

## Protein Modifications

International Edition: DOI: 10.1002/anie.201912075  
German Edition: DOI: 10.1002/ange.201912075

## Isotopically Labeled Desthiobiotin Azide (isoDTB) Tags Enable Global Profiling of the Bacterial Cysteinome\*\*

Patrick R. A. Zanon, Lisa Lewald, and Stephan M. Hacker\*

**Abstract:** Rapid development of bacterial resistance has led to an urgent need to find new druggable targets for antibiotics. In this context, residue-specific chemoproteomic approaches enable proteome-wide identification of binding sites for covalent inhibitors. Described here are easily synthesized isotopically labeled desthiobiotin azide (isoDTB) tags that shortened the chemoproteomic workflow and allowed an increased coverage of cysteines in bacterial systems. They were used to quantify 59% of all cysteines in essential proteins in *Staphylococcus aureus* and enabled the discovery of 88 cysteines that showed high reactivity, which correlates with functional importance. Furthermore, 268 cysteines that are engaged by covalent ligands were identified. Inhibition of HMG-CoA synthase was verified and will allow addressing the bacterial mevalonate pathway through a new target. Overall, a broad map of the bacterial cysteinome was obtained, which will facilitate the development of antibiotics with novel modes-of-action.

## Introduction

Infections with multidrug-resistant bacteria like methicillin-resistant *Staphylococcus aureus* (MRSA) are emerging as major threats to human health.<sup>[1]</sup> Nevertheless, very few novel classes of antibiotics have been introduced to clinics over the last decades.<sup>[1]</sup> Furthermore, almost all approved antibiotics exclusively interfere with a very limited set of bacterial targets involved in protein, nucleic acid, and cell wall biosynthesis.<sup>[1]</sup> Therefore, developing innovative methods to discover novel druggable targets for antibiotics is a pivotal task to guarantee efficient treatment of bacterial infections in the future.

Chemoproteomic approaches are extremely powerful for understanding which proteins are able to bind small molecules as ligands<sup>[2]</sup> and are particularly straightforward for covalently reactive molecules.<sup>[2a,c,d]</sup> Strikingly, covalent inhib-

itors have seen a resurgence of interest for the development of novel drugs as they can increase compound selectivity, reduce resistance formation, and target shallow protein pockets.<sup>[3]</sup> This interest has led to the recent clinical approval of several covalent kinase inhibitors.<sup>[4]</sup> Especially in the field of antibiotics, covalent inhibitors are prevalent as exemplified by  $\beta$ -lactams,<sup>[3]</sup> fosfomycin,<sup>[5]</sup> showdomycin,<sup>[6]</sup> and optimized arylomycins.<sup>[7]</sup>

Recently, methods have emerged to globally identify the exact interaction site of covalent inhibitors in a competitive fashion.<sup>[2a,b,8]</sup> In this way, many pockets that can bind covalent ligands are identified in parallel using a small library of covalently reactive molecules. This technology is especially well established for profiling cysteine residues using methods based on the isoTOP-ABPP (isotopic tandem orthogonal proteolysis activity-based protein profiling) platform (Figure 1 a).<sup>[2a]</sup> In this technology, a proteome of interest is split into two samples. One of these samples is treated with a covalent inhibitor and the other one with DMSO as a control. In the next step, both samples are treated with iodoacetamide alkyne (IA-alkyne).<sup>[9]</sup> This probe will modify many cysteines in both samples with alkynes and this reactivity will be blocked by the covalent inhibitor at its specific binding sites. The samples are next modified with isotopically labeled affinity tags using copper-catalyzed azide-alkyne cycloaddition (CuAAC).<sup>[10]</sup> The samples are combined, enriched on streptavidin beads, proteolytically digested and the modified peptides eluted for mass spectrometry (MS) based quantification. Most quantified cysteines will have ratios  $R \approx 1$  between the heavy and light channel indicating no interaction with the covalent compound (Figure 1 a). In contrast, cysteines at the specific binding sites will show ratios of  $R \gg 1$ . In this way, quantitative and site-specific interaction studies in the whole proteome are possible with unmodified covalent inhibitors that do not need to be equipped, for example, with an affinity handle.

In the last step of this protocol, the modified peptides need to be eluted from the streptavidin beads for MS-based analysis. As previous studies have utilized biotin, which binds almost irreversibly to streptavidin, as an affinity handle, various cleavable linkers have been applied to elute the peptides from the beads.<sup>[11]</sup> These linkers include those that are cleaved by proteases (Figure 1 b),<sup>[2a,9]</sup> acidic,<sup>[11b,12]</sup> or reductive conditions.<sup>[11b,13]</sup> Because of the high requirements on the orthogonality of these linkers, they need to be designed very carefully, which usually requires laborious multistep synthesis of the tags. Furthermore, the cleavage of the linker adds another step to the chemoproteomic protocol.

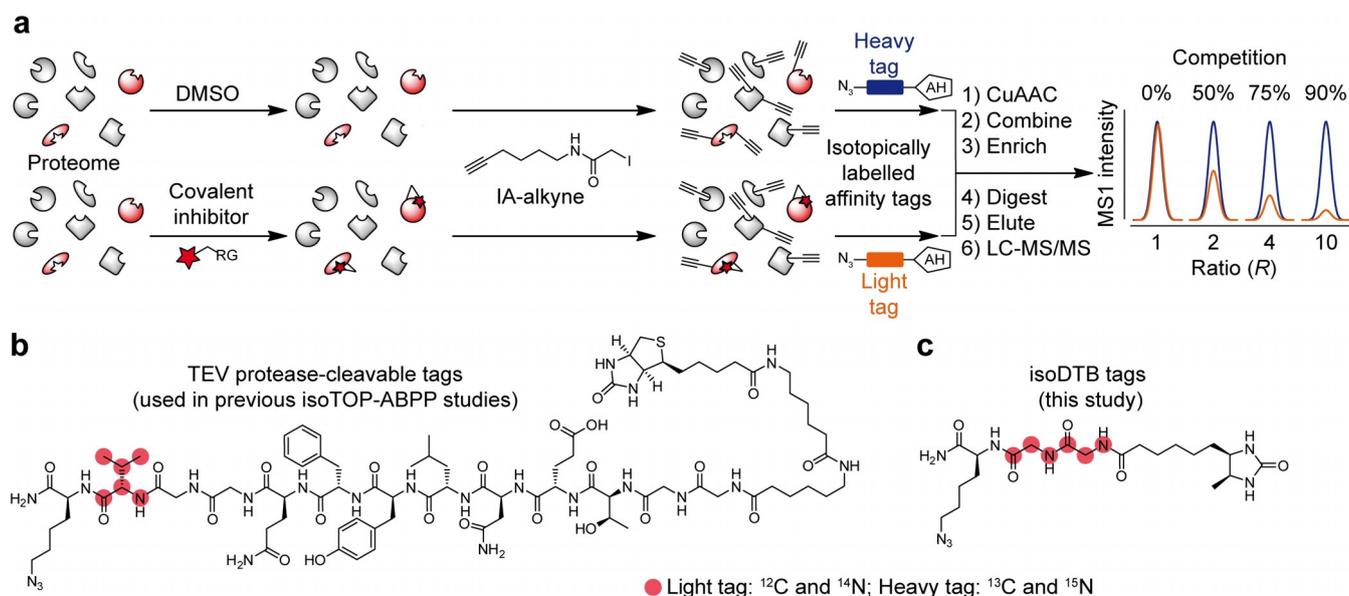
Therefore, we set out to develop isotopically labeled desthiobiotin azide (isoDTB) tags (Figure 1 c) for residue-

