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

Patrick R. A. Zanon, Lisa Lewald, Stephan M. Hacker

Submitted date: 16/09/2019 • Posted date: 20/09/2019
Licence: CC BY-NC-ND 4.0

Citation information: Zanon, Patrick R. A.; Lewald, Lisa; Hacker, Stephan M. (2019): Isotopically Labeled Desthiobiotin Azide (isoDTB) Tags Enable Global Profiling of the Bacterial Cysteinome. ChemRxiv. Preprint.

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. Here, we describe isotopically labeled desthiobiotin azide (isoDTB) tags that are easily synthesized, shorten the chemoproteomic workflow and allow an increased coverage of cysteines in bacterial systems. We quantify 59% of all cysteines in essential proteins in Staphylococcus aureus and discover 88 cysteines with high reactivity, which correlates with functional importance. Furthermore, we identify 268 cysteines that are engaged by covalent ligands. We verify inhibition of HMG-CoA synthase, which will allow addressing the bacterial mevalonate pathway through a new target. Overall, a comprehensive map of the bacterial cysteinome is obtained, which will facilitate the development of antibiotics with novel modes-of-action.

File list (7)

- manuscript_Zanon_et_al.pdf (1.20 MiB)
- SI_Zanon_et_al.pdf (1.83 MiB)
- Zanon_et_al_Table_S1.xlsx (2.45 MiB)
- Zanon_et_al_Table_S2.xlsx (190.76 KiB)
- Zanon_et_al_Table_S4.xlsx (6.82 MiB)
- Zanon_et_al_Table_S5.xlsx (865.14 KiB)
- Zanon_et_al_Table_S12.xlsx (270.33 KiB)
Isotopically Labeled Desthiobiotin Azide (isoDTB) Tags Enable Global Profiling of the Bacterial Cysteinome

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

Department of Chemistry, Technical University of Munich, Lichtenbergstrasse 4, 85748 Garching, Germany, *E-mail: stephan.m.hacker@tum.de

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. Here, we describe isotopically labeled desthiobiotin azide (isoDTB) tags that are easily synthesized, shorten the chemoproteomic workflow and allow an increased coverage of cysteines in bacterial systems. We quantify 59% of all cysteines in essential proteins in Staphylococcus aureus and discover 88 cysteines with high reactivity, which correlates with functional importance. Furthermore, we identify 268 cysteines that are engaged by covalent ligands. We verify inhibition of HMG-CoA synthase, which will allow addressing the bacterial mevalonate pathway through a new target. Overall, a comprehensive map of the bacterial cysteinome is 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.[11] Nevertheless, only a very limited number of novel classes of antibiotics has been introduced to the clinic over the last decades.[11] Furthermore, almost all approved antibiotics exclusively interfere with a very limited set of bacterial targets, mainly involved in protein, nucleic acid and cell wall biosynthesis.[11] Therefore, developing innovative methods to discover novel druggable targets for antibacterial compounds is a pivotal task to guarantee efficient treatment of bacterial infections in the future.

Chemoproteomic approaches are extremely powerful to globally understand which proteins are able to bind small molecules as ligands.[5] These strategies are particularly straightforward for covalently reactive molecules.[2a,2c,2d] Strikingly, covalent inhibitors have simultaneously seen a resurgence of interest for the development of novel drugs as they can increase compound selectivity, reduce resistance formation and target protein pockets that are hard to address using non-covalent interactions alone.[5] This has led to the recent clinical approval of several covalent kinase inhibitors.[4] Especially in the field of antibiotics, covalent inhibitors are prevalent. The most prominent examples are β-lactams,[5] but many other antibiotics, such as fosfomycin,[6] showdomycin[6] and optimized aryomycins,[7] also use a covalent mechanism-of-action.[5]

