Bioorthogonal metabolic labelling with acyl-CoA reporters

Ourailidou, Maria E.; Zwinderman, Martijn R.H.; Dekker, Frans

Published in:
MedChemCommun

DOI:
10.1039/c5md00446b

Publication date:
2016

MedChemComm

REVIEW

View Article Online
View Journal

CrossMark click for updates

Cite this: DOI: 10.1039/c5md00446b

Bioorthogonal metabolic labelling with acyl-CoA reporters: targeting protein acylation†

Maria. E. Ourailidou, Martijn R. H. Zwinderman and Frank J. Dekker*

Received 2nd October 2015,
Accepted 14th November 2015

DOI: 10.1039/c5md00446b

www.rsc.org/medchemcomm

Protein acylation is an abundant post-translational modification with a pivotal role in a plethora of biological processes. To date, metabolic labelling with functionalized precursors of acyl-CoA and subsequent bioorthogonal ligation to a complementary detection tag has offered an attractive approach for monitoring endogenous protein acylation with excellent selectivity. This review focuses on the applications of alkyne- and alkene-based bioorthogonal chemistries in the study of enzyme activity *in vitro* and summarizes the carboxylate-type chemical reporters that have enabled the visualization and identification of cellular acylated proteins. However, despite their importance, serious limitations question the use of this two-step labelling method in the quantification of the protein acylome.

1. Introduction

Chemical labelling initiated a revolution in the field of proteomics, and novel bioorthogonal reactions now offer a valuable addition to traditional detection methods, such as antibodies. Procedures involving bioorthogonal chemistry are typically based on the tagging of biomolecules with functionalities that allow for subsequent covalent attachment of an affinity label *via* a bioorthogonal chemical reaction. This reaction must be site-specific, rapid and compatible with an aqueous environment, neutral pH and room temperature. To date, an increasing number of such reactions has been developed to facilitate the study of protein function *in vitro*, among which the "click" reaction has a central position.[1,2]

The analysis of endogenous biological activity faces serious challenges due to the complexity of living systems. Nevertheless, metabolic labelling techniques in combination with bioorthogonal chemistry have paved the way for the elucidation of numerous biochemical processes.[3] Towards this aim, a wide variety of small synthetic molecules, coupled to functionalized moieties, was developed to function as structural analogues of cellular metabolites. Of special interest are terminal alkynes or alkenes whose bioorthogonal properties favour cellular uptake and use in a biological context. As a first step, these reporters are enzymatically incorporated into protein-targets *via* metabolic pathways. As a second step, these targets are visualized and identified *via* a site-specific chemical labelling with a detection tag.

Pharmaceutical Gene Modulation, Groningen Research Institute of Pharmacy (GRIP), University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands. E-mail: f.j.dekker@rug.nl; Fax: +31 50 3637953; Tel: +31 50 3638030

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5md00446b

Maria. E. Ourailidou

Maria E. Ourailidou was born in Athens, Greece in 1990. She received her Msc degree in the Department of Pharmacy at the University of Athens in 2012. Since then, she has been a PhD student at the University of Groningen, The Netherlands under the guidance of Professor Frank J. Dekker. Her research interests lie in the detection and inhibition of enzyme activity in inflammation and cancer.

Martijn R. H. Zwinderman

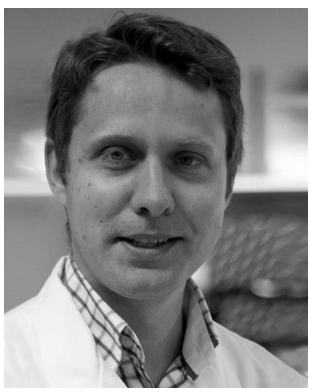
Martijn R. H. Zwinderman was born in Groningen, The Netherlands, in 1990. He received his MSc degree in Pharmacy at the University of Groningen in 2015. Since then, he has been a PhD student at the University of Groningen, The Netherlands, under the guidance of Professor Frank J. Dekker. His work focusses on the synthesis and effects of novel regulators of epigenetic processes in inflammatory diseases.

This two-step labelling strategy has found great application in shaping the profile of several post-translational modifications.[4] Among them, protein acylation, referring to the enzymatic transfer of an acyl moiety from acyl-CoA to protein residues, has received extensive attention.[5–7] For instance, histone acetylation, catalysed by Histone Acetyltransferases (HATs), plays a crucial role in chromatin remodelling and gene transcription.[8] Moreover, lysine acetylations of non-histone proteins are strongly associated with inflammatory diseases[9,10] and cancer progression.[11] Finally, protein coupling with long-chain fatty acids is known to contribute to cell proliferation, oncogenesis,[12] and HIV infection,[13] by affecting subcellular protein localization, trafficking and protein–protein interactions.[14,15]

In this review we discuss the requirements and applications of bioorthogonal chemistry to study the activity of enzymes involved in cellular protein acylations. We place special emphasis on the efforts that have been made to metabolically label such proteins by the use of bioorthogonal chemical reporters. Converging on acyl-CoA as a crucial integrator of numerous biological routes, we outline the reporters that facilitate imaging and identification of protein acylations that can be targeted in drug discovery approaches. While this has already led to promising insights in the acetylome, we conclude by highlighting the challenges that need to be overcome in order to gain a better understanding of the significance of protein acylation in pathological conditions.

2. Bioorthogonal chemistry in context

One of the oldest conventional methods for the detection of endogenous biomolecules is autoradiography, which revolves around the use of specific radioisotope-labelled substances. For example, addition of radiolabelled precursors of acetyl-CoA (Ac-CoA) allowed researchers to understand the mechanisms behind lipogenesis[16–18] and p53 acetylation *in vivo*.[19] However, this technique is frequently hampered by practical limitations considering its high risk and the time needed for execution.[20]



Frank J. Dekker

Frank J. Dekker was born in Middelburg, The Netherlands in 1977. He received his PhD degree in Medicinal Chemistry at the University of Utrecht in 2004. After doing a postdoc in the Max-Planck institute in Dortmund, Germany under supervision of Prof. H. Waldmann he moved to the University of Groningen, The Netherlands to start his own independent research line mainly focused on chemical-biological approach for detection and inhibition of enzyme activity in inflammation.

More commonly used research methodologies in the field of proteomics make use of protein-specific antibodies, such as immunoprecipitation,[21] Western blot[22] and ELISA.[23] While antibodies have become the standard choice for studying proteins in complex biological matrices, a lack of antibody selectivity due to cross reactivity with other epitopes and loss of target proteins due to poor affinity are in some cases problematic.[24,25] Moreover, for certain protein modifications antibodies do not exist. An important example is the absence of antibodies against lipid protein modifications, such as myristoylation and palmitoylation.

