130*, 1835 – 1838.
[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] "O-Acetylated Sialic Acids and Their Role in Immune Defense": R. Schauer, G. Srinivasan, D. Wipfler, B. Kniep, R. Schwartz-Albiez, in *The Molecular Immunology of Complex Carbohydrates-3. Advances in Experimental Medicine and*

Biology, *Vol. 705* (Ed.: A. Wu), Springer, Boston, **2011**, pp. 525 – 548.
[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; *Angew. Chem.* **2010**, *122*, 2931 – 2934.
[141] L. G. Meimetis, J. C. T. Carlson, R. J. Giedt, R. H. Kohler, R. Weissleder, *Angew. Chem. Int. Ed.* **2014**, *53*, 7531 – 7534; *Angew. Chem.* **2014**, *126*, 7661 – 7664.
[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; *Angew. Chem.* **2014**, *126*, 5915 – 5919.
[144] A. Vázquez, R. Dzijak, M. Dračínský, R. Rampmaier, S. J. Siegl, M. Vrabel, *Angew. Chem. Int. Ed.* **2017**, *56*, 1334 – 1337; *Angew. Chem.* **2017**, *129*, 1354 – 1357.
[145] T. Tamura, A. Fujisawa, M. Tsuchiya, Y. Shen, K. Nagao, S. Kawano, Y. Tamura, T. Endo, M. Umeda, I. Hamachi, *Nat. Chem. Biol.* **2020**, *16*, 1361 – 1367.
[146] E. Németh, G. Knorr, K. Németh, P. Kele, *Biomolecules* **2020**, *10*, 397 – 409.

Manuscript received: January 31, 2021
Accepted manuscript online: June 7, 2021
Version of record online: July 7, 2021

V. ACIDES SIALIQUES ET INGÉNIERIE MÉTABOLIQUE DES OLIGOSACCHARIDES

1. GLUCIDES ET GLYCOCONJUGUES

Dans le cadre de mon projet de thèse, je me suis particulièrement intéressé aux développements méthodologiques permettant d'appliquer la chimie click et la chimie bioorthogonale présentées ci-avant à l'étude de la sialylation, un type particulier de glycosylation. Les glucides constituent une des quatre grandes classes de biomolécules avec les protéines, les lipides et les acides nucléiques. Ils sont retrouvés dans tous les organismes vivants et possèdent une immense diversité structurale et fonctionnelle. La glycosylation est un phénomène cellulaire majeur consistant en l'assemblage de *glycoconjugués*, des biomolécules constituées d'une ou plusieurs chaînes oligo- ou polysaccharidiques (glycanes) liés de manière covalente à un substrat, notamment les protéines (glycoprotéines), les lipides (glycolipides) et les glyco-ARN récemment découverts,[38] formant ce qui est couramment appelé le *glycome* de la cellule. En comparaison des autres polymères biologiques comme les acides nucléiques et les protéines, le nombre de glycanes qu'il est possible de créer à partir d'un nombre restreint de monosaccharides est extrêmement élevé. Ceci est dû au fait que la structure des monosaccharides présente un grand degré de modifications possibles (substitutions des hydroxyles, anomérie, caractère multi-antenné des glycanes complexes, …). Ainsi, en considérant des polymères théoriques de six résidus de long, quatre nucléotides permettent 4096 combinaisons différentes, vingt acides aminés en permettent 64 millions, et dix monosaccharides près de 200 milliards.[39] La glycosylation permet ainsi véritablement d'étoffer les biomolécules et d'encoder une couche supplémentaire d'information lors de leur assemblage. Ce processus a lieu principalement à l'intérieur des cellules mammifères, dans le réticulum endoplasmique et l'appareil de Golgi.[40] Une très grande diversité de monosaccharides est également présente chez les bactéries, lesquels sont assemblés sous forme de glycoconjugués et polysaccharides, notamment au sein des parois bactériennes.

Les glycoprotéines sont divisées en deux catégories : les *N*-glycoprotéines et les *O*-glycoprotéines, couramment appelées *N*- et *O*-glycanes. Les *N*-glycanes sont tous basés sur une structure commune : un enchaînement GlcNAc-GlcNAc-Man-$(Man)_2$ formant le « noyau », lequel est lié au résidu Asparagine d'une protéine. Ce noyau peut être modifié de diverses manières, et on distingue trois types de *N*-glycanes : oligomannose, complexe, ou hybride. Les *O*-glycanes possèdent quant à eux un noyau constitué d'un résidu *N*-acétylgalactosamine (GalNAc) relié à un résidu sérine ou thréonine d'une protéine. Un résidu GalNAc seul ajouté de cette manière constitue l'antigène Tn, lequel peut alors être substitué de nombreuses manières formant différents noyaux, les principaux étant numérotés de un à quatre (Figure 8). La *O*-GlcNAcylation est un autre type de *O*-glycosylation qui consiste en l'ajout d'un simple résidu *N*-acéytlglucosamine (GlcNAc) sur une sérine ou une thréonine. C'est une modification essentielle des protéines mammifère, ubiquitaire et très dynamique, qui a de nombreux rôles, dont servir à prévenir la *O*-phosphorylation des protéines, modulant ainsi la signalisation en réponse à certains nutriments ou stress.[41]

Les glycolipides sont ubiquitaires et revêtent de nombreuses fonctions biologiques. Leur caractère amphiphile leur confère des propriétés structurales et un rôle primordial dans le maintien de l'intégrité membranes biologiques. Ils revêtent également des fonctions de communication et d'interaction cellulaire, de modulation du système immunitaire, ou encore de transport d'électrons dans les systèmes photosynthétiques.[42]

Les glycoconjugués sont distribués dans l'ensemble de la cellule, intervenant dans tous les mécanismes physiologiques primordiaux. Ils sont néanmoins particulièrement concentrés à la surface des cellules. En effet, toutes les cellules eucaryotes possèdent à leur surface une épaisse couche appelée *glycocalyx*, constituée des glycanes libres, des glycolipides et des glycoprotéines membranaires, dont le rôle, longtemps resté mystérieux, a depuis été connecté à de nombreux processus tels que le développement embryonnaire, l'immunité, ou l'infection virale et bactérienne.[43]

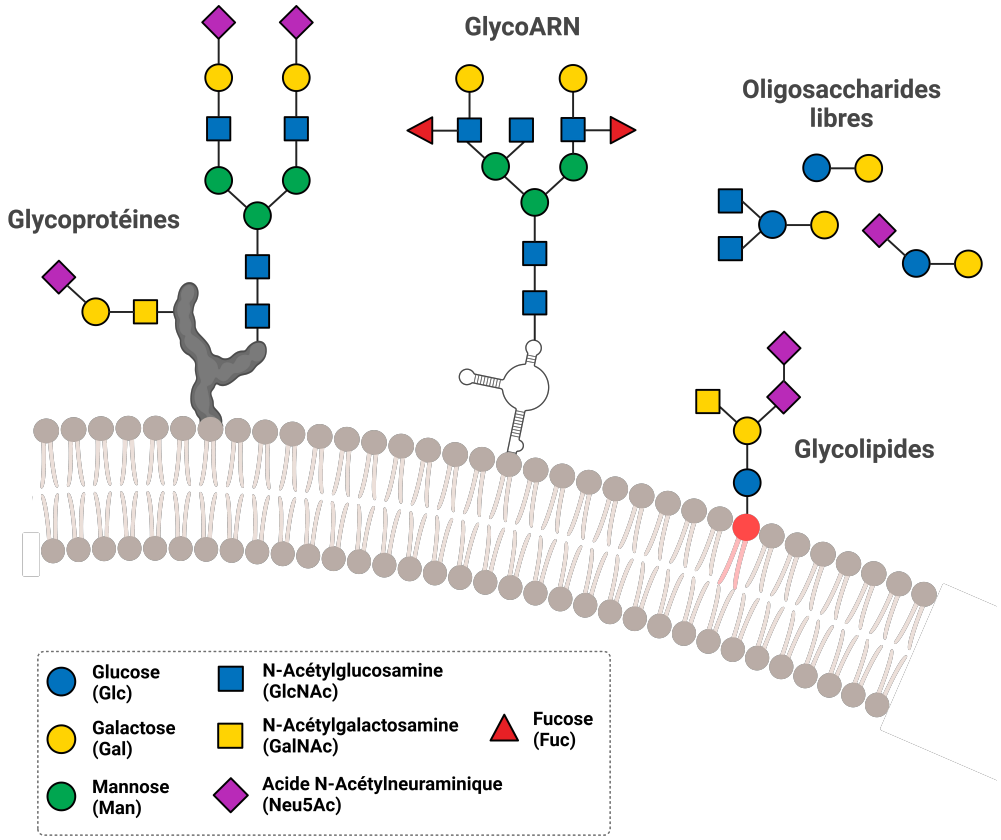


Figure 8 Différents types de glycoconjugués. Les grandes classes de biomolécules naturelles (protéines, lipides, ARN) sont étoffées par des chaînes de monosaccharides modulant leurs fonctions et propriétés.

Parmi les monosaccharides formant les parties glycaniques de ces structures complexes, on retrouve les acides sialiques. Ceux-ci composent toute une famille de glucides particuliers, et sont principalement retrouvés en position terminale des glycanes.

2. LES ACIDES SIALIQUES

Les acides sialiques tirent leur nom du grec *sialon* (salive) car ils ont été découverts pour la première fois dans des glandes salivaires. En 1936, Gunnar Blix, à l'origine de la découverte des glycolipides, isola un monosaccharide cristallin jusqu'alors inconnu à partir de glandes sous-maxillaires bovines,[44] qu'il nomma plus tard *acide sialique*. En parallèle de ces travaux, une autre équipe de recherche, celle de Klenk, isola un lipide dans le tissu cérébral d'un patient, et observa des similitudes avec le composé décrit par Blix, démontrant qu'il s'agissait en fait d'un glycolipide qu'il baptisa ganglioside (constituant aujourd'hui une large famille de glycolipides, principalement retrouvés dans le système nerveux périphérique). Quelques années plus tard, Klenk identifia un dérivé

methoxylé d'un monosaccharide très proche dans des glycolipides cérébraux et le nomma acide neuraminique. La relation entre les deux structures isolées par ces deux équipes ne sera élucidée que plus tard : Blix suggéra en 1952 que l'acide neuraminique est en réalité un analogue désacétylé de l'acide sialique, ce qui fut plus tard confirmé par le groupe de Klenk.[45] L'acide sialique « originel » est aujourd'hui appelé acide *N*-acétylneuraminique (Neu5Ac), et depuis de nombreux monosaccharides apparentés ont été identifiés. Ces molécules appartiennent à une famille particulière de cétoses à neuf carbones possédant une fonction acide carboxylique anionique à pH physiologique, dont les principaux membres et dérivés sont représentés en Figure 9.

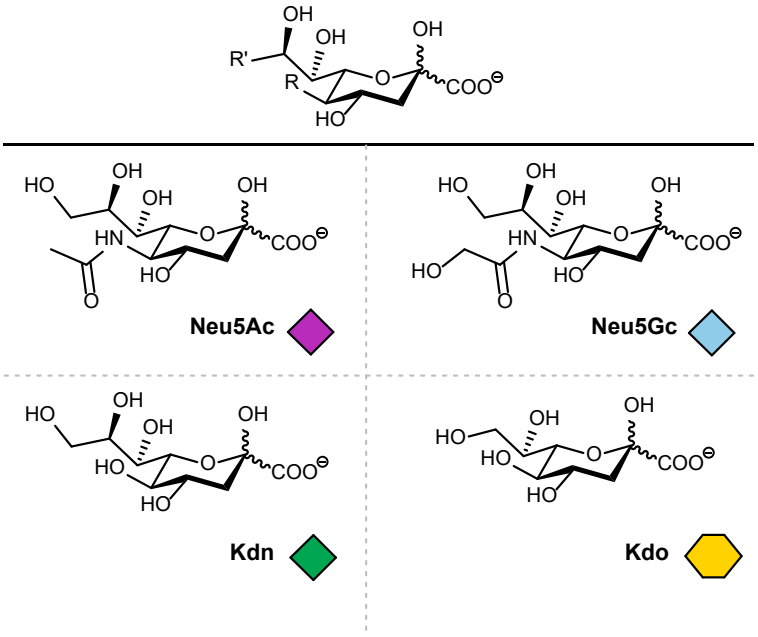


Figure 9 Structure générale des acides sialiques, et principaux représentants de cette famille retrouvés dans la nature.

Le Neu5Ac est le principal représentant de cette famille de molécules. Les deux équipes citées plus haut se sont associées pour décrire dans le journal *Nature* une nomenclature officielle selon laquelle tous les acides neuraminiques acylés sont regroupés sous le terme « acide sialique ».[46] Les acides sialiques deutérostomiens possèdent la plus grande diversité structurale ; ils peuvent être substitués de nombreuses manières (acétylation, méthylation, sulfatation, …). Ils occupent souvent les positions terminales des glycanes et sont ainsi présents à la périphérie des cellules, jouant un rôle fonctionnel extrêmement important dans les phénomènes de reconnaissance et d'interaction cellulaire, par exemple dans l'immunité innée, le développement et la reproduction.[47–49]

Il existe différents types de glycoconjugués pouvant être sialylés : chez les mammifères, ce sont principalement les glycolipides, les glycoprotéines et les acides polysialiques. Des modifications de la nature ou du taux d'expression des acides sialiques sont souvent observées lors du développement de cancers.[50,51] Certains antigènes sialylés associés aux cancers, notamment le SialylTn, constituent des cibles thérapeutiques importantes.[52,53]

Certaines bactéries possèdent également des glycoconjugués sialylés ; en effet, bien que l'on ait longtemps considéré que les acides sialiques n'existaient que chez les animaux supérieurs et chez quelques souches pathogènes particulières, une diversité structurale importante de sialoglycoconjugués a depuis été mise en évidence chez un nombre important d'espèces bactériennes.[54] Les principaux acides sialiques microbiens sont le Neu5Ac et le Kdn, ainsi que des dérivés nonulosonates apparentés, spécifiques à certaines bactéries (acide légionaminique, acide pseudaminique). Ils sont souvent associés au caractère pathogène des bactéries les exprimant (comme *Escherichia coli* K1, *Neisseria meningitidis*, *Legionella pneumophila* …).[48] Les bactéries tirent plusieurs avantages des acides sialiques, il a par exemple été montré que le LPS de bactéries du genre *Streptococcus* incorpore des motifs sialylés normalement reconnus en tant que *soi* par les neutrophiles, permettant d'échapper au système immunitaire de l'hôte.[55]

2.1. BIOSYNTHÈSE DES ACIDES SIALIQUES

Au sein des cellules mammifères, la biosynthèse des acides sialiques se déroule principalement en milieu intracellulaire. Celle-ci débute dans le cytosol, via l'action de la UDP-GlcNAc 2-épimérase (GNE) qui catalyse la conversion de l'UDP-GlcNAc en ManNAc. La réaction étant irréversible, cette enzyme n'est pas une épimérase à proprement parler. Cette enzyme existe sous deux formes dites hydrolysante et non hydrolysante (qui fournit l'UDP-ManNAc au lieu du ManNAc). Chez les mammifères, seule la version hydrolysante existe, tandis que les deux versions sont retrouvées chez les bactéries selon l'espèce étudiée, fournissant à celles-ci une voie de synthèse d'UDP-ManNAc, précurseur d'une grande variété de polysaccharides de surface, mais pas des acides sialiques.[56]

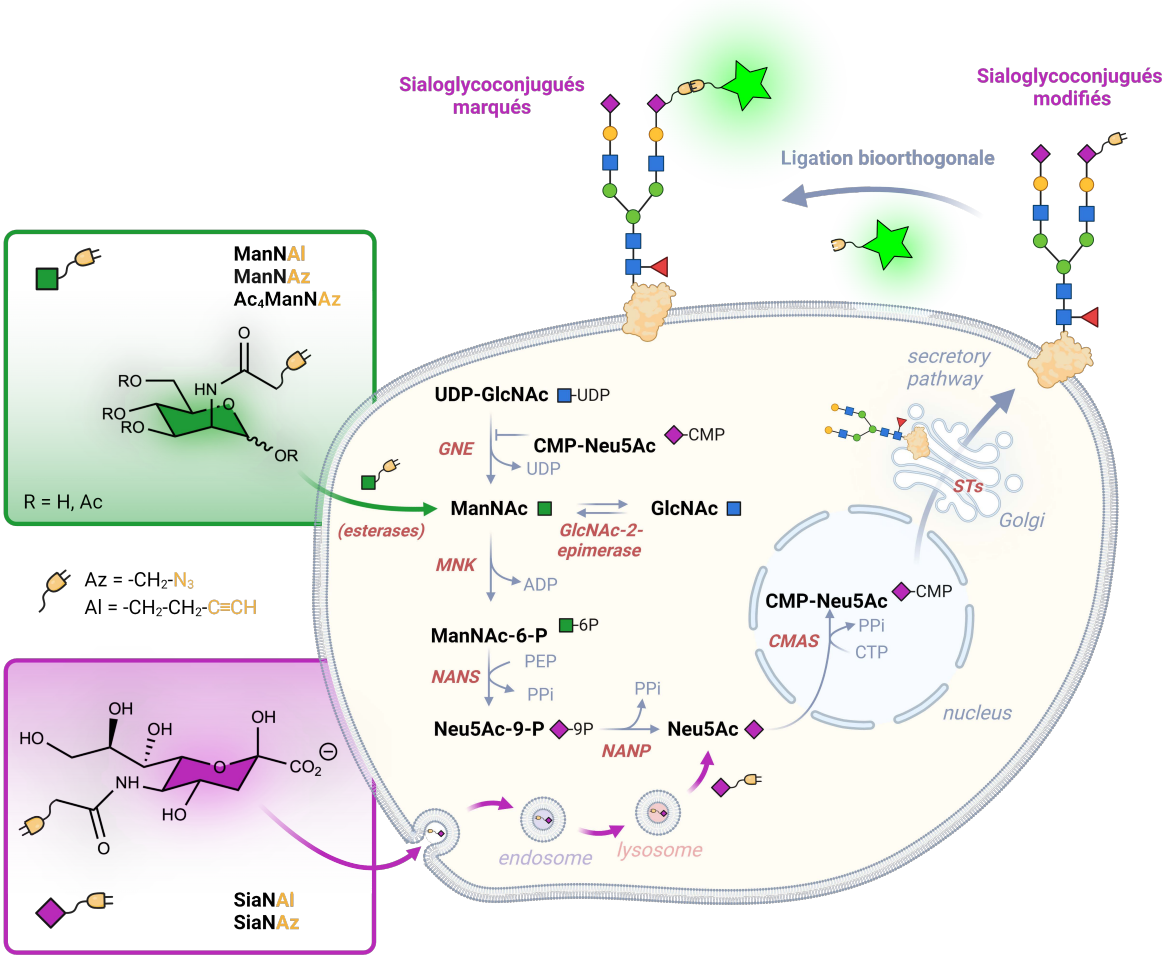


Figure 10 Biosynthèse mammifère des acides sialiques et schéma récapitulatif de la stratégie du rapporteur chimique appliquée à leur étude. Des analogues de ManNAc ou de Neu5Ac (ici portant un groupement azoture ou alcyne) peuvent être utilisés pour marquer les sialoglycoconjugués de cellules en culture.

La GNE mammifère est bifonctionnelle, elle possède également une activité kinase, et catalyse la seconde étape de la biosynthèse en phosphorylant la position 6 du ManNAc, fournissant le ManNAc-6-P. Cette étape n'est pas requise chez les bactéries qui procèdent directement à la prochaine étape à partir du ManNAc, via l'action de la Sialic acid synthase. Celle-ci condense le phosphoénolpyruvate (PEP) pour donner le Neu5Ac chez les bactéries, ou le Neu5Ac-9-P chez les mammifères, lequel subira alors l'action d'une phosphatase pour libérer le Neu5Ac. Avant d'être incorporé au sein des sialoglycoconjugués, le Neu5Ac est préalablement activé sous forme cytidine monophosphate (CMP-Neu5Ac) par la CMP-Sialic acid synthetase (CSS). Cette étape a lieu dans le noyau des cellules mammifères, à proximité de l'appareil de Golgi. Les CMP-Neu5Ac exportés du noyau sont alors pris en charge par un transporteur membranaire qui internalise le CMP-Neu5Ac dans l'appareil de Golgi, au sein duquel les acides

sialiques seront transférés sur les glycoconjugués en naissance, via l'action des sialyltransférases (Figure 10). Les acides sialiques assemblés sur les glycoconjugués peuvent être modifiés de nombreuses manières, principalement par *O*-acétylation d'un ou plusieurs groupements hydroxyles (notamment les positions 7, 8 et 9), mais également par méthylation, phosphorylation ou sulfatation,[48,57] ce qui ajoute encore un degré de précision à l'information qu'ils véhiculent. Une grande partie des glycoconjugués ainsi formés est alors exportée à la membrane des cellules, où ils jouent un rôle important dans les phénomènes d'interaction et de reconnaissance cellulaire.

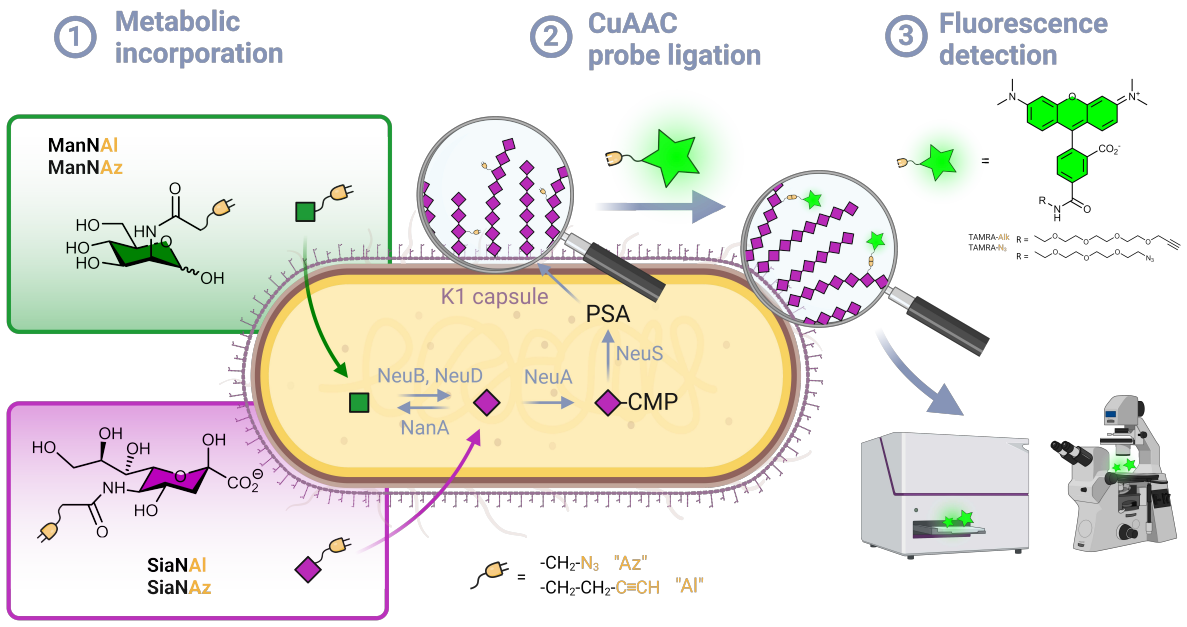


Figure 11 Biosynthèse de la capsule K1 d'E. coli et schéma récapitulatif de la stratégie du rapporteur chimique appliquée à son étude. Des analogues de ManNAc ou de Neu5Ac (ici portant un groupement azoture ou alcyne) peuvent être utilisés pour marquer les chaînes poly acide sialique de la capsule.

Si la plupart des bactéries n'expriment pas d'acides sialiques, les espèces les exprimant présentent souvent une pathogénicité exacerbée et causent de sévères maladies.[58] Chez ces espèces, toutes les étapes de la biosynthèse des acides sialiques ont lieu dans le périplasme, avant export des sialoglycoconjugués. Il est important de noter qu'en parallèle de cette voie de synthèse *de novo*, la plupart de ces bactéries possèdent des voies de *scavenging*, permettant de récupérer les acides sialiques dans leur environnement. Ceux-ci peuvent être présents sous forme libres ou après les avoir détachés des sialoglycoconjugués de l'hôte via l'action de sialidases, comme cela est observé chez de

nombreux pathogènes.[59] Certaines espèces ont même abandonné leurs voies de biosynthèse *de novo* et récupèrent uniquement les acides sialiques CMP activés des sécrétions de l'hôte, comme *Neisseria gonorrhoeae*.[60]

3. INGÉNIERIE MÉTABOLIQUE DES OLIGOSACCHARIDES

L'utilisation de la stratégie du rapporteur chimique pour incorporer des modifications au sein des glycanes en vue de les détecter ou de moduler leur activité est souvent désignée sous le terme d'ingénierie métabolique des oligosaccharides (*Metabolic Oligosaccharide Engineering*, MOE). Cette stratégie peut être appliquée à l'étude spécifique de plusieurs types de monosaccharides, dont les acides sialiques. Elle a par exemple été utilisée au sein de cellules mammifères pour mettre en évidence le mécanisme responsable d'un défaut de glycosylation dans un contexte pathologique,[61] ainsi que chez les bactéries, permettant le développement d'une méthode de marquage spécifique des bactéries Gram négatives grâce à un dérivé azoture du Kdo (Kdo-N_3), un monosaccharide apparenté aux acides sialiques spécifique du noyau interne du lipopolysaccharide (LPS).[62]

Le champ d'application de la MOE est extrêmement large et ne se limite pas qu'à l'imagerie. Par exemple, cette technique de marquage du LPS au Kdo-N_3 permet également d'isoler les bactéries viables d'un mélange, en procédant à la ligation bioorthogonale avec une biotine, puis en utilisant des billes magnétiques fonctionnalisées par des streptavidines, permettant de séparer les bactéries ayant incorporé le Kdo-N_3 du reste du mélange.[63] De la même manière, en modifiant un précurseur de l'acide légionaminique (un acide sialique spécifique de l'antigène O de *L. pneumophila*), des chercheurs ont mis au point une méthode de détection spécifique de cette bactérie au sein d'un mélange complexe, permettant d'effectuer un diagnostic de légionellose en une journée, contre dix jours en moyenne pour une identification par les méthodes classiques.[64]

Différents moyens peuvent être envisagés pour suivre la sialylation grâce à la stratégie du rapporteur chimique, basés sur deux types d'outils moléculaires : 1) on peut utiliser des acides sialiques modifiés portant un groupement rapporteur, ou 2) introduire le

groupement rapporteur au sein de précurseurs métaboliques des acides sialiques, comme la *N*-acétyl-D-mannosamine (ManNAc). Après l'étape d'incorporation métabolique, le groupement rapporteur sera ainsi retrouvé dans les glycoconjugués biosynthétisés pendant la période d'incubation avec le monosaccharide modifié (Figure 10).

La grande majorité des études de MOE a recours à des rapporteurs chimiques sous forme peracétylée, masquant les hydroxyles libres des monosaccharides et diminuant grandement leur polarité, ce qui leur permet de traverser les membranes cellulaires par diffusion passive et améliore le taux d'incorporation.[65] Une fois à l'intérieur de la cellule, ces dérivés doivent alors subir l'action d'estérases aspécifiques libérant les hydroxyles avant de pouvoir intégrer la voie de métabolisation des acides sialiques.[66] L'action de ces estérases peut néanmoins libérer des formes partiellement acétylées pouvant perturber les voies métaboliques.[67] De plus, bien que viable pour le marquage de cellules eucaryotes, cette stratégie n'est généralement pas adaptée pour le marquage des bactéries, qui ne possèdent pas une activité estérase aspécifique suffisante pour libérer les monosaccharides après diffusion dans la cellule.[68] L'utilisation de rapporteurs « libres » permet de surcroît d'étudier plus finement les mécanismes d'incorporation, ce qui a été mis à profit au laboratoire via l'utilisation conjointe d'analogues de Neu5Ac et de ManNAc portant un groupement alcyne, l'acide *N*-(4-pentynoyl) neuraminique (SiaNAl) et la *N*-(4-pentynoyl)-D-mannosamine (ManNAl) respectivement, permettant de mettre en évidence l'existence d'un transporteur actif internalisant le ManNAc au sein de cellules mammifères,[61] ce qui n'aurait pas été possible en ayant recours à des dérivés peracétylés diffusant passivement au travers des membranes. Mon projet de thèse se place en aval de ces recherches et se décline en plusieurs grands axes, que j'aborderai en trois chapitres dans cette thèse sur articles.

VI. OBJECTIFS DE LA THÈSE :

Mon projet de thèse vise à développer des outils et méthodologies chimiques permettant d'étudier les phénomènes de sialylation, et à les appliquer à des études concrètes en modèles bactériens et humains. Il se scinde en plusieurs grands axes, que j'ai séparés en 3 grands chapitres pour la rédaction de ce manuscrit.

Le premier objectif de mon projet de thèse, décrit dans le chapitre 1, est d'utiliser la MOE pour détecter la capsule polysaccharidique K1 d'*Escherichia coli*, un homopolymère de Neu5Ac conférant des capacités d'évasion au système immunitaire aux souches qui l'expriment, souvent responsables de sepsis et méningites périnatales.[69] Une telle méthode, en plus de son but premier qui est de fournir un outil méthodologique permettant d'examiner en profondeur les mécanismes de biosynthèse et d'expression de la capsule K1, pourrait être mise à profit pour l'identification de composés affectant son expression, ce qui impacterait la pathogénicité de la bactérie et non sa viabilité, fournissant une alternative potentielle aux antibactériens actuels.

Le deuxième objectif, présenté dans le chapitre 2, consiste à comparer l'incorporation des rapporteurs chimiques alcynes et azotures pour déterminer l'impact du groupement utilisé et de la nature du rapporteur (dérivé de Neu5Ac ou de ManNAc, azoture ou alcyne, peracétylé ou libre) sur l'incorporation au sein des sialoglycoconjugués. L'idée est de savoir si les groupements rapporteurs ont un impact sur la métabolisation, ou s'ils peuvent effectivement être considérés comme une modification imperceptible de la molécule naturelle. Pour effectuer cette étude, la ligation bioorthogonale sera effectuée par CuAAC, qui permet d'intervertir les groupements azotures et alcynes sur la sonde et le rapporteur chimique. Nous avons décidé d'effectuer cette étude comparative au sein de différentes lignées cellulaires mammifères, les marquages métaboliques au sein de ce modèle étant plus communs, et mieux maîtrisés au laboratoire.

Enfin, le troisième objectif, décrit dans le chapitre 3, consiste à développer des outils nouveaux pour l'étape de détection, permettant de repousser les limites actuelles de la MOE, notamment en matière de résolution, grâce au développement de sondes

organométalliques possédant un atome d'Iridium dans leur structure. En effet, la résolution spatiale maximale accessible avec les sondes fluorescentes classiques et la microscopie confocale est régie par la diffraction limite d'Abbe (environ 200nm), mais elle peut être franchie en ayant recours à d'autres techniques comme la microscopie électronique. Nous souhaitons de ce fait développer des sondes versatiles permettant à la fois de suivre l'incorporation et la métabolisation des rapporteurs grâce à la microscopie à fluorescence classique, ainsi qu'avec une résolution améliorée, grâce à des techniques telles que la microscopie associée à la fluorescence des rayons X (XRF) sous faisceau de lumière synchrotron. La présence de l'atome d'Iridium dans la structure du fluorophore modifie les propriétés de ce dernier et allonge notamment le temps de vie à l'état excité (environ 200ns, contre 10ns en moyenne pour les fluorophores classiques ou l'autofluorescence des composés biologiques).[70] Ceci permet d'avoir recours à des techniques telles que la microscopie en temps de vie (*fluorescence lifetime imaging*, FLIM) pour la détection, qui permet de discriminer les signaux par résolution temporelle en plus de la résolution spectrale déjà utilisée en microscopie à fluorescence classique, améliorant potentiellement le rapport signal sur bruit lors des acquisitions. Le métal peut également être détecté via des techniques quantitatives très précises et sensibles, comme la spectrométrie de masse à plasma à couplage inductif (ICPMS), fournissant de précieuses informations sur le taux d'incorporation des rapporteurs, et permettant potentiellement de relier des intensités arbitraires de fluorescence à des données quantitatives précises.

Au-delà des objectifs primaires de mon projet, un objectif secondaire était de participer au transfert des outils et technologies mis au point au laboratoire dans le cadre de collaborations avec des équipes de biologistes. Ainsi, j'ai synthétisé et fourni des outils chimiques qui ont pu être utilisés et valorisés sur des applications précises, ce qui sera présenté à la fin de mon manuscrit. Pour réaliser ces objectifs, et pour l'ensemble de mes travaux, j'ai recouru au même panel de rapporteurs chimiques. J'ai pour cela synthétisé différents analogues alcynes et azotures de ManNAc et de Neu5Ac, qui m'ont alors servi pour toutes mes expériences de marquage métabolique (Figure 12 et 13).

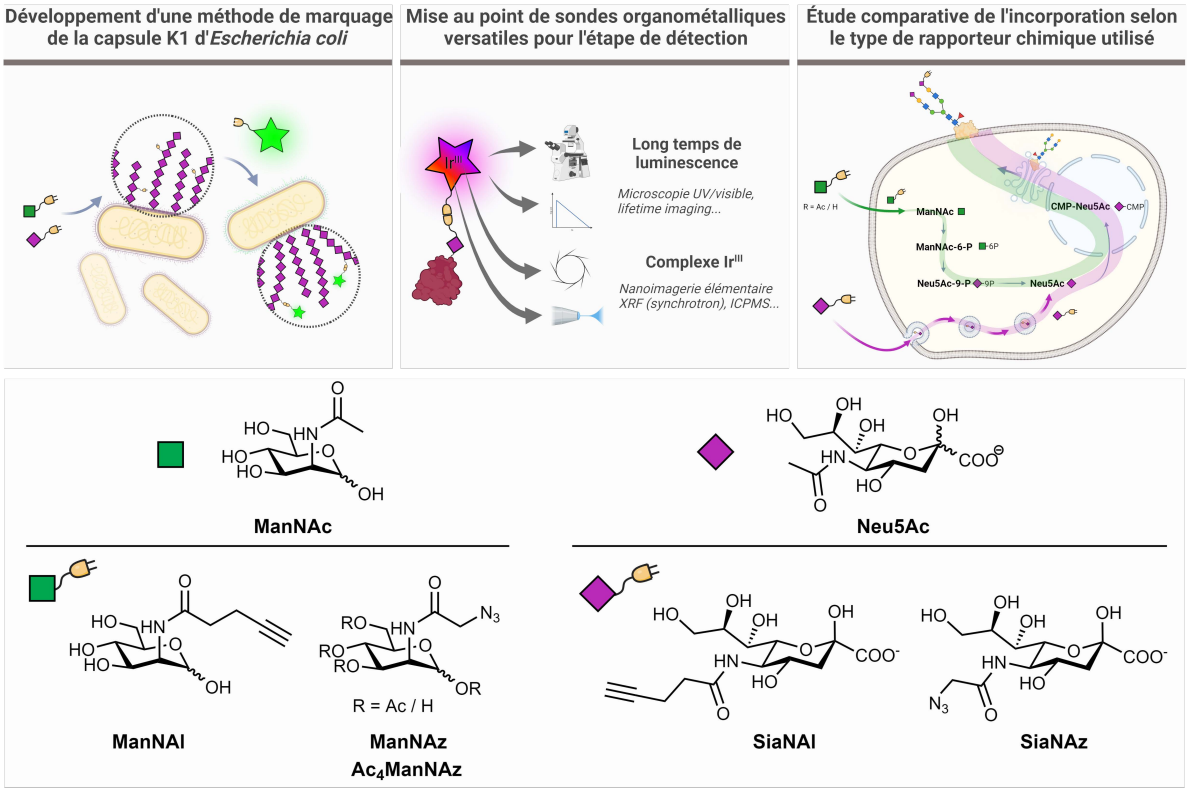


Figure 12 Schéma récapitulatif des objectifs de ce projet de thèse et structure des différents rapporteurs chimiques utilisés pour les accomplir.

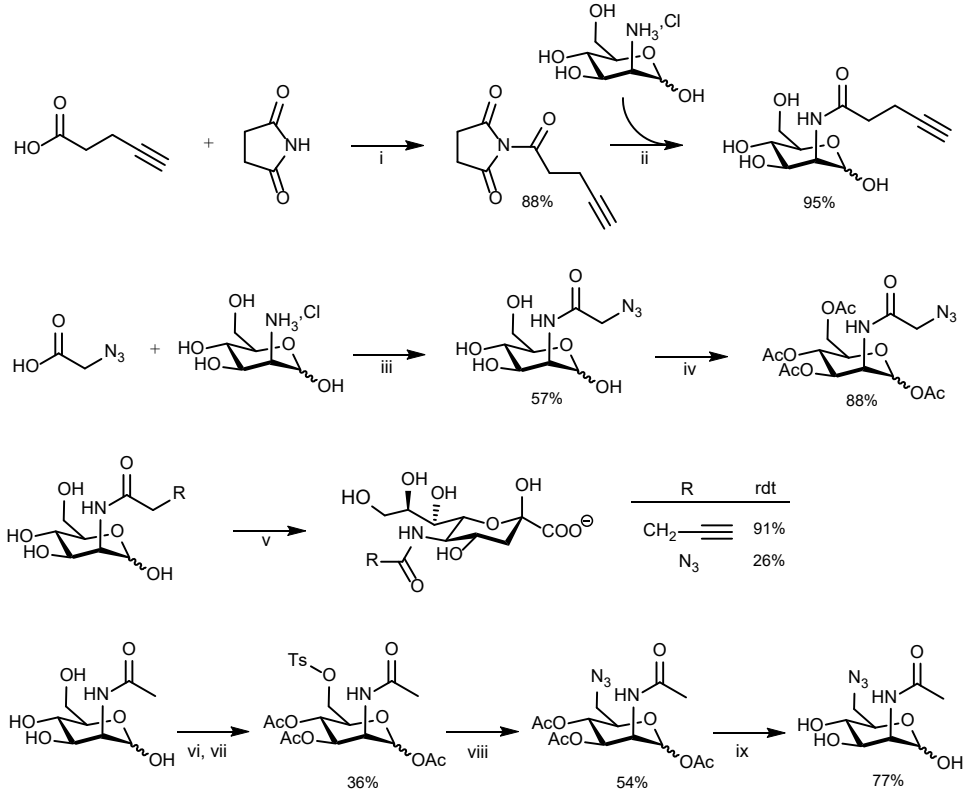


Figure 13 Synthèse des rapporteurs chimiques utilisés pour ce projet de thèse. (ii, iii) EDC, CH_2Cl_2, rt, 3h30 ; (ii) D-Mannosamine hydrochloride, Et_3N, DMF, t.a., 16h ; (iii) D-Mannosamine hydrochloride, DIC, HOBt, DIPEA, DMF, t.a., 19h ; (iv) DMAP, Ac_2O, rt, 16h ; (v) Sodium pyruvate, Neu5Ac aldolase (EC 4.1.3.3), PBS, 37°C, 18h ; (vi) Chlorure de tosyle, pyridine, 0°C, 5h ; (vii) Ac_2O, pyridine, 0°C à t.a., 14h ; (viii) nBu_4NN_3, toluène, 60°C, 6h ; (ix) NaOMe, MeOH, 0°C, 2h

COMPÉTENCES TECHNIQUES DÉVELOPPÉES AU COURS DE LA THÈSE :

Ce projet de thèse pluridisciplinaire se situe à la frontière entre la chimie organique, la biologie et la biochimie, ainsi que la microscopie et l'analyse d'image. Plutôt que de chercher à me spécialiser dans un domaine particulier, j'ai tenu lors de la réalisation de celui-ci à me familiariser avec toutes ces disciplines à la fois, et à naviguer à leur interface. Ceci me m'a permis de développer des compétences techniques nombreuses et variées, et de garder le contrôle sur toutes les étapes du projet en les effectuant moi-même. Ainsi, j'ai développé de solides compétences :

(1) En synthèse organique, nécessaire pour l'obtention, la purification et l'analyse des diverses molécules qui furent nécessaires à la réalisation de ces expériences. En plus de la synthèse organique « classique », certains pans de ma thèse m'ont également permis de me former à l'analyse photochimique des molécules, notamment pour la caractérisation de sondes photoluminescentes novatrices.

(2) En culture cellulaire, aussi bien mammifère que bactérienne (souches pathogènes niveau L2), qui fût nécessaire pour le maintien stable des lignées, pour les étapes d'incorporation métaboliques, pour les étapes de chimie *in vivo*, mais également pour l'extraction et l'analyse des biomolécules par des techniques de biochimie (électrophorèse, Western-blot, …).

(3) En microscopie, pour la préparation des échantillons, qui varie grandement selon la méthode d'observation ; l'acquisition des images, sur microscope à champ large ou confocal à fluorescence classique, résolu en temps (analyses FLIM), ou sous faisceau synchrotron (pour l'imagerie SR-XRF) ; et enfin le traitement et l'analyse informatique et statistique des images obtenues. Ainsi, je peux aujourd'hui m'adresser à un large public de spécialistes et faire le lien entre de nombreuses disciplines différentes.

Pour des raisons pratiques détaillées précédemment, la majorité des applications de la chimie bioorthogonale et de la stratégie du rapporteur chimique sont réservées au marquage membranaire des cellules mammifères. Or, au sein du laboratoire, mes encadrants avaient depuis quelques années l'idée d'aborder un autre groupe d'organismes vivants : les bactéries. L'idée était de développer de nouveaux outils chimiques et méthodologies pour le marquage des capsules bactériennes de bactéries pathogènes. Cette idée mûrit et conduit au financement du projet ANR Neuraprobe, en collaboration avec le Dr Boris Vauzeilles à l'Institut de Chimie des Substances Naturelles (ICSN), au sein duquel ma thèse prend place.

I. MEMBRANES ET PAROIS BACTÉRIENNES

Les bactéries sont couramment séparées en deux catégories selon le résultat obtenu après coloration de Gram, qui permet de distinguer les bactéries dites Gram positives et Gram négatives. Les bactéries Gram positives sont monodermes, c'est-à-dire qu'elles ne possèdent qu'une seule membrane biologique, laquelle est entourée d'une épaisse paroi (environ 20 à 80nm) formée par le peptidoglycane. Les bactéries Gram négatives sont quant à elles didermes, elles possèdent deux membranes biologiques, séparées par un espace nommé périplasme qui contient également une paroi formée par le peptidoglycane, lequel est significativement plus fin que chez les Gram positives (environ 10nm).[71] La face extérieure de la membrane externe est composée d'un lipide particulier, le lipopolysaccharide (LPS), qui est essentiel pour la survie de la bactérie et revêt de nombreuses fonctions, et qui est très immunogène. Les polysaccharides de surface sont un composant essentiel des membranes bactériennes qui sont retrouvés sous différentes formes. De nombreuses bactéries produisent de longues chaînes de polysaccharides qui s'assemblent pour former une structure capsulaire enveloppant la cellule. Cette capsule confère à la bactérie des propriétés protectrices face à des stress biologiques, chimiques,

physiques, etc. Elle permet également l'association des bactéries sous forme de biofilm et influe sur de nombreux phénomènes d'interaction avec leur environnement.[72]

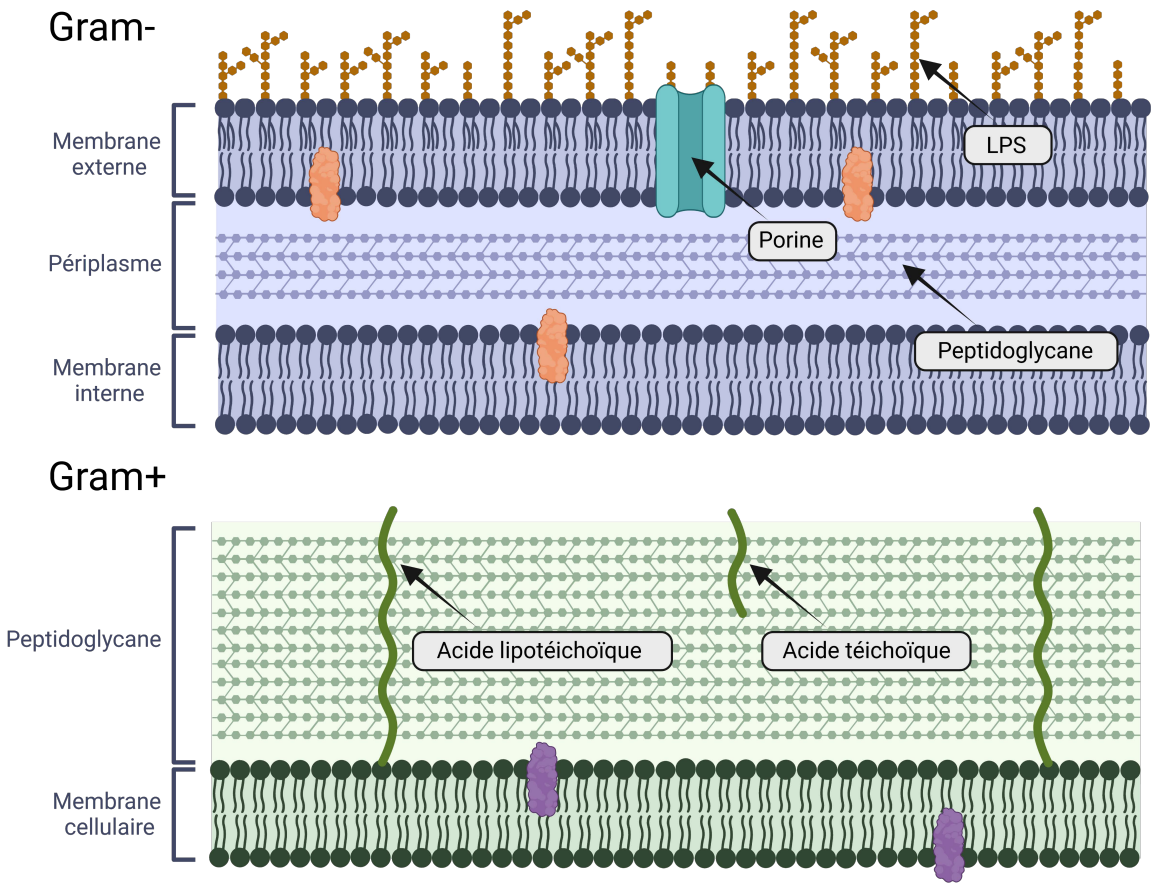


Figure 14 Différence entre la paroi bactérienne des espèces dites Gram négatives et Gram positives.

II. GLYCOSYLATION ET ACIDES SIALIQUES BACTÉRIENS

La glycosylation joue un rôle majeur dans la virulence des bactéries pathogènes et différents rôles ont été identifiés pour des glycoconjugués de surfaces bactériens, notamment dans la motilité, l'adhésion et la manipulation du système immunitaire. Les acides sialiques sont relativement rares chez les bactéries, mais les espèces en exprimant ont souvent une virulence exacerbée. Ils sont retrouvés plus fréquemment chez les bactéries pathogènes, où ils ont souvent pour rôle d'imiter les glycanes de l'hôte, échappant ainsi au système immunitaire par mimétisme moléculaire. Parmi ces facteurs de virulence, on retrouve souvent des monosaccharides ou des mécanismes de glycosylation peu usuels au sein des cellules humaines, ce qui en fait des cibles de choix pour la recherche de nouveaux antibactériens, particulièrement dans des thérapies visant

à masquer la pathogénicité des bactéries sans pour autant les éliminer, ce qui conduirait à l'apparition de résistances telles qu'observées avec les antibiotiques actuels.[73] Ainsi, plusieurs acides sialiques sont exclusifs aux bactéries, tels que l'acide légionaminique et pseudaminique.[48]

Comme vu en introduction, certaines bactéries ont développé une voie de biosynthèse *de novo* d'acides sialiques ou d'acides nonulosoniques apparentés, telles que *Mannheimia haemolytica*[74] et *Neisseria meningitidis*,[54] tandis que d'autres sont capables de prélever les acides sialiques dans leur environnement (phénomène de scavenging), comme *Neisseria gonorrhoeae* ou plusieurs espèces du genre *Trypanosoma* (responsables de la trypanosomiase africaine, couramment appelée maladie du sommeil).[54]

On retrouve ces structures au niveau du lipopolysaccharide (LPS) (antigènes O, noyau), ou encore au sein de la capsule polysaccharidique (antigènes K) de certaines souches bactériennes encapsulées (Figure 15), principalement Gram négatives. Les capsules confèrent de nombreuses propriétés aux bactéries les exprimant, souvent directement en lien avec leur virulence, telles que des capacités de camouflage vis-à-vis du système immunitaire, de résistance à la phagocytose, etc. Elles sont également retrouvées au sein d'espèces non pathogènes, où elles jouent un rôle de protection environnemental contre certains stress comme la dessication.[75] Plus de 80 capsules bactériennes de structure différentes ont été répertoriées,[76] certaines étant plus fréquentes et virulentes que d'autres. La plupart sont de nature polysaccharidiques, mais certaines capsules sont constituées d'autres molécules, à l'instar de la capsule poly acide glutamique de *Bacillus anthracis*.[77]

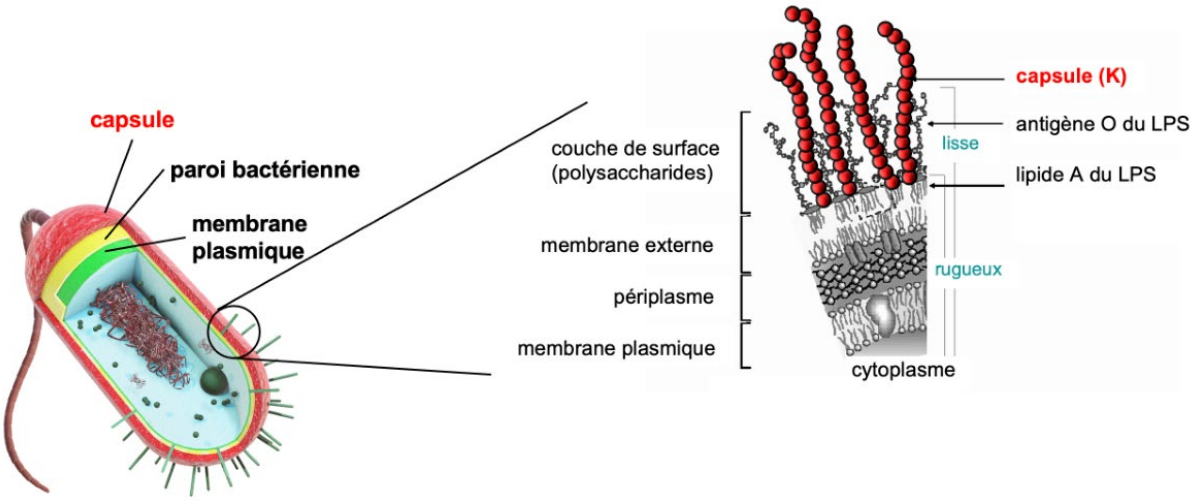


Figure 15 Représentation et structure de la capsule bactérienne.

III. ACIDE POLY-SIALIQUE ET CAPSULE K1 D'*ESCHERICHIA COLI*

La cible principale de notre projet est la capsule K1 d'*E. coli*, également appelé acide colominique. *E. coli* est un type de bactéries commensales retrouvée dans l'intestin des êtres humains et mammifères, dont la plupart des souches sont inoffensives. En revanche, les souches *E. coli* exprimant la capsule K1 sont particulièrement virulentes. Il est établi que cette capsule procure l'un des effets les plus forts sur la pathogénicité des souches l'exprimant, et elle est directement liée à de graves infections.[78] Ces souches sont responsables de méningites et sont connues pour migrer du système vasculaire, après colonisation des muqueuses, vers le système nerveux en traversant la barrière hématoencéphalique, conduisant à une mortalité élevée,[69] notamment lors d'infections néo-natales. Cette virulence accrue est due à la partie polysaccharidique de la capsule K1, appelée acide α2,8-polysialique (PSA), composée d'un maximum de 200 résidus Neu5Ac. La capsule K1 est également produite par d'autres bactéries Gram-négatives pathogènes telles que *N. meningitidis* (méningites), *M. haemolytica* (pneumonies chez le bétail) et *Moraxella nonliquefaciens* (infection des yeux).

Le PSA de la capsule a une structure linéaire, qui est également retrouvée dans le cerveau humain. En effet, le PSA est une glycosylation d'un complexe protéique jouant un rôle dans le développement du système nerveux, le Neural Cell Adhesion Molecule (NCAM), et est également exprimé sur certaines cellules immunitaires.[79–82] Ainsi, les bactéries

couvertes par le PSA sont masquées et parviennent à échapper à l'action du système immunitaire de l'hôte, en empêchant notamment la phagocytose par les cellules immunitaires.[83] Il a été montré que le ciblage thérapeutique de la capsule K1 permettant de re-sensibiliser *E. coli* à l'action du système immunitaire.[78] Le PSA est assemblé dans le cytosol de la bactérie avant d'être ancré sur un lipide, le Lyso-phosphatidylglycérol, puis exportée à la surface de la cellule. Ces lipides forment des radeaux lipidiques qui s'accumulent en un point aléatoire de la membrane externe, jusqu'à s'étendre sur l'ensemble de la cellule et former la capsule, ce qui a été montré récemment grâce à des techniques de microscopie super-résolutives.[84] Les mécanismes d'export de la capsule restent cependant méconnus à l'heure actuelle.

IV. INGÉNIERIE MÉTABOLIQUE DES GLYCANES CHEZ LES BACTÉRIES

La chimie bioorthogonale et le marquage métabolique ont principalement été appliqués au marquage des protéines chez les bactéries, en ayant recours à des rapporteurs acides aminés non-canoniques.[85,86] Les exemples d'application de la MOE sont plus rares quant à eux, notamment car cette approche a historiquement été développée et appliquée principalement pour les marquages sur cellules humaines,[87] mais aussi car le métabolisme des glycanes chez les bactéries varie largement d'une espèce à l'autre, compliquant l'implémentation de cette stratégie. Néanmoins, celle-ci a été appliquée avec succès au marquage de certains glycanes de la paroi bactérienne (LPS, peptidoglycane, …)[62,64,88–90] ou mycobactérienne.[91] En revanche, les exemples de marquage de capsule bactérienne restent très rares, avec seulement deux exemples de marquages chez des espèces non pathogène, *Bacteroides fragilis* et *vulgatus*.[92] La capsule K1 n'a quant à elle jamais été abordée.

La plupart des études de MOE chez les bactéries utilise des rapporteurs libres et non peracétylés (comme c'est fréquemment le cas pour les études sur cellules mammifères) en raison d'une activité estérase aspécifique jugée insuffisante chez les bactéries et notamment chez *E. coli*, ce qui ne permet pas de libérer les rapporteurs libres après

diffusion au travers de la membrane.[68] Il existe néanmoins quelques exemples de marquages bactériens réussis grâce à des rapporteurs peracétylés.[93]

V. OBJECTIFS ET RÉSUMÉ DES TRAVAUX

L'objectif de ces travaux est de développer une méthode de marquage métabolique de la capsule qui camoufle et protège ces bactéries. Les souches *E. coli* exprimant la capsule K1 étant pathogènes, elles nécessitent d'être manipulées au sein d'un laboratoire ayant un niveau de sécurité biologique suffisant (BSL 2). Afin de faciliter l'étude de la capsule K1, un hybride *E. coli* K12 exprimant K1 a été développé par Vimr et Troy.[94] Cet hybride, qui ne possède pas de pathogénicité associée et est donc manipulable dans un simple laboratoire de microbiologie, nous a été fourni par le Pr Sagona,[95] facilitant grandement la mise au point des protocoles pour le marquage.

Pour notre étude nous comptons utiliser des analogues alcynes et azotures de ManNAc et de Neu5Ac pour marquer la capsule K1, cependant, la voie métabolique par laquelle passe la ManNAc implique d'abord sa conversion en de nombreux monosaccharides différents, dont l'UDP-GlcNAc, conduisant potentiellement à une dilution du signal dans toutes les voies en découlant, et notamment celles des autres exopolysaccharides et composés de surface (LPS, peptidoglycane, …) (Figure 16). De plus, il a été montré qu'un rétrocontrôle négatif était exercé par le ManNAc-6P sur la synthèse de la capsule,[96] compliquant potentiellement le marquage de la capsule par incorporation d'analogues de ManNAc.

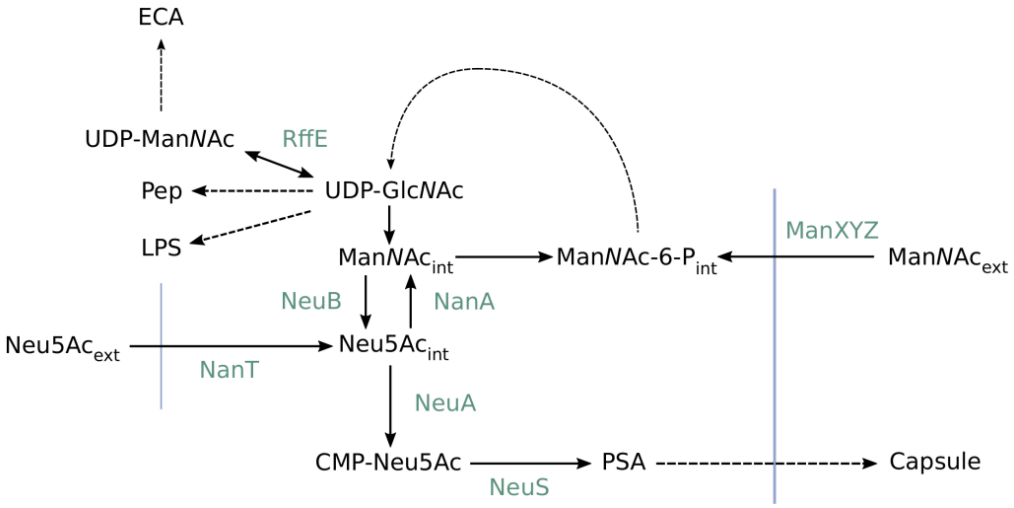


Figure 16 Métabolisme de la biosynthèse de la capsule K1. Le Neu5Ac peut être récupéré dans l'environnement ou synthétisé à partir de ManNAc. L'assemblage des chaînes poly acide sialique a lieu dans le milieu intracellulaire avant export à la surface pour former la capsule.

De nombreux essais furent nécessaires avant de trouver des conditions permettant le marquage de la capsule, or, l'observation de ces marquages par microscopie à fluorescence est assez fastidieuse, impliquant le montage délicat des bactéries sur lames de microscopie, de longues acquisitions d'image, et l'observation de relativement peu de bactéries à la fois. Afin de simplifier cet aspect, nous avons souhaité développer un test permettant une lecture de la fluorescence en microplaque. Ceci, en plus de grandement faciliter et accélérer les expériences, permet de lire la fluorescence émanant d'une suspension de plusieurs milliers de bactéries, fournissant des données bien plus robustes statistiquement comparé à l'observation de champs comportant quelques bactéries à la fois en microscopie. Cette approche est également plus facilement transférable vers la mise au point de tests automatisés de criblage à haut débit, qui viseraient à identifier des principes actifs empêchant la biosynthèse de la capsule plutôt que de tuer la bactérie, afin de déjouer les mécanismes de résistance.

Pour ces travaux, nous avons d'abord mis au point et optimisé les conditions pour l'incorporation métabolique et la ligation bioorthogonale. Bien que nous ayons dans un premier temps envisagé d'effectuer l'étape de ligation par SPAAC avec une sonde cyclooctyne, en raison de la simplicité de mise en œuvre de cette réaction, nous nous sommes heurtés à des problèmes d'aspécificité du signal trop importants. En effet, les

alcynes cycliques utilisables en SPAAC sont relativement encombrants et hydrophobes, et nous soupçonnions que l'aspect intriqué de la membrane bactérienne encapsulée rende difficile le rinçage de fluorophore qui auraient été piégés dans ces structures. D'autres essais en CuAAC avec une sonde de type tétraméthylrhodamine (TAMRA) ont fourni des résultats bien plus concluants, nous avons donc sélectionné cette réaction pour la suite de l'étude.

Nous nous sommes alors penchés sur le choix du milieu de culture. Des essais préliminaires en milieu minimum n'étaient pas concluants, la bactérie ne poussant pas de manière optimale, nous avons donc opté pour le milieu LB, un milieu riche dans lequel nous nous sommes assurés que la bactérie poussait correctement. Nous avons également vérifié que la bactérie exprimait sa capsule après croissance en présence de rapporteurs chimiques (pour s'assurer que nous n'inhibions pas ses voies de biosynthèse avec nos analogues), et que celle-ci était toujours intacte après les conditions de CuAAC impliquant des rinçages assez drastiques (nous redoutions que la capsule, structure décrite comme étant assez fragile,[97] ne soit perdue dans ce processus), en ayant recours à des marquages en immunofluorescence avec un anticorps anti-PSA.[84]

Ces vérifications effectuées, nous avons alors optimisé la concentration en fluorophore et en rapporteur chimique dans le milieu de culture avec une mesure de la fluorescence en lecteur de microplaques (statistiquement plus robuste, et bien plus commode que la microscopie pour la comparaison de dizaines de conditions), aussi bien en termes de rapport signal sur bruit obtenu après CuAAC sur cellules entières, qu'en terme d'impact sur la croissance ou la viabilité des bactéries. Un effet dose important a été observé, avec une concentration optimale en rapporteur chimique au-dessus ou en dessous de laquelle la spécificité du signal obtenu diminue grandement. Bien qu'ils ne semblent pas impacter la croissance cellulaire, les rapporteurs chimiques présentaient en revanche une certaine toxicité pour les bactéries. En comparaison des autres rapporteurs, cette toxicité était moindre pour le ManNAz. Nous avons pu démontrer la spécificité du marquage de la capsule en combinant ces expériences sur cellules entières à des marquages sur des lots

de capsules purifiés, à la fois en plaque, mais aussi après électrophorèse des extraits de capsule sur gel de polyacrylamide.

Enfin, nous avons adapté ces conditions pour une observation en microscopie à champ large, fournissant des informations qualitatives sur le marquage de la capsule. Pour cela, nous avons effectué des marquages sur la souche modèle *E. coli* EV36 mais également sur deux souches K1 sauvages et une autre souche encapsulée mais n'exprimant pas de PSA. Ceci permit de confirmer notre capacité à marquer la capsule K1 de différentes souches, et la spécificité de ce marquage. Nous observons un marquage hétérogène des bactéries en microscopie, semblant montrer qu'elles n'incorporent ou ne produisent pas toute la capsule de manière identique. Un marquage localisé à la périphérie et aux pôles des bactéries semble coïncider avec les conditions d'export de la capsule décrites dans la littérature, bien que cette observation nécessite une technique plus résolutive pour être précisée.

La mise au point et l'application de cette méthode de marquage, permettant d'introduire spécifiquement des dérivés de ManNAc cliquables dans la capsule PSA des souches *E. coli* K1 pathogènes, et de détecter la capsule K1 après ligation d'un fluorophore par CuAAC en microscopie mais aussi en lecteur de microplaques, ont fait l'objet d'une publication en 2022 dans le journal *Royal Society of Chemistry – Chemical Biology*, laquelle est insérée ci-après.[98]

J'ai également effectué une autre série d'expériences en rapport avec le marquage de la capsule K1, mais dont les résultats n'ont pas pu être intégrés de manière cohérente avec le reste des résultats présentés dans l'article. Celle-ci part du constat que la biosynthèse des acides sialiques au sein des cellules mammifères implique une phosphorylation en position 6 du ManNAc, tandis que chez les bactéries la conversion se fait directement depuis le ManNAc, sans phosphorylation intermédiaire. La biosynthèse bactérienne *de novo* du Neu5Ac est intracellulaire, et l'internalisation du ManNAc par le transporteur ManXYZ nécessite quant à elle la phosphorylation de l'hydroxyle en position 6.

Étant curieux de savoir si l'internalisation par ce transporteur était l'unique voie d'entrée du ManNAc dans la bactérie, j'ai synthétisé un analogue de ManNAc portant un groupement azoture en position 6 (empêchant donc toute phosphorylation de cette position), baptisé ManNAc-6-Az. L'idée était que, si un autre mode d'internalisation du ManNAc était existant, les bactéries seraient en mesure de métaboliser ces dérivés, nous permettant d'incorporer des groupements azotures au sein de leurs sialoglycoconjugués, contrairement aux cellules mammifères au sein desquelles le blocage de la position 6 empêcherait toute incorporation. Ceci pourrait potentiellement conduire à une méthode de marquage spécifique des bactéries dans des modèles de co-infection au moyen d'un unique rapporteur chimique.

J'ai effectué la synthèse de ce rapporteur chimique via une stratégie de protection / déprotection ; en introduisant un groupement partant tosyle sur l'hydroxyle en position 6 du ManNAc et en le peracétylant, puis en effectuant une substitution nucléophile du tosyle par un azoture, avant de saponifier les acétates fournissant le ManNAc-6-Az. J'ai alors effectué des expériences d'incorporation métabolique avec ce dérivé, au sein de cellules HeLa et de bactéries *E. coli* exprimant la capsule K1. Comme attendu, aucun marquage n'a été observé dans les cellules mammifères, mais malheureusement les bactéries n'ont pas non plus incorporé cet analogue.

Bien qu'elles n'aient pas conduit à la technique de marquage spécifique de la capsule en modèle de co-infection, ces expériences corroborent l'hypothèse selon laquelle la phosphorylation du ManNAc et son internalisation par ManXYZ constituent l'unique voie d'entrée de ce monosaccharide au sein de ces bactéries.

RSC Chemical Biology

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d2cb00219a

A bioorthogonal chemistry approach to detect the K1 polysialic acid capsule in *Escherichia coli*†

Vincent Rigolot, ‡ Yannick Rossez,*‡ Christophe Biot * and Cédric Lion *

Received 22nd October 2022,
Accepted 21st December 2022

DOI: 10.1039/d2cb00219a

rsc.li/rsc-chembio

Most *Escherichia coli* strains associated with neonatal meningitis express the K1 capsule, a sialic acid polysaccharide that is directly related to their pathogenicity. Metabolic oligosaccharide engineering (MOE) has mostly been developed in eukaryotes, but has also been successfully applied to the study of several oligosaccharides or polysaccharides constitutive of the bacterial cell wall. However, bacterial capsules are seldom targeted despite their important role as virulence factors, and the K1 polysialic acid (PSA) antigen that shields bacteria from the immune system still remains untackled. Herein, we report a fluorescence microplate assay that allows the fast and facile detection of K1 capsules with an approach that combines MOE and bioorthogonal chemistry. We exploit the incorporation of synthetic analogues of *N*-acetylmannosamine or *N*-acetylneuraminic acid, metabolic precursors of PSA, and copper-catalysed azide–alkyne cycloaddition (CuAAC) as the click chemistry reaction to specifically label the modified K1 antigen with a fluorophore. The method was optimized, validated by capsule purification and fluorescence microscopy, and applied to the detection of whole encapsulated bacteria in a miniaturized assay. We observe that analogues of ManNAc are readily incorporated into the capsule while those of Neu5Ac are less efficiently metabolized, which provides useful information regarding the capsule biosynthetic pathways and the promiscuity of the enzymes involved. Moreover, this microplate assay is transferable to screening approaches and may provide a platform to identify novel capsule-targeted antibiotics that would circumvent resistance issues.

Introduction

Although *Escherichia coli* is an important part of the commensal microbiota colonizing the digestive tract of humans and animals, certain strains of this species have developed virulence attributes and cause serious diseases. Among these pathogenic strains, the meningitis/sepsis-associated *E. coli* (MNEC) pathotype is an important cause of neonatal infections. Mostly transmitted from mother to infant during birth, they are known to migrate to the vascular system after mucosal colonization, then penetrate the blood–brain barrier. This generally leads to meningitis or to septicaemia (blood-poisoning), both life-threatening conditions with high mortality and morbidity rates.[1] 80% of MNEC strains express the K1 capsule, a mucous layer of poly-α-2,8-sialic acid (PSA) that surrounds the bacterium thereby shielding its immunogenic proteins from detection by the host.[2,3] Mimicking the human glycan structure found in the neural cell adhesion molecule (NCAM), PSA is not recognized as an external threat and allows the encapsulated bacteria to evade the immune system, and the severity of these infections is directly related to the amount of K1 antigen found at the bacterial surface.[4–7] This capsular polysaccharide is a linear homopolymeric chain constituted of *N*-acetylneuraminic acid (Neu5Ac) monomers belonging to the sialic acid (Sia) family of carbohydrates.[8] Its *de novo* biosynthesis requires the condensation of *N*-acetylmannosamine (ManNAc) and phosphoenolpyruvate to form Neu5Ac, which is then converted to the sugar nucleotide donor cytidine 5′-monophospho *N*-acetylneuraminic acid (CMP-Neu5Ac) and assembled into PSA in the cytosol prior to translocation to the cell surface[9] (Fig. 1). *E. coli* serotype K1, as well as other bacteria that produce polysialic acid capsules such as *Neisseria meningitidis*, *Pasteurella haemolytica* A2 or *Moraxella nonliquefaciens*, can also scavenge free Neu5Ac from the host.[10]

Monosaccharide analogues modified with an additional chemical moiety can be used as molecular tools to engineer sialylated glycoconjugates in metabolic oligosaccharide engineering (MOE) approaches. Reutter and co-workers first pioneered MOE, with synthetic ManNAc analogues bearing an elongated *N*-acyl side chain that were successfully metabolized into cell surface sialoglycoproteins.[11] Modern labelling methodologies combine MOE and click chemistry or bioorthogonal

Univ. Lille, CNRS, UMR 8576 – UGSF – Unité de Glycobiologie Structurale et Fonctionnelle, Lille, France. E-mail: cedric.lion@univ-lille.fr; christophe.biot@univ-lille.fr, yannick.rossez@univ-lille.fr

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d2cb00219a

‡ These authors contributed equally.

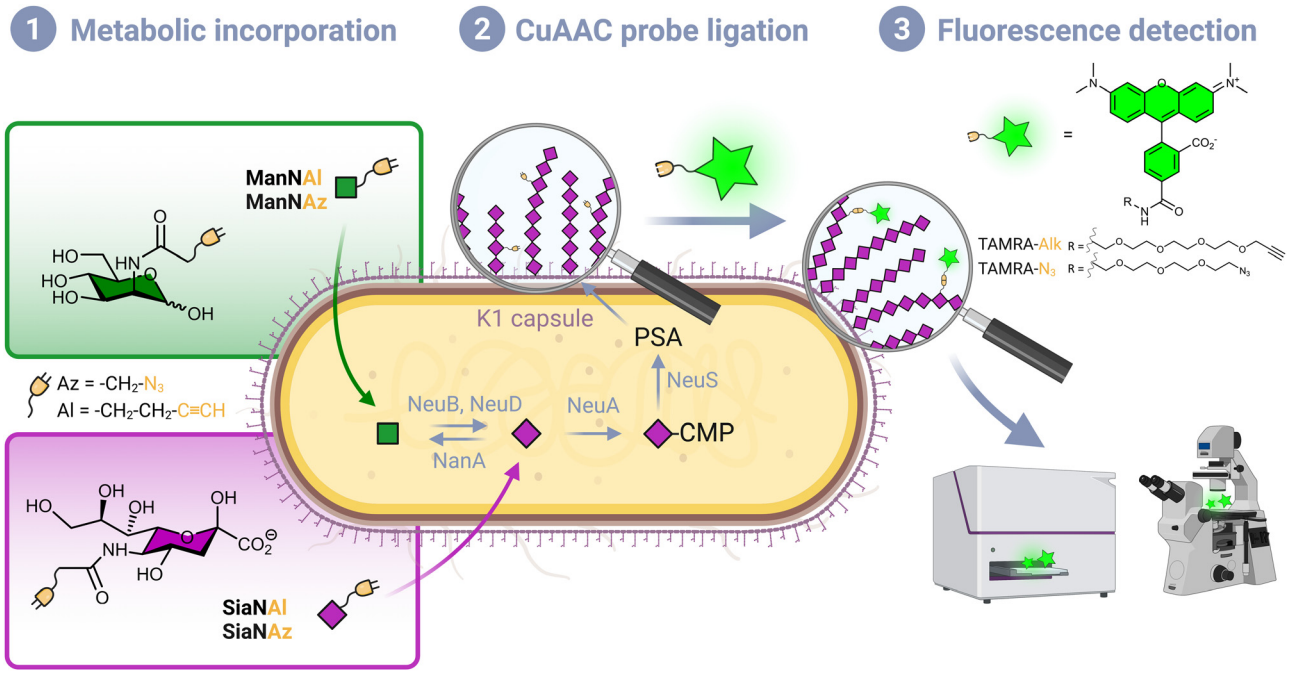


Fig. 1 Labelling strategy workflow. (1) Chemical reporters bearing an azide or alkyne moiety are incorporated in the capsule of living *E. coli* EV36 by hi-jacking the polysialic acid metabolic pathway. After being transported from the extracellular milieu to the cytosol, ManNAc analogues enter the *de novo* biosynthesis of sialic acid (Neu5Ac), while Neu5Ac (2) the incorporated analogues are specifically labelled with a fluorescent dye *in vivo* by a CuAAC reaction. (3) The labelled bacteria are analysed with the desired technique such as fluorescence microscopy or fluorescence plate reader.

chemistry, an ensemble of reactions originally pioneered by the groups of Sharpless, Meldal and Bertozzi, who were recently awarded the Nobel prize for this achievement,[12–14] and further developed by them and other groups in the past twenty years. A sugar reporter capable of entering the metabolic pathway under scrutiny and equipped with a reactive handle is first introduced in a living organism of interest, followed by the bioorthogonal ligation of an appropriately functionalized probe, most often a fluorescent dye for optical bioimaging.[15] These methods rely on biocompatible click reactions that involve π_4s + π_2s pericyclic processes, namely the copper-catalysed azide–alkyne (3+2) cycloaddition, the strain-promoted azide–alkyne (3+2) cycloaddition and the inverse electronic demand Diels–Alder reaction (CuAAC, SPAAC and IEDDA, respectively). Indeed, such mechanisms are mostly absent from living systems, thus greatly minimizing interference with biological processes.[16]

Although many bioorthogonal methods have been used in prokaryotic cells to track proteins with non-canonical amino acids,[17] efforts to apply MOE to detect glycans in prokaryotes remain rather meager by comparison, as these strategies were developed in human models first and foremost. Some glycoengineering methods were reported in various species to label bacterial glycans, which were extensively reviewed by Banahene *et al.*[18] They aim at detecting components of bacterial cell walls, such as peptidoglycans with *N*-acetylmuramic acid derivatives[19,20] or lipopolysaccharides with, for example, azide-modified analogues of *N*-acetylglucosamine,[21] 3-deoxy-D-manno-octulosonic acid[22] or legionaminic acid precursors.[23] However, very few bacterial species have been specifically labelled on their capsular polysaccharides by bioorthogonal chemistry. We could only identify two reports, in which capsules of commensal bacteria of the *Bacteroides* genus have been effectively tagged by MOE with modified *N*-acetylgalactosamine reporters.[24,25] The authors visualized the capsular component of prelabelled bacteria in live mice intestines, opening an avenue to better understand host–pathogen interactions and the evolution of related diseases.

