Bioorthogonal pro-metabolites for profiling short chain fatty acylation†

Wilson R. Sinclair,†a Jonathan H. Shrimp,a Thomas T. Zengeya,a Rhushikesh A. Kulkarni,a Julie M. Garlick,a Hans Luecke,b Andrew J. Worth,c Ian A. Blair,c Nathaniel W. Snyder d and Jordan L. Meier*eda

Short chain fatty acids (SCFAs) play a critical role in health and disease. One function of these signaling molecules is to serve as precursors for short chain fatty acylation, a class of metabolically-derived posttranslational modifications (PTMs) that are established by lysine acetyltransferases (KATs) and lysine deacetylases (KDACs). Via this mechanism, short chain fatty acylation serves as an integrated reporter of metabolism as well as KAT and KDAC activity, and has the potential to illuminate the role of these processes in disease. However, few methods to study short chain fatty acylation exist. Here we report a bioorthogonal pro-metabolite strategy for profiling short chain fatty acylation in living cells. Inspired by the dietary component tributyrin, we synthesized a panel of ester-caged bioorthogonal short chain fatty acids. Cellular evaluation of these agents led to the discovery of an azido-ester that is metabolized to its cognate acyl-coenzyme A (CoA) and affords robust protein labeling profiles. We comprehensively characterize the metabolic dependence, toxicity, and histone deacetylase (HDAC) inhibitor sensitivity of these bioorthogonal pro-metabolites, and apply an optimized probe to identify novel candidate protein targets of short chain fatty acids in cells. Our studies showcase the utility of bioorthogonal pro-metabolites for unbiased profiling of cellular protein acylation, and suggest new approaches for studying the signaling functions of SCFAs in differentiation and disease.

Introduction

Short chain fatty acids (SCFAs) play a critical role as mediators of health and disease. For example, butyrate stimulates differentiation in many cellular models of cancer and displays context-dependent effects on tumorigenesis and immune cell function in whole organisms. In addition to their role in acetyl-coenzyme A (CoA) biosynthesis via the β-oxidation pathway and activity as histone deacetylase (HDAC) inhibitors, SCFAs also constitute the requisite precursor for protein short chain fatty acylation. This family of posttranslational modifications (PTMs) derives from the metabolism of SCFAs to acyl-CoAs, which may then be used as cofactors by lysine acetyltransferase (KAT) enzymes (Fig. 1). Due to its unique reliance on both metabolism and KAT activity, studies of this PTM have significant potential to provide insights into the role of these pathways in disease. Indeed, recent studies of short chain fatty acylation in histones have found these PTMs are dynamic and are associated with distinct effects on transcription relative to acetylation. However, a limitation to our understanding of this process is the fact that relatively few methods for studying short chain fatty acylation exist. Antibody-based analyses of these modifications are limited by the poor affinity and selectivity of antibodies generated against marks

Fig. 1 Bioorthogonal pro-metabolite strategy for profiling short chain fatty acylation.
such as butyryl- and propionyl-lysin. This may be why these approaches have, to date, identified only a handful of short chain fatty acylated targets outside of histones.\textsuperscript{22–24} To address the limitations of antibody-based methods, bioorthogonal metabolic tracing has emerged as a complementary strategy to profile PTMs in living cells.\textsuperscript{25–27} The application of this approach to study SCFA-derived acylation was pioneered by Hang and coworkers, who demonstrated that alkynyl-SCFAs could be used to introduce a latent affinity handle into SCFA-modified proteins, enabling their facile visualization and enrichment.\textsuperscript{28} A variant of this approach has also been usefully applied to study lysine malonylation.\textsuperscript{29} However, a challenge of this strategy is that bioorthogonal SCFA tracers must compete with endogenous SCFAs for metabolism and KAT utilization, potentially limiting their efficacy. Thus, the development of highly sensitive bioorthogonal reporters for profiling short chain fatty acylation remains an important goal.