Over recent years, methods have emerged that identify not only the target proteins of covalent inhibitors but also the exact interaction site in a competitive fashion.[5,2a,2d] In this way, many pockets that can bind covalent ligands can be identified in parallel using a small library of covalently reactive molecules. This technology is especially well established for profiling cysteine residues as these can be investigated with highly specific acrylamide- or chloroacetamid-containing compounds.[2a] Methods that allow this global profiling of cysteines are usually based on the isoTOP-ABPP (isotopic tandem orthogonal proteolysis activity-based protein profiling) platform (Figure 1a).[2a] 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 control. The covalent inhibitor attaches to cysteines in pockets, in which there is an initial favorable non-covalent interaction. In the next step, both samples are treated with iodoacetamide alkyne (IA-alkyne).[9] This probe will modify many cysteines in both samples and this reactivity will be blocked by the covalent inhibitor at its specific binding sites. In order to read out these differences in alkynylation, the samples are next modified with isotopically labeled affinity tags using copper-catalyzed azide-alkyne cycloaddition (CuAAC).[10] As peptides originating from both samples are differentiated due to the isotopic labels, the samples can be combined, enriched on streptavidin beads, proteolytically digested and the modified peptides eluted for mass spectrometry (MS)-based quantification. Most quantified cysteines will have ratios \( R = 1 \) between the heavy and light channel indicating no interaction with the covalent compound (Figure 1a). In contrast, cysteines at the specific binding sites will show ratios of \( R >> 1 \). In this way, the method allows quantitative and site-specific interaction studies in the whole proteome while working with unmodified covalent inhibitors that do not need to be equipped e.g. 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 in order to elute the peptides from the beads.[11] These include linkers cleaved by proteases (Figure 1b),[2a,9] acidic[11b,12] or reductive conditions.[11b,13] Due to the high requirements on the stability of the linkers to various conditions and due to the need of cleaving them specifically under mild conditions, these linkers need to be designed very carefully, which usually requires laborious multi-step 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 1c) for residue-specific proteomics. As desthiobiotin still binds very strongly to streptavidin, all steps up to the proteolytic digestion can be kept the same.[14] Due to the reversibility of binding of desthiobiotin to streptavidin, in the last step, the peptides can then easily be eluted using acidic conditions with acetonitrile as the co-solvent.[14] Because a complex cleavable linker is not needed, we designed these tags with exclusively two isotopically differentiated glycine residues as the linker moiety.
Figure 1. a) Workflow for competitive, residue-specific chemoproteomic experiments.[2a] RG = reactive group; AH = affinity handle. b) Structure of the TEV protease-cleavable tags (TEV tags) originally used for residue-specific proteomics.[2a,9] c) Structure of the isoDTB tags developed in this study.

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 identify 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 with HATU as the coupling reagent were utilized. We sequentially coupled ε-azido-lysine, two glycine residues and desthiobiotin. We used glycine with the natural isotope distribution for the light isoDTB tag and glycine with two $^{13}$C atoms and one $^{15}$N atom for the heavy tag. In this way, a total mass difference between the tags of 6 Da was obtained. After the last coupling step, the isoDTB tags were cleaved from the resin with TFA and purified by RP-HPLC resulting in a yield of approximately 70% for both isoDTB tags.

In order 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[15] with 1 mM IA-alkyne and modified the two samples with the light and heavy isoDTB tag, respectively, using CuAAC. The light and heavy 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 straightforward approach. Analysis using liquid chromatography coupled to tandem MS (LC-MS/MS) and evaluation using freely available MaxQuant software[36] identified 1155 cysteines that were quantified for both conditions (Figure 2a, Supporting Information, Table S1). This analysis revealed a narrow distribution for the sample with a ratio of 1:1 around the expected value of $\log_{2}(R) = 0$ with $> 99.7\%$ of all cysteines within the preferred window of $-1 < \log_{2}(R) < 1$. For the sample with the ratio of 1:4 we also detected a narrow distribution around the expected value of $\log_{2}(R) = 2$ with $> 97\%$ of all cysteines being in the preferred window of $1 < \log_{2}(R) < 3$. The isoDTB tags therefore reliably allow quantification of cysteines in the whole bacterial proteome.