To the best of our knowledge, capsular polysaccharides constituted of sialic acids have thus never been investigated by MOE in combination with click chemistry, despite the importance of these structures as virulence factors. Herein, we report a bioorthogonal labelling method to detect the polysialic acid capsule using alkyne- and azide-modified ManNAc and Neu5Ac reporters in *E. coli* K1 strains. Although this reporter toolbox is commonly used in mammalian systems to label *N*-glycoproteins,[26,27] it is the first time that these reporters are used successfully in *E. coli*. We aimed to take advantage of the fact that, unlike non-pathogenic strains, K1 strains are capable of synthesizing Neu5Ac *de novo*, and developed a test that allows facile and specific detection of the PSA capsule. The method was miniaturized into a microplate assay for transferability to screening approaches, which could serve as an interesting platform to decipher K1 pathways and identify external factors that influence PSA biosynthesis. Such assays could help in the discovery of new capsule-targeted classes of antibiotic drugs that decrease the virulence of a pathogenic bacterium while circumventing known resistance mechanisms.

Results and discussion

Chemical reporters

Per-*O*-acetylated monosaccharides are commonly used for MOE in mammalian models. However, per-*O*-acetylation can cause issues as it may generate non-specific signal and false positives.[28,29] It may also be conducive to acidification of the intracellular milieu, metabolic perturbation due to partially acetylated forms, or cytotoxicity.[16] In addition, it has been suggested that low levels of non-specific esterase activity prevent the efficient release of the free sugar form thus thwarting the use of such reporters in some prokaryotic species including *E. coli*.[23] Conversely, this species is capable of active transport mediated ManNAc uptake *via* the ManXYZ transporter and can also scavenge exogenous sialic acid from the host *via* the sialic acid transporter NanT. We therefore decided to use unprotected reporters and synthesized *N*-(2-azidoacetyl)-D-mannosamine (ManNAz), *N*-4-pentynoyl-D-mannosamine (ManNAl), *N*-(2-azidoacetyl)-neuraminic acid (SiaNAz) and *N*-4-pentynoylneuraminic acid (SiaNAl) (see experimental section). CuAAC was chosen as the bioorthogonal reaction, owing to its fast kinetics and ease of use. In addition, the azide and terminal alkyne tags are easily interchanged thus allowing comparison of the chemical handle's impact.

CuAAC whole cell labelling of K1 expressing *E. coli* EV36

We first assessed whether these reporters could be taken up and metabolically incorporated into living *E. coli* EV36, a strain containing all fourteen genes of the *kps* PSA biosynthetic cluster that produces the K1 capsule.[30–32] EV36 cultures were incubated overnight in lysogeny broth (LB) medium complemented with the modified monosaccharides, prior to CuAAC labelling with an appropriate clickable tetramethylrhodamine dye bearing an alkyne or azide reactive moiety (TAMRA-N_3 for alkynyl reporters and TAMRA-Alk for azido reporters). The same fluorochrome was used in both instances in order to allow intensity-based comparisons. The reactive function is not π-conjugated to the rhodamine core owing to the presence of an oligoPEG spacer arm, and both dyes present the same photophysical characteristics (Φ = 0.1; ε = 92.000 M^{-1} cm^{-1}, superimposable spectra). The assay was implemented using a fluorescence microplate reader to enable facile testing of multiple conditions. The bioorthogonal ligation was first optimized to obtain a robust labelling and satisfying signal-to-noise ratio in EV36 whole cells. Various parameters were screened, with TAMRA dye concentrations ranging from 62.5 nM to 25 μM in EV36 cultures incubated with reporters at a fixed 1 mM concentration, giving the best signal to noise ratio at 0.25 μM of fluorophore for 45 minutes (Fig. 2A), and then with chemical reporter concentrations ranging from 100 μM to 2 mM at a fixed 0.25 μM TAMRA concentration, resulting in an optimal chemical reporter concentration found at 600 μM (Fig. 2B). In contrast to what has been reported in activity-based protein profiling studies in mammalian cell lysates using dye concentrations 200 to 400 fold higher,[33,34] we observed a higher level of background using TAMRA-N_3 at 0.5 μM. This could be attributed to non-covalent interactions of azide dipoles at the bacterial cell surface at higher concentrations, or to the fact that alkyne-dyes exert less non-specificity in the nanomolar range.

Next, we sought to perform statistically robust tests to determine the significance of the difference in labelling after the incorporation of the chemical reporters at 600 μM and 1 mM. In order to allow transferability of the protocol to prospective screening approaches, we developed this assay in 96-well plates, using an intensity-based readout with a fluorescence microplate reader. For all four compounds, we observed significant labelling at 600 μM but not at 1 mM, except for ManNAz being the only analogue where the labelling was significant in both cases (Fig. 2C). This suggested that ManNAl, SiaNAl and SiaNAz might exert a negative effect on the physiological state or viability of bacteria at higher concentrations.

Experiments in which the bacteria were labelled with SiaNAz and SiaNAl generally exhibited a lower fluorescence intensity when compared to ManNAz and ManNAl, which came as a surprise. Indeed, these microbial cells are generally thought to be able to scavenge, process and install Sia derivatives but to metabolize ManNAc very slowly,[35,36] although this is highly species dependent. ManNAz led to the most efficient labelling. The noticeable difference in intensity between ManNAz and ManNAl tagging indicates that the azide handle may be more compatible with the metabolic network overall, or that the alkyne moiety might induce toxicity during the metabolic incorporation step. Distinct uptake and/or metabolisms should account for differences in labelling efficiency between ManNAc or Sia derivatives bearing the same chemical tag (*e.g.*, ManNAz and SiaNAz).[36]

Reporter accumulation in the capsule

We pursued by investigating whether the observed labelling originates from reporters incorporated into the K1 capsule. To show this, *E. coli* EV36 was grown in LB medium complemented with ManNAz or ManNAc, then the K1 capsular polysaccharides were extracted and purified according to a previously described procedure.[37] A polyacrylamide gel electrophoresis of the extracted K1 capsule was performed, followed by alcian blue staining (a dye that specifically reveals acidic polysaccharides[38]) to confirm the presence and purity of PSA (Fig. 3A). To evaluate the incorporation of the chemical reporters in the isolated capsule, the extracted acidic polysaccharides were then labelled in solution by CuAAC and fluorescence intensity was measured as described hereabove. Specific labelling was observed with both ManNAc derivatives ManNAl and ManNAz, but not with the Neu5Ac analogues SiaNAl and SiaNAz (Fig. 3B). Therefore, this indicates that chemical reporters ManNAz and ManNAl enter the bacterial cell (possibly through the ManXYZ-encoded transporter) and are converted to SiaNAz and SiaNAl, respectively, in the bacterial cytoplasmic compartment.[39] The derived tagged Sia is subsequently activated as CMP-Sia, polymerized and exported at the cell surface by various enzymes and transporters.[9] *E. coli* has been described to convert ManNAc to ManNAc-6-P during uptake

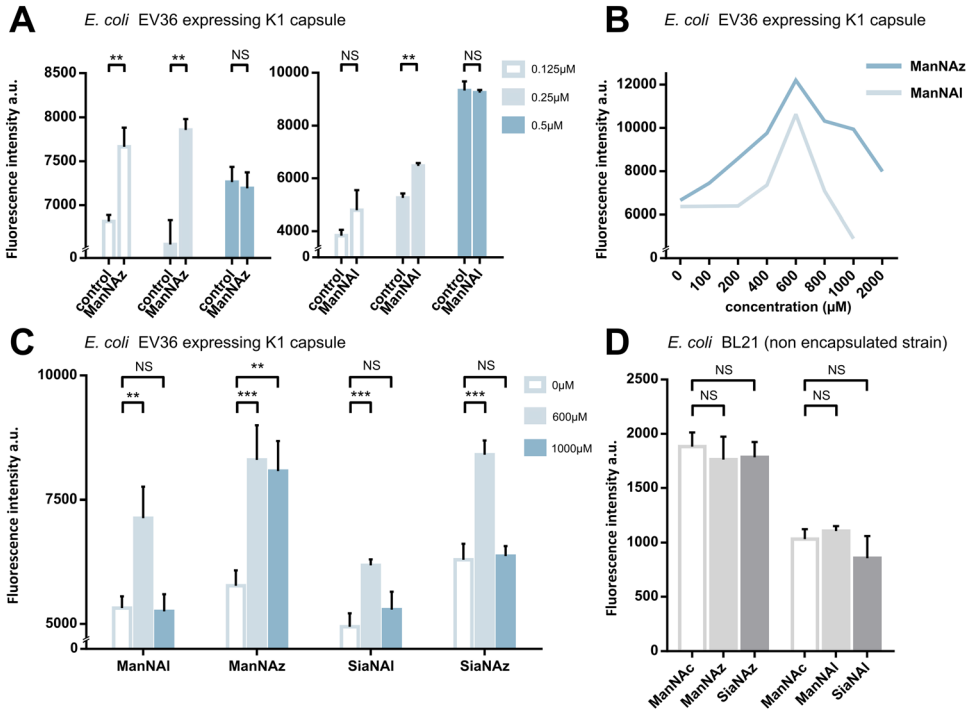


Fig. 2 Whole cell suspension fluorescence after MOE in *E. coli* EV36. (A) Optimization of the fluorophore concentration. *E. coli* EV36 were grown with 1 mM ManNAz or 1 mM ManNAl, followed by CuAAC ligation of TAMRA-Alk or TAMRA-N_3, respectively, at concentrations of 0.125, 0.25 μM or 0.5 μM. Negative control samples were grown with 1 mM ManNAc (Ctrl) and reacted in the same conditions. (B) Optimization of the chemical reporter concentration. *E. coli* EV36 were grown with ManNAl or ManNAz at concentrations ranging from 0 to 2000 μM. The CuAAC ligation was then performed using the relevant fluorophore (TAMRA-N_3 or TAMRA-Alk, respectively) at a concentration of 0.25 μM. (C) Miniaturized assay on *E. coli* EV36 in 96-well plates. CuAAC labelling was performed using the identified optimal conditions, namely 0.25 μM dye for 45 minutes in presence of the catalytic buffer. (D) Negative control experiments carried out on *E. coli* BL21, a non-encapsulated strain. 3 biological replicates with a minimum of 3 technical replicates were measured. Results are presented as mean + SD. Unpaired *t*-tests were performed to get statistical significance of difference observed compared to control condition (NS = non significant ($P > 0.05$); ** = $P < 0.01$; *** = $P < 0.001$).

to redirect the product on the degradation pathway.[39,40] Our results suggest that *E. coli* can also incorporate ManNAc into the K1 capsule biosynthesis *via* an unidentified pathway able to either transport ManNAc *via* the manXYZ phosphotransferase system, producing intracellular ManNAc-6-P which is then converted to ManNAc by an unknown phosphatase, or to transport ManNAc without ManNAc-6-P conversion. The former hypothesis carries a much stronger weight as the presence of this enzyme was previously suggested in *Neisseria meningitidis*, which also expresses K1 capsules, though no proper characterization has yet been reported.[41] Furthermore, Neu5Ac is transported by the NanT sialate permease to the interior of the bacterial cells and is known to be directly degraded.[30,35,40,42,43] Our results are in good agreement with this observation. Neu5Ac is indeed transported inside the cell as demonstrated by the whole cell fluorescence observed (Fig. 2), but the monosaccharide is not directed to the PSA biosynthesis as suggested by the lack of signal detected on the purified capsule (Fig. 3B).

To check whether constitutive biomolecules other than the capsule were labelled by our strategy, we carried out the same experiments on whole *E. coli* BL21 cells, a B strain that does not express capsules. All four reporters were tested, and no labelling was observed on this strain (Fig. 2D), thereby confirming the specific incorporation of the reporters into the K1 PSA capsule in the EV36 bacterium.

Effect of unnatural monosaccharides on the K1 capsule biosynthesis

To ensure that the capsule was correctly expressed in the different labelling conditions, we performed immunofluorescence labelling of whole *E. coli* EV36 cells with an anti-K1 antibody followed by observation in fluorescence microscopy. A neat peripheral labelling pattern was observed, confirming the presence and integrity of capsular PSA. The experiments were carried out after growth in the presence of a ManNAc or Neu5Ac analogue (ManNAz and SiaNAl, respectively), or after the thorough rinsing and shaking conditions required during the CuAAC labelling protocol (“Shaken” condition) (Fig. 4). These results demonstrate a robust expression of the K1 capsule, without side effects of the unnatural monosaccharides on the K1 pathways.

Impact of unnatural monosaccharides on the growth and physiological state of the bacteria.

To further investigate the results obtained in Fig. 2C, *E. coli* EV36 growth and viability were evaluated for ManNAz, ManNAl, SiaNAz, or SiaNAl at 1 mM and 0.6 mM in comparison to

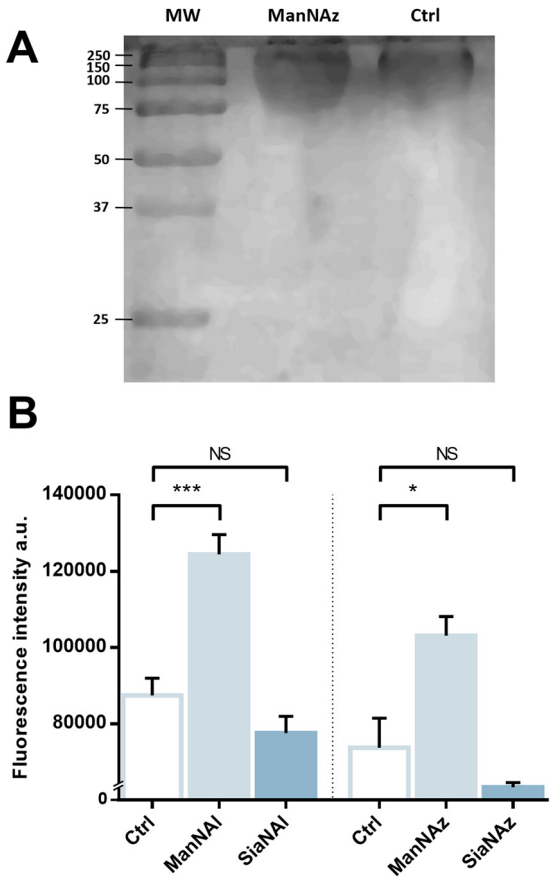


Fig. 3 Capsular extraction and fluorescence quantification – (A) The capsular polysaccharides were purified after metabolic incorporation and analysed by PAGE with alcian blue staining. Molecular weight (MW) standard is indicated in kDa. (B) Capsular extracts obtained after metabolic incorporation of chemical reporters were labelled by CuAAC with 5 μM TAMRA-N_3 or TAMRA-Alk. 3 biological replicates with a minimum of 3 technical replicates were measured. Results are presented as mean +/− standard deviation. Unpaired *t*-tests were performed to get statistical significance of the difference measured compared to the control condition (Ctrl). NS = non significant ($P > 0.05$); * = $P < 0.05$; *** = $P < 0.001$.

ManNAc and LB alone. To evaluate the impact of the chemical reporters on the bacterial growth we measured the optical density at 600 nm (OD_{600}) of LB suspensions at regular intervals over the course of several hours. No difference was observed between the various conditions, illustrating that the reporters do not impair cell growth (Fig. 5A). To evaluate the impact on long-term cell viability, serial dilutions of the suspensions were carried out after growth that were plated for viable counts. All chemical reporters except ManNAz induced rather significant toxicity on the long term when compared to ManNAc. Interestingly, when ManNAc was added to the medium the cell viability was also significantly lower by comparison to LB medium alone (Fig. 5B). ManNAz and ManNAc equally affect bacterial long-term viability to a mild extent. On the contrary, ManNAl drastically reduces the viability at both concentrations, revealing an increased toxicity linked to the appended pentynoyl moiety. This toxicity leads to decreased levels of specific PSA labelling (Fig. 2), and might enhance non-specific binding. Dead bacteria have been recently described to form nanotubes by membrane extrusion or vesicles. These structures may impact background levels.[44] Such an effect due to the nature of the bioorthogonal handle has been evidenced before in human Jurkat cells, where azidofucose derivatives proved to be more efficiently incorporated but more toxic than their alkynyl counterparts.[45] In the case of K1-expressing *E. coli*, it is the alkyne tag that decreases the viability of EV36 cells, which might be explained by the fact that a lower level of incorporation in capsular PSA leads to increased catabolism and accumulation of alkyne by-products interfering with other metabolic networks in the cell. Nevertheless, ManNAl is efficiently incorporated into the capsule (Fig. 3B).

It could thus be hypothesized that the pentynoyl sidechain is less compatible with the *E. coli* NeuB-encoded sialic acid synthase or NeuA-encoded CMP-sialic acid synthase than its 2-azido counterpart in ManNAz. This result calls attention to the use of unnatural monosaccharides to study bacterial physiology. To avoid any misinterpretation, viability assays should always be performed when testing MOE strategies on a new bacterial strain.

MOE microscopy

Given that ManNAz enables specific labelling of the capsule with no significant toxicity nor impact on growth, we chose to use this analogue for MOE followed by CuAAC labelling of the capsule for observation by fluorescence microscopy. After ManNAz incorporation, cells were treated with a CuAAC buffer whose composition was adapted from a previously reported procedure.[22] A bright, heterogeneous and specific labelling was observed after ManNAz incorporation compared to ManNAc as negative control (Fig. 6). All cells appear to incorporate the reporter. In most cells, the labelling is of rather low intensity (although significant when compared to the control) but on some others, a bright peripheric fluorescent signal was visible (Fig. 6B and C), typical of capsule localisation. Additionally, a number of cells exert a labelling pattern at the poles of the bacterium (Fig. 6D), and most cells that do appear to be undergoing or to have just undergone cell division. This could be interpreted as a physiological state in which the older labelled PSA capsule has been relocated at the polar regions while new unlabelled capsule is being produced at the septum region during binary fission. Like most capsular polysaccharides in Gram-negative bacteria, PSA assembly and export through the periplasm is controlled by a multiprotein complex. The K1 antigen belongs to group 2, which identifies strains for which this export involves an ABC-transporter.[46] Our results might indicate that this translocation to the outer membrane is spatially regulated at the equatorial region during cell division, similarly to what has been proposed for Gram-positive *Streptococcus pneumoniae*, whose capsule export is mediated by the Wzx–Wzy system.[47] Indeed, Henriques *et al.* suggested that this spatial control ensures the coordination of capsule and cell wall biogenesis at the division site, leaving the bacteria unexposed to the host's immune system throughout. The heterogeneous labelling points

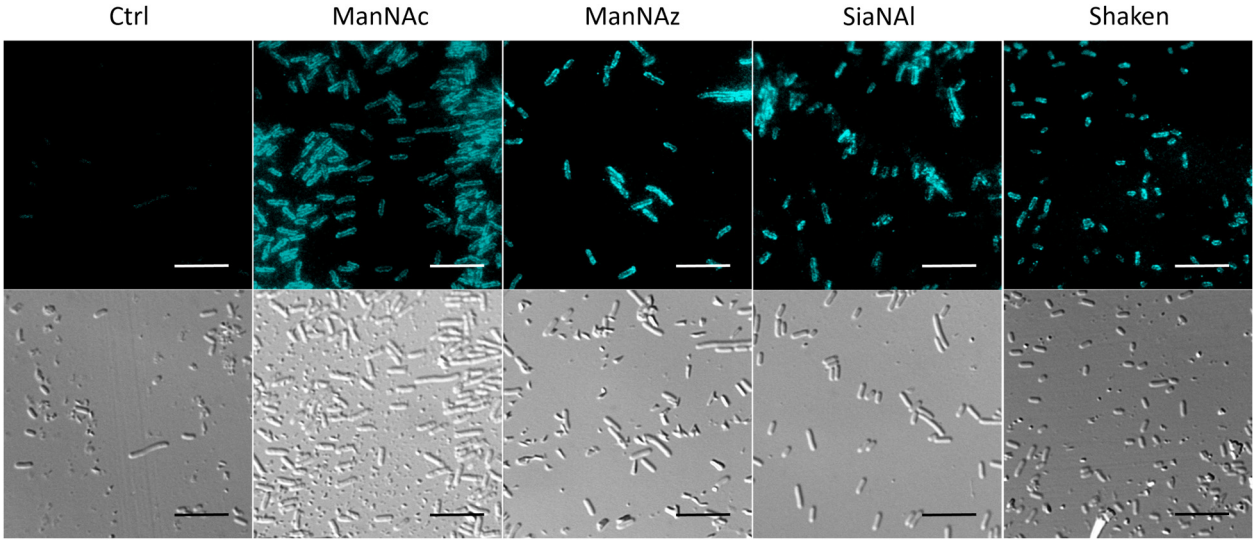


Fig. 4 Immunofluorescence microscopy of the K1 capsule from *E. coli* EV36 with antibodies. *E. coli* EV36 bacteria were grown with either ManNAc, ManNAz, or SiaNAl at 1 mM before being incubated with anti-K1 rabbit IgG antibody ENZ-ABS559-0100 recognizing homopolymers of α2,8-linked sialic acid of *E. coli* K1 (1/100, 45 min), rinsed with 200 µl PBS and treated with Alexa Fluor 488 conjugated anti-rabbit secondary antibody (1/250, 30 min). *Top*: fluorescence channel. *Bottom*: Differential interference contrast (DIC) channel. Scale bar = 10 µm. For the shaken condition, cells were submitted to the same shaking and rinsing conditions needed for CuAAC to confirm that the K1 capsule is still intact and not washed away after these treatments. For the control condition (Ctrl), the secondary antibody was omitted to get the autofluorescence background.

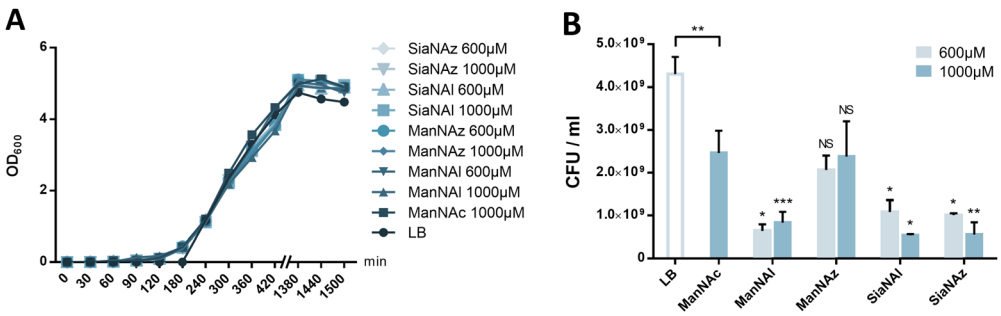


Fig. 5 Growth and viability of *E. coli* EV36 in the presence of unnatural monosaccharides. (A) Growth assay of *E. coli* EV36 cultures with different chemical reporters at 0.6 mM and 1 mM compared to the control conditions (LB). (B) Viability assay of *E. coli* EV36 cultures after growth with different chemical reporters at 0.6 mM and 1 mM. See experimental part for details. 3 biological replicates with a minimum of 3 technical replicates were measured. Results are presented as mean + SD. Unpaired *t*-tests were performed to get statistical significance of difference observed compared to ManNAc 1 mM condition (above error bars), as well as between ManNAc 1 mM and LB (above brackets) (NS = non significant ($P > 0.05$); * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$).

out the incorporation of ManNAz with a potential dependence on the physiological state of the bacteria during their growth. Similarly, the bacteria need to adjust their metabolism to ManNAc and ManNAz import and successfully export the polysialic acid at the cell surface, which may also result in heterogeneity. Finer localisation may be obtained with super-resolution microscopy, to refine the visualization of capsule export zones at the poles of the cells.[48]

In order to further confirm the specificity of our method for the K1 capsule and ascertain its applicability to various pathogenic bacteria, we then tested three other wild-type *E. coli* strains (Fig. 7). In addition to *E. coli* BL21,[49] a strain devoid of capsule mentioned hereabove as negative control (Fig. 2D), we also evaluated the ability of *E. coli* Nissle 1917 bacteria to metabolize ManNAz. The Nissle 1917 strain expresses another capsular acidic polysaccharide, namely the K5 antigen heparosan constituted of a disaccharide repeating unit that contains glucuronic acid and *N*-acetylglucosamine.[8,50] As expected, this strain was unable to incorporate ManNAz, further confirming the specificity of this assay for the K1 PSA capsule. Moreover, the two *E. coli* K1 pathogenic strains K-235 and U5/41, which produce significantly higher levels of PSA than the EV36 model, exhibited strong fluorescence (Fig. 7). This indicates an increased level of incorporation of SiaNAz units into PSA derived from the metabolization of ManNAz when compared to experiments on EV36. The heterogeneity was lesser in the K-235 and U5/41 wild-types and the pattern observed was consistent, with bacteria marked at the periphery typical of

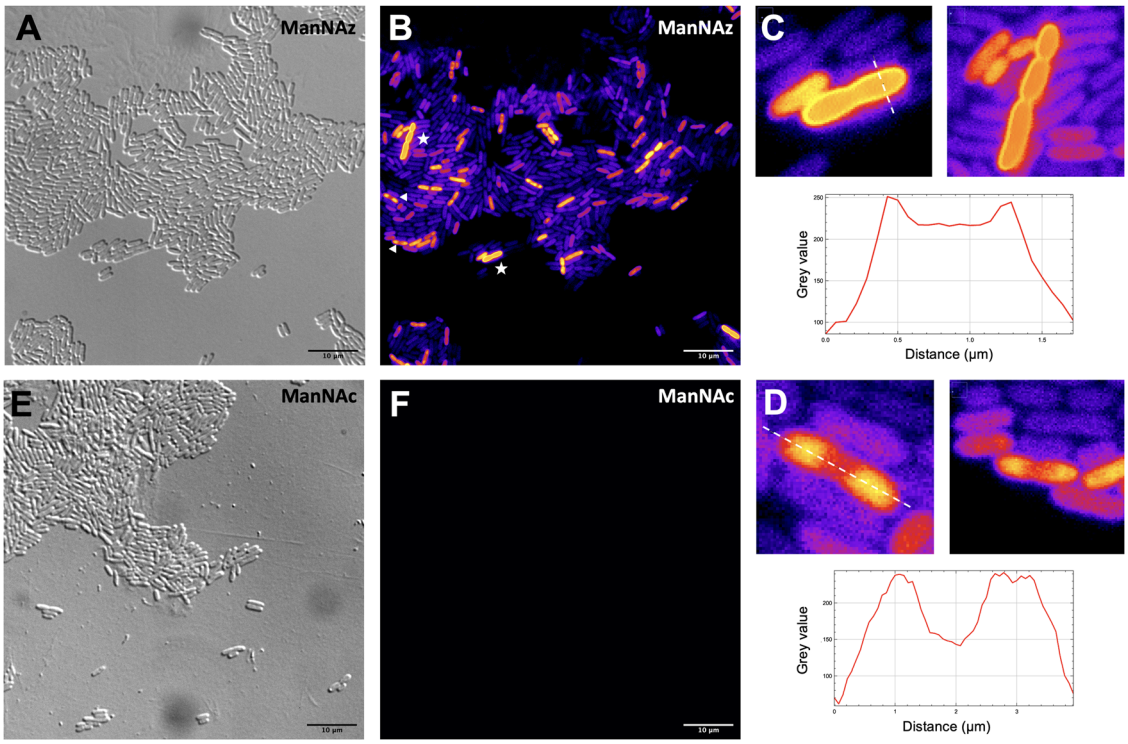


Fig. 6 Fluorescence microscopy of *E. coli* EV36 after incorporation of ManNAz at 600 μM. CuAAC was performed with the TAMRA-Alk dye at 100 mM. Brightfield channel (A) depicted in greyscale, fluorescence channel (B and F) depicted in "Fire" LUT colorscale. (C) (top) Magnified zooms on cells showing a pericellular labelling (highlighted by white stars in insert B) and (bottom) intensity profile along the transversal dashed line; (D) (top) Magnified zooms on cells showing a polar labelling (highlighted by white triangles in insert B) and (bottom) intensity profile along the longitudinal dashed line. (E and F) Negative control incubated with ManNAc and reacted in the same CuAAC conditions (TAMRA-Alk, 100 mM). Experiments were carried out as 3 biological replicates. Scale bars= 10 μm.

capsule staining and other bacteria showing localization at the poles. In contrast, the heterogenous incorporation observed in the EV36 construct might be attributed to a less efficient metabolic channeling process, induced by the expression of exogenous genes from K1 strains in a K-12 background (*e.g.*, the manXYZ operon). The presence of PSA in the K-235 and U5/41 strains and its absence in the Nissle 1917 strain were further confirmed by immunofluorescence using a fluorescent anti-K1 antibody (Fig. S1, ESI†). Thus, our method can be used to detect K1 antigen production in pathogenic strains, which could help decipher the dynamics of capsule expression and the factors that regulate it in future studies. Furthermore, it provides a platform for screening new types of antibiotics targeting K1 capsule metabolism.

Experimental

Chemical synthesis of unnatural monosaccharides

ManNAl (*N*-4-pentynoyl-D-mannosamine): 4-pent-ynoïc acid (502 mg, 5.1 mmol, 1 eq.) was solved into CH_2Cl_2 (25 ml). *N*-Hydroxysuccinimide (674 mg, 5.8 mmol, 1.14 eq.) and EDC or 1-ethyl (3-dimethylaminopropyl)carbodiimide (1.96 g, 10.2 mmol, 2 eq.) were slowly added. The mixture was stirred for 3 h 30 at room temperature. The end of reaction was confirmed by TLC (cyclohexane/AcOEt; 70 : 30). The solution was then washed four times with KHSO4 (aqueous solution 2.8%). The recombinated organic layers were dried over sodium sulfate, filtered, and then concentrated under reduced pressure to give the succinimidyl ester as a white solid (875.3 mg, 4.48 mmol, 88%). The product was used in the next step without further purification. The coupling reaction with Mannosamine hydrochloride was conducted under nitrogen atmosphere. Hydrochloride D-Mannosamine (759 mg, 3.52 mmol, 1 eq.) and the succinimidyl ester from previous step (688 mg, 3.52 mmol, 1 eq.) were solved into DMF (25 ml). Triethylamine (1.4 ml, 10.77 mmol, 3 eq.) was slowly added. The mixture was stirred overnight. DMF was then removed under reduced pressure and the crude product was purified by flash chromatography (silica column, 40 g, 30 μm) with an elution by CH_2Cl_2/EtOAc/MeOH; 45 : 45 : 10). Solvents were then eliminated to yield ManNAl as a white solid (830 mg, 3.34 mmol, 95%). Mixture of anomers ($\alpha/\beta \approx$ 60%/40%).

α ^{1}H NMR (300 MHz, D_2O): δ = 5.00 (d, J = 1.4, 1H, H_1), 4.23 (dd, J = 4.7, 1.4, 1H, H_2), 3.93 (dd, J = 9.8, 4.7, 1H, H_3), 3.80 – 3.62 (m, 3H, H_5 & H_6), 3.50 (t, J = 9.6, 1H, H_4), 2.58 – 2.32 (m, 4H, H_8 & H_9), 2.26 (t, J = 2.3, 1H, H_{11}).

^{13}C NMR (75 MHz, D_2O): δ = 175.26 (C7), 93.17 (C1), 83.43 (C10), 71.94 (C5), 70.14 (C11), 68.73 (C3), 66.72 (C4), 60.36 (C6), 53.15 (C2), 34.10 (C8), 14.40 (C9).

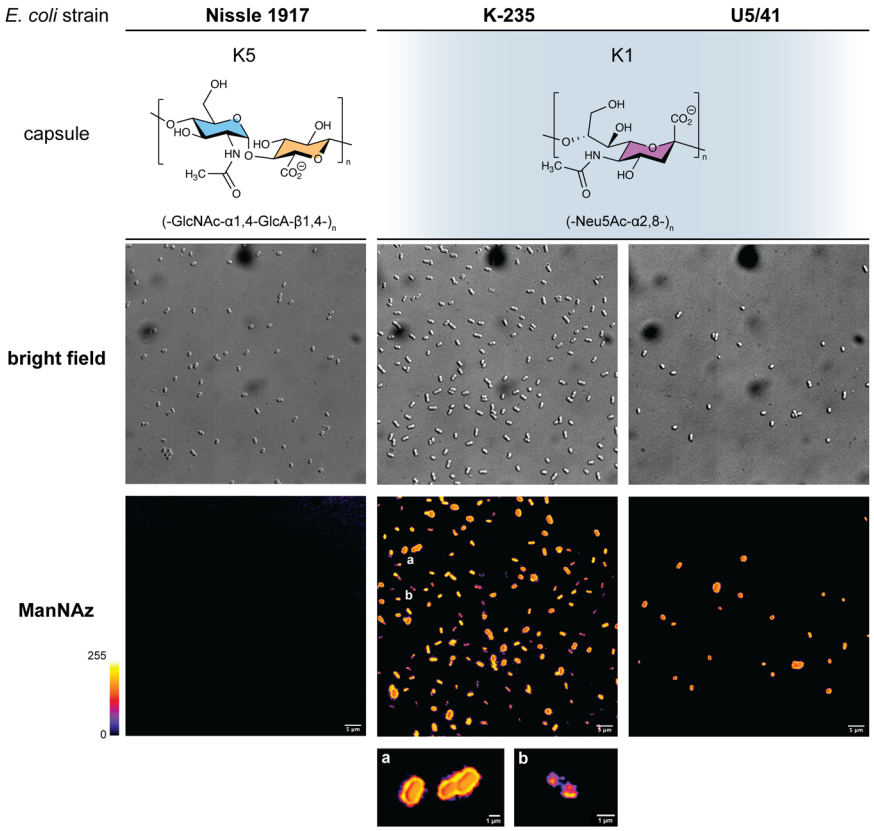


Fig. 7 Fluorescence microscopy of *E. coli* strains Nissle 1917 expressing the K5 capsular polysaccharide (heparosan), and K-235 and U5/41 expressing the K1 capsular polysaccharide (PSA), grown with ManNAz at 600 μM overnight then labelled with TAMRA-Alk – (top) brightfield channel; (bottom) fluorescence channel in Fire colorscale LUT. (a) zoom on bacteria exhibiting a peripheral pattern typical of capsule expression, (b) zoom on bacterium with labelling in the polar regions during cell division. Experiments were carried out as 3 replicates. Scale bar = 5 μm.

β ^{1}H NMR (300 MHz, D_2O): δ = 4.90 (d, J = 1.6, 1H, H_1), 4.35 (dd, J = 4.4, 1.4, 1H, H_2), 3.82 – 3.61 (m, 3H, H_3 & H_6), 3.40 (t, J = 9.8, 1H, H_4), 3.29 (ddd, J = 9.9, 4.8, 2.3, 1H, H_5), 2.53 – 2.31 (m, 4H, H_7 & H_8), 2.26 (t, J = 2.3, 1H, H_{11}).

^{13}C NMR (75 MHz, D_2O): δ = 175.98 (C7), 92.84 (C1), 83.85 (C10), 76.28 (C5), 71.96 (C3), 69.99 (C11), 66.45 (C4), 60.36 (C6), 53.98 (C2), 34.25 (C8), 14.26 (C9).

SiaNAl (*N*-4-pentynoylneuraminic acid): ManNAl, (60 mg, 0.23 mmol, 1 eq.), sodium pyruvate (46 mg, 0.345 mmol, 2 eq.) and Neuraminic5Ac aldolase (Sigma Aldrich EC 4.1.3.3) from E. Coli K12 (5 units dissolved in 0.1 ml) were added into 600 μL of phosphate buffer (KH_2PO_4, 10.3 mM, K_2HPO_4, 52.8 mM and $MgCl_2$ 20 mM, pH – 7.5) into a reaction tube which was incubated at 37.5 °C with moderate shaking (140 rpm) for 18 hours. The reaction completion was monitored by TLC (propan-1-ol/25% Ammonia/H_2O; 6 : 1: 2.5). Reaction was quenched by the addition of 30 ml of water. The product was purified on an anion exchange resin (Bio-Rad AG1X2) activated with NH_4HCO_3 0.1 M. The sample was loaded on the column which was washed with H2O (5 CV). Elution was performed with aqueous NH_4HCO_3 (0.05 M, 5 CV) and then NH_4HCO_3 (0.2 M, 5 CV). Fractions were detected by TLC (same elution as for the reaction monitoring) combined and freeze dried. After being solved into H_2O (1 ml), sample was passed through a gel filtration column (P2) and then freeze dried again. SiaNAl was obtained as white powder (73 mg, 0.21 mmol, 91%).

^{1}H NMR (600 MHz, D_2O): δ = 3.96 – 3.86 (m, 2H, H_4 & H_9), 3.83 (t, J = 10.1, 1H, H_5), 3.72 (dd, J = 11.9, 2.7, 1H, H_9), 3.64 (ddd, J = 9.2, 6.7, 2.7, 1H, H_8), 3.51 (d, J = 9.2, 1H, H_7), 3.49 – 3.43 (m, 1H, $H_{9'}$), 2.47 – 2.34 (m, 4H, H_{11} & H_{12}), 2.28 (d, J = 2.0, 1H, H_{13}), 2.10 (dd, J = 13.0, 4.9, 1H, H_{3eq}), 1.70 (dd, J = 12.7, 11.7, 1H, H_{3ax}).

^{13}C NMR (75 MHz, D_2O): δ 175.43 (C1), 172.91 (C10), 95.16 (C2), 83.54 (C13), 70.40 (C14), 70.37 (C8), 70.23 (C6), 68.29 (C7), 66.45 (C4), 63.21 (C9), 52.04 (C5), 38.79 (C3), 34.67 (C11), 14.53 (C12).

m/*z*: Calculated: [M]+ = 346.312; measured: 346.017.

ManNAz (*N*-azidoacetyl-D-mannosamine): synthesis, 2-azido-acetic acid (117 mg, 1.16 mmol, 1 eq.) was dissolved along with DIC (*N*,*N*′-diisopropylcarbodiimide, 175 mg, 1.392 mmol, 1.2 eq.), HOBt (hydroxybenzotriazole, 195 mg, 1.276 mmol, 1.1 eq.) and DIPEA (*N*,*N*-diisopropylethylamine, 141 mg, 1.392 mmol, 1.2 eq.) in DMF (dimethylformamide, 20 ml), D-mannosamine hydrochloride (250 mg, 1.16 mmol, 1 eq.) was added and the reaction was stirred at room temperature for 19 h under argon. The total consumption of mannosamine was assessed by silica thin layer chromatography (TLC)

(CH_2Cl_2/MeOH 9 : 1 v/v) before solvent removal under reduced pressure. The reaction crude was purified by silica flash column chromatography (50 μm, 40 g, dry load, CH_2Cl_2/MeOH 95 : 5 v/v), and fractions containing the product were gathered and concentrated under reduced pressure, affording the ManNAz as a white powder (175 mg, 0.67 mmol, 57%).

^{1}H NMR (300 MHz, D_2O): δ = 5.10 (d, J = 1.0 Hz, 1H, H_1 α), 5.01 (d, J = 1.4 Hz, 1H, H_1 β), 4.46 (dd, J = 4.0 Hz, 1.4Hz, 1H, H_2 β), 4.33 (dd, J = 4.4 Hz, 1.0 Hz, 1H, H_2 α), 4.10 – 3.98 (m, 5H, H_3 α + H_8 αβ), 3.88 – 3.71 (m, 5H, H_5 α + H_3 β + H_6 αβ), 3.56 (t, J = 9.5Hz, 1H, H_4 α), 3.46 (t, J = 9.8Hz, 1H, H_4 β), 3.38 (ddd, J = 9.8Hz, 4.7Hz, 2.1Hz, 1H, H_5 β).

^{13}C NMR (75 MHz, D_2O): δ = 171.87 (s, C7 α), 170.97 (s, C7 β), 92.89 (s, C1 α), 92.83 (s, C1 β), 76.40 (s, C5 β), 72.03 (s, C3 β), 71.97 (s, C5 α), 68.82 (s, C3 α), 66.76 (s, C4 α), 66.52 (s, C4 β), 60.41 (s, C6 β), 60.39 (s, C6 α), 54.28 (s, C2 β), 53.37 (s, C2 α), 51.72 (s, C8 β), 51.65 (s, C8 α).

SiaNAz (*N*-azidoacetylneuraminic acid): ManNAz (42 mg, 0.159 mmol, 1 eq.) was dissolved in 600 μl PBS (pH 7.6) with sodium pyruvate (32 mg, 0.239 mmol, 1.5 eq.) in a 5 ml capped tube. 5 units of Neu5Ac aldolase (Sigma Aldrich, EC 4.1.3.3) were added and the reaction was stirred at 37 °C for 16 h. The formation of the product was checked by TLC (1-propanol, NH_3, H_2O 6 : 1 : 2.5) with resorcinol as a reagent. The reaction crude was purified by anion exchange chromatography (BioRad, AG1X8; H_2O 5 CV, NH_4HCO_3 0.05 M 5 CV, NH_4HCO_3 0.2 M 5 CV). Fractions containing the product were identified by TLC as previously then gathered and concentrated under reduced pressure. The white powder obtained was desalted using size exclusion chromatography (P2) with ultrapure water as eluent. Fractions containing the product were gathered, concentrated under reduced pressure then freeze-dried, affording the product as a white powder (18.5 mg, 0.053 mmol, 34%).

^{1}H NMR (300 MHz, D_2O): δ 4.09–3.88 (m, 5H, H_5 + H_4 + H_3 + H_2), δ 3.78 (dd, J = 11.2Hz, 2.2Hz, 1H, H_{8cis}), δ 3.65 (ddd, J = 8.5Hz, 4.1Hz, 1.2Hz, 1H, H_7), δ 3.55 (dd, J = 11.2Hz, 5.8Hz, 1H, H_{8trans}), δ 3.45 (dd, J = 9.2Hz, 0.9Hz, 1H, H_6), δ 2.18 (dd, J = 12.5 Hz, 4.8Hz, 1H, H_1 eq.), δ 1,75 (dd, J = 12.4Hz, 11.8Hz, 1H, H_1 ax)

^{13}C NMR (75 MHz, D_2O): δ 70.8 (CH, C5), δ 69.82 (CH, C7), δ 68.43 (CH, C6), δ 67.01 (CH, C2), δ 63.10(CH2, C8), δ 52.21(CH, C3), δ 51.9 (CH2, C4), δ 39.32 (CH2, C1)

Bacterial strains and growth conditions

Five bacterial strains were used in this study. The non-pathogenic model *E. coli* EV36[30] that expresses the K1 capsule was kindly provided by Pr Antonia P. Sagona. The K1 producing pathogenic strains *E. coli* K-235 (ATCC 13027) and *E. coli* U5/41 (ATCC 11775) were purchased from ATCC. *E. coli* BL21,[49] a well-described B strain, and *E. coli* Nissle 1917 expressing the K5 capsule were used as negative controls to demonstrate the specificity of the method for K1 capsule and were kindly provided by Dr Marie Titecat. The bacteria were streaked out from −80 °C stocks on LB agar Petri dishes and cultured overnight at 37 °C. Isolated colonies grown on these plates were used to inoculate liquid cultures for the rest of the experiments.

Metabolic incorporation and CuAAC buffer preparation

E. coli liquid cultures prepared as explained above were diluted to an OD_{600} of 1. 20 μl of this solution was transferred to 10 ml LB (supplemented with the desired chemical reporters) in a 50 ml conical tube and grown overnight at 37 °C/180 rpm. For fluorescence labelling by CuAAC, fresh aqueous stock solutions of the different reagents were used to prepare the CuAAC buffer. It should be prepared several minutes to an hour prior to use and should not be kept for more than a few hours. $CuSO_4$ (1 mg ml^{-1}, final concentration 150 μM) and BTTAA (10 mg ml^{-1}, final concentration 300 μM) are first mixed and added to the fluorescent probe (TAMRA-Alk or TAMRA-N_3, 1 mg ml^{-1}, desired final concentration). K_2HPO_4 (100 mg ml^{-1}, final concentration 100 mM) and H_2O are then added. Finally, sodium ascorbate (10 mg ml^{-1}, final concentration 2.5 mM) is added right before use.

Microplate fluorescence

Overnight cultures were adjusted to OD_{600} 1 by diluting the suspensions with PBS. 200 μl of these suspensions were split in 2 ml microtubes (minimum of 3 per condition) then centrifuged (2 min, 10 000 G), and pellets were resuspended in 200 μl CuAAC buffer and agitated for 45 min, 600 rpm at room temperature in the dark. These suspensions were rinsed 3 times with 1 ml PBS (2 min, 10 000 rpm) before being resuspended in 200 μl PBS. CuAAC whole cell suspensions were split in a dark opaque 96-well plate (100 μl per well). Fluorescence (λ_{em} 535 ± 20 nm/λ_{exc} 585 ± 30 nm) was measured on a CLARIOstar Plus microplate reader.

Fluorescence microscopy

Agar pads were made by pouring 10 μl of hot LB agar on a microscopy slide and quickly covering it with a glass coverslip. Whole-cell suspensions were deposited on the hardened agar pad by gently raising the coverslip and placing it back down. These slides were then observed on a Leica AF6000 LX inverted video microscope with differential interference contrast (DIC). For immunofluorescence staining of the capsule, overnight cultures were adjusted to 1 OD_{600nm} by diluting the suspensions with PBS and treated with anti-K1 rabbit antibody (ENZ-ABS559-0100 Enzo Life Sciences) (1/100, 45 min), rinsed with 200 μl PBS and treated with anti-rabbit AF488 antibody (1/250, 30 min), rinsed with 200 μl PBS and finally mounted on agar pads and observed as described above. For CuAAC fluorescence labelling, cells were treated as described for the microplate fluorescence with the fluorophore concentration adjusted to 100 mM. After the reaction and rinsing steps, cells were resuspended in 200 μl PBS and 10 μl of these suspensions were mounted on agar pad before observation on a fluorescence videomicroscope.

Viability assay and growth curves

For viability assay, OD_{600} of overnight cultures was harmonized to 1 by diluting the suspensions with sterile PBS. 200 μl of these suspensions were transferred in a 96-well plate. Serial dilutions

ranging from 10^{-1} to 10^{-10} were performed. 20 µl of these suspensions were streaked on LB agar Petri dishes and grown overnight at 37 °C. Petri dishes showing between 5 to 250 colonies were selected, colonies counted, and the number of CFU in the original suspension was deducted from the dilution. For growth curves, from an isolated colony, 2 ml of LB liquid culture were grown in a 15 ml conical tube for 2 hours at 37 °C. 10 µl of this suspension was used to inoculate 10 ml LB supplemented with the desired chemical reporter or monosaccharide in a 50 ml conical tube. OD_{600} of these suspensions was then recorded at regular intervals.

Capsule extraction and quantification

Capsular extracts were purified by following a method described previously.[37] Overnight grown cultures were suspended in 600 µl lysis buffer (100 mM SDS, 50 mM Tris, 0.128 mM NaCl). 600 µl phenol/chloroform/isoamyl alcohol (25 : 24 : 1) were added and agitated 15 min at 65 °C. After centrifugation (16 000 g for 15 min at 4 °C), the upper phase was transferred and completed with 600 µl of ice-cold absolute ethanol. The tubes were stored overnight at −20 °C. Tubes were then centrifuged (10 min, 13 000 g) and the white precipitate was rinsed with ethanol then dried under N_2 flow. 20 µl DNAse was added (45 min, 37 °C, 300 rpm) followed by 20 µl Proteinase K (1 h, 56 °C, 300 rpm). 560 µl H_2O and 600 µl phenol/chloroform/isoamyl alcohol (25 : 24:1) were added. Tubes were centrifuged (15 min, 13 000 rpm) and the upper phase transferred in a new microtube. 200 µl H_2O, 50 µl sodium acetate 3 M and 1 m absolute ethanol were added and the tubes were stored at −20 °C overnight. The tubes were centrifuged (15 min, 13 000 rpm) and the white precipitate was rinsed with 1 ml ice cold absolute ethanol and dried under N_2 flow. Capsular extracts were finally resuspended in 50 µl of H_2O. The capsule extracted was labelled and quantified on a microplate reader as follow: 10 µl of the capsular extract solutions were added to 190 µl CuAAC buffer in a microtube and agitated 45 min, 600 G at RT. After the reaction 50 µl sodium acetate 3 M and 1 ml ice cold absolute ethanol were added and tubes were stored overnight at −20 °C. Tubes were then centrifuged (15 min, 13 000 G and the white pellet was rinsed with ice cold absolute ethanol, dried under N_2 flow, resuspended in 20 µl H_2O and transferred to a dark opaque 96-well plate for fluorescence readout on a microplate reader. The purity of the capsular extract was performed by polyacrylamide gel electrophoresis, for this 10 µl of the same capsular preparation were loaded on a 15% TBE-PAGE (Tris-Boric-EDTA polyacrylamide gel electrophoresis) as described previously.[51] The capsular polysaccharides were stained with 5% alcian blue in water for 15 min followed by three washing steps with acetic acid 5% in a solution of 50% methanol until the gel background get unstained.

Conclusions

We developed a method for the specific bioorthogonal labelling of K1 capsules in *E. coli* after the metabolic incorporation of ManNAc analogues equipped with alkyne and azide chemical handles. These sialylation reporters did not exert any significant effect on bacterial growth. ManNAz was determined as the better choice among the tested reporters, as ManNAl showed inherent long-term cytotoxicity. While both ManNAc and Neu5Ac derivatives were readily incorporated in growing *E. coli* EV36, resulting in fluorescent labelling of whole bacteria, only the signal obtained from ManNAc chemical reporters could be attributed to their incorporation into the polysialic acid capsule. This allowed us to refine our understanding of the capsule metabolic pathways. The method was miniaturized as a microplate assay amenable to screening approaches. With the ability to track the K1 capsule biosynthesis, this platform might be a useful tool for future studies aiming at impacting capsule expression, which is of great interest in the context of increasing pathogen resistance.

Author contributions

V. R. and Y. R. performed the experiments. V. R., Y. R. and C. L. analysed and interpreted the data. C. L., Y. R., and V. R. wrote the manuscript. V. R. and C. L. prepared figures. C. L. and C. B. conceptualised the project. C. B. and C. L. acquired funding and supervised the work. C. L., V. R. and Y. R. revised the paper.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the CNRS and the University of Lille. V. R. is a recipient of a research grant from ANR NEURAPROBE (ANR-18-CE07-0042) (C. B., C. L.). We thank Dr Corentin Spriet (Plateformes Lilloises en Biologie et Santé (PLBS) – UAR 2014 – US 41) and Dr Boris Vauzeilles for helpful discussions, Pr Antonia P. Sagona for providing us the EV36 strain, and Dr Olivier Vidal for training of V. R. Fig. 1 and ToC were created with BioRender.com.

Notes and references

1 K. S. Kim, *Nat. Rev. Neurosci.*, 2003, **4**, 376–385.
2 A. Alkeskas, P. Ogrodzki, M. Saad, N. Masood, N. R. Rhoma, K. Moore, A. Farbos, K. Paszkiewicz and S. Forsythe, *BMC Infect. Dis.*, 2015, **15**, 449.
3 J. B. Kaper, J. P. Nataro and H. L. T. Mobley, *Nat. Rev. Microbiol.*, 2004, **2**, 123–140.
4 K. J. Kim, J. W. Chung and K. S. Kim, *J. Biol. Chem.*, 2005, **280**, 1360–1368.
5 K. S. Kim, H. Itabashi, P. Gemski, J. Sadoff, R. L. Warren and A. S. Cross, *J. Clin. Invest.*, 1992, **90**, 897–905.
6 J. B. Robbins, G. H. McCracken, E. C. Gotschlich, F. Ørskov, I. Ørskov and L. A. Hanson, *N. Engl. J. Med.*, 1974, **290**, 1216–1220.

7 J. Sarowska, B. Futoma-Koloch, A. Jama-Kmiecik, M. Frej-Madrzak, M. Ksiazczyk, G. Bugla-Ploskonska and I. Choroszy-Krol, *Gut Pathog.*, 2019, **11**, 10.
8 B. R. Kunduru, S. A. Nair and T. Rathinavelan, *Nucleic Acids Res.*, 2016, **44**, D675–D681.
9 S. M. Steenbergen and E. R. Vimr, *Mol. Microbiol.*, 2008, **68**, 1252–1267.
10 B.-X. Lin, Y. Qiao, B. Shi and Y. Tao, *Appl. Microbiol. Biotechnol.*, 2016, **100**, 1–8.
11 O. T. Keppler, P. Stehling, M. Herrmann, H. Kayser, D. Grunow, W. Reutter and M. Pawlita, *J. Biol. Chem.*, 1995, **270**, 1308–1314.
12 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596–2599.
13 C. W. Tornøe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057–3064.
14 E. M. Sletten and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2009, **48**, 6974–6998.
15 J. A. Prescher and C. R. Bertozzi, *Nat. Chem. Biol.*, 2005, **1**, 13–21.
16 V. Rigolot, C. Biot and C. Lion, *Angew. Chem., Int. Ed.*, 2021, **60**, 23084–23105.
17 K. Lang and J. W. Chin, *Chem. Rev.*, 2014, **114**, 4764–4806.
18 N. Banahene, H. W. Kavunja and B. M. Swarts, *Chem. Rev.*, 2022, **122**, 3336–3413.
19 H. Liang, K. E. DeMeester, C.-W. Hou, M. A. Parent, J. L. Caplan and C. L. Grimes, *Nat. Commun.*, 2017, **8**, 15015.
20 A. R. Brown, K. A. Wodzanowski, C. C. Santiago, S. N. Hyland, J. L. Follmar, P. Asare-Okai and C. L. Grimes, *ACS Chem. Biol.*, 2021, **16**, 1908–1916.
21 P. Kaewsapsak, O. Esonu and D. H. Dube, *ChemBioChem*, 2013, **14**, 721–726.
22 A. Dumont, A. Malleron, M. Awwad, S. Dukan and B. Vauzeilles, *Angew. Chem., Int. Ed.*, 2012, **51**, 3143–3146.
23 J. Mas Pons, A. Dumont, G. Sautejeau, E. Fugier, A. Baron, S. Dukan and B. Vauzeilles, *Angew. Chem., Int. Ed.*, 2014, **53**, 1275–1278.
24 N. Geva-Zatorsky, D. Alvarez, J. E. Hudak, N. C. Reading, D. Erturk-Hasdemir, S. Dasgupta, U. H. von Andrian and D. L. Kasper, *Nat. Med.*, 2015, **21**, 1091–1100.
25 J. E. Hudak, D. Alvarez, A. Skelly, U. H. von Andrian and D. L. Kasper, *Nat. Microbiol.*, 2017, **2**, 17099.
26 P. A. Gilormini, C. Lion, D. Vicogne, T. Levade, S. Potelle, C. Mariller, Y. Guérardel, C. Biot and F. Foulquier, *Chem. Commun.*, 2016, **52**, 2318–2321.
27 P. A. Gilormini, C. Lion, D. Vicogne, Y. Guérardel, F. Foulquier and C. Biot, *J. Inherited Metab. Dis.*, 2018, **41**, 515–523.
28 K. N. Chuh, B. W. Zaro, F. Piller, V. Piller and M. R. Pratt, *J. Am. Chem. Soc.*, 2014, **136**, 12283–12295.
29 K. Qin, H. Zhang, Z. Zhao and X. Chen, *J. Am. Chem. Soc.*, 2020, **142**, 9382–9388.
30 E. R. Vimr and F. A. Troy, *J. Bacteriol.*, 1985, **164**, 854–860.
31 C. Møller-Olsen, T. Ross, K. N. Leppard, V. Foisor, C. Smith, D. K. Grammatopoulos and A. P. Sagona, *Sci. Rep.*, 2020, **10**, 8903.
32 C. Møller-Olsen, S. F. S. Ho, R. D. Shukla, T. Feher and A. P. Sagona, *Sci. Rep.*, 2018, **8**, 17559.
33 A. E. Speers and B. F. Cravatt, *Chem. Biol.*, 2004, **11**, 535–546.
34 B. W. Zaro, Y.-Y. Yang, H. C. Hang and M. R. Pratt, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 8146–8151.
35 J. Plumbridge and E. Vimr, *J. Bacteriol.*, 1999, **181**, 47–54.
36 C. T. Saeui, E. Urias, L. Liu, M. P. Mathew and K. J. Yarema, *Glycoconjugate J.*, 2015, **32**, 425–441.
37 Y. Talyansky, T. B. Nielsen, J. Yan, U. Carlino-Macdonald, G. Di Venanzio, S. Chakravorty, A. Ulhaq, M. F. Feldman, T. A. Russo, E. Vinogradov, B. Luna, M. S. Wright, M. D. Adams and B. Spellberg, *PLoS Pathog.*, 2021, **17**, e1009291.
38 S. S. Spicer and L. Warren, *J. Histochem. Cytochem.*, 1960, **8**, 135–137.
39 B. Revilla-Nuin, A. Reglero, M. A. Ferrero and L. B. Rodriguez-Aparicio, *FEBS Lett.*, 1999, **449**, 183–186.
40 E. R. Vimr, *ISRN Microbiol.*, 2013, **2013**, 1–26.
41 M. Petersen, W.-D. Fessner, M. Frosch and E. Lüneberg, *FEMS Microbiol. Lett.*, 2000, **184**, 161–164.
42 C. Brigham, R. Caughlan, R. Gallegos, M. B. Dallas, V. G. Godoy and M. H. Malamy, *J. Bacteriol.*, 2009, **191**, 3629–3638.
43 E. R. Vimr, K. A. Kalivoda, E. L. Deszo and S. M. Steenbergen, *Microbiol. Mol. Biol. Rev.*, 2004, **68**, 132–153.
44 J. Pospíšil, D. Vítovská, O. Kofroňová, K. Muchová, H. Šanderová, M. Hubálek, M. Šiková, M. Modrák, O. Benada, I. Barák and L. Krásný, *Nat. Commun.*, 2020, **11**, 4963.
45 M. Sawa, T.-L. Hsu, T. Itoh, M. Sugiyama, S. R. Hanson, P. K. Vogt and C.-H. Wong, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 12371–12376.
46 C. Whitfield, S. S. Wear and C. Sande, *Annu. Rev. Microbiol.*, 2020, **74**, 521–543.
47 M. X. Henriques, T. Rodrigues, M. Carido, L. Ferreira and S. R. Filipe, *Mol. Microbiol.*, 2011, **82**, 515–534.
48 S. Phanphak, P. Georgiades, R. Li, J. King, I. S. Roberts and T. A. Waigh, *Langmuir*, 2019, **35**, 5635–5646.
49 F. W. Studier and B. A. Moffatt, *J. Mol. Biol.*, 1986, **189**, 113–130.
50 **https://www.iith.ac.in/EK3D/index.php**.
51 S. Pelkonen, J. Häyrinen and J. Finne, *J. Bacteriol.*, 1988, **170**, 2646–2653.

CHAPITRE 2 – ÉTUDE DE L'INFLUENCE DU GROUPEMENT RAPPORTEUR SUR L'INCORPORATION EN CELLULES MAMMIFÈRES

La majorité des études ayant recours à la stratégie du rapporteur chimique utilise des analogues modifiés par un groupement alcyne ou azoture pour l'incorporation métabolique. Bien que considérés comme discrets et affectant peu les propriétés physico-chimiques de la biomolécule, l'introduction de ces groupements ne peut pas être considérée comme anodine. Par exemple, il a été montré que les monosaccharides peracétylés modifiés par un azoture Ac_4GlcNAz et Ac_4GalNAz sont plus aisément métabolisés que leurs homologues alcynes respectifs, et l'efficacité de leur incorporation métabolique varie en fonction du type cellulaire.[99] En revanche, la conversion métabolique de Ac_4ManNAl en SiaNAl est plus efficace que celle de Ac_4ManNAz en SiaNAz dans plusieurs lignées cellulaires.[100]

OBJECTIFS ET RÉSUMÉ DES TRAVAUX

Dans le cadre d'un travail collaboratif avec le Dr Anne-Sophie Vercoutter-Edouart (CR CNRS), nous avons étudié l'efficacité du marquage des sialoglycoconjugués de surface dans des lignées cellulaires humaines du côlon qui sont au centre des recherche de cette équipe, et pour lesquelles des analyses glycomiques avaient été réalisées préalablement.[101] Pour cette étude, nous avons comparé l'incorporation des analogues non peracétylés ManNAz, ManNAl, SiaNAz et SiaNAl, ainsi que l'analogue peracétylé Ac_4ManNAz, qui reste à ce jour le rapporteur le plus couramment utilisé en MOE. En tirant parti de la réaction de CuAAC, nous avons comparé l'incorporation métabolique de ces cinq rapporteurs dans les lignées cellulaires humaines du côlon CCD841CoN, HT29 et HCT116, ainsi que dans les deux lignées cellulaires modèles les plus utilisées en laboratoire, HEK293 et HeLa. Pour cela, nous avons eu recours à différentes techniques. Le Western-blot et la cytométrie de flux nous ont permis de comparer l'incorporation de manière quantitative au sein des glycoprotéines et des sialoglycoconjugués de surface respectivement, et la microscopie confocale nous a permis d'observer les différences en termes de localisation au sein des cellules, en colocalisant les signaux issus des

fluorophores liés aux rapporteurs incorporés et un anticorps fluorescent anti trans-Golgi. L'appareil de Golgi étant l'organelle dans lequel les glycoconjugués sont assemblés, l'accumulation des rapporteurs chimiques y est maximale, et la mesure des signaux concentrés à cet endroit est un excellent indicateur du niveau de production des sialoglycoconjugués à un instant donné.

Pour ce projet, j'ai synthétisé les différents rapporteurs chimiques et effectué les expériences d'incorporation métabolique et d'observation en microscopie confocale à fluorescence. J'ai également mis au point un programme automatisant le traitement des données issues de ces expériences, dont le code source est accessible librement.[102] Celui-ci est utilisable avec le logiciel ImageJ® (une référence pour le traitement d'images issues de la microscopie), et permet, directement à partir des images produites par le microscope, de segmenter les cellules, de les compter, de localiser l'appareil de Golgi, de mesurer les niveaux de fluorescence en cellule entière et dans le Golgi, et d'exporter automatiquement ces différentes dans un tableau. Ce script produit également différents lots d'images seuillées et des masques visuels, fournissant un moyen objectif et rapide de juger la qualité des images. Il m'a permis, ainsi qu'à nos collaborateurs à qui je l'ai transmis, de traiter facilement de nombreuses conditions et images différentes, et a grandement facilité l'analyse statistique des données.

Cette étude a montré des disparités marquées dans l'efficacité du marquage des sialoglycoprotéines selon la lignée cellulaire et le rapporteur chimique utilisé, le remplacement des étiquettes azoture et alcyne sur les analogues ayant considérablement influencé leur incorporation métabolique dans les lignées cellulaires. Nos résultats ont mis en évidence que le ManNAz était le rapporteur le plus efficacement métabolisé pour étudier la sialylation dans ces cellules.

scientific reports

Check for updates

OPEN

Switching azide and alkyne tags on bioorthogonal reporters in metabolic labeling of sialylated glycoconjugates: a comparative study

Jodie Scache[1,2], Vincent Rigolot[1,2], Cédric Lion[1], Marlène Mortuaire[1], Tony Lefebvre[1], Christophe Biot[1] & Anne-Sophie Vercoutter-Edouart[1]✉

Sialylation of cell surface glycans plays an essential role in cell–cell interaction and communication of cells with their microenvironment. Among the tools that have been developed for the study of sialylation in living cells, metabolic oligosaccharide engineering (MOE) exploits the biosynthetic pathway of sialic acid (Sia) to incorporate unnatural monosaccharides into nascent sialylatedglycoconjugates, followed by their detection by a bioorthogonal ligation of a molecular probe. Among bioorthogonal reactions, the copper-catalyzed azide-alkyne cycloaddition (CuAAC) is the only ligation where both reactive tags can be switched on the chemical reporter or on the probe, making this reaction very flexible and adaptable to various labeling strategies. Azide- and alkyne-modified ManNAc and Sia reporters have been widely used, but per-*O*-acetylated ManNAz (Ac_4ManNAz) remains the most popular choice so far for tracking intracellular processing of sialoglycans and cell surface sialylation in various cells. Taking advantage of CuAAC, we compared the metabolic incorporation of ManNAl, ManNAz, SiaNAl, SiaNAz and Ac_4ManNAz in the human colon cell lines CCD841CoN, HT29 and HCT116, and in the two gold standard cell lines, HEK293 and HeLa. Using complementary approaches, we showed marked differences in the efficiency of labeling of sialoglycoproteins between the different chemical reporters in a given cell line, and that switching the azide and alkyne bioorthogonal tags on the analogs highly impacted their metabolic incorporation in the human colon cell lines. Our results also indicated that ManNAz was the most promiscuous metabolized reporter to study sialylation in these cells.

The N-Acetyl neuraminic acid (Neu5Ac, Sia) is the most abundant member of the sialic acid family. This negatively charged monosaccharide is located at the non-reductive end of glycan chains carried by glycoproteins and glycosphingolipids at the surface of human cells. Sialylation plays an important role in cell–cell recognition and cell-microenvironment interactions in physiological or pathological context. For example, changes of sialylation are associated with cancer progression and metastasis, especially in colon cancer[1,2], and sialylated cancer-associated antigens, such as sialyl-Tn, constitute targets for therapeutic strategies[3,4]. Yet, identification of proteins whose sialylation is impaired is still of interest to better understand the molecular and cellular mechanisms that are affected during tumor progression, such as migration and invasion properties, and immune escape mechanisms[5].

During the last decade, huge efforts have been made to develop novel chemical reporters for the biocompatible labeling of sialylated glycoconjugates (GCs) using metabolic oligosaccharide engineering (MOE). Such a strategy exploits the biosynthetic pathway of sialic acid to incorporate unnatural Sia derivatives into GCs in living cells[6,7]. Sia or ManNAc analogs can both be used as two different entry points into the metabolic route (Fig. 1). Synthetic ManNAc analogs are incorporated into the de novo biosynthetic pathway thanks to the *N*-acetylmannosamine kinase (MNK) activity of the bifunctional enzyme GNE (UDP-GlcNAc 2-epimerase). In contrast, GNE is not required for Sia analogs which go through the endo-lysosomal recycling pathway before being exported into the cytosol via SLC17A5 (Fig. 1). Once synthetized in the nucleus by CMP-sialic acid synthetase (CMAS), the

[1]Université de Lille, CNRS, UMR 8576, UGSF, Unité de Glycobiologie Structurale et Fonctionnelle, 59000 Lille, France. [2]These authors contributed equally: Jodie Scache and Vincent Rigolot. ✉email: anne-sophie.vercoutter@univ-lille.fr

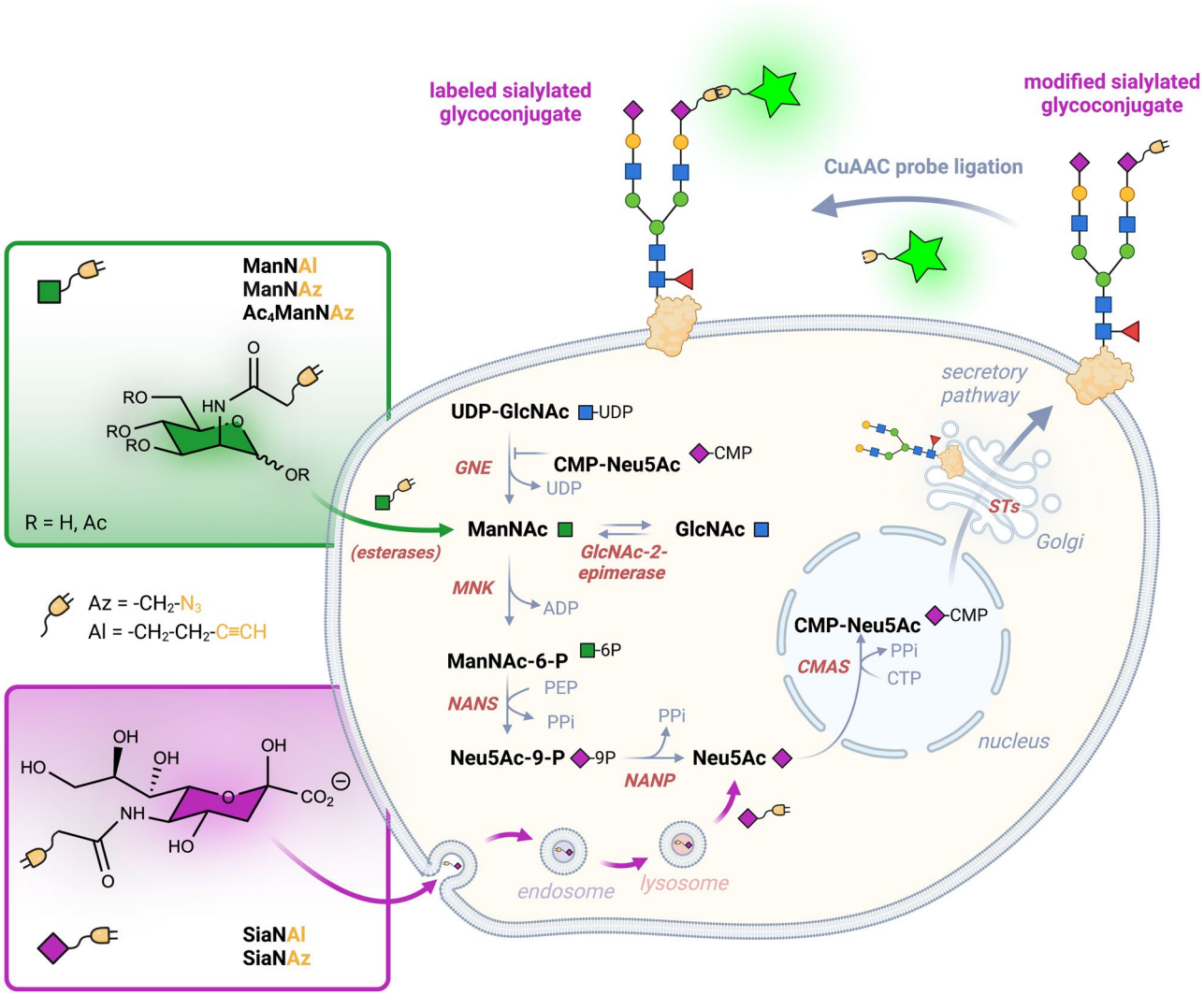


Figure 1. Metabolic incorporation of ManNAc and nonnatural ManNAc and Sia (Neu5Ac) analogs in the sialic acid biosynthetic pathway. ManNAc is synthesized from UDP-GlcNAc and GlcNAc by UDP-GlcNAc 2-epimerase (GNE) and GlcNAc-2-epimerase (RENBP), respectively. ManNAc and ManNAc analogs are phosphorylated by ManNAc 6-kinase (MNK) to yield ManNAc-6-phosphate (ManNAc-6-P). GNE/MNK is a bifunctional enzyme whose GNE activity is inhibited by the final product CMP-NeuAc. The condensation of ManNAc-6-P with phosphoenolpyruvate (PEP) is catalyzed by *N*-acetylneuraminate synthase (NANS) to produce *N*-acetyl neuraminic acid-9-phosphate (Neu5Ac-9-P), which is then dephosphorylated by sialic acid phosphatase (NANP) to yield Neu5Ac. Cytosolic pool of Neu5Ac can also be provided by the salvage pathway through endocytosis of free sialic acid and lysosomal recycling. Transport of Neu5Ac to the nucleus enables CMP-sialic acid synthetase (CMAS) to produce CMP-NeuAc. After its transport into the Golgi apparatus, CMP-NeuAc is utilized by the sialyltransferases (STs) for the terminal glycosylation of glycoconjugates that are ultimately secreted or destined for the cell surface or organelles, as illustrated here with a transmembrane *N*-glycoprotein.

functionalized CMP-Sia is then transported into the Golgi where it is used by sialyltransferases (Fig. 1). Among the reactive chemical groups, alkynes and azides are totally absent from living organisms, small, relatively stable and chemically inert in the cellular environment. This makes the metabolic chemical reporters bearing them structurally as close as possible to the natural analogs, offering great confidence in their physiological behaviors as a model. Modified alkyne- or azide ManNAc and Sia monosaccharides are usually used at high concentrations (>0.5–1 mM) because of their high polarity that limits their cellular uptake[8–10]. For that reason, per-*O*-acetylated analogs are preferentially used, due to their hydrophobic properties that facilitate their passive diffusion across the cell membrane, allowing lower working concentrations (20–100 μM)[7,11]. Both peracetylated alkyne and azide-modified ManNAc and Sia can be used. However, differences in the efficiency of incorporation and cell surface labelling were reported, as shown for example for Ac_4ManNAz and Ac_4ManNAl[12,13], or Ac_5SiaNAz and Ac_5SiaNAl (referred to as Ac_5SiaNPoc)[14,15]. Sia reporters, whose synthesis is demanding, seem to be more efficiently incorporated into sialylated GCs compared with their ManNAc analogs[9,14]. Indeed, Sia analogs enter

cells via the salvage pathway and are directly available for CMP activation in the nucleus, while the conversion of ManNAc analogs into the modified nucleotide-sugar requires three additional enzymatic steps (Fig. 1)[7]. For this reason, MOE using ManNAc and Sia analogs is useful to study defects in the sialic acid biosynthetic and recycling pathways[8,9,16].

Among the bioorthogonal ligations used to label the modified sialylated GCs with a fluorescent or biotinylated probe, cycloadditions including the Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and the strain-promoted [3 + 2] alkyne-azide cycloaddition (SPAAC), have been widely used in vitro and in vivo[7–9,12,13,17–22]. CuAAC is the only bioorthogonal reaction in which alkynes and azides can be introduced indifferently either on the metabolic analog or on the probe for the ligation, making this reaction very flexible and adaptable to various labeling strategies. In contrast, only azides are adapted for a bioorthogonal ligation via SPAAC reaction, which is easier to apply in living systems than CuAAC due to the absence of copper associated cytotoxicity, but generally suffers from poorer kinetics and specificity especially in the intracellular environment[23]. Albeit numerous works using MOE to label sialylated glycans have been reported, the comparison of alkyne- or azide-functionalized ManNAc and Sia free monosaccharides has not been done so far. Therefore, in this work, we utilized the advantage of CuAAC to swap the azide and alkyne tags between the unnatural monosaccharides and the probes. We compared the metabolic incorporation efficiency of ManNAl, ManNAz, SiaNAl, SiaNAz and peracetylated Ac_4ManNAz in several human cell lines (Fig. S1). We chose the two widely used HEK293 and HeLa cells, and the three human colon cell lines CCD841CoN, HT29 and HCT116. The latter are particularly interesting as in vitro cellular models of colon cancer progression for studying how cytosolic *O*-GlcNAcylation can interfere with glycosylation that occurs in the secretory pathway[24,25]. After ligation with the appropriate probe, the labeling of sialylated GCs was compared by flow cytometry, Western blotting and confocal fluorescent microscopy. We showed marked differences in the efficiency of cell surface glycan labeling between the five different bioorthogonal reporters in a given cell line. Moreover, although Ac_4ManNAz has been extensively used in many cell types, we found that ManNAz appears to be the bioorthogonal reporter of choice to detect sialylated GCs at the cell surface of human colon cells.