[\*] M. Sc. P. R. A. Zanon, M. Sc. L. Lewald, Dr. S. M. Hacker  
Department of Chemistry, Technical University of Munich  
Lichtenbergstrasse 4, 85748 Garching (Germany)  
E-mail: stephan.m.hacker@tum.de

[\*\*] A previous version of this manuscript has been deposited on a preprint server (<https://doi.org/10.26434/chemrxiv.9853445.v1>).

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:  
<https://doi.org/10.1002/anie.201912075>.

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial, and no modifications or adaptations are made.



**Figure 1.** a) Workflow for competitive, residue-specific chemoproteomic experiments.<sup>[2a]</sup> RG = reactive group, AH = affinity handle. b) Structure of the TEV protease-cleavable tags (TEV tags) originally used for residue-specific proteomics.<sup>[2a,9]</sup> c) Structure of the isoDTB tags developed in this study.

specific proteomics. As desthiobiotin still binds very strongly to streptavidin, all steps up to the proteolytic digestion can be kept the same.<sup>[14]</sup> Because of the reversibility of binding of desthiobiotin to streptavidin, in the last step, the peptides are then easily eluted using acidic conditions with acetonitrile as the cosolvent.<sup>[14]</sup> Because a complex cleavable linker is not needed, we designed these tags exclusively with two isotopically differentiated glycine residues as the linker moiety.

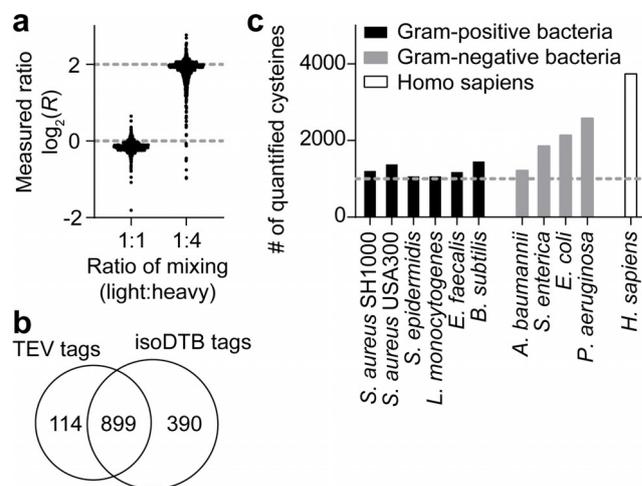
After establishing the utility of the isoDTB tags for residue-specific proteomics, we used them to globally investigate cysteines in the proteome of *S. aureus* for their reactivity and their potential to bind covalent ligands. In this way, we identified 88 highly reactive cysteines and more than 250 cysteines that can be addressed with covalent ligands. These residues are starting points for the development of antibiotics with novel modes-of-action.

## Results and Discussion

We synthesized the isoDTB tags using solid-phase peptide synthesis. For this purpose, a Rink amide resin and an Fmoc strategy were utilized. We sequentially coupled  $\epsilon$ -azido-lysine, two glycine residues, and desthiobiotin. We used glycine with the natural isotope distribution for the light isoDTB tag and glycine with two  $^{13}\text{C}$  atoms and one  $^{15}\text{N}$  atom for the heavy tag. In this way, a total mass difference between the tags of 6 Da was obtained. Purification by RP-HPLC resulted in a yield of approximately 70% for both isoDTB tags.

To establish that the tags are applicable to broadly investigate cysteines in a proteomic context, we treated two identical samples of the lysate of the methicillin-sensitive *S. aureus* (MSSA) strain SH1000<sup>[15]</sup> with 1 mM IA-alkyne and modified the two samples with the light and heavy isoDTB tag, respectively, using CuAAC. The samples were combined

either in a ratio of 1:1 or 1:4. Subsequently, we enriched the samples on streptavidin beads, digested the proteins with trypsin, and eluted the modified peptides using our straight-forward approach. Analysis using liquid chromatography coupled to tandem MS (LC-MS/MS) using a Q Exactive Plus (Thermo Fisher) mass spectrometer and evaluation using freely available MaxQuant software<sup>[16]</sup> identified 1155 cys-



**Figure 2.** a) Ratios  $R$  of all quantified cysteines in the *S. aureus* SH1000 proteome in experiments, in which the light and heavy labeled samples were both reacted with 1 mM IA-alkyne, clicked to the isoDTB tags and mixed at the indicated ratios. Expected values of  $\log_2(R)$  of 0 for the 1:1 mixture and 2 for the 1:4 mixture are indicated with dashed grey lines. b) Venn diagram comparing the number of quantified cysteines in the *S. aureus* SH1000 proteome using 1 mM IA-alkyne and the TEV tags or the isoDTB tags, respectively. c) Number of quantified cysteines in a variety of Gram-positive and Gram-negative bacteria as well as in the human cell line MDA-MB-231 using the isoDTB tags. The grey dashed line indicates 1000 quantified cysteines. All data results from duplicates.