![Figure 2](image-url)

**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. The total area of the circles is scaled to the number of cysteines quantified with the respective tag. c) Number of quantified cysteines in a variety of Gram-positive and Gram-negative bacteria as well as in the proteome of the human cell line MDA-MB-231 using the isoDTB tags. The grey dashed line indicates 1000 quantified cysteines. All data results from duplicates.
We next benchmarked our technology against a known residue-specific technology. Here, we performed an experiment with the TEV protease-cleavable biotin tags (TEV tags) that have been most broadly used to residue-specifically map proteomes. We conducted the same experiment with the TEV tags as described above and mixed the heavy and light sample at a ratio of 1:1 (Figure 2b, Supporting Information, Table S1). Our isoDTB tags outperformed the TEV tags by quantifying 27% more cysteines in the *S. aureus* proteome. Furthermore, we were able to increase the number of cysteines quantified with the isoDTB tags even more by performing additional experiments with chymotrypsin and AspN as substitutes for trypsin for the proteolytic digest (Supporting Information, Figure S1 and Table S1). This is 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 LC-MS/MS experiments. We next investigated the performance of the isoDTB tags with IA-alkyne in different Gram-positive and Gram-negative bacteria (Figure 2c, Supporting Information, Table S1) and could consistently quantify more than 1000 cysteines in each strain. Moreover, in a benchmark study in the proteome of the human lung cancer cell line MDA-MB-231, we were able to quantify more than 3500 cysteines, which is competitive with previously described methods. Therefore, our isoDTB tags not only shorten the chemoproteomic protocol by one day but also lead to increased coverage in bacterial systems compared to the widely used TEV tag technology.

Having this method in hand, we next applied it to analyze the reactivity of cysteines in the bacterial proteome (Figure 3a). As the reactivity of cysteines is linked to their functional relevance in human cells, we reasoned that this feature might also be conserved in bacteria and in this way lead to the identification of functionally important cysteine residues. In order to study cysteine reactivity, two identical samples of the proteome of the *S. aureus* strain SH1000 were treated with either a high (100 µM) or a low (10 µM) concentration of IA-alkyne. In this way, while at the high concentration many cysteines are labeled, at the low concentration only the most reactive cysteines are labeled quantitatively. After CuAAC with the light (low concentration) and heavy isoDTB tags (high concentration), respectively, the samples are analyzed in the same way as described above. Here, high ratios \( R_{10:1} \) indicate low reactivity cysteines, whereas the most reactive cysteines will 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, Supporting Information, 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 (Supporting Information, Figure S2). Interestingly, highly reactive cysteines were depleted in essential proteins 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 reactive and medium reactive cysteines at functional sites (Figure 3d). This demonstrates that cysteine reactivity is a very good measure to indicate functional importance also in bacterial proteomes. In line with this finding, the highly reactive cysteines include many residues that are directly involved in the catalytic
mechanism (e.g. C178 in the essential GTP cyclohydrolase FoiE2 (UniProt code Q2G0L1), C112 in FabH (UniProt code Q2FZS0), C88 in the probable acetyl-CoA acyltransferase (UniProt code Q2G124), C111 in the putative HMG-CoA synthase (UniProt code Q2FV76) and C165 in the 3-oxoacyl-(acyl-carrier-protein) synthase 2 (UniProt code Q2FZ99), 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) is close to the ATP-binding site and C45 in MmMG (UniProt code Q2FUG3) is close to the FAD-binding site or metal-binding sites (e.g. C145 of alcohol dehydrogenase Adh (UniProt code Q2G0G1) binds the catalytically active zinc ion and C65 of biotin synthase BioB (UniProt code Q2FVJ7) binds the iron-sulfur cluster). Therefore, residue-specific proteomics using our isoDTB tags allows global profiling of the reactivity of cysteines in the bacterial proteome and enables 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 commercially available electrophilic cysteine-directed compounds (EN001 - EN211, Supporting Information, Table S3), mainly α-chloro-acetamides. 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 in order to prioritize compounds, whose target spectrum includes essential proteins that can be addressed in intact cells. Based on an initial screen in the MSSA strains SH1000, NCTC8325 and ATCC29213, we selected 24 compounds (Supporting Information, Figure S3) based on their MIC values and structural diversity for comprehensive studies. 23 compounds had MICs of ≤ 100 µM in all three strains with six compounds having MICs of ≤ 12.5 µM. Two of these compounds showed activity (MIC ≤ 100 µM) in two tested MRSA strains (USA300 and DSM18827) with two compounds (EN085 and EN177) having an MIC ≤ 10 µM in all five tested strains. This showed that electrophilic compounds with desired biological activity can efficiently be identified from small compound libraries.