Results

Comparison of labeling of cell surface sialoglyconjugates by flow cytometry. Monosaccharides are polar molecules that are unable to cross the plasma membrane by passive diffusion. Therefore, the usual strategy uses peracetylated analogs, which are capable of passive diffusion and are then hydrolyzed in the cytosol by non-specific esterases to regenerate the free sugar. Unlike some other monosaccharides studied in metabolic glycoengineering studies, ManNAc and its derivatives are capable of crossing the plasma membrane in their free-hydroxyl form by an active transport mechanism that is yet unidentified[9], therefore unprotected ManNAc derivatives have also been used to label sialoglycoconjugates. In agreement with previous studies, we used 500 µM of the free-hydroxyl analogs[9,10]. Ac_4ManNAz is usually used in a range of 20 to 100 µM[7,11]. However, we found that 100 µM Ac_4ManNAz reduced approximately by 40% the cellular growth of CCD841CoN, HT29 and HCT116 human colon cell lines (Fig. S2), as reported in another cell lines[26]. Therefore, the experiments were carried out with 50 µM Ac_4ManNAz. Moreover, high concentrations of Ac_4ManNAz can lead to S-glyco-modification side-reactions between peracetylated monosaccharides and protein cysteine residues, inducing nonspecific labeling of proteins that cannot occur with non-peracetylated analogs[27,28]. Western blot was used to check the expression of GNE, the first enzyme required to metabolize the unnatural ManNAc analogs into the sialic acid biosynthesis pathway[29] (Fig. 1). Our result showed that GNE was expressed at a similar level in the different cell lines (Fig. S3). However, HT29 cells expressed two isoforms of GNE, GNE1 (lower band, ~ 79 kDa) and GNE2 (higher band, ~ 82 kDa) at the same level, while only GNE1 (later referred to as GNE) was strongly detected in the other cell lines[30]. Incubation of cells with ManNAc or ManNAz did not change the steady-state level of any of the two GNE isoforms (Fig. S3).

Flow cytometry analysis was performed to measure cell surface fluorescence after 24 h and 48 h of metabolic incorporation with each analog (Fig. 2). It should be noted that chelation-assisted CuAAC was carried out for the detection of cell surface alkyne-modified GCs by flow cytometry. Picolyl-azide probes (here picolyl-azide-AF594) contain an internal copper-chelating moiety that reacts faster and in a more robust manner with alkyne-modified biomolecules compared with non-chelating azide probes[21,31]. In order to point out the marked differences in MFIs, histograms are presented with raw means of fluorescence intensity (MFIs) rather than relative MFIs (Fig. 2b). This is illustrated when comparing Ac_4ManNAz and ManNAz for example: the cell surface fluorescence increased from ten-fold to 40-fold in cells cultured with ManNAz compared with the negative controls, whereas it only increased from twice to three-fold with Ac_4ManNAz (Fig. 2b). Relative MFIs and statistical analysis are shown in Fig. S4. Whatever the analog and the cell line, the differences of MFI observed between 24 and 48 h of incubation were not significant, except for ManNAl in CCD841CoN, HCT116 and HEK293 cells (Figs. 2, S4). In these cells, there was an increase of almost two times of cell surface fluorescence between 24 and 48 h of incorporation of ManNAl. Kinetics of incorporation and increase of fluorescence intensity were similar between ManNAz and ManNAl, except in HT29 cells which did not metabolize ManNAl (Fig. 2).

When comparing the metabolic incorporation of SiaNAz and SiaNAl to their ManNAc counterparts, we noted that the relative MFI obtained with SiaNAz was as strong as ManNAz in HCT116, HEK293 and HeLa cells (Fig. S4). In contrast, the SiaNAz-induced relative cell surface fluorescence was reduced by about 60% in CCD841CoN and HT29 cells compared to ManNAz (Fig. S4). The reduced metabolic conversion of Sia compared to ManNAc chemical reporters was even more obvious with the alkyne tag: Relative MFIs were 50–60% lower than the ones obtained with ManNAl in all the cell lines except HT29 cells which did not metabolize SiaNAl at all (Figs. 2, S4).

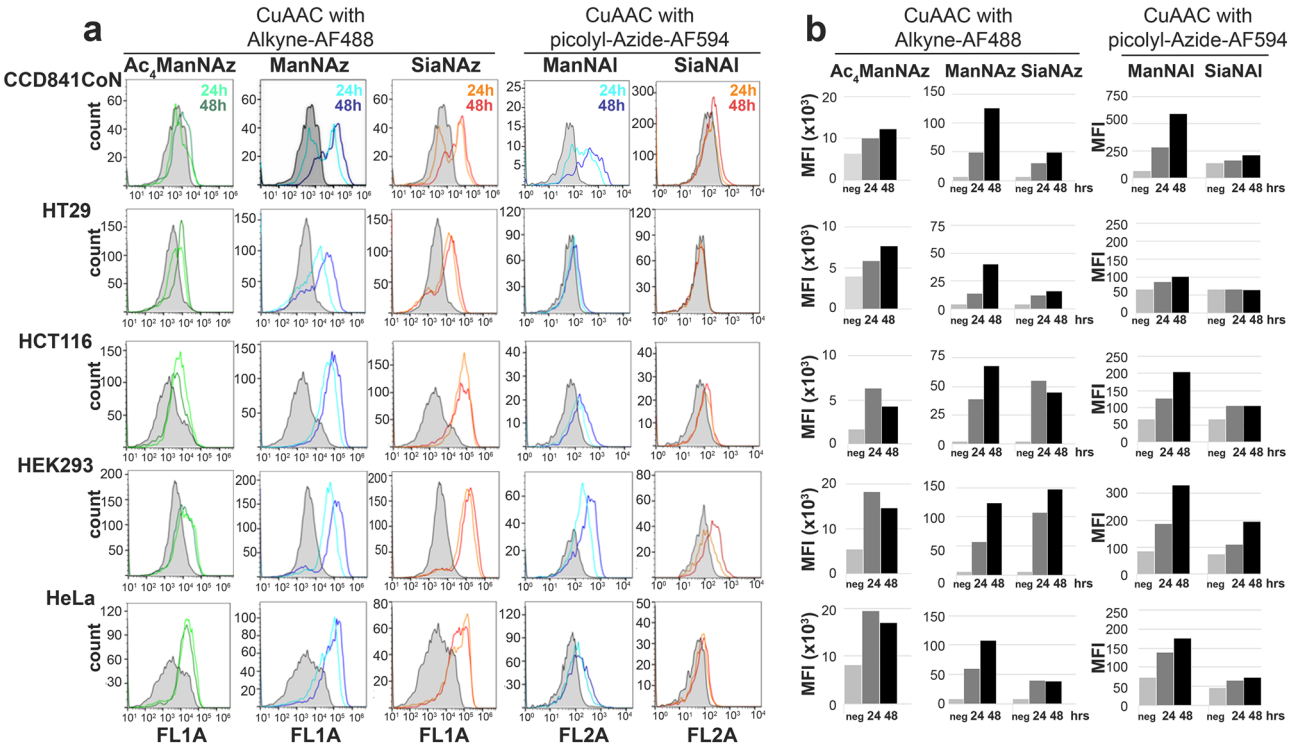


Figure 2. Flow cytometry analysis of labeled sialoglyconjugates at cell surface. Cells were incubated 24 h (in light color) or 48 h (in dark color) with the analogs before CuAAC using the appropriate fluorescent probe. For the negative controls (in grey), CuAAC was performed on cells cultured with ManNAc. (**a**) Flow cytometry histograms and (**b**) histograms showing the mean of fluorescent intensity (MFI). Results are representative of three experiments.

Comparison of in cell glycoconjugates labeling by confocal microscopy. The labeling of sialylated GCs was monitored by confocal microscopy, 16 h or 24 h following metabolic incorporation of Ac_4ManNAz, ManNAz, SiaNAz and ManNAl. SiaNAl was not included in this experiment because it gave only faint signal by flow cytometry. To compare the labeling of the different reporters within cells, we performed CuAAC with the azide-biotin or the alkyne-biotin probe on fixed and permeabilized cells. GCs were visualized using Streptavidin-DyLight488 (Strept). Sialylation occurs in the Golgi complex, making its organelle the first in which incorporated chemical reporters can be visualized onto GCs. Therefore, the Golgi marker TGN46 was used to colocalize the Strept signal in the Golgi (Fig. 3). We also quantified the intensity of fluorescence of sialylated GCs in and out of the Golgi to get an overview of the steady-state metabolic incorporation of each analog during the 16 or 24 h of continuous incubation (Fig. S5). A high integrated density (IntDen) of Strept "in the Golgi" (compared to the negative control) means that chemically modified CMP-NeuAc is efficiently transferred onto GCs by sialyltransferases. In contrast, a low and non-significant IntDen in the Golgi reflects a poor incorporation of the analog. The "out of Golgi" value reflects the steady-state trafficking of labeled GCs to the plasma membrane, their localization at the plasma membrane, but also their recycling in the endolysosomal system and the re-use of recycled non-natural CMP-NeuAc exported from the lysosomes. Because new chemical reporters are still being metabolized and processed in the Golgi apparatus in the meantime, the "in Golgi" value does not automatically decrease reciprocally.

As shown by confocal microscopy, metabolic incorporation of ManNAz, SiaNAz and ManNAl led to a strong fluorescent signal in the Golgi, as evidenced by the partial colocalization with TGN46 (Fig. 3) and the significant values of "in Golgi" IntDen of Streptavidin in the cells, except HT29 cells (Fig. S5). Also, fluorescent streptavidin was detected at the plasma membrane and in the cytoplasm (Fig. 3), leading to a significant increase of "out of Golgi" IntDen compared to the negative controls (Fig. S5). Green spots in cytoplasm might stand for vesicles from the secretory pathway, leading ultimately to the labeling of plasma membrane, and endocytic vesicles of recycling membrane glycoproteins. In HT29 cells, no specific signal was observed with ManNAl after 16 h of incorporation compared to the negative control (ManNAc, azide-biotin) (Fig. 3b). However, a slight colocalization of sialylated glycoconjugates with TGN46 was observed after 24 h, indicating that ManNAl was metabolized at a very low rate in these colon cancer cells (Fig. 3b).

Lastly, incubation of the colon cell lines with Ac_4ManNAz led to a strong and diffuse nuclear and cytosolic labeling (Fig. 3a–c), as previously reported in another cancerous cell lines[19]. There was no colocalization of Strept signal with TGN46 in these cells, as confirmed by the quantification that showed that most of the signal was quantified as "out of Golgi" (Fig. S5a–c). This indicates that the signal obtained with Ac_4ManNAz appears to be non-specific to the Golgi complex compared to the other analogs. Taken together, these data highlighted that

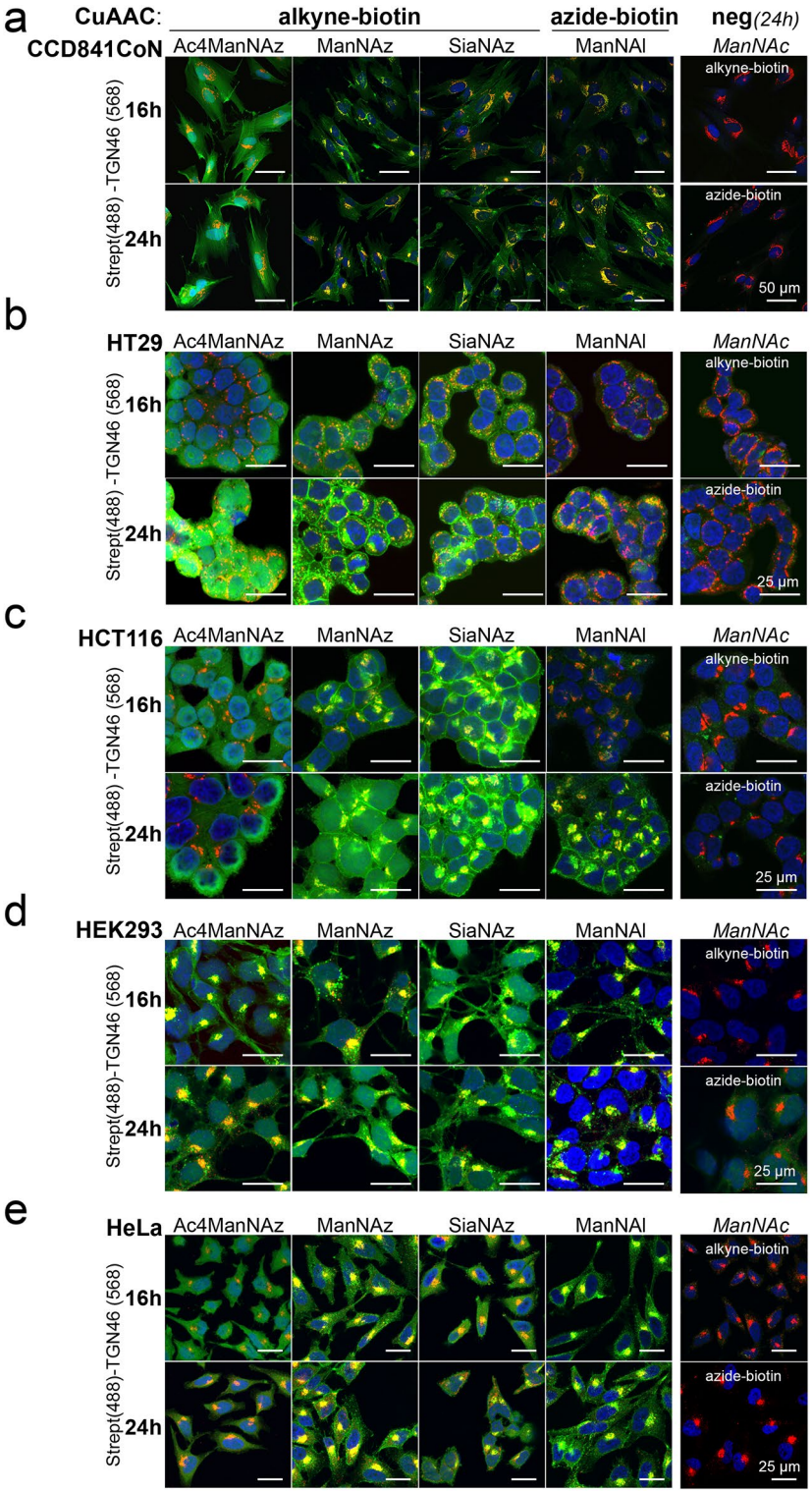


Figure 3. Confocal microscopy detection of labeled sialoglycoconjugates. Cells were incubated 16 h or 24 h with either the indicated analog or ManNAc (negative control). Then they were fixed and permeabilized before CuAAC with the appropriate biotin probe, followed by overnight incubation with anti-TGN46 (Golgi marker). Biotinylated proteins and TGN46 were revealed with Streptavidin-DyLight488 and AF568-conjugated secondary antibody, respectively. Nuclei were stained with DAPI before mounting. Two experiments were performed.

ManNAz and SiaNAz reporters appear to be the best compromise for labeling of sialylated GCs into the Golgi and on cell surface of the human colon cells.

Comparison of bioorthogonal labeling of sialylated glycoproteins by Western Blot. We further investigated the discrepancies between metabolic incorporation of the unnatural alkyne- and azide-modified ManNAc/Neu5Ac by detecting the labeled glycoproteins (GPs) by Western blot. Cells were cultured for 24 or 48 h with the chemical reporters and click chemistry reaction was performed with the appropriate biotinylated probe onto whole cells. A large range of tagged sialoglycoproteins was detected in CCD841CoN cells whatever the analog, although the use of SiaNAz gave a fainter signal for the same time of exposure (Fig. 4a). Conversely, in HT29 cells, ManNAz was the most efficient chemical reporter for the labeling of sialylated GPs compared to the others: Faint signals were observed with SiaNAz and Ac_4ManNAz, and no specific bands were detected with ManNAl and SiaNAl (Fig. 4b). In agreement with the flow cytometry analysis, incorporation of ManNAz, ManNAl, SiaNAz, and to a lesser extent Ac_4ManNAz, gave a strong labeling of GPs in HCT116, HEK293 and HeLa cells, while only faint signals were obtained with SiaNAl for similar times of exposure (Fig. 4c–e).

Ac_4ManNAz was used at 50 µM while the free hydroxyl analogs were used at 500 µM. To evaluate whether fainter signals obtained with Ac_4ManNAz could be due to this large discrepancy, we evaluated the labeling of GPs in cells incubated for 48 h with 50 µM of all the analogs. Cells incubated with 50 µM ManNAc and 500 µM ManNAz stand respectively for the negative and positive controls (Fig. S6). We observed a stronger labeling of GPs with Ac_4ManNAz compared with ManNAz and SiaNAz in HEK293 and HeLa cells, but not in the colon cell lines. No labeling was observed with 50 µM ManNAl and SiaNAl, whatever the cell lines (Fig. S6). These results indicate that high concentrations of free hydroxyl analogs are needed to force cell entry via active transport or endocytosis, whereas peracetylated reporter Ac_4ManNAz penetrates the cells by passive diffusion. The strongest labeling of GPs was obtained with 500 µM ManNAz, therefore we used this reporter for the further experiments.

We next examined the relative labeling of *N*- and *O*-glycoproteins using ManNAz in the various cell lines. CuAAC was performed as previously described with the alkyne-biotin probe before cell lysis. Lysates were then incubated with PNGase F to remove *N*-linked glycans from glycoproteins. The activity of PNGase F was first checked by lectin blotting using L-PHA which specifically binds β1,6-branched oligosaccharides of complex-type *N*-linked glycans. The L-PHA signal was rather totally lost in cell lysates treated with PNGase F compared with nontreated lysates, asserting the cleavage efficiency of PNGase F (Fig. S7). Then, treatment with PNGase F was applied on cell lysates following click chemistry reaction. The signal of biotin-tagged sialylated GPs detected by Western blot was reduced by approximatively 40% in CCD841CoN and by 70% in the other cell lines (Fig. 5a). Some of the remaining bands observed in the PNGase F-treated samples might match the signal of *O*-glycoproteins detected with jacalin, a lectin that recognizes the Galβ1,3GalNAc motif found on *O*-linked glycans (Fig. 5a).

Finally, to test whether ManNAz could be metabolized into both sialylated GPs and glycosphingolipids at the membrane of colon cells, harvested cells were incubated with bromelain, a widely acting protease. Then cells were washed and CuAAC was done using either the alkyne-biotin probe for Western blot (Fig. 5b, d), or the alkyne-AF488 probe for flow cytometry analysis (Fig. 5c). Loss of E-cadherin detection in bromelain-treated HT29 cells confirmed the cleavage activity of the protease towards cell surface GPs (Fig. S8a). Both α2,3 and α2,6 sialylation was affected by the protease, as evidenced by the decrease of cell surface fluorescence using respectively MAA and SNA lectins (Fig. S8b). We showed that detection of biotinylated sialoglycoproteins by Western blot was almost completely lost in cells incubated with bromelain compared to control cells (ManNAz, without protease) (Fig. 5b). However, flow cytometry analysis showed that the intensity of fluorescent tagged GCs was markedly reduced after treatment with bromelain compared to the controls: after bromelain treatment, cell surface fluorescence was reduced by approximately 50% (Fig. 5c). Sialylated glycosphingolipids (gangliosides) have been characterized in the three colon cell lines[25], and in HEK293 and HeLa cells[32]. To assess whether the remaining cell surface fluorescence after protease treatment could be attributed to gangliosides, we performed lipid extraction after CuAAc, and glycolipids were spotted on PDVF membrane. Dot-blot using anti-biotin showed that labeled gangliosides were detected in samples from bromelain-treated or not treated cells (Fig. 5d). This indicates that both GPs and glycosphingolipids were labeled using ManNAz, as previously shown for example with Ac_4ManNAz in Caco-2 colon cancer cells[33], or SiaNAl in human fibroblasts[8].

Discussion

Previous studies reported that metabolic incorporation of alkyne- and azide- modified monosaccharides displays a large discrepancy, but most of these studies used peracetylated precursors. For example, peracetylated GlcNAz and GalNAz are more easily metabolized than their respective alkyne counterparts, and the efficiency of their metabolic incorporation is cell-type specific [34]. In contrast, the metabolic conversion of Ac_4ManNAl into SiaNAl is more efficient than Ac_4ManNAz into SiaNAz in several cell lines[12]. Here, we used the non-protected analogs ManNAz, ManNAl, SiaNAz or SiaNAl and the peracetylated Ac_4ManNAz to investigate the efficiency of labeling of cell surface sialoglyconjugates in three human colon cell lines in which we previously carried out glycomic analysis[25], and compare them to the widely used HEK293 and HeLa cells. First, our results emphasize the cell line-dependent capability to metabolize the unnatural monosaccharides targeting sialic acid biosynthesis, as previously reported with Ac_4ManNAz and other synthetic ManNAc analogs[12,19,33,35,36]. HEK293 and HeLa efficiently metabolize the five unnatural analogs. In contrast, CCD841CoN cells, HT29 and HCT116 colon cells display a marked difference in their ability to use the chemically modified ManNAc and Neu5Ac reporters, even though they derive from the same organ. Also, switching the azide and alkyne bioorthogonal tags highly impacts their metabolic incorporation. The non-immortalized and fetal colon CCD841CoN cells are the most promiscuous ones since they efficiently metabolize any of the analogs, whatever the tag used. Conversely, the colon adenocarcinoma cells HT29 poorly metabolize ManNAl and SiaNAl reporters in comparison with ManNAz and SiaNAz.

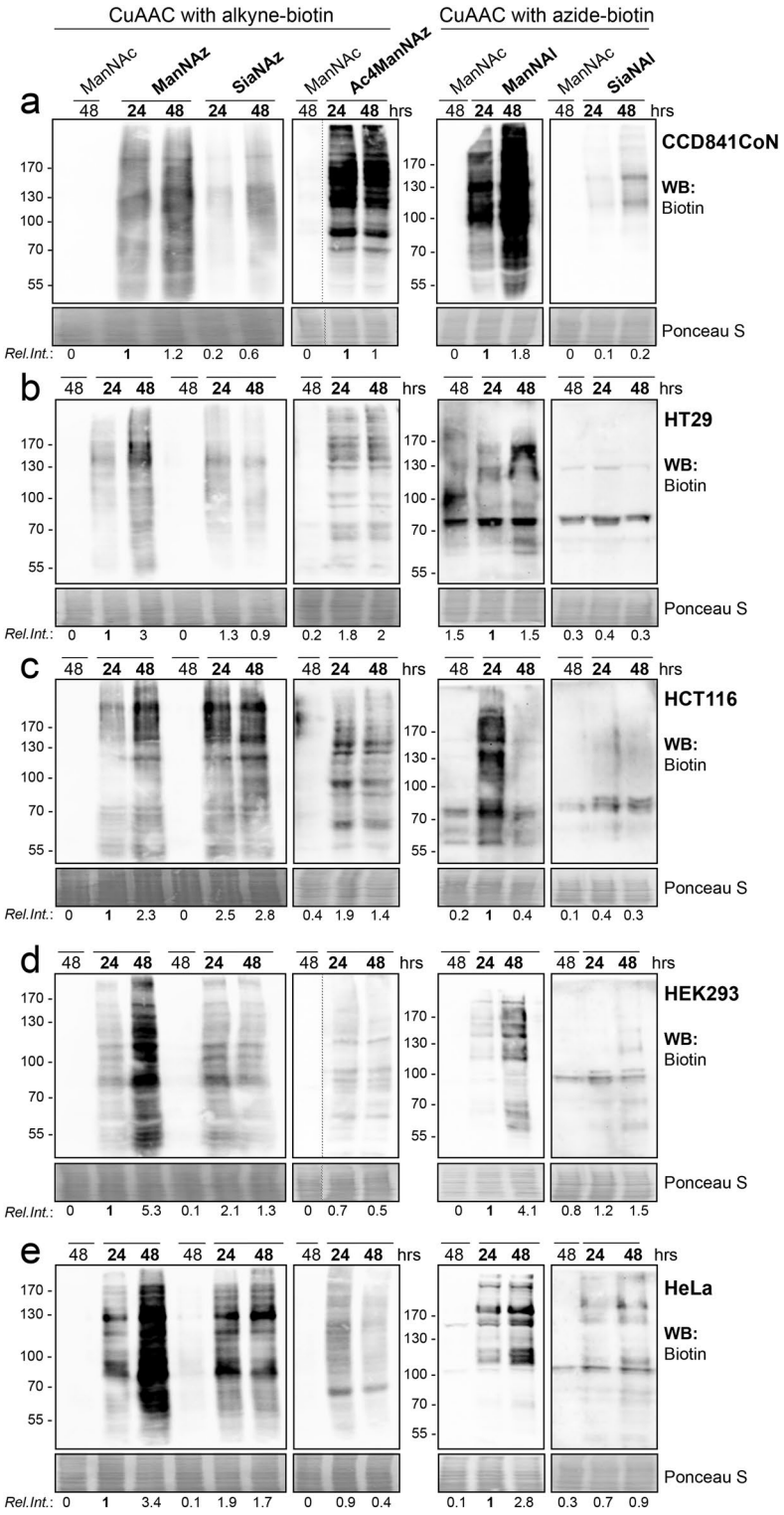


Figure 4. Comparison of the efficiency of metabolic incorporation of the analogs by Western blot. Unnatural monosaccharides were added in cell culture medium for the indicated times (at 500 μM, except for Ac_4ManNAz, 50 μM). CuAAC was then performed onto whole cells using the appropriate biotin probe. Biotinylated glycoproteins were detected by Western blotting using anti-biotin-HRP. For each analog, similar time exposures are shown for all the cell lines, except for CCD841CoN cells for which shorter times have been chosen to avoid overexposure. Equal loading was confirmed by Ponceau red staining. The relative intensity (Rel. Int.) of each lane is indicated (biotin/Ponceau red), compared to the lane '24 h ManNAz' or '24 h ManNAl' for click chemistry with the alkyne probe or the azide probe, respectively. MW standards are indicated in kDa. Results are representative of three experiments.

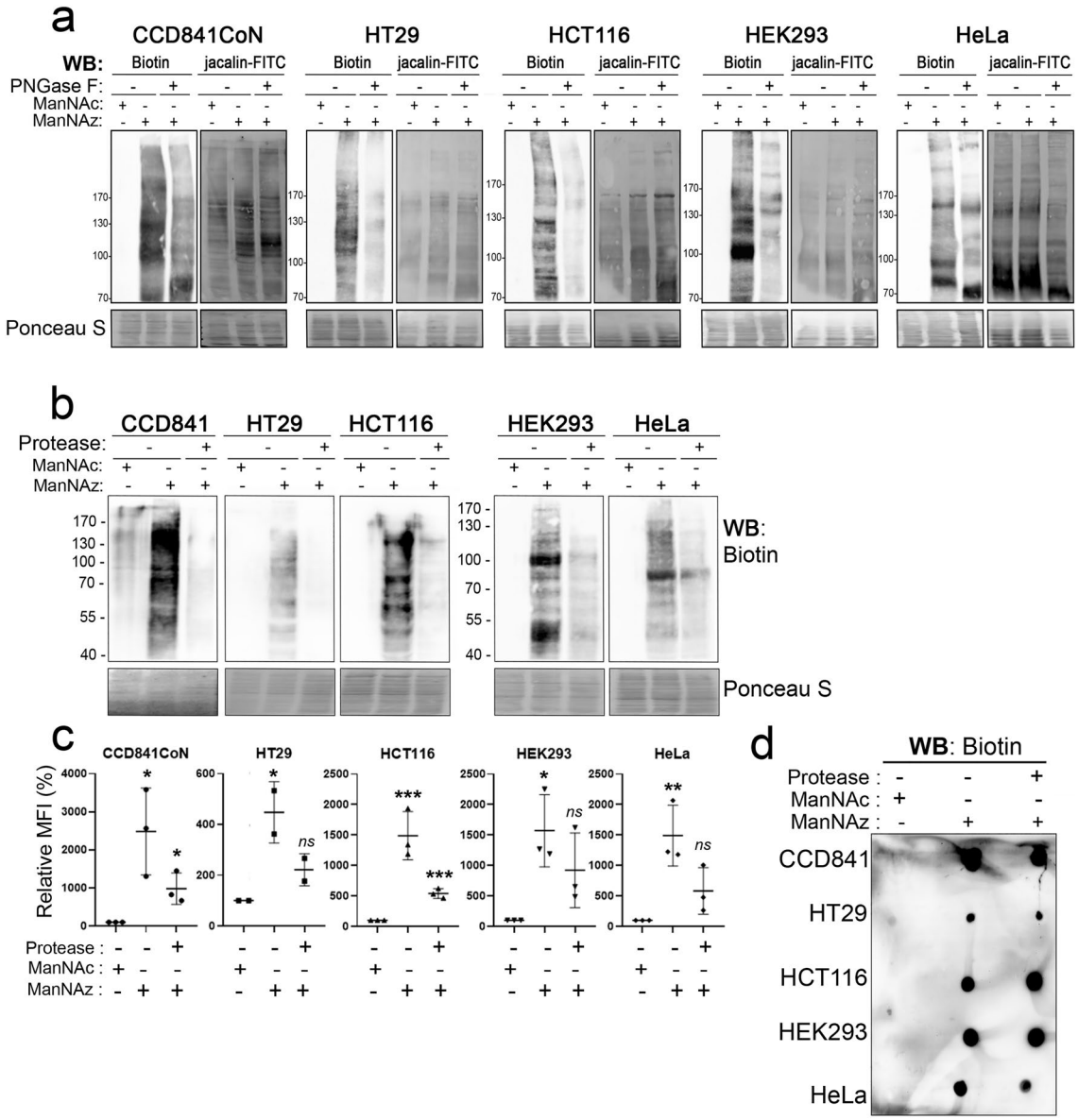


Figure 5. Sensitivity of labeling of glycoconjugates to PNGase F or protease treatment. Cells were incubated with 500 µM of either ManNAc (negative control) or ManNAz for 48 h before harvesting. **a**) After CuAAC with the alkyne-biotin, cells were lyzed and 30 µg of proteins were incubated with or without PNGase F before Western blotting using anti-biotin-HRP or Jacalin-FITC. **b**–**d**) Harvested cells were incubated for 15 min at 37 °C with or without the protease. Then CuAAC was performed onto cell pellets with either the alkyne-biotin probe (**b, d**) or the alkyne-AF488 probe for flow cytometry analysis (**c**). Labeled glycoproteins (**b**) or extracted gangliosides (**d**) were detected by Western blotting using anti-biotin-HRP. Results are representative of at least two experiments. *** $p \leq 0.005$, ** $p \leq 0.01$, * $p \leq 0.05$; *ns*, non-significant (compared to the negative control).

Impairment in the metabolic incorporation of the alkyne-modified precursors in HT29 cells might be due to a defect in the late steps of sialylation. It is unlikely to be due to a low activity of CMAS towards alkyne-substrates to produce CMP-Neu5Ac derivatives. Indeed, we previously showed that the activity of a recombinant bacterial CMAS was as effective for SiaNAz and SiaNAl as for Neu5Ac[37]. It is generally accepted that small modifications at C-5 and C-9 of the sialic acid moiety are well tolerated by sialyltransferases[7,12,38]. However, we cannot exclude that the activity rate of some sialyltransferases could be affected by the alkyne tag on the unnatural CMP-Neu5Ac. Indeed, it has been evidenced that the activity of ST8Sia-II, but not ST8Sia-IV, is affected by the use of C-5 modified sialic acids for the polysialylation of neural cell adhesion molecule (NCAM) in cells[39]. This is also the case for ST6Gal-I and ST3Gal-IV whose in vitro activity is differentially affected by C-9 modified neuraminic acids[40].

Strikingly, HT29 is the only cell line used in this work to express two isoforms of GNE, GNE1 and GNE2. GNE2 has additional N-terminal 31 amino-acids compared to the ubiquitously expressed GNE1 (commonly referred to as GNE). GNE2 displays a drastically reduced epimerase activity and is found only as a dimer[30]. In contrast, GNE1 exists in different oligomeric states: the hexameric state is the fully functional active form of GNE in vitro, but how the oligomerization of GNE affects its functional activity into cells is not yet understood[29,30].

Interestingly, in vitro assays showed that GNE has a strongly reduced kinase activity towards ManNAl (ManNPoc) or ManNBut, compared to ManNAc[41]. In HT29 cells, the presence of GNE2 into the oligomeric complexes of GNE might change the promiscuous kinase activity of GNE towards ManNAl. Further investigations are needed to explore this hypothesis.

So far, Ac_4ManNAz has been widely used for MOE targeting sialylation. Here we show that ManNAz is a good alternative for the labeling of cell surface sialylated GCs, especially for the colon cell lines whose proliferation is sensitive to high concentration of Ac_4ManNAz (≥ 100 µM). It may come from the perturbation of intracellular pH due to acetic acid release into cells, and the cross-reactivity with cysteine residues of cytosolic proteins[23,28]. In addition, Ac_4ManNAz metabolic incorporation results in a diffuse fluorescence in the cytoplasm and nucleus of HT29 and HCT116 cells. In contrast, distinct subcellular compartments such as Golgi apparatus are clearly labeled in colon cells treated with ManNAz. This may reflect the marked difference of metabolism kinetics of these two compounds. Indeed, ManNAz enters cells by an active transport mechanism that remains to be identified[9] but is then directly incorporated into the sialic acid de novo metabolic pathway. On the contrary, Ac_4ManNAz enters by passive diffusion, but requires non-specific cytosolic esterase activity to be fully deacetylated before being incorporated in the metabolic pathway. Diffuse cytoplasmic staining and relatively low labeling of GCs observed in colon cancer cells may come from an inefficient and/or partial deacetylation of Ac_4ManNAz, which might perturb its metabolic incorporation and even induce cytotoxicity in these cells. As previously performed in cultured cells and organisms, longer incubation period could address this issue[12].

Another hypothesis regarding the diffuse staining of labeled GCs into cells is supported by a recent study showing that incubation of human fibroblasts with Ac_4ManNAl (Ac_4ManNPoc) resulted in biosynthesis of not only CMP-SiaNAl, but also UDP-HexNAl[42]. The latter may come from the hijacking of unnatural ManNAc by GlcNAc-2-epimerase (RENBP), leading to the biosynthesis of chemically modified UDP-GlcNAc through the salvage pathway[43] (Fig. 1). Cytosolic UDP-GlcNAz could be further used for *O*-GlcNAc modification proteins catalyzed by the nuclear and cytosolic enzyme, *O*-GlcNAc transferase (OGT)[23]. Interestingly, OGT is overexpressed in colon cancer cells[24]; this could explain in part the higher cytosolic and nuclear signal observed in the two cancer cell lines cultured with Ac_4ManNAz, compared to the non-cancerous CCD841 cell line. It is worth noting that the low relative abundance of unnatural UDP-GlcNAc compared to its natural counterpart may not be sufficient to detect the incorporation of modified GlcNAc into the *N*- and *O*-linked glycans on cell surface, as evidenced by a glycomic study in prostate cancer cells cultured with Ac_4ManNAz[44].

Numerous unnatural ManNAc and Neu5Ac analogs are available to study sialylation in cells and organisms using MOE. However, as highlighted in this work, caution should be taken to carefully choose the chemical reporter for a given cellular model. Among the five reporters that we tested in this study, ManNAz is the most appropriate analog to label sialylated GCs that are being processed into cells and exposed at the cell surface of human colon cells. Also, the azido group offers the great advantage to be converted to an amine group in presence of a reducing agent[33,44], thus facilitating the analysis of metabolically modified glycans by mass spectrometry.

Methods

Materials. Stock solutions (50 mM) of the chemically modified monosaccharides were prepared in PBS, except for Ac_4ManNAz which was dissolved in DMSO. Acetylene-polyethylene glycol (PEG)4-Biotin (Alk-biotin, CLK-TA105, 10 mM in DMSO) and AF594-Picolyl-Azide (CLK-1296, 2 mM in DMSO) were purchased from Jena Bioscience (Jena, Germany). Azide-PEG3-biotin (Az-biotin, 762024, 10 mM in DMSO), Azide-fluor 545 (760757, 2 mM in DMSO), Fluor 488-Alkyne (761621, 2 mM in DMSO) were from Merck-Sigma-Aldrich. Streptavidin DyLight® 488 and lectins (MAA-II-FITC, FL-1344; SNA-FITC, FL-1301; Jacalin-FITC, FL-1151; L-PHA-biotin, B-1115) were from Vector Laboratories (Eurobio, Les Ulis, France).

Chemical synthesis of unnatural monosaccharides. ManNAl (*N*-(4-pentynoyl mannosamine) and SiaNAl (*N*-(4-pentynoyl) neuraminic acid) were synthesized as previously described[9].

Synthesis of ManNAz (*N*-azidoacetylmannosamine). 2-azidoacetic acid (117 mg, 1.16 mmol, 1 eq) was dissolved along with DIC (*N*–*N*'-diisopropylcarbodiimide, 175 mg, 1.392 mmol, 1.2 eq), HOBt (hydroxybenzotriazole, 195 mg, 1.276 mmol, 1.1 eq) and DIPEA (*N*,*N*-diisopropylethylamine, 141 mg, 1.392 mmol, 1.2 eq) in DMF (dimethylformamide, 20 mL), D-mannosamine hydrochloride (250 mg, 1.16 mmol, 1 eq) was added and the reaction was stirred at room temperature (r.t.) for 19 h under argon. The total consumption of mannosamine was assessed by silica thin layer chromatography (TLC) (CH_2Cl_2/MeOH 9:1 v/v) before solvent removal under reduced pressure. The reaction crude was purified by silica flash column chromatography (50 µm, 40 g, dry load, CH_2Cl_2/MeOH 95:5 v/v), fractions containing the product were gathered and concentrated under reduced pressure, affording ManNAz as a white powder (175 mg, 0.67 mmol, 57%).

^{1}H NMR (300 MHz, D_2O) δ = 5.10 (d, *J* = 1.0 Hz, 1H, H1 α), 5.01 (d, *J* = 1.4 Hz, 1H, H1 β), 4.46 (dd, *J* = 4.0 Hz, 1.4 Hz, 1H, H2 β), 4.33 (dd, *J* = 4.4 Hz, 1.0 Hz, 1H, H2 α), 4.10–3.98 (m, 5H, H3 α + H8 αβ), 3.88–3.71 (m, 5H, H5 α + H3 β + H6 αβ), 3.56 (t, *J* = 9.5 Hz, 1H, H4 α), 3.46 (t, *J* = 9.8 Hz, 1H, H4 β), 3.38 (ddd, *J* = 9.8 Hz, 4.7 Hz, 2.1 Hz, 1H, H5 β).

^{13}C NMR (75 MHz, D_2O) δ = 171.87 (s, C7 α), 170.97 (s, C7 β), 92.89 (s, C1 α), 92.83 (s, C1 β), 76.40 (s, C5 β), 72.03 (s, C3 β), 71.97 (s, C5 α), 68.82 (s, C3 α), 66.76 (s, C4 α), 66.52 (s, C4 β), 60.41 (s, C6 β), 60.39 (s, C6 α), 54.28 (s, C2 β), 53.37 (s, C2 α), 51.72 (s, C8 β), 51.65 (s, C8 α).

Synthesis of Ac_4ManNAz (1,3,4,6-tetra-*O*-acetyl-*N*-azidoacetylmannosamine). ManNAz (80 mg, 0.31 mmol, 1 eq) and DMAP (*N*,*N*-dimethylaminopyridine, catalytic amount) were dissolved in pyridine (5 mL). Acetic anhydride (189 mg, 1.85 mmol, 6 eq) was added dropwise and the reaction was stirred at r.t.

for 16 h under argon. The reaction crude was concentrated under reduced pressure and purified by silica flash column chromatography (50 µm, 40 g, cyclohexane/ethyl acetate 2:1 v/v). Fractions containing the product were gathered and concentrated under reduced pressure, affording Ac_4ManNAz as a white powder.

^{1}H NMR (300 MHz, $CDCl_3$) δ = 6.64 (d, *J* = 8.9 Hz, 1H, H1), 6.56 (d, *J* = 9.2 Hz, 1H, H1′), 6.05 (d, *J* = 1.9 Hz, 1H, NH′), 5.89 (d, *J* = 1.6 Hz, 1H, NH), 5.35 (dd, *J* = 10.1 Hz, 4.2 Hz, 1H, H3′), 5.22 (t, *J* = 10.1 Hz, 1H, H4′), 5.17 (t, *J* = 9.8 Hz, 1H, H4), 5.06 (dd, *J* = 9.8 Hz, 3.8 Hz, 1H, H3), 4.73 (ddd, *J* = 8.9 Hz, 3.8 Hz, 1.6 Hz, 1H, H2), 4.62 (ddd, *J* = 9.3, 4.2, 1.9, 1H, H2′), 4.30–4.00 (m, 5H, H5' + H6 + H6' + CH_2-N_3 + CH_2-N_3′), 3.82 (ddd, *J* = 9.8 Hz, 4.5 Hz, 2.5 Hz, 1H, H5), 2.24–1.93 (m, 12H, OAc + OAc′).

^{13}C NMR (75 MHz, $CDCl_3$) δ = 170.67–166.55 (m, CO OAc+C1+C1′), 91.31 (s, CN), 90.26 (s, CN′), 73.40 (s, C5), 71.45 (s, C3), 70.27 (s, C5′), 68.86 (s, C3′), 65.10 (s, C4′), 64.93 (s, C4), 61.75 (s, C6′), 61.67 (s, C6), 52.62 (s, CH_2-N_3), 52.45 (s, CH_2-N_3′), 49.72 (s, C2), 49.26 (s, C2'), 20.90–20.56 (m, 4 × CH3 OAc).

Synthesis of SiaNAz (*N*-azidoacetylneuraminic acid). ManNAz (60 mg, 0.228 mmol, 1 eq) was dissolved in 600 µl PBS (pH 7.6) with sodium pyruvate (92 mg, 0.456 mmol, 2 eq) in a 5 ml capped tube. 5 units of Neu5Ac aldolase (Sigma-Aldrich, EC 4.1.3.3) were added and the reaction was stirred at 37 °C for 16 h. The formation of the product was checked by TLC (1-propanol, NH_3, H_2O 6:1:2,5) with resorcinol as revelating agent. The reaction crude was purified by anion exchange chromatography (BioRad, AG1X8; H_2O 5CV, NH_4HCO_3 0.05 M 5CV, NH_4HCO_3 0.2 M 5CV). Fractions containing the product were identified by TLC as previously then gathered and concentrated under reduced pressure. The white powder obtained was desalted using size exclusion chromatography (P2) with ultrapure water as eluent. Fractions containing the product were gathered, concentrated under reduced pressure then freeze-dried, affording the product as a white powder (20.6 mg, 0.059 mmol, 26%).

^{1}H NMR (300 MHz, D_2O) δ 4.09–3.88 (m, 5H, H5 + H4 + H3 + H2), δ 3.78 (dd, *J* = 11.2 Hz, 2.2 Hz, 1H, $H8_{cis}$), δ 3.65 (ddd, *J* = 8.5 Hz, 4.1 Hz, 1.2 Hz, 1H, H7), δ 3.55 (dd, *J* = 11.2 Hz, 5.8 Hz, 1H, $H8_{trans}$), δ 3.45 (dd, *J* = 9.2 Hz, 0.9 Hz, 1H, H6), δ 2.18 (dd, *J* = 12.5 Hz, 4.8 Hz,1H, $H1_{eq}$), δ 1,75 (dd, *J* = 12.4 Hz, 11.8 Hz, 1H, $H1_{ax}$).

^{13}C NMR (75 MHz, D_2O) δ 70.8 (CH, C5), δ 69.82 (CH, C7), δ 68.43 (CH, C6), δ 67.01 (CH, C2), δ 63.10(CH_2, C8), δ 52.21(CH, C3), δ 51.9 (CH_2, C4), δ 39.32 (CH_2, C1).

Cell culture. All the cell lines were purchased from ATCC. HT29 cells (derived from a colon adenocarcinoma), HCT116 cells (derived from a colon carcinoma), HeLa (derived from a cervix adenocarcinoma) and HEK293 human cells (derived from a human embryonic kidney) were maintained in Dulbecco's modified Eagle's medium (DMEM, 4.5 g/L glucose) (Bio West, Nuaille, France). CCD841CoN cells (primary cells from fetal normal colon) were maintained in Eagle's Minimum Essential Medium (EMEM, 1 g/L glucose) (Lonza). All media were supplemented with 10% (v/v) fetal calf serum and 2 mM L-Glutamine (Lonza, Levallois-Perret, France). Cells were cultured at 37 °C in a 5% CO_2-enriched humidified atmosphere.

Metabolic incorporation of chemical reporters. CCD841CoN cells were seeded in 6-well plates (12×10^4 cells/well). HT29 (10^5/well), HCT116 (8×10^4/well), HEK293 (5×10^4/well) and HeLa (3.5×10^4/well) were seeded in 12-well plates. One day later, ManNAc (500 µM, used as negative control) or the chemical reporter was added in cell culture medium (500 µM for all compounds except for Ac_4ManNAz at 50 µM) for the indicated times. Then, cell monolayers were rinsed once with ice-cold PBS and cells were gently scrapped in 1 mL DPBS$^{-/-}$ (DPBS without calcium and magnesium, Gibco™) before centrifugation in microtubes at 600 g and 400 g for 5 min at 4 °C for respectively CCD841CoN and cancer cells. Cells were then gently resuspended in 20 µL PBS.

Click chemistry reaction. For Western blotting, CuAAC was performed onto cells using either 50 µM azido-PEG3-biotin or alkyne-PEG4-biotin diluted in a freshly prepared click buffer containing 100 mM K_2HPO_4, 150 µM $CuSO_4$ (stock solution 15 mM), 300 µM BTTAA (stock solution 6 mM), 2.5 mM sodium ascorbate (stock solution 100 mM) (80 µl/tube). For the negative control, click reaction was performed on cells cultured with ManNAc. After the biorthogonal labelling one hour at r.t., 1 mL DPBS was added in the microtube, cells were centrifuged (1000 g for 5 min, at 4 °C) and then lysed in 35 µl of 2X Laemmli buffer. Samples were boiled before SDS-PAGE. For flow cytometry analysis, CuAAC was performed 45 min at r.t. in the dark using a fluorescent probe (10 µM AlexaFluor® 594-picolyl-azide or 15 µM alkyne-AlexaFluor® 488) in the click buffer. Cells were then washed three times in 1 mL DPBS and centrifuged (600 g for 5 min, at 4 °C). Finally, cells were resuspended in 100 µL DPBS and kept in ice until flow cytometry analysis.

Protease treatment. To cleave plasma membrane proteins prior to the Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC), harvested and pelleted cells were treated with 0.1% (w/v) bromelain (B4882, Sigma-Aldrich) for 15 min at 37 °C in 100 µL or 150 µL of DPBS-0.1% (w/v) BSA, for respectively CCD841CoN and cancer cell lines[45]. As a negative control, cells were incubated without the protease in the same buffer. After incubation, cells were washed twice in 1 mL of DPBS and centrifuged (600 g, 5 min at 4 °C). Cell pellets were gently resuspended in 20 µL PBS before CuAAC.

Extraction of glycolipids after protease treatment. Cells were cultured in T25 or T75 flasks with 500 µM ManNAc (negative control) or ManNAz for 48 h, and harvested as described above (approximatively 3×10^6 CCD841CoN cells, 10^7 HeLa cells, and 5×10^6 cells for HT29, HCT116 and HEK293 cell lines).Cells were incubated for 15 min at 37 °C with or without bromelain (0.1% in 1 mL DPBS-0.1% BSA) before CuAAC with alkyne-PEG4-biotin (50 µM in 1 mL click reaction buffer per tube). Ten percent of the samples were set aside for

SDS-PAGE and Western blotting. After the last centrifugation, cell pellets were frozen in 1 mL of water and lyophilized. Then glycolipids were extracted by the sequential addition of 1 mL $CHCl_3/CH_3OH$ (2:1, v/v), $CHCl_3/CH_3OH$ (1:1, v/v), and $CHCl_3/CH_3OH/H_2O$ (1:2:0.8), followed by centrifugation (2500 g, 20 min)[25]. Combined supernatants were dried under a nitrogen stream and the resulting glycolipids were dissolved in 100 µL CH_3OH. PVDF membrane (Amersham™ Hybond™ P 0.45 PVDF, Dominique Dutscher, Illkirch, France) was dipped in methanol for 1–2 min and allowed to dry for a few minutes at r.t. To detect labeled sialylated glycolipids (gangliosides), samples (10 µL) were blotted onto the PVDF membrane and Western blotting was performed as described below.

Cell surface labeling with lectin. After protease treatment, cells were incubated for 1 h at 4 °C with 100 µL of MAA-II-FITC or SNA-FITC, diluted at 20 µg/ml in PBS with 1% (w/v) BSA. Cells were washed three times in DPBS and resuspended in 100 µL of PBS before flow cytometry analysis.

PNGase F treatment. *N*-glycans were released from cell surface glycoproteins using PNGase F (P0704, New England BioLabs, Evry, France), according to manufacturer's instructions. Briefly, after CuAAC with the alkyne-PEG4-biotin, cell pellets were resuspended in water (10 µL per 10^4 cells). 10× denaturing buffer was added to each sample and microtubes were incubated at 95 °C for 10 min. PNGase F digestion was performed on 10 µL of denaturated proteins (~ 30 µg) by adding 2 µL of 10× reaction buffer, 2 µL of 10% NP-40 solution, 1 µL of PNGase F and 4 µL of water (final volume 20 µl). After incubation at 37 °C for one hour, reaction was stopped by adding 5× Laemmli buffer and samples were heated at 95 °C for 7 min before SDS-PAGE.

Western blotting. Proteins were separated by 8% SDS-PAGE and transferred onto nitrocellulose membrane (Amersham™Protran™ supported 0.45 µm NC, Dominique Dutscher, Illkirch, France). Membranes were stained with Ponceau Red (0.1% (w/v) Red Ponceau S in 5% (v/v) acetic acid) to confirm equal loading, washed with TBST (15 mM Tris pH 8, 140 mM NaCl, 0.05% Tween-20) before blocking in TBST with 5% (w/v) nonfat dry milk. Membranes were incubated overnight at 4 °C with peroxidase-conjugated monoclonal mouse anti-biotin (200–032-211, Jackson ImmunoResearch) (1:10,000) in blocking buffer. After four washes with TBST, chemiluminescence signal (SuperSignal West Pico Plus, Life Technologies, Courtaboeuf, France) was detected on a CCD camera (FUSION Solo, Vilber Lourmat, France) using the FUSION-Capt Solo software for acquisition and processing merge images. Western blot with the following antibodies were done as previously described[46]: GNE (ab189927, Abcam; 1:1,000), E-cadherin (sc-7870, Santa Cruz;1:1000), GAPDH (sc-32233, Santa Cruz; 1:2000), alpha-tubulin (sc-23948, Santa Cruz; 1:2000). For detection of *O*-linked glycoproteins, membranes were incubated overnight at 4 °C in blocking buffer (1% (w/v) BSA in PBS-Tween 0.05% (v/v)) before incubation with Jacalin-FITC (1:100) in blocking buffer for one hour at r.t. After washes, fluorescence was read with ImageQuant™ LAS 4000 (GE Healthcare). *N*-linked glycoproteins were revealed using L-PHA-biotin (1:500 in PBST with 1% BSA), followed by incubation with the anti-biotin-HRP. Densitometry analysis was done using the ImageJ® software.

Flow cytometry analysis. Cell surface fluorescence was measured on a BD ACCURI C6 Plus flow cytometer (BD Biosciences, Rungis, France), either in the FL1 channel for labeling with the Alkyne-AF488 probe and lectin-FITC, or in the FL2 channel for labeling with the picolyl-azide-AF594 probe. Data were analyzed with FlowJo software (vX 0.7). Cell population was gated based on SSC/FSC dot plot to remove cell debris and doublets. The means of fluorescence intensity (MFI) were obtained for each condition (FL1A or FL2A), unspecific MFI was quantified from cells incubated with ManNAc (negative control). Raw values of MFI are indicated on the respective histograms in order to be able to compare the MFI obtained with a given fluorescent probe for the different analogs performed in the same experiment. Histograms with the relative MFIs compared to the negative controls are presented in Fig. S4 (mean ± standard deviation, n = 3).

Immunofluorescence and confocal microscopy. Cells were seeded in 24-wells plate on a 10 mm glass coverslip (CCD841CoN, 25×10^3/well; HT29, 30×10^3/well; HCT116 and HEK293, 23×10^3/well; HeLa 20×10^3/well). One day later, chemical reporters were added in the cell culture medium for the indicated times. Then cells were gently washed in DPBS and fixed in 4% (w/v) paraformaldehyde (PAF) for 30 min at r.t. Coverslips were then washed three times in PBS (5 min) before quenching with 100 mM glycine for 20 min at r.t. After three washes with PBS, cells were permeabilized with 0.5% (v/v) Triton X100 in PBS for 10 min. After three washes in PBS, coverslips were incubated for 30 min at r.t. in the dark and under gentle agitation with 100 µl of the freshly prepared click reaction mix containing the appropriate biotinylated probe (50 µM). After three washes with PBS, immunofluorescence was performed as previously described with the rabbit polyclonal anti-TGN46 (1:100, AHP1586, BioRad, Marnes-la Coquette, France) and the goat anti-Rabbit IgG Alexa Fluor 568® (1:600)[46]. Streptavidin-DyLight®488 (1:600) was added at the same time to reveal click chemistry reaction. Nuclei were stained with DAPI (50 µg/mL) and coverslips were mounted on microscope slides using a drop of mounting medium (Dako mounting medium, Agilent). Images were acquired on a Nikon A1R confocal microscope (Apo 60X, NA1.4, oil immersion) using NIS-Elements software, and further analyzed with ImageJ® software.

We developed a home-made macro on Image J® to quantify the relative fluorescence signal of streptavidin in whole cell vs. Golgi apparatus (https://doi.org/10.5281/zenodo.6778042). Briefly, a threshold value was first applied to confocal images based on the streptavidin channel intensity of each condition's control to get rid of the background signal emanating from the binding of the streptavidin into cells (negative controls, i.e. CuAAC performed on cells cultured with ManNAc for 24 h). Regions of interest (ROI) of the whole cells were generated on the thresholded images. Golgi ROIs were created using the TGN46 detection channel. These ROIs allowed

us to discriminate regions *in* and *out of* the Golgi apparatus, and to measure and export the streptavidin signal values (min, max and mean grey values, integrated densities) of these regions. Results are presented according to the mean of the integrated density (IntDen), representing the concentration of the signal of streptavidin in the ROI. Standard deviation was obtained from at least two images of two independent experiments.

Statistical analysis. Unpaired/two-tailed *t*-test was used. *p*-values were calculated and reported accordingly (*** $p \leq 0.005$, ** $p \leq 0.01$, * $p \leq 0.05$; ns, non-significant).

Data availability

The datasets generated for the image analysis are available from the corresponding author on request.

Received: 22 July 2022; Accepted: 15 December 2022
Published online: 22 December 2022

References

1. Holst, S., Wuhrer, M. & Rombouts, Y. Glycosylation characteristics of colorectal cancer. In *Advances in Cancer Research* vol. 126 203–256 (Elsevier, 2015).
2. Boyaval, F. *et al.* N-Glycomic signature of stage II colorectal cancer and its association with the tumor microenvironment. *Mol. Cell. Proteom.* **20**, 100057 (2021).
3. Ferreira, J. A. *et al.* Protein glycosylation in gastric and colorectal cancers: Toward cancer detection and targeted therapeutics. *Cancer Lett.* **387**, 32–45 (2017).
4. Haga, Y. & Ueda, K. Glycosylation in cancer: Its application as a biomarker and recent advances of analytical techniques. *Glycoconj. J.* **39**, 303–313 (2022).
5. Dobie, C. & Skropeta, D. Insights into the role of sialylation in cancer progression and metastasis. *Br. J. Cancer* **124**, 76–90 (2021).
6. Varki, A. & Schauer, R. Sialic Acids. In *Essentials of Glycobiology* (eds. Varki, A. et al.) (Cold Spring Harbor Laboratory Press, 2009).
7. Moons, S. J., Adema, G. J., Derks, M. T., Boltje, T. J. & Büll, C. Sialic acid glycoengineering using N-acetylmannosamine and sialic acid analogs. *Glycobiology* **29**, 433–445 (2019).
8. Vanbeselaere, J. *et al.* Alkynyl monosaccharide analogues as a tool for evaluating Golgi glycosylation efficiency: Application to Congenital Disorders of Glycosylation (CDG). *Chem. Commun. Camb. Engl.* **49**, 11293–11295 (2013).
9. Gilormini, P. A. *et al.* A sequential bioorthogonal dual strategy: ManNAl and SiaNAl as distinct tools to unravel sialic acid metabolic pathways. *Chem. Commun.* **52**, 2318–2321 (2016).
10. Gilormini, P. A. *et al.* Chemical glycomics enrichment: imaging the recycling of sialic acid in living cells. *J. Inherit. Metab. Dis.* **41**, 515–523 (2018).
11. Jones, M. B. *et al.* Characterization of the cellular uptake and metabolic conversion of acetylated N-acetylmannosamine (ManNAc) analogues to sialic acids. *Biotechnol. Bioeng.* **85**, 394–405 (2004).
12. Chang, P. V. *et al.* Metabolic labeling of sialic acids in living animals with alkynyl sugars. *Angew. Chem. Int. Ed.* **48**, 4030–4033 (2009).
13. Altmann, S. *et al.* Metabolic glycoengineering in hMSC-TERT as a model for skeletal precursors by using modified azide/alkyne monosaccharides. *Int. J. Mol. Sci.* **22**, 2820 (2021).
14. Büll, C. *et al.* Sialic acid glycoengineering using an unnatural sialic acid for the detection of sialoglycan biosynthesis defects and on-cell synthesis of siglec ligands. *ACS Chem. Biol.* **10**, 2353–2363 (2015).
15. Heise, T. *et al.* Metabolic oligosaccharide engineering with alkyne sialic acids confers neuraminidase resistance and inhibits Influenza reproduction. *Bioconjug. Chem.* **28**, 1811–1815 (2017).
16. Van Karnebeek, C. D. M. *et al.* NANS-mediated synthesis of sialic acid is required for brain and skeletal development. *Nat. Genet.* **48**, 777–784 (2016).
17. Hsu, T.-L. *et al.* Alkynyl sugar analogs for the labeling and visualization of glycoconjugates in cells. *Proc. Natl. Acad. Sci.* **104**, 2614–2619 (2007).
18. Yang, L., Nyalwidhe, J. O., Guo, S., Drake, R. R. & Semmes, O. J. Targeted identification of metastasis-associated cell-surface sialoglycoproteins in prostate cancer. *Mol. Cell. Proteomics* **10**, M110.007294 (2011).
19. Almaraz, R. T. *et al.* Metabolic oligosaccharide engineering with N-Acyl functionalized ManNAc analogs: Cytotoxicity, metabolic flux, and glycan-display considerations. *Biotechnol. Bioeng.* **109**, 992–1006 (2012).
20. Mbua, N. E. *et al.* Abnormal accumulation and recycling of glycoproteins visualized in Niemann-Pick type C cells using the chemical reporter strategy. *Proc. Natl. Acad. Sci. USA* **110**, 10207–10212 (2013).
21. Jiang, H. *et al.* Monitoring dynamic glycosylation in vivo using supersensitive click chemistry. *Bioconjug. Chem.* **25**, 698–706 (2014).
22. Layek, B., Sadhukha, T. & Prabha, S. Glycoengineered mesenchymal stem cells as an enabling platform for two-step targeting of solid tumors. *Biomaterials* **88**, 97–109 (2016).
23. Rigolot, V., Biot, C. & Lion, C. To view your biomolecule, click inside the cell. *Angew. Chem. Int. Ed Engl.* **60**, 23084–23105 (2021).
24. Steenackers, A. *et al.* Silencing the nucleocytoplasmic *O*-GlcNAc transferase reduces proliferation, adhesion, and migration of cancer and fetal human colon cell lines. *Front. Endocrinol.* **7**, 46 (2016).
25. Biwi, J. *et al.* OGT controls the expression and the glycosylation of E-cadherin, and affects glycosphingolipid structures in human colon cell lines. *Proteomics* **19**, e1800452 (2019).
26. Han, S.-S. *et al.* Physiological effects of Ac_4ManNAz and optimization of metabolic labeling for cell tracking. *Theranostics* **7**, 1164–1176 (2017).
27. Qin, W. *et al.* Artificial cysteine S-glycosylation induced by per-*O*-acetylated unnatural monosaccharides during metabolic glycan labeling. *Angew. Chem. Int. Ed Engl.* **57**, 1817–1820 (2018).
28. Qin, K., Zhang, H., Zhao, Z. & Chen, X. Protein S-glyco-modification through an elimination-addition mechanism. *J. Am. Chem. Soc.* **142**, 9382–9388 (2020).
29. Hinderlich, S. *et al.* UDP-GlcNAc 2-epimerase/ManNAc kinase (GNE): a master regulator of sialic acid synthesis. *Top. Curr. Chem.* **366**, 97–137 (2015).
30. Reinke, S. O., Eidenschink, C., Jay, C. M. & Hinderlich, S. Biochemical characterization of human and murine isoforms of UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase (GNE). *Glycoconj. J.* **26**, 415–422 (2009).
31. Uttamapinant, C. *et al.* Fast, cell-compatible click chemistry with copper-chelating azides for biomolecular labeling. *Angew. Chem. Int. Ed. Engl.* **51**, 5852–5856 (2012).
32. Fujitani, N. *et al.* Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers. *Proc. Natl. Acad. Sci. USA* **110**, 2105–2110 (2013).
33. Park, D. D. *et al.* Membrane glycomics reveal heterogeneity and quantitative distribution of cell surface sialylation. *Chem. Sci.* **9**, 6271–6285 (2018).

34. Batt, A. R., Zaro, B. W., Navarro, M. X. & Pratt, M. R. Metabolic chemical reporters of glycans exhibit cell-type-selective metabolism and glycoprotein labeling. *ChemBioChem* **18**, 1177–1182 (2017).
35. Pham, N. D. *et al.* Cellular metabolism of unnatural sialic acid precursors. *Glycoconj. J.* **32**, 515–529 (2015).
36. Dold, J. E. G. A. & Wittmann, V. Metabolic glycoengineering with azide- and alkene-modified hexosamines: Quantification of sialic acid levels. *Chembiochem. Eur. J. Chem. Biol.* **22**, 1243–1251 (2021).
37. Gilormini, P. A. *et al.* Improved workflow for the efficient preparation of ready to use CMP-activated sialic acids. *Glycobiology* **26**, 1151–1156 (2016).
38. Cheng, B., Xie, R., Dong, L. & Chen, X. Metabolic remodeling of cell-surface sialic acids: Principles, applications, and recent advances. *Chembiochem Eur. J. Chem. Biol.* **17**, 11–27 (2016).
39. Horstkorte, R. *et al.* Selective inhibition of polysialyltransferase ST8SiaII by unnatural sialic acids. *Exp. Cell Res.* **298**, 268–274 (2004).
40. Chinoy, Z. S. *et al.* Selective engineering of linkage-specific α2,6-N-linked sialoproteins using sydnone-modified sialic acid bioorthogonal reporters. *Angew. Chem. Int. Ed Engl.* **58**, 4281–4285 (2019).
41. Benie, A. J. *et al.* Characterization of ligand binding to the bifunctional key enzyme in the sialic acid biosynthesis by NMR: II. Investigation of the ManNAc kinase functionality. *J. Biol. Chem.* **279**, 55722–55727 (2004).
42. Van Scherpenzeel, M. *et al.* Dynamic tracing of sugar metabolism reveals the mechanisms of action of synthetic sugar analogs. *Glycobiology* **32**, 239–250 (2022).
43. Luchansky, S. J., Yarema, K. J., Takahashi, S. & Bertozzi, C. R. GlcNAc 2-epimerase can serve a catabolic role in sialic acid metabolism. *J. Biol. Chem.* **278**, 8035–8042 (2003).
44. Shajahan, A. *et al.* Mass spectrometric method for the unambiguous profiling of cellular dynamic glycosylation. *ACS Chem. Biol.* **15**, 2692–2701 (2020).
45. Shirure, V. S. *et al.* Gangliosides expressed on breast cancer cells are E-selectin ligands. *Biochem. Biophys. Res. Commun.* **406**, 423–429 (2011).
46. Masclef, L. *et al.* Cyclin D1 stability is partly controlled by *O*-GlcNAcylation. *Front. Endocrinol.* **10**, 106 (2019).

Acknowledgements

This work was supported by the CNRS and the University of Lille. J.S. is a recipient of a research grant from the Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation. V.R. is a recipient of a research grant from ANR NEURAPROBE (ANR-18-CE07-0042) (C.B.). We thank Dr Corentin Spriet (Plateformes Lilloises en Biologie et Santé (PLBS)—UAR 2014—US 41) for helpful discussions on Image J analysis and Dr Clément Delannoy for his technical help for glycolipid extraction.

Author contributions

J.S. and A.S.V.E. planned and designed the experiments. V.R. and C.L. performed the chemical synthesis. J.S., V.R., M.M. and A.S.V.E. carried out the experiments. J.S., V.R. and A.S.V.E. analyzed the data. A.S.V.E., V.R., C.L. and C.B. wrote the manuscript. T.L. revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-26521-3.

Correspondence and requests for materials should be addressed to A.-S.V.-E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

CHAPITRE 3 – DÉVELOPPEMENT DE SONDES ORGANOMÉTALLIQUES VERSATILES POUR LA LIGATION BIOORTHOGONALE

I. LIMITES DE LA CHIMIE BIOORTHOGONALE

Le milieu de la chimie bioorthogonale évolue rapidement et de nombreuses améliorations sont sans cesse effectuées sur différents aspects, tels que la résolution, la spécificité et la versatilité des outils disponibles.

La première peut être améliorée grâce à la nanoscopie (imagerie en fluorescence super-résolutive permettant de dépasser la diffraction limite d'Abbe), au moyen de sondes émettant à des longueurs d'onde bien en deçà du domaine de l'UV/visible (par exemple dans les rayons X), ou encore via des techniques telles que la microscopie électronique. Bien que ces techniques avancées apportent des informations bien plus fines sur la localisation des composés en comparaison de la microscopie à fluorescence classique, elles présentent l'inconvénient de ne pas être accessible en routine dans la plupart des laboratoires, ou d'être limitantes en ce qui concerne le choix des outils ou la préparation des échantillons (les études dynamiques *in vivo* ne sont par exemple pas possibles en ayant recours aux rayons X ou à la microscopie électronique).

Le manque de versatilité vient du type de sondes moléculaires utilisées, qui fournissent bien souvent un signal non adapté à l'emploi de méthodes analytiques corrélatives. Par exemple, les fluorophores employés en bioimagerie UV/vis classique fournissent des informations qualitatives basées sur la présence et la localisation des signaux, et des informations quantitatives relatives en comparant l'intensité des signaux dans différents organelles, pourvu que la différence d'intensité soit assez marquée. Lorsque ces variations sont plus subtiles ou qu'une quantification absolue est requise, ces outils ne sont en revanche pas adaptés. De ce fait, dans de nombreuses études, la bioimagerie doit être complétée par d'autres analyses coûteuses requérant parfois la lyse des cellules, ou encore l'extraction et la purification d'une certaine catégorie de biomolécules.

Enfin, la spécificité est probablement l'embûche principale dans le cadre du marquage métabolique. Le marquage aspécifique est principalement dû à la réactivité croisée des rapporteurs ou des sondes employés avec des composés naturellement présents dans les cellules, ou à l'impossibilité de retirer les sondes résiduelles après réaction, qui peuvent rester piégées dans les membranes par effet hydrophobe

II. OBJECTIFS ET RÉSUMÉ DU PROJET

De nombreux efforts sont effectués par la communauté de chercheurs pour améliorer les propriétés spectrales des sondes fluorescentes, notamment en concevant des sondes émettant de plus en plus « vers le rouge », à l'opposé des ultra-violets où l'autofluorescence des échantillons biologiques est la plus forte, ou en mettant au point des sondes pro-fluorescentes dont la fluorescence est exacerbée après réaction avec leur cible, permettant de mieux les distinguer du signal résiduel des sondes n'étant pas liées de manière spécifique.

Pour cet axe méthodologique de mon projet de thèse, j'ai synthétisé et mis au point des des sondes organométalliques cliquables polyvalentes à base d'Iridium(III), plus précisément des complexes d'iridium chélaté par des ligands de type phénylpyridine (Ir-ppy). Ces sondes sont équipées d'un bras espaceur et d'une fonction réactive en chimie bioorthogonale, et permettent de marquer les sialoglycoconjugués de cellules en culture par CuAAC (Figure 17).

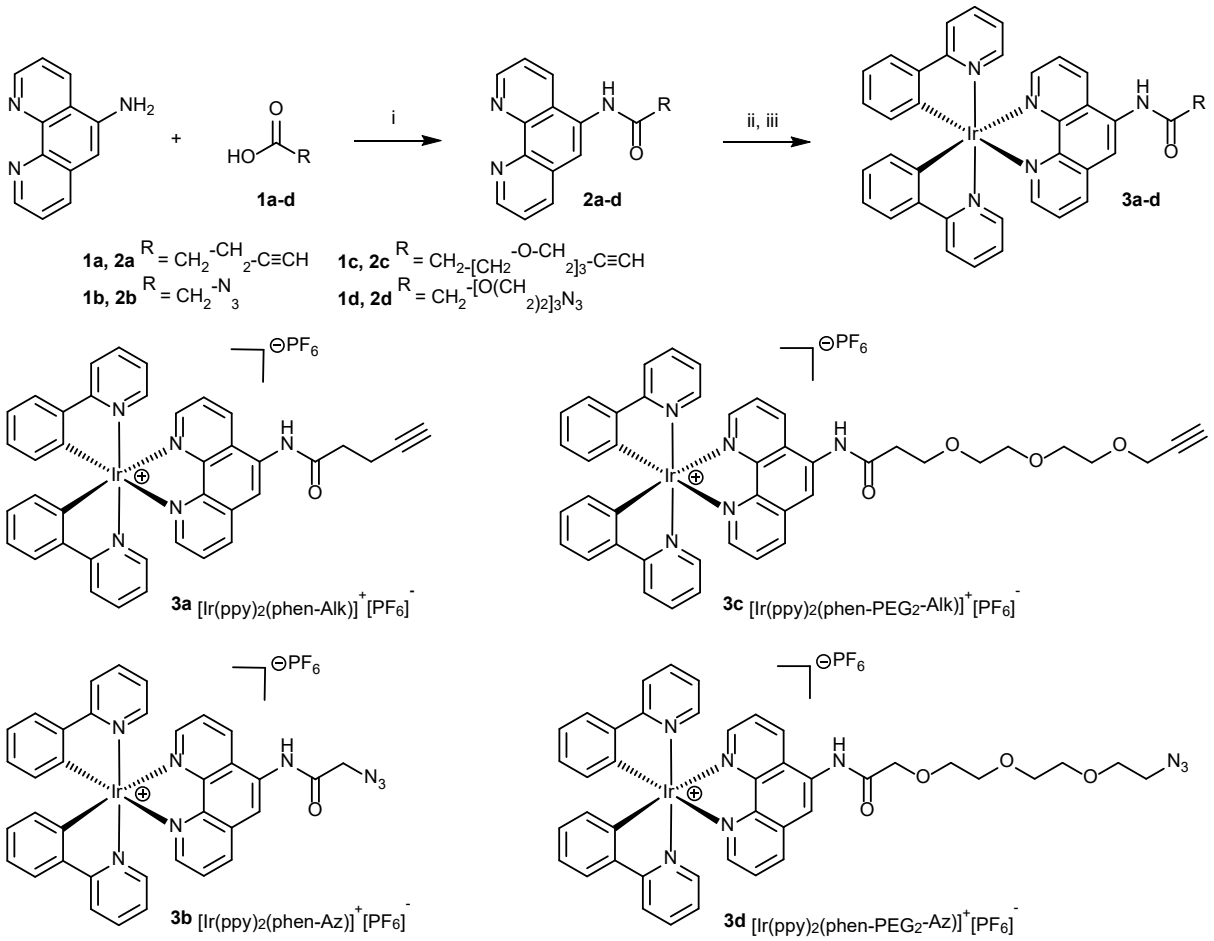


Figure 17 Synthèse et structure des sondes iridium clickables. (i) 1,10-phenanthrolin-5-amine, DCC, MeCN, *t.a.*, 20h – 2a: 95%, 2b: 78%, 2c: 74%, 2d: 72% ; (ii) $[IrCl(ppy)_2]_2$ CH_2Cl_2-MeOH, 45 à 55°C, 5 à 15h ; (iii) NH_4PF_6, *t.a.*, 15mn – 3a: 93%, 3b: 54%, 3c: 88%, 3d: 99%.

L'iridium étant totalement absent des systèmes vivants, il est possible de détecter spécifiquement le centre métallique avec une grande sensibilité et de manière quantitative par des techniques de pointe exploitant la fluorescence X et le rayonnement synchrotron. Ces sondes présentent également des propriétés de photoluminescence UV/vis intéressantes car elles possèdent un large déplacement de Stokes (écart entre le maximum d'absorption et le maximum d'émission d'un fluorophore), ce qui est avantageux car cela facilite la séparation des signaux de l'autofluorescence des cellules. De plus, elles sont très photostables et permettent d'effectuer de longues acquisitions en minimisant le photoblanchiment des sondes, et elles possèdent un temps vie de photoluminescence bien plus long que les fluorophores organiques communs. J'ai tiré profit des propriétés de fluorescence UV/vis des sondes pour mettre au point les conditions de marquage en cellules mammifères grâce à la microscopie confocale, qui a l'avantage d'être accessible

en routine au laboratoire. J'ai là encore pu automatiser une large partie de l'analyse de ces données grâce à différents scripts pour le traitement d'image, ce qui a facilité l'optimisation de nombreuses conditions (temps d'incorporation des rapporteurs, concentration en sonde, perméabilisation des cellules, étapes de rinçage).

Une fois ce protocole de marquage établi, j'ai pu l'appliquer à de nombreuses méthodes de détection en collaboration avec d'autres équipes. Ainsi, nous avons appliqué ces nouveaux outils à un panel varié de techniques de détection telles que (1) la microscopie en temps de vie de photoluminescence (photoluminescence lifetime imaging microscopy, PLIM) avec le Dr Michel Sliwa et le Dr Aude Bouchet du LASIRE à Lille, (2) la nano-imagerie synchrotron par fluorescence des rayons X (SR-XRF) avec le Dr Sylvain Bohic au European Synchrotron Radiation Facility (ESRF) à Grenoble, ou encore (3) la spectrométrie de masse couplée à plasma inductif (ICP-MS) avec la doctorante Veronica Di Battista de BASF à Ludwigshafen (Allemagne).