Some limitations of antibodies can be circumvented by using bioorthogonal chemical labelling strategies instead. One of the first bioorthogonal reactions was presented by Bertozzi in 2000 (before she coined the term in 2003) and she humbly called it the Staudinger ligation.[26] Deviating from the Staudinger reduction,[27] where an azide reacts with a triaryl phosphine to form a water-labile aza-ylide intermediate, the now-named Staudinger–Bertozzi ligation yields a stable amide bond through a triaryl phosphine carrying an electrophilic trap. The trap, a methoxycarbonyl at the *ortho*-position, stabilizes the intermediate by intramolecular cyclization.[28] Proof of principle was provided by the coupling of biotinylated phosphine molecules to cell-surface azido sialic acid chemical reporters.[26] The abiotic nature of both reaction partners, the mild reaction conditions and the favourable reaction kinetics allowed chemoselective ligation to occur for the first time in a variety of biological systems. Since the pioneering work of Bertozzi, she and others massively expanded the bioorthogonal toolkit. Over the years, some of the newer reactions show marked developments with regard to reaction kinetics and nature of the chemical reporter.[29]

2.1 Exploring post-translational protein modifications

Selecting a suitable reaction for a given application is not necessarily an easy task and requires several aspects to be taken into account. One important consideration is whether a reaction can be carried out *in vitro* or *in vivo*, and correspondingly, whether or not performing a reaction *in vivo* is truly necessary. For instance, in two-step labelling procedures with the goal of detecting proteins, the first step, comprising metabolic incorporation of chemical reporters using a cell's own metabolic machinery, typically needs to occur *in vivo*. The second step, however, may be performed *in vitro*, since proteins of interest are then already labelled. This is only feasible when metabolic labelling yields a stable chemical reporter-to-protein bond, and when adequate measures have been taken to reduce or control for background noise due to side-reactivity or toxicity of the selected reaction. Taking these requirements into consideration, we can argue that *in vivo* labelling with the appropriate probes and *in vitro* bioconjugation of isolated proteins are sufficient for monitoring the desired biological target especially when toxicity of the coupling reagents is a concerning issue.

In the context of metabolic tagging, chemical reporters must meet several criteria. The most important being the size of the chemical motif, with only small chemical motifs being widely compatible with the cellular biosynthetic machinery. This may be particularly true for smaller probes, where the relative contribution of the chemical motif to the overall probe-size becomes larger. Prime examples of small, non-perturbing bioorthogonal functional groups are terminal alkenes and alkynes.[6] Consisting of only two carbons, their bioorthogonal properties compare favourably to the larger and charged azides used in the Bertozzi–Staudinger ligation. In conclusion, terminal alkynes and alkenes are in principle ideal chemical reporters. However, solely comparing the properties of a chemical reporter is of no use since the characteristics of the corresponding bioorthogonal reaction are equally important. Consequently, we will now compare several reported reactions with terminal alkynes and alkenes.

3. Bioorthogonal reactions with terminal alkynes and alkenes

3.1 Alkyne–azide: "click"

The first bioorthogonal reaction to involve a terminal alkyne is the copper-catalysed alkyne–azide cycloaddition (CuAAC), known as "click" reaction (Fig. 1A). Based on the azide–alkyne Huisgen [3 + 2] 1,3 dipolar cycloaddition,[30] Meldal and Sharpless independently discovered that addition of copper(I)-salts increases the rate of triazole formation.[31,32] Meeting most requirements for bioorthogonality, Wang *et al.* fine-tuned the CuAAC to conjugate dye-alkynes to azide-labelled virus particles.[33] One of the adjustments needed in the context of bioconjugation was the use of TCEP (tris(2-carboxyethyl)phosphine) instead of ascorbate, since the latter induced substantial disassembly of the virus capsid. It is now understood that this may be caused by ascorbate's ability to form dehydroascorbate byproducts with arginine residues.[34] Despite of this ascorbate is the preferred reducing agent, since the azide-reducing and copper-binding properties of phosphine TCEP causes it to interfere with the CuAAC reaction. To suppress ascorbate-related side-reactions, aminoguanidine may be added or hydroxylamine can be used as the reductant.[35]

While exchanging ascorbate with TCEP turned out to be disadvantageous, an adjustment that did enhance the reaction in general was the addition of a tris(triazolyl)amine ligand.[36] Chan *et al.* observed that synthesis of polytriazoles with the CuAAC occurs at an unusual high rate (through autocatalysis). The stabilizing effect of polytriazole ligands on the Cu(I) oxidation state not only leads to an increased reaction rate, but may also protect biomolecules from hydrolysis by Cu(II) species. Additionally, polytriazoles may act as scavengers of reactive oxygen species resulting from copper or ascorbate related side-reactions, thereby minimizing oxidation of histidines and other amino acid residues.

The main advantage of the CuAAC reaction is the relatively fast reaction rate. Compared to the Bertozzi–Staudinger ligation, for instance, a 25-fold faster rate has been reported.[3] On the other hand, ascorbate's side reactivity[37] and the need for additional reagents to keep copper at the reduced state represent substantial downsides.

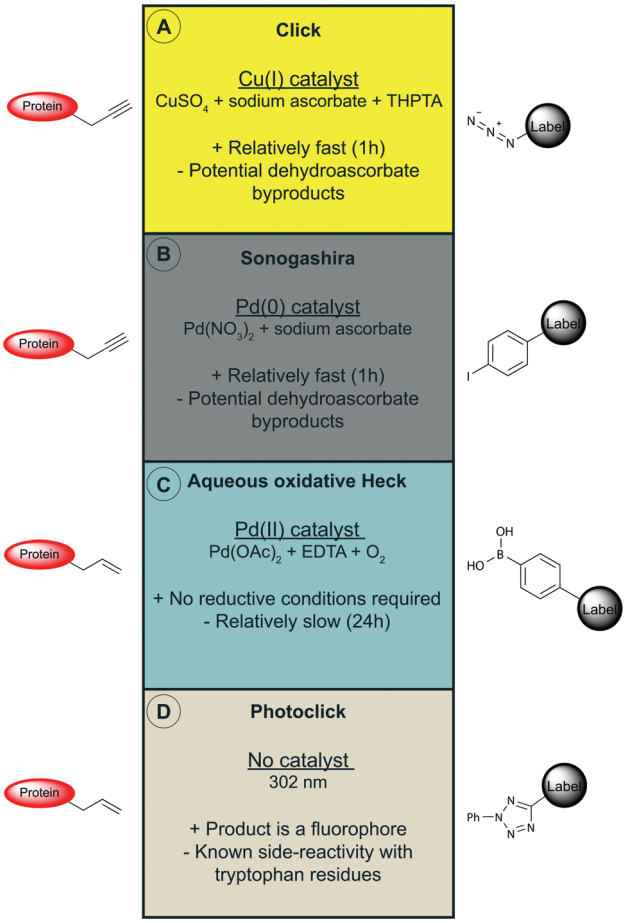


Fig. 1 Overview of bioorthogonal reactions with terminal alkyne- or alkene-tagged proteins for *in vitro* labelling.