A validated strategy to increase the cellular delivery of small molecules and metabolic tracers is through reversible masking of polar functional groups.\textsuperscript{30,37,35,36} Regarding SCFAs specifically, a number of groups have found butyrate’s biological effects to be increased upon masking its polar carboxylate moiety with an ester.\textsuperscript{37,38} In analogy with pro-drugs, we refer to such molecules as “pro-metabolites.”\textsuperscript{39} A prototypical SCFA pro-metabolite is tributyrin, a naturally-occurring butyrate ester that displays potent antineoplastic and chemopreventative activity.\textsuperscript{40} Inspired by tributyrin and other pro-metabolites, we hypothesized that masking the polar carboxylate of bioorthogonal SCFAs may similarly increase their potency. Such optimized agents have the potential to provide a robust reporter of short chain fatty acylation, and by proxy, metabolism and cellular KAT/KDAC activity. Towards the goal of defining acylation-dependent signaling in cancer, here we report a bioorthogonal pro-metabolite strategy for profiling short chain fatty acylation (Fig. 1). First, we systematically define the influence of SCFA-ester linkage and bioorthogonal chemotype on cellular protein labeling. Next, we demonstrate the ability of optimized agents to form bioorthogonal acyl-CoAs in cells and establish their cell-type and metabolism-specific labeling properties. Finally, we demonstrate the utility of bioorthogonal pro-metabolites to identify novel targets of SCFAs in cells. These studies highlight the power of pro-metabolite approaches to expand the scope of bioorthogonal metabolic tracing methods, and suggest novel strategies for studying the signaling role of short chain fatty acids in health and disease.

Results and discussion

Design of a bioorthogonal pro-metabolite library

To identify sensitive agents for profiling short chain fatty acylation, we focused on exploring three structural elements of bioorthogonal SCFAs (Fig. 2). First we varied the bioorthogonal reporter, reasoning azides (1–8) and alkynes (9–16) may demonstrate differential metabolic processing and detection sensitivities. Second, we varied the length of the alkyl chain between the SCFA carboxylate and bioorthogonal reporter. Previous studies have found SCFA chain length to be a crucial factor in the formation of SCFA-CoAs by acyl-CoA synthetases,\textsuperscript{41–43} and the utilization of these SCFMA-CoAs by KATs.\textsuperscript{44} Third, we varied the carboxylate ester moiety. Here we focused on simple ethyl esters (2, 6, 10, 14), as well as carnitine and triacylglycerol esters (Fig. 2). Carnitine esters (3, 7, 11, 15) were designed to facilitate delivery of SCFAs to the mitochondria, an organelle that shows robust acyl-CoA metabolism.\textsuperscript{45} In contrast, tributyrin-inspired triacylglycerol esters (4, 8, 12, 16) were designed to take advantage of the known increased cellular uptake of highly lipophilic molecules, as well as the ability of these esters to be specifically cleaved by endogenous triacylglycerol lipases. An additional advantage of these molecules relative to other esters is their ability to enable suprastructure delivery (three equivalents) of their cognate bioorthogonal SCFAs. This panel of pro-metabolites was readily synthesized using straightforward ester bond formation from the parent bioorthogonal carboxylates (Scheme S1†). While the majority of bioorthogonal SCFA derivatives were obtained in good yield, we found carnitine esters to be the most difficult targets due to their challenging purification. Similar optimization was required to obtain triacylglycerol SCFAs, as literature conditions yielded a complex mixture of mono-, di-, and triacylglycerides. These studies establish a systematic panel of bioorthogonal SCFAs to define the structure–activity relationship of pro-metabolites in living cells.

Evaluation of bioorthogonal pro-metabolites for \textit{in vitro} protein labeling

With these compounds in hand, we next evaluated our library for protein labeling in HepG2 human hepatocellular carcinoma cells (Fig. 3). Our choice of liver cancer as a model for evaluation of our pro-metabolites stems from the fact that the liver is a major site of SCFA metabolism\textit{in vivo}, where it utilizes
Mechanistic analysis of a lead pro-metabolite for short chain fatty acylation profiling