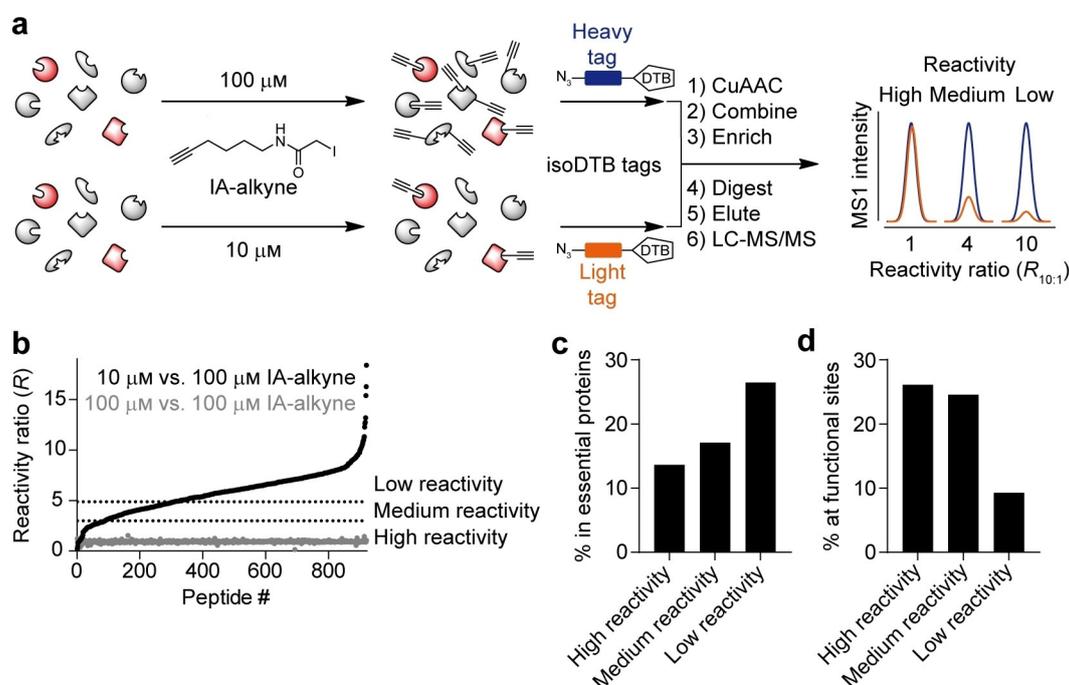
teins that were quantified for both conditions (Figure 2a; see Table S1). This analysis revealed a narrow distribution of the detected ratios for both samples around the expected values. The isoDTB tags therefore reliably allowed quantification of cysteines in the whole bacterial proteome.

We next benchmarked our technology against the TEV protease-cleavable biotin tags (TEV tags) that have been most broadly used to residue-specifically map proteomes (Figure 2b; see Table S1).<sup>[2a,9]</sup> Our isoDTB tags outperformed the TEV tags by quantifying 27% more cysteines in the *S. aureus* proteome. We increased the number of cysteines quantified with the isoDTB tags even more by additionally using chymotrypsin and AspN for the proteolytic digest (see Figure S1 and Table S1). These experiments are not possible for the TEV tags as these proteases would cleave the tag itself. In this way, we quantified a total of 1643 cysteines in the *S. aureus* proteome using the isoDTB tags in only six experiments. We next investigated the performance of the isoDTB tags with IA-alkyne in different Gram-positive and Gram-negative bacteria (Figure 2c; see Table S1), and consistently quantified more than 1000 cysteines in each strain. Moreover, we were able to quantify more than 3500 cysteines in the human cell line MDA-MB-231, which is competitive with previously described methods.<sup>[2a]</sup> Therefore, our isoDTB tags not only shortened the chemoproteomic protocol but also led to increased coverage in bacterial systems compared to the widely used TEV tag technology.

We next applied our method to analyze the reactivity of cysteines in the bacterial proteome (Figure 3a).<sup>[9a]</sup> As the reactivity of cysteines is linked to their functional relevance in human cells,<sup>[9a]</sup> we reasoned that this feature might also be

conserved in bacteria and in this way lead to the identification of functionally important cysteine residues. To study cysteine reactivity, two identical samples of the proteome of the *S. aureus* strain SH1000 were treated with either a high (100  $\mu\text{M}$ ) or a low (10  $\mu\text{M}$ ) concentration of IA-alkyne. In this way, while at the high concentration many cysteines were labeled, at the low concentration only the most reactive cysteines were labeled quantitatively. After CuAAC with the light (low concentration) and heavy isoDTB tags (high concentration), respectively, the samples were analyzed in the same way as described above. Here, high ratios ( $R_{10:1}$ ) indicate low reactivity cysteines, whereas the most reactive cysteines have  $R_{10:1} \approx 1$ . Using this procedure, we quantified 921 cysteines and identified 88 highly reactive cysteines with  $R_{10:1} < 3$  in 69 different proteins (Figure 3b; see Table S2). Another 240 cysteines showed medium reactivity ( $3 < R_{10:1} < 5$ ), whereas the remaining 593 cysteines were of low reactivity ( $R_{10:1} > 5$ ). Cysteines of all three bins of reactivity were evenly distributed throughout the different functional classes of proteins (see Figure S2).<sup>[18]</sup> Interestingly, highly reactive cysteines were depleted in essential proteins<sup>[17]</sup> in comparison to their counterparts of lower reactivity (Figure 3c). It can be speculated that evolutionary pressure has selected against highly reactive cysteines in essential proteins as these would interact with many reactive small-molecule electrophiles that occur in nature.

There is a strong enrichment of the highly and medium reactive cysteines at functional sites (Figure 3d). These cysteines include many residues that are directly involved in the catalytic mechanism (e.g. C178 in the GTP cyclohydrolase FolE2 (UniProt code Q2G0L1),<sup>[19]</sup> C112 in FabH (UniProt



**Figure 3.** a) Workflow for the measurement of cysteine reactivity using the isoDTB tags.<sup>[9a]</sup> DTB = desthiobiotin. b) Plot of the reactivity ratios ( $R_{10:1}$ ) obtained by comparing *S. aureus* SH1000 proteomes treated with high (100  $\mu\text{M}$ ) vs. low concentration (10  $\mu\text{M}$ ) of IA-alkyne (black). Ratios ( $R_{1:1}$ ) of an experiment with high concentration used for both samples (grey) are used as a control to ensure reliable quantification of all cysteines. c,d) Percentage of cysteines in the different reactivity bins that are in essential proteins (c)<sup>[17]</sup> or at functional sites (d).

code Q2FZS0), and C88 in the probable acetyl-CoA acyltransferase (UniProt code Q2G124), which are all essential proteins). Furthermore, several highly reactive cysteines are close to cofactor-binding sites (e.g. C239 of the CTP synthase PyrG (UniProt code Q2FWD1), and C45 in MnmG (UniProt code Q2FUQ3)) or metal-binding sites (e.g. C145 of alcohol dehydrogenase Adh (UniProt code Q2G0G1), and C65 of biotin synthase BioB (UniProt code Q2FVJ7)). Therefore, residue-specific proteomics using our isoDTB tags allowed global profiling of the reactivity of cysteines in the bacterial proteome and enabled the identification of functionally relevant residues.