3.2 Alkyne–halide: Sonogashira

Another interesting reaction that employs terminal alkynes is the Sonogashira cross-coupling reaction.[38] Here, alkynes react with aryl or vinylhalides to yield a carbon–carbon bond. Traditionally, the reaction proceeds at room temperature in organic solvents, employing a ligand-bound palladium(0) catalyst, a copper(I) co-catalyst and a base. Yet, Kodama *et al.* found that the reaction also proceeds in aqueous media, buffered to a slightly basic pH, with the addition of sodium ascorbate and 12% DMSO.[39] Unfortunately, the required use of ascorbate to assure reductive conditions entails that the same CuAAC related drawbacks apply.

Therefore, Li *et al.* set out to develop a copper-free Sonogashira cross-coupling reaction.[40] They succeeded with an efficient water-soluble palladium–2-amino-4,6-dihydroxypyrimidine (ADHP) complex and selectively coupled fluorescein iodide to alkyne-functionalized proteins. Moreover,

Chen's group recently reported a ligand-free Sonogashira coupling method (Fig. 1B).[41] They found that $Pd(NO_3)_2$ as a catalyst sufficiently accelerates cross-coupling of rhodamine-labelled phenyl iodide to alkyne-encoded GFP. It is worth mentioning that they demonstrated that their cross-coupling method is suitable for protein labelling inside bacterial cells without apparent toxicity of the palladium catalyst. On the other hand, non-specific metal binding to proteins has been observed, resulting in the requisite use of a metal chelator.[42]

3.3 Terminal alkynes *versus* terminal alkenes

While tremendous progress has been made in the development of alkyne-based labelling strategies, there are some anecdotal problems associated with the introduction of terminal alkynes in living systems. Problems mainly arise due to the relative acidity of the alkynic proton, which affects the stability of terminal alkynes.[43] A closely related issue is that alkyne-labelled carboxylic acids, such as 17-octadecynoic acid, can be converted into highly reactive intermediates by oxidative enzymes, which causes covalent inhibition at 10 μM, whereas metabolic labelling requires the same concentrations.[44–46] Also, alkyne homo-coupling[47] and the covalent binding of terminal alkynes to active site cysteine residues[48] (the thiol–yne reaction) have been described. While the latter may equally apply to alkenes,[49,50] the lower intrinsic reactivity of alkenes makes this less likely to occur.

3.4 Alkene–boronic acid: aqueous oxidative Heck

The potential benefits of olefinic chemical reporters are only relevant when suitable alkene-targeted reactions exist. Only some reactions that originate from the realms of organic chemistry exploit the low intrinsic reactivity of alkenes, and even fewer may find an application in proteomics. An especially promising candidate to this end is the (Mizoroki–) Heck reaction, which involves the coupling of a terminal alkene with an unsaturated halide in presence of base and a palladium(0) catalyst to form a substituted alkene.[51,52] Proof of concept was provided by Kodama *et al.* with the conjugation of vinylated biotin with iodophenyl-functionalized Ras proteins.[39] While they did not observe a significant decrease in the stability of model Ras protein during or after the reaction, the comparatively low yield of 25% is a major shortcoming. Additional work by Uemura *et al.* found that commercially available boronic acids could be used instead of unsaturated halides.[53] Also, Cho *et al.* reported an oxidative palladium(II)-catalysed reaction of boronic acids with alkenes without the need for a base and high temperatures.[54] Further optimization of the catalyst, using BIAN ligands, promoted the reaction of many different arylboronic acids with alkenes.[55] Encouraged by these results, we recently developed a Heck-type reaction to label protein-bound alkenes with arylboronic acid derivatives *in vitro*.[56] Gratifyingly, we observed that, under the optimized conditions, the oxidative Heck reaction proceeds in high yields and is chemoselective towards terminal olefins, even in complex protein mixtures. Yet, the addition of EDTA after completion of the reaction was found to be essential for chelating palladium out from protein sites. This problem was circumvented with the later development of a water-soluble Pd(II)–EDTA catalyst that effectively catalyses the reaction at room temperature in aqueous environment (Fig. 1C).[57] The reaction does not require any other additives and runs smoothly under oxygen atmosphere. A downside of the oxidative Heck reaction is the relatively long reaction time needed for the cross-coupling (24 h).

3.5 Alkene–nitrile imine: "photoclick"

Lastly, exciting new developments centre around the reaction between a nitrile imine and an alkene through a photoinducible dipolar cycloaddition.[58] Like the CuAAC, the original reaction comes from the group of Huisgen. Recently, Song *et al.* optimized this transformation in bioorthogonal labelling experiments and called it the "photoclick" reaction (Fig. 1D).[59] The reaction relies on the *in situ* activation of tetrazoles, which rapidly release nitrogen to produce nitrile imines upon exposure to UV-light. Nitrile imines subsequently react with alkenes to form fluorescent pyrazoline cycloadducts. While nitrile imines may form adducts with water, reaction with the dipolarophilic alkene generally occurs at a much faster rate. The reaction has been used in live *E. Coli* cells containing protein-bound alkenes.[60] In this study, a tetrazole was added followed by UV irradiation for 4 min. The bacterial cells were then incubated overnight to allow the cycloaddition to proceed to completion. The next day, alkene-containing cells were visualized by fluorescent imaging.

An advantage of the "photoclick" reaction is that nitrile-imines appear to be highly bioorthogonal and only slowly degrade in water. Yet, in the absence of alkene moieties, conjugation of tetrazoles to tryptophan residues is known to occur, depicting a serious drawback.[61,62] Although more recently developed tetrazoles allow for longer wavelength activation, thereby minimizing the risk of phototoxicity, they have only been shown to react with non-terminal conjugated alkenes *via* an inverse electron-demand Diels–Alder reaction.[63]

4. Metabolic labelling of protein acylations: acetyl-CoA metabolites as key intermediates

The mentioned reactions have been used to investigate a large variety of post-translational modifications.[4] Focusing on the acylome, incorporation of terminal alkenes or alkynes in chemical precursors of protein acylations enables selective monitoring of these modifications. The idea behind it is to use the cell's own metabolic machinery to install the functionalized acyl-groups in order to visualize intracellular processes. In this respect, we emphasize on Ac-CoA

intermediates that have a key role in a vast array of metabolic pathways. Generated from the catabolism of fats, sugars and proteins, Ac-CoA is oxidized *via* the Tricarboxylic Acid (TCA or Krebs) cycle to provide the main energy source for cell respiration. Moreover, its quintessential thioester bond facilitates the transfer of an acetyl group during the anabolism of essential molecules (*e.g.* fatty acids) as well as several protein acetylation reactions. In the subsequent sections we will first discuss the biosynthesis of Ac-CoA and the role of its intermediates in cellular metabolism. From this, we will examine current and future entry routes for metabolic labelling of protein acylations.