To better define the activity of pro-metabolite 6, we next evaluated it in a series of experiments using our fluorescent gel-based assay. Cellular labeling of proteins by 6 was dose-dependent, with significant labeling observed at concentrations as low as 0.5 mM (Fig. 4a). Analyzing the time-dependent activity of 6 found labeling to peak between 6 and 9 hours (Fig. S1d†). Of note, the reduced labeling observed at 24 hours suggests short chain fatty acylation implemented by 6 may be reversible, consistent with the known ability of many HDAC enzymes to deacetylate fatty acid-lysine derivatives.30,51 To assess the generality of pro-metabolite 6, we compared cell labeling in HepG2 to labeling in cell lines derived from a variety of tissues of origin, including lung (A549) and kidney (HEK293). Interestingly, each cell line demonstrated intense, but unique, patterns of protein labeling by 6 (Fig. 4b). This suggests the ability of 6 to function as a useful reporter of short chain fatty acylation in a number of tissues and cell types.

Next, we sought to better understand the metabolic determinants of short chain fatty acylation by 6 (Fig. 5). First, we explored whether labeling of proteins by 6 was sensitive to competition by endogenous short chain fatty acids such as acetate. Our hypothesis was that acetate may compete with 6 for acyl-CoA synthetase active sites,52 thereby reducing bio-orthogonal acyl-CoA biosynthesis and competing for sites of protein labeling (Fig. 5a). Consistent with this hypothesis, we observed decreased labeling of cellular proteins by 6 upon co-administration of acetate (Fig. 5b). Treatment of cells with fatty acid synthase (FASN) inhibitors did not substantially reduce labeling by 6 (Fig. S4†), indicating that metabolism of azidopropionate to long chain fatty acyl-CoAs (such as those used as palmitoyltransferases) is not a major mechanism of labeling. Next, we tested the effect of media glucose on labeling of azidopropionyl-CoA, consistent with the ability of 6 to form cofactors for protein short chain fatty acylation reactions (Fig. S3†). To our knowledge this represents first direct evidence of bioorthogonal acyl-CoA formation in living cells. These data, combined its facile synthesis, prompted us to further evaluate 6 as an optimized reporter of cellular short chain fatty acylation.

Mechanistic analysis of a lead pro-metabolite for short chain fatty acylation profiling

To better define the activity of pro-metabolite 6, we next evaluated it in a series of experiments using our fluorescent gel-based assay. Cellular labeling of proteins by 6 was dose-dependent, with significant labeling observed at concentrations as low as 0.5 mM (Fig. 4a). Analyzing the time-dependent activity of 6 found labeling to peak between 6 and 9 hours (Fig. S1d†). Of note, the reduced labeling observed at 24 hours suggests short chain fatty acylation implemented by 6 may be reversible, consistent with the known ability of many HDAC enzymes to deacetylate fatty acid-lysine derivatives.30,51 To assess the generality of pro-metabolite 6, we compared cell labeling in HepG2 to labeling in cell lines derived from a variety of tissues of origin, including lung (A549) and kidney (HEK293). Interestingly, each cell line demonstrated intense, but unique, patterns of protein labeling by 6 (Fig. 4b). This suggests the ability of 6 to function as a useful reporter of short chain fatty acylation in a number of tissues and cell types.

Next, we sought to better understand the metabolic determinants of short chain fatty acylation by 6 (Fig. 5). First, we explored whether labeling of proteins by 6 was sensitive to competition by endogenous short chain fatty acids such as acetate. Our hypothesis was that acetate may compete with 6 for acyl-CoA synthetase active sites,52 thereby reducing bio-orthogonal acyl-CoA biosynthesis and competing for sites of protein labeling (Fig. 5a). Consistent with this hypothesis, we observed decreased labeling of cellular proteins by 6 upon co-administration of acetate (Fig. 5b). Treatment of cells with fatty acid synthase (FASN) inhibitors did not substantially reduce labeling by 6 (Fig. S4†), indicating that metabolism of azidopropionate to long chain fatty acyl-CoAs (such as those used as palmitoyltransferases) is not a major mechanism of labeling. Next, we tested the effect of media glucose on labeling of azidopropionyl-CoA, consistent with the ability of 6 to form cofactors for protein short chain fatty acylation reactions (Fig. S3†). To our knowledge this represents first direct evidence of bioorthogonal acyl-CoA formation in living cells. These data, combined its facile synthesis, prompted us to further evaluate 6 as an optimized reporter of cellular short chain fatty acylation.