We next set out to study which cysteines in the *S. aureus* proteome can be targeted with covalent ligands. For this purpose, we obtained a library of 211 electrophilic cysteine-directed compounds (EN001–EN211; see Table S3), mainly  $\alpha$ -chloroacetamides. These compounds were initially screened for antibacterial activity by performing minimum inhibitory concentration (MIC) experiments. While we did not expect these small compounds to be completely specific, we used this phenotypic pre-filter to prioritize compounds, whose target spectrum includes essential proteins that can be addressed in intact cells. Based on an initial screen in three MSSA strains, we selected 24 compounds (see Figure S3) based on their MIC values and structural diversity for further studies. Interestingly, many of these compounds contain a 2-aminothiazole moiety, which seems to be beneficial for activity. 23 compounds had MICs of  $\leq 100 \mu\text{M}$  in all three strains with six compounds having MICs of  $\leq 12.5 \mu\text{M}$  in all three strains (see Figure S4). Furthermore, 14 compounds showed activity (MIC  $\leq 100 \mu\text{M}$ ) in two tested MRSA strains with two compounds (EN085 and EN177) having an MIC  $\leq 10 \mu\text{M}$  in all five tested strains. This data shows that electrophilic compounds with desired biological activity could efficiently be identified from a small compound library.

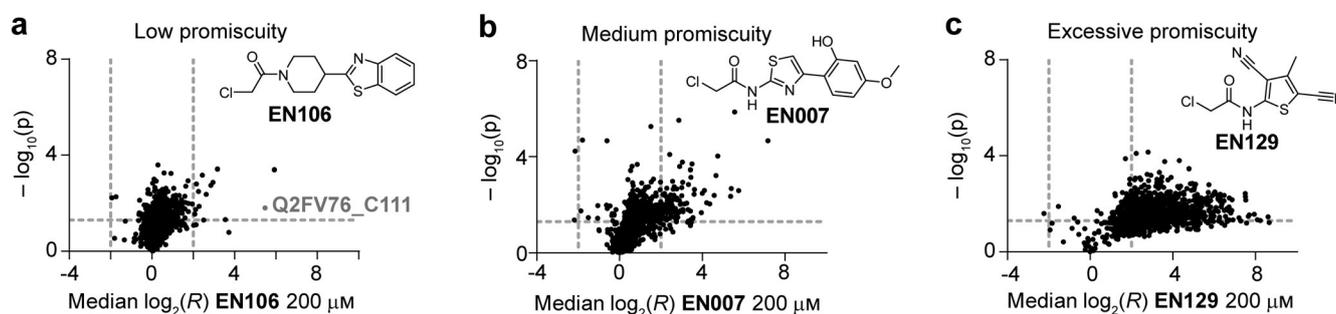
The selected 24 compounds were screened at a  $200 \mu\text{M}$  concentration in residue-specific chemoproteomic experiments using our isoDTB tags in duplicates (Figure 1a; see Figure S5 and Figure S6). For three of the compounds (EN007, EN085 and EN177), we performed an additional set of biologically independent duplicates. Given the high reproducibility between the biologically independent experi-

ments (see Figure S5), we performed the remaining profiling in duplicates and prioritized screening more compounds over performing more replicates. Five of the compounds (EN007, EN085, EN135, EN177, and EN201) that showed MIC values  $\leq 25 \mu\text{M}$  in all five tested strains were additionally tested at  $20 \mu\text{M}$  concentration (see Figure S7).

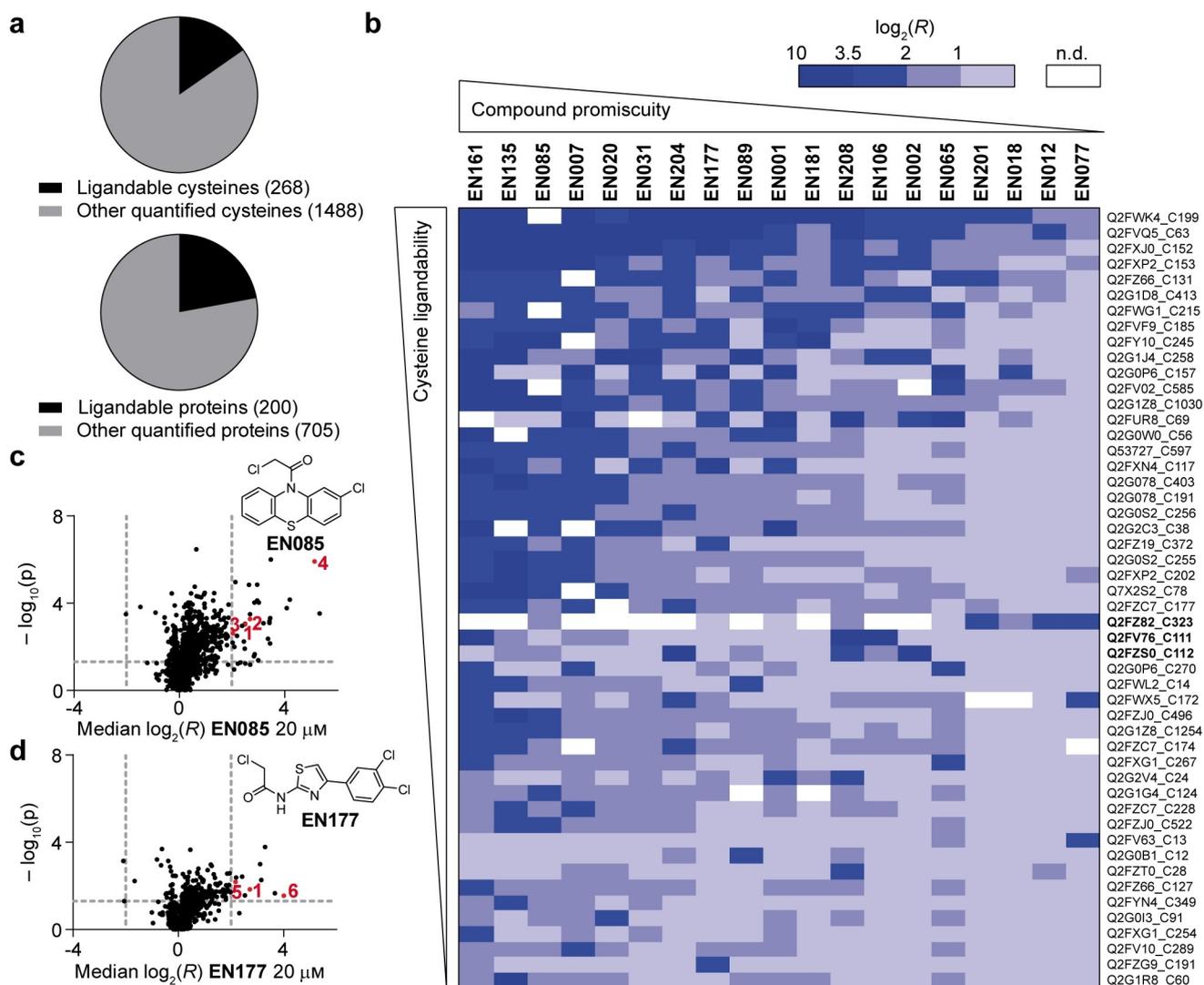
In all experiments, we consider cysteines that have a ratio of  $R > 4$  ( $\log_2(R) > 2$ ) and whose  $R$  value is statistically significantly different from  $R = 1$  ( $p$ -value  $< 0.05$  in a one-sample t-test), to be engaged by the covalent ligand. We identified a large range of values for the fraction of cysteines that are engaged by the different compounds (Figures 4a–c; see Figure S6 and Figure S8a). Nine compounds showed low promiscuity ( $< 2\%$  of all quantified cysteines are engaged, Figure 4a), ten compounds showed medium promiscuity (2% to 10%, Figure 4b), and five compounds showed excessive promiscuity ( $> 10\%$ , Figure 4c). Strikingly, no correlation between MIC and promiscuity could be observed (see Figure S8b), indicating that it is possible to identify highly active and still selective electrophiles. As we cannot rule out unspecific effects for the highly promiscuous compounds, we excluded these from all further analysis. While the low promiscuity compounds are most interesting for further compound development, the medium promiscuity compounds are most useful for the global profiling approach performed here.