4.1 Anabolism of acetyl-CoA

Formation of Ac-CoA takes place in mitochondria *via* the oxidative degradation of mainly three types of biomolecules: carbohydrates, fatty acids and amino acids.[64,65] Considering carbohydrate metabolism, pyruvate produced during the second step of aerobic glycolysis can be converted into Ac-CoA in mitochondria through oxidative decarboxylation.[66] The transformation of the three-carbon containing pyruvate to a two-carbon Ac-CoA is catalysed by the multi-enzymatic complex of pyruvate dehydrogenase (PDH) *via* stepwise coupling to CoASH, reduction of NAD^+ and release of CO_2.[67,68]

Another metabolic pathway that generates Ac-CoA is the breakdown of fatty acids, known as β-oxidation.[69–72] Initially, fatty acids are activated by conversion to the respective acyl-CoA thioesters. The responsible enzymes are called acyl-CoA synthetases and are classified according to their specificity for a certain fatty acid chain-length.[73,74] Afterwards, the CoA esters enter the mitochondrial matrix by simple diffusion (short-chain) or *via* a translocase (long-chain).[75] Each β-oxidation cycle involves four enzymatic procedures: dehydrogenation of acyl-CoA to 2-*trans*-enoyl-CoA, hydration of the double bond, oxidation to β-ketoacyl-CoA and, finally, thiolytic cleavage, affording Ac-CoA and a fatty acyl-CoA shorter by two carbon atoms. β-Oxidation can also take place in peroxisomes where long-chain and very long-chain acyl-CoA fatty esters are degraded, producing H_2O_2.[76]

In addition to carbohydrates and lipids, several amino acids, originating from protein degradation, serve as precursors of Ac-CoA.[65,77] In the case of glucogenic amino acids (*e.g.* alanine, cysteine or serine), when deaminated, their carbon skeletons are enzymatically converted to α-ketoacids (*e.g.* pyruvate) or other intermediates of the TCA cycle, leading to Ac-CoA *via* additional reactions. In this way, they supply gluconeogenesis, especially when glucose levels are very low. On the other hand, ketogenic amino acids (*e.g.* leucine or lysine) are broken down directly to Ac-CoA or its precursor acetoacetate leading to the formation of fatty acids or ketone bodies. Larger amino acids, such as tryptophan, fall into both categories.

Alternatively to the above catabolic pathways, cytosolic Ac-CoA can derive either from citrate shuttled to the cytoplasm[78,79] or from cytosolic acetate,[80] the end product of ethanol metabolism. Both reactions are ATP-dependent with the first being catalysed by ATP citrate lyase and the second by acetyl-CoA synthetase. Analogously, acyl-CoA synthetases (*vide supra*) couple CoA to several carboxylates (*e.g.* propionate or butyrylate) to generate acyl-CoA derivatives in the cytoplasm.

4.2 Cellular uses of acetyl-CoA

Apart from being the crossroad between glycolysis, lipolysis and proteolysis, Ac-CoA has a leading role in energy production *via* oxidation of valuable fuels in the TCA cycle in mitochondria. The cycle begins with the condensation of Ac-CoA with oxaloacetate to form citrate. Through a series of biochemical reactions, the high energy molecule ATP is produced while the two carbons of the acetyl group are released as CO_2.[81,82]

Furthermore, Ac-CoA is an essential building block of a wide variety of intracellular compounds. In the cytosol, Ac-CoA is first carboxylated by acetyl-CoA carboxylase to malonyl-CoA which is the starting point for fatty acids biosynthesis. Next, both acetyl- and malonyl-CoA get activated by attachment to an acyl carrier protein by 4′-phosphopantentheine transferase. A repeating cycle begins, resulting each time in an extension of the chain by two carbon atoms.[83] The first acid to be synthesized *de novo* is palmitic acid (C16). Enzymatic elongation of the latter generates long-chain acids, while unsaturated fatty acids are produced by several desaturases. Subsequently, the fatty acids are utilized in the formation of diacylglycerol and membrane lipids. In addition, through the mevalonate pathway,[84] Ac-CoA gives rise to highly important intermediates in the synthesis of isoprenoids, such as cholesterol, vitamins and steroid hormones.

In the nucleus, Ac-CoA serves as a cofactor for histone acetylation catalysed by HATs. This abundant post-translational modification occurs in lysine residues at the N-terminal regions of histones and results in transcriptionally active chromatin.[85,86] The reverse process is catalysed by histone deacetylases (HDACs) through hydrolytic removal of the acetyl group. Next to histones, hyperactivity of lysine acetyltransferases (KATs) on other proteins has been shown to lead to aberrant gene expression and pathological disorders.[10] Ac-CoA is also linked to N-terminal acetylation which is a co-translational protein modification with major impact on protein stability, degradation and localization[87] as well as oncogenesis.[88] Apart from acetylations, recent studies report the existence of other acylations such as histone propionylation and butyrylation.[7]

From keeping the balance between sugar and fat metabolism, driving the synthesis of vital biomolecules and regulating gene expression, Ac-CoA can be considered as a key metabolite for normal protein function and cell survival. As a result, a number of synthetic reporters that mimic endogenous Ac-CoA precursors have been developed with the aim of monitoring a plethora of fundamental biological processes.

5. Metabolic labelling of protein acylations *via* acyl-CoA chemical reporters

From the examination of the precursors of Ac-CoA, one can envision that a more comprehensive view of protein acylation can be obtained by the use of functionalized analogues of pyruvate, citrate, acetate, long-chain fatty acids, amino acids or Ac-CoA itself. Regarding the latter, various CoA-like functionalized reporters have been used in the study of protein acylation *in vitro*.[43,89,90] Such molecules are recognized and used directly as cofactors by several acyltransferases. For instance, Yang *et al.* managed to evaluate *in vitro* the optimal synthetic cofactors of different mutated members of the KAT family.[89] Nevertheless, a majority of the acyl-CoA reporters involve the use of functionalized carboxylates, due to their easier synthetic route and smaller size (Fig. 2). Inside a cell, these molecules are first converted to their corresponding acyl-CoA derivatives and then used in acylation reactions by the respective enzymes. It has to be noted though that in order to compete with the native compounds, cells were frequently starved before treatment with the reporters at subtoxic but, still, high concentrations (up to the mM range). Moreover, in some cases, inhibitors of the endogenous enzymes that can remove the synthetic labels were co-incubated in order to enhance the labelling efficiency. It is hence conclusive that, under these conditions, a true representation and quantification of cellular protein acylation is rather challenging.