Mechanistic analysis of a lead pro-metabolite for short chain fatty acylation profiling

To better define the activity of pro-metabolite 6, we next evaluated it in a series of experiments using our fluorescent gel-based assay. Cellular labeling of proteins by 6 was dose-dependent, with significant labeling observed at concentrations as low as 0.5 mM (Fig. 4a). Analyzing the time-dependent activity of 6 found labeling to peak between 6 and 9 hours (Fig. S1d†). Of note, the reduced labeling observed at 24 hours suggests short chain fatty acylation implemented by 6 may be reversible, consistent with the known ability of many HDAC enzymes to deacetylate fatty acid-lysine derivatives.30,51 To assess the generality of pro-metabolite 6, we compared cell labeling in HepG2 to labeling in cell lines derived from a variety of tissues of origin, including lung (A549) and kidney (HEK293). Interestingly, each cell line demonstrated intense, but unique, patterns of protein labeling by 6 (Fig. 4b). This suggests the ability of 6 to function as a useful reporter of short chain fatty acylation in a number of tissues and cell types.

Next, we sought to better understand the metabolic determinants of short chain fatty acylation by 6 (Fig. 5). First, we explored whether labeling of proteins by 6 was sensitive to competition by endogenous short chain fatty acids such as acetate. Our hypothesis was that acetate may compete with 6 for acyl-CoA synthetase active sites,52 thereby reducing bio-orthogonal acyl-CoA biosynthesis and competing for sites of protein labeling (Fig. 5a). Consistent with this hypothesis, we observed decreased labeling of cellular proteins by 6 upon co-administration of acetate (Fig. 5b). Treatment of cells with fatty acid synthase (FASN) inhibitors did not substantially reduce labeling by 6 (Fig. S4†), indicating that metabolism of azidopropionate to long chain fatty acyl-CoAs (such as those used as palmitoyltransferases) is not a major mechanism of labeling. Next, we tested the effect of media glucose on labeling of azidopropionyl-CoA, consistent with the ability of 6 to form cofactors for protein short chain fatty acylation reactions (Fig. S3†). To our knowledge this represents first direct evidence of bioorthogonal acyl-CoA formation in living cells. These data, combined its facile synthesis, prompted us to further evaluate 6 as an optimized reporter of cellular short chain fatty acylation.
on short chain fatty acylation by 6. Glucose-derived acetyl-CoA may compete with SCFA-CoAs for protein acetylation. In addition, previous studies have found that limiting glucose causes increased utilization of SCFAs for energy production via fatty acid oxidation. Importantly, this bioenergetic shift requires increased biosynthesis of SCFA-CoAs, and should therefore increase labeling by 6 (Fig. 5a) indeed, we observed increased labeling of cellular proteins by 6 when cells were grown in low glucose compared to nutrient replete media (Fig. 5c). Finally, we explored the impact of acetylation dynamics on metabolic labeling. To probe the role of KATs in short chain fatty acylation, we applied a spirocyclic EP300 inhibitor scaffold recently reported in the literature (p300i; Fig. S5a†). Inhibition of EP300 reduced the acylation of some, but not all, proteins labeled by 6 (Fig. 5d). This suggests EP300 may collaborate with other enzymatic and non-enzymatic mechanisms to establish short chain fatty acylation. Due to the lack of well-validated chemical probes for other KAT enzymes, we next assessed how labeling of proteins by 6 was impacted by HDAC inhibitors. Treatment of cells with the HDAC inhibitor SAHA was found to upregulate global acetylation, while coordinately decreasing labeling of cellular proteins by 6 (Fig. 5e). This indicates active deacetylation may be required to liberate lysines for short chain fatty acylation by bioorthogonal acyl-CoAs. These studies define pro-metabolite 6 as a general reporter of SCFA metabolism and acylation.