Taking all 25 investigated conditions together (19 compounds at  $200 \mu\text{M}$ , five compounds at  $20 \mu\text{M}$ , and a DMSO control), we compiled a competitive data table (see Table S4), which includes all cysteines that were quantified for at least three of the conditions. In this way, we obtained information on 1756 cysteines in 905 different proteins, which corresponds to a coverage of 33% of all the cysteines encoded in the *S. aureus* genome. As cysteines in essential proteins<sup>[17]</sup> are enriched in our data over the genomic background (see Figure S9a), this equates to the quantification of 59% of all cysteines in essential proteins. Each cysteine was quantified on average for 21 of the 25 conditions (see Figure S10). Therefore, our method allowed obtaining information on many cysteine residues in *S. aureus* in a reproducible manner.

268 cysteines in 200 different proteins were engaged by at least one ligand (Figure 5a). In many proteins, we detect one



**Figure 4.** a–c) Volcano plots for three representative compounds of low (a), medium (b) and excessive promiscuity (c). Plots show the  $\log_2(R)$  of the ratio between the heavy (DMSO-treated) and light (compound-treated) channels and the  $-\log_{10}(p)$  of the statistical significance in a one-sample t-test for all quantified cysteines. In plot a) the data point for the ligandable active site residue C111 of the essential putative HMG-CoA synthase (UniProt code Q2FV76) is highlighted in grey. All data results from duplicates. For EN007 an additional set of biologically independent duplicates was included.



**Figure 5.** a) Total number of ligandable and other quantified cysteines and proteins in our competitive data table. b) Heat map of the log<sub>2</sub>(R) values for a selection of ligandable cysteines with all tested compounds at 200 μM. Cysteines discussed in the text are highlighted in boldface. Compounds are sorted from left to right by decreasing promiscuity and cysteines are sorted from top to bottom by decreasing number of identified ligands. c,d) Volcano plots for compounds **EN085** (c) and **EN177** (d) at 20 μM in *S. aureus* SH1000. Plots show the log<sub>2</sub>(R) of the ratio between the heavy (DMSO-treated) and light (compound-treated) channels and the  $-\log_{10}(p)$  of the statistical significance in a one-sample t-test for all quantified cysteines. Ligandable cysteines in essential proteins that are engaged by the respective compound are highlighted in red. 1: C152 of MurC (UniProt code Q2FXJ0), 2: C199 of MnmA (UniProt code Q2FXV6), 3: C1030 of PolC (UniProt code Q2G1Z8), 4: C151 of glyceraldehyde-3-phosphate dehydrogenase (UniProt code Q2G032), 5: C410 of pyruvate kinase (UniProt code Q2FXM9), 6: C88 of probable acetyl-CoA acyltransferase (UniProt code Q2G124). All data results from duplicates. For compounds **EN007**, **EN085** and **EN177** in panel b) and all data in panels c) and d) an additional biologically independent set of duplicates was also included.

ligandable cysteine that is engaged by several compounds, while the other cysteine(s) are never engaged (see Figure S11). This observation indicates that our method measures local target engagement of the cysteines rather than global changes to the protein structure. While ligandable cysteines are enriched in enzymes, we also identify them in other functional classes of proteins (see Figure S12).<sup>[18]</sup> Ligandable cysteines are enriched at functional sites (see Figure S13) and similarly abundant in essential proteins as compared to other quantified cysteines (see Figure S9b). When we compared this data to the cysteine reactivity data (see Figure S14), we could see that, while the highly reactive

cysteines are clearly more likely to be ligandable, there are also many ligandable cysteines of medium and low reactivity, indicating that specific noncovalent interactions are important in these cases.

