2012

Chemical Reporters for Bacterial Pathogenesis and Beyond

Markus Grammel

CHEMICAL REPORTERS FOR BACTERIAL PATHOGENESIS AND BEYOND

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Markus Grammel

June 2012

CHEMICAL REPORTERS FOR BACTERIAL PATHOGENESIS AND BEYOND

Markus Grammel, Ph.D.

The Rockefeller University 2012

Microbial infection involves complex and dynamic interplay of host defense and pathogen virulence pathways that requires new methods for mechanistic investigation. Towards this goal, I have developed chemical reporters of nucleic acid and protein synthesis and posttranslational modifications to explore their functions in bacterial pathogenesis as well as basic biological pathways.

Initially, I developed amino acid chemical reporters that allow the cell-selective labeling of bacterial proteomes to facilitate the analysis of bacterial protein expression during infection of host cells. These reporters are based on unnatural amino acids that are incorporated into the bacterial proteome instead of methionine or phenylalanine when used in combination with mutant aminoacyl-tRNA-synthetases. This technology allows the visualization of bacterial protein synthesis during infection as well as the enrichment, identification and proteomic analysis of bacterial proteins from infected host cells.

Additionally, I developed a chemical reporter for the study of protein AMPylation, a key posttranslational modification used by various bacterial pathogens to rewire host-signaling pathways. This reporter is used by all of the known AMPylation enzyme classes and allows the proteomic identification of enzyme-specific substrates in cell lysates. This project also stimulated the development of a chemical reporter for RNA synthesis, which in contrast to previous RNA reporters allows labeling of mRNA polyadenylation tails.

In conclusion, I have developed chemical reporters for the cell selective labeling of bacteria in the presence of mammalian cells, a chemical reporter for protein AMPylation, and a chemical reporter for RNA synthesis and mRNA polyadenylation. These reagents should help to identify bacterial proteins that are differentially expressed during infection, new AMPylation substrates and the analysis of mRNA polyadenylation dynamics in basic cell biology and microbial pathogenesis studies in the future.

Für meine Familie

ACKNOWLEDGMENTS

First and foremost I would like to thank my mentor Howard Hang for supporting me over the past five years, providing me with a stimulating research environment and fostering independent thought. I would like to thank Mingzi Zhang for starting the amino acid project with me. Most indebted I am to Kelvin Tsou, Ann Raghavan, Yu-Ying Yang, Guillaume Charron, and Paul Dossa for many discussions about the chemical synthesis aspects of my thesis. Furthermore I would like to thank Kavita Rangan and Jacob Yount for their help, support and advice over the years. I also would like to thank my two summer students Emma Taylor-Salmon (BRSS program) and James Flexner.

In particular for the successful AMPylation project I would like to thank Kim Orth (UTSW) and Phi Luong, who executed all cloning and expression experiments.

For the RNA labeling project I would like to thank Nicholas Conrad (UTSW), who carried out all further characterization of the chemical reporter apart from the microscopy experiments.

I also want to thank the Proteomics Resource Center of The Rockefeller University and its current and past staff for doing a wonderful job on the mass spectrometry analysis and providing us with many more resources.

I would like to thank Johannes Larsch for many exciting experiments with the worm C. elegans and wonderful support over the years.

I would like to thank my former collaborators, who did not directly contribute to the projects discussed in this thesis, but accompanied me for the first legs of my graduate

career at The Rockefeller University: Marc-Werner Dobenecker, Alexander Tarakhovsky and Margaret MacDonald.

Last but not least I would like to thank my committee members Tom Muir, Erec Stebbins and Virginia Cornish.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Introduction

The genomic revolution has offered unique opportunities to understand basic biology and human disease. With increasingly high-throughput and cost-effective sequencing methods, the genomes of many organisms and mutations associated with human diseases have been annotated (Cooper & Shendure 2011). In parallel, large-scale gene deletion and overexpression studies as well as protein-protein interaction screens have uncovered complex signaling networks (S. J. Dixon et al. 2009; Choudhary & Mann 2010). These studies have not only revealed the genetic basis for many fundamental biological processes but have also presented new possibilities for personalized medicine. Nonetheless, significant challenges lay ahead for translating genetically encoded information into biological function and therapeutic development. Starting at the nucleic acid level, many biological phenotypes are controlled by multiple genes that many not be readily apparent by perturbation of individual genes or alleles. Furthermore, the phenotype of interest may only be transient or controlled by epigenetic mechanisms not evident by reading the genetic code (Goldberg et al. 2007; M. Guttman & Rinn 2012). To complex matters further, the genome of most organisms encode protein and enzyme families that exhibit overlapping biochemical activities and crosstalk (López-Otín & Hunter 2010), which are regulated by specific interactions with other factors in time and space inside of cells (Seet et al. 2006). Beyond nucleic acids and proteins, glycans, lipids and other small molecule cofactors provide important regulators of cellular function by direct modification of substrates (Walsh et al. 2005), which are directly coupled to primary metabolism that has become increasing important in human disease (Vander

Heiden et al. 2009). The resulting signaling pathways are therefore often not linear and typically comprise more complex

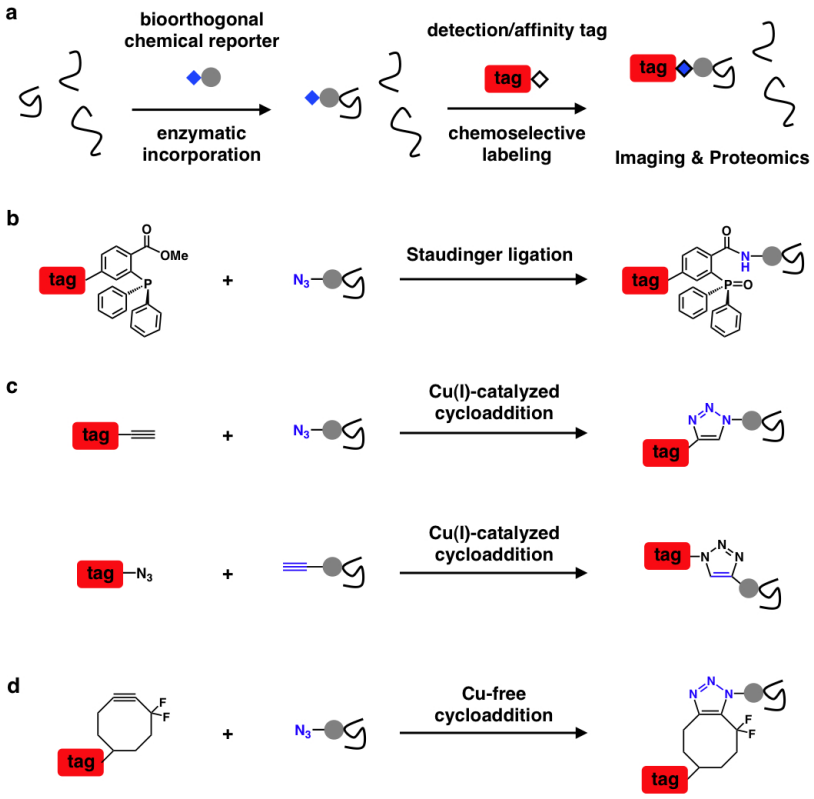


Figure 1 Two-step labeling of biomolecules. a) Bioorthogonal chemical reporters and chemoselective reactions allow covalent labeling of biomolecules with detection or affinity tags for imaging or proteomics applications, respectively. b) Staudinger ligation allows labeling of alkyl-azides with triarylphoshine reagents. c) Cu(I)-catalyzed cycloaddition between alkyl-azides and terminal alkynes affords triazole adducts, which can operate in both orientations. d) Activated cyclooctynes can react with alkyl-azides for Cu-free labeling.

positive and negative feedback networks to amplify or shutdown cellular functions (Nurse & Hayles 2011). These complex and regulated properties allow cells to response to changes their environment and interact with other cells in higher organisms that can

give rise to remarkably biological phenotypes but when misregulated can yield disease. These biological challenges demand a better understanding of the chemistry inside cells and animals that is ever changing in response to mutation and environmental cues. The integration of intrinsic factors encoded within genomes with extrinsic factors in the environment is therefore crucial and require new approaches to evaluate changes in nucleic acid biosynthesis, protein turnover and modifications. As new technology has accelerated the advances in genomics, here we highlight advances in bioorthogonal chemical reporters that have enabled the imaging and functional analysis of nucleic acids, proteins, glycans, lipids as well as other metabolites in biology.

The last decade has seen the remarkable development of highly selective chemical reactions, which have allowed specific labeling of molecules in diverse biological contexts (Fig. 1) (Prescher & Carolyn R Bertozzi 2005; Sletten & Carolyn R Bertozzi 2011). This two-step labeling and detection strategy involves metabolic or enzymatic installation of a chemical reporter that can be selectively reacted with specifically functionalized probes for the visualization or identification of target molecules (Fig. 1a). Central to this chemical labeling strategy is the introduction of azide and alkyne groups as abiotic, minimally perturbing functionalities onto small molecules and their selective covalent reaction with activated phosphine-, alkyne- or azide-modified probes (Fig. 1). Since these functional groups are foreign to living organisms and uniquely reactive with each other, this class of chemical reactions has been described as *bioorthogonal* (Prescher & Carolyn R Bertozzi 2005). The first example of these bioorthogonal reactions was the Staudinger ligation introduced by the Bertozzi laboratory (Fig. 1b) (Saxon & C R Bertozzi 2000). These studies demonstrated that alkyl-azides can react with ester-

functionalized triphenylphosphine reagents to form covalent adducts in aqueous solution and on the surface of mammalian cells (Saxon & C R Bertozzi 2000) as well as in living animals (Prescher et al. 2004). Alternatively, the Sharpless and Meldal laboratories showed that catalytic amounts of Cu(I) can accelerate the [3 + 2] cycloaddition of alkyl-azides with terminal alkynes to form stable triazole products (Fig. 1c)(Tornøe et al. 2002; Rostovtsev et al. 2002a), which elaborates on thermal reactions originally described by Huisgen and coworkers (Huisgen 1963). The Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC, Figure 58) is highly chemoselective and represents the prototypic *click chemistry* reaction, which is widely used in chemical synthesis and bioorthogonal reactions since the azide and alkyne groups can be used in either orientation, are relatively stable and readily accessible. Notably, the terminal alkyne is comparable in size with the azide and when used as a primary chemical reporter in combination with azide-functionalized probes provides the optimal signal-to-noise for bioorthogonal labeling (Fig. 1c) (Speers & Cravatt 2004; Charron et al. 2009a). While the cytotoxicity of Cu(I) can preclude *in vivo* applications, the development of activated cyclooctyne reagents has circumvented this limitation and has enabled imaging of bioorthogonal chemical reporters in living organisms (Fig. 1d) (Laughlin et al. 2008a; Laughlin & Carolyn R Bertozzi 2009b).

The significant advances made in bioorthogonal chemistry have inspired a series of azide- and alkyne-functionalized chemical reporters as molecular global positioning systems for nucleic acid and protein synthesis as well as other biomolecules. Key to this bioorthogonal labeling strategy is the ability of native or engineered enzymes to accept azide or alkyne-functionalized chemical reporters. Provided the activity and selectivity of

the enzymes are not significantly altered by chemical reporters, this approach allows the visualization and biochemical analysis of endogenous biomolecules. This approach provides several important advantages for biological applications. First, the separation of metabolic or enzymatic incorporation and detection decouples sterically-demanding fluorophores or affinity tags, which often interfere with biological activity, from enzymatically processed metabolites. Secondly, in comparison to protein-based reagents (e.g. antibodies), bioorthogonal chemical reporters provide direct detection of the modified substrates and are relative independent of adjacent functional groups, which may confound the specificity of protein reagents. Lastly, the ability to pulse-label specific populations of molecules in cells with bioorthogonal chemical reporters provides an important means to monitor dynamics akin to classic radioactive tracers, but provides much more sensitive detection and allows the identification of labeled targets. Since newly synthesized molecules can be easily separated from the steady state population in this manner, chemical reporters are excellent tools for monitoring the dynamics of biomolecules in response to extrinsic stimuli. Bioorthogonal chemistry has also facilitated the analysis of small molecule inhibitor interactions with proteins for biochemical target identification and activity-based protein profiling, which is beyond the scope of this review and summarized elsewhere (Nomura et al. 2010). Herein we will review the currently available chemical reporters and highlight how they can be applied to monitor temporal changes in cellular activity (Fig. 2a), profile different cell types, states or mutations (Fig. 2b) and identify targets of specific enzymes *in vitro* (Fig. 2c). We close with an outlook on future developments that should expand the utility of bioorthogonal chemical reporters in even more complex biological settings *in vivo*.

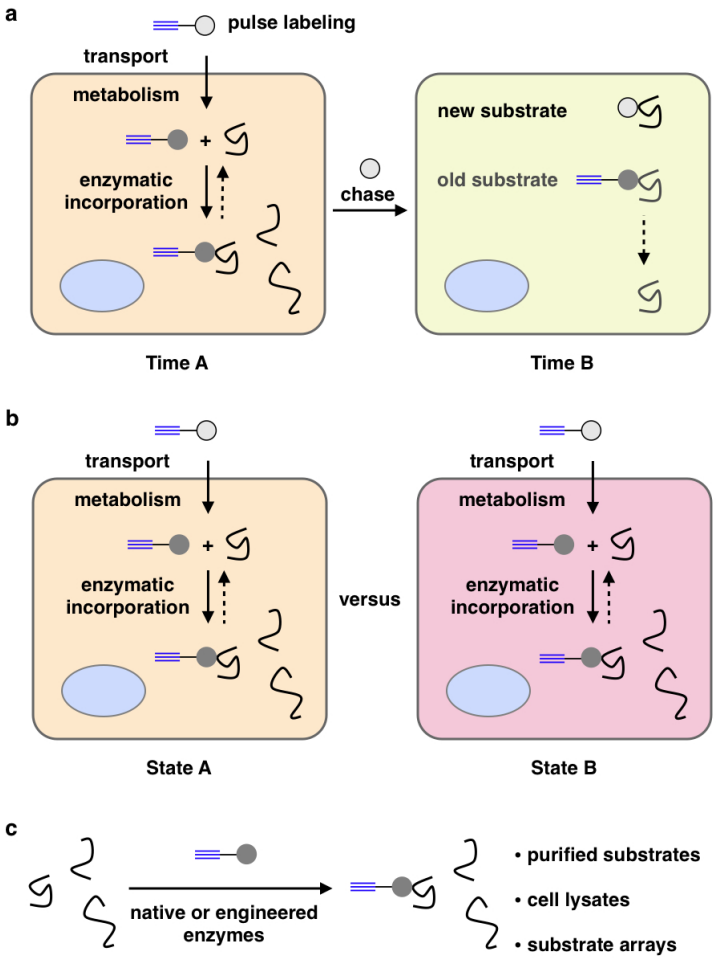


Figure 2 Chemical reporters can be used for comparative analysis *in vivo* and *in vitro*. a) Pulse-labeling of isolated cells or organisms with a chemical reporter leads to the metabolic incorporation of the reporter into a defined biomolecule population. Chasing this population with the native metabolites or differentially labeled reporter enables the monitoring of specific turnover rates. b) Pulse-labeling of two distinct cellular populations or organisms allows the comparative analysis of their metabolically labeled biomolecule pools. c) The application of chemical reporters in cell lysates, with purified substrates, or in combination with substrate arrays allows for enzyme-specific substrate detection.

1.1 Bioorthogonal chemical reporters for nucleic acid turnover and modifications

Gene expression arrays and deep-sequencing methods provide important profiles of steady state DNA and RNA expression, but understanding the rate of nucleic acid synthesis, degradation and modification is crucial for elucidating their function in different biological contexts. The modification of nucleosides with alkynes has afforded bioorthogonal chemical reporters of DNA and RNA synthesis. For example, Salic and coworkers have demonstrated that 5-ethynyl-2'-deoxyuridine (EdU) (Salic & Mitchison 2008) and 5-ethynyluridine (EU) (Jao & Salic 2008) can be metabolized by mammalian cells and incorporated into DNA and RNA for bioorthogonal detection and pulse-chase experiments, respectively (Fig. 3a,b). These nucleic acid reporters function in animals and provide sensitive reagents to compare DNA and RNA turnover in different tissues using fluorescence. More recently, fluorinated arabinosyl-EdU has also been developed to improve the sensitivity of DNA labeling with reduced cytotoxicity compared to EdU (Fig. 3c) (Neef & Luedtke 2011). In addition to uridine derivatives, 5-ethynyl-2'-deoxycytosine (EdC) (Guan et al. 2011) and N-6-propargyladenosine (N6pA) (Grammel et al. 2012) can be utilized by mammalian cells to label DNA and RNA, respectively (Fig. 3d,e). In addition, N6pA labeling allows the pulse-chase analysis of RNA polyadenylation. Alternatively, He and coworkers have employed a specific glycosyltransferase to install 6-azido-glucose onto 5-hydroxymethylcytosine (5-hmc) marks in DNA for subsequent bioorthogonal detection (Fig. 3f) (C.-X. Song et al. 2011). This chemoenzymatic bioorthogonal detection of 5-hmc has provided an important method for characterizing the mechanism of DNA oxidation and demethylation catalyzed by Tet1-3 proteins in eukaryotic cells (He et al. 2011; Ito et al. 2011). As the turnover and modification of nucleic acids is becoming increasingly relevant in dissecting epigenetic

mechanisms, these nucleic acid chemical reporters will likely become more widely used in biology.

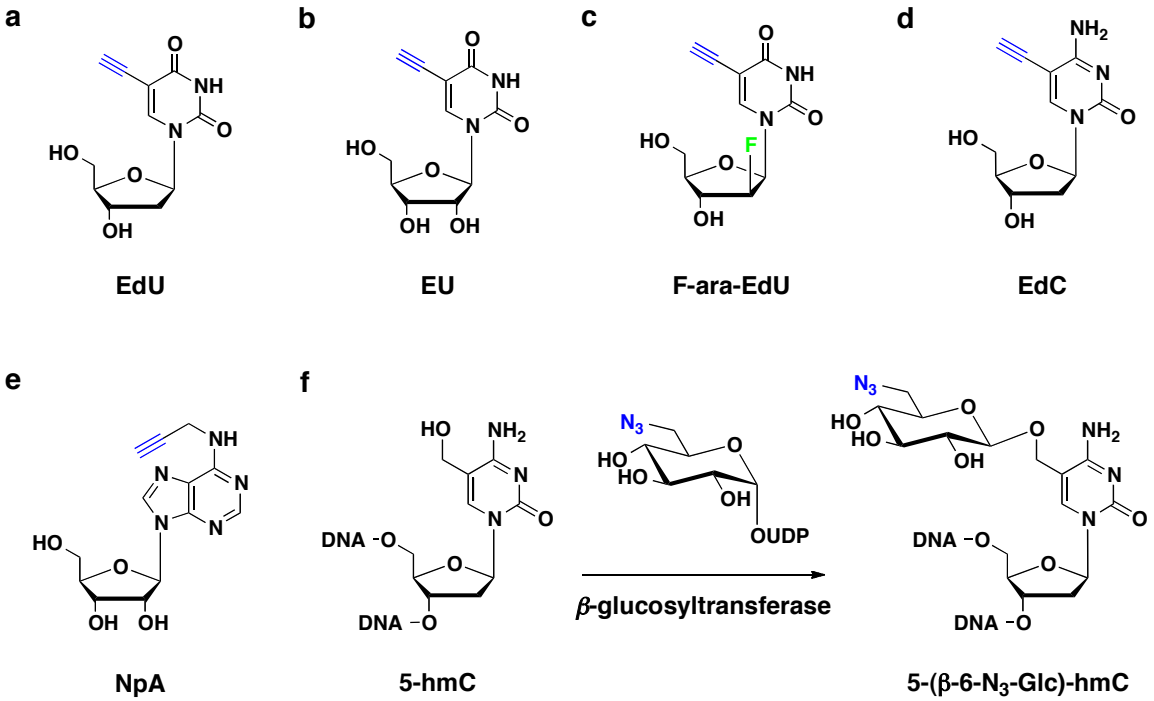


Figure 3 Chemical reporters for nucleic acid synthesis and modifications. a) 5-ethynyl-2'-deoxyuridine (EdU) is incorporated into DNA. b) 5-ethynyluridine (EU) is incorporated into RNA. c) 5-ethynyl-2'-deoxyfluorouridine (F-ara-EdU) is incorporated in to DNA. d) 5-ethynyl-2'-deoxycytosine (EdC) in incorporated into DNA. e) N-6-propargyladenosine (N6pA) is incorporated into RNA and into mRNA polyadenylation tails. f) The DNA modification 5-hydroxymethylcytosine (5-hmc) can be enzymatically labeled with 6-azido-glucose for subsequent bioorthogonal detection.

1.2 Bioorthogonal chemical reporters for protein turnover

Azide- and alkyne-bearing amino acids allow the co-translational labeling of newly synthesized proteins and provide a direct measure of the translational activity in cells (Fig. 4). Two strategies for unnatural amino acid incorporation have been developed that

enable site-specific or residue-specific labeling of proteins. The site-specific approach pioneered by the Schultz laboratory allows the incorporation of an unnatural amino acid at a predetermined position in a protein of interest using amber suppression (C.-X. Song et al. 2011). This technique is based on engineering an orthogonal triad of tRNA, aminoacyl-tRNA-synthetase (aaRS) and unnatural amino acid, which allows the site-specific installation of azide- and alkyne amino acids on individual proteins in bacteria, yeast and mammalian cells (C. C. Liu & P. G. Schultz 2010). For example, the pyrrolysine (Pyl) system has been particularly useful for site-specific incorporation of amino acid reporters such as PCL (Hancock et al. 2010) or ACPK (S. Lin et al. 2011) (Fig. 4a). For the global analysis of translation, the Tirrell laboratory has developed residue-selective incorporation of azide and alkyne amino acids such as azidohomoalanine (AHA) and homopropargylglycine (HPG) that substitute for methionine (Met) in newly synthesized proteins (Fig. 4b) (Johnson et al. 2010). AHA and HPG are readily taken up and utilized by the wild type methionyl-tRNA-synthetase (metRS) in prokaryotic and eukaryotic cells, which enables the analysis of newly synthesized proteins by bioorthogonal labeling with fluorescence and affinity probes. The availability of two uniquely reactive orientations (azide versus alkyne) also makes it possible to label and visualize two distinct protein populations with AHA and HPG, respectively (L. Wang & P. G. Schultz 2004; C. C. Liu & P. G. Schultz 2010). In contrast to metabolic labeling with isotopically labeled amino acids, AHA and HPG labeled proteins can be bioorthogonally reacted with affinity tags and selectively recovered from total cell lysates for mass spectrometry-based protein identification using gel-based methods or multi-dimensional protein identification technology (MudPIT) (Beatty &

Tirrell 2008). Quantitative elution of selectively captured proteins can be challenging using biotinylated affinity tags and streptavidin beads due to the strength of this interaction ($K_D \sim 10^{-15}$), thus several cleavable linker systems (disulfide (Nessen et al. 2009), acid-sensitive (Szychowski et al. 2010), azobenzenes (Y.-Y. Yang, Grammel, et al. 2010), and proteolytic cleavage sites (Speers & Cravatt 2005)) have been developed to facilitate protein elution and bioorthogonal chemical proteomics applications.