5.1 Short-chain fatty acid reporters

Metabolic labelling with short-chain acidic reporters, in combination with bioorthogonal techniques, has become an appealing approach in the study of protein acylation.[20,91] For example, incubation with terminal alkynyl-acetate analogues (C4-6) by Yang *et al.* resulted in successful detection of acylated proteins in cell lysates and histone extracts, when the reporters were used at concentrations higher than 2.5 mM for 6–8 h.[43] The highest signal came from the C5 precursor and further proteomic analysis of 4-pentynoyl-labelled histones revealed specific histone acylation sites. Analogously, our group employed 4-pentenoate and its methyl and thioester derivatives as an equally effective alternative to the triple bond-carrying reporters.[57] After treatment with 5 and 10 mM of each reporter and coupling to a biotinylated phenylboronic acid *via* the aqueous oxidative Heck reaction,[56] a luminescent signal of predominantly histone H3 was observed only in the case of the free 4-pentenoic acid. Moreover, inhibition studies clearly showed the effect of known HAT inhibitors on histone acylation patterns compared to the commercially available anti-acetyl-lysine antibody. Finally, lysine malonylation was recently monitored by the introduction of an alkynyl diacetoxy-methyl pro-drug that, intracellularly, releases the corresponding malonyl analogue. Subsequent "click" coupling enabled fluorescent imaging of the malonylation marks.[92]

5.2 Long-chain fatty acid reporters

The linkage of fatty acids to proteins and their further assembly into membranes is achieved by two basic types of

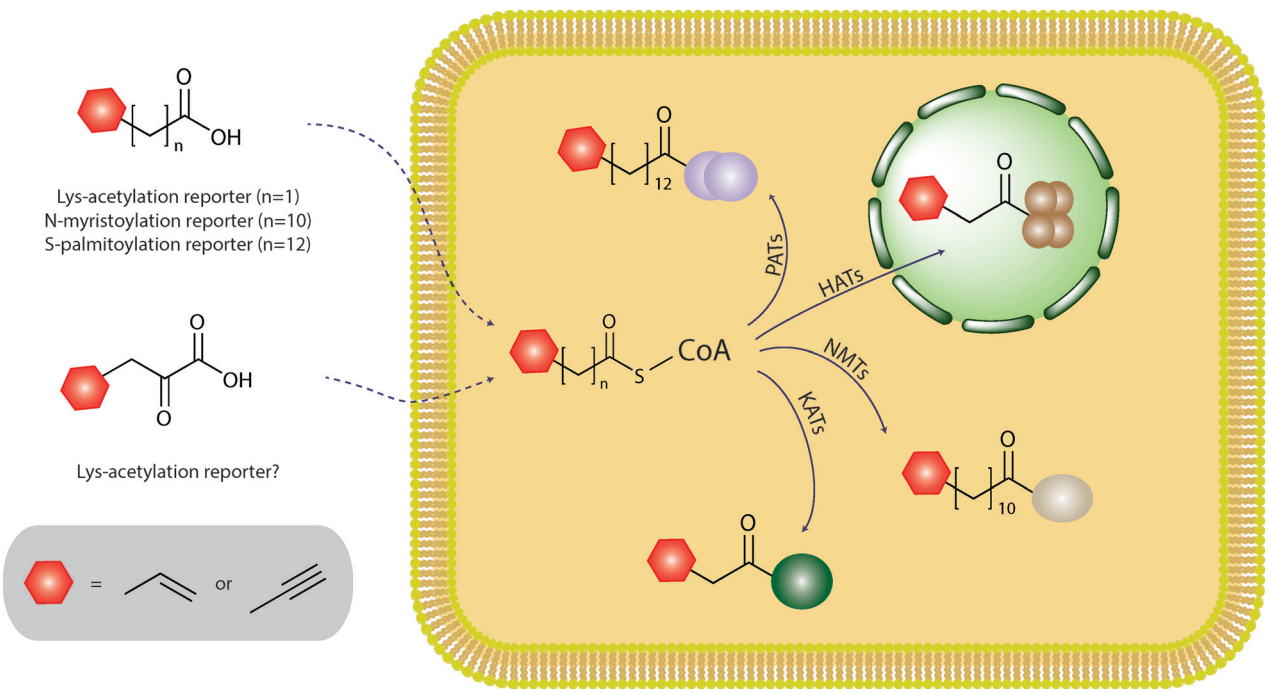


Fig. 2 Metabolic labelling of cellular protein acylations. Carboxylic acids functionalized with terminal alkynes or alkenes can be used as acyl-CoA precursors to be enzymatically incorporated into protein residues. HATs; histone acetyltransferases, NMTs; *N*-myristoyltransferases, PATs; *S*-palmitoyltransferases, KATs; lysine acetyltransferases.

lipidation: palmitoylation and myristoylation.[93] *S*-Palmitoylation, catalysed by palmitoyltransferases (PATs), is based on the post-translational covalent attachment of palmitate to cysteine residues *via* a thioester bond. In the case of *N*-myristoylation, *N*-myristoyltransferases (NMTs) co- or post-translationally transfer a myristoyl unit (C14) to N-terminal glycine residues, forming an amide bond.

Over the past few years, the use of fatty acids modified with bioorthogonal chemical reporters has provided a more detailed insight in the mechanisms that drive lipidation.[6,94,95] The first probes of fatty acylation were synthesized by Hang *et al.* as ω-azido fatty acids (C12-16). After cellular incubation with 0.1–200 μM concentrations, these analogues were effectively introduced in the metabolic pathway and converted to their respective CoA derivatives by acyl-CoA synthetase (*vide supra*). On-blot imaging of the *N*-myristoylated (C12) and *S*-palmitoylated (C14-16) proteins was then possible after biotinylation with the Staudinger ligation.[96] Similar functionalized myristic analogues were also employed, allowing fast and sensitive identification of proteins subjected to *N*-myristoylation during cell apoptosis.[97] Recently, Colquhoun *et al.* identified 27 myristoylated and 45 palmitoylated proteins in HIV-1 infected cells after treatment with myristic and palmitic azides, respectively, using "click" chemistry. The results were confirmed with sample enrichment strategies and mass spectrometric analysis.[98]

Next to the azide-carrying probes, an increasing preference for "click" over the Staudinger reaction has led to the development of different classes of ω-alkynyl-fatty acids.[99–102] Varying the chain length, these mimics were successfully incorporated in endogenous targets as the natural fatty acyl-CoA.