**Profiling the cellular targets of bioorthogonal SCFAs**

Finally, we sought to demonstrate the ability of 6 to identify specific targets of short chain fatty acylation in living cells. As a proof-of-concept, we exposed human embryonic kidney cells to either a pulse of 6 or vehicle DMSO for six hours. Following isolation of cell lysates, proteins labeled by 6 were chemoselectively ligated to a biotin alkyn via Cu-catalyzed [3 + 2] cycloaddition, and enriched over streptavidin-agarose. Enriched proteins were subjected to an on-bead tryptic digest and identified by LC-MS/MS. Candidate short chain fatty acylation targets were defined as proteins enriched greater than 2-fold from cells treated with 6, relative to a vehicle control (Table S1†). Analyzing the abundance of these targets via spectral counting identified 447 low abundance proteins (1–5 spectral counts), 121 medium abundance proteins (5–10 spectral counts), and 91 highly abundant proteins (>10 spectral counts). Of note, many of these proteins, including 41% of highly abundant enriched proteins, have previously been identified as targets of short chain fatty acylation by Hang and coworkers (Table S3†). Gene ontology analysis also supported the utility of 6 as a bioorthogonal reporter, with acetylation being the most strongly enriched term (P = 1.5 × 10−153) (Table S2†). This analysis also revealed targets of 6 were enriched in nuclear proteins, consistent with the nuclear localization of many KATs (Table S3†). Of note, among proteins identified in triplicate datasets was EP300, which has previously been observed to catalyze short chain fatty acylation. EP300 may be a target of this modification via autoacylation, and follow-up studies validated the ability of 6 to label ectopically-expressed EP300 (Fig. S5b and c†). While we defer a more detailed validation and discussion of these candidate SCFA targets for future studies, overall these findings support the ability of 6 to be applied in global proteomic studies of short chain fatty acylation.

**Conclusions**

Here we have described a bioorthogonal pro-metabolite strategy for profiling cellular short chain fatty acylation. Specifically, we find that several alkylne- and azide-SCFAs function as effective protein labeling reagents when delivered as triglyceride analogues (4, 8, 12, 16), or as ethyl esters (6). Detailed characterization of lead pro-metabolite 6 found that it can form bioorthogonal acyl-CoAs in cells, is active in several cell models, and labels proteins in a manner reflective of cellular metabolism and acetylation dynamics. Our studies suggest that pro-metabolite strategies, which have previously been used to optimize incorporation of azide- and alkyne-containing glycans, may have broad utility in the delivery of bioorthogonal metabolic tracers. It is important to highlight that our work does not stand alone, but is connected to that of other
groups that have used acetoxyethyl esters, as well as O- and N-linked hexosamines to deliver SCFAs to cells. The ability of these scaffolds to deliver bioorthogonal SCFAs has not yet been systematically determined, and in the future it may be informative to compare them to the ethyl esters and triacylglycerols investigated here. Looking forward, we anticipate several applications for the potent SCFA pro-metabolites identified in this study. First, the strong labeling afforded by suggest it may be a useful in vivo tracer for identifying endogenously short chain fatty acylated proteins associated with SCFA-driven phenotypes such as differentiation, anti-inflammation, or tumor suppression. To be most impactful, such studies will require that pro-metabolites recapitulate the phenotypic effects of endogenous SCFAs. As many of these effects have been associated with the HDAC activity of SCFAs, it is promising that our cellular studies indicate the azido- and alkynyl-SCFAs investigated here also stimulate histone acetylation (Fig. S2f). Second, bioorthogonal pro-metabolites may provide a facile readout of global KAT/KDAC activity in cellular or in vivo systems, complementing current chemical proteomic approaches, which are only applicable in cell lysates. Finally, bioorthogonal SCFA pro-metabolites may facilitate chemical genetic approaches to identify specific KAT substrates in cells.

Of note, mutations have been reported that confer KAT2A (GCN5L2) and KAT8 (MOZ) with the ability to use elongated azido- and alkynyl-CoAs. Overall, these studies represent a key methodological advance in the development of new methods to understand acetylation-dependent signaling in living systems.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgements

We thank the Laboratory of Proteomics and Analytical Technology for LC-MS/MS analyses. This work was supported by the Intramural Research Program of the NIH, National Cancer Institute (ZIA BC011488-04).

Notes and references