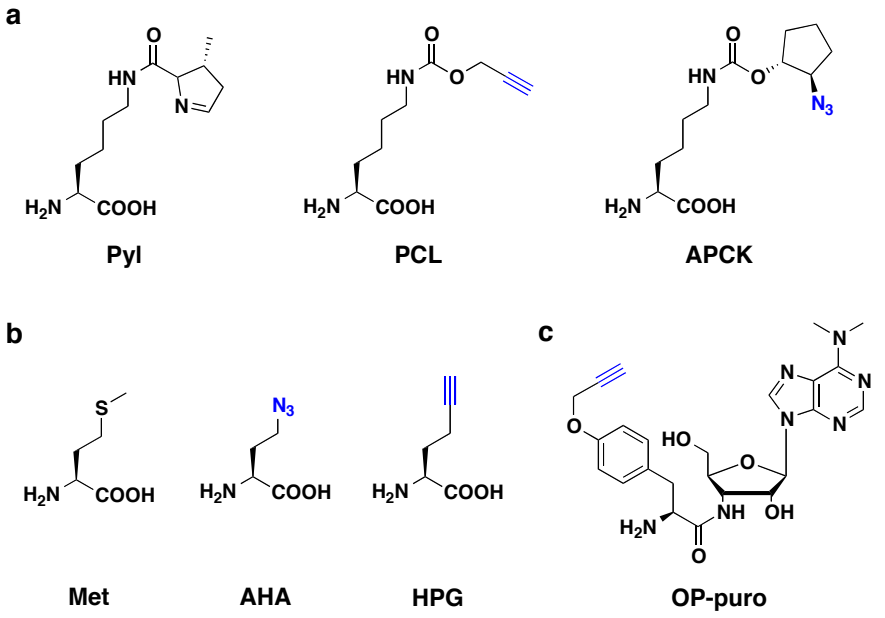


Figure 4 Amino acid reporters can co-translationally incorporated into proteins. a) Analogs of the amino acid pyrrolysine (Pyl) with alkyne (N_ε-[(2-propynyloxy)carbonyl]-lysine, PCL) or azide (azido-cyclopentyl-carboxy-lysine, APCK) groups can be site-selectively incorporated into individual proteins through the expression of orthogonal aminoacyl-tRNA-synthetase and tRNA. b) Various analogs of the amino acid methionine (Met) can be selectively incorporated instead of Met throughout the bacterial or mammalian proteome to install an alkyne (homopropargylglycine, HPG) or azide (azidohomoalanine, AHA) group. c) O-propargyl-puromycin (OP-puro) covalently labels the C-terminus of nascent peptides.

The ability to differentiate newly synthesized from pre-existing proteins with AHA and HPG has revealed important biological insights. For example, Henikoff and colleagues have used AHA pulse labeling for enrichment of newly synthesized nucleosomes (Deal et al. 2010). It is generally appreciated that nucleosome turnover erases histone marks and serves as an important epigenetic mechanism (S. Henikoff 2008). Previous studies measured histone H3.3 incorporation as a surrogate for nucleosome turnover or used overexpressed histone transgenes to follow nucleosome dynamics (Mito et al. 2005; Dion et al. 2007). In contrast, AHA allows pulse labeling of endogenous histones and provides an affinity handle for enrichment and subsequent analysis of nucleosome-associated DNA with tiling arrays. This approach, termed CATCH-IT by the authors, provides better temporal resolution compared to previous histone replacement methods and revealed more rapid turnover of nucleosomes at active genes, specific epigenetic marks and replication origins (Deal et al. 2010). Dynamic protein synthesis is also particularly important in neurons, as localized translation is believed to shape axon growth, guidance and regeneration in the brain for synaptic plasticity and memory (Gkogkas et al. 2010; Jung et al. 2011). To explore these issues, Schuman and colleagues have used AHA and HPG labeling to visualize specific sites and measure rates of protein synthesis in primary neurons after bioorthogonal labeling with fluorescent dyes and quantum dots (Dieterich et al. 2010). Another application of AHA labeling in primary neuronal cultures revealed localized protein synthesis adjacent to the transmembrane receptor DCC (Tcherkezian et al. 2010). These studies show how bioorthogonal amino acid reporters provide valuable tools for monitoring discrete protein populations and imaging protein synthesis at high resolution.

1.2.1 Cell-selective labeling of proteomes.

Studying cells in their native environment and in interaction with other cell types is crucial in many areas of biology. Methods for cell-selective proteome labeling would therefore be valuable for dissecting the activity of heterogeneous cell populations. While AHA and HPG do not provide a means to separate proteomes of individual cell types, Tirrell and colleagues have identified a number of metRS mutants (metRS*) that facilitate the incorporation of the otherwise translationally inactive amino acid reporters such as azidonorleucine (ANL) (Fig. 5a) (Link et al. 2006; Tanrikulu et al. 2009). Since ANL incorporation is dependent on metRS* activity, ANL can be selectively directed to certain cell types by expression of metRS* (Ngo et al. 2009). For example, bacterial cells expressing metRS* can be selectively labeled with ANL in the presence of mammalian cells (Ngo et al. 2009). To differentiate bacterial proteins from host proteins during infection, our laboratory developed 2-aminooctynoic acid (AOA) (Fig. 5a), which is also activated by metRS* to label the Gram-negative bacterial pathogen *Salmonella typhimurium* inside murine macrophages (Grammel et al. 2010). AOA appears to provide better signal-to-noise than ANL and enables the proteomic identification of endogenously expressed *S. typhimurium* proteins during infection (Fig. 5b) (Grammel et al. 2010). In addition to these Met reporters, an alkyne phenylalanine (Phe) analog, *para*-ethynylphenylalanine (PEP) (Fig. 5c) can also function as a cell-selective amino acid reporter in combination with a phenylalanine-tRNA-synthetase mutant (A294G of *pheS*) (Grammel et al. 2012). Amino acids such as Phe are present in higher frequencies than Met and the possible combination of these reporters may allow greater coverage of proteomes or even selective labeling of two different proteomes. An extension of these

strategies and especially translation to eukaryotic systems might allow selective labeling of individual cell types in multicellular organisms in the future.

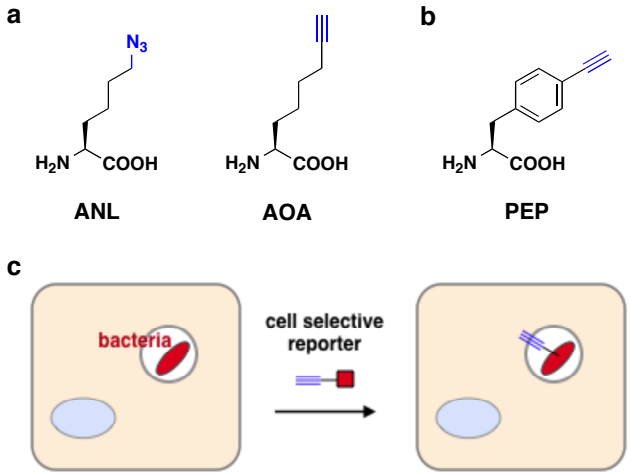


Figure 5 Cell selective proteome labeling with engineered tRNA synthetases. a) Engineered methionyl-tRNA-synthetase reporters: azidonorleucine (ANL) and 2-aminooctynoic acid (AOA). b) *para*-ethynylphenylalanine (PEP) can be incorporated residue-selectively instead of the amino acid phenylalanine (Phe) through the expression of an engineered phenylalanyl-tRNA-synthetase. c) The expression of engineered tRNA synthetases allows the selective incorporation of the bioorthogonal amino acid reporters ANL, AOA or PEP into bacterial proteomes during infection – alkynyl reporter is shown.

1.2.2 Bioorthogonal labeling and capture of nascent polypeptides.

Notwithstanding the advances made possible with amino acid reporters, some limitations should be noted. First and foremost, the global replacement of natural amino acids with unnatural counterparts might destabilize proteins and interfere with activity. Short pulse-labeling periods can minimize this issue and prevent toxicity effects. Secondly, the Met and Phe amino acid reporter incorporation is proportional to the number of residues per protein, which could misrepresent the abundance of proteins. To address these issues,

Salic and colleagues have shown that O-propargyl-puromycin (OP-puro) (Fig. 4c), an alkyne analog of the covalent translational inhibitor puromycin, can be used to label nascent polypeptide chains in mammalian cell culture as well as in mice (J. Liu et al. 2012). OP-puro labeled proteins can be imaged by fluorescence microscopy and captured to retrieve nascent polypeptide chains. In contrast to amino acid reporters, which are incorporated into the proteome with a certain bias, exactly one OP-puro is attached to each nascent polypeptide, an important feature for comparative proteomic studies. It should be noted that OP-puro-protein conjugates are rapidly degraded by the proteosome (J. Liu et al. 2012), a fact to be considered for future proteomic studies. Collectively, these bioorthogonal protein synthesis reporters have provided more sensitive methods for monitoring protein dynamics and have already revealed important biological mechanisms, which should complement recent large-scale ribosome profiling methods (Ingolia et al. 2011; Heiman et al. 2008).

1.3 Bioorthogonal chemical reporters for posttranslational modifications

In addition to nucleic acid and protein synthesis, bioorthogonal chemical reporters have been developed for diverse posttranslational modifications (PTMs) in bacteria, eukaryotes and plants (Fig. 6-8).

1.3.1 Glycosylation.

The covalent modification of proteins with mono or oligosaccharides is one of the most ubiquitous and most complex forms of PTMs in biology (Varki 1999). Glycoproteins decorate the exterior of eukaryotic cells and mediate a wide range of interactions with soluble extracellular factors, neighboring cells and microbes. Additionally, glycosylation plays an important role inside cells by regulating protein trafficking, turnover and signal

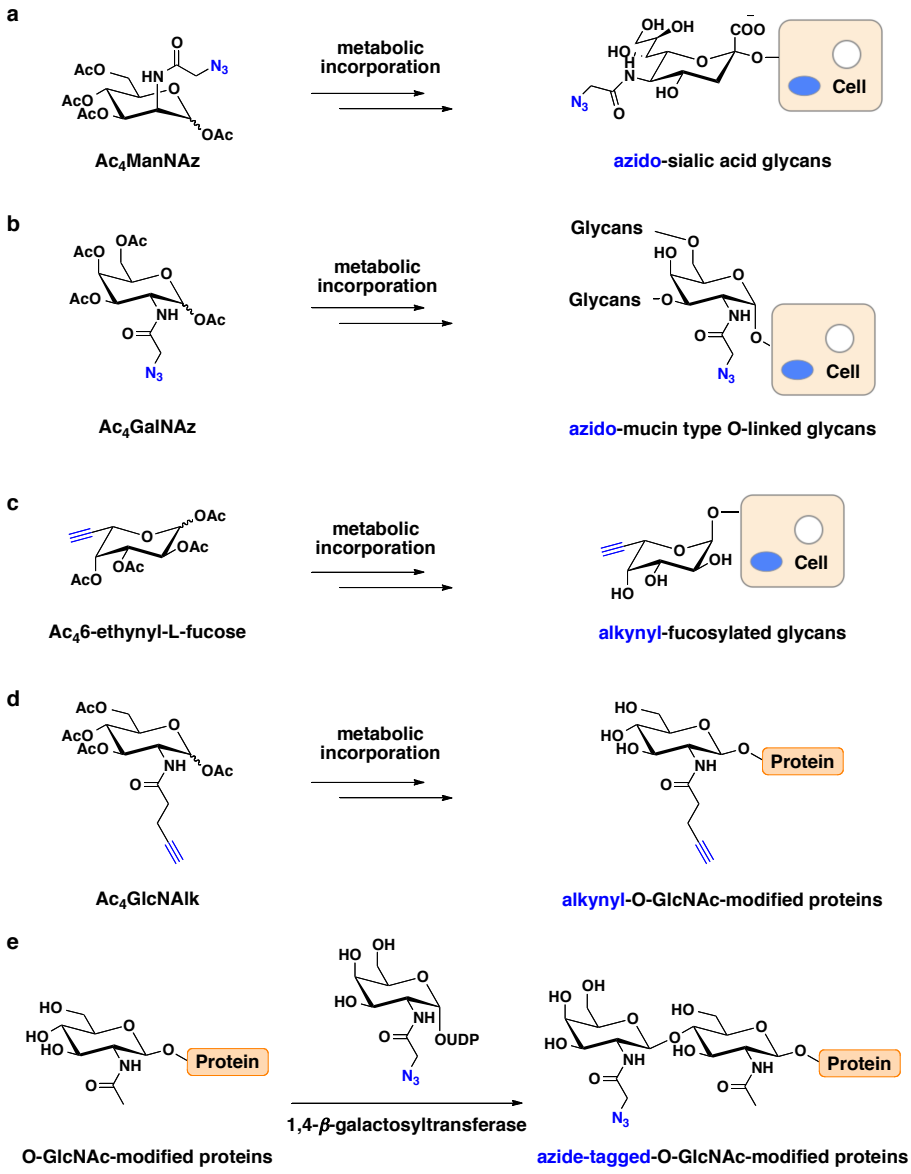


Figure 6 Chemical reporters can be used to probe different glycans. a) Peracetylated N-azidoacetyl-D-mannosamine (Ac_4ManNAz) is metabolically incorporated into sialic acid glycans. b) Peracetylated N-azidoacetyl-D-galactosamine (Ac_4GalNAz) is metabolically incorporated into mucin type O-linked glycans. c) Peracetylated 6-ethynyl-L-fucose ($Ac_4$6-ethynyl-L-fucose) is metabolically incorporated into fucosylated proteins. d) Peracetylated N-alkynylacetyl-D-glucosamine (Ac_4GlcNAlk) is metabolically incorporated into O-GlcNAc modified proteins. e) O-GlcNAc modified proteins can be chemoenzymatically labeled using a 1,4-β-galactosyltransferase that transfers GalNAz onto O-GlcNAc modified proteins.

transduction. Indeed, the complex roles of glycans and the challenges associated with their detection and synthesis was a major driving force for the development of bioorthogonal ligation reactions (Prescher & Carolyn R Bertozzi 2006; C R Bertozzi & Kiessling 2001). The Staudinger ligation reaction and the first azide monosaccharide reporter, N-azidoacetyl-D-mannosamine (ManNAz, Fig. 6a) was developed in the context of monitoring cell surface sialic acid-containing glycans (Saxon & C R Bertozzi 2000). This landmark study has spurred the development of many bioorthogonal chemical reporters highlighted in this review and in particular other azide- and alkyne-functionalized monosaccharides for sialic acid glycoconjugates (Chang et al. 2009; Luchansky et al. 2004), mucin-type O-linked glycoproteins (Fig. 6b) (Hang et al. 2003; Zaro, Bateman, et al. 2011), fucosylated glycans (Fig. 6c) (Rabuka et al. 2006; Sawa et al. 2006), and intracellular O-GlcNAc modified proteins (Fig. 6d) (Khidekel et al. 2003; Vocadlo et al. 2003; Zaro, Y.-Y. Yang, et al. 2011), which have improved the detection and proteomic analysis of these glycoprotein families (Zaro, Y.-Y. Yang, et al. 2011; Hanson et al. 2007). In addition to profiling glycoproteins, the metabolic incorporation of chemical reporters into cell surface glycans can be used to select cell populations with aberrant glycosylation levels (Yarema et al. 2001). An early example of this strategy using a ketone-functionalized sialic acid precursor and hydrazide labeling revealed mutations in glycan biosynthetic enzymes associated with congenital human diseases such as sialuria (Yarema et al. 2001). Remarkably, these glycan reporters in combination with activated cyclooctyne probes have now enabled the dynamic imaging of cell surface glycans in live developing zebrafish embryos and in worms (Laughlin & Carolyn R Bertozzi 2009a; Laughlin et al. 2008b), which opens the door for differential analysis of

glycan biosynthesis in specific tissues and cell types *in vivo*. Beyond eukaryotes, glycan chemical reporters have also been incorporated into bacteria such as *Campylobacter jejuni* (F. Liu et al. 2009), *Helicobacter pylori* (Koenigs et al. 2009) and *Escherichia coli* (Dumont et al. 2012) as well as plants (Anderson et al. 2012).

Metabolic and chemoenzymatic chemical reporters are providing new insight into O-GlcNAc modification of cytoplasmic and nuclear proteins. The discovery of GlcNAc on serine (Ser) and threonine (Thr) residues was unexpected, but this dynamic and regulated PTM is emerging as an important sensor of metabolic state of cells and is associated with diseases such as diabetes, neurodegeneration and cancer (Hart et al. 2011). Initial studies demonstrated that chemical reporter GlcNAz could be utilized by the hexosamine salvage pathway and incorporated into O-GlcNAc modified proteins in mammalian cells (Vocadlo et al. 2003) and even be used for proteomic studies (Sprung et al. 2005). More recently, an alkyne-modified GlcNAc analog (GlcNAlk, Fig. 6d) has enabled more sensitive bioorthogonal detection as well as greater proteomic coverage and uncovered glycosylation of the E3 ubiquitin ligase NEDD4 (Zaro, Y.-Y. Yang, et al. 2011). Alternatively, Hsieh-Wilson and colleagues have developed a chemoenzymatic method that allows labeling of O-GlcNAc modified proteins in tissue/cell lysates with a ketone (Khidekel et al. 2003) or azide (Clark et al. 2008) group using an engineered beta-1,4-galactosyltransferase (Fig. 6e). The tagged O-GlcNAc-modified proteins can then be evaluated by fluorescence (Clark et al. 2008), mass-shift western blotting (Rexach et al. 2010) and quantitative proteomic s(Khidekel et al. 2007), which has revealed differential dynamics of O-GlcNAc modified proteins in neurons and O-GlcNAc regulation of CREB-mediated gene expression *in vivo* (Rexach et al. 2012). Beyond these studies,

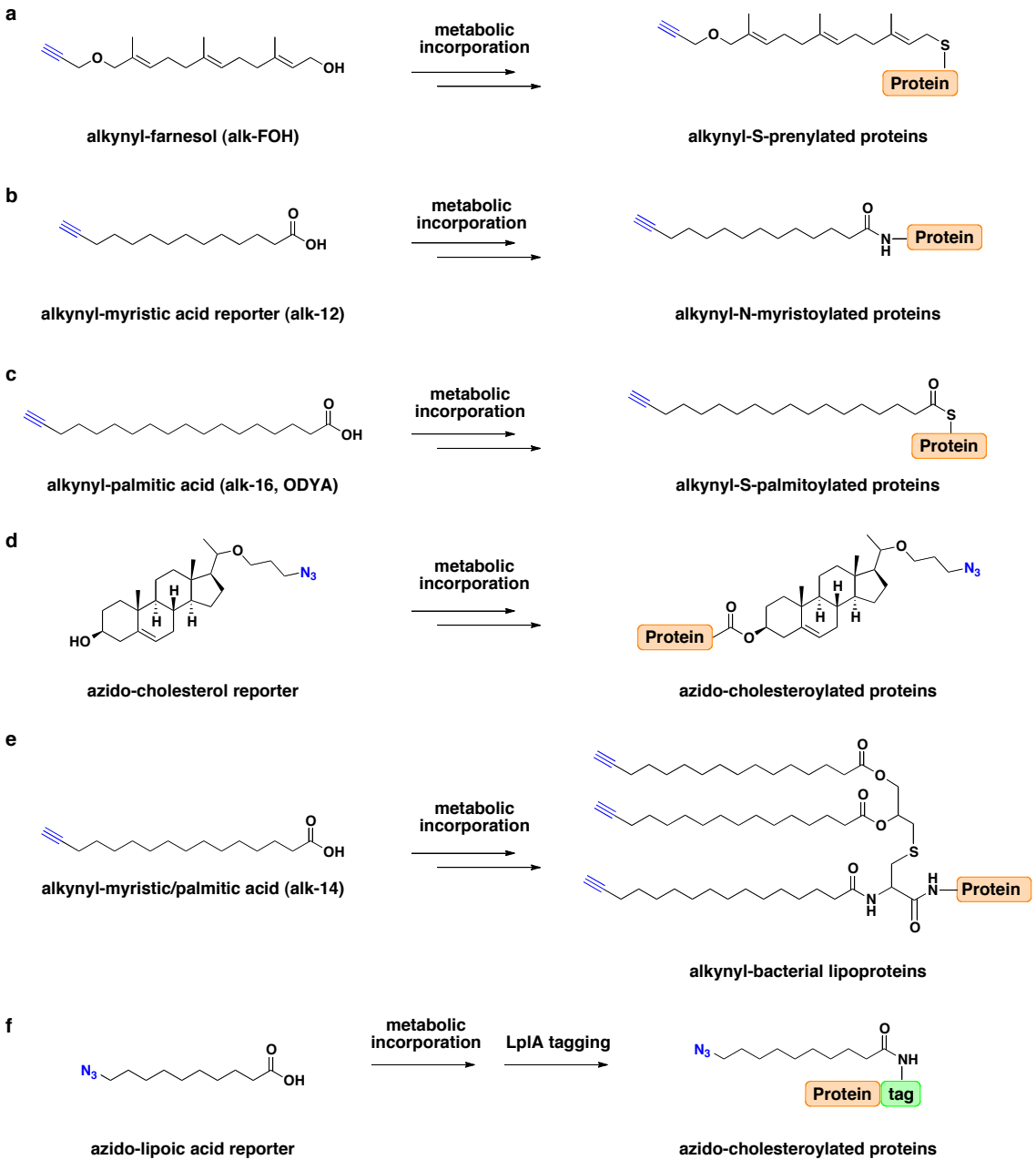
azide- and alkyne sugar nucleotide reporters have also been useful for *in vitro* profiling glycosyltransferase substrate specificities (Pratt et al. 2004) as well as *in vivo* imaging of glycans (T. Zheng et al. 2011; Baskin et al. 2010). These studies highlight the utility of bioorthogonal chemical reporters for monitoring glycan biosynthesis, glycoproteomics and *in vivo* imaging of glycans in model organisms.

1.3.2 Lipidation.

Lipid modifications of proteins modulate their membrane affinity, localization and trafficking for cell signaling (Resh 2006). Even though some forms of protein lipidation can now be predicted based on conserved amino acid motifs (Resh 2006), direct biochemical detection of lipidated proteins is still required for functional studies, discovery of unannotated substrates and novel types of lipidation. For example, S-farnesylation or geranylgeranylation of cysteine (Cys) residues on the C-terminus of proteins can be predicted by bioinformatics based on conserved CaaX or CC motifs (Resh 2006), respectively. However, biochemical detection of S-prenylation is still crucial since many of these lipidated proteins are small GTPases (e.g. K/H/N-Ras) that control cell signaling and when mutated are major drivers of oncogenesis (Karnoub & Weinberg 2008). The corresponding prenyltransferases are thus active targets of drug development for cancer and other human diseases (Berndt et al. 2011; Meta et al. 2006). To this end, azide(Kho et al. 2004) and alkyne (Charron et al. 2011; DeGraw et al. 2010) isoprenoid reporters have been developed that allow bioorthogonal detection and proteomic analysis S-prenylated proteins in mammalian cells. Interestingly, our proteomic analysis of alkyne-farnesol (alk-FOH) (Fig. 7a) modified proteins in

Figure 7 Different forms of protein lipidation can be analyzed with chemical reporters. a) Alk-FOH allows metabolic labeling of S-farnesylated and S-geranylgeranylated proteins in mammalian cells. b) und c) Alkynyl-fatty acid reporters of different chain length enable selective metabolic labeling of N-myristoylated or S-palmitoylated proteins. d) An azide analog of cholesterol can be metabolically incorporated onto the C-terminus of sonic hedgehog. e) Alkynyl-fatty acid reporters can also be metabolically incorporated into bacterial lipoproteins. f) The azido-decanoic acid can function as a substrate for lipoic acid ligase (LplA) for sequence-specific tagging of proteins.