The selected 24 compounds were screened at 200 µM concentration in residue-specific chemoproteomic experiments using our isoDTB tags in duplicates (Figure 1a, Supporting Information, Figure S5 and Figure S6). For three of the compounds (EN007, EN085 and EN177), we performed an additional set of biologically independent duplicates. Due to the high reproducibility between the biologically independent experiments (Supporting Information, 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 ≤ 25 µM in all five tested strains were additionally tested at 20 µM concentration (Supporting Information, Figure S7).

In all experiments, we consider cysteines that have a ratio of R > 4 (log2(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. When we plotted log2(R) against −log10(p) (Volcano plot), we identified a large range of values for the fraction of cysteines that are engaged by the different compounds (Figure 4a-c). Nine compounds of low promiscuity (< 2% of all quantified cysteines are engaged, representative plot in Figure 4a), ten compounds of medium promiscuity (between 2% and 10%, representative plot in Figure 4b) and five compounds of excessive promiscuity (> 10%, representative plot in Figure 4c). Strikingly, within the 24 profiled compounds no correlation between MIC and promiscuity could be observed (Supporting Information, Figure S8b) indicating that it is possible to identify highly active and still selective electrophiles. As we cannot exclude unspecific effects on the proteome or on the chemoproteomic workflow 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 (the 19 low and medium promiscuity compounds at 200 µM, five compounds at 20 µM and a DMSO control) we compiled a comprehensive competitive data table (Supporting Information, 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 are
enriched in our data over the genomic background (Supporting Information, 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 (Supporting Information, Figure S10). Therefore, our method allows obtaining comprehensive information on the cysteinome of S. aureus in a reproducible manner.

We next focused on the cysteines, for which we detect strong engagement with covalent ligands. For this purpose, we define a ligandable cysteine as a cysteine, for which engagement with at least one covalent ligand could be identified with our method ($\log_2(R) > 2$ and $p < 0.05$). In this way, more than 250 ligandable cysteines in 200 different proteins are identified (Figure 5a). In many proteins, for which we were able to quantify more than one cysteine, we detect exclusively one ligandable cysteine that is engaged by several compounds, while the other cysteine(s) are never engaged (Supporting Information, Figure S11). This 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 especially in proteins involved in gene expression and nucleic acid binding as well as among uncharacterized proteins (Supporting Information, Figure S12).