IMAGERIE PLIM

Les fluorophores classiquement utilisés et les composés biologiques autofluorescents ont un temps de vie de fluorescence très court (de l'ordre de quelques nanosecondes). Lors de la résolution spectrale (en fonction de la longueur d'onde) des signaux observés lors de l'acquisition d'images par microscopie à fluorescence, un compromis doit être effectué entre perte de sensibilité et risque de recouvrement des signaux. En effet, pour distinguer deux sondes émettant à des longueur d'ondes proches, il est nécessaire d'utiliser des systèmes de filtres qui vont « couper » le signal en dessous d'une certaine longueur d'onde. Plus l'émission des deux sondes sera proche, plus il sera nécessaire de couper radicalement les signaux, sans quoi le signal d'une sonde risquerait d'être détecté dans le canal de détection d'une autre (Figure 18).

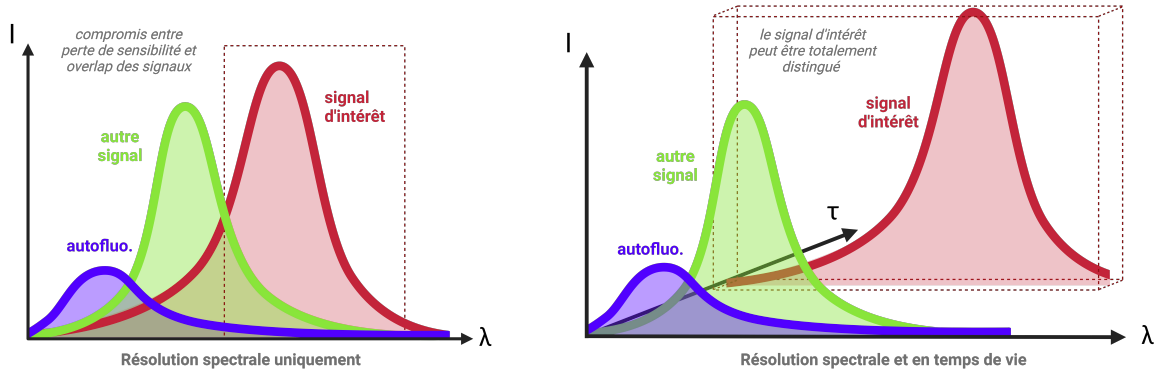


Figure 18 Illustration de l'avantage de la résolution en temps de vie pour l'observation de signaux par microscopie à fluorescence. La résolution spectrale est parfois insuffisante pour discriminer des signaux lors de marquages multiples ou en raison de l'autofluorescence des composés biologiques, ce qui peut être résolu via l'utilisation de méthodologies séparant les signaux d'intérêt selon leur temps de vie de fluorescence.

Les travaux antérieurs réalisés au laboratoire ont toujours tiré parti de l'utilisation alternée d'analogues clickable de ManNAc ou de Neu5Ac, dans des expériences de mono-marquage. Cela a permis, entre autres, de mettre en évidence leurs mécanismes d'entrée dans la cellule[103] ou bien d'identifier le mécanisme responsable d'une maladie congénitale.[61] Grâce aux sondes Ir-ppy que j'ai synthétisées, nous avons pu cette fois-ci développer une méthode de double marquage des acides sialiques. C'est la première fois que nous avons la possibilité de comparer le marquage métabolique des sialoglycoconjugués en incorporant à la fois un rapporteur ManNAc et un rapporteur Neu5Ac (un azoture et un alcyne). En effectuant deux réactions de CuAAC, l'une avec un fluorophore classique (qui a un temps de vie de fluorescence court) et l'autre avec une sonde Ir-ppy, nous pouvons exploiter la dimension temporelle supplémentaire pour séparer les signaux, sans compromettre la sensibilité. Cette approche abaisse ainsi le seuil potentiel de détection des composés marqués, et nous permet d'obtenir des signaux plus propres et fidèles par rapport à la microscopie photonique traditionnelle.

Il est courant de marquer le noyau des cellules avec le DAPI lors d'expériences de bio-imagerie. Malheureusement, cela n'est pas possible avec les sondes Ir-ppy lorsque les marquages sont observés en microscopie à fluorescence classique. En effet, le DAPI absorbe à la même longueur d'onde, il n'est donc pas possible d'exciter l'un sans exciter l'autre, et bien que la longueur d'onde d'émission maximale du DAPI soit bien plus basse que celle des sondes Ir-ppy, celle-ci est extrêmement large, ce qui rend impossible la

détection du signal de la sonde sans qu'il parasité par celui du DAPI. Cependant, le DAPI possédant un temps de fluorescence très court, nous avons là encore pu tirer profit de la PLIM, et distinguer efficacement le signal provenant du noyau des cellules de celui des sialoglycoconjugués marqués, en une seule acquisition. Ceci nous a permis de contourner un problème tout en développant une élégante méthode de double détection à partir d'une excitation laser unique. Il s'agit des premiers exemples d'application de la PLIM à la MOE. Nous prévoyons de soumettre ces travaux pour publication dans la revue *Journal of the Americal Chemical Society*. La version pour soumission est insérée ci-après.

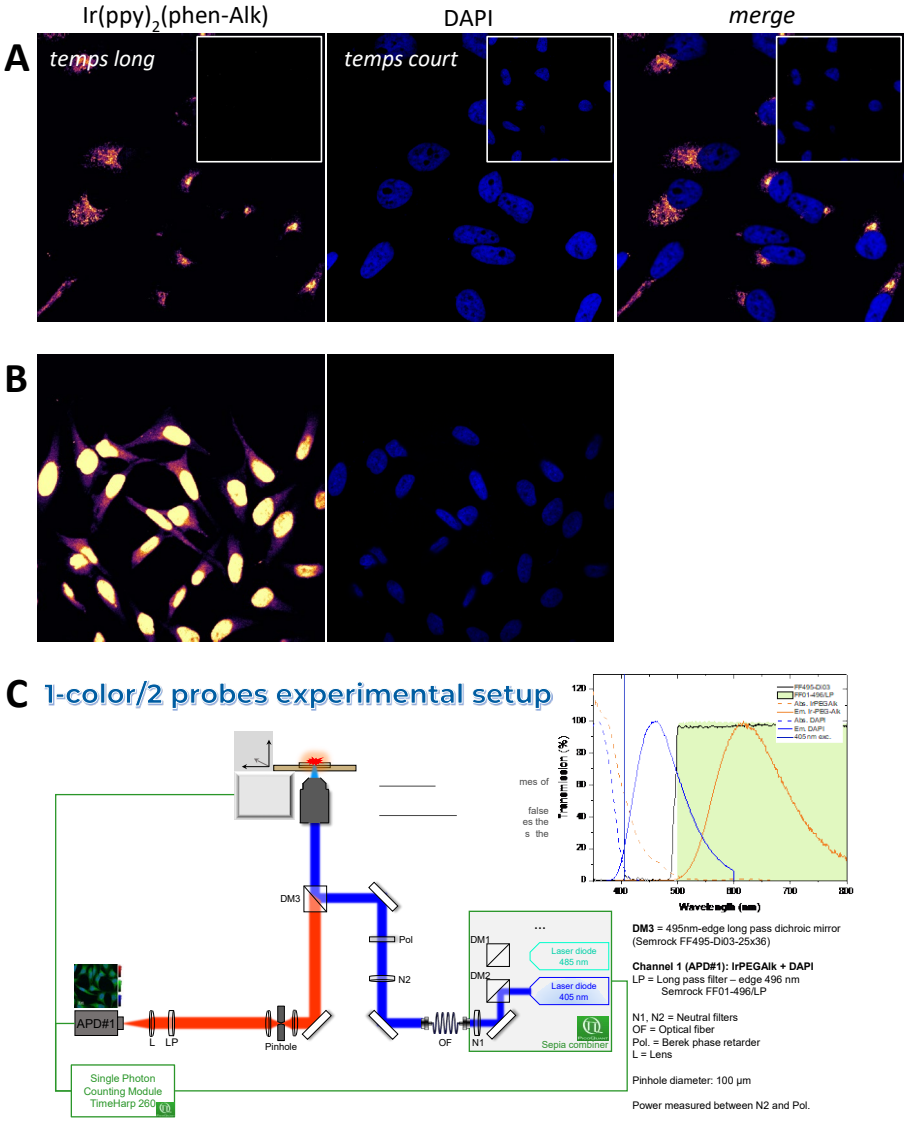


Figure 19 Expérience de double détection, excitation unique. (A) Imagerie PLIM des sialoglycoconjugués marqués par une sonde Ir-ppy et du noyau des cellules marqués au DAPI, en excitation unique (405nm). L'imagerie PLIM permet une séparation très nette des signaux. (B) Même expérience observée en microscopie confocale à fluorescence classique, le canal de détection de la sonde Ir-ppy est largement parasité par le signal du DAPI. (C) Trajet optique et montage des expériences de PLIM pour ces marquages. Une excitation laser unique à 405nm irradie les cellules, et les photons émis (>496nm) sont captés par un seul détecteur. Les signaux sont séparés uniquement sur la base de leur temps de vie de photoluminescence.

DEVELOPMENT OF BIOORTHOGONAL VERSATILE PHOTOLUMINESCENT IRIDIUM COMPLEXES FOR METABOLIC OLIGOSACCHARIDE ENGINEERING

Vincent Rigolot[1], Clémence Simon[1], Aude Bouchet[2], Veronica di Battista[3], Dmitry Karpov[4], Boris Vauzeilles[5], Sylvain Bohic[4], Michel Sliwa[2], Christophe Biot[1]* and Cedric Lion[1]*.

1: UGSF, UMR CNRS 8576, Université de Lille, France; 2: LASIRE UMR CNRS 8516, Université de Lille, France; 3: BASF SE, Ludwigshafen, Germany; 4: ID16A beamline, ESRF, Grenoble, France; 5: ICSN UPR CNRS 2301, Gif-sur-Yvette, France. * : co-corresponding authors.

ABSTRACT

Herein, we report the synthesis, photophysical characterization and validation of iridium(III)-polypyridine complexes functionalized for click chemistry and bioorthogonal chemistry labeling, as well as their versatile applications as probes in bioimaging studies exploiting metabolic labeling of intracellular glycoconjugates. The designed probes are easily conjugated to chemical reporters within cells by CuAAC ligation in a specific manner and display attractive photophysical properties in the UV-visible range. They are indeed highly photostable and emit in the far-red to near-IR region with long lifetimes and large Stoke's shifts. We demonstrate that they can be efficiently used to monitor nascent sialylated glycoconjugates in bioorthogonal MOE studies with a varied panel of optical and non-optical techniques, namely conventional UV-Vis laser scanning confocal microscopy (for routine purposes), UV-Vis time-resolved luminescence imaging (for specificity and facilitated multiplexing with nano-environment sensitivity), synchrotron radiation based X-ray Fluorescence nanoimaging (for high resolution, elemental mapping and quantification *in situ*) and inductively coupled plasma mass spectrometry (for routine quantification on cell populations with high statistical confidence). The synthesized Ir(III) complexes were utilized in single labeling experiments, as well as in dual click-labeling experiments utilizing two distinct monosaccharide reporters relevant to the same metabolic pathway.

INTRODUCTION

The advent of click chemistry and bioorthogonal chemistry combined to metabolic labeling has revolutionized chemical biology over the past two decades, culminating in the recent award of the Nobel prize in Chemistry to Carolyn Bertozzi, K. Barry Sharpless and Morten Meldal. The major reactions used *in situ* to detect biomolecules exploit π4s + π2s pericyclic mechanisms that are considerably less present in living organisms than polar reactions between nucleophiles and electrophiles, thus greatly minimizing interference with biological processes. The copper-catalyzed azide-alkyne [3+2] cycloaddition (CuAAC),[1] strain-promoted azide-alkyne [3+2] cycloaddition (SPAAC)[2] and alkene-tetrazine inverse electronic-demand [4+2] Diels-Alder cycloaddition (IEDDA)[3,4]

are the three main click chemistry reactions that exert bioorthogonal nature. Hence azide, alkyne or alkene groups have been synthetically introduced as chemical handles on metabolite reporters, which enables their covalent ligation to a molecular dye directly in the midst of a cell. Not only does the approach provide a strong alternative to genetic methods in the detection of proteins, it also gives access to various dynamic phenomena that could not easily be addressed in imaging studies, such as protein post-translational modifications (PTMs) or biomolecule metabolism tracking.[5] Numerous click-ready bioorthogonal reporters have been developed, including nucleosides and deoxynucleosides for visualizing nucleic acids, non-canonical amino-acids for visualizing newly synthesized proteins in a residue-specific or site-specific manner, or lipids for monitoring lipid metabolism or lipidation PTMs. Last but not least, a toolbox of tagged synthetic monosaccharides is now readily available to study glycosylation PTMs and glycan metabolism via metabolic oligosaccharide engineering (MOE).[6]

Pioneered by the Bertozzi group,[2,7] MOE bioorthogonal labeling strategies have raised great interest because glycosylation PTMs are crucial in countless biological processes that involve molecular recognition (*e.g.*, cell-cell interactions, host-pathogen interactions, protein localization) and impact pathologies ranging from cancer to genetic disorders to viral and bacterial infections. This method involves the incorporation of a monosaccharide chemical reporter in a specific metabolic path (*e.g.*, sialylation, fucosylation, *O*-linked β-*N*-acetylglucosaminylation), followed by biorthogonal ligation of an appropriately functionalized probe relevant to the chosen readout technique, most often a fluorescent dye for optical bioimaging. Bioorthogonal chemistry has improved our understanding of the structures, localizations and functions of glycans in eucaryotic and procaryotic cells, an on-going effort of the scientific community at the interface between chemistry and biology as much of it still remains uncharted territory.

In the context of bioimaging, the rapidly advancing field of bioorthogonal chemistry still requires improvements in key aspects that can perhaps be summarized in three words: resolution, versatility, and specificity. The first of these aspects can now be enhanced through super-resolution fluorescence imaging nanoscopy, which utilizes photonic techniques in the UV-Vis range to achieve resolutions higher than those limited by the optical diffraction barrier (*ca*. 200-250 nm).[8–10] Alternatively, improvements can be made by employing techniques involving dyes that emit at much shorter and energetic wavelengths than the UV-Vis range (such as X-rays),[11] or by utilizing non-optical techniques such as cryogenic electron microscopy (cryoEM).[12] The lack of versatility is primarily due to the type of molecular probe used, which typically provides a single read-out but may not be suitable for correlative multimodal microscopy and multianalytical methods.[13] Therefore, in many research studies, bioimaging must be supplemented with additional time-consuming bioanalytical techniques that may require cell lysis, the extraction/separation of biomolecule pools and/or the use of several

distinct labelling methods (*e.g.*, glycomic studies by LC-MS/MS). Finally, specificity is perhaps the main issue to tackle in many applications. Indeed, non-specific labelling may occur in bioimaging studies due to cross-reactivity of either reporters or probes with irrelevant off-target sites in the cell, or to non-covalent entrapment of unreacted dye molecules[6]. This increases non-specific fluorescence background and diminishes the specific signal-to-noise ratio thus hampering the experiment by decreasing its sensitivity, and in some cases, it might even generate aberrant false positive signals and bias the results when seeking identification of precise structures or quantification.[14]

If many efforts are constantly being made to improve bioorthogonal monosaccharide reporters and reactions in terms of specificity, cell entry, level of incorporation, reporter group innocuity, stability or kinetics,[15–21] the most promising advances also stem from the development of improved click-ready molecular probes. For example, near-IR dyes,[22] long lifetime dyes, dyes that rely on Forster's resonance energy transfer (FRET) interactions or quenching[23] or fluorogenic dyes can help circumventing autofluorescence and non-specific background issues.

In the search for new emissive probes to improve versatility, specificity and resolution in bioimaging studies, luminescent d^6 transition metal complexes such as iridium(III)-polypyridines are very attractive candidates. Indeed, the physico-chemical properties of d^6 metal ions and strong π-accepting ligands result in various charge transfer mechanisms that accelerate intersystem crossing from singlet to triplet excited states.[24–26] Thus, using UV/Vis excitation they often exhibit red-shifted long luminescence lifetimes (formally, $T_1 \rightarrow S_0$ phosphorescence as the major signal contributor) with large Stoke's shifts, both exploitable aspects to improve signal-to-noise ratio and isolate specific signal in multiplexed settings. Luminophores with large Stoke's shifts (hundreds of nm) indeed have the double benefit of displaying no self-quenching and of reducing signal from autofluorescence.[27] Moreover, it should also be possible to utilize a single irradiation to simultaneously identify two dyes that have overlapping absorption spectra but distinct emission spectra. As an example, a single excitation pulse from a standard 405 nm laser may generate 2-channel images with a blue/green emitting organic probe and a far-red/near-IR emitting organometallic probe. Nevertheless, the true potential of Ir(III)-based luminophores is fully realized when their lifetimes in the range of hundreds of nanoseconds to microseconds are exploited through time-resolved luminescence imaging. This technique enables the complete elimination of the sample's short-lived autofluorescence by employing time-gating. As a result, the sensitivity of the experiment and the specificity of the detected signal are significantly enhanced. Moreover, in emission lifetime imaging microscopy,[28] signals originating from multiple dyes with overlapping emission spectra but distinct lifetimes can be easily distinguished, thereby further amplifying the multiplexing capabilities. In this regard, Ir(III)-polypyridine complexes present the advantage of easily modulating their photophysical parameters (absorption, emission and lifetime)

when compared to other d^6 metal ions such as Ru(II), Os(II) or Re(I). In addition, iridium is practically absent from living systems and presents emission rays that do not overlap with other endogenous elements, which also enables elemental mapping by X-ray fluorescence (XRF).[11] This offers an opportunity for nano-imaging with enhanced resolutions compared to optical diffraction limited confocal UV-Vis microscopy, as well as straightforward quantification by XRF or by mass spectrometry.

Although iridium(III) luminophores have been attracting much attention in the bioimaging field in the past ten years due to their red shifted long emission lifetime, surprisingly such probes are scarcely used to label biomolecules with bioorthogonal click chemistry.[29] Lo and co-workers, who pioneered the development of iridium sensors for bioconjugation on many application types,[24] proposed the very first proof-of-concept for phosphorescent bioorthogonal reagents to detect azide-modified biomolecules at the cell surface, equipped with a dibenzocyclooctyne moiety for use with SPAAC labelling on living cells.[30] Highly lipophilic probes entering the cell led to strong non-specific signal and high cytotoxicity, which the authors solved by appending a carboxylate group to prevent cell entry and visualize only cell-surface glycans. The same group also developed tetrazine-based probes with luminogenic behaviour after IEDDA ligations.[31–33] Focusing on spectral properties of the probes, these reports did not exploit their lifetime, and to our knowledge only two examples used time-resolved imaging so far: in 2017, an iridium-alkynylbipyridine complex was described for photoluminescence lifetime imaging microscopy (PLIM) of protein pools after azido-homoalanine incorporation. In this article, the lifetime of the Ir-alkyne probe was significantly shortened upon formation of the triazole cycloadduct by CuAAC, thus allowing the authors to elegantly distinguish reacted probes from unreacted probes in single-labeling experiments.[34] More recently, a FRET-based sensing approach involving SPAAC and IEDDA was found to enable detection of caspase-3 activity with PLIM.[35]

In this work, we report the synthesis, photophysical characterization and validation of click-ready azide- and alkyne-functionalized iridium(III)-polypyridine complexes, and we showcase their potential as versatile probes in studies exploiting bioorthogonal metabolic labeling of glycoconjugates with a combination of optical and non-optical techniques. The designed heteroleptic Ir(III) complexes are of the type $[Ir(C\hat{}N)_2(N\hat{}N)]^+[PF_6]^-$, in which cyclometalating 2-phenylpyridine (ppy) is the C^N backbone ligand and a phenanthroline-based N^N ligand (phen) bears a spacer arm equipped with an azide or a terminal alkyne reactive moiety. These probes are specifically conjugated to chemical reporters inside cells by CuAAC ligation and display attractive photophysical properties in the UV-visible range (*i.e.*, high photostability, large Stoke's shift, long lifetime, emission in the far-red to near-IR region). We demonstrate that they can be efficiently used to monitor nascent sialylated glycoconjugates in bioorthogonal MOE studies with a varied panel of optical and non-optical techniques (Figure 1), namely conventional laser scanning confocal microscopy (LSCM) for routine purposes, time-resolved

photoluminescence lifetime imaging microscopy (PLIM) for facilitated multiplexing and high signal specificity with nano-environment sensitivity in multicolor settings), synchrotron radiation based X-ray Fluorescence nanoimaging (XRF-NI) for high resolution, elemental mapping and quantification *in situ* and inductively coupled plasma mass spectrometry (ICP-MS) for routine quantification on cell populations with high statistical confidence.

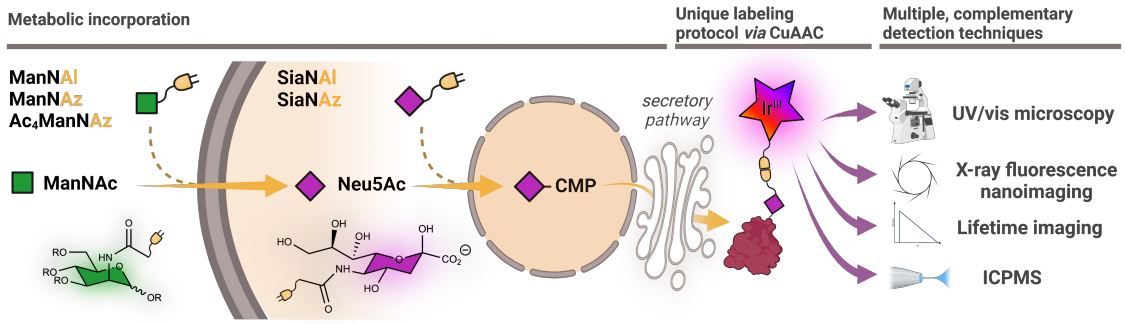


Figure 1. *Chemical reporter strategy using unprotected N-acetylmannosamine (ManNAc) or N-acetylneuraminic acid (Neu5Ac) derivatives as chemical reporters and Ir(III) luminophores as molecular probes allows time-resolved PLIM bioimaging and X-ray fluorescence nanoimaging in addition routine confocal microscopy. The probes may also be used as mass tags in techniques such as ICP-MS.*

RESULTS AND DISCUSSION

DESIGN AND SYNTHESIS OF CLICK-READY IRIDIUM(III) LUMINOPHORES

We envisaged *bis*-cyclometalated complexes of the $[Ir(N^{\wedge}C)_2(N^{\wedge}N)]^+$ structure family, as they are known to combine intense long-lived luminescence, good photostability and cell permeability while being more easily synthesized than their N_6-coordinated counterparts. The auxiliary ligand can be modified as desired to fine-tune its photophysical parameters[36] or to functionalize the complex for biological targeting [37,38] or for specific reactivity, in our case click chemistry. We thus set out to synthesize Ir(III) complexes **3a-d** with ppy as the main cyclometalating ligand, and a phenanthroline functionalized with amide-linked spacer arms and either azide or terminal alkyne linkers (Figure 2), in order to label sialoglycoconjugates inside cells after metabolic incorporation of modified sialic acid precursors. As opposed to the two ratiometric or fluorogenic approaches cited above, our objective here was to generate robust, versatile large Stoke's shift probes with long lifetime and photophysical properties that are not influenced by the bioorthogonal reaction used for labeling. Hence, the click-reactive moiety (azide or terminal alkyne) was not designed to be conjugated to the luminophore. Owing to our previous experience in intracellular sialylation studies, we considered the introduction of a small PEG spacer arm to improve solubility and allow enough flexibility for the azide/alkyne linker to react efficiently with the metabolically engineered glycans. We opted for CuAAC as the bioorthogonal

reaction, as it still remains the best candidate for detecting intracellular glycans in fixed cells. Indeed, the potential toxicity of copper is irrelevant in this context as CuAAC ligation of the probe is carried out after cell fixation and permeabilization, while its kinetics and specificity are generally superior to SPAAC MOE approaches for intracellular applications.[6] IEDDA MOE approaches are promising[39] but still suffer from the higher steric hindrance of strained alkene reporters.

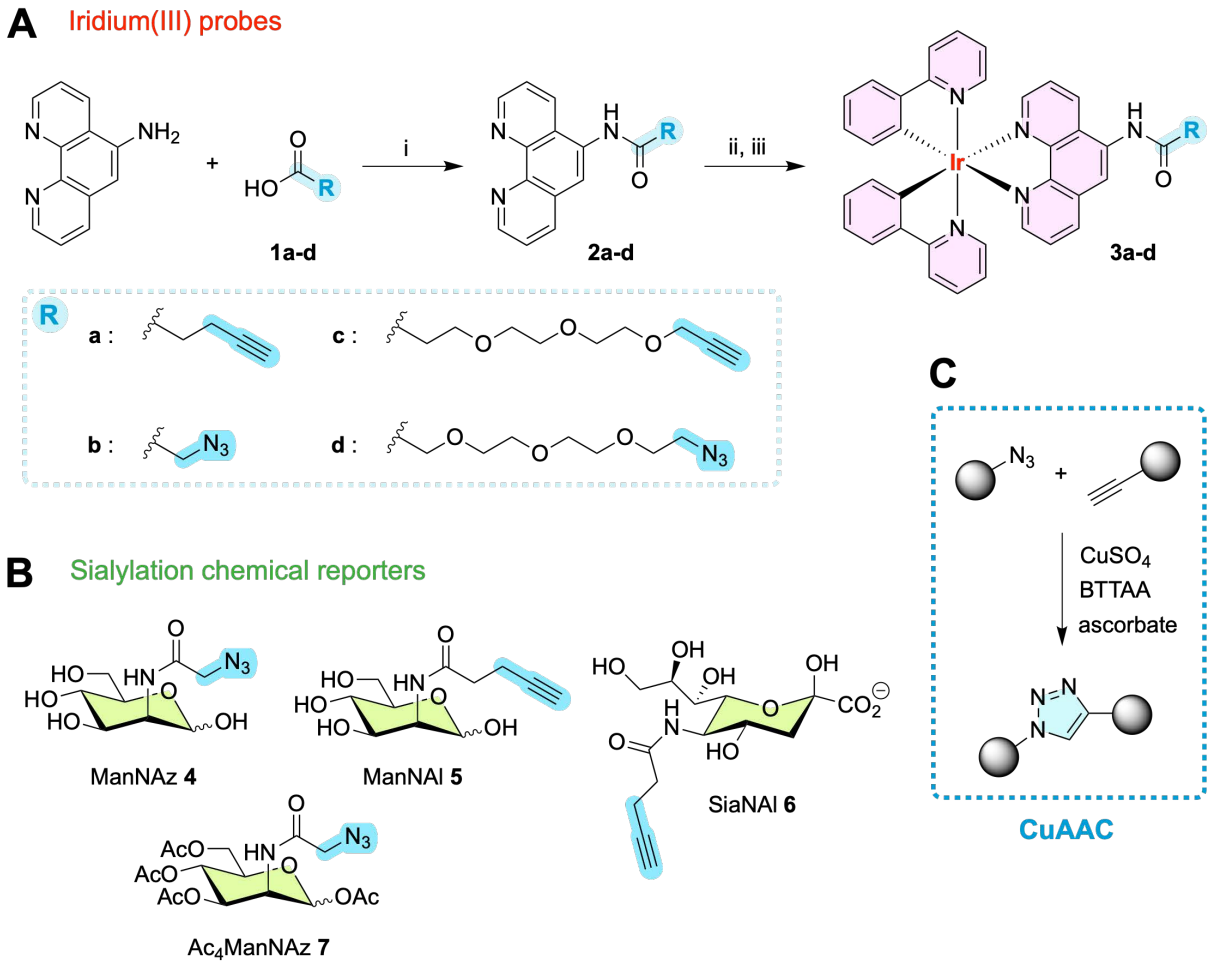


Figure 2. *A) Synthesis of Ir(III) luminophores* ***3a-d****. Reagents and conditions: (i) 1,10-phenanthrolin-5-amine, DCC, MeCN, rt, 20h.* ***2a****: 95%,* ***2b****: 78%,* ***2c****: 74%,* ***2d****: 72%. (ii) $[IrCl(ppy)_2]_2$, CH_2Cl_2/MeOH, 45 to 55°C, 5 to 15h ; (iii) NH_4PF_6, rt, 15mn.* ***3a****: 93%,* ***3b****: 54%,* ***3c****: 88%,* ***3d****: 99%. B) Structure of sialylation chemical reporters.* ***4****,* ***5*** *and* ***7*** *are derivatives of N-Acetylmannosamine and enter the metabolic pathway at early stages in the cytosol (GNE/MNK) while SiaNAl* ***6*** *enters the pathway at the activation stage (CSS) in the nucleus. C) CuAAC allows the ligation of probes onto metabolically incorporated reporters within cells.*

The synthesis of photoluminescent complexes **3a-d** and of chemical reporters *N*-azidoacetyl-D-mannosamine (ManNAz **4**), *N*-pentynoyl-D-mannosamine (ManNAl **5**) and *N*-pentynoylneuraminic acid (SiaNAl **6**) is detailed in supporting information. In brief, carboxylic acids **1a-d** were reacted with 1,10-phenanthrolin-5-amine and converted to the desired N^N bidentate auxiliary ligands **2a-d** using dicyclohexylcarbodiimide (DCC) as activating agent in acetonitrile. **1a-d** were added in a five-fold excess

to compensate for the low nucleophilicity of the amine. Complexes $[Ir(ppy)_2(phen\text{-}Alk)]^+[PF_6]^-$ **3a**, $[Ir(ppy)_2(phen\text{-}Az)]^+[PF_6]^-$ **3b**, $[Ir(ppy)_2(phen\text{-}PEG_2\text{-}Alk)]^+[PF_6]^-$ **3c** and $[Ir(ppy)_2(phen\text{-}PEG_2\text{-}Az)]^+[PF_6]^-$ **3d** were obtained after reaction of two equivalents of N^N ligands **2a-d** with one equivalent of the cyclometalated dimer $[IrCl(ppy)_2]_2$ in CH_2Cl_2/MeOH (2:1, v/v) followed by chloride anion exchange with ammonium hexafluorophosphate. Functionalized Ir(III) complexes **3a-d** were then purified by silica column chromatography prior to characterization by ^1H and ^{13}C NMR, MALDI-ToF mass spectrometry and HPLC.

PHOTOPHYSICAL CHARACTERIZATION

The photophysical characterization of probes **3a-d** was conducted in water and toluene in the presence or absence of oxygen, in order to measure their absorption/emission spectra and to determine their lifetimes and emission quantum yields.[40] The steady state absorption spectra of all four probes exhibit identical characteristics (Figure 3A) and are indicative of a mixed intraligand (ILCT) and metal-ligand charge transfer (MLCT) band. They cover a broad range from 250 nm to 500 nm, with an absorbance maximum at about 260 nm, a second maximum at 375nm and a shoulder until 500 nm. A broad emission spectrum (excitation at 375 nm) is measured covering the 500 to 800 nm range with a maximum at 615nm (large Stoke's shift of 240 nm / 10407 cm^{-1}).

As the emission spectrum is assigned to a triplet ^{3}MLCT band, we investigated whether the presence of oxygen might result in emission quenching. Therefore, the quantum yields (Φ) and lifetimes (τ) of **3a-d** were measured in both aerobic and degassed conditions. The literature also documented the potential impact of polarity, leading us to determine these photophysical properties in toluene as well as in aqueous media.

In water, the quantum yields and emission lifetimes of **3a-d** were revealed to be quite insensitive to oxygen. Φ was measured at about 3-4%, (Table 1), reasonable values corresponding to those previously described in the literature for similar structures, *i.e.* about 1-2%.[31,32,41] The emission decay τ in water is nearly monoexponential with a lifetime of *ca*. 70-80 ns. Interestingly, the lifetime in toluene was found to be longer and strongly dependent on the presence of oxygen. In degassed conditions, τ was indeed measured at *ca*. 700 ns, a 7-fold increase when compared to the value in aerobic conditions (*ca*. 100 ns). In accordance with the literature, these results point out the sensitivity of the triplet excited state to polar environments as well as to oxygen. On average, probes **3c** and **3d** bearing a PEG_3 spacer arm were found to have slightly higher τ values (ca. 15%) than probes equipped with a shorter, more hydrophobic alkyl spacer arm **3a** and **3b**. In cell labeling experiments, the obtained signal comprises two components: a fast lifetime (1-3 ns) corresponding to cell autofluorescence, and a long lifetime corresponding to the Ir(III) complex. The long lifetime of **3c** and **3d** (*ca*. 400 ns) was consistently found

to be longer than in non-degassed solutions, in between the value obtained in solution for aerobic conditions and the value in degassed toluene. This τ value is homogeneous across the cell, regardless of the localization. The complex heterogeneous polarity microenvironments in cells may shield the triplet excited states from the effect of water thus leading to longer average lifetimes when compared to experiments in aqueous solutions, particularly when considering that the tagged *N*-glycoconjugates are majoritarily anchored in hydrophobic membranes. In addition, the fact that an anti-oxidant anti-fading mounting medium is used for imaging experiments might also aid in averting quenching and contribute to lengthening the lifetime of Ir(III) dyes.

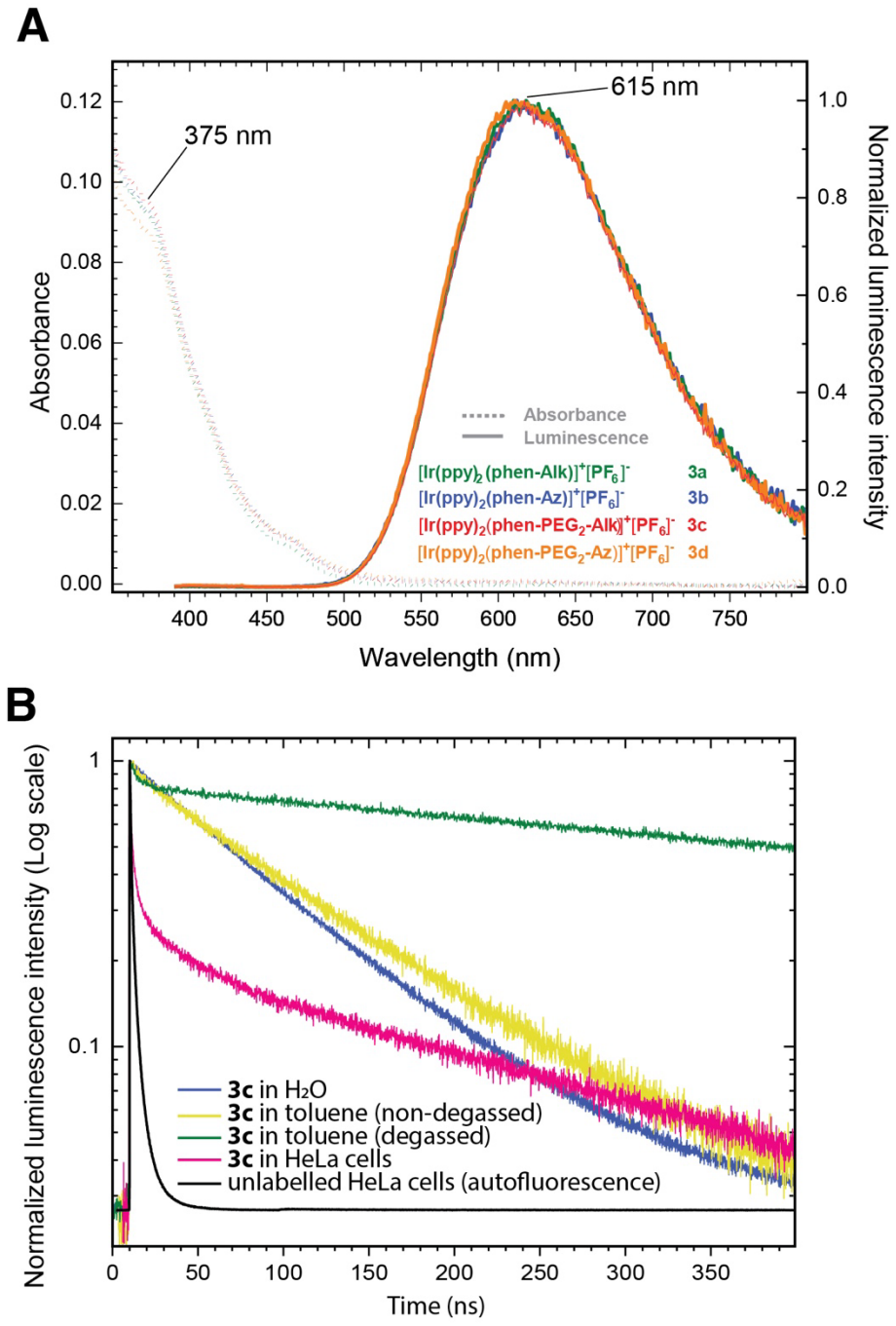


Figure 3. *A) Absorbance (dashed line) and normalized luminescence upon excitation at 375 nm (solid line) of probes* ***3a-d*** *in water. B) Emission decay at 375 nm excitation for* ***3c*** *in water and toluene (non-degassed and degassed by bubbling argon), and in cells, versus emission decay of cell autofluorescence. Note: the decay in cells presents a fast component due to cell autofluorescence and a long component due to the Ir(III) probe.*

	τ H_2O	τ H_2O *degassed*	τ toluene	τ toluene *degassed*	τ_{av} in cells	Φ H_2O
3a	68.8 ± 0.2 ns	71.4 ± 0.3 ns	98.0 ± 0.3 ns	728.0 ± 0.6 ns	-	3.8%
3b	67.8 ± 0.2 ns	68.9 ± 0.3 ns	88.0 ± 0.2 ns	454.4 ± 0.7 ns	-	3.4%
3c	81.9 ± 0.2 ns	79.7 ± 0.3 ns	97.0 ± 1.0 ns	744.6 ± 0.5 ns	387.1 ± 5.4 ns	3.9%
3d	76.0 ± 0.2 ns	74.2 ± 0.3 ns	94.0 ± 0.2 ns	696.2 ± 0.9 ns	402.6 ± 14 ns	4.0%

Table 1. *Luminescence lifetimes τ of probes **3a-d** in water and toluene (non-degassed and degassed by bubbling argon) and in cells. Quantum yields Φ of probes **3a-d** in water.*

PROBE VALIDATION *IN VITRO*

We then investigated whether probes **3a-d** could be efficiently conjugated by CuAAC to adequately functionalized monosaccharides in aqueous medium, in conditions relevant to bioorthogonal cellular assays.[42] All four probes were reacted with a corresponding sugar reporter, namely either ManNAz **4** or ManNAl **5**. The CuAAC ligations were monitored by HPLC equipped with UV and fluorescence detectors, and the formed cycloadducts were characterized by MALDI-ToF mass spectrometry of the isolated HPLC fractions. PEGylated probes **3c** and **3d** were quickly and fully converted to the desired cycloadducts as expected, indicating fast kinetics suitable for cell culture experiments. However, we observed that the reaction of probes **3a** and **3b** was significantly advanced but still incomplete after two hours, showing conversion rates of 94.5% and 77.5%, respectively (Figure S7 and S8). Although the reactivity assessed in aqueous solution with a monosaccharide is not directly transferable to the kinetic behavior of probes reacting in the midst of cells where the hydrophobic-hydrophilic balance varies and the local concentration of reporter subunits is highly heterogeneous, this hinted at a potentially lesser labelling efficiency for dyes **3a** and **3b** in biological experiments.

METABOLIC LABELING OF SIALOGLYCOCONJUGATES IN CELLS

The probes were then tested in cells to monitor the biosynthesis of sialylated glycoconjugates by MOE. ManNAz **4** and ManNAl **5** are intracellularly converted to the corresponding modified sialic acid, respectively, then incorporated into nascent sialoconjugates in the golgi apparatus after being activated as cytidine monophosphate nucleotide sugars in the nucleus. During their life cycle, the formed

sialoconjugates are then generally exported to the cell membrane via the secretory pathway, and can be re-internalized during pinocytosis and recycled. Whereas per-*O*-acetylated monosaccharides are commonly used for MOE in mammalian models, it has been shown that per-*O*-acetylation generates non-specific signal and false positives.[18] It may also lead to metabolic perturbation due to partially acetylated forms, thus for reporters that are capable of entering cells by active transport such as ManNAc or Sia derivatives, we tend to use unprotected reporters in studies aimed at investigating intracellular metabolic paths.[6] **4** or **5** were metabolically incorporated in HeLa cells that were then fixed with paraformaldehyde and permeabilized with Triton X.

Cells incubated with ManNAz **4** were then reacted with alkyne-functionalized $[Ir(ppy)_2(phen\text{-}Alk)]^+[PF_6]^-$ **3a** or $[Ir(ppy)_2(phen\text{-}PEG_2\text{-}Alk)]^+[PF_6]^-$ **3c** using the copper sulfate / ascorbate catalytic system in the presence of copper tetravalent ligand BTTAA, adapting our previous protocol, while cells incubated with ManNAl **5** were labelled with azide dyes $[Ir(ppy)_2(phen\text{-}Az)]^+[PF_6]^-$ **3b** or $[Ir(ppy)_2(phen\text{-}PEG_2\text{-}Az)]^+[PF_6]^-$ **3d**.

Azide **3a** and alkyne **3b** at concentrations ranging from 1 to 100 µM did not lead to efficient marking of sialylated structures (Figure 4 and Figure S1 A). Although HPLC monitoring of the reaction in solution indicated a slower but decent rate, we had expected these probes to bind to sugar reporters in cells. An analogous Ir-alkyne complex with no spacer arm was indeed previously reported to bind to azide-tagged protein pools after incubating cells with 2 mM azidohomoalanine in methionine-depleted medium.[34] However, in our case the non-specific background signal was consistently too high to detect any potential specific labeling. The short, rigid spacer arms of **3a** and **3b** combined to the hydrophobic nature of the complex core might not confer enough flexibility for the reactive handle to efficiently access reporter groups in the highly polar microenvironment of glycan chains. The issue might also stem from the low incorporation rate of monosaccharide reporters that leads to a lower concentration of tagged biomolecules in MOE than for residue-specific protein labeling.

Conversely, probes **3c** and **3d** bearing a more flexible PEG_3 spacer arm capable of accepting hydrogen bonds provided images with efficient specific labeling using a conventional laser scanning confocal microscope (Figure 4 and Figure S1 B). The obtained pattern was fully consistent with intracellular sialylation studies previously evidenced with small organic fluorophores such as tetramethylrhodamine (TAMRA).[43,44] The CuAAC conditions were then optimized for maximum signal-to-noise ratio (Figure S2), and alkyne-functionalized probe **3c** was selected for further studies.

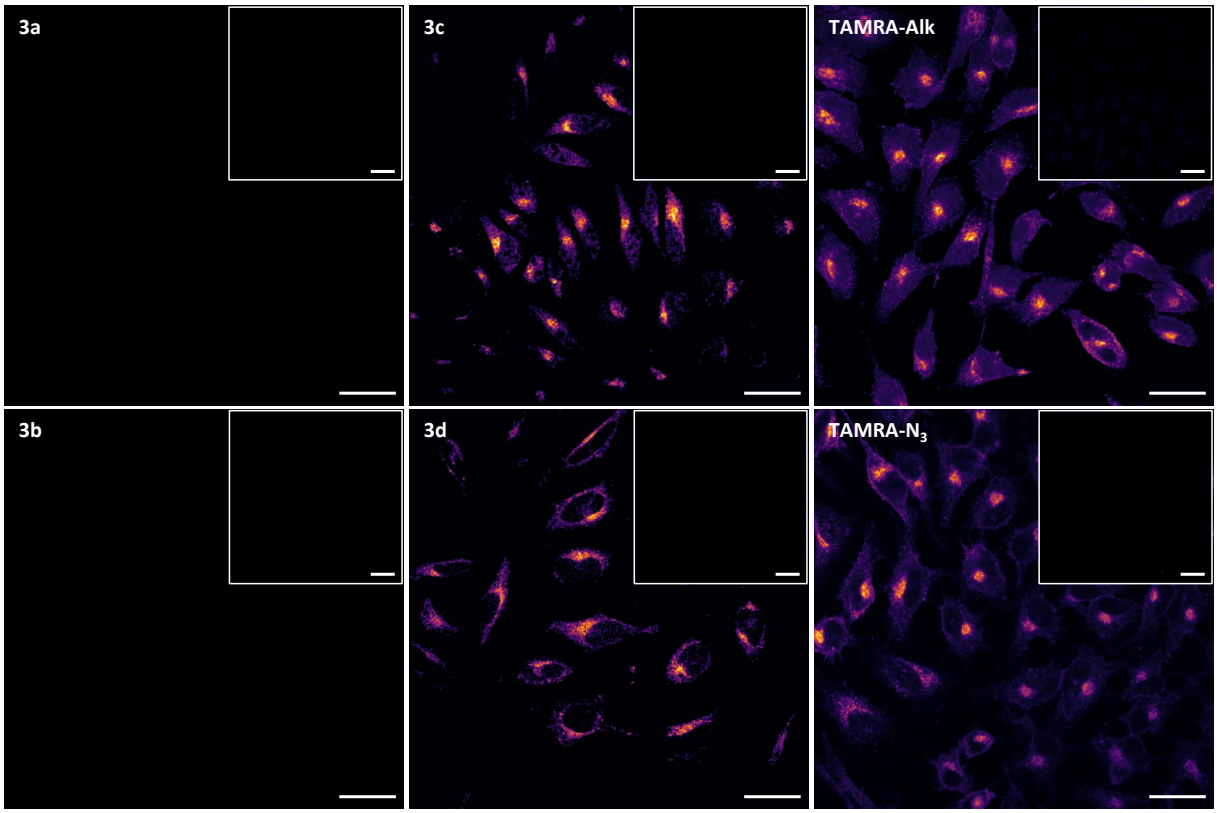


Figure 4. *Bioorthogonal labeling of sialoglycoconjugates with Ir(III)-ppy complexes observed by laser scanning confocal microscopy. Cells were incubated with ManNAz* ***4*** *(top) or ManNAl* ***5*** *(bottom) and reacted with complementary probes* ***3a-d****, or TAMRA-N_3 or TAMRA-Alk. A negative control showing cells reacted with each probe after incubation of untagged N-acetylmannosamine (no chemical reporter) is shown in the top-right corner of each image. Experiments were carried out as three biological replicates. Scale bar = 30µm.*

The photobleaching of **3c** was then assessed and compared to two dyes reputed for their photostability (Figure 5). In cells incubated with ManNAz **4** and labelled with either **3c** or Alexa Fluor 405 alkyne, 108 x 108 µm squares (512 pixels x 512 pixels) corresponding to a single cell were continuously scanned with laser irradiation at 405 nm with a dwell time of 50 µs/pixel in the presence and absence of antifade reagent. **3c** was revealed to be remarkably photostable and robust, with 85% of the mean photoluminescence intensity being retained after 80 seconds of exposure to the laser, time by which Alexa Fluor 405 was fully photobleached. Dye **3c** excited at 405 nm even proved to be more photostable than Alexa Fluor 488 excited at the lower energetic wavelength 488 nm (Figure S4). Susceptibility to hydrolysis was also tested in PBS buffer over several weeks and no degradation was observed. Altogether the designed probes displayed consistent photophysical parameters independent of the bioorthogonal handle used, with large Stoke's shift and long emission lifetime, and are chemically robust and highly photostable. They are thus ideal candidates for applications in multicolor time-gated bioimaging, whereby luminophores are distinguished using differing times of photon emission after an excitation pulse. Classical methods acquire two images with different sets of excitation light and of

emission filters to separate two probes which either emit in different regions or can be excited with two different colors. In contrast, in time-gated methods, imaging of the two labelling patterns is done in one experiment using a single excitation pulse, thus reducing photobleaching and any optical artefact that could occur changing the irradiation/detection conditions.

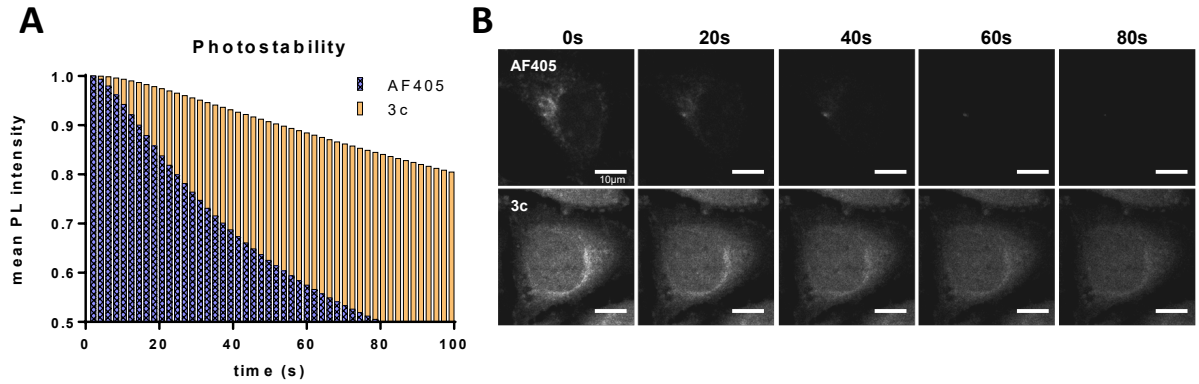


Figure 5. *Photostability of probe* ***3c*** *compared to AlexaFluor 405. The probes were linked to reporter units by CuAAC after metabolic incorporation of* ***4*** *in HeLa cells for 48h.* ***(A)*** *mean photoluminescence (PL) intensity of confocal images (108x108µm, 211nm px size, exc 405nm, em 600nm; 2s per image).* ***(B)*** *confocal images of a single cell followed upon continuous data acquisition (35x35µm, 211nm px size, exc 405nm, em 600nm).*

TIME-RESOLVED PHOTOLUMINESCENCE IMAGING MICROSCOPY

We then assessed our probe in lifetime imaging, using a single laser excitation pulse to distinguish two probes in dual labeling time-gated experiments with **4/3c** and DAPI. In this setting, the two probes are not localized in the same region of the cell: at the incubation time used for **4**, **3c** (long lifetime) is mainly localized in the golgi apparatus after bioorthogonal ligation, with a less intense secondary pool at the cell membrane. In contrast, DAPI (short lifetime, τ = 2 ns)[45] is well-known to localize in the nucleus upon binding with the minor groove of DNA, thus providing a straightforward platform to validate signal specificity in single pulse time-gated experiments. Images of HeLa cells were acquired using a 405 picosecond laser excitation and a long-pass filter, to collect all photons with a wavelength superior to 410 nm with a single avalanche photodiode detector (APD) equipped with a single photon counting module. Time-gating was then applied to unmix emitted photons corresponding to DAPI (blue channel, time-gate $0 < \tau < 20$ ns) and to sialoglycoconjugates ($40 < \tau < 500$ ns) (Figure 6).

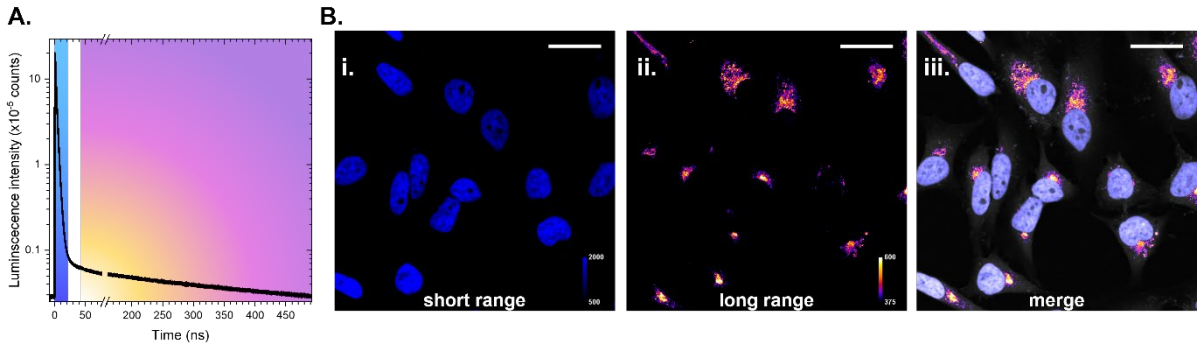


Figure 6. *HeLa cells labelled with DAPI and ManNAz **4** / **3c**. A) Emission time decay and raw image with all collected photons used to generate time-gated channels. B) Time-gated images for short lifetime range in blue (0 < τ < 20 ns, left), long lifetime range in fire scale (40 < τ < 500 ns, middle), and merge of the two channels. The marking patterns are typical of DAPI and of sialylated glycoconjugates, respectively. Experiments were carried out as three biological replicates. Scale bar = 30 µm.*

We then showed the full potential of probe **3c** in dual labeling experiments exploiting two monosaccharide reporters of the same metabolic pathway. For that purpose, we used ManNAz **4** and SiaNAl **6**, as they are indeed both incorporated as sialic acid units in glycoconjugates, but enter the metabolic network at two different points. SiaNAl **6** enters cells by pinocytosis and is directly activated in the nucleus as the sugar nucleotide donor cytidine 5'-monophospho *N*-pentynoylneuraminic (CMP-SiaNAl) acid prior to golgi export and transfer onto a nascent glycoconjugate by a sialyltransferase. ManNAz **4** must follow a longer path as it penetrates cells through a yet unknown transport mechanism and undergoes three enzymatic conversions in the cytosol (phosphorylation to ManNAz-6-P, conversion to SiaNAz-9-P by condensation with phosphoenolpyruvate, then dephosphorylation to SiaNAz) before it can be transported to the nucleus and activated as CMP-SiaNAz[43]. Therefore, the reporters are expected to be localized in the same regions of the cell, but with different kinetics of incorporation.

Hela cells incubated with both reporters were submitted to two consecutive CuAAC labeling steps. ManNAz **4** was reacted with long-lived alkyne dye **3c**, while SiaNAl **6** was ligated to short-lived Alexa Fluor 405 azide. Cells were then imaged as previously (Figure 7). Using similar time-gating analysis, we could get an image specific of Alexa 405 nm (green channel) and of **3c** (red channel), with full signal specificity of each channel. To the best of our knowledge, this is the first time that both a sialic acid reporter and a *N*-acetylmannosamine reporter are used simultaneously in the same cell to generate dual color maps of sialylation. Partial co-localization of time-resolved signals was observed, but as expected some signals are not co-localized, thus evidencing differences in metabolism kinetics between the two reporters. This dual-labelling method could provide a useful imaging platform in the future for studying the early steps of sialylation pathways, or for investigating trafficking and recycling of sialoconjugates in pathologies such as congenital disorders of glycosylation (CDGs).

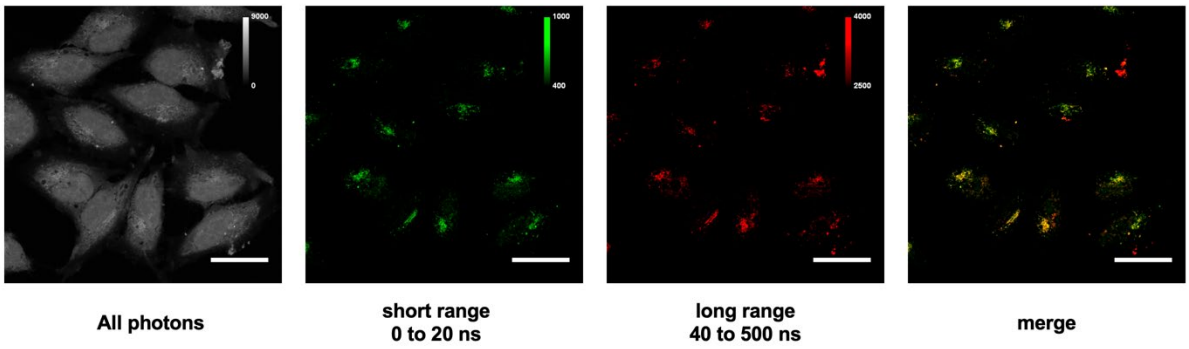


Figure 7. *Time-resolved imaging of HeLa cells dual labelled with SiaNAl* ***6*** */ Alexa Fluor 405 azide and with ManNAz* ***4*** */ Ir(III) dye* ***3c****. From left to right: raw image with full photon count, time-gated channel with photons emitted between 0 and 20 ns after the excitation pulse (specific of Alexa Fluor 405 nm), time-gated channel with photons emitted between 40 and 500 ns after the excitation pulse (specific of* ***3c****), merge image. Experiments were carried out as three biological replicates. Scale bar = 30 µm.*

Additional time-resolved experiments using a more conventional set-up using two excitation pulses at two different wavelengths to separately acquire the two channels were also successfully performed. In these experiments, ManNAz **4** and SiaNAl **6** were bioorthogonally labeled with **3c** and AlexaFluor 488, respectively (Figure 8). The advantages of clickable Ir(III) probes in multiplexed settings are showcased here: although the emission spectra of **3c** and of AlexaFluor 488 partially overlap, as well as their absorption spectra, time-gating allows full distinction of the photon's origin, ensuring the specificity of signal in each channel.

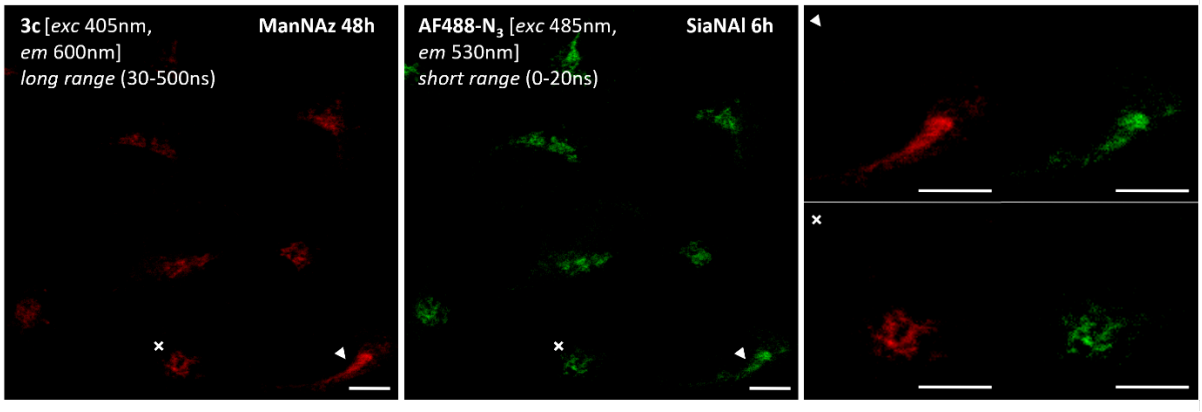


Figure 8. *Time-resolved imaging of HeLa cells dual labelled with SiaNAl* ***6*** */ Alexa Fluor 488 azide and with ManNAz* ***4*** */ Ir(III) dye* ***3c****. Experiments were carried out with three biological replicates. Scale bar = 30 µm.*

SYNCHROTRON-RADIATION BASED X-RAY FLUORESCENCE NANOIMAGING

While optical microscopy is the method of choice for imaging organelles or proteins within the context of a living cell, X-ray fluorescence nanoimaging (XRF-NI) techniques developed at some synchrotron facilities provide a highly complementary subset of information. This requires an adequately designed

XRF sensitive metabolic probe for quantitative co-localization purposes. Indeed, XRF-NI now allows high-spatial resolution (30-50 nm) of the elements within a sub-cellular structure with a high degree of sensitivity, as shown for Zn in neuronal structures reaching very low limit of detection of 130 atoms of zinc within a 40 nm pixel size.[46] This is encouraging in the quest for quantifying intracellular metabolic probes that are expected to report locally for a very low number of molecules at the nanoscale.

Iridium is an atom of choice for nanoscale XRF-NI, having no biological role and an extremely low abundance in living organisms (below 1 ng/L in serum for example).[47] Thus, endogenous Ir is not detected by XRF within cells and does not induce any X-ray autofluorescence background. Furthermore, main X-ray fluorescence emission rays can be observed for Ir at 9.175 keV (Lα) and 10.708 keV (Lβ), a region nearly devoid of emission lines from most biologically important elements observed in cells. This ensures a minimal background signal. In our samples, only the Zn-Kβ X-ray emission line at 9.57 keV might slightly overlap with the Ir-Lα emission line, while the Ir-Lβ remains in a spectral region free of other element contribution. Spectral fitting with a dedicated software allows the deconvolution of any such potential spectral overlap as well as the removal of any background contributions, to get the net peak area for all elements detected. From these, concentrations can be calculated using the fundamental parameter method as already reported in similar XRF-NI studies.[11,48,49]

HeLa cells were grown for 48h on silicon nitride membranes with either ManNAz **4**, Ac_4ManNAz or ManNAc as control and submitted to CuAAC with **3c** as described hereabove. XRF-NI is a low throughput technique and the limited access to these specific scarce nanoimaging instruments hosted by large scale facilities, *i.e.* synchrotrons, further reduces the availability of the method and limits the number of sample to be analyzed. Thus, we restricted our reporter panel to these two azide derivatives which are commonly used in MOE approaches. The benefit of a high resolution and high elemental sensitivity provided by XRF-NI should help to decipher differences in the signal obtained with ManNAz and Ac_4ManNAz, as the peracetylation of monosaccharide chemical reporters is a technical point of importance in MOE.[18] Silicon nitride membranes were directly observed by confocal fluorescence microscopy in PBS with an immersion objective to minimize any damages to the cells and/or the Si_3N_4 membranes. This allowed to take wide field images and to observe almost every cell on each membrane prior to XRF experiments at the synchrotron. Cells were kept under PBS at 4°C until they were quickly rinsed with an ammonium acetate buffer, vitrified by plunge-freezing in liquid ethane and kept under liquid nitrogen before cryo-transfer for cryo-XRF imaging.[50] In the present case, the sample analysis at cryogenic temperature limits the impact of the intense focused X-ray beam and radiation damages on cells during acquisition of XRF data.[51]

The online transmitted light microscope allowed us to precisely select the cells of interest and generate a direct comparison between confocal UV/vis phosphorescence and XRF images (Figure 9). We observed a strong correlation between both experiments, with clear co-localization of the probes by UV/vis phosphorescence and the iridium XRF signal. While the UV/vis signal is more faint or blurry, the XRF allows finer details to be observed, which is especially visible in the ManNAz condition where a punctiform pattern around the golgi apparatus is revealed by the cryo-XRF. A slightly different labeling profile is observed with the free-hydroxyles reporter ManNAz **4** compared to the peracetylated Ac_4ManNAz **6**, for which the pattern seems more diffuse and scattered, possibly due to non-specific incorporation of the reporter,[15,18] or to hydrophobic trapping. XRF-NI experiments give access to the Iridium surface mass expressed in ng/mm^2 which, combined with surface measurements, allows us to precisely determine the quantity of Iridium atoms and thus the number of probe molecules bound within the cells for each condition analyzed. In multimodal approaches, XRF-NI combined with UV-Vis luminescence microscopy may thus be used to link the arbitrary luminescence units to actual quantitative values in future studies. Incidentally, in the incubation conditions used here (48h for both sugars in HeLa cells), the surface density of Ir(III) probe was shown to be higher after ManNAz incubation than after Ac_4ManNAz incubation, an unexpected result confirming that unprotected ManNaz has a high potential as sialylation reporter, as it avoids most of the non-specific labeling caused by peracetylated reporters. Indeed, whereas incorporation of peracetylated reporters is usually thought to be more efficient than free-hydroxyl reporters because they enter cells by passive diffusion, in our case their actual metabolic incorporation in glycoconjugates was less efficient.

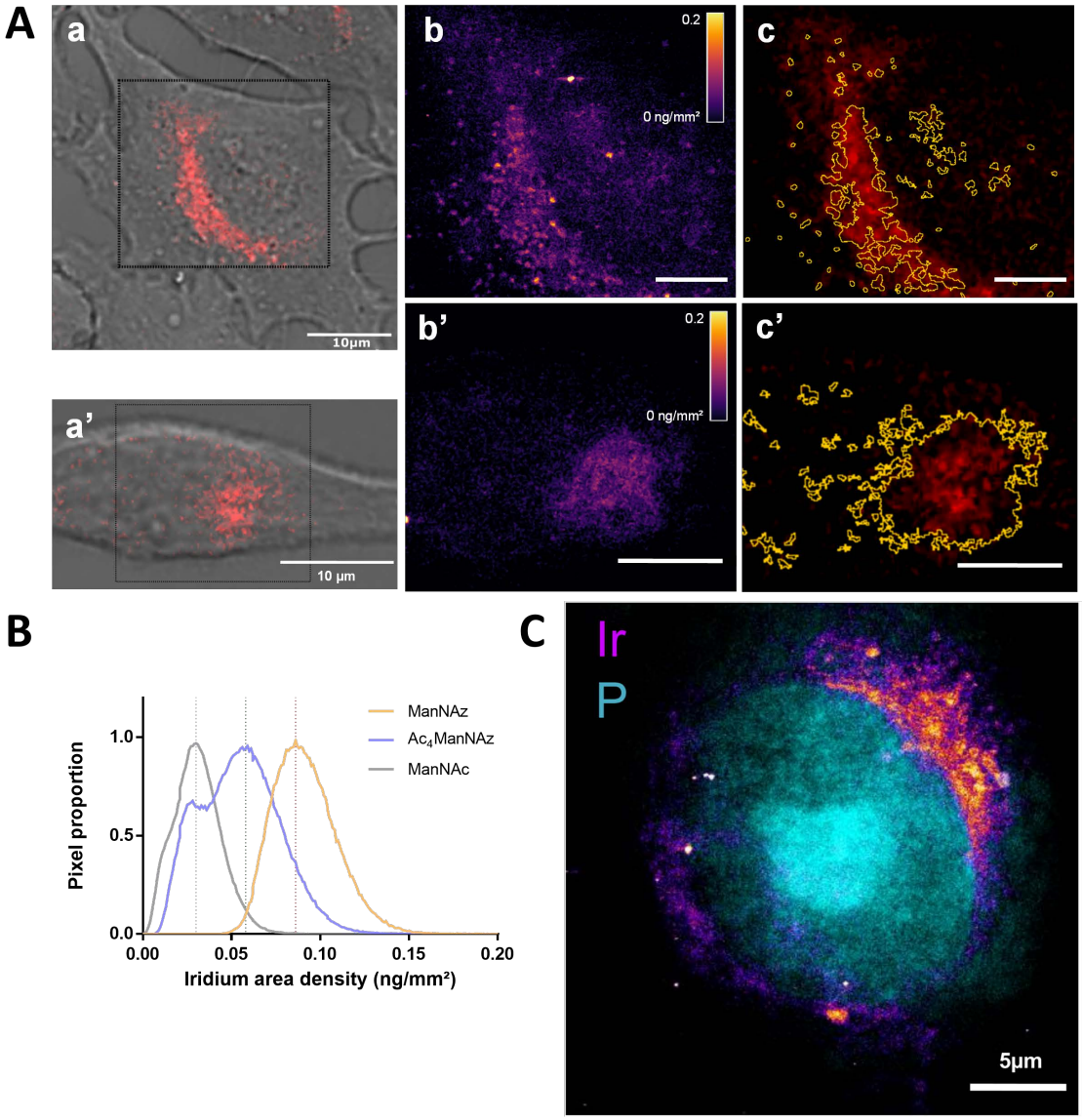


Figure 9. *X-ray fluorescence nanoimaging (XRF-NI).* ***(A)*** *Dual observation of HeLa cells after incorporation of ManNAz* ***4*** *or Ac_4ManNAz* ***7*** *and CuAAC with* ***3c****, by UV/vis confocal microscopy (a, a', respectively) and XRF-NI (b, b'). XRF-NI analysis allowed creation of masks showing great correlation between UV/vis and XRF signal (c, c'). Scale-bar = 5µm.* ***(B)*** *Comparison of Iridium area density (ng/mm²) in XRF analysed cell surfaces after 48h ManNAz, Ac_4ManNAz and ManNAc (control) incorporation.* ***(C)*** *XRF elemental imaging of Iridium (fire colorscale) and Phosphorus (cyan) in HeLa cell after 48h ManNAz incorporation shows 16-fold improved resolution compared to UV/vis confocal microscopy (50nm/px).*

INDUCTIVELY COUPLED PLASMA MASS SPECTROMETRY

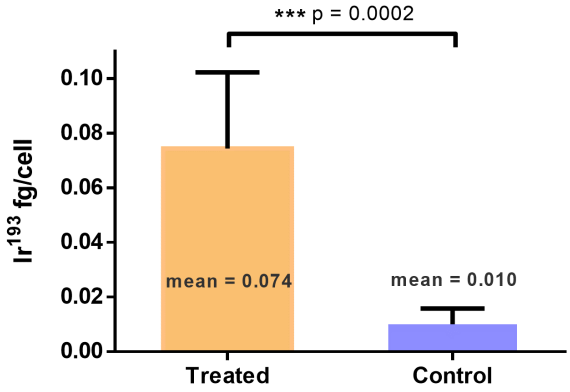


Figure 10 *– ICP-MS. HeLa cells grown with ManNAz 4 or ManNAc (control) for 48h and submitted to CuAAC with 3c and digested in nitric acid before concentration of iridium ions was determined by ICP-MS. A seven-fold increase in iridium content is observed for ManNAz condition (*** p < 0.001 T-test)*

Since elemental mapping by XRF-NI is not a routine technique, we then wanted to check whether ICP-MS could be used for iridium quantification as a routine technique exploitable in higher throughput, to allow bulk statiscal analysis on a larger pool of samples. HeLa cells after 48h incorporation of ManNAz **4** or ManNAc as control and submitted to CuAAC with **3c** as described above were digested in nitric acid. The concentration of iridium ion was then determined by ICP-MS. On average, a seven-fold increase in iridium concentration was found in ManNAz 4 treated cells (0.074 fg / cell, Figure 10) compared to control, with high-statistical confidence.