Figure 7


a
metabolic incorporation
Protein
alkynyl-farnesol (alk-FOH)
alkynyl-S-prenylated proteins
b
metabolic incorporation
Protein
alkynyl-myristic acid reporter (alk-12)
alkynyl-N-myristoylated proteins
c
metabolic incorporation
Protein
alkynyl-palmitic acid (alk-16, ODYA)
alkynyl-S-palmitoylated proteins
d
metabolic incorporation
Protein
azido-cholesterol reporter
azido-cholesteroylated proteins
e
metabolic incorporation
Protein
alkynyl-myristic/palmitic acid (alk-14)
alkynyl-bacterial lipoproteins
f
metabolic incorporation
LplA tagging
Protein
tag
azido-lipoic acid reporter
azido-cholesteroylated proteins

macrophages has enabled large-scale profiling of S-prenylated proteins and revealed isoform-specific S-farnesylation of zinc-finger antiviral protein (ZAP) is crucial for the membrane targeting and antiviral activity of this interferon-induced effector (Charron G et al submitted).

Azide and alkyne fatty acids are also utilized by native enzymes in bacteria and eukaryotes(Hang et al. 2011). Myristoylation of N-terminal glycine (Gly) residues by N-myristoyltransferases that occurs primarily cotranslationally (Farazi et al. 2001), is readily labeled and visualized by myristic acid reporters in mammalian cells (Fig. 7b) (Hang et al. 2007; Charron et al. 2009a; D. D. O. Martin et al. 2008). In contrast to N-myristoylation, which is stable and constitutive, the S-palmitoylation of Cys residues can be reversible, regulated by a family of DHHC-palmitoyltransferases (DHHC-PATs), not readily predicted by primary amino acid sequence and particularly challenging to biochemical analyze (Linder & Deschenes 2007). Metabolic labeling with alkyne palmitic acid reporters (Fig. 7c) in combination with bioorthogonal detection by our laboratory (Charron et al. 2009a) and others (B. R. Martin & Cravatt 2009; Hannoush & Arenas-Ramirez 2009; Yap et al. 2010) has largely solved this problem, enabling fluorescence visualization and proteomic analysis of S-palmitoylated proteins (B. R. Martin & Cravatt 2009; B. R. Martin et al. 2011; J. P. Wilson et al. 2011; Yount et al. 2010). Notably, bioorthogonal palmitoylome profiling and hydroxylamine-mediated acyl biotin exchange (ABE) proteomics (Wan et al. 2007; Roth et al. 2006) has significantly increased the diversity of S-palmitoylated proteins in eukaryotes (Hang & Linder 2011) and even revealed new fatty-acylated proteins in innate immunity (Yount et al. 2010) and neuronal activity (Kang et al. 2008). In addition to lipidation of cytoplasmic proteins,

secreted proteins are also lipid-modified and readily labeled by chemical reporters as highlighted by fatty acid and cholesterol reporter labeling of Wnt (Gao et al. 2011) and sonic hedgehog proteins (Fig. 7d) (Heal et al. 2011), respectively. It should be noted that bioorthogonal chemical reporters have also been developed for monitoring the metabolism and trafficking of free, unconjugated lipids in mammalian cells (Neef & Carsten Schultz 2009; Milne et al. 2010).

Bioorthogonal lipid reporters have been especially useful for monitoring protein S-palmitoylation dynamics. For instance, tandem pulse-chase labeling of T cells with protein synthesis and palmitoylation reporters provides a robust method for quantifying S-acylation dynamics relative to protein turnover, which revealed accelerated depalmitoylation of signaling kinases such as Lck upon T cell activation (M. M. Zhang et al. 2010). In addition, the large-scale analysis of S-palmitoylation dynamics using bioorthogonal labeling and SILAC showed that many proteins are stably fatty-acylated while other proteins involved in cell signaling undergo rapid S-acylation cycles (B. R. Martin et al. 2011). To understand how specific S-palmitoylated proteins are regulated, our laboratory has employed palmitate chemical reporters in the model system *Schizosaccharomyces pombe* (fission yeast) to evaluate the distinct functions of different DHHC-PATs. Through pulse labeling of synchronized cell populations and proteomic analysis of wild type and mutant *S. pombe* strains, we have discovered that the regulated expression of a single DHHC-PAT can control S-palmitoylation of specific protein substrates, cell signaling and ultimately the ability of fission yeast to undergo meiosis or cellular differentiation (Mingzi M Zhang n.d.). These early applications of bioorthogonal

lipid reporters are beginning to reveal the differential modification of S-palmitoylated proteins and their mechanisms of regulation as well as their impact on cellular function.

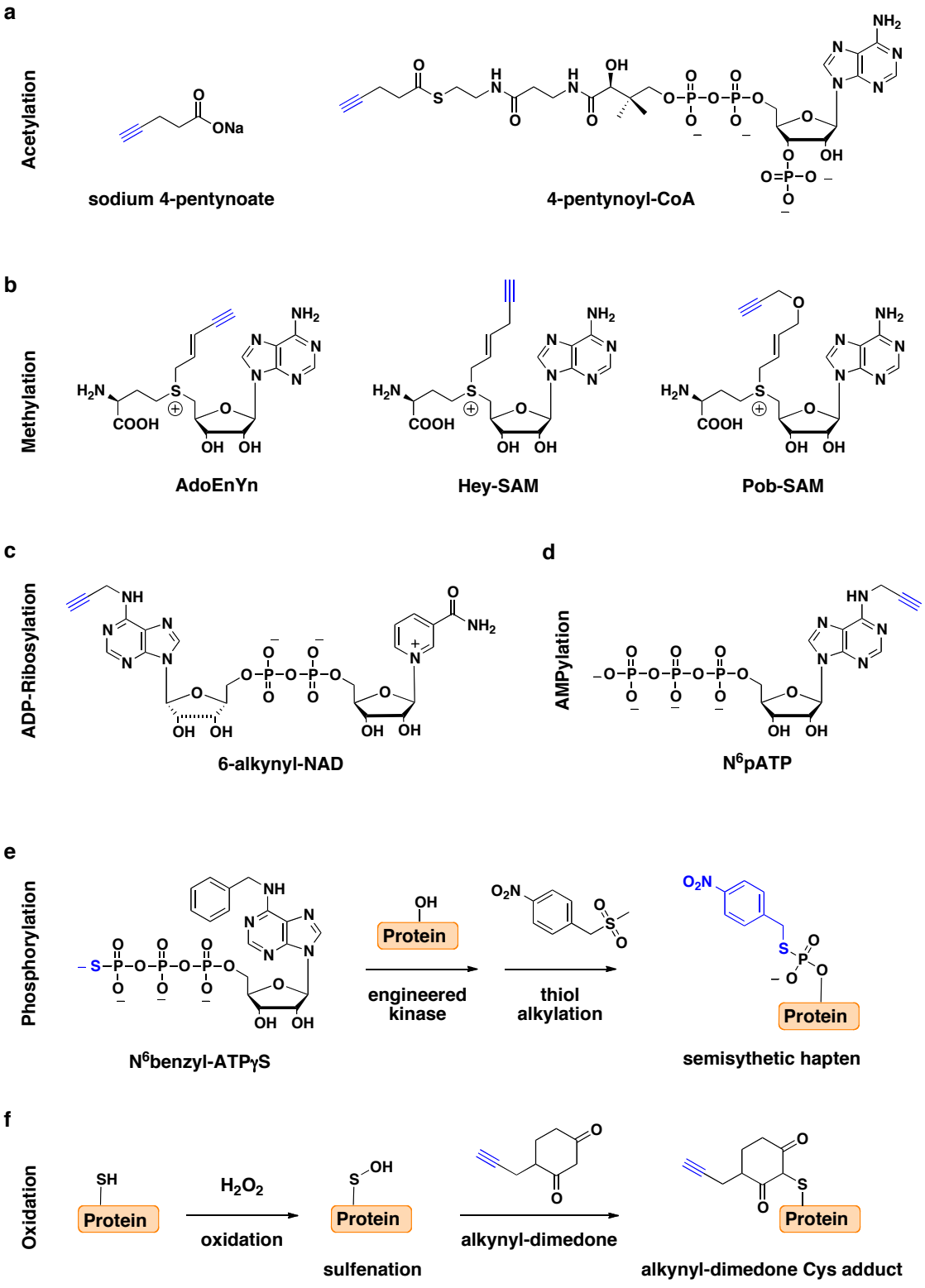
Lipid reporters also allow labeling of bacterial proteins and sequence-specific tagging of engineered proteins. Protein lipidation in bacteria involves the installation of a diacylglyceryl group onto N-terminal Cys residue within a well-defined lipobox-motif (Sankaran & Wu 1994). Although algorithms are available for the prediction of canonical bacterial lipoproteins(Babu et al. 2006), the majority of lipidated proteins in bacteria are poorly characterized and may exist in other forms. For example, the *E. coli* protein toxin hemolysin is activated through fatty-acylation of a lysine (Lys) residue (Issartel et al. 1991). In this regard, the application of fatty acid reporters to bacteria allows both the rapid profiling of canonical bacterial lipoproteins as well as the discovery of unconventionally lipidated proteins in bacteria (Fig. 7e) (Rangan et al. 2010). In addition to fatty acids, phosphopantetheinyl reporters can be directly transferred onto bacterial proteins involved in polyketide biosynthesis (Mercer et al. 2009). For protein tagging applications, these lipid reporters also allows site-specific labeling of recombinant proteins containing N-myristoylation and S-prenylation motifs by co-expression of the associated lipid transferases in bacteria (Heal et al. 2008; Gauchet et al. 2006; Wollack et al. 2009; Yi et al. 2010). Alternatively, lipoic acid ligase tagging has been developed for sequence-specific tagging of proteins in mammalian cells for fluorescence imaging applications (Fig. 7f) (Yao et al. 2012; Uttamapinant et al. 2010). As specific antibodies for lipid modifications have been difficult to generate, bioorthogonal chemical reporters have been valuable for the detection and discovery of lipidated proteins, analysis of lipidation dynamics and regulation as well as protein engineering (Hang et al. 2011).

1.3.3 Acetylation.

Protein acetylation of Lys residues plays important roles in epigenetic regulation of transcription, signaling and metabolism (X.-J. Yang & Seto 2008; Kouzarides 2007). With the explosion of Lys acetylated proteins from recent proteomic studies in mammalian cells and bacteria (Choudhary et al. 2009; Q. Wang, Y. Zhang, et al. 2010), unraveling what substrates are regulated by specific Lys acetyltransferases (KATs) and deacetylases (KDACs) in various signaling pathways is a major challenge. In this regard, alkyne analogs of acetate and acetyl-CoA such as 4-pentynoate and 4-pentynoyl-CoA can function as chemical reporters of Lys acetylation in cells and *in vitro*, respectively (Fig. 8a) (Hang & Linder 2011). Metabolic labeling of mammalian cells with 4-pentynoate readily labels known Lys acetylated proteins such as histones, which allows simple fluorescence profiling of acetylated proteins in different cell types (Y.-Y. Yang, Ascano & Hang 2010). In addition, 4-pentynoyl-CoA is utilized by the KAT p300 for *in vitro* bioorthogonal profiling of candidate protein substrates in cell lysates (Y.-Y. Yang, Ascano & Hang 2010). As acetylation also occurs on the N-terminus of proteins(Polevoda & Sherman 2003) and Ser/Thr residues (Mukherjee et al. 2007), acetylation reporters should also facilitate the analysis of these modifications and their associated enzymes. Other forms of Lys acylation such as propionylation, butrylation, malonylation, succinylation and crotylation have also been reported (M. Tan et al. 2011; Z. Zhang et al. 2011; Chen et al. 2007; Du et al. 2011), which may also benefit from specific bioorthogonal reporters or labeling strategies in the future.

Figure 8 Many posttranslational modifications can be analyzed with cofactor analogs or selective chemical modification reagents. a) Sodium-4-pentynoate allows metabolic labeling of acetylated proteins, while 4-pentynoyl-CoA can be used for *in vitro* analysis of acetyltransferase substrates. b) Numerous S-adenosyl-L-methionine reporters have been developed (AdoEnYn, Hey-SAM, Pob-SAM), whose extended side chains are transferred by protein methyltransferases onto their native substrates. c) 6-alkynyl-NAD is a chemical reporter for ADP-mono and poly-ribosylation. d) N^6pATP serves as a chemical reporter for protein AMPylation. e) N^6benzyl-ATPγS serves as a substrate for engineered kinases and allows chemical substrate modification with a nitrobenzyl-hapten for antibody detection. f) The oxidative modification of cysteines to sulfenic acid can be monitored by reaction with an alkynyl-dimedone reagent for subsequent bioorthogonal detection of protein sulfenation.

Figure 8


a
Acetylation
sodium 4-pentynoate
4-pentynoyl-CoA
b
Methylation
AdoEnYn
Hey-SAM
Pob-SAM
c
ADP-Ribosylation
6-alkynyl-NAD
d
AMPylation
N6pATP
e
Phosphorylation
N6benzyl-ATPγS
Protein
engineered kinase
thiol alkylation
semisythetic hapten
f
Oxidation
Protein
H2O2
oxidation
sulfenation
alkynyl-dimedone
alkynyl-dimedone Cys adduct

1.3.4 Methylation.

The modification of Lys or arginine (Arg) residues with one or more methyl groups is a common PTM in epigenetics and various signaling pathways (Y.-Y. Yang et al. 2011). The general cofactor for nearly all cellular methylation reactions is S-adenosyl-L-methionine (AdoMet or SAM), from which the methyl group is transferred onto proteins, nucleic acids or other substrates. The development of bioorthogonal methylation reporters has therefore attracted considerable interest by many groups. Initially, Weinhold and colleagues demonstrated that various DNA methyltransferases are able to transfer allylic or propargylic groups from SAM reporters onto DNA substrates (Bedford & Clarke 2009; Rathert et al. 2008). The authors hypothesized that the improved reactivity of these reagents could be explained by the conjugative stabilization of the forming *p* orbital on the reactive carbon during the *S*N2-reaction. This observation was instrumental in the development of the first alkyne SAM reporter for site-selective labeling of protein methyltransferase substrates such as histone H3 (AdoEnYn) (Fig. 8b) (Dalhoff et al. 2006). AdoEnYn possesses an extended unsaturated pentameric side chain, with a double bond in the β-position relative to the sulfonium center and a terminal alkyne group. A suite of additional AdoMet reporters capable of labeling Lys and Arg methyltransferase substrates with terminal alkyne groups was subsequently reported (W. Peters et al. 2010). A propargyl AdoMet reporter with the smallest alkyne substituent was utilized by the Lys methyltransferase SETDB1 but unfortunately not active with other tested methyltransferases such as SET7, PRMT1, PRDM8 (Binda et al. 2011; R. Wang, W. Zheng, et al. 2011; Islam et al. 2011). Alternatively, Luo and colleagues have identified protein methyltransferase mutants that can utilize other SAM chemical reporters. After screening a small panel of cofactor analogs against a series of active site

methyltransferase mutants, a G9a mutant that is able to accept *(E)*-hex-2-en-5-ynyl-SAM (Hey-SAM) and modify recombinant histone H3 was identified (Fig. 8b) (Binda et al. 2011). In addition to Lys methyltransferases, a double mutant of protein Arg methyltransferase PRMT1 that can utilize 4-propargyloxy-but-2-enyl-SAM (Pob-SAM) was developed (Fig. 8b) (Islam et al. 2011). Interestingly, when added to cell lysates, the double mutant of PRMT1 in combination with Pob-SAM specifically labeled a number of proteins. In both cases the most active SAM reporters carried an β-unsaturation consistent with the studies by Weinhold and coworkers. This approach in combination with mass spectrometry-based protein identification should facilitate the identification of specific methyltransferase substrates in the future.

1.3.5 ADP-Ribosylation.

Proteins can be reversibly modified with mono-ADP-ribose on a range of different residues (R. Wang, W. Zheng, et al. 2011). This modification was initially described as the virulence mechanism for a number of different bacterial toxins, but is now appreciated as a common PTM with many important biological functions in eukaryotes as well (Hottiger et al. 2010; Hassa et al. 2006). Protein mono-ADP-ribosylation is carried out by a diverse set of mono-ADP-ribosyltransferases (MARTs) that require the cofactor nicotinamide adenine dinucleotide (NAD), from which the ADP-ribose moiety is transferred onto the protein substrate. Additionally, in eukaryotes the ADP-ribose modification can be extended by addition of ADP-ribosyl groups to the 2'-OH of the ribose moieties to form poly-ADP-ribosylated proteins. This modification is carried out by poly-ADP-ribosylpolymerases (PARPs) that also use NAD as donor for the ADP-ribosyl group. The growing list of ADP-ribosylation substrates and potential new

ribosylation enzyme classes has motivated the development of ADP-ribosylation reporters for substrate detection and identification (Hassa et al. 2006). Lin and colleagues initially generated 8-alkyne-NAD and 6-alkyne-NAD (Fig. 8c) and showed that both of these chemical reporters are accepted by diphteria toxin (DT), a well-characterized MART that modifies elongation factor eEF-233. The analysis of these chemical reporters with PARPs, PARP-1 and Tankyrase-1, revealed 6-alkyne-NAD was most effective for monitoring the modification of p53 and TRF-1, respectively (Du et al. 2009; Jiang et al. 2010). Incubation of cell lysates with recombinant PARP-1 and 6-alkyne-NAD enabled bioorthogonal detection and identified of poly-ADP-ribosylated proteins, demonstrating the utility of these reagents for ADP-ribosylation analysis.

1.3.6 AMPylation.

Though originally discovered in the context of glutamine synthetase regulation in *E. coli* (Jiang et al. 2010), AMPylation (adenylylation) has become more widely appreciated since the discovery of secreted bacterial effector enzymes that modify host small GTPases during infection (Kingdon et al. 1967, Woolery et al. 2010). AMPylation describes the enzymatic transfer of an adenosine monophosphate (AMP) from ATP onto hydroxyl protein side chains under the formation of a phosphodiester bond. The conserved protein domains (fic domain, adenylyl transferase domain) responsible for the AMPylation activity are widely present among prokaryotes and eukaryotes, suggesting a ubiquitous role for this PTM (Yarbrough et al. 2009; Worby et al. 2009; Müller et al. 2010). Recent studies have also revealed specific phosphoesterases that can remove the AMP moiety suggesting that this modification can be dynamic and regulated (Kinch et al. 2009). The re-emergence and prevalence of potential AMPylation enzymes has motivated

the development of specific antibodies (Hao et al. 2011), mass spectrometry detection methods (Li et al. 2011) as well as an bioorthogonal AMPylation reporter from our laboratory. Modification of ATP at the N^6-position with a propargyl group provides an alkyne-decorated AMPylation reporter (N^6pATP) (Fig. 8d) that can be enzymatically transferred onto specific sites of cognate protein substrates by all known classes of AMPylation enzymes (Yan Li et al. 2011). Furthermore, N^6pATP can be used to proteomically identify AMPylation substrates at endogenous levels in cell lysates (Grammel et al. 2011), which should facilitate the characterization of AMPylated proteins in diverse organisms.

1.3.7 Phosphorylation.

The reversible addition of phosphate onto Ser, Thr or tyrosine (Tyr) residues on proteins are perhaps the best studied PTMs (Pawson & Scott 2005; Tarrant & Cole 2009), for which specific antibodies and quantitative proteomic methods have been developed(Macek et al. 2009). Nonetheless, the large number of phosphoproteins and regulatory enzymes (kinases and phosphatases) with potential overlapping biochemical activity makes dissecting phosphorylation-dependent signaling pathways still very challenging. Robust biochemical methods for directly monitoring enzyme-specific phosphorylation events are desirable. The identification of adenosine triphosphate (ATP) analogs that can be utilized by engineered kinase mutants by Shokat and coworkers has been instrumental in breaking the degeneracy of kinase activities and allowed the identifying specific protein substrates (Fig. 8e) (Shogren-Knaak et al. 2001; C. Zhang et al. 2005). To distinguish kinase-specific phosphorylation events from existing phosphoproteins in cell lysates, γ-thioATP was employed as substrate such that the

unique reactivity of the sulfhydryl group could be selectively alkylated with detection and enrichment reagents (Carlson et al. 2010; Allen et al. 2005; Allen et al. 2007; Blethrow et al. 2008). The alkylation of thiophosphate groups with a nitrobenzyl-hapten allowed specific antibody recognition of this semisynthetic epitope over alkylated Cys residues for identification of kinase-specific substrates (Allen et al. 2005; Allen et al. 2007). Although this system does not utilize alkyne- or azide-functionalized reagents, this two-step labeling protocol is itself *bioorthogonal* and showcases the utility of bump-hole approaches for identifying enzyme-specific substrates in systems with overlapping biochemical activity.

1.3.8 Oxidation.

The oxidation of amino acid side chains can result in significant rearrangements of protein structure as in the case of the green fluorescent protein (GFP) or more subtle changes such as hydroxylation of amino acid residues. These PTMs are perhaps the least amendable to bioorthogonal labeling methods, nonetheless, specific chemical reagents have been identified that enable bioorthogonal detection of some protein oxidation reactions. For example, the functionalization of dimedone with an azide or alkyne enabled the selective labeling sulfenic acid modified proteins in cells and subsequent bioorthogonal detection and proteomic analysis (Fig. 8f) (Leonard et al. 2009; Paulsen et al. 2012). Interestingly, Carroll and coworkers demonstrated that stimulation of the epidermal growth factor receptor (EGFR) results in elevated hydrogen peroxide levels which in turn oxidizes Cys797 of EGFR to enhance its tyrosine kinase activity. These observations suggest an interestingly positive feedback loop to amplify EGFR-dependent signaling (Paulsen et al. 2012). By taking advantage of the unique chemical reactivity,

bioorthogonal reporters can be developed for specific forms of protein oxidation and reveal previously unappreciated mechanisms of cell signaling.

1.3.9 Summary and Future Outlook.

Bioorthogonal reactions in conjunction with chemical reporters are providing powerful ways of tracking biomolecules in living systems and illuminating new aspects of biology. While the past decade has seen incredible advances in bioorthogonal reactions and specific chemical reporters (Carolyn R Bertozzi 2011), equally momentous leaps will be necessary to apply these chemical tools to complex biological problems. Some of the limitations of bioorthogonal detection are already being addressed with faster and more selective chemical reactions (Jewett & Carolyn R Bertozzi 2010) as well as label-free chemical imaging methods (Pezacki et al. 2011). On the chemical reporter end, many of the cofactor-based reporters only work with enzymes *in vitro* and would benefit from the development of caged precursors or engineered enzymes for *in vivo* labeling applications. Moreover, the generation of enzyme-specific chemical reporters akin to the systems developed for kinases is needed for several PTMs to characterize specific protein targets. These two-component systems for bioorthogonal chemical reporters could then be introduced into model organisms for cell-specific imaging and proteomics applications. By virtue of their specific and dynamic incorporation into biomolecules, bioorthogonal chemical reporters can provide snapshots of the chemistry inside of cells that integrate the activity of metabolic pathways and specific enzymes in response to environmental cues. The development of new bioorthogonal chemical reporters and their application should help translate genomic information into biological function and uncover key mechanisms of human diseases.