Next, we looked at the binding of our covalent ligands to a selection of ligandable cysteines in more detail (Figure 5b). While the most ligandable cysteines tend to be engaged by the most promiscuous compounds, there is clear evidence for more specific interactions between less ligandable cysteines with more selective compounds. For example, the active site residue C112 of FabH (UniProt code Q2FZS0), an essential enzyme in fatty acid synthesis,<sup>[20]</sup> is exclusively targeted by

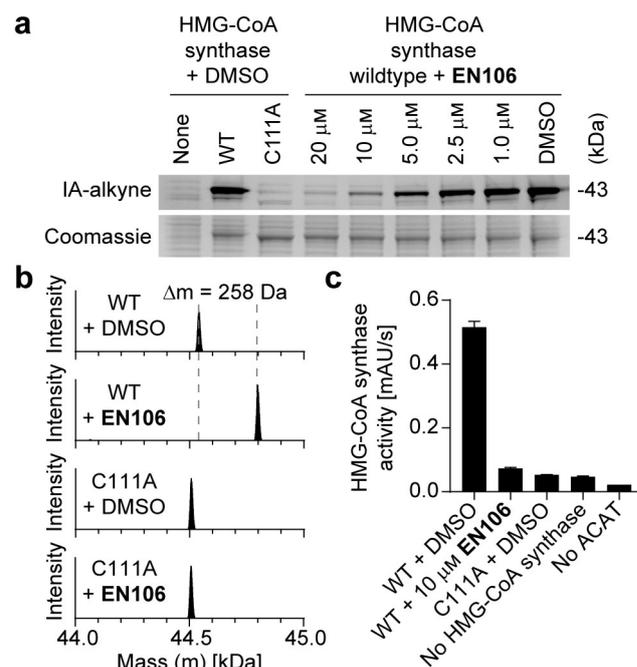
three compounds of tempered promiscuity (**EN002**, **EN204** and **EN208**, Figure 5b). This residue has previously been shown to be covalently modified for example, by the inhibitors 4,5-dichloro-1,2-dithiol-3-one and cerulenin.<sup>[21]</sup> Furthermore, the residue C323 in the isoleucine-tRNA ligase IleS (UniProt code Q2FZ82) is only targeted by another set of three compounds (**EN012**, **EN077** and **EN201**, Figure 5b), which could open up the possibility of inhibiting bacterial translation through a novel target. Overall, we detect many binding events that are strongly dependent on the compound and on the targeted cysteine indicating that our method can detect specific ligand-binding events.

Looking at the targets of the two compounds that showed the best antibacterial activity in the initial MIC assays (**EN085** and **EN177**, Figure 5c,d), we saw that both compounds show engagement of several cysteines at 20  $\mu\text{M}$  (31 for **EN085**, 10 for **EN177**). Both compounds strongly target C152 of MurC (UniProt code Q2FXJ0), which is a key enzyme essential for cell wall synthesis.<sup>[22]</sup> **EN177** additionally binds to C410 of pyruvate kinase (UniProt code Q2FXM9) and C88 of the essential probable acetyl-CoA acyltransferase (UniProt code Q2G124). The latter cysteine forms an acyl-thioester intermediate during catalysis.<sup>[23]</sup> **EN085** binds to C1030 in the DNA polymerase PolC (UniProt code Q2G1Z8), C199 in the tRNA-specific methyl transferase MnmA (UniProt code Q2FXV6), which forms a cysteine persulfide intermediate during catalysis,<sup>[23]</sup> as well as the catalytically active nucleophile C151 in glyceraldehyde-3-phosphate dehydrogenase (UniProt code Q2G032).<sup>[24]</sup> Both compounds, therefore, bind to several essential target proteins that have the potential to become novel targets of covalent antibiotic compounds.

To investigate if the results obtained in the MSSA strain SH1000 are transferable to other *S. aureus* strains, compounds **EN085** and **EN177** were additionally screened at 20  $\mu\text{M}$  in the MRSA strain USA300, for which they show MIC values of 6.3  $\mu\text{M}$  and 3.1  $\mu\text{M}$ , respectively (see Figure S15 and Table S5). We detect a very good correlation of the data obtained in the two different strains (see Figure S16). All cysteines in essential proteins discussed above were also engaged by the same compound in USA300. While no new engaged cysteines were identified for **EN177**, we identified five additional engaged cysteines for **EN085** in USA300 that were not quantified at all in SH1000. Among those, two cysteines are in essential proteins.<sup>[17]</sup> **EN085** binds to the active site C119 in MurA (UniProt code A0A0H2XGP3), which is a key enzyme in cell wall biosynthesis<sup>[22]</sup> and also targeted by fosfomycin.<sup>[5]</sup> Additionally, C565 in the aspartate-tRNA ligase AspS (UniProt code Q2FG97) is modified by **EN085**, which opens up the possibility to target translation through a novel mechanism. The highly reproducible results between the MSSA and MRSA strains demonstrate that our data delivers a broadly applicable map of ligandable cysteines in the *S. aureus* proteome that will guide the design of antibiotics with novel modes-of-action.

To validate the interaction of a selected compound with an identified ligandable cysteine, we investigated C111 of the putative HMG-CoA synthase (UniProt code Q2FV76), which is an essential enzyme in the mevalonate pathway and in this way might open up targeting bacteria through this so far

clinically unexplored pathway.<sup>[25]</sup> In gel-based experiments (Figure 6a; see Figure S17), strong labeling by IA-alkyne was observed for the recombinant wildtype protein, but not for the C111A mutant. This data is in good agreement with the high reactivity of C111 ( $R_{10:1} = 0.84$ ) in the reactivity experiments (Figure 3; see Table S2). Furthermore, the low promiscuity compound **EN106** that we identified to target HMG-CoA synthase (Figure 4a) blocked labeling at low micromolar concentrations, indicating covalent binding of this compound to C111. Using intact protein MS (IPMS, Figure 6b; see Figures S18 and S19), we detected quantitative single modification of the HMG-CoA synthase wildtype with **EN106**. No modification of the C111A mutant was detectable,



**Figure 6.** a) Result of gel-based labeling experiments with HMG-CoA synthase. 1  $\mu\text{M}$  recombinant wild-type (WT) HMG-CoA synthase was added into 1  $\text{mg mL}^{-1}$  soluble lysate of *S. aureus* SH1000. As controls, 1  $\mu\text{M}$  of the HMG-CoA synthase mutant (C111A) or no HMG-CoA synthase (none) were added. The samples were treated with the indicated concentrations of **EN106** or with DMSO as control. The samples were labeled with IA-alkyne and modified with TAMRA-azide using CuAAC. Analysis using SDS-PAGE with subsequent in-gel fluorescence scanning and Coomassie staining is shown. b) IPMS analysis of the modification of HMG-CoA synthase by **EN106**. 1  $\mu\text{M}$  HMG-CoA synthase wildtype (WT) or mutant (C111A) was treated with DMSO as control or 10  $\mu\text{M}$  **EN106**. Deconvoluted IPMS spectra are shown. The mass difference between the wildtype treated with **EN106** or DMSO ( $\Delta m = 258$  Da) exactly corresponds to the modification of the protein with one molecule of **EN106**. c) Results of activity assays with HMG-CoA synthase. 1  $\mu\text{M}$  HMG-CoA synthase wildtype (WT) was treated with 10  $\mu\text{M}$  **EN106** or DMSO as a control. Acetyl-CoA, acetyl-CoA-acyl-transferase (ACAT) and Ellman's reagent were added and the reaction progress was followed by measuring the absorbance at 410 nm over time. HMG-CoA synthase activity was calculated by a linear fit of the linear portion of this curve. Controls with the HMG-CoA synthase mutant (C111A), no HMG-CoA synthase or no acetyl-CoA-acyl-transferase (no ACAT) were included. The graph shows mean  $\pm$  standard deviation. All data results from triplicates. mAU: milli absorbance units.