Ligandable cysteines are enriched at functional sites (Supporting Information, Figure S13) and similarly abundant in essential proteins as compared to other quantified cysteines (Supporting Information, Figure S9b). When we compared this data to the cysteine reactivity data (Supporting Information, 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 non-covalent 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 Q2FZ50), an essential enzyme in fatty acid synthesis,[22] 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 e.g. by the inhibitors 4,5-dichloro-1,2-dithiol-3-one and cerulenin.[23] Another example is the ligandable residue C111 of the essential putative HMG-CoA synthase (UniProt code Q2FV76), which forms the thioester intermediate during catalysis.[20,24] This enzyme is essential for ligandable residue C111 of the essential putative HMG-CoA synthase IleS (UniProt code Q2FZ82) is only targeted by Figure 5b). Furthermore, the residue C323 in the isoleucine-tRNA specific methyl transferase (UniProt code Q2FXJ0), which is a key enzyme in cell wall synthesis.[25] 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 µM HMG-CoA synthase wildtype (WT) or mutant (C111A) was treated with DMSO as control or 10 µM EN106. Deconvoluted IPMS spectra are shown. The mass difference between the wildtype treated with EN106 or DMSO (∆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 µM HMG-CoA synthase wildtype (WT) was treated with 10 µM EN106 or DMSO as a control. Acetyl-CoA, acetyl-CoA-acetyl-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) or no HMG-CoA synthase or no acetyl-CoA-acetyl-transferase (no ACAT) were included. The graph shows mean ± standard deviation. All data results from triplicates. mAU: milli absorbance units.

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 µM (31 for EN085, 10 for EN177). Both compounds strongly target C152 of MurC (UniProt code Q2FZXJ0), which is a key enzyme for cell wall synthesis.[28] EN177 additionally binds to C410 of pyruvate kinase (UniProt code Q2FXM9) and C88 of the essential probable acetyl-CoA acetyltransferase (UniProt code Q2G124). The latter cysteine forms an acyl-thioester intermediate during catalysis,[29] as well as the catalytically active nucleophile C151 in glyceraldehyde-3-phosphate dehydrogenase (UniProt code Q2G032).[30] Both compounds, therefore, bind to several essential target proteins that have the potential to become novel targets of covalent antibiotic compounds.

In order 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 µM in the MRSA strain USA300, for which they show MIC values of 6.3 µM and 3.1 µM, respectively (Supporting Information, Figure S15 and Table S5). We detect a very good correlation of the data obtained in the two different strains (Supporting Information, 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.[17] EN085 binds to the active site C119 in MurA (UniProt code A0A0H2XGP3), which is a key enzyme in cell wall biosynthesis.[26] C119 is also the target of the known MurA inhibitor fosfomycin.[16] C656 in the aspartate-tRNA ligase AspS (UniProt code Q2FG97) is also targeted 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 comprehensive and broadly applicable map of ligandable cysteines in the S. aureus proteome that will guide the design of antibiotics with novel modes-of-action.

In order to validate the interaction of a selected compound with an identified ligandable cysteine and to establish the functional outcome of this interaction, we investigated C111 of the putative HMG-CoA synthase (UniProt code Q2FV76), which is an essential enzyme in the mevalonate pathway in this way might open up targeting bacteria through this so far clinically unexplored pathway.[20] In gel-based experiments (Figure 6a, Supporting Information, Figure S17), strong
labeling by IA-alkyne is observed for the recombinant wildtype protein. In contrast the C111A mutant is not labeled. This is in good agreement with the high reactivity of C111 (R10.1 = 0.84) in the reactivity experiments performed here (Figure 3, Supporting Information, Table S2). Furthermore, the low promiscuity compound EN106 that was identified as a ligand for HMG-CoA synthase (Figure 4a) blocked labeling of the wildtype protein at low micromolar concentrations indicating covalent binding of this compound to C111. Using intact protein MS (IPMS, Figure 6b, Supporting Information, Figure S18, S19), we detected quantitative modification of the HMG-CoA synthase wildtype with EN106 with no double modification detected. No modification of the C111A mutant was detectable, strongly indicating that C111 is the site of covalent modification with EN106. In order to study the activity of HMG-CoA synthase (Figure 6c, Supporting Information, Figure S20) we set up a coupled assay with acetyl-CoA-acyl transferase (ACAT).[28] In the first step, ACAT catalyzes the synthesis of acetocetetyl-CoA from two molecules of acetyl-CoA, which is accompanied by the liberation of one molecule of free CoA-SH. Acetocetetyl-CoA is the substrate for HMG-CoA synthase, which uses an additional equivalent of acetyl-CoA to give HMG-CoA and another equivalent of free CoA-SH. The free thiol group in the released CoA-SH is detected using Ellman’s reagent and the resulting absorbance at 410 nm. As the ACAT reaction releases one equivalent of free CoA-SH, there is a reaction in the absence of HMG-CoA synthase that is strongly reduced in the absence of ACAT (Figure 6c). Wildtype HMG-CoA synthase strongly increases the formation of free CoA-SH. This activity is reduced to the level without HMG-CoA synthase when the C111A mutant is used or when the wildtype is pretreated with 10 µM EN106. Covalent modification of HMG-CoA synthase at C111 with compound EN106 therefore leads 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