2 Cell-selective labeling of *Salmonella typhimurium* with 2-amino-7-octynoic acid (AOA) during infection[1]

2.1 Abstract

2.2 Background

Bacterial pathogens have evolved sophisticated mechanisms to evade host defenses and cause disease (Bhavsar et al. 2007). The emergence of new and antibiotic-resistant bacterial pathogens demands a better understanding of virulence mechanisms for antibacterial drug discovery. While the discovery of bacterial toxins, quorum sensing and protein secretion pathways has revealed some key virulence mechanisms, the precise mechanisms by which intracellular bacterial pathogens subvert host immune responses are still unclear (Bhavsar et al. 2007). The analysis of individual virulence factors has demonstrated that bacterial pathogens alter their protein expression to infect and replicate in host tissues (Haraga et al. 2008). However, the system-wide identification and analysis of bacterial proteins that are uniquely expressed or secreted during infection is paramount for understanding mechanisms of bacterial pathogenesis (Haraga et al. 2008; Rodland et al. 2008). Comparative genomics and mutagenesis studies have revealed bacterial genes that are important for infection, but their precise biochemical mechanisms and temporal expression pattern can be elusive due to post-transcriptional regulation. Direct biochemical analysis of bacterial proteomes during infection is needed. The large excess

---

[1] This work was carried out in collaboration with Mingzi M Zhang (The Rockefeller University), who cloned metG from *E. coli* into pCOLAduet-1. Mass spectrometry runs for proteomic analysis were carried out by the Proteomics Resource Center of The Rockefeller University.

of host proteins in mixed pathogen-host lysates presents a significant challenge for proteomic analysis of bacterial proteins during infections and even after physical isolation of intact bacteria significant amounts of host proteins still remain (Rodland et al. 2008; Shi et al. 2006). This is particularly important since many bacterial virulence factors are often expressed at low levels (Haraga et al. 2008). New strategies are therefore required to selectively enrich bacterial proteins from host proteomes for their analysis during infection.

The incorporation of unnatural amino acids in bacteria has provided new methods to differentiate bacterial proteins from host proteomes. For example, the incorporation of phenylalanine analogs in mycobacteria by amber stop codon suppression technology has enabled the selective labeling of green fluorescent protein expressed in *M. tuberculosis* during intracellular infection of macrophages (F. Wang et al. 2010). Alternatively, alkyne- or azide-functionalized methionine (Met) surrogates can be incorporated by the endogenous methionyl-tRNA synthetase (MetRS) into bacterial proteomes (Beatty et al. 2005; Kiick et al. 2002). These amino acid chemical reporters allow the metabolic labeling of newly synthesized proteins, which in combination with bioorthogonal ligation methods, such as Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) can be used to detect and identify Met-containing proteins (Dieterich et al. 2006). Additionally, MetRS mutants have been identified in *E. coli* (MetG) that can incorporate azidonorleucine (ANL, Figure 9a), a Met surrogate that is not efficiently activated by the wild-type MetRS or other endogenous aminoacyl-tRNA synthetases (Link et al. 2006; Tanrikulu et al. 2009). ANL can therefore be used as an orthogonal amino acid reporter to selectively label proteins in non-pathogenic *E. coli* by bacterial expression of MetRS mutants in the

presence of mammalian cells (Ngo et al. 2009). While these studies have demonstrated selective targeting of bacterial proteomes in the presence of host cells, the analysis of endogenously expressed bacterial proteins during infection has not been achieved using unnatural amino acid reporters (F. Wang et al. 2010; Ngo et al. 2009). The following chapter of this thesis describes a new orthogonal alkynyl-amino acid reporter (2-amino-7-octynoic acid, AOA, Figure 9a) for specific imaging and enrichment of bacterial proteomes during infection of mammalian cells, which we applied to the intracellular bacterial pathogen *Salmonella typhimurium* (Figure 9b).

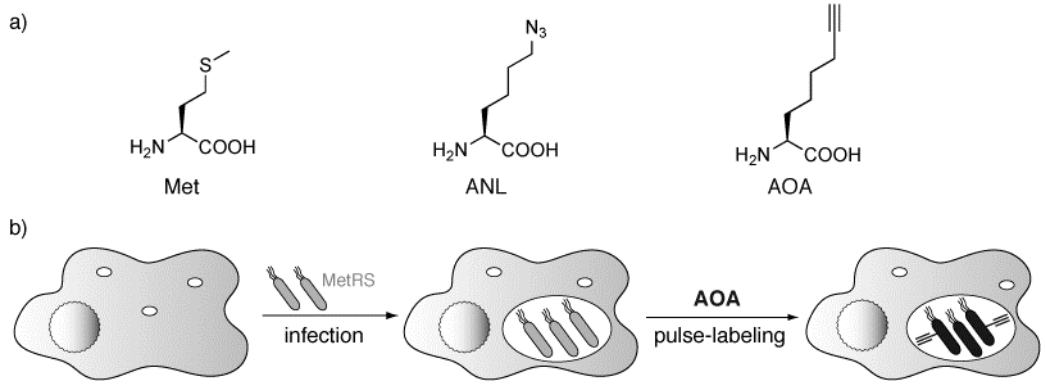


Figure 9 Orthogonal (cell-selective) amino acid reporters for bacterial translation.

2.3 Results

2.3.1 Expression of MetG in *Escherichia coli* and *in vitro* labeling with 2-amino-7-octynoic acid (AOA)

Based on the reported superior selectivity of azide- compared to alkyne-functionalized secondary CuAAC reagents (Speers & Cravatt 2004) and our own experience with fatty acid chemical reporters (Charron et al. 2009) we evaluated whether an alkynyl-isostere of ANL, 2-amino-7-octynoic acid (AOA, Figure 1a), could be accepted by previously

reported MetRS mutants (Link et al. 2006; Tanrikulu et al. 2009). MetRS mutants were generated from the *E. coli metG* gene (L13G, L13N Y260L H301L: NLL, L13P Y260L H301L: PLL) by site-directed mutagenesis and ligated into pCOLAduet-1 for over-expression in *E. coli* (only L13G and NLL were cloned in pCOLAduet-1, PLL was directly cloned in pWSK29, see below). AOA was synthesized by alkylation of diethyl acetamidomalonate with hex-5-ynyl-4-methylbenzenesulfonate followed by saponification, decarboxylation, and deacetylation to yield the racemic product (Figure 10).[2]

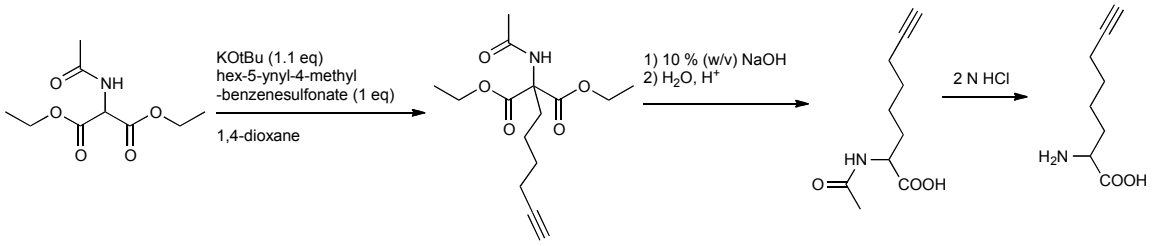


Figure 10 Synthesis of 2-amino-7-octynoic acid (AOA).

E. coli pCOLAduet_MetG$_{NLL}$ and *E. coli* pCOLAduet_MetGL13G overnight cultures were diluted 1:30 into M9 minimal and grown for 3 h. The cultures were pelleted, resuspended in an equal volume of M9 minimal medium with or without IPTG and with or without AOA or ANL, and grown for 3 h at 37 °C with aeration. Cells were lysed with SDS lysis buffer and reacted with azido-rhodamine (az-rho) or alkyne-rhodamine (alk-rho) by CuAAC and analyzed by SDS-PAGE and in-gel fluorescence scanning (Charron et al. 2009). Over-expression of MetG$_{NLL}$ or MetG$_{L13G}$ facilitates

[2] Due to the racemic nature of AOA, any indicated concentration refers to the effective concentration of the L-isomere, unless otherwise noted.

incorporation of AOA into the *E. coli* proteome (Figure 11) (further analysis of AOA incorporation was carried out in *S. typhimurium*, see below). As reported previously, $MetG_{NLL}$ provides better ANL incorporation than $MetG_{L13G}$, however this difference is not as pronounced with AOA, which is incorporated less efficiently in *E. coli* than ANL. These results were encouraging to test AOA incorporation in the bacterial pathogen *S. typhimurium*.

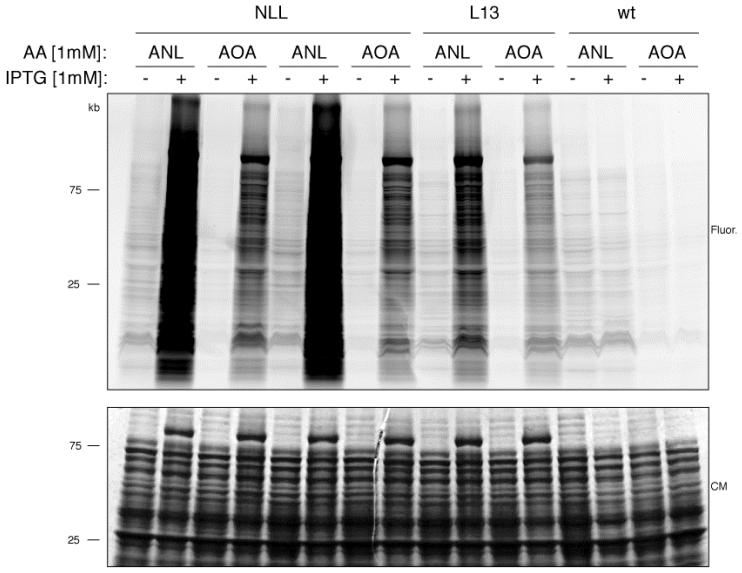


Figure 11 Expression of $MetG_{NLL}$ or $MetG_{L13}$ facilitates AOA incorporation in *E. coli*. *E. coli* pCOLAduet_MetGNLL and *E. coli* pCOLAduet_MetGL13 (induced with 1 mM IPTG) was labeled with 1 mM azidonorleucine (ANL) or 1 mM 2-amino-7-octynoic acid (AOA) for 3 h in M9 minimal medium. Cells were lysed, reacted with az-rho or alk-rho by CuAAC and analyzed by in-gel florescence screening. AA: amino acid, IPTG: Isopropyl β -D-1-thiogalactopyranoside, NLL: $MetG_{NLL}$, L13: $MetG_{L13G}$, Fluor.: fluorescence, CM: coomassie blue.

2.3.2 Expression of MetG in *Salmonella typhimurium* and *in vitro* labeling with 2-amino-7-octynoic acid (AOA)

Since the metG gene in *E. coli* and *S. typhimurium* are 95 % identical, overexpression of *E. coli* metG in *S. typhimurium* should facilitate amino acid activation. To allow mutant MetG overexpression in *S. typhimurium* $MetG_{L13G}$, $MetG_{NLL}$, and $MetG_{PLL}$ were cloned in the low copy number plasmid pWSK29 under the expression control of the lac promoter, which provides constitutive expression in *S. typhimurium*. All plasmids were transformed into the *S. typhimurium* strain IR715. For the *in vitro* analysis of the *S. typhimurium* strains expressing MetRS mutants, bacteria were grown in full Luria-Bertani (LB) medium to stationary phase and diluted into M9 minimal medium containing Met, AOA, or ANL. Protein lysates were reacted with the az-rho detection tag by CuAAC and analyzed by SDS-PAGE and in-gel fluorescence scanning.

To determine the most efficient MetG mutant and orthogonal amino acid reporter combination for selective labeling of the *S. typhimurium* proteome, we compared the three MetG mutants $MetG_{L13G}$, $MetG_{NLL}$, and $MetG_{PLL}$ for ANL and AOA incorporation in *S. typhimurium in vitro*. All three MetRS mutants permit incorporation of ANL and AOA into the *S. typhimurium* proteome (Figure 12a). Expression level of all MetG mutants were comparable as judged by S-tag Western blot analysis of the epitope tag fused to all MetRS constructs (Figure 12a). Both triple mutants (NLL, PLL) conveyed more efficient incorporation of ANL and AOA over the single mutant (L13G). In addition, AOA showed superior signal-to-noise ratios compared to labeling with ANL for all three MetRS mutants (Figure 12b). These data demonstrate that both orthogonal amino acid reporters function in *S. typhimurium* expressing MetG mutants, but $MetG_{NLL}$ or $MetG_{PLL}$

in combination with AOA affords the optimal orthogonal enzyme-substrate pair for metabolic labeling of *S. typhimurium* proteins.

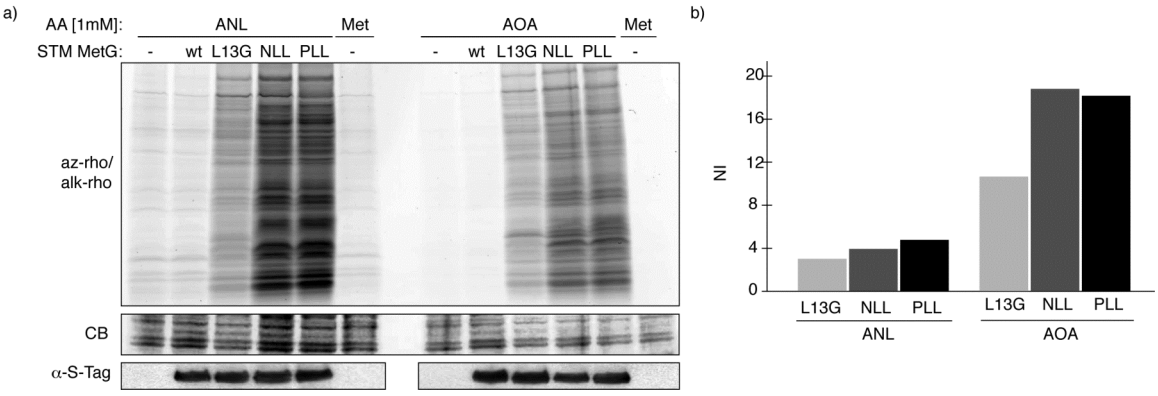


Figure 12 Triple mutants ($MetG_{NLL}$ and $MetG_{PLL}$) in combination with 2-amino-7-octynoic acid (AOA) provide optimal signal-to-noise in *S. typhimurium*. a) *S. typhimurium* pWSK29_MetG, pWSK29_$MetG_{L13G}$, pWSK29_$MetG_{NLL}$, or pWSK29_$MetG_{PLL}$ was labeled with ANL or AOA in minimal medium. b) Quantification of fluorescence signal. AA: amino acid, STM: *S. typhimurium* IR715, az-rho: azido-rhodamine, alk-rho: alkyne-rhodamine, CB: coomassie blue, α-S-Tag: anti-S-Tag Western blot, NI: normalized fluorescence intensity, normalized for negative control and coomassie blue staining.

To confirm selectivity of AOA for newly synthesized Met-containing proteins, we conducted protein synthesis inhibitor and Met competition experiments. Pre- and co-incubation of *S. typhimurium* pWSK29_$MetG_{NLL}$ with the protein synthesis inhibitor tetracycline (Tet) effectively abrogated AOA-labeling (Figure 13a). Met efficiently competed away AOA-incorporation in a dose-dependent manner (Figure 13b). The synthetic route of AOA results in a racemic mixture. Some D-amino acids can be translationally incorporated into proteins and D-amino acids can be toxic to various organisms (Champney & Jensen 1970; Friedman 1999). Therefore we wanted to know

whether D-AOA interferes with AOA labeling of *S. typhimurium*. Pig acylase I was used to kinetically resolve the racemic mixture of AOA (Chenault et al. 1989), which afforded enantiomerically enriched (ee) L-AOA (Figure 14). The purity of ee L-AOA was not further quantified, but the NMR analysis revealed that L-AOA was the predominant species present. Racemic AOA was used at 2 mM and compared to 1 mM ee L-AOA. Both preparations showed identical labeling efficiencies at comparable effective concentrations (Figure 13c) and we could not detect any growth differences of *S. typhimurium* between the racemic and the enantiomerically pure AOA preparation. These data demonstrate that AOA can selectively label newly synthesized Met-containing proteins in gram-negative bacterial pathogens such as *S. typhimurium* upon expression of $MetG_{NLL}$.

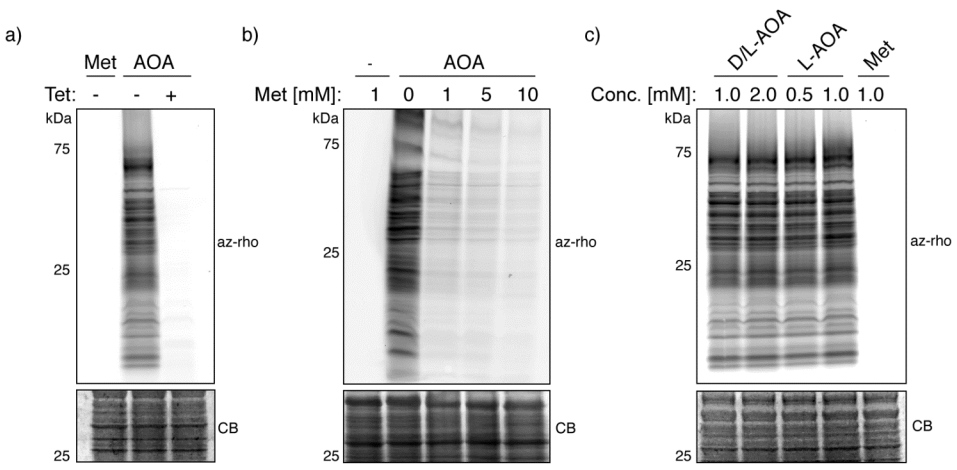


Figure 13 2-amino-7-octynoic acid (AOA) serves as a co-translational chemical reporter for methionine (Met) in *S. typhimurium*. a) *S. typhimurium* pWSK29_MetG_{NLL} was treated with Met (1 mM) or AOA (1 mM), in the presence or absence of tetracycline (Tet) for 30 min. b) *S. typhimurium* pWSK29_MetG_{NLL} was incubated with increasing concentrations of Met and AOA (1 mM). c) Comparison of the labeling efficiency of D/L-AOA and L-AOA. CB: Coomassie blue, az-rho: azido-rhodamine.

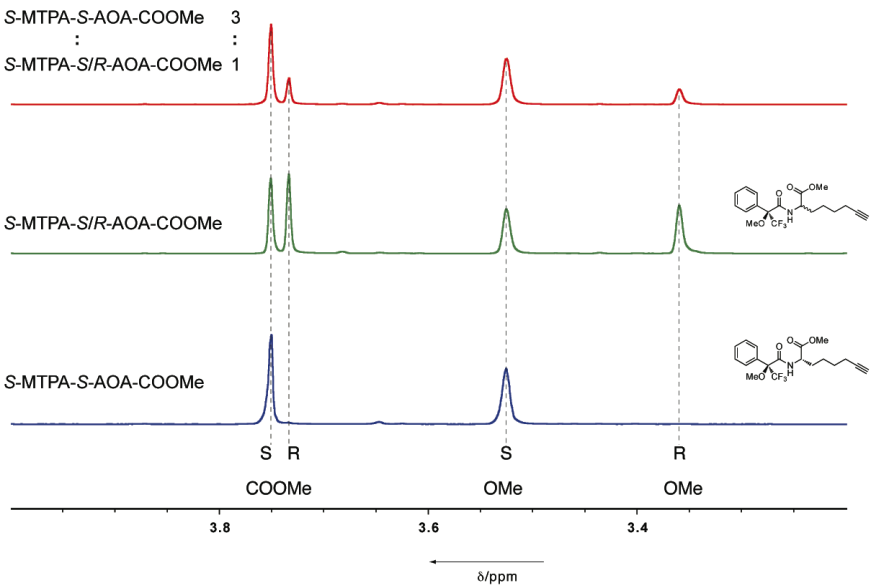


Figure 14 Determination of the absolute stereochemistry of kinetically resolved 2-amino-7-octynoic acid (AOA). The methyl ester of AOA (racemic mixture) and kinetically resolved AOA were formed using thionyl chloride in anhydrous methanol. Both methyl esters were reacted with R-α-methoxy-α-(trifluoromethyl)phenylacetyl chloride (MTPA chloride) to form S-MTPA-S/R-1-COOMe and S-MTPA-S-1-COOMe (as determined subsequently by 1H NMR), respectively. ^{1}H NMR spectra were taken of both MTPA derivatives as well as a 1:3 mix of the two. The singlet signal of the methyl ester (COOMe) protons experiences a characteristic chemical shift depending on the absolute stereochemistry of the chiral center of the amino acid moiety. The signal of the S-S-diastereomer experiences a greater chemical shift than the S-R-diastereomer.

2.3.3 Labeling *Salmonella typhimurium* with 2-amino-7-octynoic acid (AOA) during infection of mammalian cells

We then evaluated the efficiency of ANL and AOA-labeling of *S. typhimurium* proteomes during intracellular infection of mammalian cells. To analyze selective incorporation of orthogonal amino acid reporters in intracellular *S. typhimurium*, RAW264.7 murine macrophages were infected with *S. typhimurium* pWSK29_$MetG_{NLL}$ at a multiplicity of infection (MOI) of 100 for 30 minutes. Following infection, extra-cellular bacteria were killed by addition of gentamicin, a cell-impermeable antibiotic. After two hours, infected and non-infected macrophages were labeled with different concentrations of Met, ANL or AOA in the growth medium for 3 hours. Cell pellets containing bacteria and mammalian cells were lysed with 4% SDS lysis buffer, reacted with alk-rho or az-rho and analyzed by in-gel fluorescence. Samples infected with *S. typhimurium* pWSK29_$MetG_{NLL}$, validated by S-tag Western blot analysis, and labeled with ANL or AOA displayed a strong fluorescence signal above background in contrast to the non-infected samples labeled with ANL or AOA (Figure 15a). Comparison of the Met control samples revealed the considerably higher background signal of alk-rho, as observed *in vitro*. Close inspection of AOA-treated non-infected macrophages showed marginal concentration dependent labeling with AOA, while ANL incorporation could not be observed. Quantification of the relative fluorescence intensities of ANL and AOA suggested that AOA, at an effective concentration of 1 mM, affords the most sensitive labeling of *S. typhimurium* proteins during infection (Figure 15b). These results strongly suggested that ANL and AOA can be selectively incorporated into *S. typhimurium* during infection of mammalian cells by the cell selective expression of $MetG_{NLL}$. We found that

AOA in combination with azide-detection reagents displayed superior labeling sensitivity at all tested concentrations (Figure 15b).

To confirm selective labeling of *Salmonella* with AOA inside mammalian cells, we performed fluorescence imaging studies of infected cells. HeLa cells were infected with *S. typhimurium* pWSK29_$MetG_{NLL}$ (MOI = 100), 16 h post-infection cells were pulse-labeled with AOA or Met for 1 hour, fixed, permeabilized, reacted with az-rho and stained with anti-*Salmonella* serum. Imaging of rhodamine fluorescence demonstrated robust and selective labeling of intracellular bacteria with AOA (Figure 16a). No rhodamine fluorescence was observed in host cells or *S. typhimurium* pWSK29_$MetG_{NLL}$ -infected cells treated with Met under these conditions. AOA-labeled *S. typhimurium* co-localized with anti-*Salmonella* serum, validating selective incorporation of AOA into the bacterial proteome and exclusion from host proteins. In addition, we analyzed *S. typhimurium*-infected RAW264.7 macrophages by LAMP-1 antibody staining, a marker for the *Salmonella* containing vacuole (SCV) of infected host cells (Haraga et al. 2008). AOA-labeled *S. typhimurium* pWSK29_$MetG_{NLL}$ inside RAW264.7 macrophages were enclosed in LAMP-1 positive compartments (Figure 16b), suggesting that AOA labeling does not significantly disturb intracellular trafficking of *Salmonella* in host cells. These data further validate the utility of AOA for selective metabolic labeling of *Salmonella* proteins inside mammalian cells.