strongly indicating that C111 is the site of covalent modification with **EN106**. To study the activity of HMG-CoA synthase (Figure 6c; see Figure S20) we set up a coupled assay with acetyl-CoA-acyl transferase (ACAT).<sup>[26]</sup> ACAT forms acetoacetyl-CoA from two molecules of acetyl-CoA. HMG-CoA synthase then catalyzes the reaction with another molecule of acetyl-CoA to give HMG-CoA. The free thiol group of the CoA-SH liberated in both steps was detected using Ellman's reagent (Figure 6c). Addition of wild-type HMG-CoA synthase to the assay strongly increased the formation of free CoA-SH over the ACAT background reaction. This activity was reduced to the level without HMG-CoA synthase, when the C111A mutant was used or when the wildtype was pretreated with 10  $\mu\text{M}$  **EN106**. **EN106** had no effect on the detected ACAT activity in absence of HMG-CoA synthase or presence of the inactive C111A mutant (see Figure S21), showing that **EN106** does not inhibit the ACAT reaction or hinder detection by alkylation of the free CoA-SH or the product of the Ellman's reagent. Furthermore, inhibition remained after gel-filtration to remove excess free **EN106**, showing irreversible inhibition and further excluding interference of **EN106** with other components of the assay (see Figure S22). Covalent modification of HMG-CoA synthase at C111 with compound **EN106** therefore led to effective inhibition of its activity. HMG-CoA synthase is therefore a promising target for the development of novel antibiotics that interfere with the essential mevalonate pathway.

## Conclusion

We describe the synthesis of isotopically labeled desthiobiotin azide (isoDTB) tags and their application in chemoproteomic experiments. These tags were easily synthesized by solid-phase peptide synthesis in high yields and showed excellent physicochemical properties. By using desthiobiotin, these tags circumvented the need to use complex cleavable linkers<sup>[11a]</sup> for peptide elution and thus significantly shortened the chemoproteomic protocol, while increasing the coverage of cysteines in the proteome of *S. aureus*. The isoDTB tags allowed quantification of many cysteines across different Gram-positive and Gram-negative bacterial proteomes and gave results comparable to the TEV tags, also in the human proteome.<sup>[2a]</sup> Because of the easy synthesis of the tag, the shortened workflow, the use of freely available MaxQuant data evaluation software,<sup>[16]</sup> and the excellent performance, this technology will make residue-specific proteomics applicable in many laboratories not specialized in chemoproteomics.

The isoDTB tags were applied to study the reactivity of cysteines in the proteome of *S. aureus*. We identified 88 highly reactive cysteine residues that are strongly enriched at functional sites of proteins. This enrichment indicates that the reactivity of cysteines is a proxy for the functional relevance of certain residues also in bacterial proteomes. Interestingly, highly reactive cysteines were less likely to be found in essential proteins, pointing to the fact that evolution may have selected against highly reactive cysteines in

essential proteins to protect bacteria from the influence of reactive electrophiles occurring in nature either during metabolism or as environmental chemicals.

Finally, we applied the isoDTB tags to broadly understand which cysteines in the bacterial proteome can be engaged with covalent ligands. For this purpose, we compiled competitive data for 19  $\alpha$ -chloroacetamides and profiled 1756 cysteines, including 59% of all cysteines in essential proteins. We identified 268 cysteines that can bind covalent ligands in 200 different proteins. The targeted cysteines include many functionally relevant residues in essential proteins involved in many different pathways. In this way, the data presented will be the starting point for more specific covalent inhibitors to develop antibiotics with novel modes-of-action. The presented isoDTB tags will allow monitoring of the on- and off-target effects of the compounds and in this way streamline the development process.

We investigated inhibition of HMG-CoA synthase in more detail. Modification at the ligandable cysteine residues was detected and this interaction led to inhibition of the enzyme activity. HMG-CoA synthase inhibition by modification of C111 using the human HMG-CoA synthase inhibitor hymeglusin has been described, but this inhibitor suffers from a short half-life of the thioester in the covalent protein adduct.<sup>[27]</sup> Therefore, permanent covalent inhibition by the low promiscuity compound **EN106** is a very promising starting point to explore the antibiotic potential of this protein.<sup>[25]</sup> This case study demonstrates that our map of ligandable cysteines is an excellent resource to quickly identify residues that can be targeted in a functionally relevant manner.

Taken together, our isoDTB tags are important new tools for residue-specific proteomics in bacterial systems. They allowed the investigation of the bacterial cysteinome globally and should be transferable to studying other amino acids in a straightforward manner.<sup>[8]</sup> The cysteines that were characterized to bind to covalent ligands in this study serve as the foundation for the development of covalent inhibitors that could lead to antibiotics with totally new modes-of-action.

## Acknowledgements

We gratefully acknowledge Prof. Dr. Stephan A. Sieber and his group for their generous support. We thank Dr. Markus Lakemeyer and Dr. Franziska Mandl for critical reading of the manuscript. SMH and PRAZ acknowledge funding by the Fonds der Chemischen Industrie through a Liebig Fellowship and a Ph.D. fellowship. We acknowledge financial support by the TUM Junior Fellow Fund.

## Conflict of interest

The authors declare no conflict of interest.