As expected, experiments in human and mouse cell lines demonstrated that, according to their sensitivity to alkali, the reporters with 10, 11, 16 and 18 carbons showed some specificity over the *S*-palmitoylation, whereas C14 probes were mainly coupled to proteins *via* an amide bond, serving *N*-myristoylation.[99] Furthermore, studies with the C16 reporter, gave a general idea of the dynamics of protein fatty acylation.[99] A major improvement in sensitivity was obtained when Charron *et al.* presented a robust method based on in-gel fluorescent imaging after coupling of alkyne-labelled acylated proteins with rhodamine-azide.[100] More recently, a clearer picture of the myristoylation pattern of human cells and zebrafish was obtained by the use of an alkynyl myristic analogue.[103] Lastly, the introduction of a photo-crosslinking group to an alkyne-functionalized palmitic probe empowered not only the labelling, but also the investigation of the interactions of an antiviral protein with other proteins.[104]

5.3 Experimental work with α-keto acids

As mentioned above, a direct source of Ac-CoA is pyruvate, formed during aerobic glycolysis. Previously, isotope-labelled α-keto acids, serving as metabolic building blocks of several amino acids, were used to study the structure of high-molecular weight proteins by high-resolution NMR spectroscopy.[105–107] Inspired by this, we set out to employ functionalized pyruvic acid mimics in an attempt to gain a deeper understanding of protein modifications where Ac-CoA plays a defining role. However, despite our expectations, our efforts proved fruitless as no tagging of protein acylation was achieved. More specifically, for our purposes, we synthesized the olefinic reporter 2-oxo-hex-5-enoic acid and its ethyl ester, which could function as a pro-drug (see the ESI†). RAW264.7 cells were treated for 6 h with 12.5 mM of each of these analogues and biotinylation was performed *via* the oxidative Heck reaction. Unfortunately, no luminescence was observed in either the cell lysates or the extracted core of histones, apart from a signal corresponding to endogenously biotinylated proteins (Fig. S1†). Possible explanations could be a low incorporation rate due to competition with natural pyruvate, or a more favourable catabolic pathway by other intracellular enzymes. It is then possible that these reporters are not recognized by PDH and, hence, their acyl tail is not incorporated into CoA.

Outlook

It goes without saying that metabolic labelling techniques and subsequent bioorthogonal chemistries have considerably contributed in the investigation of protein modifications. Herein, we described various functionalized chemical reporters and site-specific cross-coupling methods that enabled the visualization of a surprisingly large number of acylated proteins involved in diseases. However, we also shape the challenges associated with this two-step labelling approach. The need for high concentrations of chemical reporters reveals the hobbling tagging efficiency in the cellular physiological environment. Thus, metabolic labelling can serve as a tool useful only for identification and not for quantification of endogenous protein acylation. Another important point is the extent to which enzymes involved in metabolism and acylation show substrate promiscuity. KATs, for instance, show remarkable promiscuity towards propionyl and butyryl substrates.[108] While this lack of substrate specificity has been exploited to study protein acetylations, the procedure can ultimately not distinguish between acetylation, propionylation and butyrylation. High throughput mass spectrometric analysis is therefore needed to quantify these kind of post-translational modifications.[109,110] Contrarily, as we have shown in the previous section, enzymes catalysing post-translational modifications with long-chain fatty acids display a higher degree of substrate specificity but are still sufficiently loose to allow incorporation of small bioorthogonal functional groups. In conclusion, when devising an experiment to investigate protein acylation, one should bear the advantages and disadvantages of metabolic labelling and of available bioorthogonal reactions in mind. Thus, there is an imperative request for serious improvements in order to gain clear insights in metabolic processing of acyl-CoA and, eventually, a better understanding of the role of protein acylation in disease processes.

Acknowledgements

This work was financially supported by a VIDI grant (016.122.302) from the Netherlands Organization for Scientific Research and an ERC starting grant (309782) from the European Union to F. J. D.