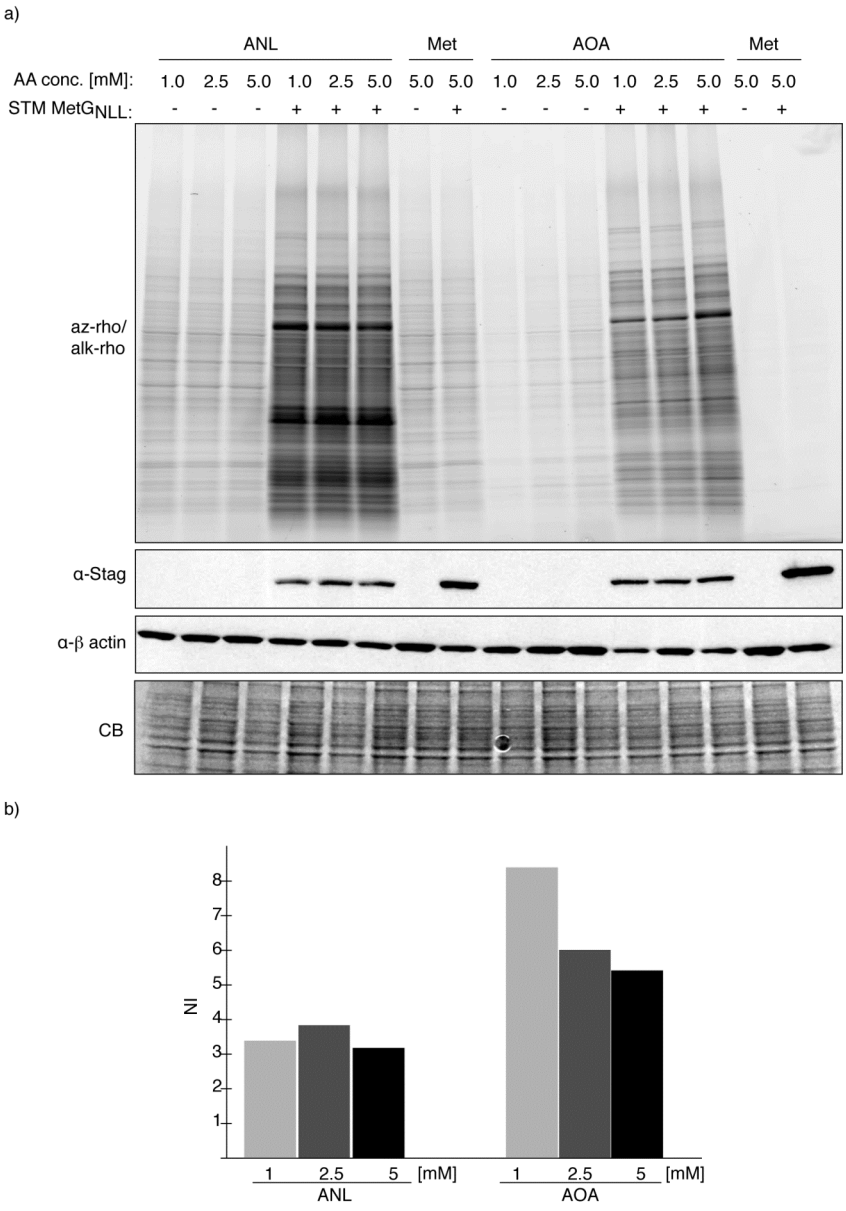


Figure 15 AOA and ANL allow labeling of *S. typhimurium* pWSK29_MetG$_{NLL}$ infected macrophages. a) RAW264.7 cells were infected with *S. typhimurium* pWSK29_MetG$_{NLL}$ and pulse-labeled with different concentrations of AOA or ANL 16 hours post-infection. b) Quantification of fluorescence signal. AA: amino acid, STM: *S. typhimurium* IR715, az-rho: azido-rhodamine, alk-rho: alkyne-rhodamine, CB: coomassie blue, α-S-Tag: anti-S-Tag Western blot, NI: normalized fluorescence intensity, normalized for negative control and coomassie blue staining.

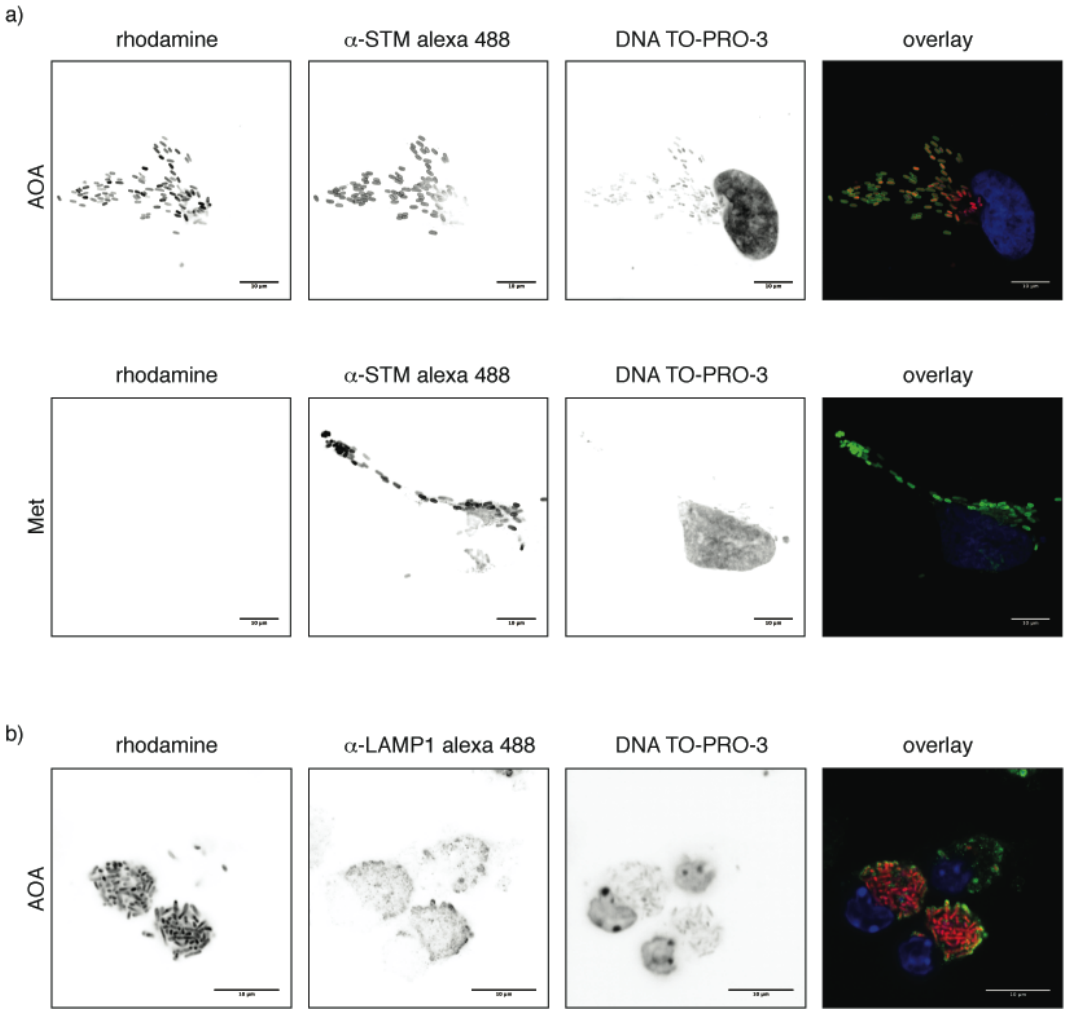


Figure 16 2-amino-7-octynoic acid (AOA) allows selective labeling of *S. typhimurium* in mammalian cells during infection.

2.3.4 Proteomic analysis of 2-amino-7-octynoic acid (AOA) labeled proteins of *Salmonella typhimurium*

After selectivity and efficiency of AOA labeling was established, we compared *S. tyhimurium* proteins that are metabolically labeled by AOA during growth in liquid culture to *S. tyhimurium* proteins that are labeled during infection of macrophages. For this comparative analysis, *S. typhimurium* pWSK29_$MetG_{NLL}$ grown in liquid culture (*in vitro*) (Figure 18b) and RAW264.7 macrophages infected with *S. typhimurium* pWSK29_$MetG_{NLL}$ (16 h post infection) (Figure 18c) were pulse-labeled with AOA or Met for 1 hour. Generated cell lysates were reacted with azido-azo-biotin (Y.-Y. Yang, Ascano & Hang 2010), affinity purified with streptavidin and eluted from beads for protein identification by gel-based proteomics (Figure 18a). The generated raw data was searched against a concatenated mouse-*S. typhimurium* database for protein identification. Only proteins that contained two unique peptides and were not detected in the Met control were considered for this study. The analysis of infected RAW264.7 macrophages identified a total of 218 proteins that sufficed our filter criteria (Appendix Table 3). Amongst these proteins, 185 were assigned to the *S. typhimurium* proteome (85%), the remainder were mouse proteins (Figure 18d). Additionally, we identified 472 proteins from *S. typhimurium* pWSK29_$MetG_{NLL}$ pulse-labeled with AOA in minimal medium (Appendix Table 4). Between the *in vitro Salmonella* and the intracellular *Salmonella* datasets we could identify 96 common proteins, while 89 proteins (48%) were exclusively observed in the sample derived from infected macrophages (Figure 18e). Interestingly, at least 5 *Salmonella* proteins (SodM, SsrB, SseA, PipB2 and PhoP) from infected RAW264.7 cells were previously described as virulence factors (J. Yang et al. 2008).

Compared to other proteomic datasets generated from macrophages infected with *S. typhimurium* (Shi et al. 2006), our results demonstrate that AOA-pulse-labeling in conjunction with bioorthogonal ligation methods allows greater enrichment of bacterial proteins (85% *Salmonella* proteins with AOA-labeling compared to ~45% under comparable conditions without AOA) and reveals additional bacterial proteins that may be preferentially expressed during infection (Figure 18f). Although the use of AOA and $MetG_{NLL}$ only targets Met-containing proteins, 96% of the *S. typhimurium* LT2 proteome contains one or more Met residues in addition to the N-terminal initiator Met residue. Hence, the majority of the *Salmonella* proteome is amenable to labeling with AOA (Figure 17). Our preliminary proteomic studies demonstrate that AOA-pulse-labeling in combination with CuAAC allows efficient enrichment and identification of endogenously expressed bacterial proteins that are differentially synthesized by *Salmonella* during infection of host cells.

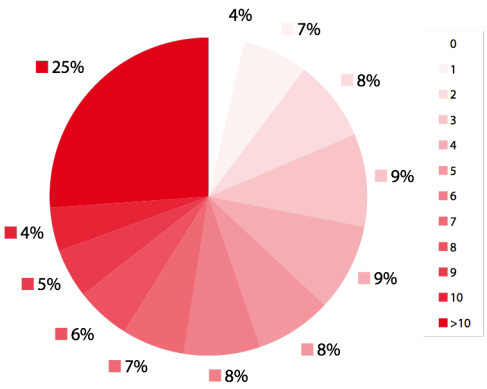
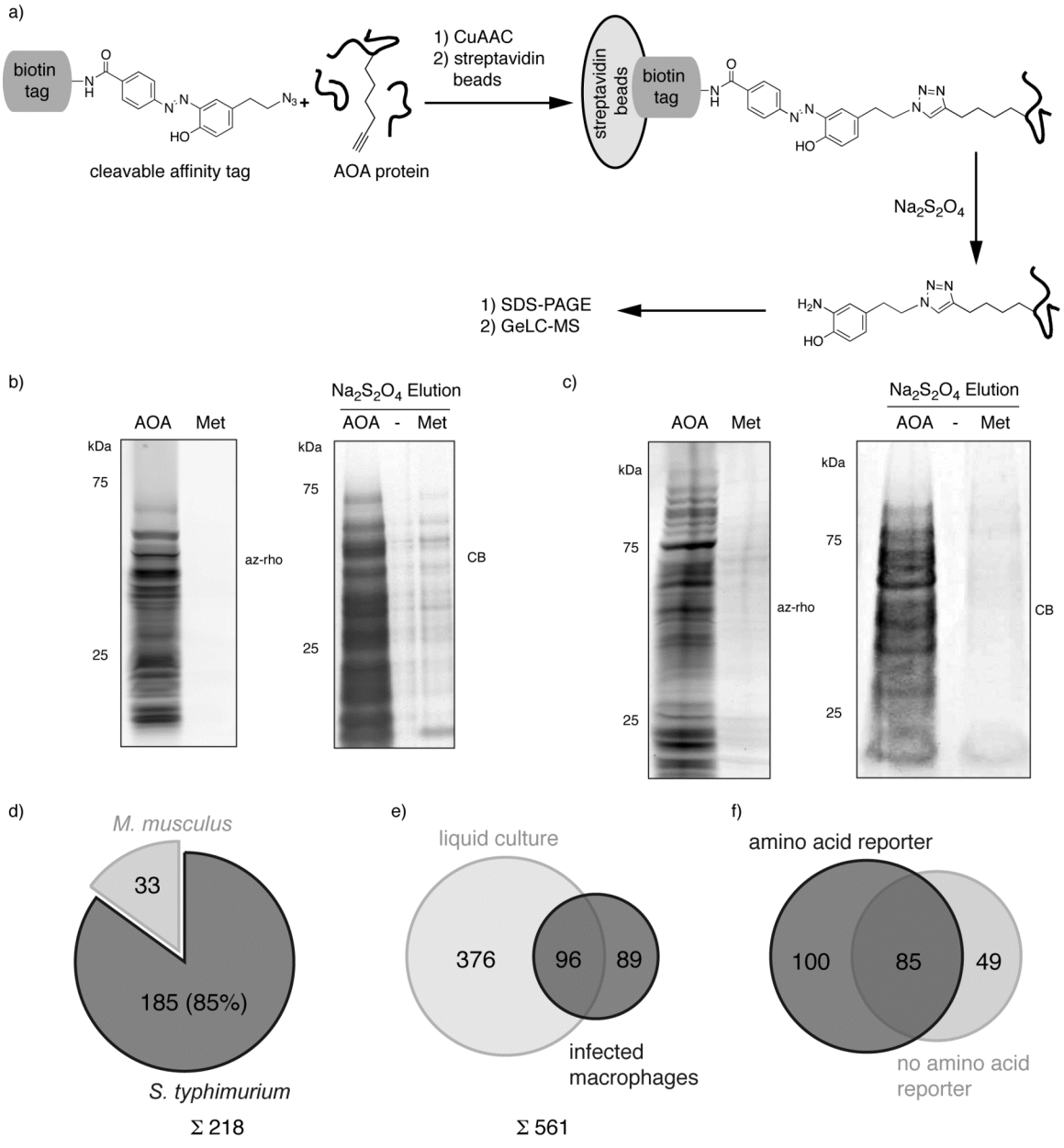


Figure 17 Methionine distribution in *S. typhimurium*. Chart shows *S. typhimurium* proteome fractions binned by number of methionines per protein (exclusive the N-terminal methionine). Data based on genome of *S. typhimurium* LT2 SGSC1412 (Washington University).

Figure 18 2-amino-7-octynoic acid (AOA) allows enrichment and identification of *S. typhimurium* proteins from infected mammalian cells. a) Cell lysates were reacted by CuAAC with azido-azo-biotin (cleavable affinity tag), affinity purified on streptavidin beads, and selectively eluted with $Na_2S_2O_4$ by reduction of the azo-bond. The eluate was separated by SDS-PAGE and subsequently analyzed by mass spectrometry based proteomics. b) *S. typhimurium* pWSK29_$MetG_{NLL}$ was labeled with 1 mM AOA or 1 mM Met in MgM minimal medium. Cell lysates were analyzed by CuAAC mediated in-gel fluorescence scanning. c) RAW264.7 cells were infected with *S. typhimurium* pWSK29_$MetG_{NLL}$ and labeled with 1 mM AOA or 1 mM Met 16 hours post-infection. Cell lysates were analyzed by CuAAC mediated in-gel fluorescence scanning. Cell lysates were also reacted by CuAAC with azido-azo-biotin, affinity purified and selectively eluted by $Na_2S_2O_4$. Eluate was separated by SDS-PAGE and subsequently analyzed by mass spectrometry based proteomics. d) A total of 218 *S. typhimurium* proteins were identified from infected RAW264.7 cells. Of these proteins, 185 (85%) were *Salmonella* proteins and 33 were mouse proteins. e) Comparison of *Salmonell*a proteins identified from infected macrophages and Salmonella grown in minimal liquid culture: 376 proteins were identified in liquid culture only, 89 proteins were identified in infected macrophages only, and 96 were identified in both samples. f) Comparison of proteins identified from AOA-pulse-labeled infected RAW264.7 cells (amino acid reporter) and a previously published dataset (Shi et al. 2006) of *Salmonella* proteins identified from infected RAW246.7 cells under comparable conditions without AOA labeling (no amino acid reporter): 100 proteins were exclusively identified in our dataset, 85 proteins were identified in both datasets, and 49 proteins were only identified in the previously published dataset. az-rho: azido-rhodamine, CB: coomassie blue.

Figure 18


a)
biotin tag
cleavable affinity tag
AOA protein
1) CuAAC
2) streptavidin beads
streptavidin beads
biotin tag
$Na_2S_2O_4$
1) SDS-PAGE
2) GeLC-MS
b)
AOA Met
$Na_2S_2O_4$ Elution
AOA - Met
kDa
75
25
az-rho
CB
c)
AOA Met
$Na_2S_2O_4$ Elution
AOA - Met
kDa
75
25
az-rho
CB
d)
M. musculus
33
185 (85%)
S. typhimurium
Σ 218
e)
liquid culture
376
96
89
infected macrophages
Σ 561
f)
amino acid reporter
100
85
49
no amino acid reporter

2.4 Materials and Methods

2.4.1 General Methods and Materials

Unless otherwise noted all chemical reagents were purchased from Sigma-Aldrich or Fisher Scientific. ^{1}H and ^{13}C NMR spectra were acquired on a Bruker DPX-400 or Bruker UltraShield-600 Plus instrument. Chemical shifts are reporter in ppm, and J values are reported in Hz. MALDI-TOF mass spectra were acquired on a Applied Biosystems Voyager-DE mass spectrometer or a LC-MS. In-gel fluorescence scanning was performed with a Amersham Biosciences Typhoon 9400 variable mode imager (excitation 532 nm, 580 nm filter, 30 nm band-pass). All contrast or brightness adjustments were applied to the entire gel or blot, including the different controls. Images adjustments were done in Adobe Photoshop CS.

2.4.2 Synthesis of 2-amino-7-octynoic acid (AOA)

2.4.2.1 Diethyl 2-acetamido-2-(hex-5-yn-1-yl)malonate

Diethyl acetamidomalonate (1.05 g, 4.75 mmol) was dissolved in 30 ml anhydrous 1,4-dioxane in a previously flame-dried round bottom flask equipped with a condenser and set under Argon atmosphere. Potassium *tert*-butoxide (0.59 g, 5.23 mmol) was added to the vigorously stirred solution and it was incubated for 2 hours at 60 °C. Hex-5-ynyl-4-methylbenzenesulfonate was added drop wise to the stirring solution and the reaction was refluxed for 24 h. The suspension was cooled to room temperature and solid material was filtered and washed with diethyl ether. Filtrate was evaporated under reduced pressure, resulting in a brown oil. This material was purified by column chromatography on silica gel (30% ethyl acetate/70% hexanes to 40% ethyl acetate/60% hexanes) to afford 520 mg pure product **1** (37%). R_f = 0.4; ^{1}H NMR (600 MHz, $CDCl_3$): δ = 1.15-1.20 (m, 8H),

1.44-1.49 (quint., J=7.2 Hz), 1.85 (s, 1H), 1.97 (s, 3H), 2.10-2.13 (m, 2H), 2.25-2.28 (m, 2H), 4.16-4.20 (q, J=7.2 Hz, 4H), 6.75 (s, 1H); ^{13}C NMR (600 MHz, $CDCl_3$): δ = 13.88, 17.95, 22.41, 22.86, 27.80, 31.43, 62.34, 66.36, 68,50, 83.77, 167.97, 168.92; MS (ESI): calc. for $C_{15}H_{23}NO_5$ $[M-H]^+$ requires 298.16, found 298.05.

2.4.2.2 2-acetamidooct-7-ynoic acid

Diethyl 2-acetamido-2-(hex-5-yn-1-yl)malonate (520 mg, 1.75 mmol) was dissolved in 10 ml 10% NaOH and refluxed for 5 hours. The solution was cooled to room temperature and acidified to pH 2 with 6 N HCl. Solvent was evaporated under reduced pressure and the resulting residue was taken up in 10 ml of H_2O and refluxed for 5 hours. The reaction mix was cooled to ambient temperature, acidified to pH 2 with 1 N HCl and extracted with three portions of ethyl acetate. The organic layers were combined, dried over $MgSO_4$, filtered, and evaporated under reduced pressure, yielding 240 mg of pure product **2** (70%). ^{1}H NMR (600 MHz, CD_3OD): δ = 1.46-1.53 (m, 4H), 1.64-1.70 (m, 1H), 1.80-1.86 (m, 1H), 1.96 (s, 3H), 2.15-2.17 (m, 3H), 4.33-4.35 (dd, *J*=5.4, 3.6 Hz, 1H); ^{13}C NMR (600 MHz, CD_3OD): δ = 17.50, 21.02, 24.61, 27.80, 30.76, 52.20, 68.38, 83.36, 172.00, 174.13; MS (ESI): calc. for $C_{10}H_{15}NO_3$ $[M-H]^+$ requires 198.11, found 198.03.

2.4.2.3 2-amino-7-octynoic acid (AOA)

2-acetamidooct-7-ynoic acid (240 mg, 1.22 mmol) was dissolved in 10 ml 2 N HCl and refluxed for 2 hours. The solution was cooled to ambient temperature and loaded on an equilibrated DOWEX 50WX4-100 cation-exchange resin. The resin was extensively washed with water and eluted with 1 M NH_4OH. The eluate was evaporated under reduced pressure to afford 138 mg of **1** (73%). ^{1}H NMR (600 MHz, D_2O): δ = 1.40-1.49 (m, 4H), 1.82-1.93 (m, 2H), 2.11-2.16 (m, 2H), 2.27-2.28 (t, *J*=2.4 Hz, 1H), 3.98-4.00 (t,

J=6 Hz, 1H); ^{13}C NMR (600 MHz, D_2O): 17.21, 23.49, 27.22, 29.97, 54.73, 69.34, 85.75, 175.00; MS (ESI): calc. for $C_8H_{13}NO_2$ $[M-H]^+$ requires 156.10, found 155.98.