Herein, we describe the synthesis of isotopically labeled desthiobiotin azide (isoDTB) tags and their application in chemoproteomic experiments. These tags are easily synthesized by solid-phase peptide synthesis in high yields and show excellent physicochemical properties. By using desthiobiotin that binds very strongly but reversibly to streptavidin beads, these tags circumvent the need to use complex cleavable linkers[114] for peptide elution and thus significantly shorten the chemoproteomic protocol (e.g. by one day in comparison to the traditionally used TEV tags[2a,9]), while increasing the coverage of cysteines in the proteome of S. aureus. Furthermore, due to the simple structure of the linker, additional proteases can be used for proteolytic digest, which further increases the number of quantified cysteines. The isoDTB tags do allow quantifying many cysteines across different Gram-positive and Gram-negative bacterial proteomes and give results comparable to the TEV tags also in the human proteome.[24] Due to the easy synthesis of the tag, the shortened workflow, the use of freely available MaxQuant data evaluation software[19] 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 identify 88 highly reactive cysteine residues that are strongly enriched at functional sites of proteins. This indicates that 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. This points to the fact that evolution may have selected against highly reactive cysteines in essential proteins in order to protect bacteria from the influence of reactive electrophiles occurring in nature 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 the competitive data for 19 chloroacetamides and were able to profile 1756 cysteines including 59% of all cysteines in essential proteins. We identify more than 250 cysteines that are able to bind covalent ligands in 200 different proteins. The targeted cysteines include many functionally relevant residues in essential proteins involved e.g. in glycolysis (glycer-aldehyde-3-phosphate dehydrogenase), fatty acid synthesis (FabH), the mevalonate pathway (HMG-CoA synthase), cell wall biosynthesis (MurA and MurC), protein translation (IleS and AspS), DNA replication (PolC) and tRNA modification (MnmA). In this way, the data presented will be the starting point for more specific covalent inhibitors of these proteins in order to develop antibiotics with novel modes-of-action. The presented isoDTB tags will allow monitoring the on- and off-target effects of the compounds closely and in this way streamline the development process.

As case study to verify that the identified covalent ligands can modulate protein function, we investigated HMG-CoA synthase. Modification of the recombinant protein at the identified cysteine residues is detected and this interaction leads 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.[29] Therefore, permanent covalent inhibition by the low promiscuity compound EN106 is a very promising starting point to explore the antibiotic potential of this protein that is key to the essential mevalonate pathway.[29] This case study demonstrates that our map of cysteines that can be engaged with covalent ligands is an excellent resource to quickly identify cysteines 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 allow studying the bacterial cysteinome globally and should be transferable to studying other amino acids like lysine, for which broadly reactive alkyne probes are available, in a straightforward manner.[16] The cysteines that are characterized to be able to bind to covalent ligands in this study are the foundation for the development of covalent inhibitors that could lead to novel 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.

Supporting Information

Figures S1-S22, Tables S1-S12 and experimental details are provided in the Supporting Information accompanying this article.