References

1 J. M. Baskin and C. R. Bertozzi, *QSAR Comb. Sci.*, 2007, **26**, 1211–1219.
2 M. D. Best, *Biochemistry*, 2009, **48**, 6571–6584.
3 J. A. Prescher and C. R. Bertozzi, *Nat. Chem. Biol.*, 2005, **1**, 13–21.
4 E. M. Sletten and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2009, **48**, 6974–6998.
5 C. Choudhary, B. T. Weinert, Y. Nishida, E. Verdin and M. Mann, *Nat. Rev. Mol. Cell Biol.*, 2014, **15**, 536–550.
6 M. Grammel and H. C. Hang, *Nat. Chem. Biol.*, 2013, **9**, 475–484.
7 Y. Chen, R. Sprung, Y. Tang, H. Ball, B. Sangras, S. C. Kim, J. R. Falck, J. Peng, W. Gu and Y. Zhao, *Mol. Cell. Proteomics*, 2007, **6**, 812–819.
8 F. J. Dekker and H. J. Haisma, *Drug Discovery Today*, 2009, **14**, 942–948.
9 F. J. Dekker, T. van Den Bosch and N. I. Martin, *Drug Discovery Today*, 2014, **19**, 654–660.
10 M. Ghizzoni, H. J. Haisma, H. Maarsingh and F. J. Dekker, *Drug Discovery Today*, 2011, **16**, 504–511.
11 X. J. Yang, *Nucleic Acids Res.*, 2004, **32**, 959–976.
12 M. H. Wright, W. P. Heal, D. J. Mann and E. W. Tate, *J. Mater. Chem. A*, 2010, **3**, 19–35.
13 A. Tashiro, S. Shoji and Y. Kubota, *Biochem. Biophys. Res. Commun.*, 1989, **165**, 1145–1154.
14 C. Aicart-Ramos, R. A. Valero and I. Rodriguez-Crespo, *Biochim. Biophys. Acta, Biomembr.*, 2011, **1808**, 2981–2994.
15 T. Iwanaga, R. Tsutsumi, J. Noritake, Y. Fukata and M. Fukata, *Prog. Lipid Res.*, 2009, **48**, 117–127.
16 R. O. Brady and S. Gurin, *J. Biol. Chem.*, 1950, **186**, 461–469.
17 H. Singh, M. Brogan, D. Johnson and A. Poulos, *J. Lipid Res.*, 1992, **33**, 1597–1605.
18 L. Rivière, P. Moreau, S. Allmann, M. Hahn, M. Biran, N. Plazolles, J.-M. Franconi, M. Boshart and F. Bringaud, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 12694–12699.
19 W. Gu and R. G. Roeder, *Cell*, 1997, **90**, 595–606.
20 Y. Y. Yang and H. C. Hang, *ChemBioChem*, 2011, **12**, 314–322.
21 M. Mann and O. N. Jensen, *Nat. Biotechnol.*, 2003, **21**, 255–261.
22 K. Webley, J. A. Bond, C. J. Jones, J. P. Blaydes, A. Craig, T. Hupp and D. Wynford-Thomas, *Mol. Cell. Biol.*, 2000, **20**, 2803–2808.
23 B. Dai, F. Dahmani, J. A. Cichocki, L. C. Swanson and T. P. Rasmussen, *J. Visualized Exp.*, 2011, 1–4.
24 A. M. Silverstein, *Nat. Immunol.*, 2004, **5**, 1211–1217.
25 M. A. Helsby, J. R. Fenn and A. D. Chalmers, *F1000Research*, 2013, **2014**, 1–13.
26 E. Saxon and C. R. Bertozzi, *Science*, 2000, **287**, 2007–2010.
27 H. Staudinger and J. Meyer, *Helv. Chim. Acta*, 1919, **2**, 612–618.
28 F. L. Lin, H. M. Hoyt, H. van Halbeek, R. G. Bergman, C. R. Bertozzi and L. Berkeley, *J. Am. Chem. Soc.*, 2005, **127**, 2686–2695.
29 D. M. Patterson, L. A. Nazarova and J. A. Prescher, *ACS Chem. Biol.*, 2014, **9**, 592–605.
30 R. Huisgen, *Angew. Chem., Int. Ed. Engl.*, 1963, **2**, 565–598.
31 C. W. Tornøe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057–3064.
32 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596–2599.
33 Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless and M. G. Finn, *J. Am. Chem. Soc.*, 2003, **125**, 3192–3193.
34 O. Reihl, M. O. Lederer and W. Schwack, *Carbohydr. Res.*, 2004, **339**, 483–491.
35 S. I. Presolski, V. P. Hong and M. G. Finn, *Curr. Protoc. Chem. Biol.*, 2011, **3**, 153–162.
36 T. R. Chan, R. Hilgraf, K. B. Sharpless and V. V. Fokin, *Org. Lett.*, 2004, **6**, 2853–2855.
37 L. M. Gaetke and C. K. Chow, *Toxicology*, 2003, **189**, 147–163.
38 K. Sonogashira, *J. Organomet. Chem.*, 2002, **653**, 46–49.
39 K. Kodama, S. Fukuzawa, H. Nakayama, K. Sakamoto, T. Kigawa, T. Yabuki, N. Matsuda, M. Shirouzu, K. Takio, S. Yokoyama and K. Tachibana, *ChemBioChem*, 2007, **8**, 232–238.
40 N. Li, R. K. V. Lim, S. Edwardraja and Q. Lin, *J. Am. Chem. Soc.*, 2011, **133**, 15316–15319.
41 J. Li, S. Lin, J. Wang, S. Jia, M. Yang, Z. Hao, X. Zhang and P. R. Chen, *J. Am. Chem. Soc.*, 2013, **135**, 7330–7338.
42 P. Tsiveriotis and N. Hadjiliadis, *Coord. Chem. Rev.*, 1999, **190–192**, 171–184.
43 Y. Y. Yang, J. M. Ascano and H. C. Hang, *J. Am. Chem. Soc.*, 2010, **132**, 3640–3641.
44 A. L. Blobaum, *Drug Metab. Dispos.*, 2006, **34**, 1–7.
45 S. Shak, N. O. Reich, I. M. Goldstein and P. R. Ortiz de Montellano, *J. Biol. Chem.*, 1985, **260**, 13023–13028.
46 A. P. Zou, Y. H. Ma, Z. H. Sui, P. R. Ortiz de Montellano, J. E. Clark, B. S. Masters and R. J. Roman, *J. Pharmacol. Exp. Ther.*, 1994, **268**, 474–481.
47 C. D. Hein, X.-M. Liu and D. Wang, *Pharm. Res.*, 2008, **25**, 2216–2230.
48 R. Ekkebus, S. I. van Kasteren, Y. Kulathu, A. Scholten, I. Berlin, P. P. Geurink, A. de Jong, S. Goerdayal, J. Neefjes, A. J. R. Heck, D. Komander and H. Ovaa, *J. Am. Chem. Soc.*, 2013, **135**, 2867–2870.
49 K. L. Killops, L. M. Campos and C. J. Hawker, *J. Am. Chem. Soc.*, 2008, **130**, 5062–5064.
50 A. B. Lowe, *Polym. Chem.*, 2010, **1**, 17.
51 T. Mizoroki, K. Mori and A. Ozaki, *Bull. Chem. Soc. Jpn.*, 1971, **44**, 581–581.
52 R. F. Heck and J. P. Nolley, *J. Org. Chem.*, 1972, **37**, 2320–2322.