2.4.2.4 Enantioselective hydrolysis of 2-acetamidooct-7-ynoic acid with acylase I

2-acetamidooct-7-ynoic acid (750 mg) was dissolved in 70 ml H_2O and pH was adjusted to pH 7.5 with 1 M LiOH. 6 mg acylase I (acylase I from porcine kidney) was added and solution was slowly stirred at 37 °C for 24 hours. The material was lyophilized and 50% mass was resuspended in 10 ml H_2O. Activated charcoal (45 mg) was added and incubated for 5 min. Charcoal was filtered over Buchner funnel. Aqueous filtrate was washed twice with diethyl ether. Aqueous phase was evaporated under reduced pressure at 45 °C. Resulting residue was dissolved in 1 M HCl and purified over DOWEX 50WX4-100 cation-exchange resin. Eluate was evaporated under reduced pressure at 45 °C. ^{1}H NMR (600 MHz, D_2O) data was identical to AOA.

2.4.2.5 Determination of the absolute stereochemistry of kinetically resolved AOA

The absolute stereochemistry of enantioselectively hydrolyzed 2-acetamidooct-7-ynoic acid was determined by formation of the MTPA-amide (*s*-α-methoxy-α-(trifluoromethyl)phenylacetyl-amide, Mosher-amide) of the methylester of enantioselectively hydrolyzed 2-acetamidooct-7-ynoic acid (Dale & Mosher 1973). The methylester of enantioselectively hydrolyzed 2-acetamidooct-7-ynoic acid was formed as follows. Enantioselectively hydrolyzed 2-acetamidooct-7-ynoic acid was dissolved in anhydrous methanol under Argon atmosphere. Solution was cooled to 0 °C and 50 equivalents of thionyl chloride were added. Reaction mix was allowed to warm up to room temperature over night. Reaction was quenched with H_2O and solvent was evaporated under reduced pressure. Crude material was purified by column

chromatography on silica gel (10% methanol/89% dichloromethane/1% acetic acid). The methylester of AOA was formed accordingly, however, the reaction was refluxed over night with 30 equivalents of thionyl chloride.

The *S*-MTPA-amides of both species were formed as follows. The methyl ester was dissolved in anhydrous pyridine under argon atmosphere. Three equivalents of *R*-MTPA chloride was added and the reaction was stirred at room temperature for 90 min. Reaction completion was assessed by absence of ninhydrin signal. The reaction was quenched with H_2O and the aqueous phase was extracted with diethyl ether. The organic phase was washed twice with 10% HCl and twice with sat. $NaHCO_3$. The organic phase was dried over $MgSO_4$, filtered, and evaporated under reduced pressure. The material was directly analyzed by ^1H NMR.

2.4.3 Synthesis of *S*-2-amino-6-azidohexanoic acid (azidonorleucine, ANL)

2.4.3.1 *S*-6-azido-2-((tert-butoxycarbonyl)amino)hexanoic acid

S-6-azido-2-((tert-butoxycarbonyl)amino)hexanoic acid was basically synthesized as previously reported with modifications (Link et al. 2004). Instead of trifluoro-methanesulfonyl azide we used imidazol-1-sulfonyl azide hydrochloride as the diazotransfer reagent (Goddard-Borger & Stick 2007). Imidazol-1-sulfonyl azide hydrochloride (3.06 g, 14.61 mmol) was added to a stirred solution of BOC-lysine (3 g, 12.18 mmol), K_2CO_3 (4.55 g, 32.89 mmol), and $CuSO_4 \cdot 5H_2O$ (30.4 mg, 0.12 mmol) in 60 ml of methanol. The reaction mix was stirred at ambient temperature over night and diluted with phosphate buffer (0.25 M KH_2PO_4, 0.25 M K_2HPO_4). The solution was carefully acidified to pH 3 and extracted with 3 portions ethyl acetate. The organic layers were combined, dried over $MgSO_4$, filtered, and evaporated under reduced pressure. The

resulting crude material was purified by column chromatography on silica gel (25% ethyl acetate/74% hexanes/1% acetic acid) and afforded 1.95 g product. ^{1}H NMR (600 MHz, $CDCl_3$): δ = 1.43-1.48 (m, 11H), 1.59-1.65 (m, 2H), 1.67-1.71 (m, 1H), 1.86-1.89 (m, 1H), 3.26-3.28 (t, *J*=6.6 Hz, 2H), 4.10-4.11 (m, 0.4H), 4.32-4.33 (m, 0.6H), 5.20-5.21 (d, 0.6H), 6.69-6.70 (d, 0.40).

2.4.3.2 *S*-2-amino-6-azidohexanoic acid (ANL)

ANL was prepared from *S*-6-azido-2-((tert-butoxycarbonyl)amino)hexanoic acid as previously reported (Link et al. 2004). ^{1}H NMR (600 MHz, D_2O): δ = 1.36-1.50 (m, 2H), 1.55-1.60 (m, 2H), 1.83-1.96 (m, 2H), 3.27-3.29 (t, *J*=6.6Hz, 2H), 3.98-4.00 (t, *J*=6Hz, 1H).

2.4.4 Plasmids and bacterial strains

S. enterica typhimurium IR715 was transformed with different pWSK29 derived plasmids, expressing different mutants of the *E. coli* methionyl-tRNA synthetase gene (MetRS, *metG*). *MetG* was amplified from *E. coli* K12 (GAAGATCTTCATG ACTCAAGTCGCGAAGAA, CTCGAGTTTCACCTGATGACCCGG) and cloned into pCOLAduet-1 (Novagen) with BglII and XhoI to create a C-terminal S-tagged fusion of MetRS and the plasmid pMetG1. Three mutant versions of MetRS were created using the QuikChange XL Site-Directed Mutagenesis Kit or the QuikChange Multi Site-Directed Mutagenesis Kit (Stratagen): L13G (MetRS-L13G), L13N-Y260L-H301L (MetRS-NLL), and L13P-Y260L-H301L (MetRS-PLL). MetRS was amplified from pMetG1 (CCTGAGCTCCATGACTCAAGTCGCGAAG, CCGGAATTCATTACGAGT CCATGTGCTGGCGTTC) and cloned into pWSK29 using SacI and EcoRI, creating

pMetG2. pMetG2 encodes a C-terminal S-tag fusion of MetRS under the expression control of the lac promoter.

2.4.5 Infection of mammalian cells

HeLa cells (5 x 10^4 cells/well) or RAW264.7 macrophages (5 x 10^5 cells/well) were split into a 24 well plate 10 hours before infection. *S. typhimurium* was grown over night in full LB with 100 μg/ml Carbenicillin (Cb). Overnight cultures were diluted 1:100 into fresh LB (Cb) and grown to stationary phase (OD_{600} = 2). Bacteria were washed once in PBS and incubated for 30 min in DMEM (10% FBS) on ice. Bacteria were added to mammalian cells in DMEM (10 % FBS) at MOI = 100 and centrifuged (400 g, 5 min, 4 °C) for spin infection. Subsequently, cells were incubated for 25 min at 37 °C, 5% CO_2. Cells were washed 4x with PBS and DMEM (10% FBS) with 100 μg/ml gentamicin was added for 90 min to kill all extracellular bacteria. After 90 min medium was exchanged to DMEM (10% FBS) with 10 μg/ml gentamicin for the remainder of the experiment.

2.4.6 Metabolic labeling (in vitro, during infection)

Over night cultures of bacteria were grown in full LB with 100 μg/ml Carbenicillin (Cb). Over night cultures were diluted 1:100 into fresh LB (Cb) and grown to stationary phase. Bacteria were collected by centrifugation (10,000 rpm, 3 min) and resuspended in MgM (100 mM MES pH 5, 1 mM KH_2PO_4, 10 μM $MgCl_2$, 5 mM KCl, 7.5 mM NH_4OH, 0.5 mM K_2SO_4, 38 mM glycerol) minimal medium or M9 (1x M9 salts, 20 mM glucose, 2 mM MgSO4, 100 μM $CaCl_2$) minimal medium (Cb). Bacteria were grown in minimal medium for 15 min and collected by centrifugation. Bacterial pellets corresponding to 500 μl culture volume were resuspended in MgM or M9 minimal medium (Cb) containing chemical reporter. Bacteria were labeled for 1 hour and collected by

centrifugation, washed once in PBS, and stored at -80 °C. For tetracycline (Tet) protein synthesis inhibition, bacteria were grown for 30 min in LB (40 μg/ml Tet) before being labeled for 30 min in M9 minimal medium (40 μg/ml Tet).

For metabolic labeling of infected mammalian cells, cells were labeled with amino acid reporter in DMEM (methionine free) for 3 hours, 90 min after infection or labeled for 1 hour (fluorescence microscopy), 16 hours after infection. Subsequently, medium was aspirated and cells were washed multiple times with PBS.

For proteomic analysis infected macrophages were pulse-labeled with 1 mM AOA or 1 mM Met 16 h after infection for 1 h pulse time. For proteomic analysis of *S. typhimurium* labeled in liquid culture, bacteria were grown as described above in MgM minimal medium and pulse labeled for 1 h with 2 mM AOA or 1 mM Met.

2.4.7 Cell lysis

Bacterial cell pellets and cell pellets of infected mammalian cells were lysed with CuAAC compatible SDS lysis buffer as previously reported with slight modifications (Rangan et al. 2010). For this study cells were lysed without lysozyme present in the lysis buffer and lysed cells were centrifuged at 20,000 g for 15 min at 18 °C to remove cell debris

2.4.8 CuAAC for fluorescence in-gel profiling

50 μg of cell lysate was brought up to 15 μl with CuAAC buffer (4% SDS, 150 mM NaCl, 50 mM TEA). CuAAC components were added as a master mix to adjust the total volume to 50 μl with final concentrations as reported previously. CuAAC was carried out at room temperature for 1 hour. Protein was precipitated by adding >10 volumes of methanol and incubation at -20 °C for at least 1 hour. Precipitated protein was pelleted by

centrifugation at 20,000 g for 15 min at 4 °C. Protein pellets were washed once with cold acetone and dried on air. Protein pellets were resuspended in 1x Laemmli buffer, denatured at 95 °C for 5 min and separated by SDS-PAGE (4-20% Tris-HCl, Bio-Rad Criterion).

2.4.9 Fluorescence microscopy

For fluorescence imaging experiments infections were carried out as described above, however cells were grown on cover slips. 16 h after infection cells were labeled with 2 mM AOA, or Met for 1 hour in DMEM (methionine free). After metabolic labeling cells were fixed for 10 min with 3.7% PFA in PBS. Cells were washed 3x with PBS and permeabilized with 0.1% Triton-X-100, 0.2% saponin in PBS for 10 min. Cells were washed 3x with PBS and blocked with 10% FBS in PBS for 1 hour at room temperature. Cells were washed 3x with PBS and subsequently tagged by CuAAC. 300 μl of a CuAAC master mix (20 μM azRho, 1 mM TCEP (tris(2-carboxyethyl)phosphine hydrochloride), 100 μM TBTA (tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine), 1 mM $CuSO_4$) was added to each well and CuAAC was carried out for 1 h at room temperature. Subsequently, cells were washed with 1% Tween20, 0.5mM EDTA in PBS and PBS. Cells were incubated with primary antibody (LAMP-1, 1:1000. Abcam 24170; anti-*Salmonella* serum 1:250) in 0.2% saponin in PBS for 30 min. Cells were washed 3x with 0.2% saponin in PBS. Then cells were incubated with secondary antibody (Alexa Fluor® 488 goat anti-rabbit IgG (H+L), 1:1000 , Invitrogen) in 0.2% saponin in PBS for 30 min. Cells were washed 3x with 0.2% saponin in PBS and finally incubated with TO-PRO-3 (1:1000) in 0.2% saponin in PBS, before being washed 3x with 0.2% saponin in PBS. Cover slips were mounted on ProLong Gold (Invitrogen).

2.4.10 Quantification of in-gel fluorescence scanning

In-gel fluorescence profiles were quantified in ImageJ 1.40g by measuring the mean intensity of an equally sized square of each gel lane. The observed intensity was normalized by protein load as measured by coomassie stain intensity and the F.I. (normalized fluorescence intensity) was formulated as the quotient of this normalized intensity and the normalized intensity of the empty vector control or the uninfected sample with equal amino acid reporter concentration.

2.4.11 CuAAC for capture and retrieval

Cell lysates were brought up to a final concentration of 1 mg/ml with 50 mM TEA, 4% SDS and 150 mM NaCl. CuAAC components were added to 100 μM azido-azo-biotin, 1 mM TCEP, 100 μM TBTA, 1 mM $CuSO_4$ final concentrations and incubated for 1.5 hours at room temperature. Subsequently, protein was precipitated by methanol-chloroform precipitation. Before streptavidin bead capture, protein pellets were initially resuspended in 4% SDS in PBS and finally diluted to 0.2% SDS, 10 mM EDTA in PBS.

2.4.12 Streptavidin bead capture

Streptavidin beads (100 μl/5 mg protein; high capacity streptavidin agarose resin - Thermo Scientific) were washed 3x with PBS and equilibrated with 0.2% SDS, 10 mM EDTA in PBS. Dissolved protein pellets were added to beads in 0.2% SDS, 10 mM EDTA in PBS and incubated on turning wheel for 1 h at room temperature. Beads were washed with 2x 10 ml 2% SDS, 10 mM EDTA in PBS, 2x 10 ml of 6 M urea in 250 mM ABC (ammonium bicarbonate), 2x 10 ml 1 M NaCl in PBS, and 2x 50 mM ABC. Bound proteins were reduced with 10 mM TCEP in 50 mM ABC for 30 min at room temperature and subsequently alkylated with 20 mM iodacetamide in 50 mM ABC for 30

min at room temperature. Proteins were specifically eluted with 400 μl 50 mM $Na_2S_2O_4$, 0.5% SDS in 50 mM ABC for 1 h in shaker at room temperature. Flow through was collected and remaining proteins were eluted a second time for 30 min under equal conditions. Combined eluates were concentrated in Microcon YM-10 centrifugal filters (Millipore) as per manufactures protocol. Microcon filters were rinsed twice with 50 μl 1% SDS/75 mM β-mercaptoethanol (BME) to collect proteins. Samples were dried under reduced pressure. Samples were resuspendend in 25 μl 1x LDS sample buffer (5% BME), heated for 5 min at 95 °C and separated by SDS-PAGE. Gels were stained with GelCode Blue Stain Reagent (Thermo).

2.4.13 In-gel digestion

Each gel lane was cut into 15 pieces. Each piece was cut into cubes of approximately 1 mm^3, washed once in 50 mM ABC, twice in 1:1 50 mM ABC to acetonitrile (ACN), and once in ACN. Gel pieces were dried under reduced pressure. 1/50 (w/w) of trypsin in 50 mM ABC was added to cover all gel pieces with buffer and incubated over night at 37 °C. Supernatant was collected and peptides were extracted twice with 1:1 50 mM ABC to ACN, and twice with 100% ACN. All supernatants were pooled and dried under reduced pressure. For LC-MS/MS analysis samples were dissolved in 25 μl 0.1% TFA.

2.4.14 MS analysis

LC-MS analysis was performed with a Dionex 3000 nano-HPLC coupled to an LTQ-Orbitrap ion trap mass spectrometer (ThermoFisher). Peptides were pressure loaded onto a home-made 75 μm diameter, 15 cm C18 reverse phase column and separated with a gradient running from 95% buffer A (HPLC water with 0.1% formic acid) and 5% buffer B (HPLC grade acetonitrile with 0.1% formic acid) to 55% B over 30 min, next ramping

to 95% B over 10 min and holding 95% B for 10 min. One full MS scan (300-2000 MW) was followed by 3 data dependent scans of the most intense ions with dynamic exclusion enabled. The spray voltage was set to 1.94 kV and the flow rate through the column was set to 0.25 µL/min.

2.4.15 Data analysis

For protein identification all generated .raw files were searched using the SQUEST search algorithm under BioWorks 3.3.1 SP1. All searches were performed against a concatenated indexed database containing the international protein index (IPI) mouse database v3.49, the UniProtKB/Swiss-Prot HAMAP (http://ca.expasy.org/sprot/hamap/SALTY.html) database for *S. typhimurium* LT2/SGSC1412/ATCC700720, as well as the most common protein lab contaminants. The database was indexed in BioWorks 3.3.1 with carbamidomethylation as static modification, trypsin as proteolytic enzyme, 'fully enzymatic' as enzyme limit, 250-5000 as MW range, and 2 missed cleavage sites. SEQUEST searches were conducted with monoisotopic precursor and fragment masses, 50.00 ppm peptide tolerance, 1.0 AMU fragment ion tolerance, and B and Y ion series calculated. Carbamidomethylation was selected as static modification, methionine/tryptophan oxidation as well as asparagines/glutamine deamindation was selected as variable modification. The generated unified search files .srf were opened in Scaffold 2 (Proteome Software) and further analyzed. All searches were filtered in Scaffold 2 for 99.9% protein probability, 2 unique peptides, and 95 % peptide probability. Data was exported into Excel and further analyzed. All identifications with hits in the Met control sample were discarded, as well as all identified contaminations (e.g. keratins).

3 A phenylalanine-based orthogonal amino acid reporter[3]

3.1 Abstract

Orthogonal amino acid reporters (unnatural amino acids carrying an azide or alkyne motif, which can be cell-selectively incorporated) allow the selective labeling of different cell types in heterogenous populations through the expression of engineered aminoacyl-tRNA synthetases. Previously described orthogonal amino acid reporters (AOA and ANL, see chapter *2*) are limited to proteins that contain at least one methionine. Other amino acids are present more frequently and their substitution could allow more quantitative proteome labeling. The following chapter discusses the development of a phenylalanine-based orthogonal amino acid reporter, which allows labeling of the bacterial pathogen *S. typhimurium* during infection of mammalian cells.

3.2 Introduction

Radiolabeled or stable isotope-labeled amino acids, coupled with mass spectrometry, allow the differentiation of newly synthesized from pre-existing proteins (Schwanhäusser et al. 2011). Nonetheless, it is still challenging to analyze changes in protein expression of individual cell types from heterogeneous populations since these labeled amino acids are incorporated into all proteomes. The differentiation of individual proteomes could significantly aid in the analysis of complex cell mixtures, such as bacterial cells during

---

[3] This work was carried out in collaboration with Paul Dossa (The Rockefeller Univeristy) and Emma Taylor-Salmon (Bard College). Paul Dossa and Emma Taylor-Salmon conducted some of the *in vitro* labeling experiments.

host infection. Robust methodologies are thus needed to differentiate the proteomes of individual cell types and to separate them from protein contaminations of other cell types. Two recently introduced orthogonal amino acid reporters provide a means for cell-selective labeling of newly synthesized proteins with azide or alkyne reporter groups through the expression of mutant methionyl-tRNA synthetases (MetRS*) (Ngo et al. 2009; Grammel et al. 2010). Expression of MetRS* in *S. typhimurium* allows the cell-selective co-translational incorporation of the orthogonal amino acid reporters azidonorleucine (ANL, Figure 19a) and 2-amino-7-octynoic acid (AOA, Figure 19a) into newly synthesized bacterial proteins in place of methionine (Met) during infection of mammalian cells. Since ANL and AOA are not activated by endogenous tRNA synthetases, they are excluded from the mammalian proteome. The incorporation of AOA permits the subsequent visualization and proteomic analysis of AOA-labeled proteins using the Cu(I)-catalyzed alkyne-azide cycloaddition (CuAAC) with an azido-rhodamine fluorescence dye and azobenzene-biotin affinity tags, respectively (Y.-Y. Yang, Grammel, et al. 2010). Here we describe the use of para-ethynylphenylalanine (PEP, Figure 19a) for efficient cell-selective labeling of *S. typhimurium* during infection of host cells (Figure 19b).

Even though the majority of proteins contain at least one Met (e.g. 96% in *S. typhimurium* LT2), in addition to the initiation Met, other amino acids are present more frequently (e.g. 2.8% Met vs. 3.9% phenylalanine) and their replacement could permit greater penetration of the bacterial proteomes with a corresponding amino acid reporter. We therefore explored the possibility of alternative orthogonal amino acid reporters that would replace one of the other proteinogenic amino acids present in higher frequency

than Met. Interestingly, an *E. coli* phenylalanyl-tRNA synthetase (PheRS) mutant (A294G, PheRS*) with increased substrate tolerance has been described that activates a wide range of para-substituted phenylalanine (Phe) derivatives (Ibba et al. 1994; Kirshenbaum et al. 2002). One of these derivatives, PEP, contains an alkyne chemical handle for CuAAC and could thus be used as a chemical reporter.

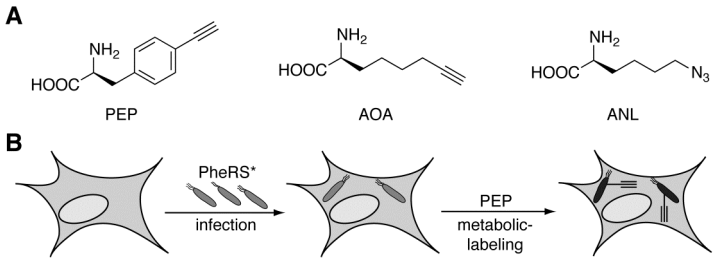


Figure 19 A) PEP: para-ethynylphenylalanine, AOA: aminooctynoic acid, ANL: azidonorleucine B) Infection of mammalian cells with *Salmonella typhimurium* expressing a mutant phenylalanyl-tRNA synthetase (PheRS*) allows selective metabolic labeling of the bacterial proteome with PEP during infection.

3.3 Results

3.3.1 Development of *para*-ethynylphenylalanine (PEP) as an orthogonal amino acid reporter for the analysis of the *S. typhimurium* proteome during infection

To achieve incorporation of PEP into *S. typhimurium*, we expressed the reported mutant *E. coli* alpha chain of PheRS (PheS), containing the A294G mutation ($PheS_{A294G}$), under the control of the constitutively active lac-promoter using the low-copy plasmid pWSK29 (pWSK29_$PheS_{A294G}$). Due to the high sequence conservation between the pheS gene in *E. coli* and *S. typhimurium* (98%), we anticipated that the over-expressed mutant *E. coli*

alpha chain would form a hybrid hetero tetramer with the endogenous copy of the *S. typhimurium* PheRS beta chain (pheT) to generate active mutant PheRS*.

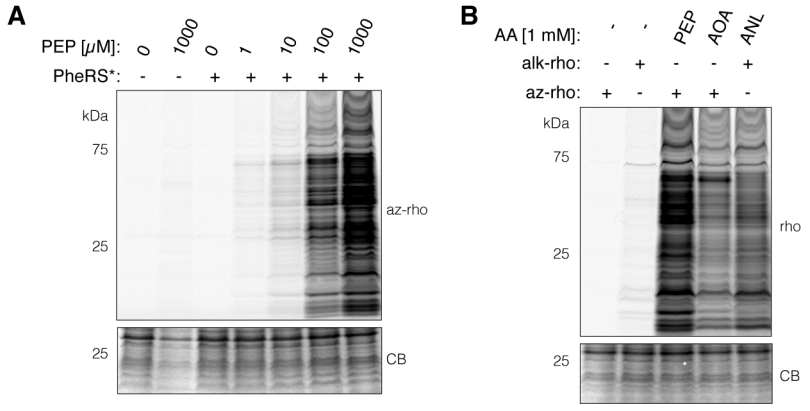


Figure 20 In-gel fluorescence analysis of PEP incorporation into the *S. typhimurium* proteome in M9 minimal medium. a) *S. typhimurium* or S. *typhimurium* pWSK29_PheS$_{A294G}$ was labeled with PEP in M9 minimal medium. b) *S. typhimurium* pWSK29_PheS$_{A294G}$ was labeled with PEP, *S. typhimurium* pWSK29_MetG$_{NLL}$ (see chapter 2) was labeled with AOA or ANL, in M9 minimal medium. PheRS*: expression of PheS$_{A294G}$, az-rho: azido-rhodamine, alk-rho: alkyne-rhodamine, rho: azido-rhodamine or alkyne-rhodamine, CB: coomassie blue.