SUPPORTING INFORMATION

MICROSCOPY

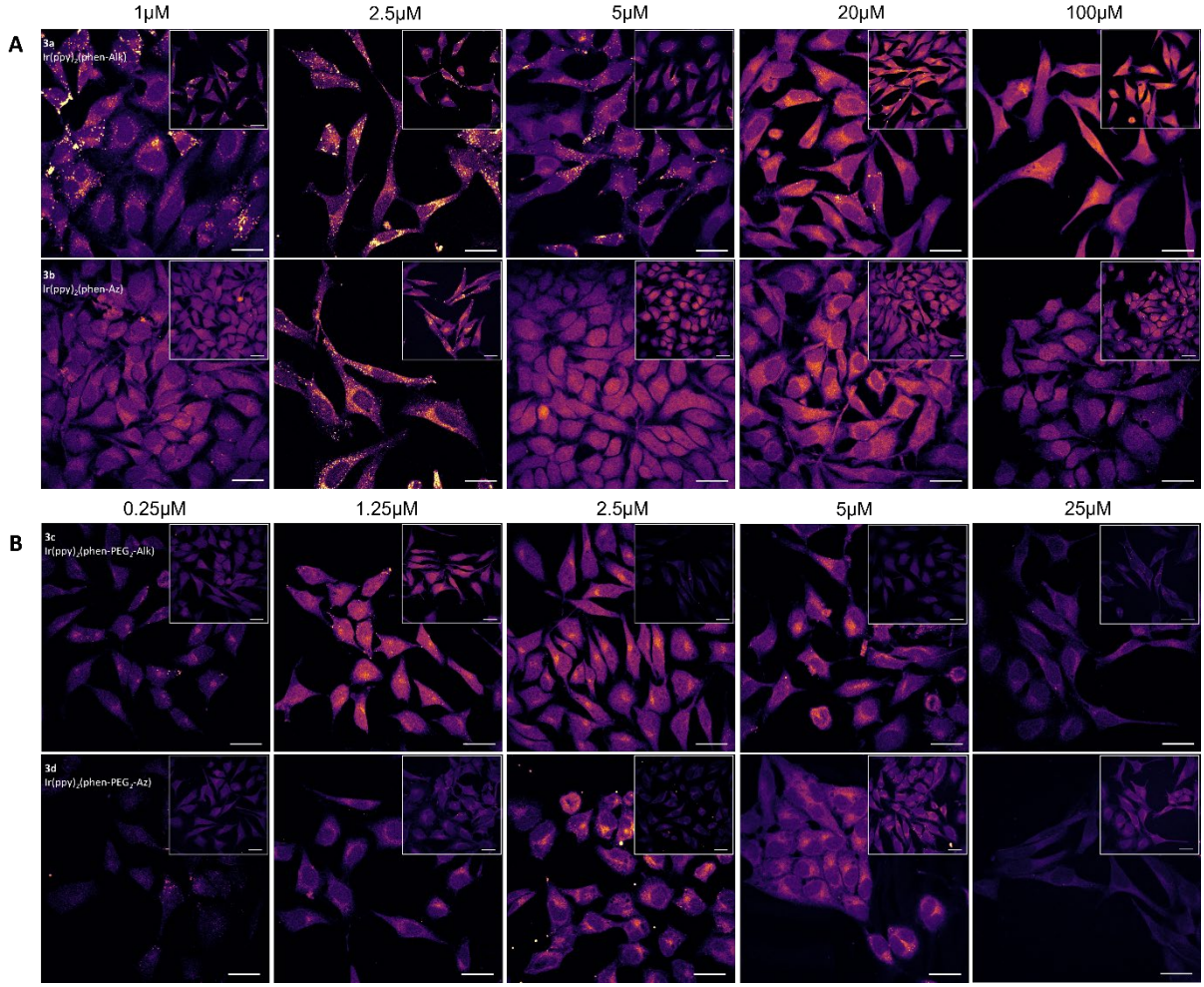


Figure S1 : *Probe concentration optimization for cell labeling via CuAAC.* ***(A)*** *Optimization of the non-PEG complexes 3a and 3b with concentrations ranging from 1 to 100µM after 48h incorporation of ManNAz or ManNAc (control, corner inserts).* ***(B)*** *Optimization of the PEG-spacer bearing complexes 3c and 3d with concentrations ranging from 0.25 to 25µM after 48h incorporation of ManNAz or ManNAc (control, corner inserts). (exc 405nm, em 570-620nm). Scale bar = 30 µm.*

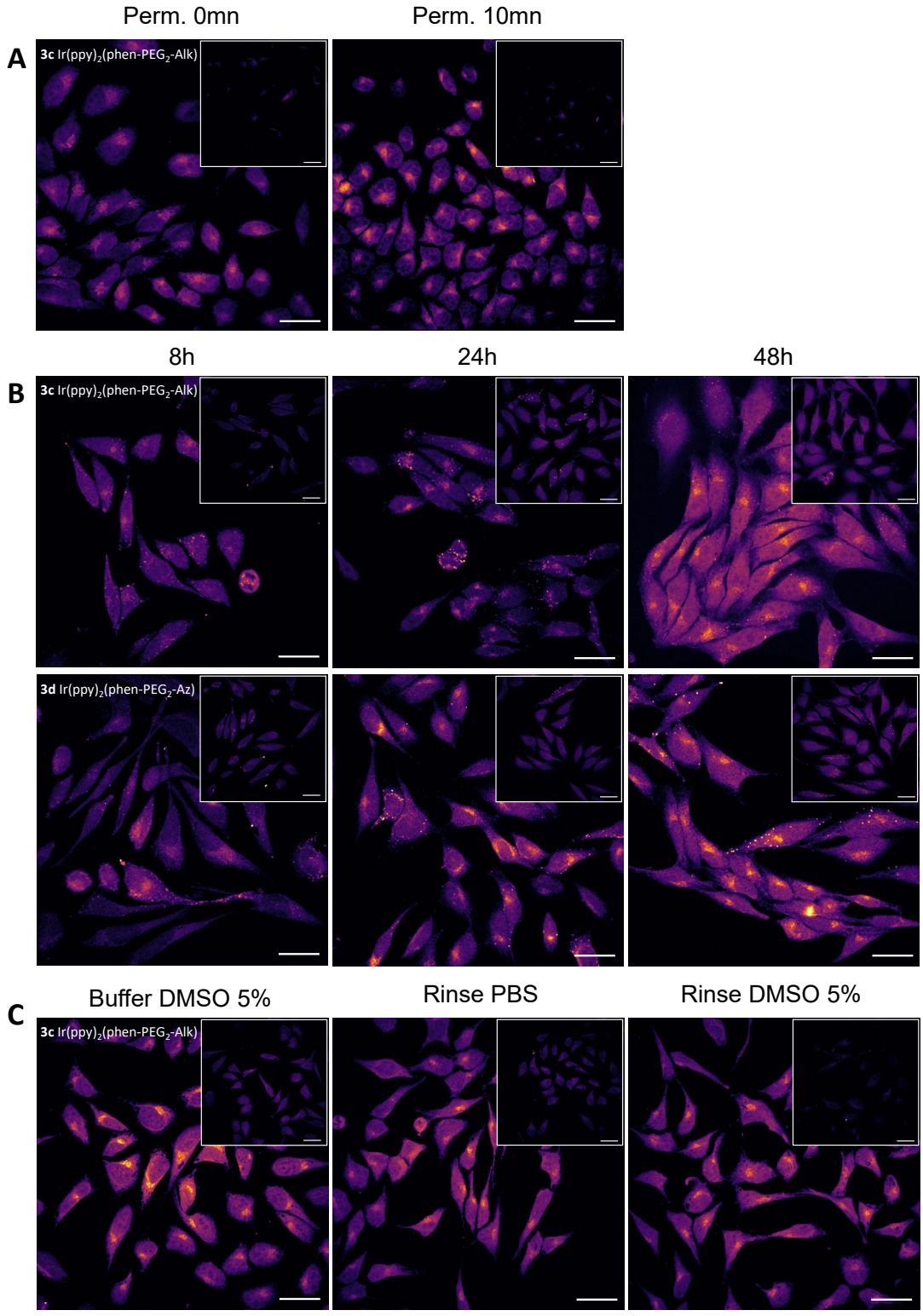
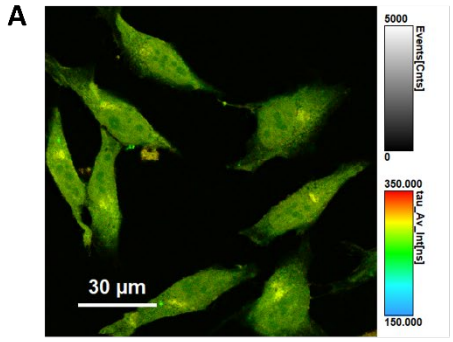
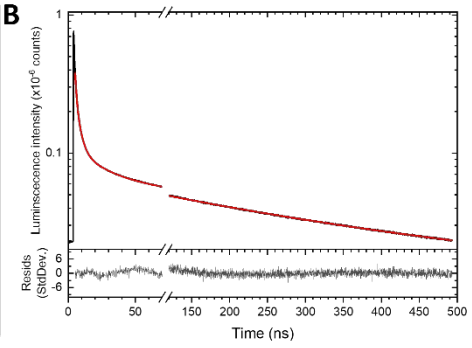


Figure S2 : *Optimization of other labelling parameters after incorporation of ManNAz or ManNAc (control, corner inserts).* ***(A)*** *Optimization of cell membranes permeabilization step with Triton-X100 0.5% for 0 or 10mn.* ***(B)*** *Optimization of ManNAz incorporation time (8, 24 or 48h).* ***(C)*** *Optimization of the CuAAC buffer composition (in PBS or PBS with 5% DMSO) and the rinsing step (pure PBS or PBS with 5% DMSO). 3*

PHOTOLUMINESCENCE LIFETIME IMAGING MICROSCOPY





C

FLIM parameters		Fractional intensities
τ_1	316 ± 0.99	0.914
τ_2	3.28 ± 0.033	0.029
τ_3	21 ± 0.15	0.050
τ_4	0.971 ± 0.016	0.007
A_1	52.9 ± 0.033	
A_2	164 ± 1.6	
A_3	44 ± 0.2	
A_4	124 ± 0.78	
$\tau_{averint}$	289.8 ± 0.93	
$\tau_{averamp}$	47.5 ± 0.18	
χ^2	2.0922	

Figure S3: (A) *PLIM image of cells labelled with 3c after fitting the signal by a sum of 4 exponential functions. The color code represents the intensity-averaged lifetime obtained from the 4 time constants arising from the fit.* ***(B)*** *Number of occurrences per photoluminescence lifetime of 3c labelling images.* ***(C)*** *FLIM parameters used for fitting 3c labelling images*

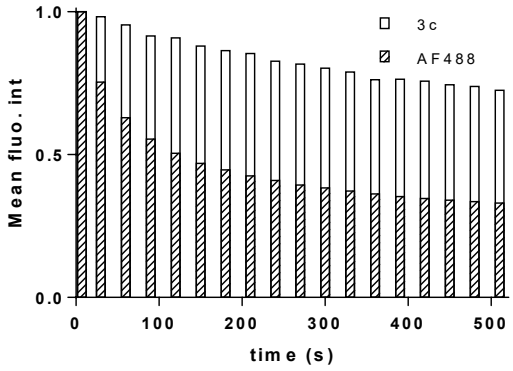


Figure S4 *– Photostability of* ***3c*** *excited at 405nm and AF488 excited at 488nm. The acquisition was made on a widefield fluorescence microscope with constant excitation. Mean fluorescence intensities of whole cells was measured and plotted over time for comparison.*

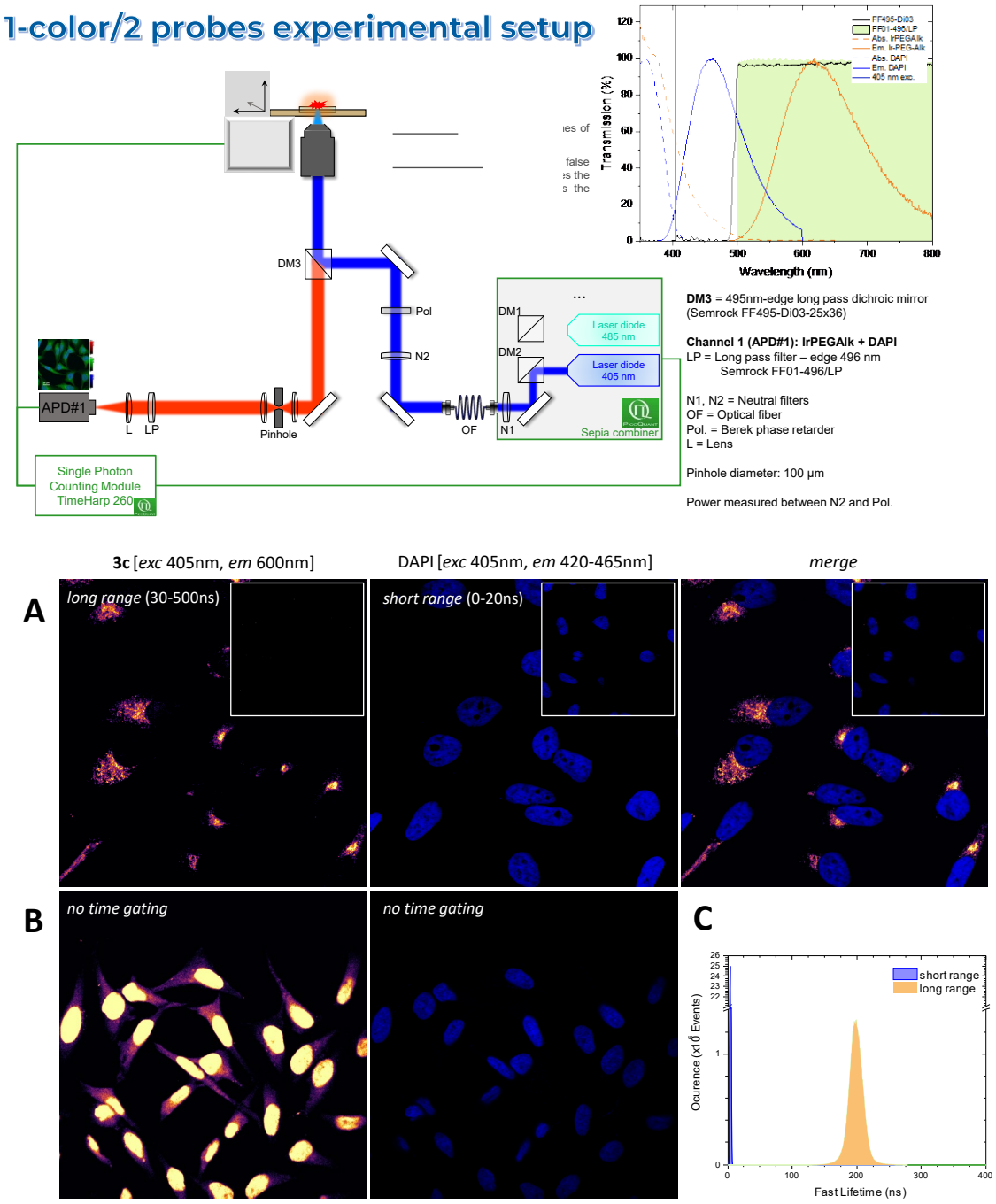


Figure S5 : *"1 color, 2 probes" experimental setup and photoluminescence lifetime imaging (PLIM). (Top) Schematic of the experimental setup. (Bottom) (A) Time-resolved, single excitation wavelength PLIM of HeLa cells after 48h ManNAz or ManNAc (control, corner inserts) incorporation, CuAAC with* ***3c*** *and nucleus staining with DAPI. (left) 3c detection channel (exc 405nm, em 600nm, τ 30-500ns), (middle) DAPI detection channel (exc 405nm, em 420-465nm, τ 0-20ns), (right) merge. Scale bar = 30µm. (B) Non-time resolved images of HeLa cells after 48h ManNAz incorporation, CuAAC with* ***3c*** *and nucleus staining with DAPI. (C) Number of occurences and photoluminescence lifetime histogram of time-resolved "1 color / 2 probes" PLIM images.*

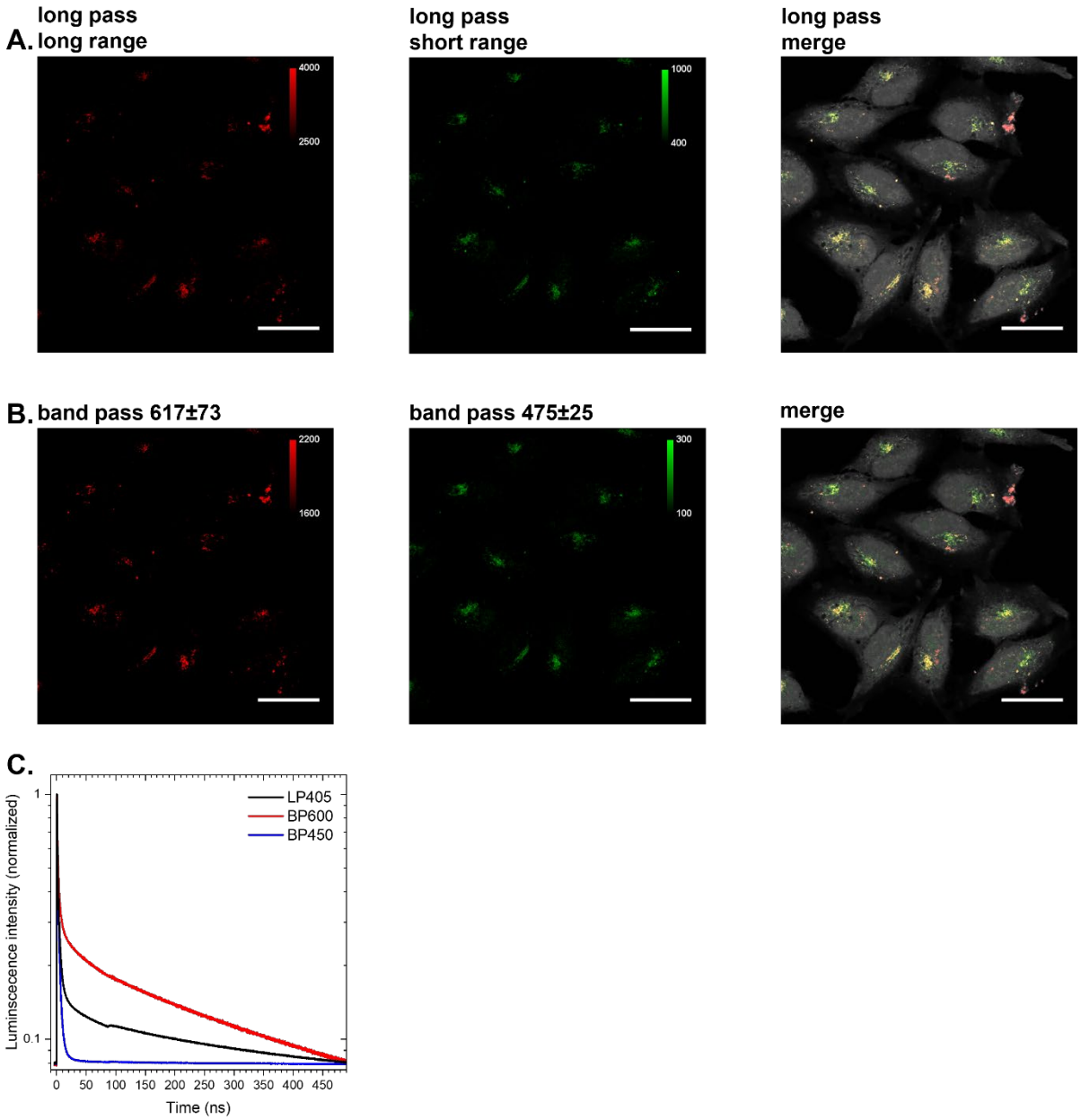


Figure S6: *Cells labelled with SiaNAz / Alexa Fluor 405 and ManNAz* ***4*** */ 3c and images with 405 nm picosecond laser excitation and* ***(A)*** *long pass emission filter > 410 nm or* ***(B)*** *band pass emission filters 617±73 nm and 475 ±25nm. Images in counts (scale bar: 30 um) considering photons emitted between 0 and 20 ns (short range) and 40 and 500 ns (long range) after the excitation pulse.* ***(C)*** *Normalized intensity curve of photoluminescence lifetime images with either a long-pass emission filter or band-pass emission filters.*

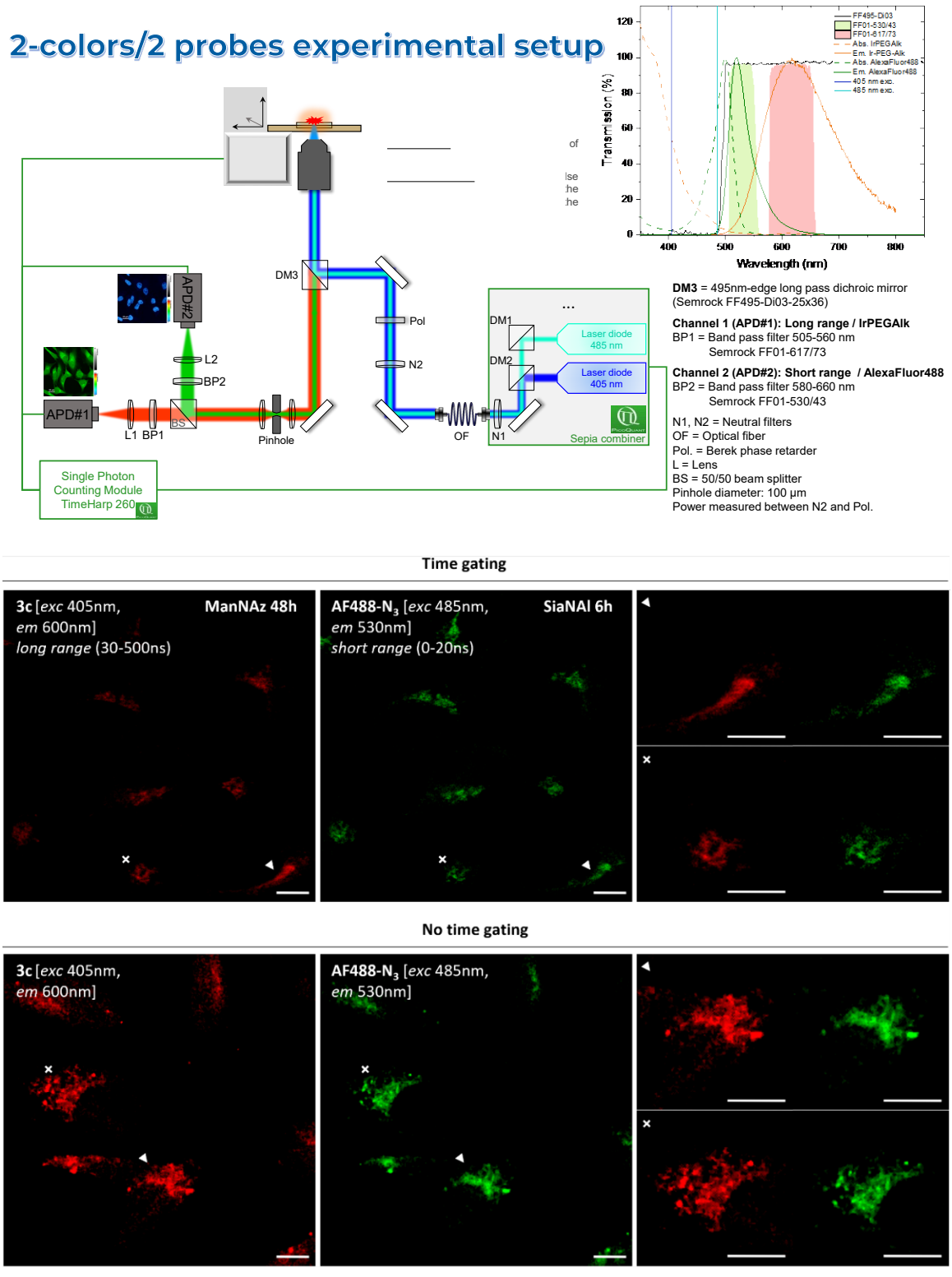


Figure S7 : *"2 color, 2 probes" experimental setup and photoluminescence lifetime imaging (PLIM) of cells labeled with SiaNAl / AF488-N_3 and ManNAz / 3c. (Top) Schematic of the experimental setup. (Bottom) (top) Time-resolved, dual excitation wavelength PLIM of HeLa cells after 48h incorporation ManNAz and 6h SiaNAl incorporation and CuAAC with 3c and AF488-N_3. (bottom) Same conditions without time resolution. Scale bar = 10 µm.*

FLUORESCENCE HPLC-MS ANALYSIS

Monitoring of the CuAAC reaction between 3a-d and ManNAz/ManNAl in solution.

The disappearance of probes **3a-d** (500 µM) was monitored by reverse phase HPLC-Fluo (Fig S5) and HPLC UV (Fig S6) when in presence of a 10-fold excess of reporters ManNAl or ManNAz (5mM). The reaction was carried out at room temperature in 200 µl CuAAC buffer ($CuSO_4$ 150 µM, BTTAA 300 µM, Na ascorbate 2.5 mM). Probes **3a-d** were prepared from a 10% DMSO stock solution. 20 µl injections were performed after 2 h of reaction (t_f). To get the initial state of the reaction (t_0) the Na ascorbate was omitted in the CuAAC buffer. Experiments were carried out with a C18 Uptisphere 4.6 x 250 mm / 5 µm column on a Shimadzu LC-20 system equipped with a RF20A fluorescence detector (λ_{exc} = 375 nm; λ_{em} = 590 nm) and on an Ultimate 3000 system equipped with a DAD-3000RS UV detector (λ = 254 nm). Data was processed in LabSolutions LITE v.5.93 and Chromeleon v.6.80.

Eluant A: H_2O/MeCN 95:5 + 0.1% TFA Eluant B: H_2O/MeCN 5 :95 + 0.1% TFA

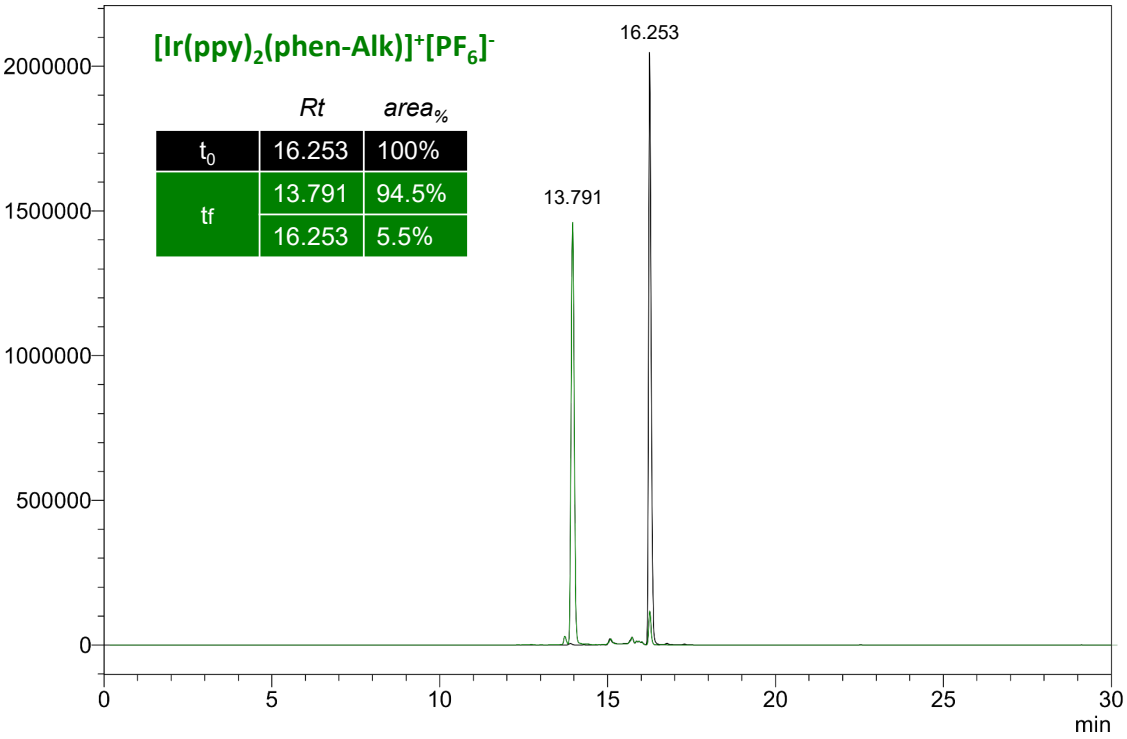
Elution method: *t(0 – 2 min): isocratic A/B 95:5 ; t(2 min – 21 min): gradient A/B 95:5 to 5:95 ; t(21 min – 26 min) isocratic A/B 5:95 ; t(26 min – 27 min): gradient A/B 5:95 to 95:5 ; t(27 min – 37 min): isocratic A/B 95:5.*

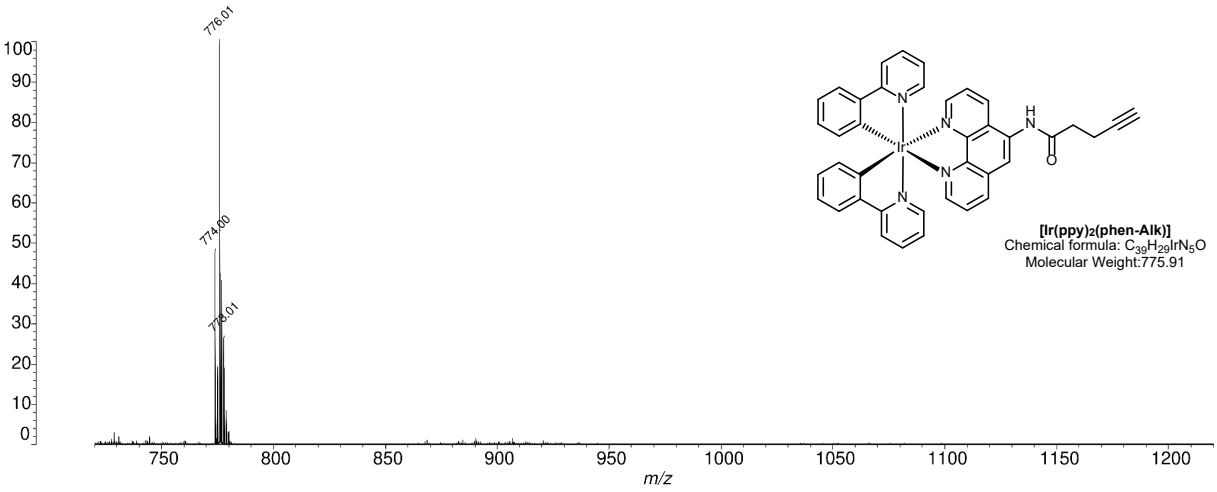
In the presence of an excess of ManNAl/ManNAz, probes **3a-d** are consumed and the formation of a peak of cycloadduct is observed (Fig. S6). Probes **3a** and **3b** show a partial reactivity as the probe's peak is still observed after 2 hours of reaction, while probes **3c** and **3d** react completely. The kinetic parameters in solution and within the biological environment are not transposable. For metabolic labeling experiments, the reaction time was therefore optimized directly in cell samples. A previous report presents a kinetic study of the CuAAC reaction between a similar Iridium complex and benzyl-azide carried out by NMR.[30]

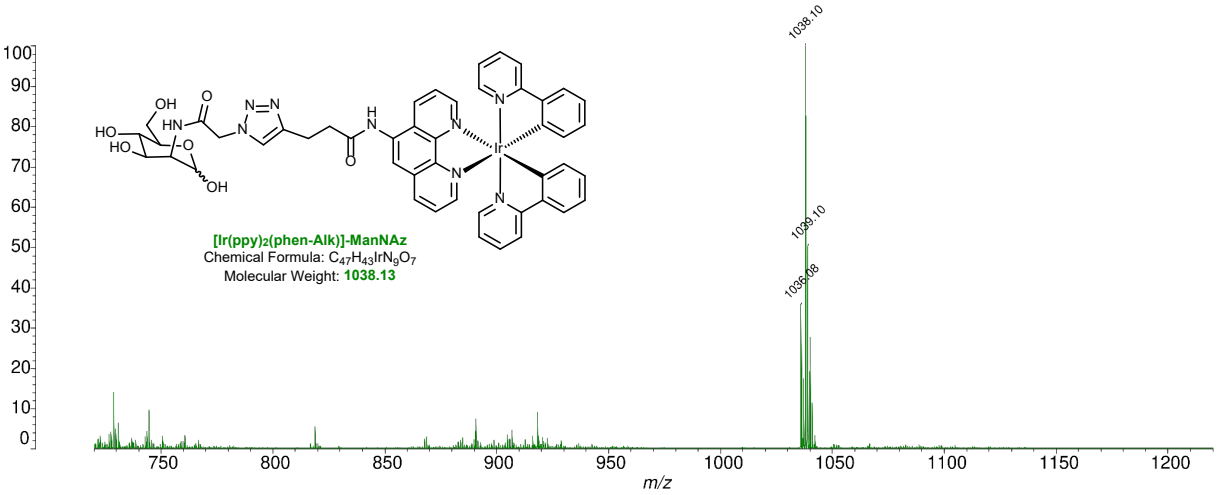
Identification of cycloadducts by mass spectrometry

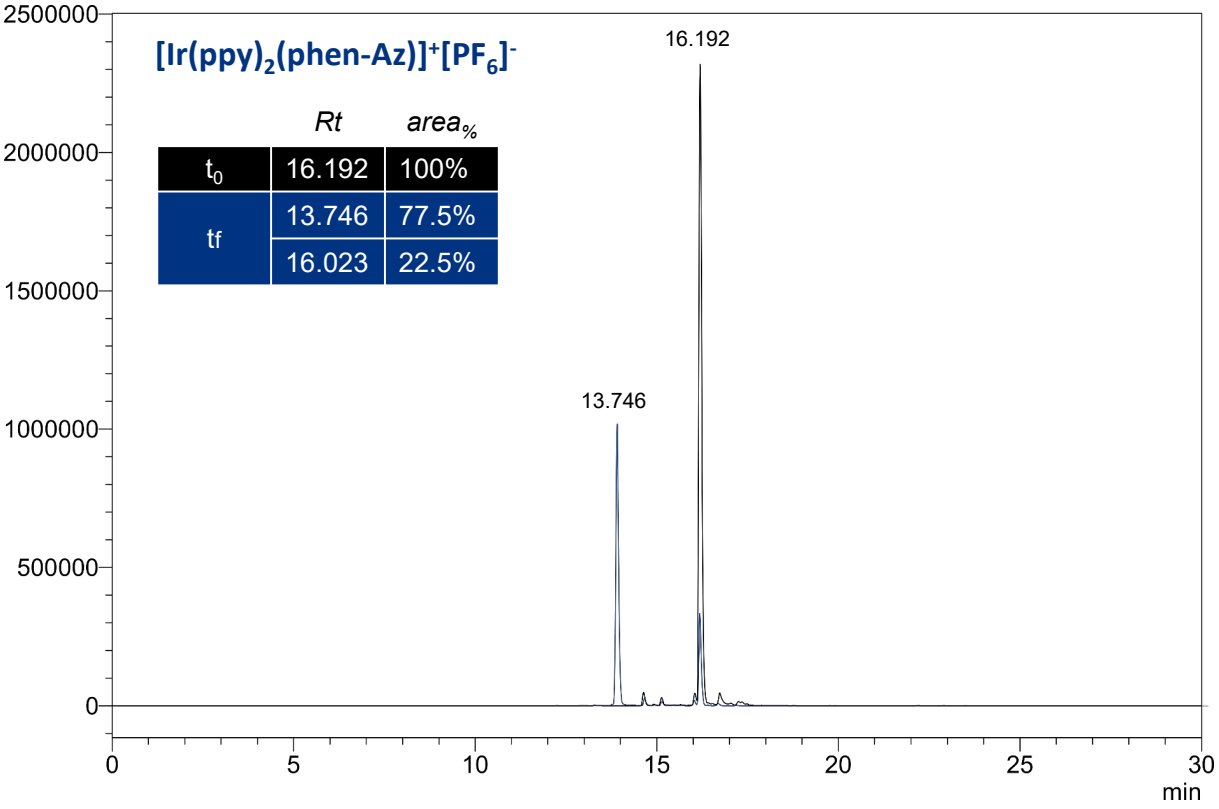
The probes and the cycloadducts fractions from the HPLC experiments were characterized by MALDI-ToF mass spectrometry using dihydroxybenzoic acid as matrix (Fig. S7):

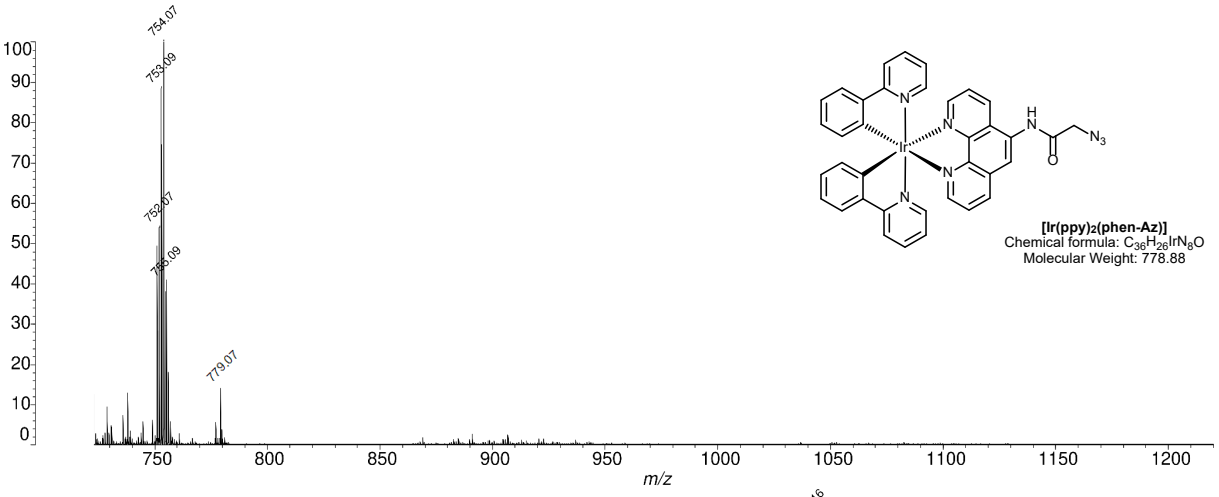
	probe	cycloadduct
3a	calcd: **775.91**; found: **776.01**	calcd: **1038.13**; found: **1038.10**
3b	calcd: **778.88**; found: **779.07**	calcd: **1038.13**; found: **1038.16**
3c	calcd: **894.04**; found: **894.18**	calcd: **1156.27**; found: **1156.27**
3d	calcd: **911.04**; found: **911.21**	calcd: **1170.29**; found: **1170.30**

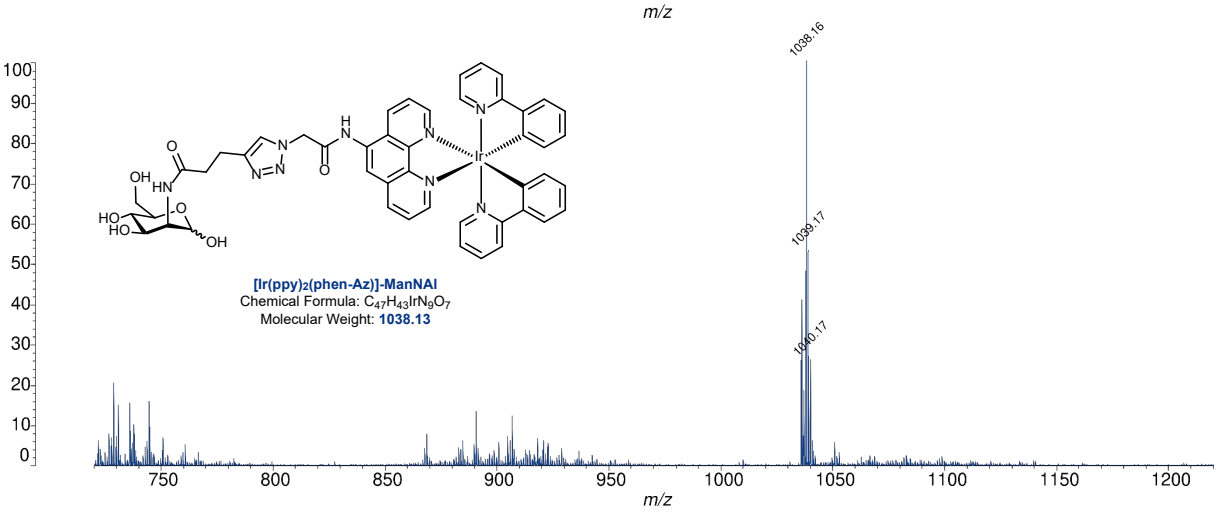

[Ir(ppy)2(phen-Alk)]+[PF6]-
Rt
area%
t0
16.253
100%
tf
13.791
94.5%
16.253
5.5%
16.253
13.791
2000000
1500000
1000000
500000
0
0
5
10
15
20
25
30
min

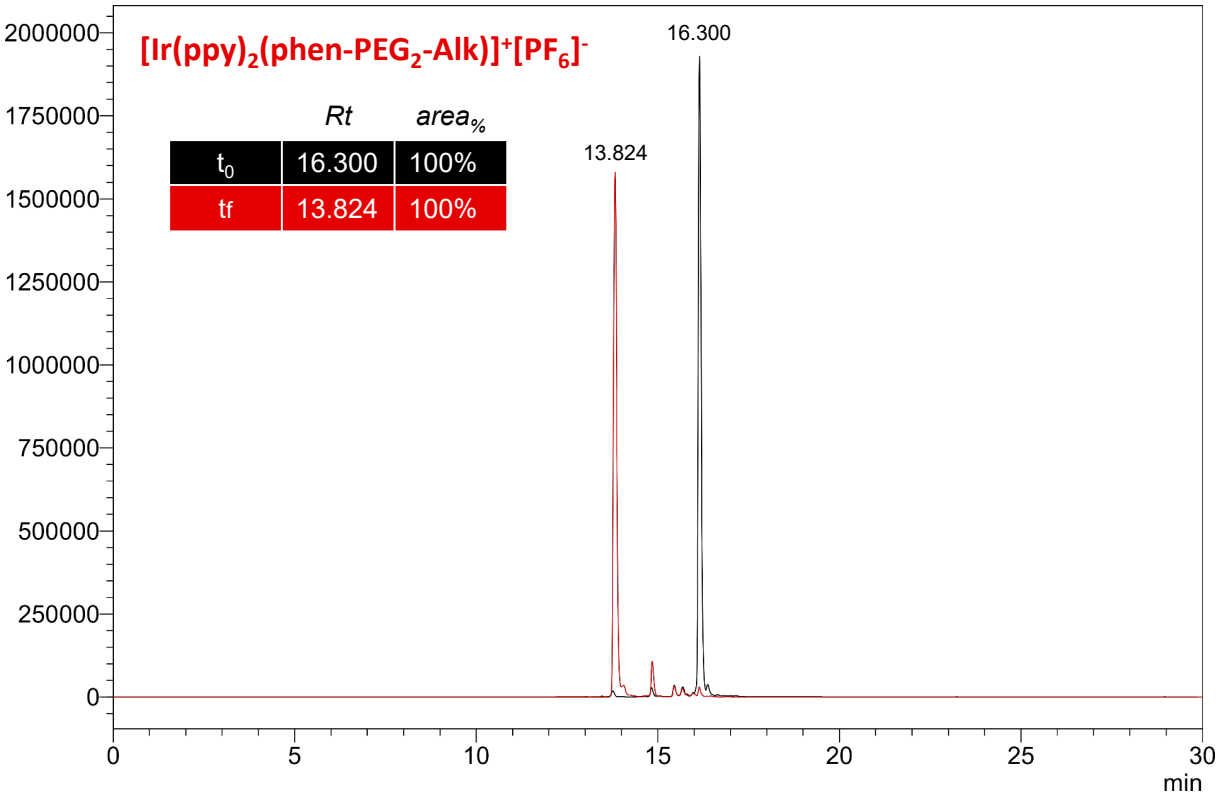

776.01
774.00
778.01
[Ir(ppy)2(phen-Alk)]
Chemical formula: C39H29IrN5O
Molecular Weight:775.91
m/z

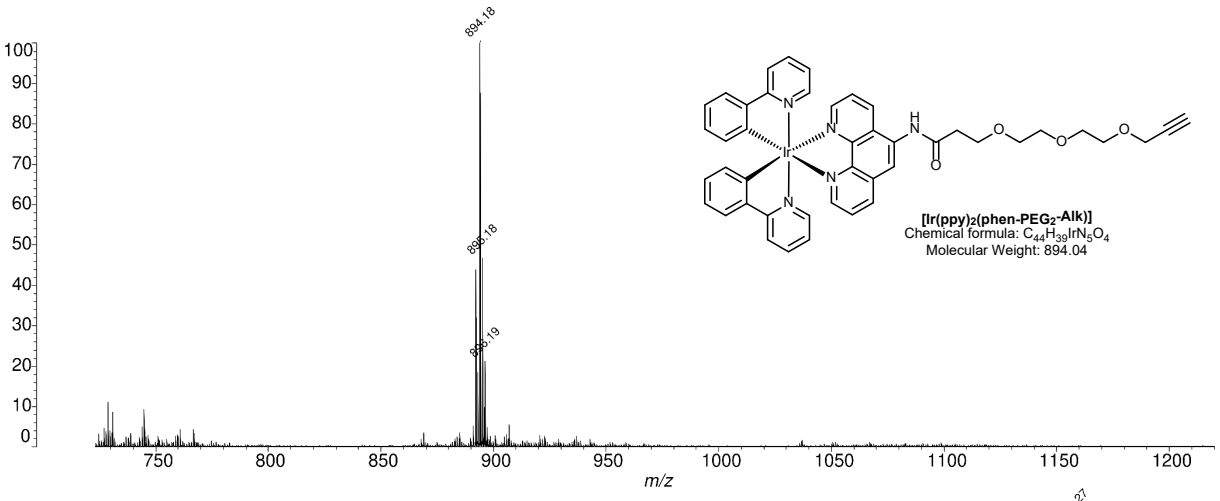

1038.10
1039.10
1036.08
[Ir(ppy)2(phen-Alk)]-ManNAz
Chemical Formula: C47H43IrN9O7
Molecular Weight: 1038.13
m/z

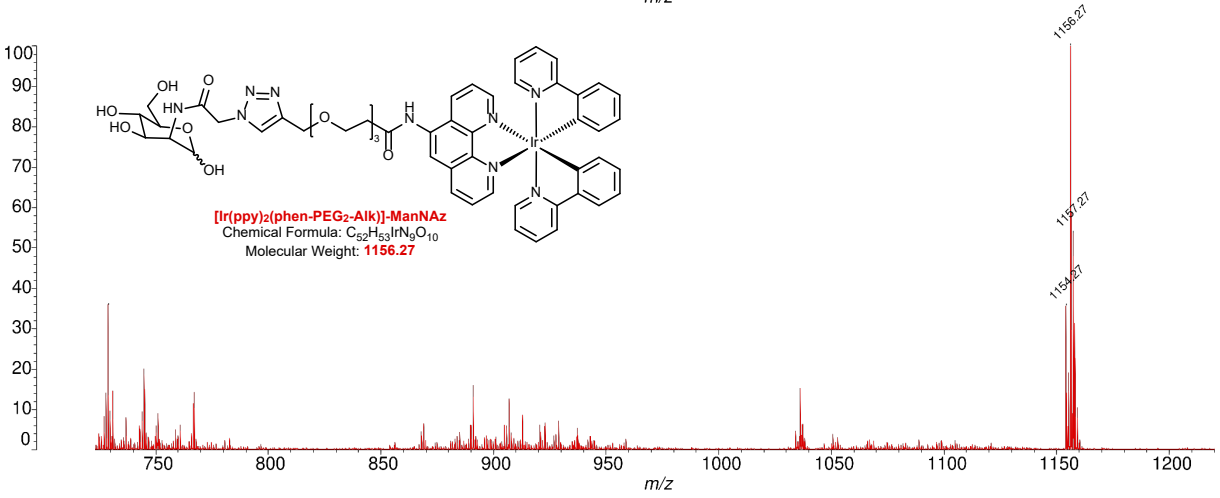

[Ir(ppy)2(phen-Az)]+[PF6]-
Rt
area%
t0
16.192
100%
tf
13.746
77.5%
16.023
22.5%
16.192
13.746
2500000
2000000
1500000
1000000
500000
0
0
5
10
15
20
25
30
min


754.07
753.09
752.07
755.09
779.07
100
90
80
70
60
50
40
30
20
10
0
750
800
850
900
950
1000
1050
1100
1150
1200
m/z
[Ir(ppy)2(phen-Az)]
Chemical formula: C36H26IrN8O
Molecular Weight: 778.88


1038.16
1039.17
1040.17
100
90
80
70
60
50
40
30
20
10
0
750
800
850
900
950
1000
1050
1100
1150
1200
m/z
[Ir(ppy)2(phen-Az)]-ManNAl
Chemical Formula: C47H43IrN9O7
Molecular Weight: 1038.13


[Ir(ppy)2(phen-PEG2-Alk)]+[PF6]-
Rt
area%
t0
16.300
100%
tf
13.824
100%
16.300
13.824
2000000
1750000
1500000
1250000
1000000
750000
500000
250000
0
0
5
10
15
20
25
30
min


894.18
896.18
896.19
100
90
80
70
60
50
40
30
20
10
0
750
800
850
900
950
1000
1050
1100
1150
1200
m/z
[Ir(ppy)2(phen-PEG2-Alk)]
Chemical formula: C44H39IrN5O4
Molecular Weight: 894.04


1156.27
1157.27
1154.27
100
90
80
70
60
50
40
30
20
10
0
750
800
850
900
950
1000
1050
1100
1150
1200
m/z
[Ir(ppy)2(phen-PEG2-Alk)]-ManNAz
Chemical Formula: C52H53IrN9O10
Molecular Weight: 1156.27

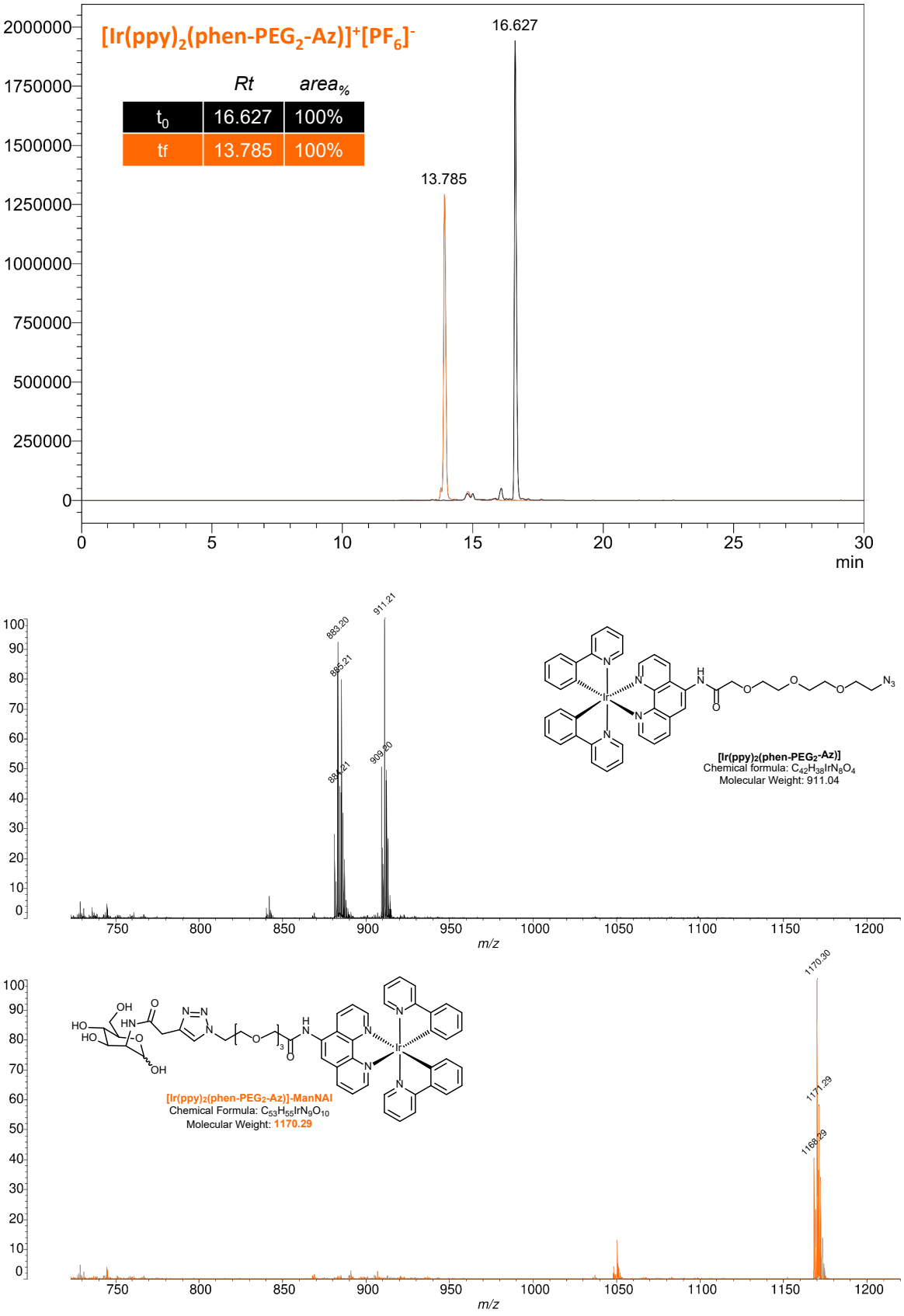
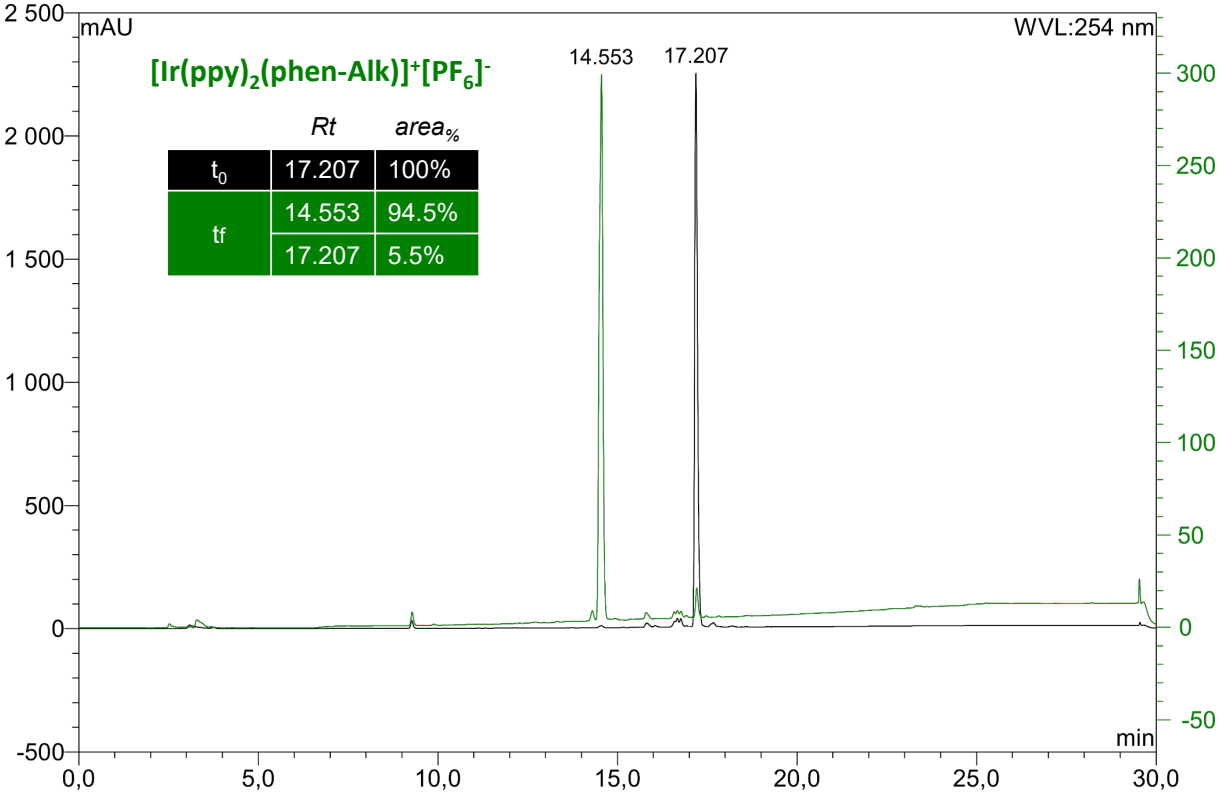
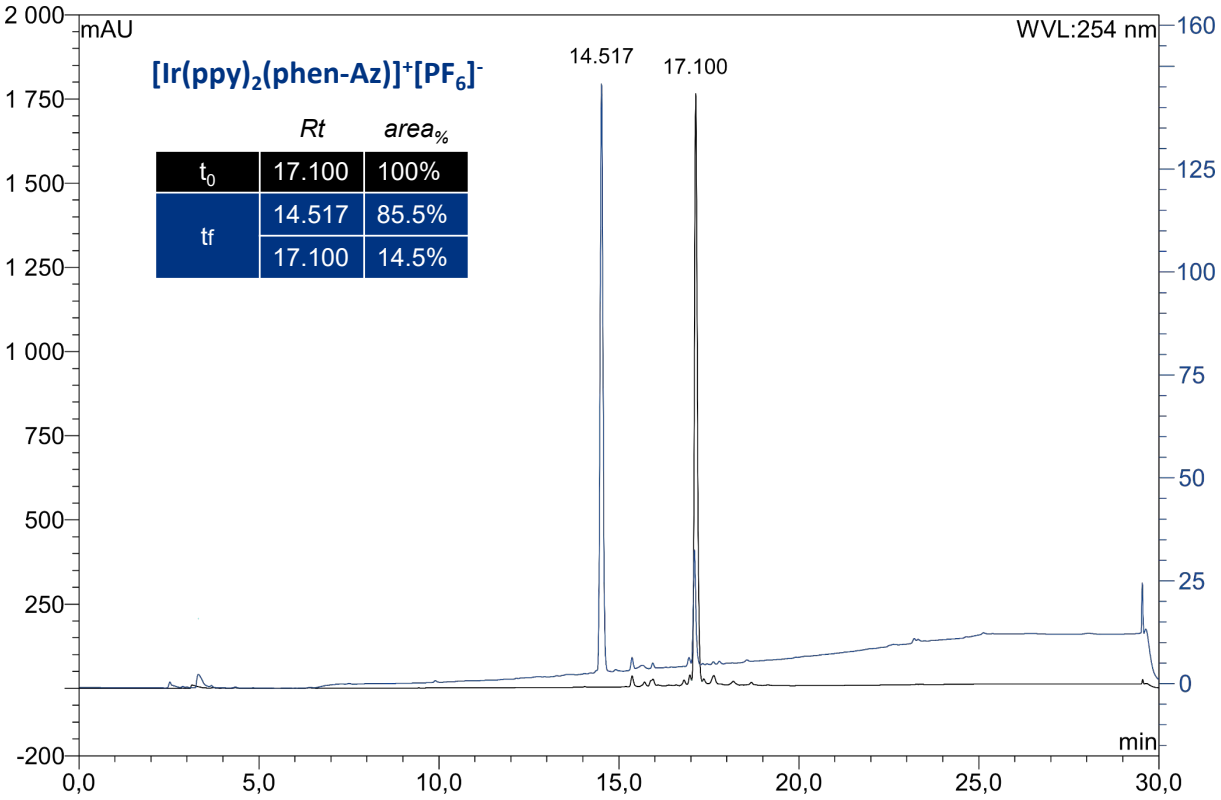


Figure S7 *– Fluorescence HPLC chromatograms and MALDI mass spectrometry spectra of eluted fractions of complexes 3a-d before and after CuAAC ligation with ManNAz.*

UV HPLC ANALYSIS



	Rt	area%
t_0	17.207	100%
tf	14.553	94.5%
	17.207	5.5%



	Rt	area%
t_0	17.100	100%
tf	14.517	85.5%
	17.100	14.5%

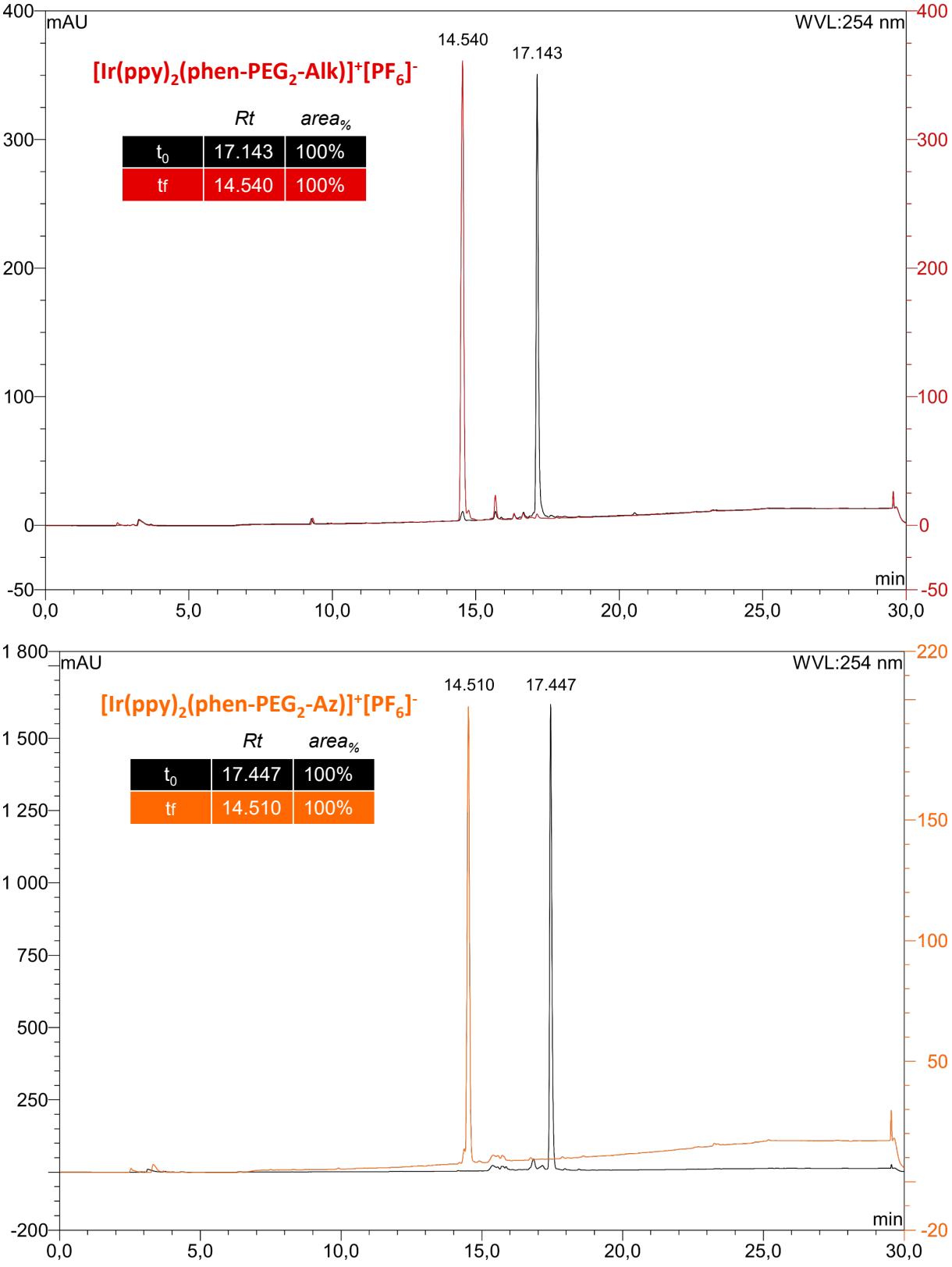


Figure S8 *– UV HPLC chromatograms of complexes 3a-d before and after CuAAC ligation with ManNAz.*

ICP MS ANALYSIS:

Hela cells were digested in HNO3 69%v/v for 40 min at 120 degrees in an autoclave. (Elara 11D class B, tabletop stainless steel from Tuttnauer) The concentration of iridium ion was determined by inductively coupled plasma-mass spectrometry (ICP-MS Perkin Elmer Nexion 2000b) with a limit of detection of 0.1 ppb. The analysis was carried out in standard mode. The results are an average of two technical replicates for each biological replicate (2 biological replicates in total).

BIBLIOGRAPHY :

[1] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *14*, 2596–2599.

[2] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.

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

[4] N. K. Devaraj, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2008**, *19*, 2297–2299.

[5] R. E. Bird, S. A. Lemmel, X. Yu, Q. A. Zhou, *Bioconjugate Chem.* **2021**, *32*, 2457–2479.

[6] V. Rigolot, C. Biot, C. Lion, *Angew. Chem. Int. Ed.* **2021**, *60*, 23084–23105.

[7] S. T. Laughlin, J. M. Baskin, S. L. Amacher, C. R. Bertozzi, *Science* **2008**, *320*, 664–667.

[8] S. Hell, *Science* **2007**, *316*, 1153–1158.

[9] F. Balzarotti, Y. Eilers, K. C. Gwosch, A. H. Gynnå, V. Westphal, F. D. Stefani, J. Elf, S. W. Hell, *Science* **2017**, *355*, 606–612.

[10] M. Weber, H. von der Emde, M. Leutenegger, P. Gunkel, S. Sambandan, T. A. Khan, J. Keller-Findeisen, V. C. Cordes, S. W. Hell, *Nat Biotechnol* **2023**, *41*, 569–576.

[11] J. J. Conesa, A. C. Carrasco, V. Rodríguez-Fanjul, Y. Yang, J. L. Carrascosa, P. Cloetens, E. Pereiro, A. M. Pizarro, *Angew. Chem. Int. Ed.* **2020**, *59*, 1270–1278.

[12] J. Tan, X. Zhang, X. Wang, C. Xu, S. Chang, H. Wu, T. Wang, H. Liang, H. Gao, Y. Zhou, Y. Zhu, *Cell* **2021**, *184*, 2665–2679.

[13] A. Walter, P. Paul-Gilloteaux, B. Plochberger, L. Sefc, P. Verkade, J. G. Mannheim, P. Slezak, A. Unterhuber, M. Marchetti-Deschmann, M. Ogris, K. Bühler, D. Fixler, S. H. Geyer, W. J. Weninger, M. Glösmann, S. Handschuh, T. Wanek, *Front. Phys.* **2020**, *8*, 47.

[14] 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.

[15] K. N. Chuh, B. W. Zaro, F. Piller, V. Piller, M. R. Pratt, *J. Am. Chem. Soc.* **2014**, *136*, 12283–12295.

[16] 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.

[17] 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.

[18] K. Qin, H. Zhang, Z. Zhao, X. Chen, *J. Am. Chem. Soc.* **2020**, *142*, 9382–9388.

[19] F. Wang, Y. Zhang, Z. Liu, Z. Du, L. Zhang, J. Ren, X. Qu, *Angew. Chem. Int. Ed.* **2019**, *58*, 6987–6992.

[20] Y. You, F. Cao, Y. Zhao, Q. Deng, Y. Sang, Y. Li, K. Dong, J. Ren, X. Qu, *ACS Nano* **2020**, *14*, 4178–4187.

[21] T. Deb, J. Tu, R. M. Franzini, *Chem. Rev.* **2021**, *121*, 6850–6914.

[22] S. M. Usama, F. Inagaki, H. Kobayashi, M. J. Schnermann, *J. Am. Chem. Soc.* **2021**, *143*, 5674–5679.

[23] L. Wu, C. Huang, B. P. Emery, A. C. Sedgwick, S. D. Bull, X.-P. He, H. Tian, J. Yoon, J. L. Sessler, T. D. James, *Chem. Soc. Rev.* **2020**, *49*, 5110–5139.

[24] K. K.-W. Lo, S. P.-Y. Li, K. Y. Zhang, *New J. Chem.* **2011**, *35*, 265–287.

[25] E. Baggaley, J. A. Weinstein, J. A. G. Williams, *Coordination Chemistry Reviews* **2012**, *256*, 1762–1785.

[26] Q. Zhao, C. Huang, F. Li, *Chem. Soc. Rev.* **2011**, *40*, 2508–2524.

[27] V. Fernández-Moreira, F. L. Thorp-Greenwood, M. P. Coogan, *Chem. Commun.* **2010**, *46*, 186–202.

[28] L. Murphy, A. Congreve, L.-O. Pålsson, J. A. G. Williams, *Chem. Commun.* **2010**, *46*, 8743–8745.

[29] L. C.-C. Lee, K. K.-W. Lo, *J. Am. Chem. Soc.* **2022**, *144*, 14420–14440.

[30] K. K.-W. Lo, B. T.-N. Chan, H.-W. Liu, K. Y. Zhang, S. P.-Y. Li, T. S.-M. Tang, *Chem. Commun.* **2013**, *49*, 4271–4273.

[31] S. P.-Y. Li, A. M.-H. Yip, H.-W. Liu, K. K.-W. Lo, *Biomaterials* **2016**, *103*, 305–313.

[32] T. S.-M. Tang, H.-W. Liu, K. K.-W. Lo, *Chem. Commun.* **2017**, *53*, 3299–3302.

[33] N. Hidalgo, J. J. Moreno, I. García-Rubio, J. Campos, *Angew Chem Int Ed* **2022**, *61*, e202206831.

[34] J. Wang, J. Xue, Z. Yan, S. Zhang, J. Qiao, X. Zhang, *Angew Chem Int Ed* **2017**, *56*, 14928–14932.

[35] Q. Wu, K. Y. Zhang, P. Dai, H. Zhu, Y. Wang, L. Song, L. Wang, S. Liu, Q. Zhao, W. Huang, *J. Am. Chem. Soc.* **2020**, *142*, 1057–1064.

[36] D.-L. Ma, H.-Z. He, K.-H. Leung, S.-H. D. Chan, C.-H. Leung, *Angew. Chem. Int. Ed.* **2013**, *52*, 7666–7682.

[37] C. Liu, C. Yang, L. Lu, W. Wang, W. Tan, C.-H. Leung, D.-L. Ma, *Chem. Commun.* **2017**, *53*, 2822–2825.

[38] K. Vellaisamy, G. Li, C.-N. Ko, *Chem. Sci.* **2018**, *9*, 1119–1125.

[39] L. M. Haiber, M. Kufleitner, V. Wittmann, *Front. Chem.* **2021**, *9*, 654932.

[40] C. Würth, M. Grabolle, J. Pauli, M. Spieles, U. Resch-Genger, *Nat Protoc* **2013**, *8*, 1535–1550.

[41] A. I. Solomatina, S.-H. Su, M. M. Lukina, V. V. Dudenkova, V. I. Shcheslavskiy, C.-H. Wu, P. S. Chelushkin, P.-T. Chou, I. O. Koshevoy, S. P. Tunik, *RSC Adv.* **2018**, *8*, 17224–17236.

[42] C. Simon, C. Lion, H. Ahouari, H. Vezin, S. Hawkins, C. Biot, *Chem. Commun.* **2021**, *57*, 387–390.

[43] 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.

[44] P. A. Gilormini, C. Lion, D. Vicogne, Y. Guérardel, F. Foulquier, C. Biot, *J Inherit Metab Dis* **2018**, *41*, 515–523.

[45] A. K. Estandarte, S. Botchway, C. Lynch, M. Yusuf, I. Robinson, *Sci Rep* **2016**, *6*, 31417.

[46] F. Domart, P. Cloetens, S. Roudeau, A. Carmona, E. Verdier, D. Choquet, R. Ortega, *eLife* **2020**, *9*, e62334.

[47] I. Rodushkin, E. Engström, A. Stenberg, D. C. Baxter, *Anal Bioanal Chem* **2004**, *380*, 247–257.

[48] F. Fus, Y. Yang, H. Z. S. Lee, S. Top, M. Carriere, A. Bouron, A. Pacureanu, J. C. da Silva, M. Salmain, A. Vessières, P. Cloetens, G. Jaouen, S. Bohic, *Angew Chem Int Ed* **2019**, *58*, 3461–3465.

[49] F. Dubar, S. Bohic, C. Slomianny, J.-C. Morin, P. Thomas, H. Kalamou, Y. Guérardel, P. Cloetens, J. Khalife, C. Biot, *Chem. Commun.* **2012**, *48*, 910–912.

[50] C. Bissardon, S. Reymond, M. Salomé, L. André, S. Bayat, P. Cloetens, S. Bohic, *JoVE* **2019**, 60461.

[51] C. Gramaccioni, Y. Yang, A. Pacureanu, N. Vigano, A. Procopio, P. Valenti, L. Rosa, F. Berlutti, S. Bohic, P. Cloetens, *Anal. Chem.* **2020**, *92*, 4814–4819.

CHAPITRE 4 – AUTRES APPLICATIONS ET PROJETS ANNEXES

I. SYNTHÈSE DE RAPPORTEURS ACIDES SIALIQUES CMP ACTIVÉS POUR LA CARACTÉRISATION DE SIALYLTRANSFÉRASES

Comme vu précédemment, les rapporteurs chimiques de type monosaccharides sont des outils puissants, qui peuvent fournir des informations précieuses sur les voies métaboliques : leur niveau d'expression dans un contexte donné, leurs mécanismes de régulations, ou la manière dont elles sont impliquées dans le mécanisme de certaines pathologies. Les applications de la chimie bioorthogonale en cellules vivantes ou en tissus vivants sont nombreuses, et les développements technologiques visent souvent à améliorer les propriétés physico-chimiques, biologiques ou photophysiques des rapporteurs métaboliques et sondes moléculaires utilisées. Cependant, les applications de cette chimie à des études in vitro, en biochimie, sont également très pertinentes, bien que moins mises en lumière dans les communications de diffusion au grand public. En effet, elles permettent souvent de donner des éléments de réponse à des questions biologiques parfois plus précises que les études en bioimagerie : on peut par exemple étudier la spécificité de substrat ou la cinétique d'une enzyme donnée en exploitant la chimie bioorthogonale, combinée à des techniques biochimiques classiques. C'est dans cette optique que j'ai également participé à un projet collaboratif entre notre équipe et celle du Dr Anne Harduin-Lepers à l'UGSF, portant sur la caractérisation de sialyltransférases de différents organismes.

Au cours de leur métabolisation, les acides sialiques sont activés sous forme d'esters de cytidine monophosphate, avant d'être transférés sur un glycane ou un polysaccharide en cours de biosynthèse par une sialyltransférase. Les CMP-Sia sont des intermédiaires métaboliques transitoires clés. Ils sont très réactifs, et par conséquent peu stables et sensibles à l'hydrolyse/solvolyse. Les méthodes de synthèse disponibles jusque-là dans la littérature requéraient des étapes de purification longues et fastidieuses (chromatographie d'échange d'ion, suivi de chromatographie d'exclusion stérique) en chambre froide pour réduire la dégradation par hydrolyse. Les CMP-Sia ainsi obtenus sont souvent peu purs

car partiellement hydrolysés, en mélange avec de la cytidine monophosphate, avec des rendements pauvres. Ces dernières années, une méthodologie permettant de synthétiser des acides sialiques activés (CMP-Sia) purs, sans aucune étape de purification, et avec un rendement quantitatif a été mise au point par notre équipe au laboratoire. Exploitant un suivi cinétique de la synthèse chimio-enzymatique des CMP-Sia par RMN du phosphore ^{31}P à partir de l'acide sialique correspondant et de cytidine triphosphate, elle permet l'accès à toutes sortes de dérivés CMP-Sia naturels ou modifiés.[104]

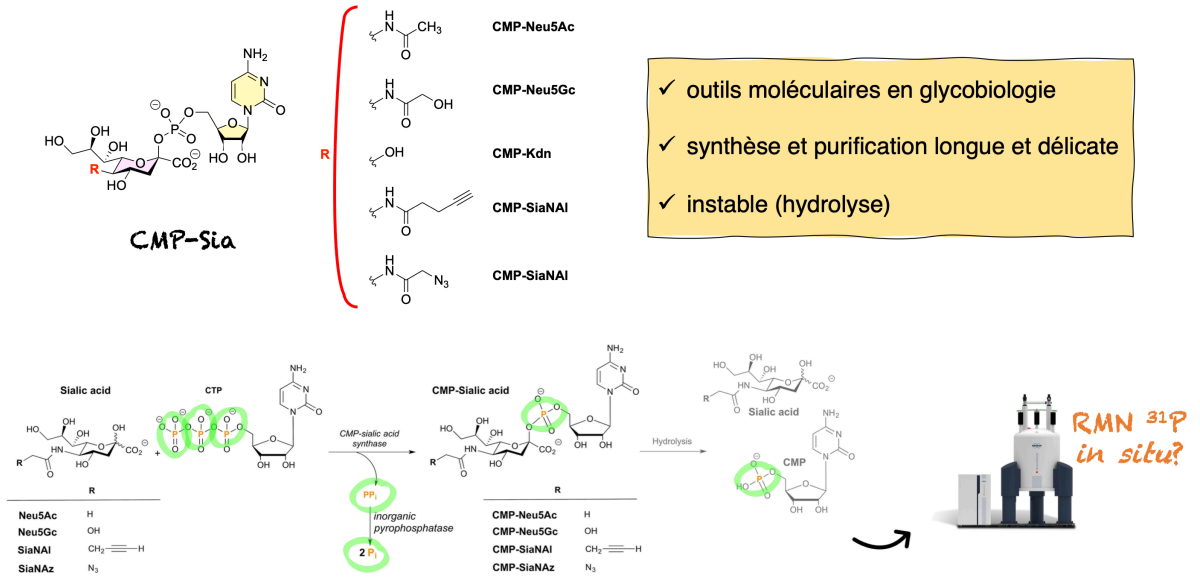


Figure 20 Rapporteurs chimiques de type CMP-acides sialiques utilisés pour l'étude de la biosynthèse des chaînes poly acides sialiques. Ces rapporteurs chimiques ont été synthétisés via une réaction enzymatique *in vitro*. La RMN du phosphore ^{31}P permet de suivre la formation du lien phosphodiester et ainsi l'avancée de la réaction.

Notre équipe a également mis au point un assay en microplaque permettant d'utiliser des dérivés CMP-Sia équipés d'un tag bioorthogonal pour tester l'activité des sialyltransférases.[105] C'est ainsi que j'ai pu générer une série d'acides sialiques activés à partir des acides sialiques synthétisés pendant ma thèse (SiaNAl, SiaNAz) et à partir d'acides sialiques commerciaux (Neu5Ac, Neu5Gc, Kdn), qui ont ensuite été utilisés par l'équipe du Dr Harduin-Lepers pour caractériser des sialyltransférases de poissons salmonidés, aux propriétés encore inconnues.