53 C. S. Cho and S. Uemura, *J. Organomet. Chem.*, 1994, **465**, 85–92.
54 K. S. Yoo, C. H. Yoon and K. W. Jung, *J. Am. Chem. Soc.*, 2006, **128**, 16384–16393.
55 A. L. Gottumukkala, J. F. Teichert, D. Heijnen, N. Eisink, S. van Dijk, C. Ferrer, A. van den Hoogenband and A. J. Minnaard, *J. Org. Chem.*, 2011, **76**, 3498–3501.
56 M. E. Ourailidou, J. Y. van der Meer, B. J. Baas, M. Jeronimus-Stratingh, A. L. Gottumukkala, G. J. Poelarends, A. J. Minnaard and F. J. Dekker, *ChemBioChem*, 2014, **15**, 209–212.
57 M. E. Ourailidou, P. Dockerty, M. Witte, G. J. Poelarends and F. J. Dekker, *Org. Biomol. Chem.*, 2015, **13**, 3648–3653.
58 Y. Wang, C. I. Rivera Vera and Q. Lin, *Org. Lett.*, 2007, **9**, 4155–4158.
59 W. Song, Y. Wang, J. Qu, M. M. Madden and Q. Lin, *Angew. Chem., Int. Ed.*, 2008, **47**, 2832–2835.
60 W. Song, Y. Wang, J. Qu and Q. Lin, *J. Am. Chem. Soc.*, 2008, **130**, 9654–9655.
61 W. Siti, A. K. Khan, H.-P. M. de Hoog, B. Liedberg and M. Nallani, *Org. Biomol. Chem.*, 2015, **13**, 3202–3206.
62 Y. Zhang, W. Liu and Z. K. Zhao, *Molecules*, 2013, **19**, 306–315.
63 P. An, Z. Yu and Q. Lin, *Org. Lett.*, 2013, **15**, 5496–5499.
64 T. Vorapreeda, C. Thammarongtham, S. Cheevadhanarak and K. Laoteng, *Microbiology*, 2012, **158**, 217–228.
65 G. Mariño, F. Pietrocola, T. Eisenberg, Y. Kong, S. Malik, A. Andryushkova, S. Schroeder, T. Pendl, A. Harger, M. Niso-Santano, N. Zamzami, M. Scoazec, S. Durand, D. Enot, Á. Fernández, I. Martins, O. Kepp, L. Senovilla, C. Bauvy, E. Morselli, E. Vacchelli, M. Bennetzen, C. Magnes, F. Sinner, T. Pieber, C. López-Otín, M. Maiuri, P. Codogno, J. Andersen, J. Hill, F. Madeo and G. Kroemer, *Mol. Cell*, 2014, **53**, 710–725.
66 J. T. Pronk, H. Y. Steensma and J. P. van Dijken, *Yeast*, 1996, **12**, 1607–1633.
67 M. S. Patel and T. E. Roche, *FASEB J.*, 1990, **4**, 3224–3233.
68 L. J. Reed, *Acc. Chem. Res.*, 1974, **7**, 40–46.
69 G. R. Drysdale and H. A. Lardy, *J. Biol. Chem.*, 1953, **202**, 119–136.
70 S. J. Wakil, D. E. Green, S. Mii and H. R. Mahler, *J. Biol. Chem.*, 1954, **207**, 631–638.
71 J. R. Stern, A. del Campillo and I. Raw, *J. Biol. Chem.*, 1956, **218**, 971–983.
72 F. Lynen and S. Ochoa, *Biochim. Biophys. Acta*, 1953, **12**, 299–314.
73 S. Eaton, K. Bartlett and M. Pourfarzam, *Biochem. J.*, 1996, **320**(Pt 2), 345–357.
74 H. Schulz, *Biochem. Lipids, Lipoproteins Membr.*, 2008, 131–154.
75 J. Bremer, *Physiol. Rev.*, 1983, **63**, 1420–1480.
76 J. K. Reddy and T. Hashimoto, *Perspective*, 2001, 193–230.
77 A. E. Harper, R. H. Miller and K. P. Block, *Annu. Rev. Nutr.*, 1984, **4**, 409–454.
78 B. L. Fatland, J. Ke, M. D. Anderson, W. I. Mentzen, L. W. Cui, C. C. Allred, J. L. Johnston, B. J. Nikolau and E. S. Wurtele, *Plant Physiol.*, 2002, **130**, 740–756.
79 M. J. Hynes and S. L. Murray, *Eukaryotic Cell*, 2010, **9**, 1039–1048.
80 B. Schwer, J. Bunkenborg, R. O. Verdin, J. S. Andersen and E. Verdin, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 10224–10229.
81 E. Meléndez-Hevia, T. G. Waddell and M. Cascante, *J. Mol. Evol.*, 1996, **43**, 293–303.
82 H. Kornberg, *Nat. Rev. Mol. Cell Biol.*, 2000, **1**, 225–228.
83 S. J. Wakil, J. K. Stoops and V. C. Joshi, *Annu. Rev. Biochem.*, 1983, **52**, 537–579.
84 J. L. Goldstein and M. S. Brown, *Nature*, 1990, **343**, 425–430.
85 M. Grunstein, *Nature*, 1997, **389**, 349–352.
86 B. D. Strahl and C. D. Allis, *Nature*, 2000, **403**, 41–45.
87 T. Arnesen, *PLoS Biol.*, 2011, **9**, e1001074.
88 T. V. Kalvik and T. Arnesen, *Oncogene*, 2012, **32**, 269–276.
89 C. Yang, J. Mi, Y. Feng, L. Ngo, T. Gao, L. Yan and Y. G. Zheng, *J. Am. Chem. Soc.*, 2013, **135**, 7791–7794.
90 M. A. Kostiuk, M. M. Corvi, B. O. Keller, G. Plummer, J. A. Prescher, M. J. Hangauer, C. R. Bertozzi, G. Rajaiah, J. R. Falck and L. G. Berthiaume, *FASEB J.*, 2007, **22**, 721–732.
91 E. Thinon and H. C. Hang, *Biochem. Soc. Trans.*, 2015, **43**, 253–261.
92 X. Bao, Q. Zhao, T. Yang, Y. M. E. Fung and X. D. Li, *Angew. Chem., Int. Ed.*, 2013, **52**, 4883–4886.
93 J. E. Smotrys and M. E. Linder, *Annu. Rev. Biochem.*, 2004, **73**, 559–587.
94 R. N. Hannoush and J. Sun, *Nat. Chem. Biol.*, 2010, **6**, 498–506.
95 H. C. Hang, J. P. Wilson and G. Charron, *Acc. Chem. Res.*, 2011, **44**, 699–708.
96 H. C. Hang, E. Geutjes, G. Grotenbreg, M. Pollington, M. J. Bijlmakers, H. L. Ploegh, M. Cells and A. M. Pollington, *J. Am. Chem. Soc.*, 2007, **129**, 2744–2745.
97 D. D. O. Martin, G. L. Vilas, J. A. Prescher, G. Rajaiah, J. R. Falck, C. R. Bertozzi and L. G. Berthiaume, *FASEB J.*, 2008, **22**, 797–806.
98 D. R. Colquhoun, A. E. Lyashkov, C. U. Mohien, V. N. Aquino, B. T. Bullock, R. R. Dinglasan, B. J. Agnew and D. R. M. Graham, *Proteomics*, 2015, **15**, 2066–2077.
99 R. N. Hannoush and N. Arenas-Ramirez, *ACS Chem. Biol.*, 2009, **4**, 581–587.
100 G. Charron, M. M. Zhang, J. S. Yount, J. Wilson, A. S. Raghavan, E. Shamir and H. C. Hang, *J. Am. Chem. Soc.*, 2009, **131**, 4967–4975.
101 M. C. Yap, M. A. Kostiuk, D. D. O. Martin, M. A. Perinpanayagam, P. G. Hak, A. Siddam, J. R. Majjigapu, G. Rajaiah, B. O. Keller, J. A. Prescher, P. Wu, C. R. Bertozzi, J. R. Falck and L. G. Berthiaume, *J. Lipid Res.*, 2010, **51**, 1566–1580.
102 R. N. Hannoush, *Curr. Protoc. Chem. Biol.*, 2011, **3**, 15–26.
103 M. Broncel, R. A. Serwa, P. Ciepla, E. Krause, M. J. Dallman, A. I. Magee and E. W. Tate, *Data Brief*, 2015, **4**, 379–383.
104 T. Peng and H. C. Hang, *J. Am. Chem. Soc.*, 2015, **137**, 556–559.

105 R. J. Lichtenecker, N. Coudevylle, R. Konrat and W. Schmid, *ChemBioChem*, 2013, **14**, 818–821.
106 R. J. Lichtenecker, K. Weinhäupl, W. Schmid and R. Konrat, *J. Biomol. NMR*, 2013, **57**, 327–331.
107 R. J. Lichtenecker, *Org. Biomol. Chem.*, 2014, **12**, 7551–7560.
108 B. N. Albaugh, K. M. Arnold and J. M. Denu, *ChemBioChem*, 2011, **12**, 290–298.
109 L.-M. P. Britton, M. Gonzales-Cope, B. M. Zee and B. A. Garcia, *Expert Rev. Proteomics*, 2011, **8**, 631–643.
110 M. D. Plazas-Mayorca, B. M. Zee, N. L. Young, I. M. Fingerman, G. LeRoy, S. D. Briggs and B. A. Garcia, *J. Proteome Res.*, 2009, **8**, 5367–5374.