To assess metabolic incorporation of PEP into the proteome of *S. typhimurium*, *S. typhimurium* pWSK29_PheS$_{A294G}$ and *S. typhimurium* pWSK29 was grown in liquid culture and incubated with different concentrations of PEP in minimal medium. Cell lysates were reacted with an azido-rhodamine fluorescent dye (az-rho) and analyzed by in-gel fluorescence scanning. Expression of PheS$_{A294G}$ facilitated incorporation of PEP into the *S. typhimurium* proteome, while wild-type *S. typhimurium* did not incorporate PEP (Figure 20a). Comparison of PEP with the Met-based amino acid reporters AOA and ANL showed comparable or even increased incorporation efficiency (Figure 20b), as anticipated owing to the higher frequency of phenylalanine in the *S. typhimurium*

proteome. PEP, indeed, serves as a phenylalanine surrogate in *S. typhimurium* (Figure 21a) and its incorporation is dependent on active ribosomal protein synthesis, since tetracycline inhibition abrogated the observed fluorescence signal (Figure 21b).

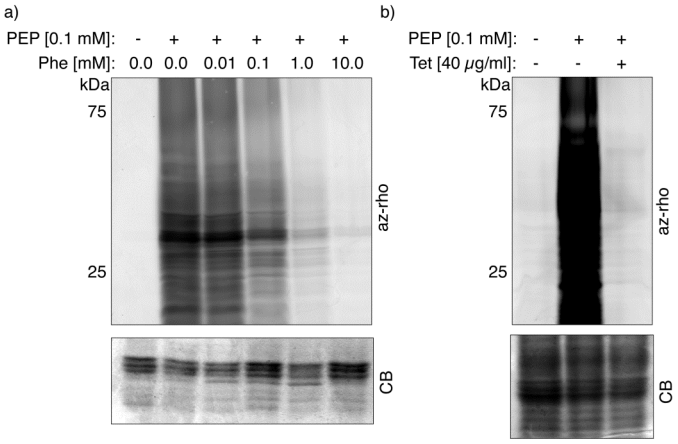


Figure 21 *Para*-enthynylphenylalanine (PEP) serves as a co-translational chemical reporter for phenylalanine in *S. typhimurium*. a) *S. typhimurium* pWSK29_PheS$_{A294G}$ was incubated with PEP and increasing concentrations of phenylalanine (Phe) in M9 minimal medium. b) *S. typhimurium* pWSK29_PheS$_{A294G}$ was incubated with PEP in the presence of tetracycline (Tet) in M9 minimal medium. az-rho: azido-rhodamine, CB: coomassie blue.

To evaluate the PEP/PheRS* system for cell-selective labeling, we analyzed *S. typhimurium* infections of mammalian cells. Initial PEP labeling experiments in RAW264.7 murine macrophages with minimal media revealed low level mammalian proteome labeling, but incubation of macrophages with PEP in full medium containing 10% fetal bovine serum did not lead to any detectable amino acid incorporation (Figure 22a). These observations suggests that PEP can be utilized by the endogenous translational machinery of mammalian cells at low levels, but it is readily competed away in the presence of exogenous Phe. We then proceeded with macrophage infection

experiments using *S. typhimurium* pWSK29_PheS$_{A294G}$. After 16 hours of infection, infected macrophages were pulse-labeled with PEP or Phe for 1 hour in full medium. The cell lysates were then reacted with az-rho and analyzed for PEP incorporation by in-gel fluorescence. Infection of macrophages with *S. typhimurium* pWSK29_PheS$_{A294G}$ resulted in PEP labeling of numerous proteins (Figure 22b), while incubation of uninfected macrophages with PEP did not yield any significant incorporation, suggesting selective labeling of the bacterial cells inside the infected macrophages.

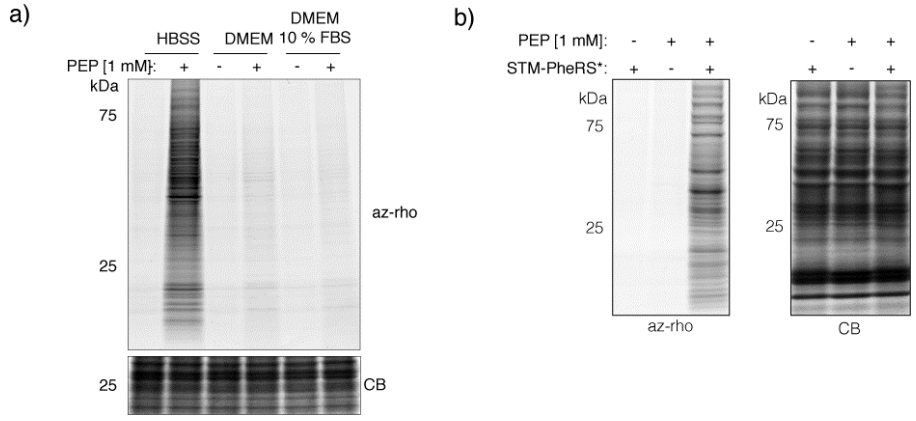


Figure 22 PEP incorporation into mammalian proteomes can be suppressed. a) In-gel fluorescence analysis of *para*-ethynylphenylalanine (PEP) labeling of RAW264.7 murine macrophages. HBSS: Hank's buffered saline solution, DMEM: Dulbecco's Modified Eagle Medium, FBS: fetal bovine serum. b) In-gel fluorescence analysis of PEP-labeled *S. typhimurium* pWSK29_PheS$_{A294G}$-infected RAW264.7 murine macrophage cell lysates. PEP: para-ethynylphenylalanine, STM-PheRS*: *S. typhimurium* pWSK29_PheS$_{A294G}$, az-rho: azido-rhodamine fluorescence, CB: coomassie blue.

To confirm selective incorporation of PEP into *S. typhimurium* pWSK29_PheS$_{A294G}$ during intracellular infection, we performed fluorescence microscopy following CuAAC. In this case, infected RAW264.7 or HeLa cells were fixed after pulse labeling with PEP and reacted with az-rho. Rhodamine fluorescence was

only observed in the samples pulsed with PEP and was absent from Phe-treated samples (Figure 23). The rhodamine signal co-localized well with anti-*S. typhimurium* antibody staining and was absent from the mammalian cytosol, which demonstrates cell-selective labeling of bacterial proteomes using the PEP/PheRS* system. While it is possible that these amino acid reporters alter protein function or cell viability (Beatty et al. 2005; Ngo & Tirrell 2011), in the context of our pulse-labeling experiments we did not observe significant defects in protein expression, recovery, or cell viability.

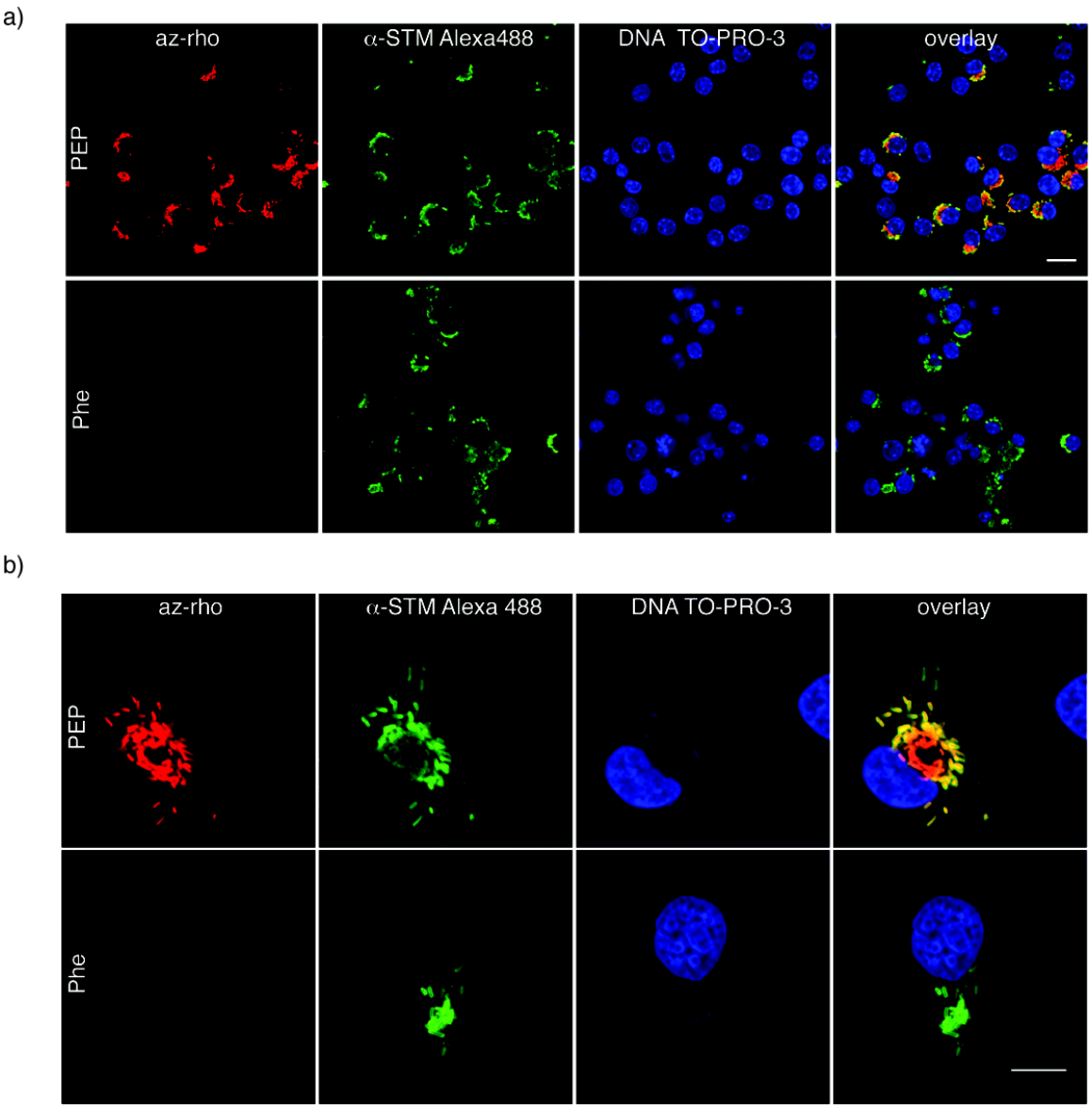


Figure 23 a) Fluorescence microscopy of PEP-labeled *S. typhimurium* pWSK29_PheS$_{A294G}$-infected RAW264.7 murine macrophages. b) Fluorescence microscopy analysis of infected PEP pulse-labeled HeLa cells. HeLa cells were infected with *S. typhimurium* pWSK29_PheS$_{A294G}$ and pulse-labeled with PEP 16 h after infection. PEP: *para*-ethynylphenylalanine, Phe: phenylalanine, az-rho: azido-rhodamine (red), a-STM Alexa488: anti-*S. typhimurium* antibody stain visualized with secondary Alexa488 conjugate (green), DNA TO-PRO-3: DNA stain (blue), scale bar indicates 10 μm.

3.3.2 Double labeling of *S. typhimurium* with 2-amino-7-octynoic acid (AOA) and *para*-ethynylphenylalanine (PEP)

PEP and AOA (or ANL) (see chapter 2) can be selectively incorporated into bacterial pathogens instead of phenylalanine and methionine, respectively. We have also learned that both systems share the intrinsic limitation to the corresponding subproteome, containing either phenylalanine or methionine. A combination of the two reporter systems might allow wider coverage of the proteome and a higher reporter incorporation density per protein. The concomitant expression of $PheS_{A294G}$ and MetRS* (more specifically $MetG_{NLL}$, see chapter 2) in the same bacterium should allow for simultaneous incorporation of PEP and AOA (Figure 19a) into the same proteome (Figure 24a). In addition, the exclusive nature of azides and alkynes could provide a means of targeting individual bacterial population in the same experiment by using PEP (alkyne) in combination with ANL (azide) (Figure 19a, Figure 24b). The expression of $MetG_{NLL}$ in one population would facilitate the incorporation of ANL, while the expression of $PheS_{A294G}$ in a second population would allow the incorporation of PEP.

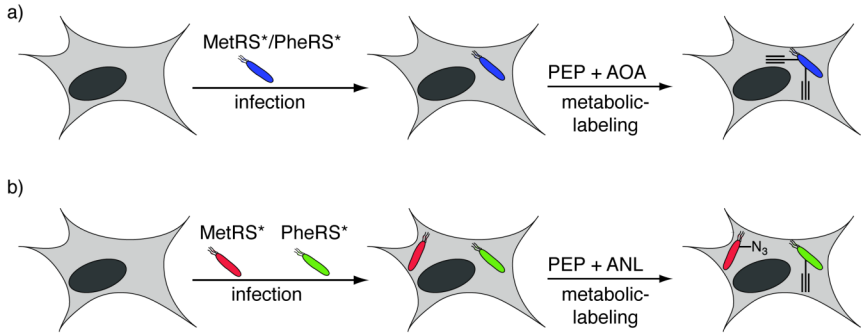


Figure 24 Double labeling with cell-selective phenylalanine and methionine surrogates a) Double expression of MetRS* and PheRS* allows concomitant incorporation of AOA and PEP. b) Individually MetRS* and PheRS* expressing strains can be differentiated.

To provide concomitant MetRS* and PheRS* activity in the same *S. typhimurium* strain a bicistronic expression construct was generated. $PheS_{A294G}$was amplified by PCR and a N-terminal HA-tag as well as a 5' ribosomal binding site (RBS) was introduced with the primers. The generated fragment was cloned into the previously described pWSK29_$MetG_{NLL}$ construct (see chapter 2) to generate the bicistronic dual expression construct pWSK29_$MetRS_{NLL}$_$PheRS_{A294G}$ (Figure 25a). This construct was transformed into *S. typhimurium* and incorporation of PEP and AOA was tested *in vitro*. Overnight cultures of *S. typhimurium* pWSK29_$MetG_{NLL}$, *S. typhimurium* pWSK29_$PheS_{A294G}$, and *S. typhimurium* pWSK29_$MetRS_{NLL}$_$PheRS_{A294G}$ were diluted 1:30 into fresh LB medium and grown for 3 hours. Aliquots of these cultures were pelleted and resuspended in equal volumes of M9 minimal medium with PEP, AOA, or PEP and AOA. The dual bicistronic expression construct facilitates PEP and AOA incorporation, even though a diminished PEP incorporation can be observed in comparison to the pWSK29_$PheS_{A294G}$ construct (Figure 25b). This might be due to a reduced expression level of $PheS_{A294G}$, since it is the second ORF on the bicistronic messanger RNA. Both proteins are expressed ($MetG_{NLL}$ as well as $PheS_{A294G}$), as controlled for by Western blot analysis (Figure 25b). The concomitant labeling of the S. *typhimurium* pWSK29_$MetRS_{NLL}$_$PheRS_{A294G}$ strain with AOA and PEP yields increased labeling in comparison to all mono-labeling combinations (Figure 25b). These experiments indicate that the simultaneous labeling with PEP and AOA in combination with dual expression of $MetG_{NLL}$ and $PheS_{A294G}$ improves the overall labeling amount.

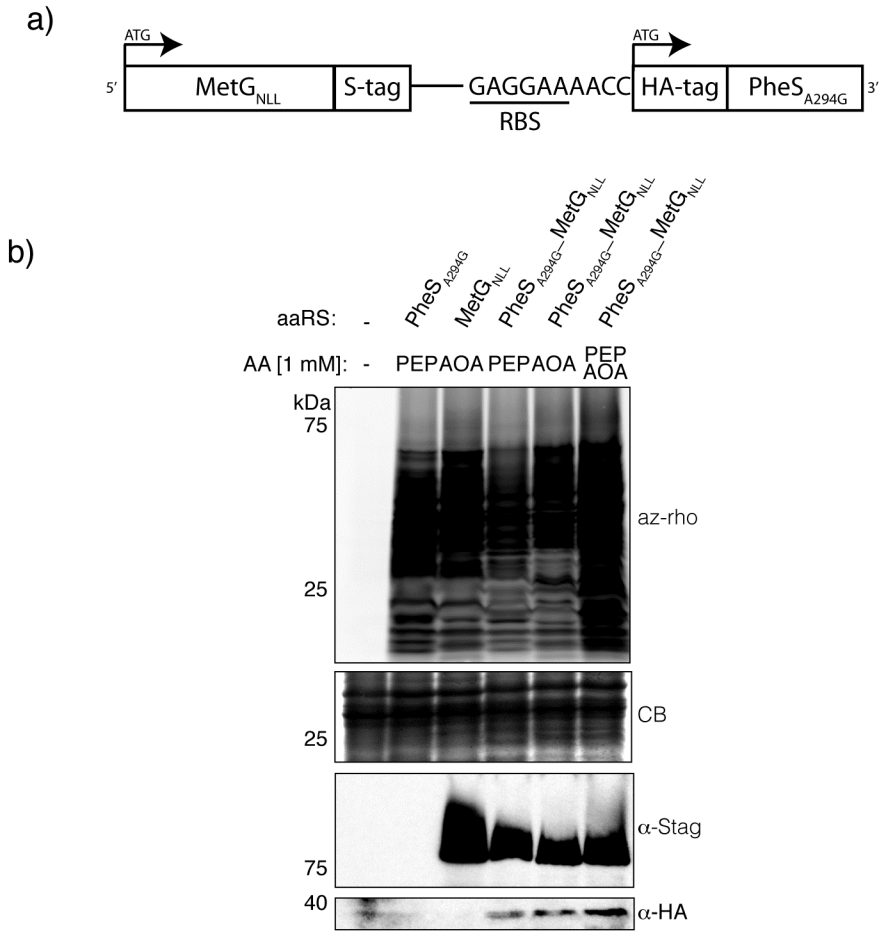


Figure 25 Dual labeling with PEP and AOA for improved labeling intensity. a) A bicistronic expression construct for dual expression of $MetG_{NLL}$ and $PheS_{A294G}$ pWSK29_$MetG_{NLL}$_$PheS_{A294G}$. b) In-gel fluorescence analysis of *S. typhimurium* strains labeled with PEP, AOA, or PEP/AOA in minimal medium. aaRS: aminoacyl-tRNA-synthetase, AA: amino acid, az-rho: azido-rhodamine, CB: coomassie blue, α-Stag: anti-Stag antibody staining, α-HA: anti-HA tag antibody staining, RBS: ribosomal biding site.

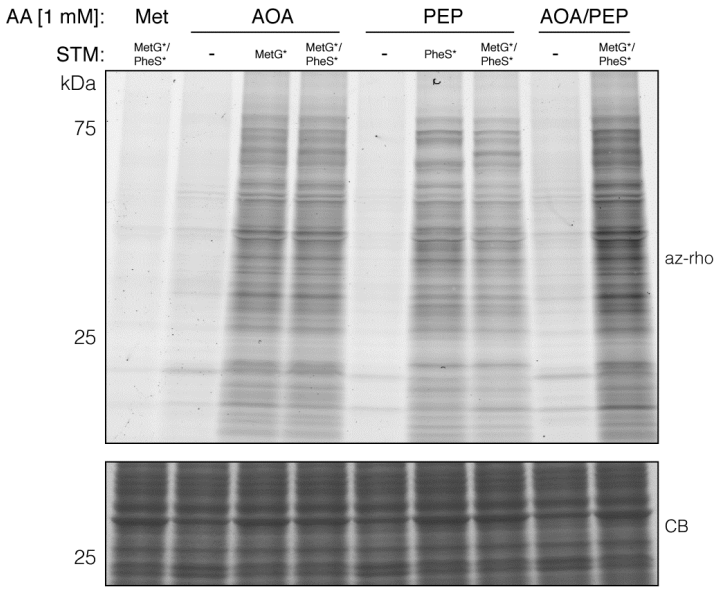


Figure 26 Dual labeling of infected cells. RAW264.7 In-gel fluorescence analysis of cells, infected with *S. typhimurium* strains and pulse-labeled with AOA, PEP, or AOA/PEP 16 hours post-infection. AA: amino acid, STM: *S. typhimurium* IR715, MetG*: pWSK29_$MetG_{NLL}$, PheS*: pWSK29_$PheS_{A294G}$, MetG*/PheS*: pWSK29_$MetG_{NLL}$_$PheS_{A294G}$, az-rho: azido-rhodamine fluorescence, CB: coomassie blue.

To investigate the labeling efficiency of the dual labeling system during infection, RAW264.7 macrophages were infected with stationary phase *S. typhimurium* pWSK29_$MetG_{NLL}$, *S. typhimurium* pWSK29_$PheS_{A294G}$, or *S. typhimurium* pWSK29_$MetG_{NLL}$_Phe_{SA294G}, at an MOI of 50. 16 hours post-infection, the infected macrophages were pulse labeled with 1 mM PEP, 1 mM AOA, or 1 mM PEP/1 mM AOA for 1 hour. Analysis of total cell lysates of infected cells, generated by SDS lysis, by CuAAC and in-gel fluorescence analysis, revealed that the dual expression construct yields almost equal overall labeling intensities as the mono-expression constructs (Figure 26). Dual labeling with AOA and PEP of *S. typhimurium* pWSK29_$MetG_{NLL}$_$PheS_{A294G}$

infected cells seems to provide a slightly elevated overall labeling intensity (Figure 26), however this result was not constantly reproducible (data not shown), indicating a large degree of variability in the infection setting, potentially due to amino acid uptake, or aaRS expression issues.