Plus précisément, les polysialyltransférases humaines ST8Sia II et ST8Sia IV catalysent le transfert de plusieurs résidus de Neu5Ac sur les glycoprotéines, formant des homopolymères jouant un rôle essentiel dans différents processus physiologiques. Chez

les salmonidés, un ensemble hétérogène de polymères d'acides sialiques a été décrit dans l'ovaire et à la surface des œufs, et trois gènes, st8sia4, st8sia2-r1 et st8sia2-r2, ont été identifiés comme pouvant être impliqués dans ces hétéropolymères. Les trois polysialyltransférases du *Coregonus maraena*, un salmonidé, ont été clonées, exprimées de manière recombinante dans les cellules HEK293, et la ST8Sia IV a été caractérisée sur le plan biochimique. L'essai de sialyltransférase en microplaque et le substrat donneur non naturel CMP-SiaNAl ont été utilisés pour démontrer l'activité enzymatique et optimiser les réactions de polysialylation. La polysialylation a également été réalisée avec des substrats donneurs naturels CMP-Neu5Ac, CMP-Neu5Gc et CMP-Kdn, synthétisés par la méthode décrite ci-dessus, dans des essais en solution et avec des cellules. Des analyses structurales des produits polysialylés ont également été réalisées à l'aide de l'anticorps monoclonal anti-polySia 735 et des approches d'endoneuraminidase N et de HPLC. Cette étude a mis en évidence des différences de spécificité de substrat entre les polysialyltransférases humaines et de salmonidés, avec notamment des différences marquées dans l'utilisation des donneurs et la capacité des enzymes de poisson à générer des hétéropolymères. Cette étude suggère en outre une évolution des fonctions biologiques du polySia.

scientific reports

Check for updates

OPEN

Salmonid polysialyltransferases to generate a variety of sialic acid polymers

Mathieu Decloquement[1], Marzia Tindara Venuto[2], Virginie Cogez[1], Anna Steinmetz[2], Céline Schulz[1], Cédric Lion[1], Maxence Noel[1], Vincent Rigolot[1], Roxana Elin Teppa[1], Christophe Biot[1], Alexander Rebl[3], Sebastian Peter Galuska[2]✉ & Anne Harduin-Lepers[1,4]✉

The human polysialyltransferases ST8Sia II and ST8Sia IV catalyze the transfer of several Neu5Ac residues onto glycoproteins forming homopolymers with essential roles during different physiological processes. In salmonids, heterogeneous set of sialic acids polymers have been described in ovary and on eggs cell surface and three genes *st8sia4*, *st8sia2-r1* and *st8sia2-r2* were identified that could be implicated in these heteropolymers. The three polysialyltransferases from the salmonid *Coregonus maraena* were cloned, recombinantly expressed in HEK293 cells and the ST8Sia IV was biochemically characterized. The MicroPlate Sialyltransferase Assay and the non-natural donor substrate CMP-SiaNAl were used to demonstrate enzyme activity and optimize polysialylation reactions. Polysialylation was also carried out with natural donor substrates CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn in cell-free and cell-based assays and structural analyses of polysialylated products using the anti-polySia monoclonal antibody 735 and endoneuraminidase N and HPLC approaches. Our data highlighted distinct specificities of human and salmonid polysialyltransferases with notable differences in donor substrates use and the capacity of fish enzymes to generate heteropolymers. This study further suggested an evolution of the biological functions of polySia. *C. maraena* ST8Sia IV of particular interest to modify glycoproteins with a variety of polySia chains.

The nine-carbon α-keto acidic monosaccharides *N*-acetylneuraminic acid (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc) and 2-keto-3-deoxy-nonulosonic acid (Kdn) are the three most common sialic acids (Sias) found in vertebrates, except in human tissues predominantly featuring Neu5Ac residues[1]. Sias at the outmost end of glycolipids and glycoproteins are linked to galactose (Gal) and *N*-acetylgalactosamine (GalNAc) via α-2,3 or α-2,6 bonds or linked to α-2,8 on other Sias forming anionic linear polymers named α-2,8-polysialic acids (polySia). The structural diversity and complexity of polySia is based on their degree of polymerization (DP), the nature of Sias and their possible modifications[2]. Homopolymers of Neu5Ac residues found on the neural cell adhesion molecule (NCAM) has been intensively studied in the mammalian central nervous system where it plays an essential role during post-natal brain development[3,4]. However, polysialylation takes place also in other organs[5]. For instance, in the male reproductive system, polySia is involved in the postnatal differentiation process of peritubular smooth muscle cells[6,7]. Moreover, polySia was found in semen and may influence immunological processes in the female reproductive tract[8].

In contrast to mammals, a broad polySia diversity has been described in salmonid fish. Seminal glycomics studies have shown the presence of homo- and hetero-polymers containing Neu5Ac and/or Neu5Gc or Kdn residues with a DP ≤ 25 Sia residues[9–11] potentially capped by Kdn[12]. In cortical alveoli of salmonid eggs, the polysialylated glycoprotein (PSGP) seems to be the major carrier of polySia[9–11,13] and these polySia chains have been implicated in sperm acrosomal reaction preventing polyspermy during fertilization[14,15]. Moreover, these polySias, which are α2,6-linked to the inner GalNAc residue of *O*-glycans, provide a protective effect against pathogens in the reproductive system[16]. More recent studies pointed to the presence of large quantities of polySia in fish serum[17] and also on various fish oogonia[18] despite different poly-α-2,8-sialyltransferases setups[19].

[1]Univ. Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, 59000 Lille, France. [2]Institute of Reproductive Biology, Research Institute for Farm Animal Biology (FBN), Wilhelm-Stahl-Allee 2, 18196 Dummerstorf, Germany. [3]Institute of Genome Biology, Research Institute for Farm Animal Biology FBN, Wilhelm-Stahl-Allee 2, 18196 Dummerstorf, Germany. [4]Unité de Glycobiologie Structurale et Fonctionnelle, UMR CNRS 8576, Faculté des sciences et Technologies, Univ. Lille, 59655 Villeneuve d'Ascq, France. ✉email: galuska.sebastian@fbn-dummerstorf.de; anne.harduin-lepers@univ-lille.fr

Biosynthetic enzymes (*i.e.* sialyltransferases) catalyzing the transfer of Sia residues from an activated donor substrate CMP-sialic acid (CMP-Sia, Fig. 1) onto glycolipids or glycoproteins are classified into ST3GAL, ST6GAL, ST6GALNAC and ST8SIA[20,21]. Phylogenetic distribution of sialyltransferases was assessed in Metazoan and found to be extremely diverse in teleost fishes[19,22,23]. Two poly-α-2,8-sialyltransferase genes *ST8SIA2* and *ST8SIA4* of the ST8SIA family have been described in the human genome[22], whereas homologous genes in fish genomes show an uneven distribution[19]. In salmonids like the whitefish *Coregonus maraena* (*C. maraena*) or *Oncorhynchus mykiss* (*O. mykiss*), in addition to *st8sia4*, two *st8sia2*-related genes were identified as a result of the species-specific genome duplication event R4-SGD that took place 80 million years ago. These genes showed distinct expression profiles in fish tissues with *st8sia2r-1* and *st8sia2r-2* genes displaying their highest mRNA levels in gonads and telencephalon, and in gonads and spleen, respectively, and *st8sia4* gene showing low mRNA levels in the brain, head kidney, gills, gonads and spleen[19]. Changes in the distribution of these genes compared to their human counterpart, and a weaker expression of the *st8sia4* gene further suggested that these fish genes likely define new subfamilies with novel and as yet unknown enzymatic specificities[19,24,25]. On the other hand, the orthologue of the *st8sia4* gene was lost in Neoteleosteii like the medaka *Oryzias latipes* and the orthologue of the *st8sia2* gene was lost in Siluriformes like *Ictalurus punctatus* and Esociformes like *Esox lucius*[19,26]. This further denoted evolution of these enzymes with potentially new functions for the remaining paralogue. The protein sequence of sialyltransferases is characterized by the presence of conserved sialylmotifs L, S, III and VS involved in substrates binding and in catalysis[22,27]. In addition, ST8Sia II and ST8Sia IV show unique conserved PolyBasic Region (PBR, 35 amino acids (aa)) and PolySialylTransferase Domain (PSTD, 32 aa) required for the selective recognition of NCAM and the catalytic activity of the human enzymes, respectively[28–30]. Preliminary sequence-based analysis using multiple sequence alignments and 3D-modeling of the PSTD unveiled potential changes of function in the salmonids polysialyltransferases[19].

To better understand the molecular evolution of polysialyltransferases that potentially led to changes in substrate recognition and acquisition of new biological functions in human, each gene identified in the *C. maraena* genome (*Cma*ST8Sia II-r1, *Cma*ST8Sia II-r2 and *Cma*ST8Sia IV) was cloned and produced as a recombinant enzyme. Functional characterization of the salmonid enzymes highlighted unique catalytic specificities of *Cma*ST8Sia IV for the in vitro synthesis of various polymers of Sias and the exopolysialylation of the cell surface glycans.

Results and discussion

Molecular cloning, production and enzymatic activities of the Cma polysialyltransferases. The deduced protein sequence of the newly identified *Cma*ST8Sia II-r1 and *Cma*ST8Sia II-r2 shared 63.8% and 62.7% sequence identity with their human homolog *Hsa*ST8Sia II, and *Cma*ST8Sia IV shared 78.8% sequence identity with the *Hsa*ST8Sia IV and all presented the four sialylmotifs and the polysialylmotifs PBR and PSTD characteristic of polysialyltransferases[28,31]. To achieve functional analyses, cDNA of *st8sia2r-1*, *st8sia2r-2* and *st8sia4* were PCR amplified from *C. maraena* gonads and brain, and soluble forms of the enzymes lacking their cytoplasmic and transmembrane domains were constructed in the p3 × FLAG-CMV9 expression vector. Since no good transfection yield could be achieved in the homologous salmonid fish cell line CHSE-214 system, secretion of the recombinant protein was optimized outside transiently transfected HEK293 cells, which remain the preferred host to produce difficult to express glycosyltransferases[32]. Cell lysates and culture medium collected 72 h post-transfection were run on SDS-PAGE and the recombinant *N*-terminally FLAG-tagged Δ28ST8Sia IV was detected with an anti-FLAG antibody both in the medium (M) and cell lysate (L) by Western Blot (WB) confirming a good level of protein production and secretion efficiency in this system (Fig. 2A; Supplementary Fig. S1). No signal could be detected with the empty plasmid transfected cells, and the relative expression and secretion levels of each fish ST8Sia II-related protein was low, as already described for the human glycoenzymes[32]. Several bands ranging from 40 to 52 kDa were detected in L and a major band around 52 kDa in M suggesting post-translational modification of the recombinant fish proteins. *N*-glycosidase F (PNGase F) treatment induced a shift to the expected molecular weight confirming the prediction (Fig. 2A; Supplementary Fig. S1).

Then, the enzymatic activities of these recombinant proteins were assessed with the newly set up Microplate Sialyltransferase Assay (MPSA) and the crude enzyme sources. The chemo-enzymatically synthesized

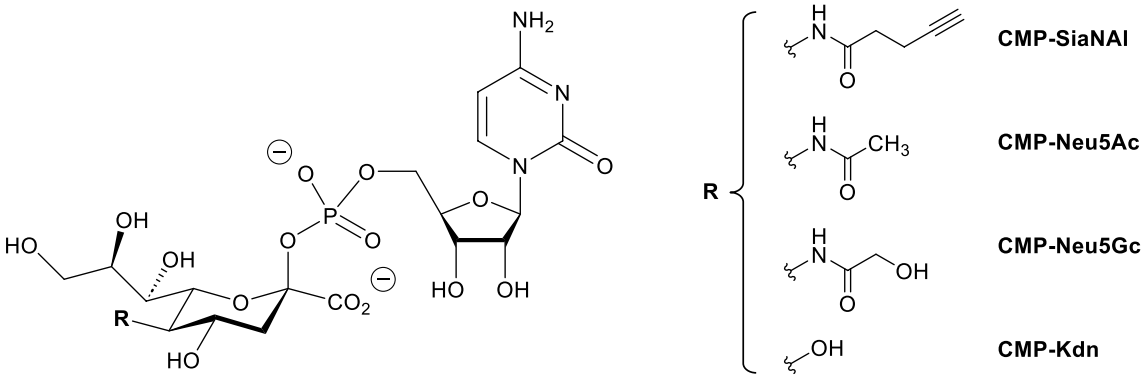


Figure 1. Chemical structure of the various CMP-Sias synthetized and used in this study: CMP-Neu5Ac, CMP-Neu5Gc, CMP-Kdn and CMP-SiaNAl.

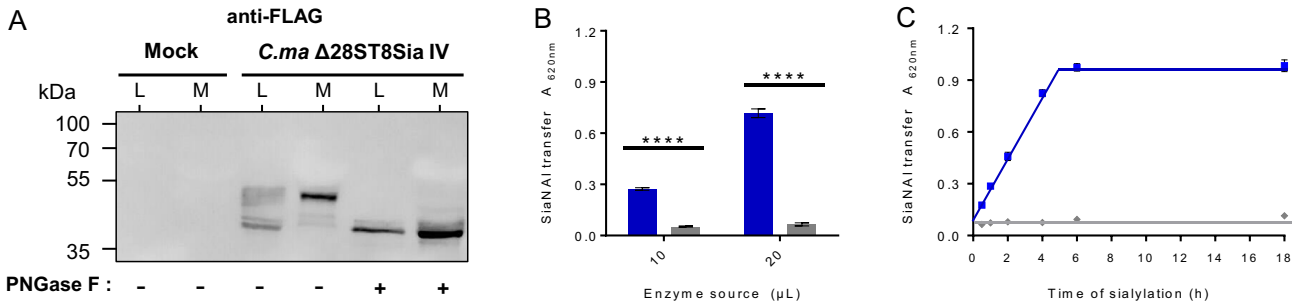


Figure 2. (**A**) Production of *Cma*ST8Sia IV in transfected HEK293 cells. HEK293 cells were transiently transfected with either an empty vector (Mock) or with the p3 × FLAG-CMV9-Δ28ST8Sia IV (*CmaST8Sia IV*). Seven µg of proteins from each cell lysate (L) and 20 µL of each cell culture medium (M) with or without 50 U PNGase F treatment were separated on 8% SDS-PAGE and proteins were transferred on nitrocellulose membrane. WB was carried out with the anti-3 × FLAG antibody BioM2 (1 µg/mL). Original blot is presented in supplementary Fig. S9. Molecular weight markers are indicated on the left side. (**B**) Enzyme activity of recombinant *Cma*ST8Sia IV using the MPSA assay. A MPSA was carried out 4 h at 27 °C with 400 ng of the fish CD166/ALCAM, 100 µM CMP-SiaNAl and 10 or 20 µL of either the recombinant *Cma*ST8Sia IV (blue) or Mock (grey). Azido-PEG3-biotin was covalently attached to the alkyne group of transferred SiaNAl and biotin detected by the anti-biotin-HRP antibody. TMB (3,3′,5,5′-Tetramethylbenzidine) substrate was used and absorbance was measured at 620 nm[33]. Error bars in black represent SEM (n = 5), *p*-value < 0.0001. (**C**) Time course of sialylation by *Cma*ST8Sia IV in MPSA. Sialylation reactions were conducted at 27 °C using 400 ng of CD166/ALCAM and 100 µM CMP-SiaNAl using 20 µL of the fish enzyme *Cma*ST8Sia IV (blue squares) or of the Mock (grey diamonds) for 0.5, 1, 2, 4, 6 or 8 h (n = 2).

CMP-SiaNAl reporter molecule (Fig. 1) was added as a substrate donor, as already described for the human ST6Gal-I and ST3Gal-I[33]. Since only a limited set of natural polysialylated proteins like NCAM have been described in the past[4], we assayed a fish glycoprotein substrate that could serve as a suitable acceptor for these vertebrate polysialyltransferases, namely the fish CD166/ALCAM with sialylated bi-, tri- and tetra-antennae *N*-glycans (Supplementary Table S-1). This glycoprotein was coated on the microplate and sialylation reactions were carried out 4 h at 27 °C with each enzyme. Azido-PEG3-biotin was then covalently attached to the alkyne group of transferred SiaNAl units and biotin was detected using an HRP-conjugated anti-biotin antibody. As shown in Fig. 2B, significant and reproducible data were obtained on both acceptors showing that the fish ST8Sia IV is active, stable and efficiently transfers SiaNAl from CMP-SiaNAl. Initial velocity was proportional to the enzyme amount up to 20 µL and sialylated product appearance was linear up to 4 h of sialylation (Fig. 2C). However, the fish ST8Sia II-r1 and ST8Sia II-r2 demonstrated no chemoenzymatic modification of these substrates under these conditions (Supplementary Fig. S1). This is in line with a previous study by Kitajima's group reporting very low levels of enzymatic activity of recombinant polysialyltransferases from *O. mykiss*[24]. Although still speculative, this observation also corroborates the previously reported inhibition of the human ST8Sia II using a propyl-C5 modified Sia precursor in vivo[34].

Therefore, we focused on the fish ST8Sia IV to assess the robustness of the enzyme activity and set up optimum polysialylation conditions, the influence of pH (4.8–8.6, measured at 27 °C) and temperature (0–60 °C) was studied on enzyme velocity. An optimal temperature of 27 °C was found for the fish ST8Sia IV, which remained highly active at high temperatures (Fig. 3A) suggesting that during evolution, fish sialyltransferases could have acquired a higher level of structural flexibility to adapt to variable environment[35]. Our data demonstrated a conserved optimum pH around 6.4 for this fish polysialyltransferase (Fig. 3B), comparable to those determined in another enzymatic assay for the human ST8Sia II[36]. Using assay conditions set at pH 6.4, 27 °C, kinetic parameters of the recombinant fish enzyme towards CMP-SiaNAl were calculated. An apparent Michaelis constant (Km) value of 17.60 ± 2.52 µM was calculated for fish ST8Sia IV according to the Michaelis–Menten model (Fig. 3C). This value is in the same order of magnitude as found for the human ST6Gal I and ST3Gal I towards CMP-SiaNAl[33]. To assess the enzyme specificity and the acceptor substrate preference of the fish ST8Sia IV, several glycoproteins with various *N*- and *O*-glycan structures (Supplementary Table S-1) were used as acceptors and coated on microplates for sialylation assays. *Cma*ST8Sia IV was active transferring SiaNAl on bovine fetuin, fish CD166/ALCAM, bovine submaxillary mucin (BSM), human DNase I, human NRP2, fish polysialoglycoprotein PSGP-H, PSGP-L and human orosomucoid (α(1)-acid glycoprotein) (Fig. 3D).

Structural assessment of polySias generated by the Cma polysialyltransferases. To explore the polysialyltransferases' ability to use natural CMP-Sia donors and successively add several Sia residues, we achieved sialylation reaction and detected polysialylated products after SDS-PAGE and WB with the mAb735 antibody, which recognizes polyNeu5Ac with a DP ≥ 8 residues[37]. We firstly checked that no pre-existing polySia could be detected with the mAb735 on the various reagents prior sialylation reaction (Fig. 4A). Then, sialylation reactions were performed with the fish ST8Sia IV and its human orthologue, 100 µM of chemo-enzymatically synthesized CMP-Neu5Ac (Fig. 1) and 6 µg of CD166/ALCAM for 0.5 to 8 h (Supplementary Fig. S2). Within 30 min, an extensive smear ranging from 70 to more than 250 kDa could be detected confirming that both the human and fish ST8Sia IV could synthetize polyNeu5Ac chains. In addition, we verified the polysialylation

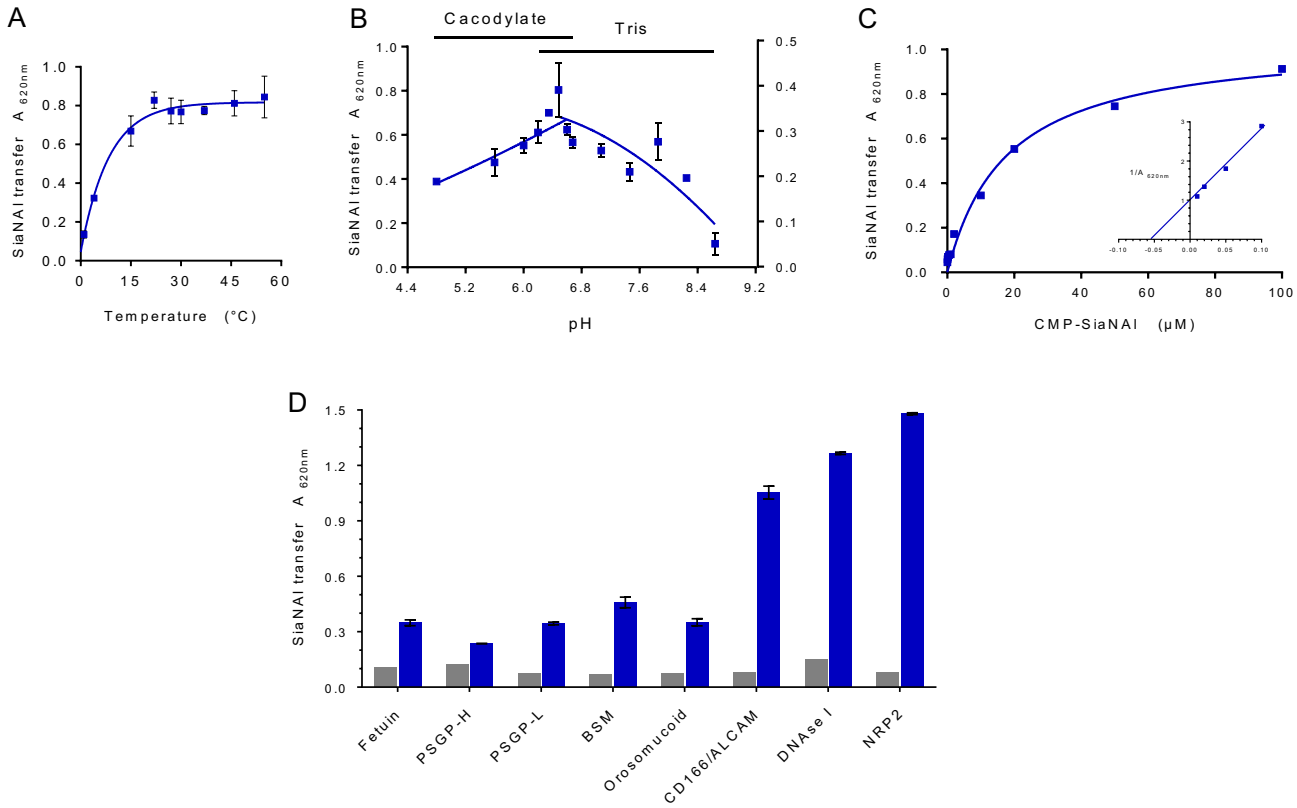


Figure 3. Optimization of sialylation conditions in the MPSA. (**A**) Effect of temperature on the sialylation by *Cma*ST8Sia IV. Sialylation was performed 4 h on 400 ng of CD166/ALCAM with 100 µM CMP-SiaNAl and 20 µL of *Cma*ST8Sia IV at different temperatures from 0 °C to 60 °C (n = 2). (**B**) Effect of pH on the sialylation by *Cma*ST8Sia IV. Sialylation reactions were performed 4 h at 27 °C on 400 ng of CD166/ALCAM with 100 µM CMP-SiaNAl and 20 µL of *Cma*ST8Sia IV at different pH from 4.8 to 8.6 (n = 2). (**C**) Determination of kinetic parameters of ST8Sia IV towards donor substrates. Sialylation reactions were carried out in cacodylate buffer pH 6.4, at 27 °C for 4 h with 20 µL of fish ST8Sia IV. CMP-SiaNAl was used at various concentrations ranging from 0 to 100 µM with 400 ng of CD166/ALCAM (n = 3). Michaelis–Menten constants were determined graphically with GraphPad. *Cma*ST8Sia IV shows an apparent Km of 17.60 + / − 2.52 µM towards donor substrate. (**D**) Enzymatic activities on several acceptors in the MPSA. Sialylation reactions were performed 4 h at 27 °C using 400 ng of various acceptors (fetuin, PSGP-H, PSGP-L, BSM, orosomucoid, CD166/ALCAM, DNase I and NRP2) with 100 µM of CMP-SiaNAl and with 20 µL of *Cma*ST8Sia IV (blue) or Mock (grey). Error bars represent SEM (n = 2).

activity of fish ST8Sia IV by HPLC analysis. As a first step, polysialylated CD166/ALCAM was enriched with an enzymatically inactive form of the endoneuraminidase N (endoN) coupled to magnetic beads. This is possible because endoN contains an oligo/polySia-binding domain[38]. The polysialylation and affinity-precipitation were monitored by WB against polySia using an aliquot of the eluate (5%) (Fig. 4B and C). The remaining sample (90%) was analyzed by a HPLC approach to determine the chain length. To this end, polySia chains were released and labeled with 1,2-diamino-4,5-methylendioxy-benzene (DMB). The resulting fluorescently labeled polymers were separated according to their DP by anion exchange chromatography. In line with the WB results, polySia chains consisting of more than 7 Sias were observed (Fig. 4D). Thus, the polysialylation activity of fish ST8Sia IV was verified by two independent methods.

To evaluate the fish and human ST8Sia IV ability to use various Sia donors, sialylation reactions were performed in cell-free or cell-based assays with 100 µM natural CMP-Sia. Firstly, to avoid background and sialylation inhibitor side products due to CMP-Sia hydrolysis, we chemo-enzymatically synthesized various natural donor substrates CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn (Fig. 1) using either the CMP-Sia synthetase (CSS) from *Neisseria meningitidis*[39] or the recently described rainbow trout CMP-Kdn synthetase (rtCSS)[40] and checked efficiency of the reaction by ^{31}P NMR (Supplementary Fig. S3). After sialylation reaction using CD166/ALCAM as an acceptor for polySia chains, sialylated products were run on SDS-PAGE, transferred on nitrocellulose membranes and detected with the mAb735. In addition, endoN treatment was used as an additional control to remove the newly formed polySias with a DP ≥ 8[38]. Broad signals between 100 and 250 kDa were visualized as expected for both enzymes with CMP-Neu5Ac on WB using the mAb735 (Fig. 5Aa). Remarkably, a smear of polySia could be detected to variable extents with the mAb735 after polysialylation with CMP-Neu5Gc (Fig. 5Ab) or CMP-Kdn (Fig. 5Ac) and the *Cma*ST8Sia IV, which could be removed by endoN treatment (Fig. 5A). In contrast to

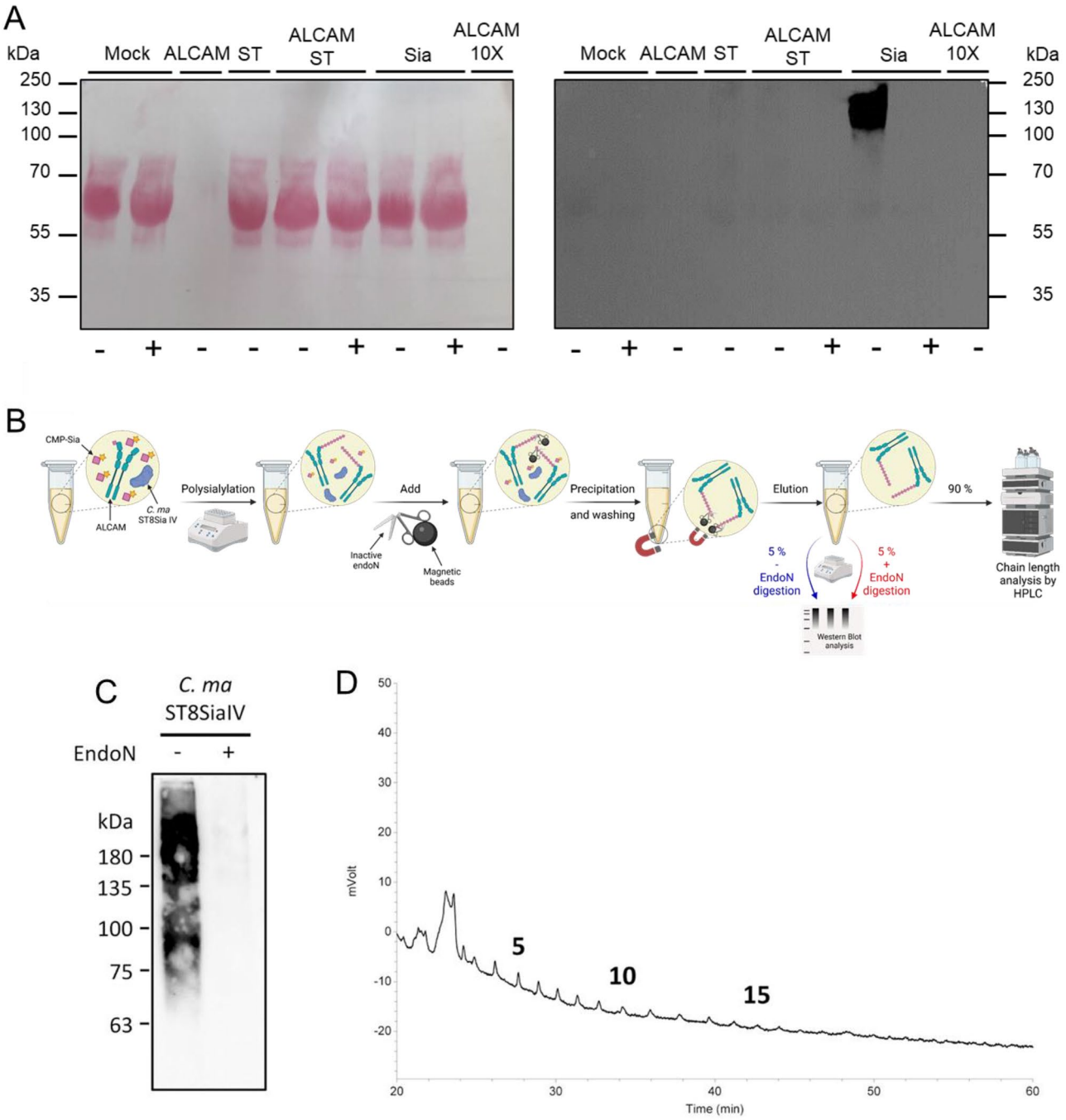


Figure 4. Chain length analysis of CD166/ALCAM after polysialylation with CMP-Neu5Ac by *Cma*ST8Sia IV. (**A**) Prior to sialylation, no polySia can be detected by the mAb735 on the various actors of the sialylation reaction *i.e.* CD166/ALCAM (ALCAM) or concentrated ALCAM (ALCAM 10×), HEK293 cell culture medium (Mock) or *Cma*ST8Sia IV (ST). PolySia is detected only after 4 h of polysialylation of CD166/ALCAM with the *Cma*ST8Sia IV (Sia). Original blot is presented in supplementary Fig. S9 (**B**) Illustration of the precipitation-workflow to enrich polySia for HPLC analysis using inactive endoN. Created with BioRender.com. (**C**) An aliquot (5% per lane) of the eluate was used for WB to monitor the polysialylation and precipitation. Original blot is presented in supplementary Fig. S9 (**D**) The remaining precipitate (90%) was used for "mild" DMB labeling. Resulting fluorescently labeled sialic acid polymers were separated via anion-exchange chromatography. The chain length is given for selected peaks.

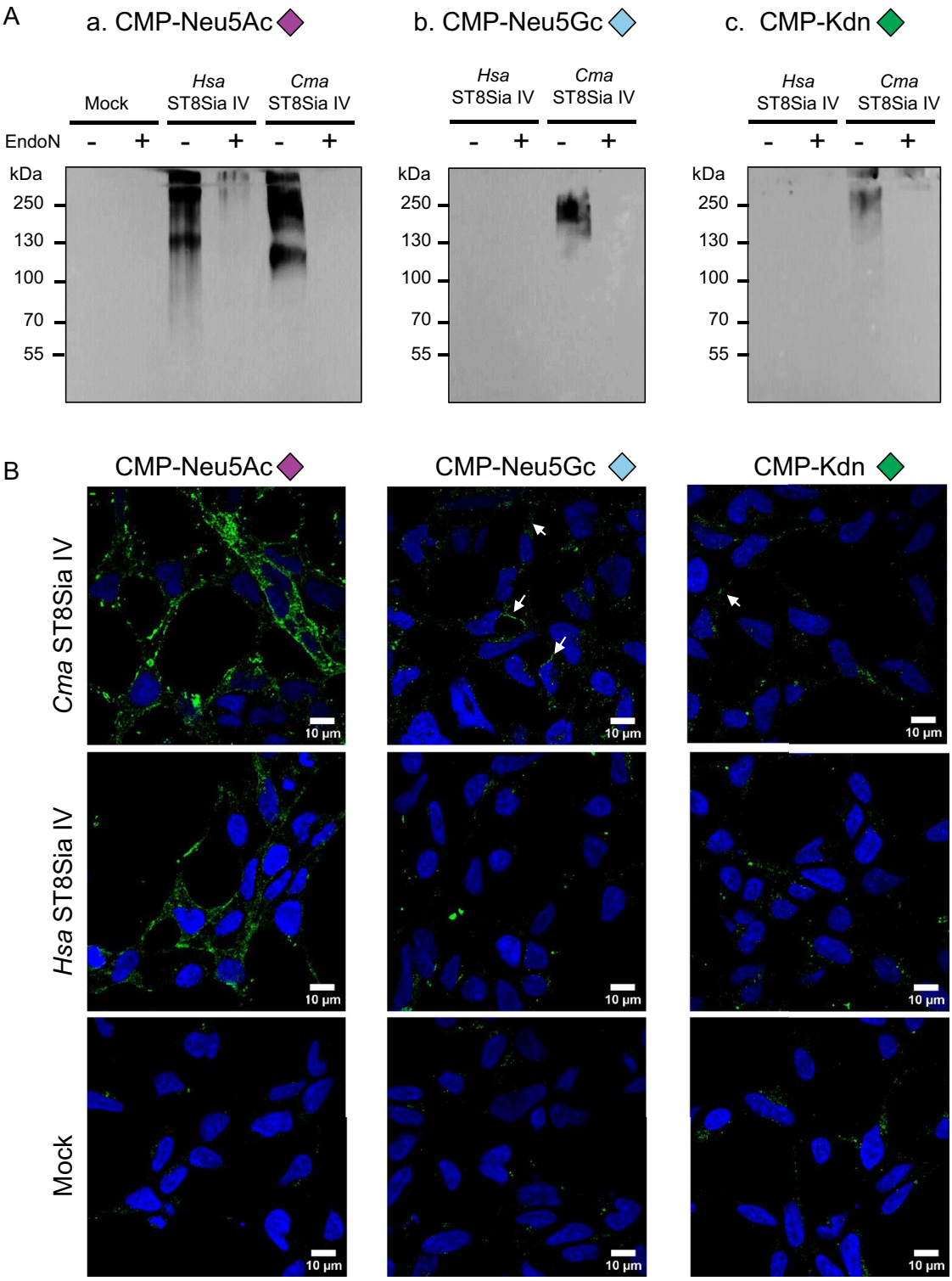


Figure 5. CD166/ALCAM polysialylation achieved with human and *Cma* polysialyltransferases and different CMP-sialic acids. (**A**) Polysialylation with ST8Sia IV enzymes visualized on WB with the mAb735**.** Sialylation reactions were performed 4 h at 27 °C with 100 µM of CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn (purple, light blue and green diamonds, respectively) on 6 µg of CD166/ALCAM with 150 ng of *Hsa*ST8Sia IV or 20 µL of either *Cma*Δ28ST8Sia IV or Mock. Treatment with endoN for 1 h at RT was used as a control to remove polysialylation. Sialylation products were heated 5 min at 60 °C in Laemmli buffer, separated on 6% SDS-PAGE and transferred onto PVDF membranes. Blots were probed with mAb735 (0.5 µg/mL). Original blots are presented in supplementary Fig. S9. Molecular weight markers are indicated on the left side. Immunoblots were incubated with femtoECL and detected after 5 s exposure on for a) CMP-Neu5Ac, 30 s for b) CMP-Neu5Gc and 180 s for c) CMP-Kdn panels. (**B**) Exosialylation of HEK293 cells with *Hsa*ST8Sia IV and *Cma*ST8Sia IV**.** After EndoN treatment, sialylation reactions were carried out 4 h at 27 °C on fixed cells with 40 µL *Cma*ST8Sia IV or 200 ng of the *Hsa*ST8Sia IV or 40 µL of the Mock and 500 µM of donor substrates: a) CMP-Neu5Ac, b) CMP-Neu5Gc and c) CMP-Kdn. PolySia on cell surface was detected by mAb735 (green), nuclei were stained with DAPI (blue). Fluorescence was detected using a Carl Zeiss confocal microscope. Scale bars: 10 µm (Zoom 2×). White arrows indicate polySia formed by *Cma*ST8Sia IV with CMP-Neu5Gc and CMP-Kdn.

fish ST8Sia IV, no polySia could be detected after polysialylation with CMP-Neu5Gc (Fig. 5Ab) or CMP-Kdn (Fig. 5Ac) by the human polysialyltransferase denoting distinct substrate specificities between each orthologue.

Comparable data were obtained with a cell-based approach using the human and fish enzymes to selectively introduce various natural Sia donors on cell surface glycoconjugates of HEK293 cells (Fig. 5B). Prior to exogenous polysialylation, we ensured that HEK293 cells displayed NCAM and checked its polysialylation status (Supplementary Fig. S4). Then, polySia of paraformaldehyde fixed HEK293 cells was degraded by endoN. After exogenous polysialylation, polySia could be detected with mAb735 using CMP-Neu5Ac for both enzymes, whereas lower levels of polySia could be detected using CMP-Neu5Gc only with the *Cma*ST8Sia IV, but not with the *Hsa*ST8Sia IV (Supplementary Figs. S5–S7). No conclusive results could be obtained after exo-polysialylation with CMP-Kdn and the fish, human or mock enzyme sources (Fig. 5B).

As our experimental data pointed to the ability of mAb735 to detect polyNeu5Gc and polyKdn, and endoN to cleave all polySias, we further checked the reliability of each polymer interactions with mAb735 through computational modeling. A previous study reported crystallization of the single chain Fv fragment of mAb735 (scFv735) in complex with an octasialic acid polymer of α2,8-linked Neu5Ac (PDB entry 3WBD)[37]. In this complex, two antibody molecules (A and B) facing each other are associated with an octasialic acid. As illustrated in Fig. 6A, one scFv735 molecule (B) interacts with Sia2–Sia4, whereas the other (A) interacts with Sia6 –Sia8. The antigen-binding site is comprised of six complementarity-determining regions (CDR) L1, L2, L3 and H1, H2, and H3 interacting with three consecutive Sias through direct and indirect hydrogen bonds. Six amino acids are involved in direct hydrogen bonds; four of them Tyr37, Arg55, Tyr159 and Asp232 are critical for antigen recognition, whereas Tyr160 and Tyr179 are less important. In addition, indirect interactions involve 11 structured water molecules. To investigate to what extent analogous interactions could be found between the antibody and polySias made of eight Neu5Gc or Kdn, we modified the Neu5Ac to Neu5Gc/Kdn molecules in the PDB 3WBD. We showed that in addition to the six direct hydrogen bond interactions described for polyNeu5Ac (Fig. 6B), five others could be established between Tyr39, Ala230, Gly227, Arg225 and Asn33 for polyNeu5Gc (Fig. 6C), whereas five out of the six aa critical for polyNeu5Ac recognition were retrieved for polyKdn (Fig. 6D, Supplementary Fig. S8) supporting our experimental data. Altogether, we provided here for the first time, computational and experimental evidences for the recognition to various levels of various polySia chains using the mAb735.

To check the transfer of Neu5Ac, Neu5Gc and Kdn by *Cma*ST8Sia IV with an antibody independent approach, a HPLC-based strategy was applied (Fig. 7A). To this end, polysialylation of CD166/ALCAM was performed in centrifugal filter units. After polysialylation, remaining CMP-Sia was removed before polySia degradation by endoN. Untreated samples (without endoN treatment) were used as negative control. Polysialylation as well as degradation by endoN were successfully performed in the centrifugal filter units as checked by WB (Fig. 7B). After a final filtration step, the flow-throughs were analyzed by HPLC. Newly added Sia residues could be determined comparing the Sia content in the flow-through of untreated and endoN treated samples. When CMP-Neu5Ac was used as substrate, a strong increase of Neu5Ac was detected after endoN digestion indicating that *Cma*ST8Sia IV efficiently transfered Neu5Ac to CD166/ALCAM (Fig. 7C). When CMP-Neu5Gc was used, an increase of Neu5Gc was also observed although lower than that obtained with CMP-Neu5Ac, which is in line with the WB and the cell-based data obtained. However, the transfer of several Kdn residues could not be proven since the Kdn values of the endoN and the untreated sample were comparable. Since slightly higher amounts, but not statistically significant amounts of Neu5Ac were detectable in endoN treated samples of the CMP-Neu5Gc and CMP-Kdn setups (Fig. 7C), it is proposed that pre-existing oligo-Neu5Ac are degraded by endoN, which were not detectable by mAb735. The elongation of already existing Neu5Ac-oligomers using CMP-Kdn let suggest that the transfer of a few Kdn residues was likely not sufficient to be detected by HPLC, but enabled binding of mAb735, a hypothesis that is supported by the WB data (Fig. 5). Altogether, these results demonstrated that CMP-Neu5Ac is the preferred substrate donor of *Cma*ST8Sia IV; however, in constrast to the human enzyme, this fish ST8Sia IV enzyme does also accommodate CMP-Neu5Gc and to a lesser extent CMP-Kdn.

In summary, this study highlights original enzymatic activities and specificities of the salmonid ST8Sia IV likely responsible for the synthesis of diverse polySias. Fish polysialyltransferases show functional divergence acquired during vertebrate evolution with specific characteristics and kinetic properties distinct from those of human enzymes. Since no crystal 3D-structure of vertebrate polysialyltransferase has been obtained yet, molecular phylogeny, modeling and docking strategies will be developed to shed light on the molecular basis of these differential substrate recognition by vertebrate polysialyltransferases and better understand their biological function. In addition, our data brought new evidences of the usefulness of the molecular tools mAb735 and endoN towards a variety of natural polySia polymers.

Materials and methods

Materials. Fetuin, orosomucoid, bovine submaxillary mucin (BSM), 3,3',5,5'-tetramethylbenzidine (TMB), azide-PEG3-biotin, monoclonal Anti-FLAG® BioM2 antibody (1 mg/mL), sialic acid aldolase from *Escherichia coli* K-12, CSS from *Neisseria meninigitidis*, pyrophosphatase from *Saccharomyces cerevisiae* (PPase), Neu5Ac, Neu5Gc and Kdn were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). Sodium cytidine-5′-triphosphate (CTP) was obtained from TCI chemicals. Human Neuropilin-2 (NRP2) was purchased from Interchim. Horseradish peroxidase (HRP) conjugated anti-biotin antibody (monoclonal fraction of mouse IgG; 0.8 mg/mL) was purchased from Jackson Immunoresearch. 2-[4-((bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino)methyl)-1H-1,2,3-triazol-1-yl]acetic acid (BTTAA) was synthesized in our laboratory as described previously[33]. The human recombinant enzymes Δ23ST8Sia II and Δ39ST8Sia IV were from R&D Systems (Rennes, France). Human Recombinant Deoxyribonuclease I (DNase I) was purchased from Boster Bio (Pleasanton, CA, USA). PageRuler Plus Prestained Protein Ladder was from Thermo Scientific. Anti-polySia

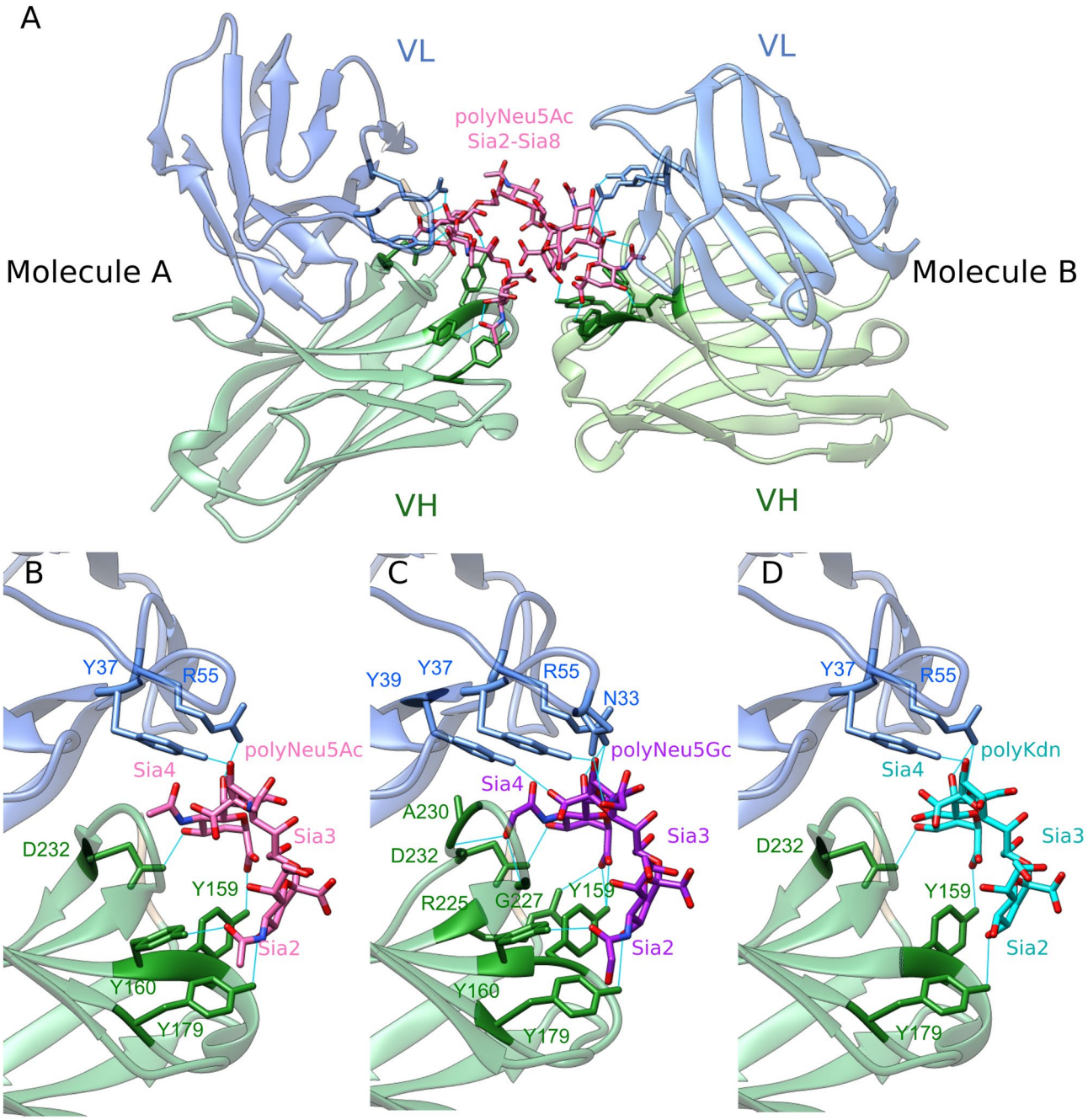


Figure 6. Crystal structure of scFv735 (PDB 3WBD) in complex with α2,8-linked octasialic acids. (**A**) Overall crystal structure of scFv735 in complex with polyNeu5Ac, variable light domain (V_L) and variable heavy domain (V_H) are coloured in blue and green, respectively. The Sia residues interacting with two scFv735 molecules are highlighted in pink. Hydrogen bonds between Sia and the antibody are shown in cyan. Residues involved in direct hydrogen bonds are shown in stick representation. Close-up view of the antigen recognition site of scFv735 in complex with trisialic acid of (**B**) Neu5Ac, (**C**) Neu5Gc and (**D**) Kdn.

mAb735 (Rabbit IgGκ, 1 mg/mL) was obtained from Enzo Life Sciences. Sodium cacodylate buffer was from Prolabo (Paris, France).

Whitefish *C. maraena* were provided by the Institute of Fisheries of the Mecklenburg Western Pomerania Research Center for Agriculture and Fisheries (Born, Germany). The present study exclusively utilised samples from previous experiments[41,42] following the 3R principles in science. These experiments had been approved by the Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei, Mecklenburg-Vorpommern, according to the German animal welfare law (approval ID: LALLF M-V/TSD/7221.3-1-069/18; November 2018) and all the methods were carried out in accordance with relevant Institutional guidelines and regulations and with the ARRIVE guidelines. Brain and kidney tissues from whitefish were snap frozen in liquid nitrogen and kept at − 80 °C for RNA isolation.

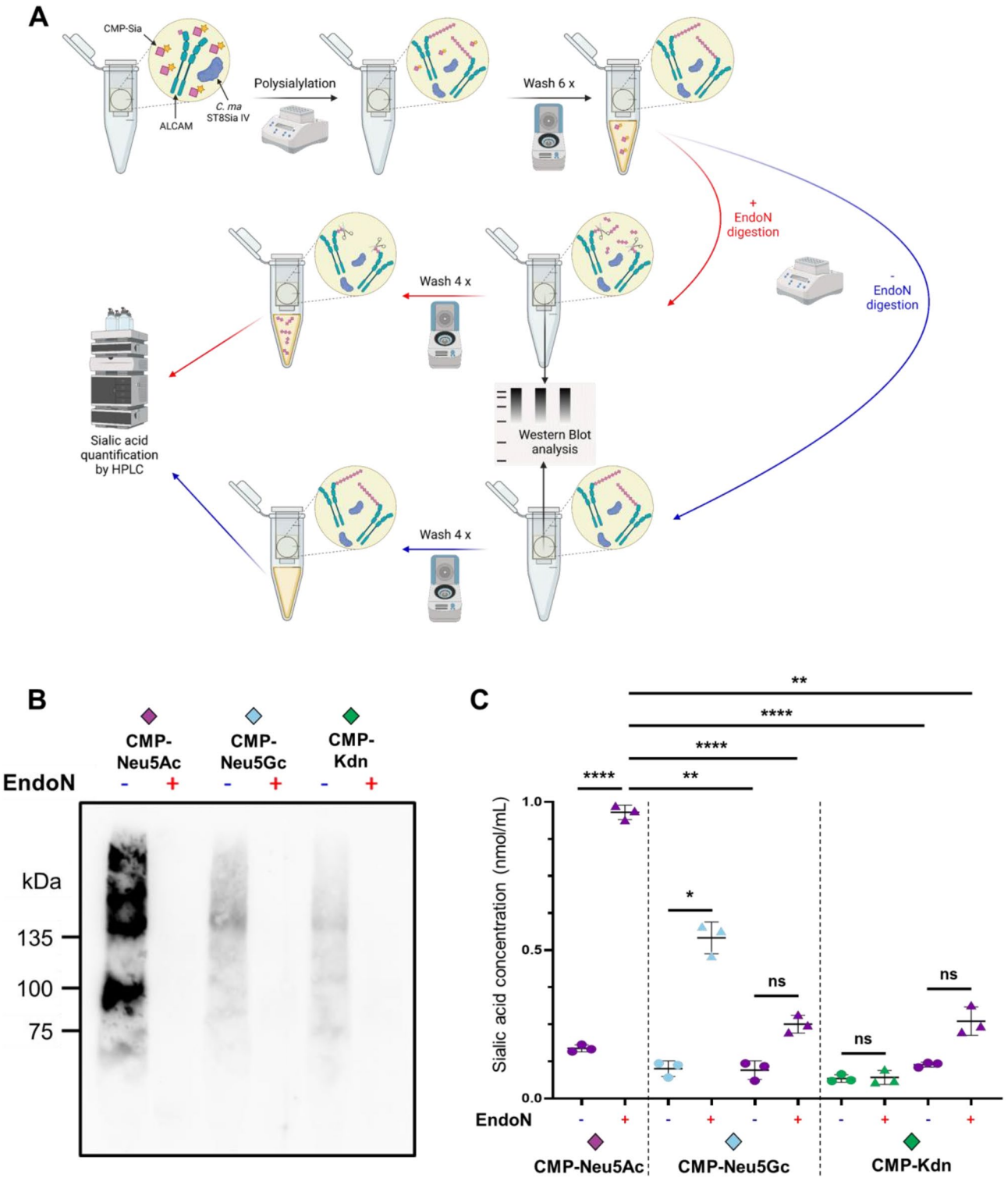


Figure 7. Composition analysis of polySia on CD166/ALCAM. (**A**) Workflow of the centrifugal filter polysialylation assay to quantitatively analyze the transfer of different Sia residues to CD166/ALCAM by *Cma*ST8Sia IV. Created with BioRender.com. (**B**) The polysialylation reaction using different CMP-Sia (CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn) was controled by WB using the mAb735. Original blot is presented in supplementary Fig. S9. (**C**) In the final flow-throughs of untreated (circle) and endoN treated (triangle) samples, sialic acids were analyzed by HPLC for each experimental set-up: donor substrate CMP-Neu5Ac, quantification of Neu5Ac (purple); donor substrate CMP-Neu5Gc: quantification of Neu5Gc (blue) and Neu5Ac (purple); donor substrate CMP-Kdn, quantification of Kdn (green) and Neu5Ac (purple). The means and standard deviations of three independent experiments are displayed. Statistical differences are indicated: *$p < 0.05$; **$p < 0.01$; ****$p < 0.0001$; *ns* not significant.

Bioinformatics analyses of polysialyltransferases sequences and molecular docking. Nucleotide sequences of *st8Sia2r-1/st8sia2a* (NCBI nucleotide accession code: XM_041891035), *st8sia2r-2/st8sia2b* (XM_041836579) and *st8sia4* (XM_041897387) genes from closely related *Coregonus clupeaformis* were synchronised with the transcriptome of *Coregonus maraena* (Bioproject ID: 302,355) previously[19] and compared with the orthologous sequences in human ST8SIA2 (NM_006011) and ST8SIA4 (NM_005668).

Multiple sequence alignments of the corresponding protein sequences were achieved using ClustalW at PRABI-Gerland (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_clustalw.html). The amino acid sequence analysis was performed using the software of Expert Protein Analysis System (ExPASy; Swiss Institute of Bioinformatics, Switzerland; website (https://www.expasy.org/, last accessed on: 04 November 2022). Hydropathy analyses and determination of potential N-glycosylation sites were performed using the servers TMHMM—2.0: Prediction of Transmembrane Regions (https://services.healthtech.dtu.dk/service.php?TMHMM-2.0, last accessed on: 04 November 2022) and the NetNGlyc 1.0 Internet program (https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0, last accessed on: 04 November 2022) of ExPASy. Sequences were also submitted to Compute pI/MW (ExPASy) analysis to determine their theoretical molecular weight.

To investigate to what extent analogous interactions could be found between the antibody and polySia chains made of eight Neu5Gc or Kdn, we took as reference direct and indirect interactions described for mAb735 complexed with polyNeu5Ac in ProteinDataBank (PDB) 3WBD and modified the Neu5Ac to Neu5Gc/Kdn molecules in the PDB 3WBD using the biocomputing software UCSF Chimera. The complex was minimized using the AMBER force field, and charges were computed using ANTECHAMBER[43].

Construction of Cma polysialyltransferases expression vectors. Total RNA was extracted from homogenized brain or kidney tissue using the Nucleospin RNA Plus (Macherey–Nagel, Düren, Germany) and quantified with the NanoDrop ®ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Total RNA was reverse-transcribed using the Maxima First strand cDNA synthesis for RT-qPCR kit according to the manufacturer protocol (Thermo Scientific). To obtain a full-length cDNA, an initial RT-PCR was performed with each kidney and brain cDNA preparation, Q5 high fidelity DNA polymerase (New England Biolabs), sense and antisense oligonucleotide primers containing *Hind*III and *Kpn*I restriction sites respectively (listed in Table S-2). PCR reactions were run for 2 min at 98 °C followed by 35 cycles (98 °C for 30 s, 65 °C for 30 s, 72 °C for 15 s) and an extension step of 2 min at 72 °C. The resulting 1074 bp amplified *st8sia4* cDNA was subcloned into the pJET vector of Clone JET PCR cloning 1.2 kit (Thermo Scientific), cut out with *Hind*III and *Kpn*I and inserted into the *Hind*III and *Kpn*I sites of the plasmid p3 × FLAG-CMV10 expression vector. The resulting 1075 bp amplified *st8sia2r*-1 cDNA was inserted directly into p3 × FLAG-CMV10. For s*t8sia2*r-2, a synthetic vector based on its full-length sequence was constructed using the pCDNA3.1 plasmid at GeneArt Instant Designer (ID project: 2020AAEFFD, Thermo Fisher Scientific). A cDNA encoding a truncated form of the *Cmast8sia2r*-1, *Cmast8sia2r*-2 and *Cmast8sia4* lacking their first 35, 35 or 28 aa residues of the open reading frame, respectively were amplified by PCR using the previous construct p3 × FLAG-CMV10 for *st8sia4* and *st8sia2*r-1 and pcDNA3.1 for *st8sia2*r-2 as templates. Sense and antisense primers are described in Supplementary Table S-2. Touchdown reactions were run for 2 min at 94 °C followed by 10 cycles (94 °C for 30 s, 60 °C (minus 1 °C per cycle) for 30 s, 72 °C for 70 s) and then 20 cycles (94 °C for 30 s, 55 °C for 30 s, 72 °C for 70 s) and an extension step of 5 min at 72 °C. The resulting PCR amplified cDNAs were inserted into the *Hind*III and *Kpn*I sites of the p3 × FLAG-CMV9 expression vector. These last constructs encoded a FLAG-tagged protein with a signal peptide sequence, the FLAG octapeptide and the three Δ35ST8Sia II-r1, Δ35ST8Sia II-r2 and Δ28ST8Sia IV sequences deleted of their cytoplasmic tail and transmembrane domain.

Cell culture, Transfection and Exosialylation of HEK293 surface using the human and fish ST8Sia IV and confocal microscopy. HEK293 cells (ATCC CRL-1573) and CHO-K1 (ATCC CCL-61) cells were grown in Dulbecco's modified eagle's medium (DMEM, Dutscher) supplemented with fetal calf serum (10% FCS, Biowest) at 37 °C, 5% CO_2. Confluent cells (~ 70%) were transiently transfected with either 2 or 10 µg of purified expression constructs in 6 well plates or in 100 mm petri dishes with Lipofectamine 2000 (Thermo Scientific) in UltraMEM medium (12-743F, Lonza) according to the manufacturer's instructions. An empty p3 × FLAG-CMV9 plasmid was used as a control (Mock). Transiently transfected cells and cell culture media were collected 72 h after transfection. Recombinant polysialyltransferases produced in DMEM-FCS culture media were used as a crude enzyme source for enzymatic assays. The p3 × FLAG-CD166/ALCAM plasmid was transiently transfected in CHO-K1 cells and the recombinant protein was secreted in the UltraMEM medium. This medium was collected 48 h after transfection and centrifuged 30 min at 10,000 *rpm* and at 4 °C to remove cell debris. The CD166/ALCAM content of the UltraMEM medium of the p3 × FLAG-CD166/ALCAM or mock transfected CHO-K1 cells was assessed with the micro-BCA protein assay kit (23,235, Thermo Scientific) and WB with the anti-FLAG (see below).

For the exosialylation experiments, HEK293 cells were grown on glass coverslips for 24 h, then washed three times with PBS and fixed in 4% PFA for 30 min at room temperature (RT). After three washes with PBS, cells were treated with endoN at 100 ng/mL for 1 h at 37 °C. Excess of endoN was eliminated by five PBS washes 5 min under gently agitation. Sialylation reactions were performed at 27 °C for 4 h with the chosen enzymatic source of enzyme and 500 µM of CMP-Sia donor substrate in 100 µL final volume. The rest of the experiment consists of immunofluorescence to visualize de novo sialylation. Briefly, after three washes with PBS, the cells were incubated with blocking buffer (2% BSA in PBS) for 1 h at RT and then overnight at 4 °C in humid chamber with the primary antibodies diluted in blocking buffer: 2 µg/mL anti-polySia mAb735 (Enzo Lifes Sciences) and 5 µg/mL anti-CD56 (NCAM, clone 123C3) (Invitrogen, ThermoFisher Scientific). After three washes with PBS, cells were incubated 1 h at RT in the dark with AlexaFluorTM conjugated secondary antibodies: respectively

goat anti-rabbit 488 and goat anti-mouse 568 (Invitrogen, ThermoFisher Scientific) diluted at 1: 600 (3.33 µg/mL) in blocking buffer. Finally, cells were washed three time with PBS, the nuclei were stained with DAPI (5 µg/mL in PBS) and the coverslips were mounted on glass slides with Mowiol®. The fluorescence was detected using the inverted Zeiss LSM780 confocal microscope with a 40 × oil immersion and data were collected with the ZEN 2010 software (Zeiss, Oberkochen, Germany). The images were analyzed with ImageJ software.

Western blot analyses. WB was used to visualize and quantify tagged recombinant proteins secreted from transiently transfected HEK293 or CHO-K1 cells. For that purpose, 20 µL of culture media of transfected cells were boiled 5 min at 95 °C in 4 × Laemmli buffer (235 mM Tris–HCl pH 6.8, 8% SDS, 40% glycerol, 10% β-mercaptoethanol, 0.01% bromophenol blue), and resolved by 8% SDS-PAGE. Proteins were transferred onto Amersham Protran nitrocellulose membrane (GE Healthcare Life Sciences) for 75 min at 200 mA and checked with Ponceau red staining (5% acetic acid, 0.1% Ponceau). Membranes were washed, then blocked using 5% non-fat milk in TBS-T (TBS with 0.05% Tween 20). Detection of recombinant proteins was achieved with the primary mouse anti-FLAG® M2 antibody (1 µg/ml; Sigma-Aldrich) in TBS-T 0.05% overnight at 4 °C. After 3 washes with TBS-T 0.05%, the membrane was incubated with an anti-mouse antibody coupled to horseradish peroxidase (HRP) (0.1 µg/mL; InVitrogen) for 1 h then was washed 5 times with TBS-T 0.05%. Blots were developed using enhanced chemiluminescence (ECL West Pico Plus or ECL West Femto, Thermo Scientific). The images were acquired using a CCD camera (Fusion Solo, Vilber Lourmat) and the Fusion software. Quantification and densitometry analysis were done with ImageJ and GraphPad Prism 6.

WB strategy was also used after polysialylation reaction achieved on CD166/ALCAM with the natural substrate donors (see below). Polysialylated CD166/ALCAM products treated or not with endoN were heated 5 min at 60 °C in 4 × Laemmli buffer, separated on an 6% SDS-PAGE gel for 90 min at 95 V (1 × Tris–Glycine buffer (Euromedex)/ 20% Methanol) then transferred onto nitrocellulose membrane. Membranes were saturated in 5% non-fat milk in TBS-T 0.05% for 1 h at room temperature (RT) and incubated with the anti-polySia mAb735 antibody (1 µg/mL) or endoN (6.7 µg/mL) overnight at 4 °C. After 3 washes with TBS-T 0.05%, the membrane was incubated with secondary anti-rabbit antibody coupled to HRP (0.1 µg/mL; InVitrogen) for 1 h and was washed 5 times with TBS-T 0.05%. Detection was achieved by chemiluminescence as reported above.

Chemo-enzymatic synthesis of activated CMP-sialic acids and sialylation assays. *N*-4-pentynoylneuraminic acid (SiaNAl) and the various natural CMP-Sia were synthetized as previously described[39]. In brief, *N*-mannosamine was converted to *N*-4-pentynoylmannosamine by coupling with succinimidyl 4-pentynoate, which was then reacted with sodium pyruvate in the presence of sialic acid aldolase from *E. coli* K12 in pH 7.5 phosphate buffer and purified by anion exchange chromatography (Dowex 1 × 8) and gel filtration chromatography (P2 resin) to yield pure SiaNAl (86% overall yield over 2 steps).

The syntheses of cytidine-5′-monophospho-N-sialic acids (CMP-Sia) CMP-Neu5Ac, CMP-Neu5Gc, CMP-Kdn and CMP-SiaNAl were carried out in equimolarity of CTP and each Sia (1:1) with 0.3 U CSS from *N. meningitidis*[39] or the rainbow trout CMP-Kdn synthetase (rtCSS)[40] and 0.5 U PPase in 100 mM Tris–HCl, 20 mM $MgCl_2$ buffer (pH 8.5) at 37 °C. The reactions were carried out in a 5 mm NMR tube and monitored by ^{31}P NMR spectroscopy in a Brüker Avance II 400 MHz spectrometer (Supplementary Fig. S2). Upon completion, the pure formed CMP-Sia was immediately stored at − 80 °C until use[33,39].

Enzymatic assays were performed using the MicroPlate Sialyltransferase Assay (MPSA) as described previously[33]. In brief, 400 ng of glycoprotein acceptors in 100 µL of sodium bicarbonate buffer (20 mM, pH 9.6) were adsorbed into the bottom of 96-well plates (F8 MaxiSorp Loose Nunc-Immuno Module ThermoScientific) at 4 °C overnight. After three washes with 150 µL of phosphate buffered saline containing 0.05% Tween 20 (PBS-T 0.05%), saturation was carried out for 1 h at RT with 100 µL of oxidized BSA at 0.05% diluted in bicarbonate buffer. The sialylation transfer reaction was carried out for one to several hours at 27 °C with the chosen enzymatic source and 100 µM of CMP-SiaNAl in 100 mM cacodylate buffer ($MnCl_2$ 10 mM, Triton CF-54 0.2%, pH 6.2) in a final volume of 100 µL. After sialylation, the wells were washed with PBS-T 0.05%, then the CuAAC labeling reaction was performed by adding 100 µL of a solution containing 300 µM $CuSO_4$, 600 µM BTTAA, 2.5 mM sodium ascorbate and 250 µM azide-PEG3-biotin in PBS[44]. After 1 h at 37 °C, the reaction was stopped by washing three times with PBS-T 0.05%, then 100 µL of an HRP-conjugated anti-biotin antibody (32 ng/mL) were added for 1 h at 37 °C. After washing, 100 µL of TMB were added and incubated for 20 min at RT in the dark. Finally, the absorbance was measured at 620 nm with a spectrophotometer (SpectroStar Nano, BMG Labtech). The data were analyzed with GraphPad using the statistical ANOVA test between samples.

Polysialylation assays were also performed at 27 °C for 4 h in 100 mM cacodylate buffer with 100 µM CMP-Sia (*i.e.* CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn), 4 µg of the CD166/ALCAM and 230 µL of enzymatic source in a total volume of 1 mL. The enzymatic source with empty vector p3 × FLAG-CMV9 (Mock) and recombinant human ST8Sia IV were used as controls. Samples were cooled on ice, dialyzed overnight on 10 kDa dialysis membrane in ammonium bicarbonate buffer 50 mM and lyophilized. Samples were resuspended in 18 µL of RIPA buffer (Tris HCl 10 mM; NaCl 150 mM; Triton X-100 1%; pH 6.4) and treated or not with endoN (6.7 µg/mL) 1 h, 37 °C. CD166/ALCAM polysialylation was visualized by WB. Total samples were boiled for 5 min at 60 °C and resolved on a 6% SDS-PAGE as described previously.

Chain length analysis of polySia by HPLC. Fish ST8Sia IV was used to polysialylate CD166/ALCAM as described in chapter Chemo-enzymatic synthesis of activated CMP-sialic acids and sialylation assays. The polysialylated fraction was isolated using affinity precipitation. Therefore, inactive endoN was coupled to tosyl-activated magnetic Dynabeads M-280 (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions as described previously[17,45]. Inactive endoN binds to polySia, but is not able to degrade the polymer. The samples

were dialyzed against TBS for 2 h using Spectra/Por® Biotech CE Tubing (MWCO: 50 kDa; Repligen, Rancho Dominguez, CA). Subsequently, the dialyzed samples were incubated with the beads for 30 min. After washing, polySia was eluted using 100 mM triethylamine, 150 mM NaCl and dried in a vacuum concentrator. To control the affinity precipitation, 10% of the eluate was analyzed by WB against polySia. The remaining 90% of the samples were used for HPLC analysis.

The polySia chain length determination of polysialylated CD166/ALCAM was performed using the 1,2-diamino-4,5-methylene-dioxybenzene (DMB)-HPLC strategy[46,47]. Briefly, the isolated polySia-CD166/ALCAM samples were dissolved in 80 µL DMB reagent (9 mM sodium hydrosulfite, 0.5 M mercaptoethanol, 20 mM trifluoroacetic acid (TFA), 0.61 mg/ml DMB) and incubated at 11 °C overnight. The reaction was stopped with 20 µL of 1 M NaOH for 1 h. The fluorescently labeled polySia chains were separated by anion exchange chromatography using a DNAPac PA-100 column (4 × 250 mm; 13 µm; Dionex). The eluents, MilliQ water (E1) and 2 M ammonium acetate (E2) were used with a flow rate of 1 mL/min following the gradient: 0 min = 0% E2, 5 min = 0% E2, 15 min = 13% E2, 30 min = 21% E2, 55 min = 33% E2, 100 min = 43% E2, 101 min = 100% E2, 110 min = 100% E2, 111 min = 0% E2, 145 min = 0% E2. Fluorescent signals were detected using an extinction wavelength of 372 nm and an emission wavelength of 456 nm.

Quantitative analysis of polySia composition by HPLC. To quantitatively analyze the transfer of different CMP-Sia residues onto CD166/ALCAM by *Cma*ST8Sia IV, we established an endoN approach using centrifugal filter units. The analytical strategy is based on Ref.[48]. Microcon® centrifugal filter units (MRCPRT010, Merck Millipore Ltd.; NMWL: 10 kDa) were prepared according to the manufacturer's instructions. Polysialylation reaction was performed overnight at 37 °C in cacodylate buffer containing 100 µl fish ST8Sia IV culture medium, 150 ng CD166/ALCAM and 100 µM CMP-Sia (CMP-Neu5Ac or CMP-Neu5Gc or CMP-Kdn) in a total volume of 250 µl. To remove remaining free CMP-Sia, the samples were centrifuged with 14,000 × *g* at 4 °C until almost all the liquid was filtered. The samples were subsequently washed 5 more times with 250 µl 50 mM NH_4HCO_3 followed by centrifugal-filtration step. Thereafter, the volume was adjusted to 110 µl with 50 mM NH_4HCO_3 buffer. EndoN was added (1.34 µg/ml) to degrade polySia. In parallel, untreated samples (without endoN) of each CMP-Sia preparation were used as negative controls. Aliquots (10 µl) of all samples were collected for WB after 1 h at 37 °C. The remaining samples were centrifuged as described above and washed three more times. For quantitative Sia analysis, the complete flow-through (containing degraded sialic acid chains) was dried in a vacuum concentrator and dissolved in 0.2 N TFA. Hydrolysis was performed at 80 °C for 4 h. The hydrolyzed samples were dried and dissolved in 80 µl DMB reagent. After 2 h at 55 °C, the reaction was stopped with 20 µl 0.2 N NaOH. Resulting fluorescently labeled sialic acid residues were injected into a HPLC system (Nexera, Shimadzu) and separated using a Superspher® 100 RP-18 end-capped column (250 mm × 40 mm, Merck-Hitachi, Darmstadt, Germany). A gradient with the eluents 92% MilliQ water, 4% ACN, 4% methanol, 0,1% TFA (E1) and 10% MilliQ water, 45% acetonitrile (ACN), 45% methanol, 0,1% TFA (E2) was applied with a flowrate of 0.25 ml/min as follows: 0 min = 0% E2, 2 min = 0% E2, 25 min = 2% E2, 35 min = 5% E2, 40 min = 50% E2, 45 min = 100% E2, 50 min = 100% E2, 51 min = 0% E2, 60 min = 0% E2. Fluorescent signals were detected using an extinction wavelength of 372 nm and an emission wavelength of 456 nm. Sialic acid standards were used to obtain a calibration line for quantification.

Statistical analyses. The statistical analyses were performed with GraphPad Prism software (version 9.5.1) using One-way ANOVA and a multiple comparison Tukey test. Following labels are used: $p < 0.05$ (*); $p < 0.01$ (**); $p < 0.001$ (***); $p < 0.0001$ (***); ns, not significant ($p \geq 0.05$).

Ethics statements. These experiments had been approved by the Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei, Mecklenburg-Vorpommern, according to the German animal welfare law (approval ID: LALLF M-V/TSD/7221.3-1-069/18; November 2018) and all the methods were carried out in accordance with relevant Institutional guidelines and regulations and with the ARRIVE guidelines.

Data availability

Data and materials can be shared upon request to corresponding authors. The *C. maraena* transcriptome can be accessed at the NCBI Sequence Read Archive (Bioproject ID: PRJNA302355).

Received: 29 June 2023; Accepted: 5 September 2023
Published online: 20 September 2023

References

1. Angata, T. & Varki, A. Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chem. Rev.* **102**, 439–469 (2002).
2. Sato, C. & Kitajima, K. Polysialylation and disease. *Mol. Aspects Med.* **79**, 100892. https://doi.org/10.1016/j.mam.2020.100892 (2021).
3. Schnaar, R. L., Gerardy-Schahn, R. & Hildebrandt, H. Sialic acids in the brain: Gangliosides and polysialic acid in nervous system development, stability, disease, and regeneration. *Physiol. Rev.* **94**, 461–518. https://doi.org/10.1152/physrev.00033.2013 (2014).
4. Villanueva-Cabello, T. M., Gutiérrez-Valenzuela, L. D., Salinas-Marín, R., López-Guerrero, D. V. & Martínez-Duncker, I. Polysialic acid in the immune system. *Front. Immunol.* https://doi.org/10.3389/fimmu.2021.823637 (2022).
5. Galuska, C. E., Lutteke, T. & Galuska, S. P. Is polysialylated NCAM not only a regulator during brain development but also during the formation of other organs?. *Biology (Basel)* **6**, 27. https://doi.org/10.3390/biology6020027 (2017).
6. Hachem, N. E. *et al.* The loss of polysialic acid impairs the contractile phenotype of peritubular smooth muscle cells in the postnatal testis. *Cells* **10**, 134. https://doi.org/10.3390/cells10061347 (2021).