Keywords: Antibiotics • Covalent Inhibitors • Isotopic labeling• Protein modifications • Proteomics

[1] M. Lakemeyer, W. Zhao, F. A. Mandl, P. Hammann, S. A. Sieber, Angew. Chem. Int. Ed. 2018, 57, 14440-14475.
[2] a) K. M. Backus et al., Nature 2016, 534, 570-574; b) C. G. Parker et al., Cell 2017, 168, 527-541.e520; c) C. Y. Chung et al., Nat. Chem. Biol. 2019, 15, 776-785; d) C. A. Berdan et al., Cell Chem. Biol. 2019, 26, 1027-1035.e1022.
[3] J. Singh, R. C. Petter, T. A. Baillie, A. Whitty, Nat. Rev. Drug. Discov. 2011, 10, 307-317.
[4] A. Chaikuad, P. Koch, S. A. Laufer, S. Knapp, Angew. Chem. Int. Ed. 2018, 57, 4372-4385.
[5] D. Hendlin et al., Science 1969, 166, 122-123.
[6] H. Nishimura et al., J. Antibiot. 1964, 17, 148-155.
[7] [12] J. Szychowski et al., 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; b) Y. Qian et al., ChemBioChem 2013, 14, 1410-1414.
[14] M. P. Patricelli et al., Chem. Biol. 2011, 18, 699-710.
[15] M. J. Horsburgh et al., J. Bacteriol. 2002, 184, 5457-5467.
[16] J. Cox, M. Mann, Nat. Biotechnol. 2008, 26, 1367-1372.
[17] R. R. Chaudhuri et al., BMC Genomics 2009, 10, 291.
[18] P. Gaudet, M. S. Livstone, S. E. Lewis, P. D. Thomas, Brief. Bioinform. 2011, 12, 449-462.
[19] N. Paranagama et al., Biochem. J. 2017, 474, 1017-1039.
[20] N. Campobasso et al., J. Biol. Chem. 2004, 279, 44883-44888.
[21] T. Osawa et al., Structure 2009, 17, 713-724.
[22] J. Yao, C. O. Rock, Biochim. Biophys. Acta Mol. Cell Biol. Lipids 2017, 1862, 1300-1309.
[23] a) J. B. Parsons, J. Yao, M. W. Frank, C. O. Rock, Antimicrob. Agents Chemother. 2015, 59, 849-856; b) A. G. Ekstrom, V. Kelly, J. Marles-Wright, S. L. Cockcroft, D. J. Campopiano, Org. Biomol. Chem. 2017, 15, 6310-6313.
[24] Y. Matsumoto et al., Sci. Rep. 2016, 6, 22894.
[25] T. O. Bugge, D. Braddick, C. G. Dowson, D. I. Roper, Trends Biotechnol. 2011, 29, 167-173.
[26] The UniProt Consortium, Nucleic Acids Res. 2019, 47, D506-D515.
[27] S. Mulherjee, D. Dutta, B. Saha, A. K. Das, J. Mol. Biol. 2010, 401, 949-968.
[28] D. A. Skaff, H. M. Miziorko, Anal. Biochem. 2010, 396, 96-102.
[29] D. A. Skaff et al., Biochemistry 2012, 51, 4713-4722.
### Other files

| File Name                  | Size       | View on ChemRxiv | Download file |
|----------------------------|------------|------------------|---------------|
| SI_Zanon_et_al.pdf         | 1.83 MiB   |                  |               |
| Zanon_et_al_Table_S1.xlsx  | 2.45 MiB   |                  |               |
| Zanon_et_al_Table_S2.xlsx  | 190.76 KiB |                  |               |
| Zanon_et_al_Table_S4.xlsx  | 6.82 MiB   |                  |               |
| Zanon_et_al_Table_S5.xlsx  | 865.14 KiB |                  |               |
| Zanon_et_al_Table_S12.xlsx | 270.33 KiB |                  |               |