**Keywords:** antibiotics · covalent inhibitors · isotopic labeling · protein modifications · proteomics

**How to cite:** *Angew. Chem. Int. Ed.* **2020**, *59*, 2829–2836  
*Angew. Chem.* **2020**, *132*, 2851–2858

- [1] M. Lakemeyer, W. Zhao, F. A. Mandl, P. Hammann, S. A. Sieber, *Angew. Chem. Int. Ed.* **2018**, *57*, 14440–14475; *Angew. Chem.* **2018**, *130*, 14642–14682.
- [2] a) K. M. Backus, B. E. Correia, K. M. Lum, S. Forli, B. D. Horning, G. E. Gonzalez-Paez, S. Chatterjee, B. R. Lanning, J. R. Tejjaro, A. J. Olson, D. W. Wolan, B. F. Cravatt, *Nature* **2016**, *534*, 570–574; b) C. G. Parker, A. Galmozzi, Y. Wang, B. E. Correia, K. Sasaki, C. M. Joslyn, A. S. Kim, C. L. Cavallaro, R. M. Lawrence, S. R. Johnson, I. Narvaiza, E. Saez, B. F. Cravatt, *Cell* **2017**, *168*, 527–541.e529; c) C. Y. Chung, H. R. Shin, C. A. Berdan, B. Ford, C. C. Ward, J. A. Olzmann, R. Zoncu, D. K. Nomura, *Nat. Chem. Biol.* **2019**, *15*, 776–785; d) C. A. Berdan, R. Ho, H. S. Lehtola, M. To, X. Hu, T. R. Huffman, Y. Petri, C. R. Altobelli, S. G. Demeulenaere, J. A. Olzmann, T. J. Maimone, D. K. Nomura, *Cell Chem. Biol.* **2019**, *26*, 1027–1035.e1022.
- [3] J. Singh, R. C. Petter, T. A. Baillie, A. Whitty, *Nat. Rev. Drug Discovery* **2011**, *10*, 307–317.
- [4] A. Chaikuad, P. Koch, S. A. Laufer, S. Knapp, *Angew. Chem. Int. Ed.* **2018**, *57*, 4372–4385; *Angew. Chem.* **2018**, *130*, 4456–4470.
- [5] D. Hendlin, E. O. Stapley, M. Jackson, H. Wallick, A. K. Miller, F. J. Wolf, T. W. Miller, L. Chaiet, F. M. Kahan, E. L. Foltz, H. B. Woodruff, J. M. Mata, S. Hernandez, S. Mochales, *Science* **1969**, *166*, 122–123.
- [6] H. Nishimura, M. Mayama, Y. Komatsu, H. Kato, N. Shimaoka, Y. Tanaka, *J. Antibiot.* **1964**, *17*, 148–155.
- [7] P. A. Smith et al., *Nature* **2018**, *561*, 189–194.
- [8] S. M. Hacker, K. M. Backus, M. R. Lazear, S. Forli, B. E. Correia, B. F. Cravatt, *Nat. Chem.* **2017**, *9*, 1181–1190.
- [9] a) E. Weerapana, C. Wang, G. M. Simon, F. Richter, S. Khare, M. B. Dillon, D. A. Bachovchin, K. Mowen, D. Baker, B. F. Cravatt, *Nature* **2010**, *468*, 790–795; b) A. E. Speers, B. F. Cravatt, *J. Am. Chem. Soc.* **2005**, *127*, 10018–10019.
- [10] 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.
- [11] a) Y. Yang, M. Fonovic, S. H. Verhelst, *Methods Mol. Biol.* **2017**, *1491*, 185–203; b) A. J. Rabalski, A. R. Bogdan, A. Baranczak, *ACS Chem. Biol.* **2019**, *14*, 1940–1950.
- [12] J. Szychowski, A. Mahdavi, J. J. Hodas, J. D. Bagert, J. T. Ngo, P. Landgraf, D. C. Dieterich, E. M. Schuman, D. A. Tirrell, *J. Am. Chem. Soc.* **2010**, *132*, 18351–18360.
- [13] a) S. H. Verhelst, M. Fonovic, M. Bogyo, *Angew. Chem. Int. Ed.* **2007**, *46*, 1284–1286; *Angew. Chem.* **2007**, *119*, 1306–1308; b) Y. Qian, J. Martell, N. J. Pace, T. E. Ballard, D. S. Johnson, E. Weerapana, *ChemBioChem* **2013**, *14*, 1410–1414.
- [14] M. P. Patricelli et al., *Chem. Biol.* **2011**, *18*, 699–710.
- [15] M. J. Horsburgh, J. L. Aish, I. J. White, L. Shaw, J. K. Lithgow, S. J. Foster, *J. Bacteriol.* **2002**, *184*, 5457–5467.
- [16] J. Cox, M. Mann, *Nat. Biotechnol.* **2008**, *26*, 1367–1372.
- [17] R. R. Chaudhuri, A. G. Allen, P. J. Owen, G. Shalom, K. Stone, M. Harrison, T. A. Burgis, M. Lockyer, J. Garcia-Lara, S. J. Foster, S. J. Pleasance, S. E. Peters, D. J. Maskell, I. G. Charles, *BMC Genomics* **2009**, *10*, 291.
- [18] P. Gaudet, M. S. Livstone, S. E. Lewis, P. D. Thomas, *Briefings Bioinf.* **2011**, *12*, 449–462.
- [19] N. Paranagama, S. A. Bonnett, J. Alvarez, A. Luthra, B. Stec, A. Gustafson, D. Iwata-Reuyl, M. A. Swairjo, *Biochem. J.* **2017**, *474*, 1017–1039.
- [20] J. Yao, C. O. Rock, *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **2017**, *1862*, 1300–1309.
- [21] a) J. B. Parsons, J. Yao, M. W. Frank, C. O. Rock, *Antimicrob. Agents Chemother.* **2015**, *59*, 849–858; b) A. G. Ekström, V. Kelly, J. Marles-Wright, S. L. Cockcroft, D. J. Campopiano, *Org. Biomol. Chem.* **2017**, *15*, 6310–6313.
- [22] T. D. Bugg, D. Braddick, C. G. Dowson, D. I. Roper, *Trends Biotechnol.* **2011**, *29*, 167–173.
- [23] The UniProt Consortium, *Nucleic Acids Res.* **2019**, *47*, D506–D515.
- [24] S. Mukherjee, D. Dutta, B. Saha, A. K. Das, *J. Mol. Biol.* **2010**, *401*, 949–968.
- [25] N. Campobasso, M. Patel, I. E. Wilding, H. Kallender, M. Rosenberg, M. N. Gwynn, *J. Biol. Chem.* **2004**, *279*, 44883–44888.
- [26] D. A. Skaff, H. M. Mizioro, *Anal. Biochem.* **2010**, *396*, 96–102.
- [27] D. A. Skaff, K. X. Ramyar, W. J. McWhorter, M. L. Barta, B. V. Geisbrecht, H. M. Mizioro, *Biochemistry* **2012**, *51*, 4713–4722.

Manuscript received: September 23, 2019

Accepted manuscript online: November 29, 2019

Version of record online: January 7, 2020