7. Simon, P. *et al.* Polysialylation of NCAM characterizes the proliferation period of contractile elements during postnatal development of the epididymis. *PLoS One* **10**, e0123960. https://doi.org/10.1371/journal.pone.0123960 (2015).
8. Simon, P. *et al.* Polysialic acid is present in mammalian semen as a post-translational modification of the neural cell adhesion molecule NCAM and the polysialyltransferase ST8SiaII. *J. Biol. Chem.* **288**, 18825–18833. https://doi.org/10.1074/jbc.M113.451112 (2013).
9. Iwasaki, M. & Inoue, S. Structures of the carbohydrate units of polysialoglycoproteins isolated from the eggs of four species of salmonid fishes. *Glycoconj. J.* **2**, 209–228 (1985).
10. Kitajima, K., Inoue, Y. & Inoue, S. Polysialoglycoproteins of Salmonidae fish eggs. Complete structure of 200-kDa polysialoglycoprotein from the unfertilized eggs of rainbow trout (Salmo gairdneri). *J. Biol. Chem.* **261**, 5262–5269 (1986).
11. Sato, C. *et al.* Structural diversity in the alpha 2–>8-linked polysialic acid chains in salmonid fish egg glycoproteins. Occurrence of poly(Neu5Ac), poly(Neu5Gc), poly(Neu5Ac, Neu5Gc), poly(KDN), and their partially acetylated forms. *J. Biol. Chem.* **268**, 23675–23684 (1993).
12. Nadano, D. *et al.* A naturally occurring deaminated neuraminic acid, 3-deoxy-D-glycero-D-galacto-nonulosonic acid (KDN). Its unique occurrence at the nonreducing ends of oligosialyl chains in polysialoglycoprotein of rainbow trout eggs. *J. Biol. Chem.* **261**, 11550–11557 (1986).
13. Kitajima, K., Inoue, S., Inoue, Y. & Troy, F. Use of a bacteriophage-derived endo-N-acetylneuraminidase and an equine antipolysialyl antibody to characterize the polysialyl residues in salmonid fish egg polysialoglycoproteins. Substrate and immunospecificity studies. *J. Biol. Chem.* **263**, 18269–18276 (1988).
14. Fliniaux, I. *et al.* Diversity of sialic acids and sialoglycoproteins in gametes and at fertilization. *Front. Cell Dev. Biol.* **10**, 982931. https://doi.org/10.3389/fcell.2022.982931 (2022).
15. Taguchi, T. *et al.* Structural studies of a novel type of pentaantennary large glycan unit in the fertilization-associated carbohydrate-rich glycopeptide isolated from the fertilized eggs of Oryzias latipes. *J. Biol. Chem.* **269**, 8762–8771 (1994).
16. Angata, T. *et al.* Identification, characterization, and developmental expression of a novel alpha 2–>8-KDN-transferase which terminates elongation of alpha 2–>8-linked oligo-polysialic acid chain synthesis in trout egg polysialoglycoproteins. *Glycoconj. J.* **11**, 493–499 (1994).
17. Zlatina, K. *et al.* Polysialic acid in human plasma can compensate the cytotoxicity of histones. *Int. J. Mol. Sci.* **19**, 1679. https://doi.org/10.3390/ijms19061679 (2018).
18. Venuto, M. T. *et al.* Characterization of the polysialylation status in ovaries of the salmonid fish coregonus maraena and the percid fish sander lucioperca. *Cells* **9**, 2391. https://doi.org/10.3390/cells9112391 (2020).
19. Venuto, M. T. *et al.* Vertebrate alpha 2,8-sialyltransferases (ST8Sia): A teleost perspective. *Int. J. Mol. Sci.* **21**, 513. https://doi.org/10.3390/ijms21020513 (2020).
20. Harduin-Lepers, A. *et al.* The human sialyltransferase family. *Biochimie* **83**, 727–737. https://doi.org/10.1016/s0300-9084(01)01301-3 (2001).
21. Harduin-Lepers, A. The vertebrate sialylation machinery: structure-function and molecular evolution of GT-29 sialyltransferases. *Glycoconj. J.* https://doi.org/10.1007/s10719-023-10123-w (2023).
22. Harduin-Lepers, A. Comprehensive analysis of sialyltransferases in vertebrate genomes. *Glycobiol. Insights* **2**, 29–61. https://doi.org/10.4137/GBI.S3123 (2010).
23. Harduin-Lepers, A., Mollicone, R., Delannoy, P. & Oriol, R. The animal sialyltransferases and sialyltransferase-related genes: A phylogenetic approach. *Glycobiology* **15**, 805–817. https://doi.org/10.1093/glycob/cwi063 (2005).
24. Asahina, S. *et al.* Involvement of the a2,8-polysialyltransferases II/STX and IV/PST in the biosynthesis of polysialic acid chains on the O-linked glycoproteins in rainbow trout ovary. *J. Biochem. (Tokyo)* **140**, 687–701. https://doi.org/10.1093/jb/mvj200 (2006).
25. Chang, L. Y. *et al.* Novel zebrafish Mono-alpha2,8-sialyltransferase (ST8Sia VIII): An evolutionary perspective of alpha2,8-sialylation. *Int. J. Mol. Sci.* **20**, 622. https://doi.org/10.3390/ijms20030622 (2019).
26. Harduin-Lepers, A. *et al.* Evolutionary history of the alpha2,8-sialyltransferase (ST8Sia) gene family: Tandem duplications in early deuterostomes explain most of the diversity found in the vertebrate ST8Sia genes. *BMC Evol. Biol.* **8**, 258. https://doi.org/10.1186/1471-2148-8-258 (2008).
27. Datta, A. K. & Paulson, J. C. Sialylmotifs of sialyltransferases. *Indian J. Biochem. Biophys.* **34**, 157–165 (1997).
28. Foley, D. A., Swartzentruber, K. G. & Colley, K. J. Identification of sequences in the polysialyltransferases ST8Sia II and ST8Sia IV that are required for the protein-specific polysialylation of the neural cell adhesion molecule. *J. Biol. Chem.* **284**, 15505–15516. https://doi.org/10.1074/jbc.M809696200 (2009).
29. Zapater, J. L. & Colley, K. J. Sequences prior to conserved catalytic motifs of polysialyltransferase ST8Sia IV are required for substrate recognition. *J. Biol. Chem.* **287**, 6441–6453. https://doi.org/10.1074/jbc.M111.322024 (2012).
30. Liao, S. M. *et al.* Molecular interactions of the polysialytransferase domain (PSTD) in ST8Sia IV with CMP-sialic acid and polysialic acid required for polysialylation of the neural cell adhesion molecule proteins: An NMR study. *Int. J. Mol. Sci.* **21**, 1590. https://doi.org/10.3390/ijms21051590 (2020).
31. Nakata, D., Zhang, L. & Troy, F. A. Molecular basis for polysialylation: a novel polybasic polysialyltransferase domain (PSTD) of 32 amino acids unique to the a2,8-polysialyltransferases is essential for polysialylation. *Glycoconj. J.* **23**, 423–436 (2006).
32. Moremen, K. W. *et al.* Expression system for structural and functional studies of human glycosylation enzymes. *Nat. Chem. Biol.* **14**, 156–162. https://doi.org/10.1038/nchembio.2539 (2018).
33. Noel, M. *et al.* MicroPlate sialyltransferase assay: A rapid and sensitive assay based on an unnatural sialic acid donor and bioorthogonal chemistry. *Bioconjug. Chem.* **29**, 3377–3384. https://doi.org/10.1021/acs.bioconjchem.8b00529 (2018).
34. Vogt, J. *et al.* Homeostatic regulation of NCAM polysialylation is critical for correct synaptic targeting. *Cell. Mol. Life Sci.* **69**, 1179–1191. https://doi.org/10.1007/s00018-011-0868-2 (2012).
35. Siddiqui, K. S. & Cavicchioli, R. Cold-adapted enzymes. *Annu. Rev. Biochem.* **75**, 403–433. https://doi.org/10.1146/annurev.biochem.75.103004.142723 (2006).
36. Guo, X. *et al.* An efficient assay for identification and quantitative evaluation of potential polysialyltransferase inhibitors. *Analyst* **145**, 4512–4521. https://doi.org/10.1039/D0AN00721H (2020).
37. Nagae, M. *et al.* Crystal structure of anti-polysialic acid antibody single chain Fv fragment complexed with octasialic acid: Insight into the binding preference for polysialic acid. *J. Biol. Chem.* **288**, 33784–33796. https://doi.org/10.1074/jbc.M113.496224 (2013).
38. Stummeyer, K., Dickmanns, A., Mühlenhoff, M., Gerardy-Schahn, R. & Ficner, R. Crystal structure of the polysialic acid–degrading endosialidase of bacteriophage K1F. *Nat. Struct. Mol. Biol.* **12**, 90–96. https://doi.org/10.1038/nsmb874 (2005).
39. Gilormini, P. A. *et al.* Improved workflow for the efficient preparation of ready to use CMP-activated sialic acids. *Glycobiology* **26**, 1151–1156. https://doi.org/10.1093/glycob/cww084 (2016).
40. Wu, D. *et al.* A novel C-domain-dependent inhibition of the rainbow trout CMP-sialic acid synthetase activity by CMP-deaminoneuraminic acid. *Biochem. Biophys. Res. Commun.* **617**, 16–21. https://doi.org/10.1016/j.bbrc.2022.05.031 (2022).
41. Martorell-Ribera, J. *et al.* Experimental handling challenges result in minor changes in the phagocytic capacity and transcriptome of head-kidney cells of the salmonid fish coregonus maraena. *Front. Vet. Sci.* **9**, 889635. https://doi.org/10.3389/fvets.2022.889635 (2022).
42. Martorell-Ribera, J. *et al.* Time-dependent effects of acute handling on the brain monoamine system of the salmonid coregonus maraena. *Front. Neurosci.* **14**, 591738. https://doi.org/10.3389/fnins.2020.591738 (2020).

43. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25**, 247–260. https://doi.org/10.1016/j.jmgm.2005.12.005 (2006).
44. Noel, M. *et al.* Probing the CMP-sialic acid donor specificity of two human beta-d-galactoside sialyltransferases (ST3Gal I and ST6Gal I) selectively acting on O- and N-glycosylproteins. *ChemBioChem* **18**, 1251–1259. https://doi.org/10.1002/cbic.201700024 (2017).
45. Kuhnle, A. *et al.* Polysialic acid interacts with lactoferrin and supports its activity to inhibit the release of neutrophil extracellular traps. *Carbohydr. Polym.* **208**, 32–41. https://doi.org/10.1016/j.carbpol.2018.12.033 (2019).
46. Hinterseher, J. *et al.* Milk polysialic acid levels rapidly decrease in line with the N-acetylneuraminic acid concentrations during early lactation in dairy cows. *Biology (Basel)* **12**, 5. https://doi.org/10.3390/biology12010005 (2022).
47. Inoue, S. & Inoue, Y. Ultrasensitive analysis of sialic acids and oligo/polysialic acids by fluorometric high-performance liquid chromatography. *Methods Enzymol.* **362**, 543–560. https://doi.org/10.1016/S0076-6879(03)01036-X (2003).
48. Vaill, M., Chen, D. Y., Diaz, S. & Varki, A. Improved methods to characterize the length and quantity of highly unstable PolySialic acids subject category:(Carbohydrates, chromatographic techniques). *Anal. Biochem.* **635**, 114426. https://doi.org/10.1016/j.ab.2021.114426 (2021).

Acknowledgements

We gratefully thank Martina Mühlenhoff (MHH, GER) for the gift of endoN and mAb735, Joachim Bentrop (KIT, GER) for the activated-leukocyte cellular adherence molecule (ALCAM) 3×FLAG tagged recombinant plasmid, Wu Di for the rainbow trout CSS (rtCSS) and Lan-Yi Chang for the PSGP-L and PSGP-H preparation. In addition, the authors thank for the excellent technical help of Gesine Krüger.

Author contributions

A.H.-L. and S.P.G. designed the experiments, administrated the project, cured data and acquired funding; M.T.V. and A.R. conducted fish tissues extractions and RNA purification. M.D. M.T.V. and V.C. performed molecular biology and obtained molecular constructs; M.N. and M.T.V. performed initial expression and initial biochemical characterization of the fish enzyme. M.D. conducted production and biochemical characterization of all the fish polysialyltransferases. M.D. and A.S. conducted polySia synthesis and structural characterization. A.S. and S.P.G. developed the polysialation on centrifugal filter units strategy and conducted structural HPLC characterization of polySias. V.R., C.B. and C.L. chemo-enzymatically synthetized SiaNAl and performed NMR experiments. M.D., M.N. and C.L. synthetized natural and unnatural CMP-Sias and M.D. prepared glycoproteinic acceptors. M.N., C.L., C.B. and A.H.L. developped the MPSA. R.E.T. performed molecular modeling and structural studies. C.S. conducted cell culture, exopolysialylation and confocal microscopy analyses. A.H–L, V.C., A.R. and S.P.G supervised experiments, A.H–L, S.P.G. and M.D. wrote initial manuscript. All authors contributed to editing the manuscript.

Funding

M.D. is a recipient of a doctoral fellowship from the French Ministry and Region Hauts-de-France. The contribution of the COST Action CA18103-INNOGLY supported by the European Cooperation in Science and Technology (COST) is greatly acknowledged. This work was supported by the CNRS, the ANR-21-CE44-0032 (project PsaMar), the German Academic Exchange Service (DAAD) (PN:57446225), a grant of the Deutsche Forschungsgemeinschaft (GA 1755 5–1), a PHC Procope grant (project n°42533RC) and the University of Lille.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-42095-0.

Correspondence and requests for materials should be addressed to S.P.G. or A.H.-L.

II. STRATÉGIE DU RAPPORTEUR CHIMIQUE APPLIQUÉE AU SUIVI D'UNE MODIFICATION POST-TRADUCTIONNELLE : LA O-GLCNACYLATION

À l'heure actuelle, aucune technologie n'est disponible pour effectuer le suivi au sein de cellules vivantes d'une protéine ciblée présentant une modification post-traductionnelle (MPT) spécifique. Cette limitation représente un obstacle majeur pour l'investigation desdites modifications dans divers contextes physiopathologiques. Parmi les MPT, la O-GlcNAcylation est très répandue dans le règne animal[106] (Figure 8). Ce processus de glycosylation implique l'ajout réversible, par liaison β-O-glycosidique, d'un résidu unique de D-*N*-acétylglucosamine (GlcNAc) sur les groupements hydroxyles de sérines ou de thréonines présentes dans les protéines localisées dans le cytosol, le noyau et la mitochondrie. La O-GlcNAcylation, ainsi que les enzymes qui régulent cette modification, à savoir la O-GlcNAc transférase (OGT) et la O-GlcNAcase (OGA), sont perturbées dans de nombreuses pathologies, telles que les maladies cardiovasculaires et les cancers.[107] À ce jour, plusieurs milliers de protéines ont été identifiées comme étant O-GlcNAcylées, parmi lesquelles on retrouve la β-caténine.[108]

Notre équipe ainsi que celle du Pr. T. Lefebvre et la société E-zyvec se sont rassemblées pour un projet dans ce contexte. Nous avons conçu un senseur permettant le marquage et la visualisation de la β-caténine dans sa forme O-GlcNAcylée en utilisant la technologie de transfert d'énergie par résonance intermoléculaire de Förster (FRET) au niveau cellulaire.[109] Notre stratégie repose sur l'utilisation de rapporteurs chimiques bioorthogonaux, combinant un double marquage de la β-caténine par une protéine fluorescente verte (GFP) pour la séquence protéique et une sonde chimiquement cliquée, ce qui permet de réaliser un test FRET qualitatif rapide et facile à mettre en œuvre. Nous avons validé cette technique en visualisant le profil de O-GlcNAcylation de la β-caténine dans les cellules HeLa. Les modifications dans la O-GlcNAcylation de la β-caténine ont été induites en perturbant les niveaux globaux de O-GlcNAc dans la cellule à l'aide d'inhibiteurs de l'O-GlcNAc transférase (OGT) et de l'O-GlcNAcase (OGA).

Dans le cadre de ce travail, ma contribution a porté sur l'étude visant à déterminer la concentration optimale d'Ac_4GalNAz (le rapporteur chimique de la GlcNAc) pour le marquage cellulaire et l'analyse FRET. Les cellules HeLa ont été exposées à des concentrations croissantes d'Ac_4GalNAz (10, 20, 50, 100, 200, 300, 400, 500, 600, 800, 1000 µM) afin d'introduire l'O-GlcNAz au niveau des protéines intracellulaires O-GlcNAcylées (voir Figure 4 de l'article). Par la suite, nous avons évalué la viabilité cellulaire en utilisant le test de prolifération cellulaire MTS. Aucun impact significatif sur la prolifération n'a été observé jusqu'à des concentrations atteignant 200 μ M d'Ac4GalNAz, ce qui suggère que le traitement n'a pas engendré d'effets cytotoxiques.

molecules

MDPI

Article

Exploring the Potential of β-Catenin *O*-GlcNAcylation by Using Fluorescence-Based Engineering and Imaging

Angelina Kasprowicz [1], Corentin Spriet [1,2], Christine Terryn [3], Vincent Rigolot [1], Stephan Hardiville [1], Matthew G. Alteen [4], Tony Lefebvre [1] and Christophe Biot [1,*]

1 Univ. Lille, CNRS, UMR 8576–UGSF–Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France; angelina.kasprowicz@univ-lille.fr (A.K.); corentin.spriet@univ-lille.fr (C.S.); vincent.rigolot.etu@univ-lille.fr (V.R.); stephan.hardiville@univ-lille.fr (S.H.); tony.lefebvre@univ-lille.fr (T.L.)
2 Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41–UMS 2014–PLBS, F-59000 Lille, France
3 PICT Platform, University of Reims Champagne-Ardenne, 51 rue Cognacq-Jay, 51100 Reims, France; christine.terryn@univ-reims.fr
4 Department of Chemistry, Simon Fraser University, Burnaby, BC V5A 1S6, Canada; malteen@sfu.ca
* Correspondence: christophe.biot@univ-lille.fr; Tel.: +33-(0)3-20-43-61-41

Received: 16 July 2020; Accepted: 24 September 2020; Published: 1 October 2020

Abstract: Monitoring glycosylation changes within cells upon response to stimuli remains challenging because of the complexity of this large family of post-translational modifications (PTMs). We developed an original tool, enabling labeling and visualization of the cell cycle key-regulator β-catenin in its *O*-GlcNAcylated form, based on intramolecular Förster resonance energy transfer (FRET) technology in cells. We opted for a bioorthogonal chemical reporter strategy based on the dual-labeling of β-catenin with a green fluorescent protein (GFP) for protein sequence combined with a chemically-clicked imaging probe for PTM, resulting in a fast and easy to monitor qualitative FRET assay. We validated this technology by imaging the *O*-GlcNAcylation status of β-catenin in HeLa cells. The changes in *O*-GlcNAcylation of β-catenin were varied by perturbing global cellular *O*-GlcNAc levels with the inhibitors of *O*-GlcNAc transferase (OGT) and *O*-GlcNAcase (OGA). Finally, we provided a flowchart demonstrating how this technology is transposable to any kind of glycosylation.

Keywords: bioorthogonal chemistry; fluorescence; glycosylation; metabolic incorporation; GFP; β-catenin

1. Introduction

The complete set of glycans, i.e., the glycome, represents a large diversity of structures and functions and plays a central role in biology and health [1]. Glycans are involved in many biological processes, such as protein folding and trafficking, cell-to-cell recognition, immune defense, and cell protection [1]. Glycans are also specific markers for numerous pathologies, including cancers [2]. This functional diversity is linked to a great structural diversity of glycans. Indeed, a wide variety of different forms of glycosylation are possible, and, moreover, variation in glycosylation of the same protein within tissues can occur, leading to various "glycoforms" of a given protein [3].

The various forms of glycosylation are essential for cell homeostasis. In this regard, it is estimated that 2% of the human genome is devoted to glycosylation. The most complex forms of glycosylation, namely *N*-linked protein, *O*-linked protein, and *O*-linked mucin-type glycosylation, take place mainly in the endoplasmic *reticulum* and the Golgi apparatus where the requisite biosynthetic enzymes reside. Another widespread form of glycosylation is *O*-linked β-*N*-acetylglucosaminylation (*O*-GlcNAcylation) [4]. *O*-GlcNAcylation is much simpler than glycosylation occurring within the

secretory pathway and consists of the modification of cytoplasmic, nuclear, and mitochondrial proteins with a single residue of *N*-acetylglucosamine (GlcNAc). Furthermore, in contrast to the more complex glycosylation of the secretory pathway, *O*-GlcNAcylation is highly dynamic, its versatility being managed by a unique pair of enzymes—*O*-GlcNAc transferase (OGT) and *O*-GlcNAcase (OGA)—which, respectively, install and remove the GlcNAc moiety on and off targeted substrate proteins [5]. *O*-GlcNAcylation can compete with phosphorylation at the same or at neighboring sites, leading to a reciprocal interplay between the two post-translational modifications (PTMs) [4,6].

During the last three decades, metabolic oligosaccharide engineering (MOE, also termed metabolic glycoengineering (MGE)) has paved the way for the manipulation of biosynthetic pathways responsible for oligosaccharide and glycoconjugate production [7–10]. MOE exploits the substrate promiscuity of different glycosylation pathways to incorporate non-natural monosaccharides into cellular glycans. Upon metabolization, the non-natural saccharides are detected with a complementary probe through bioorthogonal ligation. This technology has been mostly used for labeling cell surface-exposed glycoconjugates, especially with fluorescent probes [7,8], but there are some intracellular MOE applications, notably targeting *O*-GlcNAc.

However, these strategies usually aim to report the general glycosylation status of a cell, and, to date, only a few MOE methods have been reported that account for the glycosylation pattern of a single specific glycoprotein. The five examples are based on dual-labeling strategies, exploiting Förster resonance energy transfer (FRET). Haga and coworkers used azido sialic acid (SiaNAz) labeling of GLUT4-GFP to image cell surface glycoproteins [3]. Belardi and coworkers developed a technique using Ac_4ManNAz (the precursor of SiaNAz) to study the vitronectin receptor $\alpha_V\beta3$ after targeting with a donor fluorophore-labeled Fab fragment [11]. The dual-labeling of membrane proteins developed by Lin and coworkers relies on two bioorthogonal chemical reporters delivered to both the protein and the glycan [12]. Surprisingly, only modest progress has been made in studying intracellular *O*-GlcNAcylation. The two examples to visualize *O*-GlcNAcylation in a protein-specific manner have been reported by Lin and coworkers [13] and Doll and coworkers [14].

Herein, we presented optimized methods for probing the glycosylation state of β-catenin within HeLa cells using fluorescence-based imaging. β-catenin plays many critical functions in animals. First, Wnt/β-catenin signaling—for which the protein is the central component—is actively involved in embryonic development from flies to mammals [15]. β-catenin is also crucial for cell cycle progression and for the maintenance of the adherens junctions in epithelia. β-catenin is a proto-oncoprotein, its upregulation being responsible for tumorigenesis in a wide variety of cancers. The fate of the protein is well established as being controlled by phosphorylation of a D-box located at the *N*-termini: phosphorylation at S33, S37, T41, and S45 triggers the poly-ubiquitination of β-catenin and its subsequent targeting to the 26S proteasome [16]. Therefore, mutations of the D-box or deletions of the N-termini of β-catenin are frequently found in patient tumors. We previously demonstrated that the *O*-GlcNAcylation of β-catenin reduced its susceptibility to proteasomal degradation. In particular, the modification of T41 with *O*-GlcNAc was found to interfere with the sequential phosphorylation and ubiquitinylation, which led to the destruction of this protein [6]. Given these important biological roles and defined functions for specific post-translational modifications, β-catenin appeared to us as the model of choice for the development of our strategy to visualize specific glycosylated proteins in cells.

2. Results

Among the in-situ visualization techniques, the use of GFP and its derivatives is the gold standard to study genetically encoded biomolecules, such as proteins. However, for all other non-genetically encoded biomolecules, such as glycans or lipids, GFP-based imaging techniques are not feasible. Therefore, it is necessary to exploit novel techniques to enable visualization of these compounds.

We opted for a bioorthogonal chemical reporter strategy to detect the *O*-GlcNAc modification state of β-catenin. This technology is based on the incorporation of a chemically-modified metabolic

precursor of UDP-GlcNAc [17], bearing a pendant bioorthogonal handle at the 2-position of the pyranose ring. This chemically reactive analog permits conjugation with a fluorophore to generate an imaging probe. This approach allows the in vivo study of *O*-GlcNAc-modified β-catenin without interfering with native biological processes (Figure 1). The technology we proposed involves the dual-labeling of β-catenin with a genetically encoded GFP tag, followed by the bioorthogonal installation of a second fluorophore on the *O*-GlcNAc unit. The combined use of two fluorescent reporters in close proximity to each other enables a convenient and efficient FRET assay suitable for imaging cells.

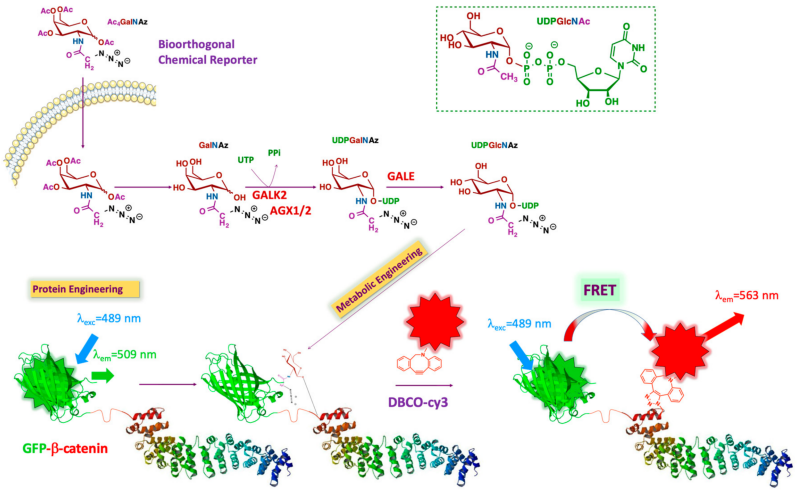


Figure 1. Schematic representation of the strategy: FRET-based imaging of dual-labeling of protein and glycans. FRET, Förster resonance energy transfer.

β-catenin is genetically labeled with GFP (donor). The intracellular *O*-GlcNAcylated proteins are labeled by metabolic oligosaccharide engineering (MOE), followed by bioorthogonal chemistry with a fluorophore (acceptor). As the distance between donor and acceptor is less than 10 nm, the only acceptor bound to the glycans on β-catenin is excited via intramolecular FRET. The β-catenin structure is adapted from Huber et al. [17].

Briefly, the chemical structure of UDP-GlcNAc is shown in the upper right in Figure 1 (green box). The different colors represent the various metabolic origins of the nucleotide-sugar, as follows: brown, carbohydrate metabolism; blue, glutamine metabolism; purple, metabolism of fatty acids and ketogenic amino acids; green, nucleotide metabolism. The unnatural GlcNAc analog used for our tracing strategy was Ac_4GalNAz (see below). This reporter enters the cell by passive diffusion. Thereafter, Ac_4GalNAz is deacetylated by esterases and then further metabolized to generate UDP-GalNAz. After isomerization into UDP-GlcNAz by the epimerase GALE (UDP-glucose 4-epimerase), OGT transfers GlcNAz to a set of proteins, including GFP-β-catenin. The addition of a compatible alkyne-labeled fluorophore results in the conjugation of the fluorophore to the azide handle of the GlcNAz unit through a strain-promoted azide-alkyne cycloaddition (SPAAC). This permits FRET between the GFP group and the fluorophore conjugated to the sugar, making it possible to specifically visualize *O*-GlcNAcylated β-catenin in cells.

2.1. Optimization of Fusion Protein Linker

The linker between β-catenin and GFP is an important consideration in order to satisfy the requirements for FRET, including the distance between fluorophore pairs (<10 nm) and dipole orientation. Therefore, four different fusion proteins were investigated with either no linker or 1-to-3 α-helix linkers (1a, 2a, and 3a) inserted between the GFP and the β-catenin domains: (1) GFP-β-catenin, (2) GFP-1a-β-catenin, (3) GFP-2a-β-catenin, and (4) GFP-3a-β-catenin (Figure 2a). As a lack of rigidity of these linkers can be a limitation for FRET-based systems [18], a semi-flexible peptide linker

$(EAAKEAAAKEAAAKEAAAKA)_{1\text{-}3}$, which adopts an alpha-helix conformation, was chosen to bridge the domains. We envisioned this linker would strike a balance in the degree of mobility between the two domains, maximizing the potential for efficient FRET.

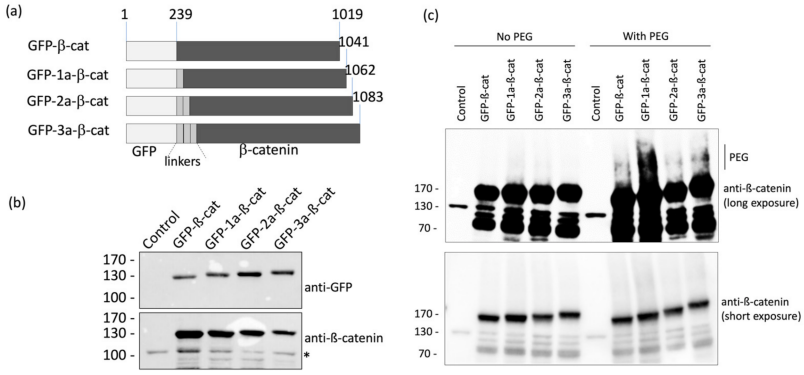


Figure 2. Design and validation of fluorescent β-catenin fusion proteins. (**a**) Schematic representation of the different GFP-β-catenin constructs used in this study. (**b**) HeLa cells were transfected with the various GFP-β-catenin expressing vectors. Cell lysates were analyzed by western blots using anti-GFP and anti-β-catenin antibodies. The asterisk indicates the detection of the endogenous form of β-catenin. (**c**) Protein extracts of HeLa cells were labeled with 4.4 kDa DBCO-PEG mass tag or incubated with DMSO. β-catenin proteins were detected by Western blot.

The different N-terminal GFP-tagged β-catenins were produced in HeLa cells and analyzed by Western blotting in order to assess the expression level of each form (Figure 2b). We found that all four fusion protein constructs were suitably expressed and, therefore, chose to assess each variant in microscopy experiments. Note that the insertion of linkers increased the apparent molecular weight of GFP-β-catenin depending on the number repeats between the two domains.

In order to ascertain that the different GFP-β-catenin constructs were modified following the incubation of HeLa cells with $Ac_4GalNAz$, a mass tag labeling was performed, as described in the materials and methods session. As shown in Figure 2c, we observed that the different constructs were PEGylated, as indicated by the shift in the upper molecular weight when compared to the control condition (no PEG). This experiment asserted that the GFP-β-catenin constructs were labeled with the GlcNAz moiety and then clicked by the DBCO-PEG.

We next evaluated the four GFP-β-catenin constructs via fluorescent microscopy using fixed Hela cells (Figure 3). We found that all constructs produced acceptable levels of fluorescent signal and subcellular localization. To ensure that FRET will be related to the proximity between donor and acceptor and not to the relative orientation of each modification, we chose the most flexible construct (GFP-3a-β-catenin) for FRET experiments. Indeed, computational simulations showed that the orientation factor converged to a constant in a FRET sensor where the donor/acceptor pair was presumed to be freely mobile. Thus, the FRET signal of the chosen construct would be related only to the donor/acceptor distance [19].

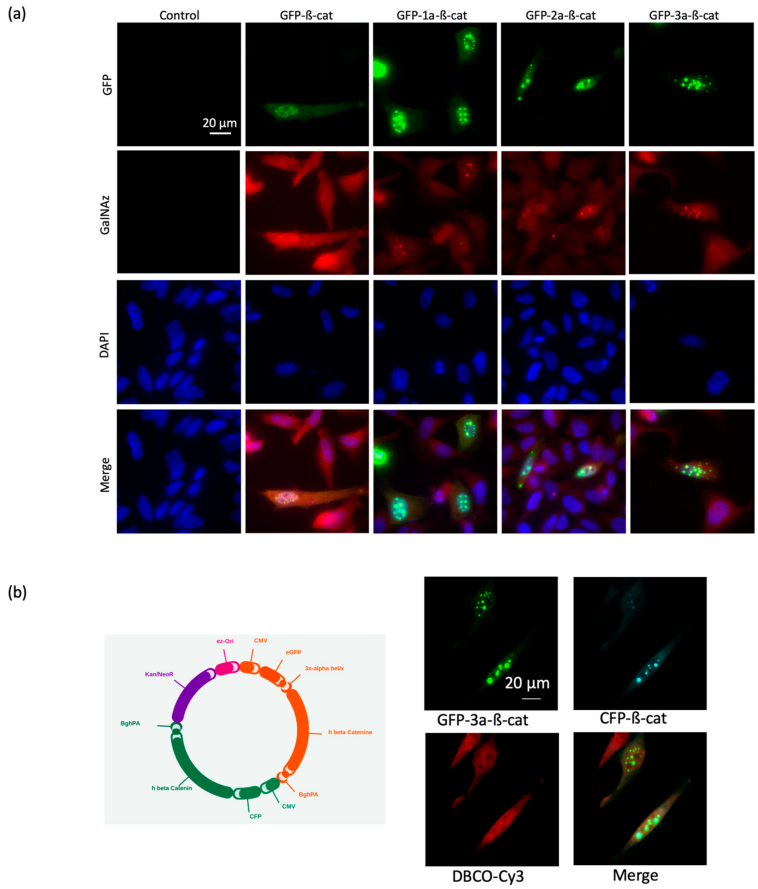


Figure 3. Evaluation of the GFP-β-catenin constructs via fluorescent microscopy. (**a**) Fluorescence images of HeLa cells expressing GFP-linker-β-catenins, treated with 200 µM Ac_4GalNAz and labeled with DBCO-cy3. (**b**) Fluorescence images of HeLa cells expressing GFP-3a-β-catenin and CFP-β-catenin, treated with 200 µM Ac_4GalNAz and labeled with 10 µM DBCO-cy3. The map of the vector encoding both the GFP-3a-β-catenin and the CFP-β-catenin is shown on the left. CFP, cyan fluorescent protein.

Interestingly, a di-systronic expression vector was designed and produced in order to combine the direct synthesis and secretion of both the GFP-3a-β-catenin and a cyan fluorescent protein β-catenin construct (CFP-β-catenin). Fluorescence colocalization microscopy demonstrated that both proteins localized in the same subcellular compartments (Figure 3b). We chose to proceed with the GFP-β-catenin construct in future experiments of the improved spectral overlap of GFP emission with the excitation band of the Cy3 acceptor fluorophore. We also concluded that the length of the linker between the GFP and β-catenin domains had no dramatic effect on the integrity of the fusion protein.

2.2. Optimization of Metabolic Oligosaccharides Engineering (MOE)

The ligation reactions used in the metabolic oligosaccharides engineering (MOE) technology are based on the principle of "click chemistry" introduced by the group of Sharpless [20]. A wide range of bioorthogonal reactions has been developed, including nucleophilic substitution reactions, carbonyl chemistry, and cycloaddition reactions between unsaturated compounds, such as Diels-Alder reactions or 1,3-dipolar cycloaddition reactions. Here, we chose the strain-promoted alkyne-azide cycloaddition (SPAAC) due to its ability to label glycans in live cells [21]. Being one of the most widely used cyclooctyne derivatives due to its robust stability and reactivity, we selected a fluorophore carrying a dibenzylcyclooctyne (DBCO), also known as ADIBO (azadibenzocyclooctyne) or DIBAC (dibenzoazacyclooctyne) [22].

As N-azidoacetylglucosamine (GlcNAz) turned out to be a weak metabolic labeling reagent [17], we instead used N-azidoacetylgalactosamine (GalNAz), which is more efficiently metabolized into UDP-N-GlcNAz in cells. In its peracetylated form (Ac_4GalNAz), this derivative has the property of being cell-permeable.

To address the optimum Ac_4GalNAz concentration for cell labeling and FRET analysis, HeLa cells were treated with increasing Ac_4GalNAz concentration (10, 20, 50, 100 and 200, 300, 400, 500, 600, 800, 1000 μM) to introduce the O-GlcNAz onto intracellular O-GlcNAcylated proteins (Figure 4a). The viability of cells was then assessed using an MTS cell proliferation assay. No significant impact on proliferation was observed at concentrations up to 200 μM Ac_4GalNAz, indicating that treatment did not induce cytotoxic effects.

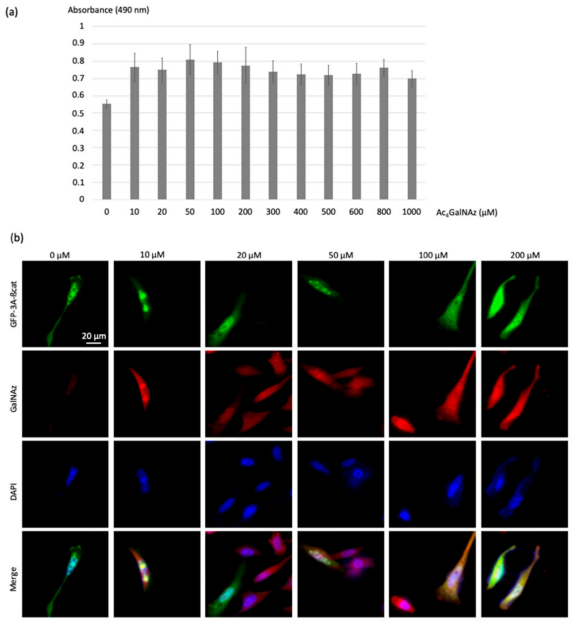


Figure 4. Optimization of the concentration of the chemical reporter. (**a**) The histogram represents the cytotoxicity evaluation of the chemical reporter Ac_4GalNAz; Hela cells were treated at a concentration of 0 to 1000 μM. Twenty-four hours later, the cell viability was determined using the MTS assay. (**b**) HeLa cells transfected with the GFP-3a-β-catenin expressing vectors were treated with increasing Ac_4GalNAz concentration (10, 20, 50, 100, and 200 μM) and labeled with DBCO-cy3 (10 μM).

The azidosugar was then labeled with the DBCO-cy3 FRET acceptor (10 μM) using the SPAAC reaction, permitting visualization of *O*-GlcNAz-modified β-catenin. Colocalization of green and red signals produced a yellow color, validating the dual-labeling (Figure 4b).

As shown in Figure 4b, no modification of the morphology of the cells was observed even at the highest dose of 200 μM of the chemical reporter, showing that the toxicity of the azidosugar concentration was negligible even at the highest concentrations. Acetic acid release inside cells upon enzymatic deacetylation of Ac_4GalNAz did not seem to induce cytotoxicity, as was previously reported for Ac_4ManNAz (>50 μM) [23]. In order to promote high FRET efficiencies, we chose the highest non-toxic Ac_4GlcNAz concentration of 200 μM.

2.3. SLiM-FRET Readout Assay

To carry out a quantitative analysis of the FRET between the GFP and the cy3-labeled GlcNAc, a spectral fluorescence lifetime imaging microscopy (SLiM) acquisition system was used, thus acquiring, simultaneously, 16 photon decay curves along the fluorescence emission spectrum, and permitting unambiguous FRET measurements (Figure 5) [24]. In the absence of the chemical reporter, a mean fluorescence lifetime of 2.8 +/− 0.06 ns was observed in GFP-3a-β-catenin expressing cells (Figure 5a).

We then imaged GFP-3a-β-catenin expressing cells that were treated with Ac_4GalNAz and click-labeled with DBCO-cy3. Using these conditions, the mean fluorescence lifetime of the GFP reporter in Ac_4GalNAz-treated cells decreased to 2.5 +/− 0.17 ns, which indicated that some of the GFP energy was transferred to the nearby cy3 fluorophore (Figure 5a). A FRET event could thus be observed with a p-value < 0.05.

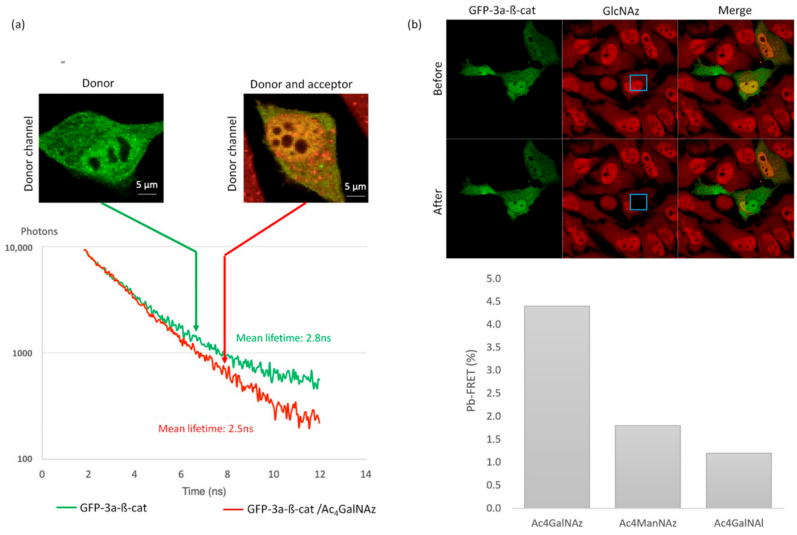


Figure 5. Proof of feasibility: FRET imaging of *O*-GlcNAcylated GFP-3a-β-catenin. (**a**) SLiM-FRET experiment using two-photon excitation: representative photon decay curves and associated fluorescence mean lifetime extracted from only the GFP emission channel. (**b**) Donor dequenching after acceptor photobleaching. The inserted blue box indicates the photobleached area. Histograms correspond to calculated pb-FRET inside the cells after photobleaching in comparison to Ac_4GalNAl and Ac_4ManNAz, which serve as a negative control.

These experiments allowed us to observe variations of up to 300 ps in lifetime experiments. This range corresponded to a deviation from the GFP-3a-β-catenin basal *O*-GlcNAcylation level to the non-*O*-GlcNAcylated state (Figure 5b). This dynamic range was sufficient to decipher subtle variations in *O*-GlcNAcylation levels.

We thus demonstrated that our GFP-3a-β-catenin GalNAz-cy3 SLiM-FRET strategy is sufficiently sensitive for use in live cells. This important point at this level of the study can be considered as a proof of concept for this new tool.

2.4. Pb-FRET Readout Assay

While SLiM allows quantitative and sensitive FRET measurements, it requires a dedicated acquisition system, and it is highly time-consuming. After validation of the biosensor using SLiM, we wanted to demonstrate the robustness of our method using a traditional confocal microscope to monitor FRET through a photobleaching FRET readout assay, a fast and widely available technology.

As depicted in Figure 5b, DBCO-cy3 was specifically photobleached within a defined region of the cell in order to keep an internal control [24]. Using pb-FRET, we observed a measurable FRET efficiency of 4.4% +/− 0.1% on GFP-3a-β-catenin expressing cells when treated with Ac_4GalNAz and click-labeled with DBCO-cy3.

In order to confirm that measured FRET resulted from intramolecular FRET, we performed control experiments with GFP-3a-β-catenin and the alkyne-containing reporter Ac_4GalNAl [19] (Figure 5b). This analog is unable to react with DBCO-cy3, allowing the evaluation of the possible contributions from intermolecular FRET between the two fluorophores. More than a thousand intracellular proteins are post-translationally modified with *O*-GlcNAc and thus susceptible to the incorporation of GlcNAz,

followed by DBCO-cy3. However, due to the relatively low expression of GFP-β-catenin in our model system, we hypothesized that the concentration of donor and acceptor fluorophores would be low enough that contributions from intermolecular FRET would be negligible [25]. In support of this assumption, no measurable FRET compared to GFP-β-catenin alone was monitored, as confirmed by experiments presented in Figure 5b. Indeed, a low background signal of 1.2% +/− 0.5% was recorded when cells were treated with Ac_4GalNAl.

Further, as control with another tagged monosaccharide (ManNAz), which cannot be metabolized into GlcNAz, we observed a low FRET efficiency of 1.8% +/− 0.7% on GFP-3a-β-catenin expressing cells when treated with Ac4ManNAz and click-labeled with DBCO-cy3 (Figure 5b). This is a typical example of the advantage of GalNAz labeling of β-catenin.

2.5. Pharmacology

Having developed an efficient, robust, and easy to use the technique to process Hela cells for pb-FRET imaging, we used this approach to evaluate the pharmacological effects of thiamet-G and $Ac_4$5SGlcNAc. Thiamet-G [26] and $Ac_4$5SGlcNAc [16] are effective cell-active inhibitors of O-GlcNAcase (OGA) and O-GlcNAc transferase (OGT), respectively. OGT catalyzes the addition of O-GlcNAc to proteins [27]; OGA removes the modification [28]. Thiamet-G is a potent (in vitro K_i = 2.1 nM, cell-based EC_{50} = 21 nM) stable mimic of the oxazoline-like transition state used by OGA active site during the OGA-catalyzed hydrolysis of O-GlcNAc, leading to an increase of O-GlcNAc modification of proteins. Within cells, the prodrug $Ac_4$5SGlcNAc generates the OGT inhibitor UDP-5SGlcNAc (in vitro K_i = 5 μM) and decreases cellular O-GlcNAcylation (cell-based EC_{50} = 5 μM). Both compounds have been evaluated in SK-N-AS cells for imaging the changes of tau *O*-GlcNAc [12], but to the best of our knowledge, this is the first study to compare the inhibitory effect of thiamet-G and $Ac_4$5SGlcNAc upon *O*-GlcNAcylation on β-catenin in living cells (Figure 6).

When GFP-3a-β-catenin-transfected HeLa cells were treated with 1 μM thiamet-G 2 h before or after incubation with 200 μM Ac_4GalNAz, and finally labeled with DBCO-cy3, a decrease of the pb-FRET signal was recorded consistent with a higher rate of occupancy of O-GlcNAc moieties on β-catenin in response to OGA inhibition (Figure 6).

Conversely, when GFP-3a-β-catenin-transfected HeLa cells were incubated with 100 μM $Ac_4$5SGlcNAc for 2 h prior to the treatment with Ac_4GalNAz and labeled with DBCO-cy3, a large increase (2.1%; see Figure 6) of the pb-FRET signal was recorded compared to the untreated cells. This observation was consistent with previous studies [16,29] that showed a major reduction in cellular *O*-GlcNAc when using this inhibitor. Indeed, the expression of OGT increased in response to its inhibition [30–32]. Thus, the blockade of its activity by UDP-5SGlcNAc resulted, as expected, in higher OGT expression, as previously demonstrated and also seen with the OGT inhibitor OSMI-1 [30] or upon glucose deprivation [31,32]. It should also be noted here that the substrate was present in an excess amount, compared to the inhibitor in our step-up, as Ac_4GalNAz concentration was higher than $Ac_4$5SGlcNAc concentration, and effective inhibition nevertheless still occurred. On the other hand, adding the OGT inhibitor after treating the cells with Ac_4GalNAz had no effect, with the same pb-FRET signal being recorded as that of the control cells incubated only with the azido-sugar. The absence of an increase in signal was most likely due to the fact that the incubation time with the inhibitor was too short to observe an increase in the expression of OGT, as we assumed for the previous condition.

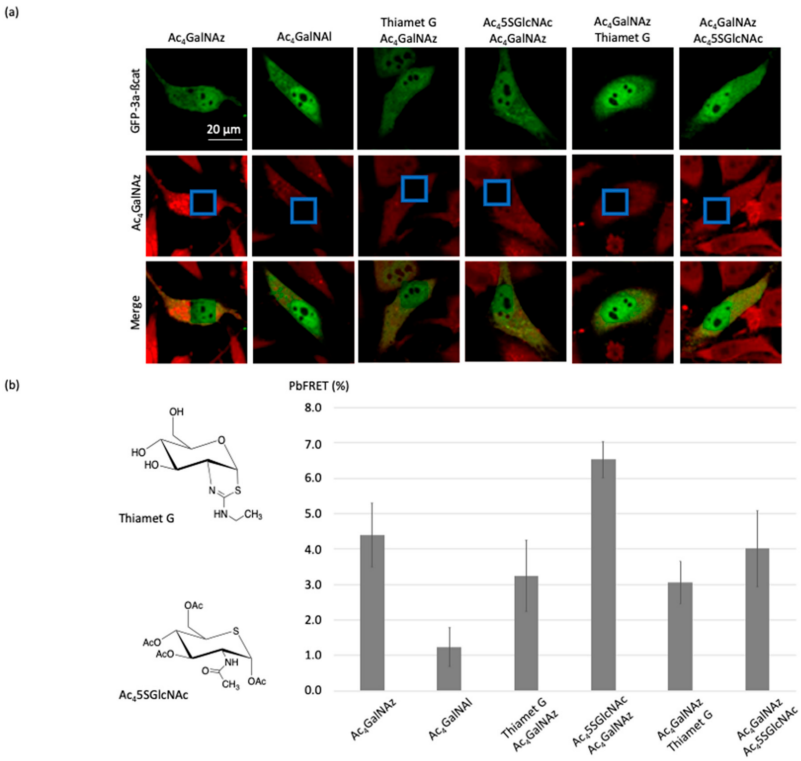


Figure 6. Imaging the changes of pb-FRET for GFP-3a-β-catenin in HeLa cells induced by the OGT and OGA inhibitors. (**a**) Fluorescence images of the *O*-GlcNAcylated glycoform of β-catenin before and after photobleaching using Thiamet-G or $Ac_4$5SGlcNAc. (**b**) Histograms corresponding to calculated pb-FRET inside the cells after photobleaching using thiamet-G or $Ac_4$5SGlcNAc, whose structures are shown at the left.

3. Discussion

The Human Genome Project consortium revealed that human DNA coded for only a third of the 100,000 genes expected [33]. Since then, this number is continuing to be lowered with a current estimate of about 21,000 genes. Yet, the human proteome is very broad since it is estimated to be capable of regulating more than 500 million different biological activities. In addition to the mechanisms of alternative splicing of the transcribed primary proteins, post-translational modifications increase exponentially the number of possible protein isoforms [4]. These PTMs alter proteins functions through various mechanisms, such as changes in protein-protein interactions, altered stability, or adaptation of catalytic activity to the environmental context. Owing to the vast diversity and the complexity of PTMs, studying them is usually challenging; the difficulty comes with the specific tagging of one precise form of the protein of interest and the lack of imaging tools to permit tracking of this protein form [3,7,11,13]. In this regard, the FRET technologies described in this work may help with visualizing the glycosylation state of specific glycoproteins and also provide the bases for characterizing glycomes with molecular precision.

Our present work refers to the specific *O*-GlcNAcylation PTM that occurs mostly intracellularly. Indeed, *O*-GlcNAcylation regulates a large set of biological functions in a nutrient-dependent manner [4,34]. Dysregulation of *O*-GlcNAcylation processes is observed in a variety of diseases, including neurodegenerative diseases, metabolic disorders, cardiovascular disorders, and cancers.

A better understanding of the *O*-GlcNAcylation function, protein by protein, would certainly translate in significant progress in the knowledge and treatment of these diseases. However, there are few existing molecular tools that allow a rapid and easy determination of the *O*-GlcNAcylation status of a protein of interest in live cells, rendering such progress difficult to make.

In this study, we successfully adapted the dual-labeling method initially developed by Lin and co-workers [13] to detect a β-catenin glycoform using FRET techniques. As the selection of a suitable linker of GFP-tagged proteins is often neglected and underexplored, we focused a part of our study in the design and the choice of the linker between GFP and β-catenin. As FRET is contingent upon the ability to precisely introduce a suitable pair of donor and acceptor fluorophores into the protein of interest, we concomitantly optimized not only the linker but also the chemical reporter concentration. We validated our approach using a combination of azido sugar labeling and two-photon fluorescence lifetime imaging spectroscopy (SLiM). As a negative control, we also used an alkyne sugar analog (which is unable to react with the selected fluorophore via a SPAAC reaction) to confirm intramolecular FRET. Thereafter, we turned our attention to demonstrate the robustness of our technology using a conventional confocal microscope to monitor FRET through a pb-FRET readout assay. With this technology in our hands, we evaluated changes of β-catenin O-GlcNAcylation by inhibiting OGA or OGT. Finally, we provided a flowchart demonstrating how these technologies could be meaningfully combined (Figure 7).

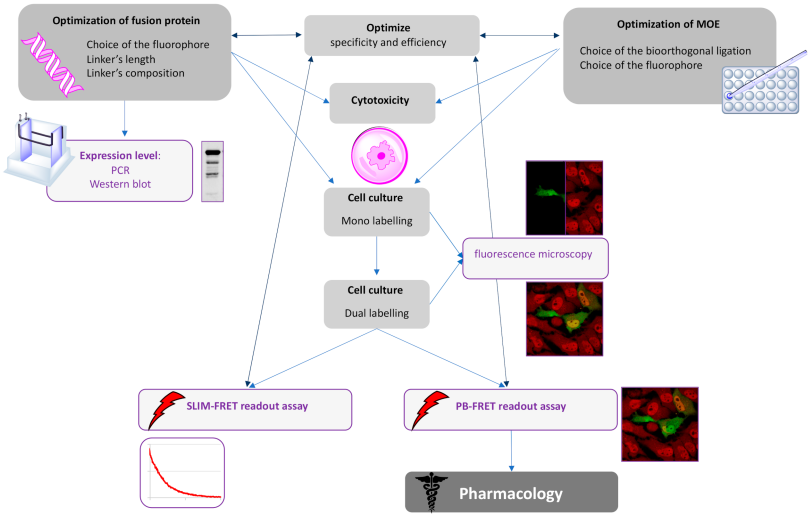


Figure 7. Flowchart demonstrating how to combine the technologies for FRET-based imaging of dual-labeling of glycoproteins.

We expect that the proof of concept established by our study will serve as a general and useful approach to studying the role of *O*-GlcNAcylation on specific protein targets. Our strategy permits a more precise way of studying these effects compared with pharmacological inhibition, which results in global perturbations of *O*-GlcNAc levels. Additionally, our approach using β-catenin as an acceptor protein represents an effective cell-based reporter assay of OGT activity, which could be further optimized for a variety of applications. For example, we believe that our methodology could be adapted for future studies to screen for inhibitors of OGT or OGA.

This methodology, developed on a precise cell system, protein, and PTM, should open up interesting perspectives in various fields of cell biology. This system should be fairly easily adaptable to any type of glycosylation, given that MOE is an established technology with a broad scope of sugars that can be incorporated in vitro. We also expect that in the future, this strategy will be useful for the study of other PTMs after the requisite optimization towards a particular type of modification. In view of the speed of coordination between chemists and biologists, such tools should soon emerge for many modifications other than glycosylations.

4. Materials and Methods

4.1. Chemical Compounds

Ac_4GalNAz, Ac_4GAlNAl, and DBCO-Cy3 were purchased from Click Chemistry Tools (Scottsdale, USA). Thiamet-G and $Ac_4$5SGlcNAc were prepared, as previously described [26,30].

The tetraacetylated *N*-azidoacetyl-mannosamine (Ac_4ManNAz) was synthesized from ManNAz. The crude product was purified by silica column chromatography with Cyclohexane/Ethyl acetate (2:1 *v*/*v*) and characterized by NMR ^{1}H, COSY, and HSQC experiments (yield 88%). The obtained white powder corresponded to a pure mixture of non-assessed α/β anomers (60/40).

^{1}H NMR (300 MHz, $CDCl_3$) δ = 6.64 (d, J = 8.9, 1H'), 6.56 (d, J = 9.2, 1H), 6.05 (d, J = 1.9, 1H), 5.89 (d, J = 1.6, 1H'), 5.34 (dt, J = 11.7, 5.9, 1H), 5.27–5.12 (m, 1H + 1H'), 5.10−5.00 (m, 1H'), 4.73 (ddd, J = 9.0, 3.8, 1.6, 1H'), 4.62 (ddd, J = 9.3, 4.2, 1.9, 1H), 4.25 (dt, J = 12.4, 4.4, 1H + 1H'), 4.19−3.99 (m, 4H + 3H'), 3.82 (ddd, J = 9.6, 4.5, 2.5, 1H'), 2.19 (s, 3H + 3H'), 2.12 (s, 3H + 3H'), 2.05 (s, 3H + 3H'), 2.01 (s, 3H + 3H').

4.2. Cell Culture

HeLa cells were obtained from the American Tissue Culture Collection. Cells were grown in Dulbecco's modified Eagle's medium (Lonza) supplemented with 10% (*v*/*v*) of fetal-calf serum (Lonza). Cells were maintained at 37 °C in a humidified atmosphere containing 5% (*v*/*v*) CO_2.

4.3. Cell Toxicity

HeLa were seeded in 96-well plate (15,000 cells/well) in 100 µL DMEM. After 24 h, the medium was replaced with DMEM containing Ac_4GalNAz at various concentrations (0, 10, 20, 50, 100, 200, 300, 400, 500, 600, 800, and 1000 µM). Each condition was replicated 6 times. Twenty-four hours later, MTS assay was performed: 20µL of MTS/PMS solution was added in each well, and cells were incubated for 2 h at 37 °C/5% (*v*/*v*) CO_2. The absorbance was then read at 490 nm using a microplate reader.

4.4. Plasmids

vCMVp-GFP-β-catenin, vCMVp-GFP-1a-β-catenin, vCMVp-GFP-2a-β-catenin, vCMVp-GFP-3a-β-catenin, vCMVp-GFP-3A-β-catenin-CFP-β-catenin were generated by e-Zyvec (www.e-zyvec.com, Loos, France). Linkers between GFP and β-catenin were an alanine-rich amino acid sequence repeated one, two, or three times: 5'-EAAAKEAAAKEAAAKEAAAKA-3'.

4.5. Transfections

For microscopy, cells were grown on glass coverslips and transfected at 70% confluency. Transfections were performed using 1 µL of Jetoptimus (Polyplus), 1 µg of the plasmid in 900 µL DMEM, according to the manufacturer's instruction. The transfections mix was replaced 4 h later with fresh medium containing Ac_4GAlNAz, Ac_4GalNAl, or Ac_4ManNAz (200 µM), with or without $Ac_4$5SGlcNAc (100 µM) or Thiamet G (1 µM), for 24 h.

4.6. Western Blot Analyses

Cells were washed once in ice-cold PBS and lysed in RIPA buffer (Tris/HCL 50mM, NaCl 150 mM, NP40 0,5% (*v*/*v*), EDTA 1 mM, Na_3VO_4 1mM, NaF 5 mM, pH 7.9) supplemented with a protease inhibitors mix (Roche Diagnostics) for 20 min on ice. The lysate was centrifuged at 5000 rpm, 4 °C for 10 min, and the supernatants were collected. Protein quantitation was determined by using the BCA assay (ThermoFischer). Equal quantities (15–20 µg) of proteins were loaded and separated by 10% SDS-PAGE. Proteins were transferred onto a nitrocellulose membrane (GE Healthcare) at 200 mA for 2 h. Membranes were blocked for 1 h at room temperature with 5% (*w*/*v*) nonfat dry milk in

Tris-buffered saline (TBS) Tween (TBST) buffer (15 mM Tris, 140 mM NaCl, 0.05% (*v*/*v*) Tween20, pH 8.0). Membranes were incubated overnight at 4 °C with anti-β-catenin (1:2000) (H102, Santa Cruz Biotechnology, Santa Cruz, CA, USA) or anti-GFP (1:1000) monoclonal antibodies. After three washes with TBST, membranes were incubated with corresponding secondary HRP-linked antibodies (1:10,000 in TBST) for 1 h at room temperature. After three washes in TBST, the detection was performed with a CCD camera (Fusion Solo, Vilber Lourmat, Marne-la-Vallee, France).

4.7. O-GlcNAc Mass Tag Labeling

The experimental procedure we followed was previously described by Leturcq et al. 2018 [34]. Briefly, HeLa cells were transfected with the different GFP-β-catenin constructs (Figure 2a) and treated with 200 μM of Ac_4GalNAz for 24 h. Then, 50 μg of the labeled proteins were resuspended in 20 mM HEPES and 1% (*w*/*v*) SDS, pH 7.9, and incubated for 1 h at room temperature with a 4.4 kDa DBCO-PEG (polyethylene glycol) at a final concentration of 10 mM in DMSO. DBCO-PEG was synthesized, as previously described [35]. One volume of PEGylated proteins was precipitated using the mix of chloroform:methanol:water (3:0.75:2) in order to remove the excess of DBCO-PEG. Proteins were heated at 95 °C for 10 min, resolved by SDS-PAGE, and analyzed by Western blot.

4.8. Sample Preparation and Bioorthogonal Ligation for Fluorescence Microscopy

Cells cultured on coverslips were washed three times in Dulbecco's Phosphate Buffer Saline (Lonza) containing calcium and magnesium. Cells were fixed with 4% (*w*/*v*) paraformaldehyde (pH 7.3) for 30 min at room temperature and washed with ice-cold PBS thrice. Coverslips were switched in a humid chamber. For SPAAC, 10 μM DBCO-Cy3 was applied for 1 h. Cells were washed three times with ice-cold PBS, and the coverslips were mounted in mounting medium (Dako).

4.9. Spectral Fluorescence Lifetime Imaging Microscopy

Lifetime measurements were performed using an MW-FLIM detector and an SPC 150 photocounting card from Becker & Hickl (Becker & Hickl, Berlin, Germany) adapted on a laser scanning microscope LSM 710 NLO Zeiss (Zeiss SAS, Jena, Germany) and coupled with a Chameleon TiSa accordable 80 MHz pulsed laser (COHERENT, Santa Clara, USA). For more details on the acquisition and analysis procedure for FRET, please see [24]. For GFP lifetime measurements, acquisitions were performed at 860 nm and detection between 500 and 560 nm. Lifetime values were extracted from a minimum of 10 cells per experiment.

4.10. Photobleaching FRET

The samples were observed under the A1 Nikon confocal microscope with a 60X oil immersion objective. The green fluorescence (GFP) was acquired with λex = 488 nm and λ_{em} = 500–530 nm, and the red fluorescence (Cy3) was acquired with λex = 561.6 nm and λ_{em} =570–620 nm. Photobleaching of the acceptor was achieved by increasing laser power from 10% to 100%, increasing the zoom from 1 to 16, and by scanning the selected area 10 times. Images acquired before and after photobleaching were processed, according to [36], with ImageJ. Photobleaching FRET values were extracted from a minimum of 10 cells per experiment.

Author Contributions: Conceptualization, C.B. and T.L.; methodology, A.K., C.S., C.T., V.R., S.H., M.G.A., T.L. and C.B.; writing—original draft preparation, A.K., C.S., T.L. and C.B.; writing—review and editing, A.K., C.S., T.L. and C.B.; funding acquisition, T.L. and C.B. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the French government through the Programme Investissement d'Avenir (I-SITE ULNE/ANR-16-IDEX-0004 ULNE) managed by the Agence Nationale de la Recherche.

Acknowledgments: The different four expression vectors were synthesized by e-Zyvec (www.e-zyvec.com). We are indebted to "UMS 2014-US 41—Plateformes Lilloises en Biologie & Santé" for providing the technical environment conducive to achieving this work. David Vocadlo is acknowledged for helpful discussion.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Varki, A.; Cummings, R.D.; Esko, J.D.; Stanley, P.; Hart, G.W.; Aebi, M.; Darvill, A.G.; Kinoshita, T.; Packer, N.H.; Prestegard, J.H.; et al. (Eds.) *Essentials of Glycobiology*, 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2015.
2. Silsirivanit, A. Glycosylation markers in cancer. *Adv. Clin. Chem.* **2019**, *89*, 189–213. [CrossRef] [PubMed]
3. Haga, Y.; Ishii, K.; Hibino, K.; Sako, Y.; Ito, Y.; Taniguchi, N.; Suzuki, T. Visualizing specific protein glycoforms by transmembrane fluorescence resonance energy transfer. *Nat. Commun.* **2012**, *3*, 1–7. [CrossRef] [PubMed]
4. Yang, X.; Qian, K. Protein O-GlcNAcylation: Emerging mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 452–465. [CrossRef] [PubMed]
5. King, D.T.; Males, A.; Davies, G.J.; Vocadlo, D.J. Molecular mechanisms regulating O-linked *N*-acetylglucosamine (O-GlcNAc)-processing enzymes. *Curr. Opin. Chem. Biol.* **2019**, *53*, 131–144. [CrossRef]
6. Olivier-Van Stichelen, S.; Dehennaut, V.; Buzy, A.; Zachayus, J.-L.; Guinez, C.; Mir, A.-M.; El Yazidi-Belkoura, I.; Copin, M.-C.; Boureme, D.; Loyaux, D.; et al. O-GlcNAcylation Stabilizes β-catenin Through Direct Competition With Phosphorylation at Threonine 41. *FASEB J.* **2014**, *28*, 3325–3338. [CrossRef]
7. Dube, D.H.; Bertozzi, C.R. Metabolic oligosaccharide engineering as a tool for glycobiology. *Curr. Opin. Chem. Biol.* **2003**, *7*, 616–625. [CrossRef]
8. Gross, H.J.; Brossmer, R. Enzymatic introduction of a fluorescent sialic acid into oligosaccharide chains of glycoproteins. *Eur. J. Biochem.* **1988**, *177*, 583–589. [CrossRef]
9. Wratil, P.R.; Horstkorte, R.; Reutter, W. Metabolic Glycoengineering with N-Acyl Side Chain Modified Mannosamines. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 9482–9512. [CrossRef]
10. Gilormini, P.A.; Lion, C.; Vicogne, D.; Guérardel, Y.; Foulquier, F.; Biot, C. Chemical glycomics enrichment: Imaging the recycling of sialic acid in living cells. *J. Inherit. Metab. Dis.* **2018**, *41*, 515–523. [CrossRef]
11. Belardi, B.; de la Zerda, A.; Spiciarich, D.R.; Maund, S.L.; Peehl, D.M.; Bertozzi, C.R. Imaging the Glycosylation State of Cell Surface Glycoproteins by Two-Photon Fluorescence Lifetime Imaging Microscopy. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 14045–14049. [CrossRef]
12. Lin, W.; Du, Y.; Zhu, Y.; Chen, X. A cis-membrane FRET-based method for protein-specific imaging of cell-surface glycans. *J. Am. Chem. Soc.* **2014**, *136*, 679–687. [CrossRef] [PubMed]
13. Lin, W.; Gao, L.; Chen, X. Protein-Specific Imaging of O-GlcNAcylation in Single Cells. *ChemBioChem* **2015**, *16*, 2571–2575. [CrossRef] [PubMed]
14. Doll, F.; Buntz, A.; Späte, A.-K.; Schart, V.F.; Timper, A.; Schrimpf, W.; Hauck, C.R.; Zumbusch, A.; Wittmann, V. Visualization of Protein-Specific Glycosylation inside Living Cells. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 2262–2266. [CrossRef] [PubMed]
15. Clevers, H.; Nusse, R. Wnt/β-catenin signaling and disease. *Cell* **2012**, *149*, 1192–1205. [CrossRef]
16. Liu, C.; Li, Y.; Semenov, M.; Han, C.; Baeg, G.-H.; Tan, Y.; Zhang, Z.; Lin, X.; He, X. Control of β-Catenin Phosphorylation/Degradation by a Dual-Kinase Mechanism. *Cell* **2002**, *108*, 837–847. [CrossRef]
17. Boyce, M.; Carrico, I.S.; Ganguli, A.S.; Yu, S.-H.; Hangauer, M.J.; Hubbard, S.C.; Kohler, J.J.; Bertozzi, C.R. Metabolic cross-talk allows labeling of O-linked β-N-acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway. *PNAS* **2011**, *108*, 3141–3146. [CrossRef]
18. Chen, X.; Zaro, J.L.; Shen, W.-C. Fusion protein linkers: Property, design and functionality. *Adv. Drug Deliv. Rev.* **2013**, *65*, 1357–1369. [CrossRef]
19. Sipieter, F.; Vandame, P.; Spriet, C.; Leray, A.; Vincent, P.; Trinel, D.; Bodart, J.-F.; Riquet, F.B.; Héliot, L. Chapter Five—From FRET Imaging to Practical Methodology for Kinase Activity Sensing in Living Cells. In *Progress in Molecular Biology and Translational Science*; Morris, M.C., Ed.; Fluorescence-Based Biosensors Academic Press: London, UK, 2013; Volume 113, pp. 145–216.
20. Kolb, H.C.; Finn, M.G.; Sharpless, K.B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 2004–2021. [CrossRef]
21. Jewett, J.C.; Bertozzi, C.R. Cu-free click cycloaddition reactions in chemical biology. *Chem. Soc. Rev.* **2010**, *39*, 1272–1279. [CrossRef]

22. Kuzmin, A.; Poloukhtine, A.; Wolfert, M.A.; Popik, V.V. Surface Functionalization Using Catalyst-Free Azide−Alkyne Cycloaddition. *Bioconjug. Chem.* **2010**, *21*, 2076–2085. [CrossRef]
23. Han, S.S.; Lee, D.E.; Shim, H.E.; Lee, S.; Jung, T.; Oh, J.H.; Lee, H.A.; Moon, S.H.; Jeon, J.; Kim, K.; et al. Physiological Effects of Ac4ManNAz and Optimization of Metabolic Labeling for Cell Tracking. *Theranostics* **2017**, *7*, 1164–1176. [CrossRef] [PubMed]
24. Terryn, C.; Paës, G.; Spriet, C. FRET-SLiM on native autofluorescence: A fast and reliable method to study interactions between fluorescent probes and lignin in plant cell wall. *Plant Methods* **2018**, *14*, 74. [CrossRef] [PubMed]
25. Batt, A.R.; Zaro, B.W.; Navarro, M.X.; Pratt, M.R. Metabolic Chemical Reporters of Glycans Exhibit Cell-Type-Selective Metabolism and Glycoprotein Labeling. *Chembiochem* **2017**, *18*, 1177–1182. [CrossRef] [PubMed]
26. Yuzwa, S.A.; Macauley, M.S.; Heinonen, J.E.; Shan, X.; Dennis, R.J.; He, Y.; Whitworth, G.E.; Stubbs, K.A.; McEachern, E.J.; Davies, G.J.; et al. A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo. *Nat. Chem. Biol.* **2008**, *4*, 483–490. [CrossRef]
27. Haltiwanger, R.S.; Holt, G.D.; Hart, G.W. Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-acetylglucosamine:peptide beta-N-acetylglucosaminyltransferase. *J. Biol. Chem.* **1990**, *265*, 2563–2568.
28. Dong, D.L.; Hart, G.W. Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol. *J. Biol. Chem.* **1994**, *269*, 19321–19330.
29. Gloster, T.M.; Zandberg, W.F.; Heinonen, J.E.; Shen, D.L.; Deng, L.; Vocadlo, D.J. Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells. *Nat. Chem. Biol.* **2011**, *7*, 174–181. [CrossRef]
30. Decourcelle, A.; Loison, I.; Baldini, S.; Leprince, D.; Dehennaut, V. Evidence of a compensatory regulation of colonic O-GlcNAc transferase and O-GlcNAcase expression in response to disruption of O-GlcNAc homeostasis. *Biochem. Biophys. Res. Commun.* **2020**, *521*, 125–130. [CrossRef]
31. Cheung, W.D.; Hart, G.W. AMP-activated Protein Kinase and p38 MAPK Activate O-GlcNAcylation of Neuronal Proteins during Glucose Deprivation. *J. Biol. Chem.* **2008**, *283*, 13009–13020. [CrossRef]
32. Taylor, R.P.; Geisler, T.S.; Chambers, J.H.; McClain, D.A. Up-regulation of O-GlcNAc Transferase with Glucose Deprivation in HepG2 Cells Is Mediated by Decreased Hexosamine Pathway Flux. *J. Biol. Chem.* **2009**, *284*, 3425–3432. [CrossRef]
33. Venter, J.C.; Adams, M.D.; Myers, E.W.; Li, P.W.; Mural, R.J.; Sutton, G.G.; Smith, H.O.; Yandell, M.; Evans, C.A.; Holt, R.A.; et al. The sequence of the human genome. *Science* **2001**, *291*, 1304–1351. [CrossRef] [PubMed]
34. Biwi, J.; Biot, C.; Guerardel, Y.; Vercoutter-Edouart, A.-S.; Lefebvre, T. The many ways by which O-GlcNAcylation may orchestrate the diversity of complex glycosylations. *Molecules* **2018**, *23*, 2558. [CrossRef] [PubMed]
35. Leturcq, M.; Mortuaire, M.; Hardivillé, S.; Schulz, C.; Lefebvre, T.; Vercoutter-Edouart, A.S. O-GlcNAc transferase associates with the MCM2–7 complex and its silencing destabilizes MCM–MCM interactions. *Cell. Mol. Life Sci.* **2018**, *75*, 4321–4339. [CrossRef] [PubMed]
36. Pang, Y.; Liu, J.; Qi, Y.; Li, X.; Chilkoti, A. A Modular Method for the High-Yield Synthesis of Site-Specific Protein–Polymer Therapeutics. *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 10296–10300. [CrossRef]

Sample Availability: Samples of the chemical reporters are available from the authors.

Les travaux exposés dans cette thèse s'articulent principalement autour de développements méthodologiques pour l'étude du métabolisme des acides sialiques chez divers organismes. Prenant origine dans les travaux précédemment effectués au laboratoire sur cellules humaines, nous avons d'abord voulu durant ce projet étendre la stratégie du rapporteur chimique exploitant les rapporteurs libres, non protégés, à des souches bactériennes. En soi, cet objectif représentait un challenge technologique non négligeable.

En effet, bien que de nombreuses méthodes bioorthogonales aient été développées dans les bactéries pour suivre les protéines avec des acides aminés non canoniques, les efforts visant à appliquer la MOE pour détecter des glycanes chez les procaryotes demeurent plutôt limités en comparaison.

Il existe plusieurs raisons à ce manque : les stratégies d'édition génomique modifiable ont historiquement été développées en premier lieu dans des modèles humains. De plus, l'application du marquage métabolique aux bactéries est techniquement exigeante et demande une conceptualisation et une mise en œuvre minutieuses, étant donné que le métabolisme des procaryotes est complexe. L'absorption cellulaire des monosaccharides est variable et les flux métaboliques ne peuvent pas être ajustés avec des sucres peracétylés comme c'est le cas dans les cellules humaines, car les bactéries présentent généralement des niveaux faibles d'estérases non spécifiques et sont donc incapables de générer le sucre libre à partir d'un précurseur peracétylé. Des techniques à haute résolution sont également nécessaires pour générer de nouvelles connaissances en raison de la petite taille des bactéries (quelques micromètres).

Associée à la MOE, la chimie bioorthogonale a été utilisée avec succès pour marquer certains glycanes constitutifs des parois cellulaires bactériennes (par exemple, les lipopolysaccharides, les peptidoglycanes, les glycolipides). Cependant, les études de marquage métabolique ciblant les capsules bactériennes sont très rares et n'avaient été

signalées que chez des espèces non pathogènes. Or, de nombreuses bactéries pathogènes ont évolué la capacité de produire des capsules de polysaccharides qui les entourent, formant une couche protectrice. Ces structures sont ainsi des facteurs de virulence, car elles imitent souvent des glycanes de l'hôte et agissent comme un bouclier vis à vis du système immunitaire. Empêchant la destruction des bactéries par les macrophages, ces facteurs de virulence jouent donc un rôle critique dans les infections.

Le premier objectif de cette thèse, dans le cadre du projet ANR Neuraprobe AA-18-CE07-0042-01, était de développer une méthode de détection de la capsule K1, présente chez certaines souches pathogènes d'*E. coli* responsables de méningites et de septicémies. Nous avons atteint ce but, avec une méthode exploitant des rapporteurs chimiques faciles d'accès, une ligation par CuAAC, et une observation double en bioimagerie et en lecteur de microplaque. C'est la première fois qu'une méthode de marquage d'un tel antigène K facteur de virulence est publié. Cet outil offre des perspectives en aval pour une meilleure compréhension du mécanisme de formation de la capsule K1 et sur sa dynamique d'évolution au cours du temps dans des conditions physiologiques différentes (stress, interactions hôtes-pathogènes, formation de biofilm). Un portage de la méthode en microplaque pour le screening haut-débit pourrait également être intéressant, car cela pourrait permettre la découverte de nouvelles familles d'antibiotiques ciblant non pas la destruction de la bactérie, mais la diminution de sa virulence par modulation de sa capsule. Cette stratégie paraît prometteuse pour découvrir de nouveaux traitements évitant l'apparition de résistances. Elle pourrait également permettre de mettre en lumière des voies métaboliques méconnues, en comparant l'incorporation de différents types de rapporteurs. Dans cette optique, j'ai effectué des essais en synthétisant un analogue de ManNAc modifié en position 6, le ManNAc-6-Az, et en comparant son incorporation au sein de cellules mammifères ou de bactéries E. coli K1 pour tenter de tirer profit d'une différence dans la biosynthèse des acides sialiques au sein de ces modèles, impliquant ou non une phosphorylation de cette position.

De manière à faciliter l'obtention d'images à très hautes résolutions, mais aussi de permettre une quantification du marquage (un point faible des méthodes existantes), un deuxième axe de ma thèse portait sur le développement de sondes organométalliques versatiles. Une famille de complexes d'iridium(III), portant des ligands polypyridines et phenanthroline, a ainsi été développée et caractérisée d'un point de vue photophysique. Ces sondes s'avèrent très intéressantes, car elles peuvent être détectées non seulement en bioimagerie en temps résolu (permettant d'aller vers des méthodes de marquages multiples tout en s'assurant à 100% de la spécificité de signal, ou encore d'aller vers des techniques dynamiques et super-résolutives exploitant le FRET et la chimiométrie), mais aussi une détection par fluorescence des rayons X. J'ai ainsi eu la chance pendant ma thèse d'effectuer 3 mobilités au synchrotron européen ESRF sur la ligne ID16A de nanoimagerie par rayons X pour des expériences de SR-XRF. Cette méthode peut atteindre une résolution spatiale xy de 25-30 nm (environ 50x supérieure à la microscopie confocale classique en UV-Vis), mais également de générer des cartes élémentaires, avec une quantité absolue d'iridium en ng/mm^2 pour chaque pixel. Les expériences de SR-XRF nous ont notamment permis de relier des intensités de fluorescence arbitraire à des valeurs quantitatives très précises. En effet, j'ai pu observer préalablement les cellules au microscope confocal avant de les apporter à l'ESRF où nous avons été en mesure de les identifier et d'observer les mêmes régions sous faisceau synchrotron, ce qui a représenté un véritable challenge technique et logistique. Ces sondes ont été testées et appliquées sur cellules humaines dans le cadre de cette thèse, car il n'était pas envisageable de mener cet axe et le marquage de la capsule K1 l'un après l'autre, ils devaient être menés de front. Ces résultats très positifs sur cellules humaines ont pour but dans un deuxième temps d'être transposés en bactéries.

En effet, les mécanismes par lesquels les bactéries, enfermées dans une capsule polysaccharidique, parviennent encore à percevoir leur environnement restent à élucider. En particulier, les processus dynamiques par lesquels les appendices bactériens tels que les flagelles et les pili, essentiels à la motilité et à l'adhésion cellulaire, peuvent être synthétisés à l'intérieur, interagir avec et traverser la capsule, sont d'une importance

primordiale. Cette investigation des interactions capsule-pili, capsule-hôte, ou pili-hôte en présence de capsule pourrait s'étendre également à la compréhension de la perception de signaux chimiques ou mécaniques à la surface bactérienne malgré la présence de la capsule. En effet, cette capacité à percevoir leur environnement en surface joue un rôle critique dans le déclenchement de divers comportements, notamment la formation de biofilms par la production de pili chaperone-usher (CUP) (Busch et Waksman - Phil Trans R Soc - 2012 – 367 – 1112-1122). La microscopie en temps de vie avec les sondes iridium marquant la capsule, exploitée en FRET en combinaison avec un marquage des pili par protéine fluorescente, serait d'une aide cruciale dans cette optique.

La plupart des études en MOE a recours à des rapporteurs chimiques azotures ou alcyne, deux groupements souvent considérés comme inertes et n'impactant pas l'incorporation des molécules sur lesquelles ils sont introduits. En parallèle du développement des méthodologies de marquage de la capsule et des sondes iridium, j'ai également appliqué les rapporteurs chimiques dérivés de ManNAc et de Neu5Ac pour un autre projet, en collaboration avec l'équipe du Dr Anne-Sophie Vercoutter-Edouart, au cours duquel nous avons étudié l'impact de ces groupements sur la métabolisation des rapporteurs chimiques, au sein de différentes lignées cellulaires du cancer du colon, ainsi que dans les lignées modèles HeLa et HEK293. Nous avons également comparé l'incorporation de rapporteurs non protégés avec celle de l'Ac_4ManNAz, qui est le rapporteur chimique le plus couramment utilisé en MOE à l'heure actuelle. En observant à la fois le marquage à la surface des cellules par cytométrie de flux, le marquage des différents types de glycoprotéines par Western-blot, et le marquage des cellules en croissance par microscopie confocale, nous avons pu mettre en exergue la grande variabilité dans le profil d'incorporation selon la lignée cellulaire, et l'impact du groupement alcyne et azoture sur l'incorporation. Cette étude a montré que le ManNAz était le rapporteur chimique au caractère le plus universel et le plus fiable pour le marquage des sialoglycoconjugués.

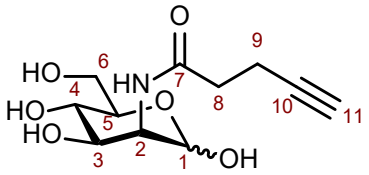
Nous avons pu comparer les profils de marquage obtenus après incorporation de ManNAz ou d'Ac_4ManNAz, et avons observé un marquage plus diffus avec le rapporteur

peracétylé, ce qui pourrait s'expliquer par le fait que ces rapporteurs sont plus affins pour les membranes des cellules et sont sujets à des réactions parasites contrairement à leurs homologues non protégés. Ceci est en adéquation avec les résultats issus des expériences de SR-XRF où nous avons également comparé ces deux rapporteurs. Ces informations, en plus de m'avoir été salutaires pour l'aboutissement des différents projets initiés lors de ma thèse, seront également précieuses pour la communauté mondiale grandissante de chémoglycobiologistes.

En marge de mes projets principaux, j'ai également participé au transfert des méthodologies de la chimie bioorthogonale auprès de différentes équipes de biologistes, et à leurs applications concrètes. Ma participation à ces différents projets a principalement consisté à synthétiser différents outils chimiques, et à apporter mon expertise pour le marquage métabolique et l'imagerie. J'ai ainsi généré une série d'acides sialiques CMP-activés à partir des analogues d'acides sialiques synthétisés pendant ma thèse (SiaNAl, SiaNAz) et à partir d'acides sialiques commerciaux (Neu5Ac, Neu5Gc, Kdn), qui ont servi à caractériser des sialyltransférases de salmonidés. Cette approche est un bon exemple d'applications purement biochimiques de la MOE, qui sont souvent moins connues que les applications en bio-imagerie, alors même qu'elles reposent souvent sur des bases quantitatives plus robustes statistiquement. J'ai également participé à l'étude d'un autre type de glycosylation, la O-GlcNAcylation, en synthétisant le dérivé Ac_4ManNAz qui servit de contrôle, et en étudiant son impact ainsi que celui de l'Ac_4GalNAz sur la viabilité cellulaire. L'utilisation de cet outil en combinaison avec une version fluorescente chimérique de la B-caténine a permis de mettre au point une méthode de marquage spécifique de la O-GlcNAcylation de cette protéine, démontrant la puissance de la stratégie du rapporteur chimique pour repousser les limites du marquage des biomolécules.

SYNTHÈSE ORGANIQUE DES DÉRIVÉS

RAPPORTEURS CHIMIQUES MONOSACCHARIDES



ManNAl (*N*-4-pentynoyl-D-mannosamine)

L'acide pent-4-ynoïque (502 mg, 5.1 mmol, 1 eq) et le N-hydroxysuccinimide (NHS, 674 mg, 5.8 mmol, 1,14 eq) sont dissous dans 20 mL de dichlorométhane (DCM) anhydre dans un ballon de 50 mL. Le ballon est placé 20 mn dans un bain de glace, puis le 1-éthyl-3-(-3-diméthylaminopropyl)carbodiimide (EDC, 1.96 g, 10.2 mmol, 2 eq) est ajouté par petites portions au moyen d'une spatule. Le ballon est fermé par un septum et l'air chassé par du diazote. Après 20 mn le bain de glace est retiré et la réaction est maintenue sous agitation à température ambiante 3 h supplémentaire. La consommation totale de l'acide pent-4-ynoïque est déterminée par CCM (éther de pétrole / acétate d'éthyle 7:3 ; révélateur acide phosphomolybdique (PMA)). La solution est lavée par une solution de $KHSO_4$ 2.8% (3 x 20 mL) puis la phase organique est séchée sur sulfate de sodium, filtrée et évaporée sous vide (40°C, 451 mbar). Le produit est obtenu sous la forme d'une poudre blanche (875.3 mg, 4.48 mmol, 88%). Il est utilisé dans la réaction suivante sans purification supplémentaire.

Le chlorhydrate de D-mannosamine (759mg, 3.52 mmol, 1 eq) et l'ester de succinimidyl obtenu à l'étape précédente sont dissous dans 20 mL de diméthylformamide (DMF) anhydre dans un ballon de 50 mL. Le ballon est fermé par un septum et l'air est chassé par du diazote. La triéthylamine (1.4 mL, 10.77 mmol, 3 eq) est ajoutée goutte à goutte et le mélange est maintenu sous agitation à température ambiante pendant 16h30. La fin de la réaction est déterminée par CCM (DCM / acétate d'éthyle / méthanol 45:45:10 ;

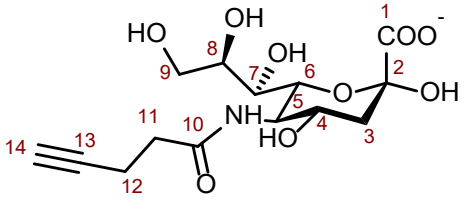
révélation PMA) en observant la consommation totale de la D-mannosamine. Le DMF est évaporé à l'évaporateur rotatif (40°C, 1 mbar) donnant une huile brune laquelle est resolubilisée dans du méthanol et mélangée à de la Celite. Le méthanol est alors évaporé et la poudre obtenue est tassée dans une cartouche pour dépôt solide. Le produit est purifié par chromatographie flash (Si, 40 g, 50 µM ; éluant DCM / acétate d'éthyle / méthanol 45:45:10). Les fractions contenant le produit sont identifiées par CCM, rassemblées et concentrées sous vide. Le produit est obtenu sous la forme d'un solide cristallin blanc (830 mg, 3.34 mmol, 95%).

RMN ^{1}H α (300 MHz, D_2O): δ = 5.00 (d, J=1.4, 1H, H_1), 4.23 (dd, J=4.7, 1.4, 1H, H_2), 3.93 (dd, J=9.8, 4.7, 1H, H_3), 3.80 – 3.62 (m, 3H, H_5 & H_6), 3.50 (t, J=9.6, 1H, H_4), 2.58 – 2.32 (m, 4H, H_8 & H_9), 2.26 (t, J=2.3, 1H, H_{11}).

RMN ^{13}C α (75 MHz, D_2O): δ = 175.26 (C7), 93.17 (C1), 83.43 (C10), 71.94 (C5), 70.14 (C11), 68.73 (C3), 66.72 (C4), 60.36 (C6), 53.15 (C2), 34.10 (C8), 14.40 (C9).

RMN ^{1}H β (300 MHz, D2O): δ = 4.90 (d, J=1.6, 1H, H_1), 4.35 (dd, J=4.4, 1.4, 1H, H_2), 3.82 – 3.61 (m, 3H, H_3 & H_6), 3.40 (t, J=9.8, 1H, H_4), 3.29 (ddd, J=9.9, 4.8, 2.3, 1H, H_5), 2.53 – 2.31 (m, 4H, H_7 & H_8), 2.26 (t, J=2.3, 1H, H_{11}).

RMN ^{13}C β (75 MHz, D2O): δ = 175.98 (C7), 92.84 (C1), 83.85 (C10), 76.28 (C5), 71.96 (C3), 69.99 (C11), 66.45 (C4), 60.36 (C6), 53.98 (C2), 34.25 (C8), 14.26 (C9).



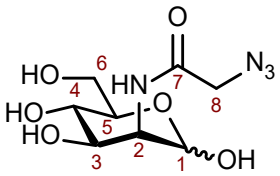
SiaNAl (N-4-pentynoylneuraminic acid)

Le ManNAl (60 mg, 0,23 mmol, 1 eq), le pyruvate de sodium (2 eq) et 22,5U de Neu5Ac aldolase de E. coli K12 (Sigma Aldrich EC 4.1.3.3) sont dissous dans 1,5mL de PBS dans un tube à méthanolyse. Le tube est refermé et placé sous agitation (50 rpm) à 37°C. La réaction est suivie par CCM (propan-1-ol / ammoniaque 25% / eau 6:1:2,5 ; révélateur résorcinol), après 48h la formation du produit et la consommation du ManNAl sont observées. La réaction est arrêtée par ajout de 40 mL d'eau distillée. La solution est déposée sur une colonne de résine échangeuse d'anions (BioRad AG1X8). La colonne est lavée par 5 volumes de colonne (CV) d'eau ultrapure, puis l'élution est effectuée par 5 CV de NH_4HCO_3 0.05M puis 5 CV de NH_4HCO_3 0.2M (débit 2mL/mn). Les fractions contenant le produit sont identifiées par CCM (voir conditions ci- dessus), rassemblées et concentrées sous vide (30°C, 20 mbar). Le solide blanc obtenu est solubilisé dans un minimum d'eau ultrapure et déposé sur une colonne de gel filtration (P2), les fractions contenant le produit sont identifiées et concentrées comme précédemment. Le solide blanc obtenu est solubilisé dans un minimum d'eau ultrapure et lyophilisé. Le SiaNAl est obtenu sous la forme d'un solide cristallin translucide (73 mg, 91%).

RMN ^{1}H (600 MHz, D_2O): δ = 3.96 – 3.86 (m, 2H, H_4 & H_9), 3.83 (t, J=10.1, 1H, H_5), 3.72 (dd, J=11.9, 2.7, 1H, $H_{9'}$), 3.64 (ddd, J=9.2, 6.7, 2.7, 1H, H_8), 3.51 (d, J=9.2, 1H, H_7), 3.49 – 3.43 (m, 1H, $H_{9'}$), 2.47 – 2.34 (m, 4H, H_{11} & H_{12}), 2.28 (d, J=2.0, 1H, H_{14}), 2.10 (dd, J=13.0, 4.9, 1H, H_{3eq}), 1.70 (dd, J=12.7, 11.7, 1H, H_{3ax}).

RMN ^{13}C (75 MHz, D_2O): δ 175.43 (C1), 172.91 (C10), 95.16 (C2), 83.54 (C13), 70.40 (C14), 70.37 (C8), 70.23 (C6), 68.29 (C7), 66.45 (C4), 63.21 (C9), 52.04 (C5), 38.79 (C3), 34.67 (C11), 14.53 (C12).

m/z: calculated: [M]+ = 346.312 ; measured: 346.017

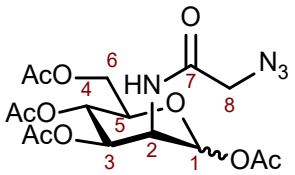


ManNAz (*N*-azidoacetyl-D-mannosamine)

L'acide 2-azidoacétique () est dissous avec le *N,N'*-diisopropylcarbodiimide (DIC, 175mg, 1.392mmol, 1.2eq), l'hydroxybenzotriazole (HOBt, 195mg, 1.276mmol, 1.1eq) et la *N,N*-diisopropylethylamine (DIPEA, 141mg, 1.392mmol, 1.2eq) dans le DMF (20ml). Le chlorhydrate de D-mannosamine (250mg, 1.16mmol, 1eq) est ajouté et la réaction est agitée à t.a. pendant 19h sous argon. La consommation totale de la mannosamine est validée par CCM (DCM / méthanol ; 9 : 1) puis le solvant est évaporé sous pression réduite. Le brut réactionnel est purifié par chromatographie flash (Si, 50 µm, 40g, dépôt solide, DCM / méthanol ; 20 : 1) et les fractions contenant le produit sont rassemblées et concentrées sous pression réduite. Le ManNAz est obtenu sous la forme d'une poudre blanche (175 mg, 0.67 mmol, 57%).

RMN ^{1}H (300 MHz, D_2O): δ = 5.10 (d, J=1.0 Hz, 1H, H_1 α), 5.01 (d, J=1.4 Hz, 1H, H_1 β), 4.46 (dd, J=4.0 Hz, 1.4Hz, 1H, H_2 β), 4.33 (dd, J=4.4 Hz, 1.0 Hz, 1H, H_2 α), 4.10 – 3.98 (m, 5H, H_3 α + H_8 αβ), 3.88 – 3.71 (m, 5H, H_5 α + H_3 β + H_6 αβ), 3.56 (t, J=9.5Hz, 1H, H_4 α), 3.46 (t, J=9.8Hz, 1H, H_4 β), 3.38 (ddd, J=9.8Hz, 4.7Hz, 2.1Hz, 1H, H_5 β).

RMN ^{13}C (75 MHz, D_2O): δ = 171.87 (s, C7 α), 170.97 (s, C7 β), 92.89 (s, C1 α), 92.83 (s, C1 β), 76.40 (s, C5 β), 72.03 (s, C3 β), 71.97 (s, C5 α), 68.82 (s, C3 α), 66.76 (s, C4 α), 66.52 (s, C4 β), 60.41 (s, C6 β), 60.39 (s, C6 α), 54.28 (s, C2 β), 53.37 (s, C2 α), 51.72 (s, C8 β), 51.65 (s, C8 α).

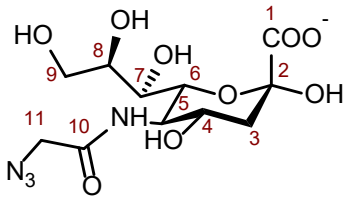


Ac_4ManNAz (1,3,4,6-tetra-*O*-acetyl-*N*-azidoacetylmannosamine)

Le ManNAz (80mg, 0.31 mmol, 1 eq) et la *N*,*N*-dimethylaminopyridine (DMAP, quantité catalytique) sont dissous dans la pyridine (5 mL). L'anhydride acétique (189 mg, 1.85 mmol, 6 eq) est ajouté goutte à goutte et la reaction est agitée à t.a. pendant 16h sous argon. Le brut réactionnel est concentré sous pression réduite et purifié par chromatographie flash (Si, 50 µm, 40 g, cyclohexane / acétate d'éthyle ; 2 : 1). Les fractions contenant le produit sont rassemblées et concentrées sous pression réduite. L'Ac_4ManNAz est obtenu sous forme d'une poudre blanche (117mg, 0.27mmol, 88%).

RMN ^{1}H (300 MHz, $CDCl_3$) δ = 6.64 (d, *J*=8.9Hz, 1H, H1), 6.56 (d, *J*=9.2Hz, 1H, H1'), 6.05 (d, *J*=1.9Hz, 1H, NH'), 5.89 (d, *J*=1.6Hz, 1H, NH), 5.35 (dd, *J*=10.1Hz, 4.2Hz, 1H, H3'), 5.22 (t, *J*=10.1Hz, 1H, H4'), 5.17 (t, *J*=9.8Hz, 1H, H4), 5.06 (dd, *J*=9.8Hz, 3.8Hz, 1H, H3), 4.73 (ddd, *J*=8.9Hz, 3.8Hz, 1.6Hz, 1H, H2), 4.62 (ddd, *J*=9.3, 4.2, 1.9, 1H, H2'), 4.30 – 4.00 (m, 5H, H5' + H6 + H6' + CH_2-N_3 + CH_2-N_3'), 3.82 (ddd, *J*=9.8Hz, 4.5Hz, 2.5Hz, 1H, H5), 2.24 – 1.93 (m, 12H, OAc + OAc').

RMN ^{13}C (75 MHz, $CDCl_3$) δ = 170.67 – 166.55 (m, CO OAc + C1 + C1'), 91.31 (s, C7), 90.26 (s, C7'), 73.40 (s, C5), 71.45 (s, C3), 70.27 (s, C5'), 68.86 (s, C3'), 65.10 (s, C4'), 64.93 (s, C4), 61.75 (s, C6'), 61.67 (s, C6), 52.62 (s, C8), 52.45 (s, C8'), 49.72 (s, C2), 49.26 (s, C2'), 20.90 – 20.56 (m, 4 x CH3 OAc).

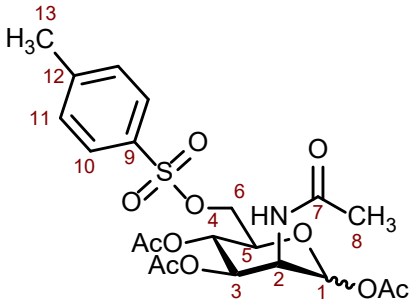


SiaNAz (*N*-2-azidoacetylneuraminic acid)

Le ManNAz (60 mg, 0.228 mmol, 1 eq) est dissous dans 600 µL de PBS (pH 7.6) avec le pyruvate de sodium (92 mg, 0.456 mmol, 2 eq) dans un tube à méthanolyse fermé de 5 mL. 5 unités de Neu5Ac aldolase (Sigma-Aldrich, EC 4.1.3.3) sont ajoutés et la réaction est agitée à 37 °C pendant 16 h. La formation du produit est suivie par CCM (1-propanol, NH_3, H_2O 6 : 1 : 2,5, révélateur résorcinol). Le brut réactionnel est purifié par chromatographie échangeuse d'anions (BioRad, AG1X8; H_2O 5CV, NH_4HCO_3 0.05 M 5CV, NH_4HCO_3 0.2 M 5CV). Les fractions contenant le produit sont identifiées par CCM, rassemblées et concentrées sous pression réduite. La poudre blanche obtenue est alors purifiée par chromatographie d'exclusion (pour retirer l'excédent de sels) avec de l'eau ultrapure comme éluant. Les fractions contenant le produit sont rassemblées et concentrées sous pression réduite. Le SiaNAz est obtenu sous la forme d'une poudre blanche (20.6 mg, 0.059 mmol, 26%).

RMN ^{1}H (300 MHz, D_2O) δ 4.09–3.88 (m, 5H, H5 + H4 + H3 + H2), 3.78 (dd, J = 11.2 Hz, 2.2 Hz, 1H, H8cis), 3.65 (ddd, J = 8.5 Hz, 4.1 Hz, 1.2 Hz, 1H, H7), 3.55 (dd, J = 11.2 Hz, 5.8 Hz, 1H, H8trans), 3.45 (dd, J = 9.2 Hz, 0.9 Hz, 1H, H6), 2.18 (dd, J = 12.5 Hz, 4.8 Hz,1H, H1eq), 1,75 (dd, J = 12.4 Hz, 11.8 Hz, 1H, H1ax).

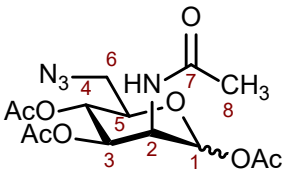
RMN ^{13}C (75 MHz, D2O) δ 70.8 (CH, C5), 69.82 (CH, C7), 68.43 (CH, C6), 67.01 (CH, C2), 63.10(CH2, C8), 52.21(CH, C3), 51.9 (CH2, C4), 39.32 (CH2, C1).



6-OTs-Ac3ManNAc (6-*O*-tosyl-1,3,4-tri-*O*-acetyl-*N*-2-acetyl-D-mannosamine)

Le ManNAc (3,0 g, 13,6 mmol, 1 eq) est dissout dans 40 mL de pyridine dans un ballon de 250 mL fermé par un septum et sous agitation. L'air est chassé par de l'argon. Le chlorure de tosyle (6,46 g, 33,9 mmol, 3 eq) repris dans 20 mL de pyridine est ajouté goutte à goutte sur 1h. La réaction est maintenue à 0°C pendant 5h. Par CCM (dichlorométhane / méthanol 9:1) on observe la consommation totale de la ManNAc. 20 mL d'anhydride acétique (Ac_2O) sont ajoutés, puis le bain de glace est retiré. La solution est maintenue sous agitation à température ambiante pendant 14h. 75 mL d'acétate d'éthyle sont ajoutés et la solution est versée dans une ampoule à décanter. La phase organique est lavée par 2 x 50 mL d'acide chlorhydrique 1M puis 2 x 50 mL d'une solution de bicarbonate de sodium saturée. La phase organique est séchée sur sulfate de sodium puis évaporée (40°C, 105-20 mbar). La pyridine restante est retirée par une série de co-évaporations au toluène. Le résidu brun obtenu est purifié par chromatographie flash (colonne de silice prépackée 80 g, 15-40 µm, éluant acétate d'éthyle / cyclohexane 1:1 → 4:1). La poudre beige obtenue (2,4 g, 36%) contient le produit sous forme d'un mélange d'anomères (1,15 : 1), l'anomère majoritaire a été isolé pour caractérisation.

RMN ^{1}H (300 MHz, $CDCl_3$) : δ 7.81 (d, J = 8.3 Hz, 2H, H10), 7.38 (d, J = 8.1 Hz, 2H, H11), 6.06 (d, J = 9.3 Hz, 1H, NH), 6.03 (d, J = 1.7 Hz, 1H, H1), 5.37 – 5.23 (m, 2H, H3 + H4), 4.64 (ddd, J = 9.3, 4.0, 1.9 Hz, 1H, H2), 4.31 (dd, J = 11.3, 1.9 Hz, 1H, H5), 4.19 – 4.02 (m, 2H, H6), 2.46 (s, 3H; H13), 2.17, 2.07, 2.03, 2.02 (OAc + H8).

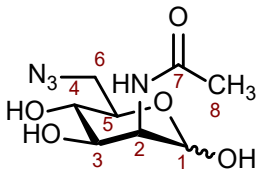


Ac_3ManNAc-6-Az (6-desoxy-6-azido-1,3,4-tri-*O*-acetyl-*N*-2-acetyl-D-mannosamine)

Le 6-*O*-tosyl-Ac_3ManNAc (1,0 g, 2,0 mmol, 1 eq) est dissous dans le toluène (25 mL) dans un ballon de 100 mL. Le Bu_4NN_3 est ajouté, le ballon est fermé et l'air est chassé par de l'argon. Un septum est placé en haut d'un réfrigérant à eau lequel est placé sur le ballon puis l'air est de nouveau chassé par de l'argon. Après 6h de réaction à reflux on constate la consommation du réactif par CCM (éther de pétrole / acétate d'éthyle 1 : 3). 20 mL d'acétate d'éthyle sont ajoutés et cette phase est lavée par 2 x 20 mL d'eau distillée. La phase organique est séchée sur sulfate de sodium puis évaporée fournissant une huile brune, laquelle est purifiée par chromatographie sur gel de silice (Si, 40 g, éluant : acétate d'éthyle / cyclohexane 1 : 1 puis 1 : 0). On obtient le produit sous la forme d'une poudre blanche (404 mg, 54%), correspondant à un mélange d'anomères (4 : 1). L'anomère majoritaire est isolé pour caractérisation.

RMN ^{1}H (300 MHz, $CDCl_3$) : δ 6.53 (d, J = 8.8 Hz, 1H, H1), 6.07 (s, 1H, NH), 5.39 – 5.30 (m, 1H, H3), 5.22 (td, J = 10.0, 2.9 Hz, 1H, H4), 4.71 – 4.63 (m, 1H, H2), 4.05 – 3.96 (m, 1H, H5), 3.48 – 3.27 (m, J = 13.5, 4.0 Hz, 2H, H6), 2.23 – 1.98 (m, 12H, NHAc + OAc)

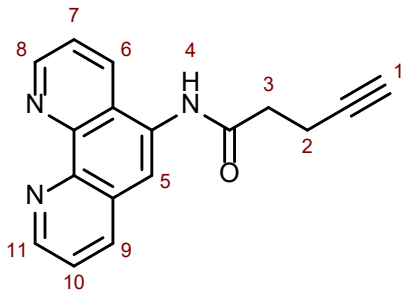
Infrarouge : 2104 cm-1 (moyen, mince, N3) ; 1747 cm-1 (fort, mince, CO esters)



ManNAc-6-Az (6-desoxy-6-azido-N-2-acetyl-D-mannosamine)

Le Ac_3ManNAc-6-Az (307 mg, 0.82 mmol) est dissous dans du méthanol anhydre. Du méthanolate de sodium (NaOMe) en solution dans le méthanol est ajouté goutte à goutte jusqu'à ce que le pH soit égal à 9. La réaction est agitée à 0°C. Le suivi est effectué par CCM. Après 2h, la réaction est neutralisée par ajout d'Amberlite (résine acide), filtrée puis évaporée sous pression réduite. L'huile obtenue est purifiée par chromatographie sur colonne de silice (Si, 40g, 15-40µm, éluant dichlorométhane / méthanol 20 : 1 puis 7 : 1). Les fractions contenant le produit sont identifiées par CCM, rassemblées et évaporées sous pression réduite. Le produit est obtenu sous la forme de cristaux blancs (156mg, 0.64 mmol, 77%).

RMN ^{1}H (300MHz, D_2O) : δ = 5.01 (d, J = 3.47 Hz, 1H, H1), 3.85 (ddd, J = 9.5, 5.9, 2.4 Hz, 1H, H5), 3.76 (dd, J = 10.7, 3.5 Hz, 1H), 3.58 (dd, J = 10.6, 8.8 Hz, 1H), 3.42 (dd, J = 13.1, 2.5 Hz, 1H, H4), 3.30 (dd, J = 13.1, 5.9 Hz, 2H, H6)

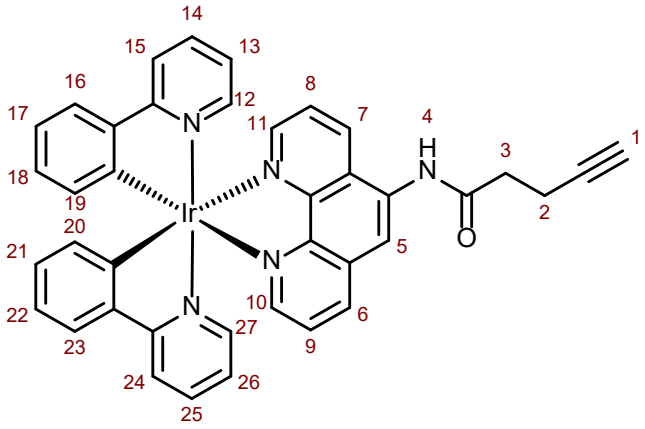


_N_-(1,10-phenanthrolin-5-yl) pent-4-ynamide

L'acide pent-4-ynoïque (194mg, 1.98 mmol, 9.7 eq), le 1,3-dicyclohexylcarbodiimide (DCC) (210mg, 1.02 mmol, 5eq) sont ajoutés à la 1,10-phenanthrolin-5-amine (40mg, 0.204mmol, 1 eq) dans un ballon de 100 mL. 60mL d'acétonitrile sont ajoutés et l'air est chassé par de l'argon. La réaction est agitée à 22°C toute la nuit. La consommation de l'acide est suivie par CCM (DCM / méthanol ; 1 : 20 + 0.5% DIPEA). Le brut réactionnel est filtré (verre fritté porosité 3) pour retirer le précipité de dicyclohexylurée (DCU) puis évaporé sous pression réduite. La poudre jaune obtenue est purifiée par chromatographie flash (Si, 40g, 50 µm, dépôt solide, éluant DCM / méthanol 2% à 6% v/v). Le produit est obtenu sous forme d'une poudre jaune (53 mg, 1.88 mmol, 95%)

RMN ^{1}H (300 MHz, CDCl3): δ 8.95(m, 2H, H8, H11), 8.79 (s, NH, H4), 8.38 (dd, J=7.9 Hz, J=1.8 Hz, H9), 7.96(m, 2H, H5, H6), 7.46(m, 2H H7, H10), 2.65 (m, 5H, H3, H2), 2.00(t, 1H, J=2.1 Hz, H1)

RMN ^{13}C (75 MHz, CDCl3): δ 168.34 (Cq, C=O), 150.10 (CH, C11), 151.22(CH, C8), 136.41(CH, C6), 121.22(CH, C5), 132.51(CH, C9), 124.12(CH, C7), 125.12(CH, C10), 71.21 (CH, C1), 36.32 (CH2, C3), 15.07 (CH2, C2).

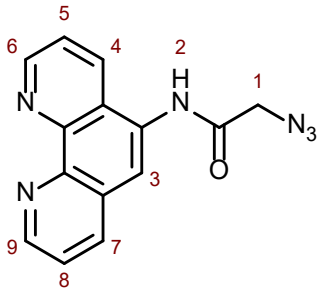


[Ir(ppy)$_2$(phen-Alk)]$^+$[PF$_6$]$^-$

Le [Ir(ppy)$_2$Cl]$_2$ (149 mg, 0.14 mmol, 1eq) et le *N*-(1,10-phenanthrolin-5-yl) pent-4-ynamide obtenu à l'étape précédente (77 mg, 0.28 mmol, 2 eq) sont dissous dans 15 ml de DCM / méthanol (2 : 1) anhydre dans un ballon de 50mL. Le mélange est porté à reflux à 45°C pendant 5h. La solution orange obtenue est refroidie à t.a. puis un excès de NH_4PF_6 (227 mg, 1.39 mmol, 10 eq) est ajouté. La réaction est agitée pendant 15 min, la couleur de la solution passe au jaune. Le solvant est évaporé sous pression réduite et le brut réactionnel est purifié par chromatographie flash (Si, 25g, 50µm, injection liquide, éluant DCM / méthanol / NH_4OH 10 : 1 : 0.1). Le produit [Ir(ppy)$_2$(phen-Alk)]$^+$[PF$_6$]$^-$ est obtenu sous la forme d'une poudre jaune (119.7 mg, 0.13 mmol, 93%)

RMN ^{1}H (300 MHz, CDCl3): δ 8.81(dd, J= 7.7 Hz, J= 2.2 Hz, 2H, H10, H11), 8.59 (s, NH, H4), 8.46(dd, J = 7.8 Hz, J= 2.1 Hz, 1H, H7), 8.21(dd, J=7.2 Hz, J=2.0 Hz, 2H, H27, H12), 7.90 (dd, 2H, H23, H16), 7.79 (dd, J=7.1 Hz, J= 1.8 Hz, 1H, H6), 7.72 (m, 5H, H15, H24, H8, H9, H5), 7.28(d, J=7.3 Hz, 2H, H20, H19), 7.15(m, 2H, H17, H22), 6.95 (m, 2H, H18, H21), 6.87 (m, 2H, H14, H25), 6.46(dt, J= 7.2 Hz, J= 1.8 Hz, 2H, H13, H26), 2.92(t, J=7.5 Hz, 2H, H3), 2.69 (dt, J=7.5 Hz, J= 1.3Hz, 2H H2), 2.08(t, J=1.3 Hz, 1H, H1)

m/z calculé : 776 ; mesuré : 776.01

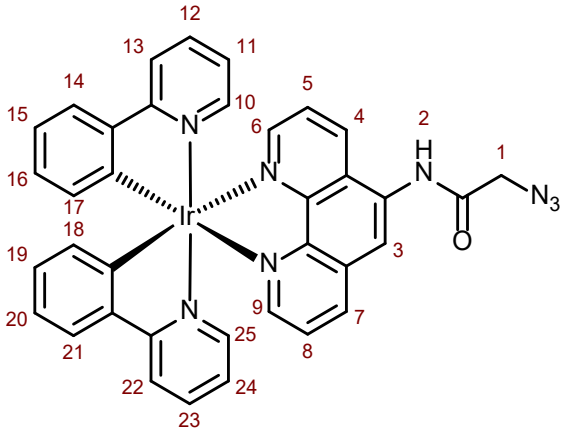


2-azido-*N*-(1,10-phenanthrolin-5-yl)acetamide

L'acide 2-azidoacétique (200 mg, 2.039 mmol, 9.7 eq) et le DCC (218 mg, 1.05 mmol, 5 eq) sont ajoutés à la 1,10-phenanthrolin-5-amine (41 mg, 0.210 mmol, 1 eq) dans un ballon de 100 mL. 60mL d'acétonitrile sont ajoutés et l'air est chassé par de l'argon. La réaction est agitée à 22°C toute la nuit. La consommation de l'acide est suivie par CCM (DCM / méthanol ; 1 : 20 + 0.5% DIPEA). Le brut réactionnel est filtré (verre fritté porosité 3) pour retirer le précipité de dicyclohexylurée (DCU) puis évaporé sous pression réduite. La poudre jaune obtenue est purifiée par chromatographie flash (Si, 40g, 50 µm, dépôt solide, éluant DCM / méthanol 2% à 6% v/v). Le produit est obtenu sous forme d'une poudre jaune (46 mg, 1.59 mmol, 78%)

RMN ^{1}H (300 MHz, DMSO-d6): δ 10.45(s, NH, H2), 9.09 (dd, 2H H6, H9), 8.52 (dd, 2H H4, H7), 8.17(s, 1H, H3), 7.81(m, 2H H5, H8), 4.26 (s, 2H, H1).

RMN ^{13}C (75 MHz, DMSO-d6): δ 168.23 (Cq, C=O), 150.10 (CH, C9), 151.22(CH, C6), 136.41(CH, C4), 132.51(CH, C7),) 124.12(CH, C8), 123.12(CH, C5), 121.22(CH, C3), 51.31 (CH2, C1)

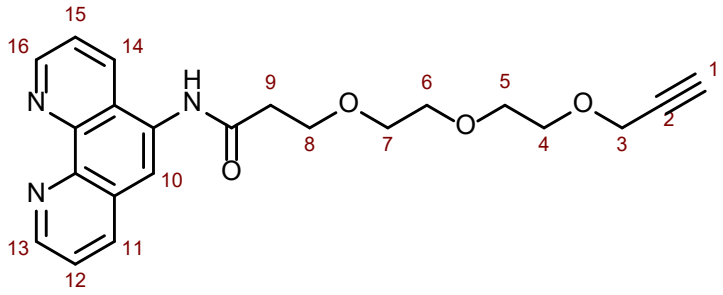


[Ir(ppy)$_2$(phen-Az)]$^+$[PF$_6$]$^-$

Le [Ir(ppy)$_2$Cl]$_2$ (58 mg, 0.054 mmol, 1 eq) et le 2-azido-*N*-(1,10-phenanthrolin-5-yl)acetamide (30 mg, 0.108 mmol, 2 eq) obtenu à l'étape précédente sont dissous dans 15mL de DCM / méthanol (2 : 1) anhydre dans un ballon de 50mL. Le mélange est porté à reflux à 55°C pendant 5h. La solution obtenue est refroidie à t.a. puis un excès de NH_4PF_6 (88 mg, 0.54 mmol, 10 eq) est ajouté. La réaction est agitée pendant 15 min, la couleur de la solution passe au jaune. Le solvant est évaporé sous pression réduite et le brut réactionnel est purifié par chromatographie flash (Si, 25g, 50µm, injection liquide, éluant DCM / acétate d'éthyle 10% à 30% v/v). Le produit [Ir(ppy)$_2$(phen-Az)]$^+$[PF$_6$]$^-$ est obtenu sous la forme d'une poudre jaune (54 mg, 0.076 mmol, 54%).

RMN ^{1}H (300 MHz, CDCl3): δ = 8.84 (dd, J=8.4Hz, 0.8Hz, 2H, H6, H9), 8.55 (s, 1H, H2), 8.39 (dd, J=10.7, 8.2, 2H, H4, H7), 8.13 (dd, 12.8, 5.9, 2H, H25, H10), 7.89 – 7.73 (m, 3H, H21, H14, H11), 7.63 (m, 4H, H22, H13, H8, H5), 7.24 (d, J=5.8, 2H, H18, H17), 7.01 (dd, J=9.3, 4.7, 2H, H20, H15), 6.95 – 6.85 (m, 2H, H23, H12), 6.81 (dd, J=12.0, 6.0, 2H, H19, H16), 6.36 – 6.26 (m, 2H, H24, H11), 5.23 (s, 1H, H3), 4.28 (s, 2H, H1).

m/z calculé : 779 ; mesuré 779.09

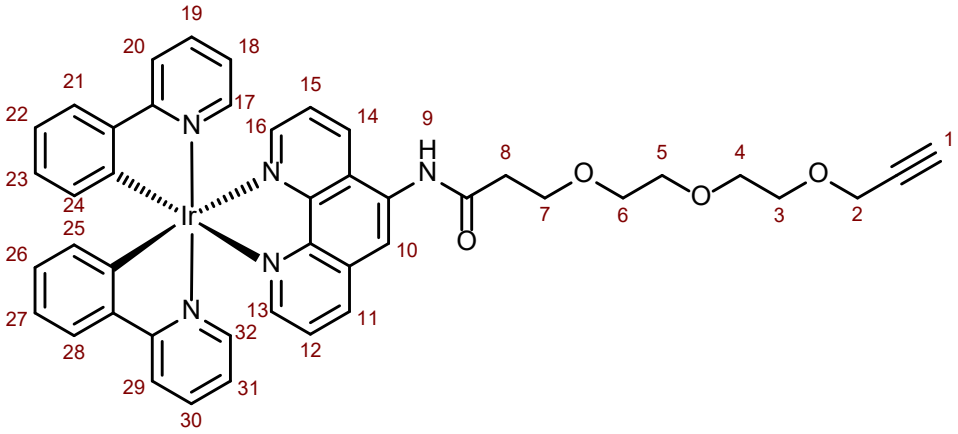


Phenanthrolin-PEG$_3$-Alk *N*-(1,10-phenanthrolin-5-yl)-3-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)propanamide

L'acide PEG$_3$-propargylique (250 mg, 1.16 mmol, 5 eq) et le DCC (119 mg, 0.58 mmol, 2.5 eq) sont dissous dans 25mL d'acétonitrile anhydre dans un ballon de 50 mL. L'air est chassé par de l'argon et la réaction est agitée à t.a. pendant 4h30. La consommation totale de l'acide est suivie par CCM (DCM / méthanol ; 20 : 1). Le brut réactionnel est filtré (verre fritté porosité 3) pour retirer le précipité de DCU. La 1,10-phénanthrolin-5-amine (45 mg, 0.232 mmol, 1 eq) est dissoute dans 3 mL d'acénotrile anhydre et ajoutée au filtrat obtenu précédemment. L'air est chassé par de l'argon et la réaction est agitée à t.a. toute la nuit. La consommation du réactif est suivie par CCM (facilement identifiable car le produit apparaît bleu vif à $\lambda = 365$ nm tandis que le réactif est jaune vif). Le brut réactionnel est purifié par chromatographie flash (Si, 25g, 50 µm, injection liquide, éluant DCM / méthanol 20 : 1). Le produit est obtenu sous la forme d'une huile translucide (91.3 mg, 0.86 mmol, 74%).

RMN ^{1}H (300 MHz, CDCl3) : δ 9.37 (br s, 1H, NH), 9.22 (dd, J= 4.3 + 1.5 Hz, 1H, H16), 9.15 (dd, J = 4.4 + 1.7 Hz, 1H, H13), 8.50 (dd, 8.4 + 1.6 Hz, H14), 8.41 (s, H10), 8.27 (dd, J = 8.2 + 1.7Hz, H11), 7.72 (dd, J = 8.4 + 4.3Hz, H15), 7.65 (dd, J = 8.1 + 4.4 Hz, H12), 3.73 (m, 12H, PEGs), 2.89 (m, 2H, H3), 2.68 (t, J = 6.2 Hz, 1H, H1)

RMN ^{13}C (75 MHz, CDCl3) : δ 171.09 (CO), 150.05 (C16), 149.20 (C13), 136.46 (C11), 131.35 (C_{IV}), 130.60 (C14), 128.66 (C_{IV}), 123.55 (C12), 123.00 (C15), 118.58 (C10), 70.42 ($CH2_{PEG}$), 70.35 ($CH2_{PEG}$), 70.18 ($CH2_{PEG}$), 69.13 (C_{IV} ?), 68.86 ($CH2_{PEG}$), 67.17 ($CH2_{PEG}$), 66.46 (C_{IV}), 37.77 (C3), 34.65 (C1)

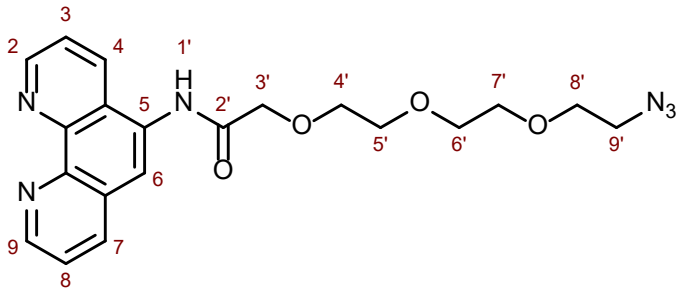


[Ir(ppy)$_2$(phen-PEG$_2$-Alk)]$^+$[PF$_6$]$^-$

Le Phenanthrolin-PEG$_3$-Alk (67 mg, 0.172 mmol, 2 eq) obtenu à l'étape précédente et le Ir(ppy)$_2$Cl]$_2$ (92 mg, 0.086 mmol, 1 eq) dans dissous dans 15mL de DCM/ méthanol anhydre (2:1) 15mL de DCM / méthanol (2 : 1) anhydre dans un ballon de 50mL. Le mélange est porté à reflux à 55°C pendant 5h. La consommation totale du réactif est suivie par CCM (DCM / méthanol, 20 : 1), facilement identifiable car le réactif apparaît bleu vif sur la plaque CCM à λ = 365 nm, tandis que le produit est orange vif. La solution obtenue est refroidie à t.a. puis un excès de NH_4PF_6 (280 mg, 1.72 mmol, 20 eq) est ajouté. La solution couleur vert olive obtenue est évaporée sous pression réduite. La poudre jaune obtenue est purifiée par chromatographie flash (Si, 40 g, 50 µm, éluant DCM / méthanol 50 : 1). Les fractions contenant le produit sont rassemblées et évaporées sous pression réduite. Le produit est obtenu sous la forme d'une poudre jaune foncé (157 mg, 0.151 mmol, 88%).

RMN ^{1}H (300 MHz, CDCl3) : δ = 8.85 (dd, J=8.6, 1.3, 2H, H16, H13), 8.58 (s, 1H, H9), 8.49 (dd, J=8.3, 1.4, 2H, H14 H11), 8.27 (dd, J=5.0, 1.2, 2H, H32, H17), 8.20 (dd, J=5.0, 1.4, 1H, H11), 7.91 (dt, J=8.5, 7.2, 2H, H28, H21), 7.78 – 7.68 (m, 5H, H29, H20, H15, H12, H10), 7.41 (dd, J=5.9, 0.7, 2H, H25, H24), 7.35 (dd, J=5.8, 0.8, 2H, H27, H22), 7.10 (tdd, J=7.7, 2.8, 1.2, 2H, H21, H28), 7.03 – 6.95 (m, 2H, H30, H19), 6.89 (ddd, J=7.3, 5.9, 1.4, 2H, H31, H18), 6.42 (dd, J=7.5, 0.9, 2H), 2.93 (t , J=5.5Hz, 1H, H1).

m/z calculé 894, mesuré 894.18

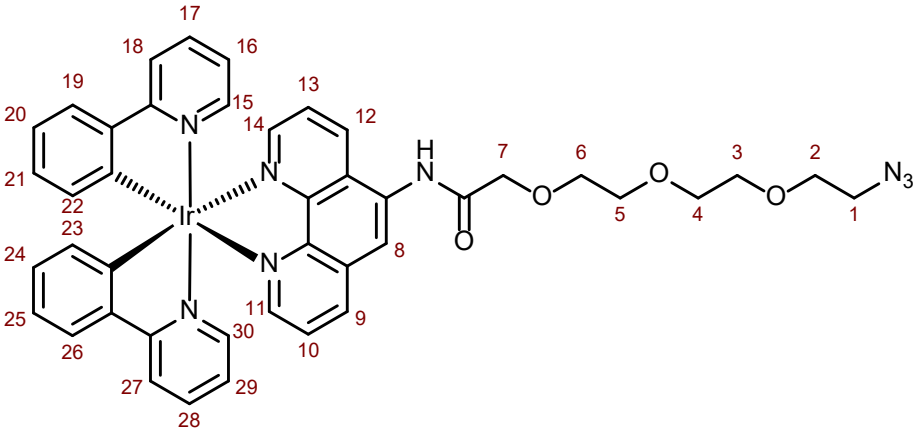


Phenanthrolin-PEG$_3$-Az 2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)-*N*-(1,10-phenanthrolin-5-yl)acetamide

L'acide PEG$_3$-azidoacétique (250mg, 1.07mmol, 5eq) et le DCC (110.6mg, 0.54mmol, 2.5eq) sont dissous dans 25mL d'acétonitrile anhydre dans un ballon de 50 mL. L'air est chassé par de l'argon et la réaction est agitée à t.a. pendant 4h. La consommation totale de l'acide est suivie par CCM (DCM / méthanol ; 4 : 1). Le brut réactionnel est filtré (verre fritté porosité 3) pour retirer le précipité de DCU. La 1,10-phénanthrolin-5-amine (42 mg, 0.214 mmol, 1 eq) est dissoute dans 5 mL d'acénotrile anhydre et ajoutée au filtrat obtenu précédemment. L'air est chassé par de l'argon et la réaction est agitée à t.a. toute la nuit. La consommation du réactif est suivie par CCM (facilement identifiable car le produit apparaît bleu vif à $\lambda = 365$ nm tandis que le réactif est jaune vif). Le brut réactionnel est purifié par chromatographie flash (Si, 25g, 50 µm, injection liquide, éluant DCM / méthanol 20 : 1). Les fractions contenant le produit sont rassemblées et concentrées sous pression réduite. Le produit est obtenu sous la forme d'une poudre jaune foncé (96.7 mg, 0.201 mmol, 94%).

RMN ^{1}H (300MHz, CDCl3) : δ 9.38 (br s, NH, H1'), 9.24 (dd, J = 4.3+ 1.5 Hz, 1H, H9), 9.15 (dd, J = 4.3 + 1.6 Hz, 1H, H2), 8.47 (dd, J = 8.4 + 1.5 Hz, H7), 8.47 (s, H6), 8.26 (dd, J = 8.1 + 1.6 Hz, H4), 7.71 (dd, J = 8.4 + 4.3 Hz, H8), 7.64 (dd, J = 8.1 + 4.4Hz H3), 4.32 (s, 2H, H3'), 3.92 (m, 2H, H5'), 3.84 (m, 2H, H4'), 3.69 (m, 2H, H6'), 3.50 (m, 2H, H7'), 3.36 (m, 2H, H8'), 3.13 (m, 2H, H9')

RMN ^{13}C (75 MHz, CDCl3) : δ 169.32 (CO), 150.27 (C9), 149.90 (C2), 136.02 (C4), 129.98 (C7), 123.57 (C3), 122.85 (C8), 119.18 (C6), 71.67 (C5'), 70.83 (C3'), 70.65 (C6'), 70.49 (C7'), 70.32 (C4'), 69.90 (C8'), 50.43 (C9').



[Ir(ppy)$_2$(phen-PEG$_2$-Az)]$^+$[PF$_6$]$^-$

Le Phenanthrolin-PEG$_3$-Az (82 mg, 0.2 mmol, 2 eq) obtenu à l'étape précédente et le Ir(ppy)$_2$Cl]$_2$ (107 mg, 0.1 mmol, 1 eq) dans dissous dans 15mL de DCM/ méthanol anhydre (2:1) 15mL de DCM / méthanol (2 : 1) anhydre dans un ballon de 50mL. Le mélange est porté à reflux à 55°C pendant 5h. Après 15h la consommation totale du réactif est suivie par CCM (DCM / méthanol, 20 : 1), facilement identifiable car le réactif apparaît bleu vif sur la plaque CCM à λ = 365 nm, tandis que le produit est orange vif. La solution obtenue est refroidie à t.a. puis un excès de NH_4PF_6 (326 mg, 2 mmol, 20 eq) est ajouté et la solution est agitée 15 min. La solution obtenue est évaporée sous pression réduite. La poudre jaune obtenue est purifiée par chromatographie flash (Si, 40 g, 50 µm, dépôt solide, éluant DCM / méthanol 33 : 1). Les fractions contenant le produit sont rassemblées et évaporées sous pression réduite. Le produit est obtenu sous la forme d'une poudre jaune vif (110 mg, 0.198 mmol, 99%).

RMN ^{1}H (300 MHz, CDCl3) : δ = 8.70 (dd, J=8.5, 1.3, 2H, H11 + H14), 8.47 (s, 1H, NH), 8.43 (dd, J=8.4, 1.4, 2H, H15 + H30), 8.20 (dd, J=5.0, 1.2, 2H, H9 + H12), 8.13 (dd, J=5.0, 1.4, 2H, H17 + H28), 7.88 – 7.75 (m, 4H, H10 + H13), 7.64 (m, 2H, H16 + H29), 7.32 (d, J=5.1, 2H, H22 + H23), 7.28 (dd, J=5.8, 0.8, 2H, H19 + H26), 7.05 – 6.96 (m, 4H, H20 + H25), 6.95 – 6.86 (m, 6H, H21 + H24), 6.85 – 6.77 (m, 1H), 6.33 (d, J=7.0, 4H), 3.88 – 3.82 (m, 4H, PEG), 3.78 (m, 4H, PEG), 3.68 (m, 4H, PEG), 3.62 – 3.56 (m, 4H, PEG), 3.53 – 3.45 (m, 4H), 3.20 (m, 4H)

m/z calculé : 911, mesuré : 911.21

PROTOCOLES DE CULTURE CELLULAIRE ET MARQUAGE MÉTABOLIQUE

CELLULES MAMMIFERES ADHÉRENTES

CULTURE CELLULAIRE ET INCORPORATIONS MÉTABOLIQUES

Les cellules humaines adhérentes ont été maintenues dans du milieu de culture Dulbecco's Modified Eagle Medium (DMEM, 4,5 g/L de glucose) complété avec 10% de sérum de veau fœtal (v/v) et 2 mM de L-glutamine. Les cellules ont été cultivées à 37 °C dans une atmosphère humidifiée enrichie en CO_2 à 5 %.

INCORPORATIONS MÉTABOLIQUES

Les cellules sont ensemencées dans des plaques 12 ou 24 puits contenant une lamelle de verre (ou une membrane en nitrure de silicium pour les observations en SR-XRF). Un jour plus tard, le milieu est supplémenté avec le rapporteur chimique désiré (500 µM pour tous les composés, sauf l'Ac_4ManNAz à 50 µM) pendant la durée désirée. Lorsque les expériences nécessitent différentes durées d'incorporation, les cellules sont ensemencées en même temps, et le milieu est supplémenté en rapporteurs chimiques à différents moments, de sorte à ce que toutes les cellules aient le même temps d'incubation. Le milieu est alors éliminé plusieurs lavages au DPBS sont effectués pour mettre fin au marquage. Après l'élimination du DPBS, 500 µL de solution de paraformaldéhyde à 4 % sont ajoutés pendant 15 minutes dans chaque puits pour l'étape de fixation. Les cellules sont alors lavées 3 fois avec du PBS, puis conservées dans le PBS à 4°C jusqu'à l'étape de ligation.

CUAAC

Les lamelles portant les cellules sont retirées des puits et placées (face portant les cellules vers le haut) dans une boîte hermétique et opaque. Chaque incubation se réalise par dépôt de la solution sur la lamelle, permettant de former un dôme de liquide au-dessus des cellules. Les membranes en nitrure de silicium étant très fragiles, elles sont laissées au fond des puits pour ces différentes étapes. La perméabilisation est effectuée en

incubant les cellules avec une solution de Triton X100 à 0,5 % pendant 10 minutes. Les cellules sont ensuite lavées avec du PBS (4 x 500 ml). La réaction de CuAAC sur les cellules adhérentes est réalisée en appliquant sur chaque lamelle portant des cellules une solution contenant : K_2HPO_4 100 mM ; Ascorbate de sodium 2,5 mM ; $CuSO_4$ 150 µM ; BTTAA 300 µM ; Sonde désirée (concentration désirée). Les lamelles de verre sont montées sur des lames de microscope en utilisant une goutte de milieu de montage (milieu de montage Dako, Agilent) puis séchées à l'obscurité, elles sont alors prêtes à être observées.

PARTICULARITÉS POUR LES EXPÉRIENCES DE SR-XRF :

Pour les observations en fluorescence des rayons X sous faisceau synchrotron, les cellules sont ensemencées sur des lamelles en nitrure de silicium plutôt que des lamelles en verre (qui est trop impur pour l'imagerie élémentaire). Les étapes de fixation, perméabilisation et CuAAC sont directement effectuées dans les puits dans lesquels les cellules ont poussé, afin de minimiser les manipulations des lamelles qui sont très fragiles. Enfin, les cellules sont simplement conservées dans de l'eau ultrapure plutôt que dans du liquide de montage avant préparation pour la cryo-microscopie.

CULTURE D'*ESCHERICHIA COLI*

Les différentes souches de bactéries *E. coli* sont ensemencées à partir de stocks à -80°C sur des boîtes de Pétri contenant du milieu LB gélifié, et sont cultivées pendant la nuit à 37°C. Des colonies isolées de ces boîtes de Pétri sont utilisées pour inoculer des cultures liquides en milieu LB qui sont alors utilisées pour la suite des expériences.

CULTURE ET INCORPORATION MÉTABOLIQUE ESCHERICHIA COLI :

Les cultures liquides d'*E. coli* préparées comme expliqué ci-dessus sont diluées à une densité optique densité optique à 600nm (DO_{600}) de 1. 20 µl de cette solution sont transférés dans un tube conique de 50 ml contenant 10 ml de LB (complété avec les marqueurs chimiques désirés) et sont cultivés pendant la nuit à 37°C/180 tr/min. Pour

le marquage par fluorescence par CuAAC, des solutions stock fraîches des différents réactifs sont utilisées pour préparer le tampon CuAAC. Il doit être préparé plusieurs minutes à une heure avant utilisation, et ne doit pas être conservé plus de quelques heures. Le $CuSO_4$ (1 mg/ml, concentration finale 150 mM) et le BTTAA (10 mg/ml, concentration finale 300 mM) sont d'abord mélangés, puis ajoutés à la sonde fluorescente (TAMRA-Alk ou TAMRA-N_3, 1 mg/ml, concentration finale: 0.25 µM pour lecture de la fluorescence en plaque, 100mM pour microscopie à fluorescence). Le K_2HPO_4 (100 mg/ml, concentration finale 100 mM) et l'eau sont ensuite ajoutés. L'ascorbate de sodium (10 mg/ml, concentration finale 2,5 mM) est ajouté en dernier, juste avant utilisation.

CUAAC

Les cultures liquids ayant poussé toute la nuit sont ajustées à $DO_{600} = 1$ en diluant la suspension dans du PBS. 200 µL de cette suspension sont répartis dans des microtubes de 2mL puis centrifugés à 10 000g pendant 2mn. Les culots sont resuspendus dans 200 mL de tampon CuAAC et agité pendant 45mn à 600 rpm à t.a. dans l'obscurité. Les suspensions sont rincées 6 fois avec 1mL PBS (2 min, 10 000g) avant d'être resuspendues dans 200 µL PBS. Pour la lecture en microplaque, les suspensions de cellules marquées par CuAAC sont réparties dans les puits d'une plaque 96 puits opaque (100 µL par puits).

PRÉPARATION DES BACTÉRIES POUR OBSERVATION AU MICROSCOPIE

Des pads d'agar sont fabriqués en versant 10 µL de LB agar chaud sur une lame de microscopie rapidement recouverts par une lamelle en verre. Après refroidissement, les suspensions de cellules sont délicatement versées sur l'agar en soulevant légèrement la lamelle en verre puis en la replaçant. Les lamelles sont alors scellées avec du vernis à ongle. Elles sont alors prêtes à être observées au microscope.

[1] J. A. Prescher, C. R. Bertozzi, *Nat Chem Biol* **2005**, *1*, 13–21.

[2] P. V. Chang, C. R. Bertozzi, *Chem. Commun.* **2012**, *48*, 8864–8879.

[3] O. Boutureira, G. J. L. Bernardes, *Chem. Rev.* **2015**, *115*, 2174–2195.

[4] I. Dovgan, O. Koniev, S. Kolodych, A. Wagner, *Bioconjugate Chem.* **2019**, *30*, 2483–2501.

[5] J. Kalia, R. Raines, *Curr Org Chem.* **2010**, *14*, 138–147.

[6] C. R. Bertozzi, M. Meldal, K. B. Sharpless, *NobelPrize.org*, www.nobelprize.org/prizes/chemistry/2022/press-release, **2022**.

[7] L. K. Mahal, K. J. Yarema, C. R. Bertozzi, *Science* **1997**, *276*, 1125–1128.

[8] H. J. Gross, A. Bunsch, J. C. Paulson, R. Brossmer, *Eur. J. Biochem.* **1987**, *168*, 595–602.

[9] E. M. Sletten, C. R. Bertozzi, *Angew. Chem. Int. Ed.* **2009**, *48*, 6974–6998.

[10] H. Staudinger, J. Meyer, *Helvetica Chimica Acta* **1919**, *2*, 635–646.

[11] E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007–2010.

[12] C. Bednarek, I. Wehl, N. Jung, U. Schepers, S. Bräse, *Chem. Rev.* **2020**, *120*, 4301–4354.

[13] R. Huisgen, *Angew. Chem. Int. Ed.* **1963**, *2*, 565–598.

[14] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.

[15] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *14*, 2596–2599.

[16] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064.

[17] L. Gaetke, *Toxicology* **2003**, *189*, 147–163.

[18] 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.

[19] 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.

[20] F. S. Ekholm, H. Pynnönen, A. Vilkman, J. Koponen, J. Helin, T. Satomaa, *Org. Biomol. Chem.* **2016**, *14*, 849–852.

[21] V. Flon, M. Bénard, D. Schapman, L. Galas, P.-Y. Renard, C. Sabot, *Biomolecules* **2020**, *10*, 619–631.

[22] 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.

[23] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.

[24] S. T. Laughlin, J. M. Baskin, S. L. Amacher, C. R. Bertozzi, *Science* **2008**, *320*, 664–667.

[25] J. C. Jewett, E. M. Sletten, C. R. Bertozzi, *J. Am. Chem. Soc.* **2010**, *132*, 3688–3690.

[26] S. T. Laughlin, C. R. Bertozzi, *ACS Chem. Biol.* **2009**, *4*, 1068–1072.

[27] 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.

[28] R. van Geel, G. J. M. Pruijn, F. L. van Delft, W. C. Boelens, *Bioconjugate Chem.* **2012**, *23*, 392–400.

[29] C. Zhang, P. Dai, A. A. Vinogradov, Z. P. Gates, B. L. Pentelute, *Angew. Chem. Int. Ed.* **2018**, *57*, 6459–6463.

[30] N. K. Devaraj, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2008**, *19*, 2297–2299.

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

[32] A.-C. Knall, C. Slugovc, *Chem. Soc. Rev.* **2013**, *42*, 5131–5142.

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

[34] 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.

[35] A. Darko, S. Wallace, O. Dmitrenko, M. M. Machovina, R. A. Mehl, J. W. Chin, J. M. Fox, *Chem. Sci.* **2014**, *5*, 3770–3776.

[36] D. M. Patterson, L. A. Nazarova, B. Xie, D. N. Kamber, J. A. Prescher, *J. Am. Chem. Soc.* **2012**, *134*, 18638–18643.

[37] V. Rigolot, C. Biot, C. Lion, *Angew. Chem. Int. Ed.* **2021**, *60*, 23084–23105.

[38] R. A. Flynn, K. Pedram, S. A. Malaker, P. J. Batista, B. A. H. Smith, A. G. Johnson, B. M. George, K. Majzoub, P. W. Villalta, J. E. Carette, C. R. Bertozzi, *Cell* **2021**, *184*, 3109–3124.

[39] D. B. Werz, R. Ranzinger, S. Herget, A. Adibekian, C.-W. Von Der Lieth, P. H. Seeberger, *ACS Chem. Biol.* **2007**, *2*, 685–691.

[40] S. Groux-Degroote, F. Foulquier, S. Cavdarli, P. Delannoy, *m/s* **2021**, *37*, 609–617.

[41] N. Zachara, Y. Akimoto, M. Boyce, G. Hart, in *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), **2022**.

[42] R. C. R. Jala, S. Vudhgiri, C. G. Kumar, *Carbohydrate Research* **2022**, *516*, 108556.

[43] L. Möckl, *Front. Cell Dev. Biol.* **2020**, *8*, 253.

[44] G. Blix, *Z Physiol Chem* **1936**, *240*, 43–54.

[45] A. Lundblad, *Upsala Journal of Medical Sciences* **2015**, *120*, 104–112.

[46] F. Blix, A. Gottschalk, E. Klenk, *Nature* **1957**, *179*, 1088.

[47] A. Varki, *Nature* **2007**, *446*, 1023–1029.

[48] T. Angata, A. Varki, *Chem. Rev.* **2002**, *102*, 439–469.

[49] I. Fliniaux, G. Marchand, C. Molinaro, M. Decloquement, A. Martoriati, M. Marin, J.-F. Bodart, A. Harduin-Lepers, K. Cailliau, *Front. Cell Dev. Biol.* **2022**, *10*, 982931.

[50] S. Holst, M. Wuhrer, Y. Rombouts, in *Advances in Cancer Research*, Elsevier, **2015**, pp. 203–256.

[51] F. Boyaval, R. Van Zeijl, H. Dalebout, S. Holst, G. Van Pelt, A. Fariña-Sarasqueta, W. Mesker, R. Tollenaar, H. Morreau, M. Wuhrer, B. Heijs, *Molecular & Cellular Proteomics* **2021**, *20*, 100057.

[52] J. A. Ferreira, A. Magalhães, J. Gomes, A. Peixoto, C. Gaiteiro, E. Fernandes, L. L. Santos, C. A. Reis, *Cancer Letters* **2017**, *387*, 32–45.

[53] Y. Haga, K. Ueda, *Glycoconj J* **2022**, *39*, 303–313.

[54] E. R. Vimr, K. A. Kalivoda, E. L. Deszo, S. M. Steenbergen, *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 132–153.

[55] A. F. Carlin, S. Uchiyama, Y.-C. Chang, A. L. Lewis, V. Nizet, A. Varki, *Blood* **2009**, *113*, 3333–3336.

[56] M. E. Tanner, *Bioorganic Chemistry* **2005**, *33*, 216–228.

[57] G. Reuter, H.-J. Gabius, *Biol. Chem.* **1996**, *377*, 325–342.

[58] B. Dudek, J. Rybka, G. Bugla-Płoskońska, A. Korzeniowska-Kowal, B. Futoma-Kołoch, A. Pawlak, A. Gamian, *Front. Microbiol.* **2022**, *13*, 1028796.

[59] T. Corfield, *Glycobiology* **1992**, *2*, 509–521.

[60] D. M. Shell, L. Chiles, R. C. Judd, S. Seal, R. F. Rest, *Infect Immun* **2002**, *70*, 3744–3751.

[61] P. A. Gilormini, C. Lion, D. Vicogne, Y. Guérardel, F. Foulquier, C. Biot, *J Inherit Metab Dis* **2018**, *41*, 515–523.

[62] A. Dumont, A. Malleron, M. Awwad, S. Dukan, B. Vauzeilles, *Angew. Chem. Int. Ed.* **2012**, *51*, 3143–3146.

[63] E. Fugier, A. Dumont, A. Malleron, E. Poquet, J. Mas Pons, A. Baron, B. Vauzeilles, S. Dukan, *PLoS ONE* **2015**, *10*, e0127700.

[64] J. Mas Pons, A. Dumont, G. Sautejeau, E. Fugier, A. Baron, S. Dukan, B. Vauzeilles, *Angew. Chem. Int. Ed.* **2014**, *53*, 1275–1278.

[65] S. T. Laughlin, C. R. Bertozzi, *Nat Protoc* **2007**, *2*, 2930–2944.

[66] A. Sarkar, T. Fritz, W. Taylor, J. Esko, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3323–3327.

[67] R. Schauer, G. V. Srinivasan, D. Wipfler, B. Kniep, R. Schwartz-Albiez, in *The Molecular Immunology of Complex Carbohydrates-3* (Ed.: A.M. Wu), Springer US, Boston, MA, **2011**, pp. 525–548.

[68] A. K. Antonczak, Z. Simova, E. M. Tippmann, *J. Biol. Chem.* **2009**, *284*, 28795–28800.

[69] K. S. Kim, *Nat Rev Neurosci* **2003**, *4*, 376–385.

[70] K. Vellaisamy, G. Li, C.-N. Ko, *Chem. Sci.* **2018**, *9*, 1119–1125.

[71] A. Mai-Prochnow, M. Clauson, J. Hong, A. B. Murphy, *Sci Rep* **2016**, *6*, 38610.

[72] C. Whitfield, S. S. Wear, C. Sande, *Annu. Rev. Microbiol.* **2020**, *74*, 521–543.

[73] L. Yakovlieva, J. A. Fülleborn, M. T. C. Walvoort, *Front. Microbiol.* **2021**, *12*, 745702.

[74] S. Barrallo, A. Reglero, B. Revilla-Nuin, H. Martínez-Blanco, L. B. Rodríguez-Aparicio, M. A. Ferrero, *FEBS Letters* **1999**, *445*, 325–328.

[75] O. Rendueles, M. Garcia-Garcerà, B. Néron, M. Touchon, E. P. C. Rocha, *PLoS Pathog* **2017**, *13*, e1006525.

[76] E. coli K antigen 3-dimensional structure database (EK3D), https://www.iith.ac.in/EK3D/index.php

[77] J. W. Ezzell, S. L. Welkos, *J Appl Microbiol* **1999**, *87*, 250–250.

[78] S. Arredondo-Alonso, G. Blundell-Hunter, Z. Fu, R. A. Gladstone, A. Fillol-Salom, J. Loraine, E. Cloutman-Green, P. J. Johnsen, Ø. Samuelsen, A. K. Pöntinen, F. Cléon, S. Chavez-Bueno, M. A. De La Cruz, M. A. Ares, M. Vongsouvath, A. Chmielarczyk, C. Horner, N. Klein, A. McNally, J. N. Reis, J. R. Penadés, N. R. Thomson, J. Corander, P. W. Taylor, A. J. McCarthy, *Nat. Commun.* **2023**, *14*, 3294.

[79] W. Egan, T.-Y. Liu, D. Dorow, J. S. Cohen, J. D. Robbins, E. C. Gotschlich, J. B. Robbins, *Biochemistry* **1977**, *16*, 3687–3692.

[80] G. T. Barry, W. F. Goebel, *Nature* **1957**, *179*, 206–206.

[81] R. L. Schnaar, R. Gerardy-Schahn, H. Hildebrandt, *Physiol Rev* **2014**, *94*, 461–518.

[82] T. M. Villanueva-Cabello, L. D. Gutiérrez-Valenzuela, R. Salinas-Marín, D. V. López-Guerrero, I. Martínez-Duncker, *Front. Immunol.* **2022**, *12*, 823637.

[83] S. Sarkar, G. C. Ulett, M. Totsika, M.-D. Phan, M. A. Schembri, *PLoS ONE* **2014**, *9*, e94786.

[84] S. Phanphak, P. Georgiades, R. Li, J. King, I. S. Roberts, T. A. Waigh, *Langmuir* **2019**, *35*, 5635–5646.

[85] K. Lang, J. W. Chin, *Chem. Rev.* **2014**, *114*, 4764–4806.

[86] J. Trouve, A. Zapun, C. Arthaud, C. Durmort, A. M. Di Guilmi, B. Söderström, A. Pelletier, C. Grangeasse, D. Bourgeois, Y.-S. Wong, C. Morlot, *Current Biology* **2021**, *31*, 2844–2856.

[87] C. T. Saeui, E. Urias, L. Liu, M. P. Mathew, K. J. Yarema, *Glycoconj J.* **2015**, *32*, 425–441.

[88] N. Banahene, H. W. Kavunja, B. M. Swarts, *Chem. Rev.* **2022**, *122*, 3336–3413.

[89] H. Liang, K. E. DeMeester, C.-W. Hou, M. A. Parent, J. L. Caplan, C. L. Grimes, *Nat. Commun.* **2017**, *8*, 15015.

[90] A. R. Brown, K. A. Wodzanowski, C. C. Santiago, S. N. Hyland, J. L. Follmar, P. Asare-Okai, C. L. Grimes, *ACS Chem. Biol.* **2021**, *16*, 1908–1916.

[91] T. J. Fiolek, N. Banahene, H. W. Kavunja, N. J. Holmes, A. K. Rylski, A. A. Pohane, M. S. Siegrist, B. M. Swarts, *ChemBioChem* **2019**, *20*, 1282–1291.

[92] N. Geva-Zatorsky, D. Alvarez, J. E. Hudak, N. C. Reading, D. Erturk-Hasdemir, S. Dasgupta, U. H. von Andrian, D. L. Kasper, *Nat. Med.* **2015**, *21*, 1091–1100.

[93] P. Kaewsapsak, O. Esonu, D. H. Dube, *ChemBioChem* **2013**, *14*, 721–726.

[94] E. R. Vimr, F. A. Troy, *J. Bacteriol.* **1985**, *164*, 854–860.

[95] C. Møller-Olsen, T. Ross, K. N. Leppard, V. Foisor, C. Smith, D. K. Grammatopoulos, A. P. Sagona, *Sci. Rep.* **2020**, *10*, DOI 10.1038/s41598-020-65867-4.

[96] B. Revilla-Nuin, Á. Reglero, H. Martínez-Blanco, I. G. Bravo, M. A. Ferrero, L. B. Rodríguez-Aparicio, *FEBS Letters* **2002**, *511*, 97–101.

[97] L. B. Rodríguez-Aparicio, A. Reglero, A. I. Ortiz, J. M. Luengo, *Appl Microbiol Biotechnol* **1988**, *27*, 474–483.

[98] V. Rigolot, Y. Rossez, C. Biot, C. Lion, *RSC Chem. Biol.* **2023**, *4*, 173–183.

[99] A. R. Batt, B. W. Zaro, M. X. Navarro, M. R. Pratt, *ChemBioChem* **2017**, *18*, 1177–1182.

[100] P. V. Chang, X. Chen, C. Smyrniotis, A. Xenakis, T. Hu, C. R. Bertozzi, P. Wu, *Angew Chem Int Ed* **2009**, *48*, 4030–4033.

[101] J. Biwi, C. Clarisse, C. Biot, R. P. Kozak, K. Madunic, M. Mortuaire, M. Wuhrer, D. I. R. Spencer, C. Schulz, Y. Guerardel, T. Lefebvre, A. Vercoutter-Edouart, *Proteomics* **2019**, *19*, 1800452.

[102] V. Rigolot, MOE_Golgi_Analyzer, www.doi.org/10.5281/zenodo.6778042, **2022**

[103] 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.

[104] P.-A. Gilormini, C. Lion, M. Noel, M.-A. Krzewinski-Recchi, A. Harduin-Lepers, Y. Guérardel, C. Biot, *Glycobiology* **2016**, *26*, 1151–1156.

[105] M. Noel, P.-A. Gilormini, V. Cogez, C. Lion, C. Biot, A. Harduin-Lepers, Y. Guérardel, *Bioconjugate Chem.* **2018**, *29*, 3377–3384.

[106] J. Biwi, C. Biot, Y. Guerardel, A.-S. Vercoutter-Edouart, T. Lefebvre, *Molecules* **2018**, *23*, 2858.

[107] P. Umapathi, O. O. Mesubi, P. S. Banerjee, N. Abrol, Q. Wang, E. D. Luczak, Y. Wu, J. M. Granger, A.-C. Wei, O. E. Reyes Gaido, L. Florea, C. C. Talbot, G. W. Hart, N. E. Zachara, M. E. Anderson, *Circulation* **2021**, *143*, 1687–1703.

[108] V. Korinek, N. Barker, P. J. Morin, D. Van Wichen, R. De Weger, K. W. Kinzler, B. Vogelstein, H. Clevers, *Science* **1997**, *275*, 1784–1787.

[109] A. Kasprowicz, C. Spriet, C. Terryn, V. Rigolot, S. Hardiville, M. G. Alteen, T. Lefebvre, C. Biot, *Molecules* **2020**, *25*, 4501.