In addition to the simultaneous labeling of a single bacterial strain with two alkyne reporters (AOA and PEP) for improved proteome coverage, we also explored the individual labeling of two different bacterial strains in the same experiment with distinguishable reporters (PEP and ANL). The expression of MetRS* in one strain should facilitate the exclusive incorporation of ANL while the expression of PheRS* on the other strain should exclusively allow the incorporation of PEP. For this purpose RAW264.7 macrophages were infected with an equal ratio of *S. typhimurium* pWSK29 *$MetG_{NLL}$ and S. typhimurium* pWSK29_$PheS_{A294G}$. 16 hours post-infection, the dually infected cells were pulse-labeled with a 1:1 mix of 1 mM PEP and 1 mM AOA for 1 hour. Subsequently the cells were prepared for fluorescence microscopy and successively reacted with two orthogonal fluorescent dyes by CuAAC - alk-CyFur (a 2-dicyanomethylene-3-cyano-2,5-dihydrofuran based dye) (Tsou et al. 2009) and az-rho, which allow independent visualization of the tagged species my florescence microscopy. Since the visualization of ANL incorporation in *S. typhimurium* during infection with alk-CyFur has never been carried out before, singly infected cells with *S. typhimurium* pWSK29_$MetG_{NLL}$ were pulse-labeled with ANL and imaged with alk-CyFur in parallel. ANL was selectively incorporated into *S. typhimurium* pWSK29_$MetG_{NLL}$ inside macrophages as was demonstrated by co-localization of the alk-CyFur fluorescence

signal with anti-*Salmonella* antibody staining (Figure 27a). No alk-CyFur fluorescence signal could be observed in the methionine labeled control sample. In the double infected,

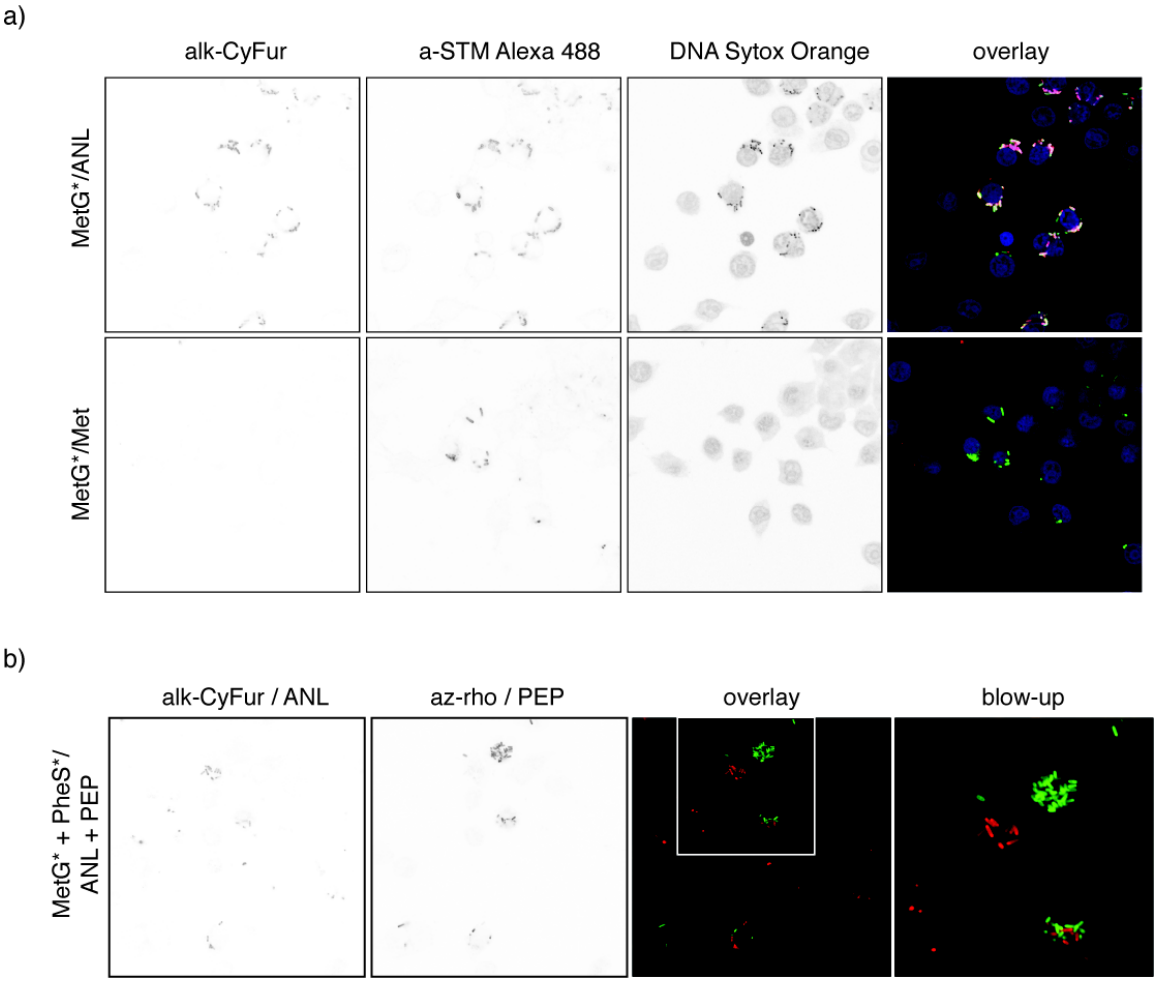


Figure 27 Visualization of two independent bacterial populations using PEP and ANL. a) RAW264.7 cells were infected with *S. typhimurium* pWSK29_MetG$_{NLL}$ and pulse-labeled with ANL. Infected cells were visualized with alk-Cy-Fur (Tsou et al. 2009). Blue: DNA Sytox Orange, green: anti-*Salmonella* antibody stain, red: alk-CyFur. b) RAW264.7 cells were infected with *S. typhimurium* pWSK29_PheS$_{A294G}$ and *S. typhimurium* pWSK29_MetG$_{NLL}$ and pulse-labeled with PEP and ANL. Cells were visualized with alk-CyFur and az-rho. Red: alk-CyFur, green: az-rho. (MetG*: pWSK29_MetG$_{NLL}$, PheS*: pWSK29_PheS$_{A294G}$, az-rho: azido-rhodamine, alk-CyFur: alkyne-cyanofuran based dye.)

PEP and AOA co-labeled macrophages, two distinct populations of bacteria could be identified. The fluorescence images clearly showed alk-CyFur positive bacteria, indicating ANL incorporation (red, Figure 27b) and therefore pWSK29_$MetG_{NLL}$ identity, as well as az-rho positive bacteria (green, Figure 27b), indicating PEP incorporation and therefore pWSK29_$PheS_{A294G}$ identity. No bacteria showed co-localization of both fluorescence dyes, indicating highly selective incorporation of the two orthogonal amino acid reporters. Interestingly, most infected mammalian cells harbored only one bacterial strain (red or green), only in a few isolated cases both strains were present in the same macrophage (Figure 27b, blow-up). Such a scenario could arise from a secondary infection over the course of the replication period or from two synchronous infections during the infection period.

3.4 Materials and Methods

3.4.1 Chemical reagents

Azido-rhodamine (az-rho), alkyne-rhodamine (alk-rho), alk-CyFur, and TBTA (tris((1-benzyl-1H-1,2,3-triazol-4-yl)methyl) amine) were generated by the Hang laboratory as previously described (Charron et al. 2009; Tsou et al. 2009). PEP, AOA, and ANL were generated as previously described ((Kirshenbaum et al. 2002; Grammel et al. 2010). All other chemical reagents were purchased from Sigma-Aldrich.

3.4.2 Plasmids and bacterial strains

Salmonella typhimurium IR715 was transformed with pWSK29 encoding the A294G mutant of the *E. coli pheS* gene. *PheS* was amplified from *E. Coli* K12 by PCR (primer 1 and primer 2) and sub-cloned into pJET1.2. The resulting fragment contained a mutation relative to the reported UniProt sequence (P08312, D258N, GAC to AAC), which was re-

mutated to the reference sequence together with the introduction of mutation (A294G) by use of the QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene, primer 3, primer 4). The fragment was excised from pJET1.2 and inserted into pWSK29 with SacI and KpnI to generate pWSK29_$PheS_{A294G}$.

To generate pWSK29_$MetG_{NLL}$_$PheS_{A264G}$, $PheS_{A264G}$ was amplified from pWSK29_$PheS_{A264G}$ with the primer 5 and primer 6. The resulting fragment was cloned into pWSK29_$MetG_{NLL}$(Grammel et al. 2010) with SalI and KpnI.

Table 1 Primers for $PheS_{A264G}$ expression constructs.

1	GAGCTCCATGTCACATCTCGCAGAACTGG
2	GGTACCCGCATCGCTATCAATCGCC
3	GAACCTTCTGCAGAAGTGGACGTCATGGGTAAAAACG
4	GTTTACTCTGGTTTCGGCTTCGGGATGGGGATG
5	GTCGACTGAGGAAAACCATGTACCCGTACGACGTGCCGGATTACGCG TCACA
6	GGTACCCGCATCGCTATCAATCGCCGG

3.4.3 Bacterial pulse labeling in liquid culture

Overnight cultures of bacteria were grown in full LB with 100 μg/ml Carbenicillin (Cb). These cultures were then diluted 1:30 into fresh LB (Cb) and grown to stationary phase. Bacteria were collected by centrifugation (10,000 rpm, 3 min) and resuspended in M9 (1x M9 salts, 20 mM glucose, 2 mM $MgSO_4$, 100 μM $CaCl_2$) minimal medium (Cb). Bacteria were grown in minimal medium for 15 min and collected by centrifugation. Bacterial pellets corresponding to 1 ml of culture volume were resuspended in M9 minimal

medium (Cb) containing the chemical reporter. Bacteria were labeled for 1 hour and collected by centrifugation, washed once in PBS, and stored at -80 °C. For tetracycline (Tet) protein synthesis inhibition, bacteria were grown for 30 min in LB (40 μg/ml Tet) before being labeled for 30 min in M9 minimal medium (40 μg/ml Tet).

3.4.4 Infection of mammalian cells and pulse labeling

For fluorescence microscopy experiments, 3 x 10^5 RAW264.7 cells were split into 12 well plates 24 h before infection on round glass cover slips. For subsequent cell lysis and in-gel fluorescence experiments, 1.5 x 10^6 RAW264.7 cells were split into 6 well plates 24 h before infection. An LB overnight culture of *Salmonella* was diluted 1:30 into fresh LB with 100 μg/ml carbenicillin and grown to stationary phase (OD600 > 2.0). Bacteria were collected by centrifugation and washed once with cold PBS. For a multiplicity of infection (MOI) of 100, an appropriate volume of bacteria in PBS was added to DMEM (10% FBS) and put on ice for 30 min. Bacteria in DMEM (10% FBS) were added to RAW264.7 cells and centrifuged at 1000g for 5 min at 4 °C. The infection was allowed to proceed for 25 min at 37 °C at 5% CO_2. The wells were then washed 4 times with PBS before the cells were covered in DMEM (10% FBS) with 100 μg/ml gentamicin and incubated or 90 min at 37 °C at 5% CO_2 to kill all extracellular bacteria. Subsequently, wells were washed 2 times with PBS and filled with DMEM (10% FBS) with 12.5 μg/ml gentamicin. The cells were further incubated for 14.5 h at 37 °C at 5% CO_2. The cells were pulse labeled with 1 mM PEP in full medium (DMEM 10% FBS) for 1 h at 37 °C at 5% CO_2. After 1 h the cells were washed 4 times with cold PBS.

3.4.5 Cell lysis, click-chemistry, and in-gel fluorescence analysis

For cell lysis, infected RAW264.7 cells were collected by scraping in ice-cold PBS and centrifugation at 1000g for 5 min at 4 °C. Cell pellets were lysed by resuspending them in 0.1% SDS in PBS with Roche complete protease inhibitor, 1 mM PMSF, and 1 μl benzonase/ml (Sigma E1014), and then sonicating them for 30 s in a water bath sonicator, followed by 10 min incubation on ice. Samples were filled up with 12% SDS in PBS to a final concentration of 4% SDS and sonicated again. Then, remaining debris was removed by centrifugation at 20,000g for 10 min at room temperature. Protein concentrations were determined by BSA assay (Pierce). For click-chemistry, 50 μg of cell lysate from each sample was filled up to the smallest common volume with 4% SDS in PBS. A click chemistry master mix was prepared and added to each sample: 100 μM azido-rhodamine (az-rho) from a 10 mM stock solution in DMSO, 1 mM TCEP (tris(2-carboxyethyl)phosphine) from a 50 mM stock solution in water (freshly prepared), 100 μM TBTA (tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine) from a 2 mM stock solution in 4:1 n-butanol:DMSO, and 1 mM $CuSO_4$ from a 50 mM stock solution in water (freshly prepared).

The components were added in the order listed above to the master mix and vigorously vortexed. Then the master mix was added to each individual protein sample to yield a final volume of 50 μl at the final concentrations indicated. The samples were incubated for 1 h at room temperature.

The samples were precipitated after click chemistry to remove small molecules. For this purpose, 50 μl of water was added to each sample, followed by 400 μl of methanol, and 100 μl of chloroform. The samples were vigorously vortexed and 300 μl of water was added. The samples were vortexed again and centrifuged at 20,000 g for 5 min at room

temperature. The upper phase was removed carefully and 1 ml of ice-cold aetone was added. The samples were vortexed and centrifuged at 20,000 g for 15 min at 4 °C. The organic supernatant was discarded and the samples briefly dried on air. Then, the samples were resuspended in 30 μl 4% SDS in PBS and briefly sonicated. 10 μl of 4 times Laemmli buffer was added to each sample together with 2 μl of beta-mercaptoethanol, samples were incubated at 95 °C for 5 min, and run for analysis on a 4-20% Tris-HCl Criterion™ Precast Gel (Biorad). Protein gels were rinsed with deionized water and incubated for multiple hours in 50% water, 40% methanol, 10% acetic acid at 4 °C. Before in-gel fluorescence analysis, protein gels were transferred to deionized water and incubated for at least 30 min at room temperature. Protein gels were scanned on an Amersham Bioscience Typhoon 9400 variable mode imager (excitation 532 nm, 580 nm filter, 30 nm band-pass). All image adjustments were done on the entire depicted image for all samples equally in Adobe® Photoshop® CS3. After fluorescence scanning, protein gels were stained with GelCode Blue Stain Reagent (Thermo) or Coomassie Brilliant Blue (BioRad). Both stains are indicated as coomassie blue (CB) in all figures.

3.4.6 Sample preparation for fluorescence microscopy

For fluorescence microscopy, cells were fixed and permeabilized in 3.7% paraformaldehyde, 0.2% Triton X-100 in PBS for 15 min at room temperature. The cells were then washed at least 4 times for 5 min each with PBS. A click-chemistry master mix was prepared from stock solutions listed above to yield the following final concentrations in PBS (20 μM az-rho, 1 mM TCEP, 100 μM TBTA, 1 mM $CuSO_4$). The cells were covered in click chemistry master mix and incubated for 1 h at room temperature. Then, the cells were blocked for 1 h with 2% FBS, 1% Tween 20 in PBS at room temperature.

The cells were washed once with PBS and anti-*S. typhimurium* serum was added (1:250) in 0.2% saponin in PBS for 30 min at room temperature. The cells were washed 4 times with PBS and subsequently incubated with anti-rabbit Alexa488-conjugated antibody in 0.2% saponin in PBS for 30 min at room temperature. Cells were washed in PBS and finally incubated with TO-PRO-3 DNA stain (1:1000) in 0.2% saponin in PBS. The cells were washed once with 0.2% saponin in PBS and 4 times in PBS before mounting on microscope slides in ProLong Gold (Invitrogen). The images were taken on an upright LSM510 laser scanning confocal microscope (Zeiss) equipped with a Krypton/Argon laser with 488 and 568 lines and a HeNe laser with a 633 line, using a 100x oil immersion objective.

For dual labeling experiments, sample preparation was conducted equally with small changes. After permeabilization and fixation the first CuAAC with alk-CyFur was carried out with 200 μM alk-CyFur for 1 hours at room temperature. The samples were washed two times with 20% methanol in PBS for 10 min to remove any un-reacted CuAAC reagents. The samples were washed with PBS and the second CuAAC with 20 μM az-rho was carried out. Since the emission spectrum of TO-PRO-3 overlaps with the emission spectrum of alk-CyFur DNA was visualized with the DNA stain Sytox Orange (1:1000 dilution of a 5 mM stock).

4 Methionine amino acid reporters based on homoserine, homocysteine, and cysteine

4.1 Abstract

Orthogonal amino acid reporters, such as 2-amino-7-octynoic acid (AOA), azidonorleucine (ANL), or *para*-phenylalanine (PEP) enable the cell-selective labeling of bacterial populations in the presence of mammalian cells. Their incorporation is dependent on the heterologous expression of mutant *E. coli* aminoacyl-tRNA-synthetases (aaRS), which requires stable transformation with a plasmid encoding the aaRS or stable integration of the aaRS gene into the bacterial chromosome. Far from all bacterial species are amenable to these genetic manipulations. Therefore it would be highly beneficial for the study of bacterial pathogenesis, if amino acid reporters could be developed with intrinsic selectivey for the bacterial translational machinery. Here, I present the initial progress towards such an intrinsically selective amino acid reporter.

4.2 Introduction

Protein expression is a highly dynamic process that is tightly regulated by intrinsic cellular signals as well as external environmental cues. In particular, bacterial pathogens adapt to environmental changes during the course of a host infection by dramatically altering their protein expression. Many pathogens produce the necessary gene products for infection (virulence factors) just in time as they encounter different host environments (e.g. gut, cell surface, cytosol, intracellular vacuolar compartments). Capturing these dynamic changes is indispensable for a complete understanding of the pathogen-host interaction and promises to be helpful in combating the incipient downturn of effective antimicrobials.

To this end selective labeling methods have been developed that can take snap-shots of the newly synthesized bacterial proteome during infection of mammalian cells. Orthogonal amino acid reporters (unnatural amino acids) with alkyne or azide reporter groups can be selectively and co-translationally incorporated into the bacterial proteome through the over-expression of engineered aminoacyl-tRNA-synthetases (aaRS). These reporters allow tagging of newly synthesized proteins with fluorescence dyes for visualization or affinity enrichment tags for mass spectrometry-based protein identification by the Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC).

Two methionine surrogates, azidonorleucine (ANL) and 2-amino-7-octynoic acid (AOA) can be activated by various *E. coli* methionyl-tRNA-synthetase (MetRS) mutants (MetRS*) and selectively incorporated into *E. coli* or intracellularly replicating *S. typhimurium* in the presence of translationally active host cells. In addition, *para*-ethynylphenylalanine (PEP) can also be used as an orthogonal amino acid reporter, since its activation is dependent on a mutant *E. coli* phenylalanyl-tRNA synthetase (PheRS*). We have learned that the over-expression of these *E. coli* aaRS mutants facilitate the required enzymatic activity in the distinct, but closely related species, *S. typhimurium*. However, in an evolutionarily more distant species the heterologous *E. coli* aaRS might not recognize the cognate tRNA anymore and therefore would not activate ANL or AOA. Furthermore, various bacterial pathogens are not yet genetically tractable and would not permit heterologous over-expression of a mutant aaRS. In our studies of alternative alkynyl orthogonal amino acid reporters we came across a potential solution for these problems, an intrinsically selective orthogonal amino acid reporter.

4.3 Results

Even tough AOA provides better signal-to-noise than the isosteric azide reporter (ANL) in *S. typhimurium*, the low yield of its synthetic route (<10% of L-AOA) led us to investigate alternative reporter scaffolds for methionine. The direct alkylation of homocysteine or homoserine with alkynyl electrophiles should provide facile synthetic routes for the large-scale synthesis of enantiomerically pure potential amino acid reporters for methionine, with a sulfur or oxygen at the same position as in methionine. Therefore we synthesized (*S*)-2-amino-4-(prop-2-yn-1-ylthio)butanoic acid (EME) from a N-*tert*-butoxycarbonyl-protected L-homocysteine thiolactone and (*S*)-2-amino-4-(prop-2-yn-1-yloxy)butanoic acid (PHS) from N-*tert*-butoxycarbonyl-protected L-homoserine. EME constitutes an ethynyl-derivative of methionine displaying the sulfur-heteroatom at the same position as methionine. PHC, the isoster of EME, has oxygen at the position of sulfur. We further wanted to test whether MetRS might still accept more diverse derivatives, with a switched sulfur-position, and therefore generated (*R*)-2-amino-3-(but-3-yn-1-ylthio)propanoic acid (BCY) from N-acetyl-protected L-cysteine and (*R*)-2-amino-3-(prop-2-yn-1-ylthio)propanoic acid (PCY) fom N-acetyl-protected L-cysteine.

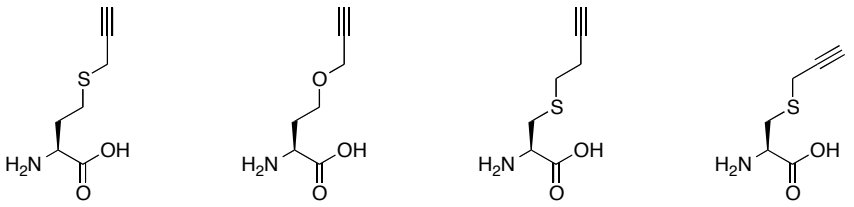


Figure 28 Potential methionine surrogates. (S)-2-amino-4-(prop-2-yn-1-ylthio)butanoic acid (EME), (S)-2-amino-4-(prop-2-yn-1-yloxy)butanoic acid (PHS), (R)-2-amino-3-(but-3-yn-1-ylthio)propanoic acid (BCY), (R)-2-amino-3-(prop-2-yn-1-ylthio)propanoic acid (PCY).

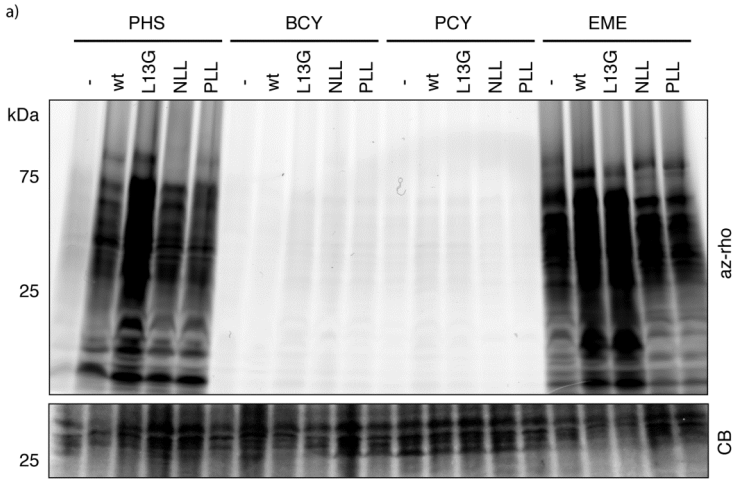


Figure 29 PHS and EME are incorporated into the *S. typhimurium* proteome. a) *S. typhimurium* pWSK29, pWSK29_MetG, pWSK29_MetG$_{L13G}$, pWSK29_MetG$_{NLL}$, or pWSK29_MetG$_{PLL}$ was incubated with PHS, BCY, PCY, or EME in M9 minimal medium, lysed, and analyzed by SDS-PAGE and in-gel fluorescence scanning. az-rho: azido-rhodamine, CB: coomassie blue.

Next, we tested incorporation of EME, PHC, BCY, and PCY in *S. typhimurium* as a function of heterologous MetRS activity. EME, PHC, BCY, and PCY was added to *S. typhimurium,* expressing different MetG mutants, grown in M9 minimal medium and analyzed the generated cell lysates by CuAAC with azido-rhodamine (az-rho) and in-gel fluorescence scanning. This analysis revealed that EME and PHC are incorporated into the *S. typhimurium* proteome upon expression of the MetRS mutants L13G, NLL, and PLL (Figure 29). Furthermore the mere over-expression of wild-type MetRS also sufficed to incorporate EME or PHC. But most interestingly, EME was even accepted by wild-type *S. typhimurium* without the over-expression of any MetRS (Figure 29). EME serves as an amino acid reporter for methionine, which is incorporated co-translationally.

Provided that mammalian cells do not accept EME, it could therefore serve as an intrinsically selective amino acid reporter for wild-type *S. typhimurium* and potentially other bacterial pathogens.

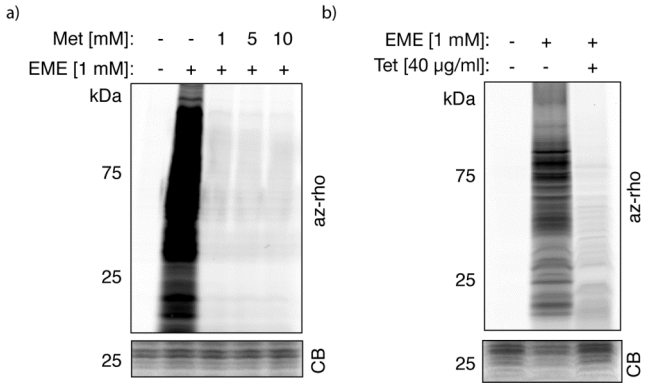
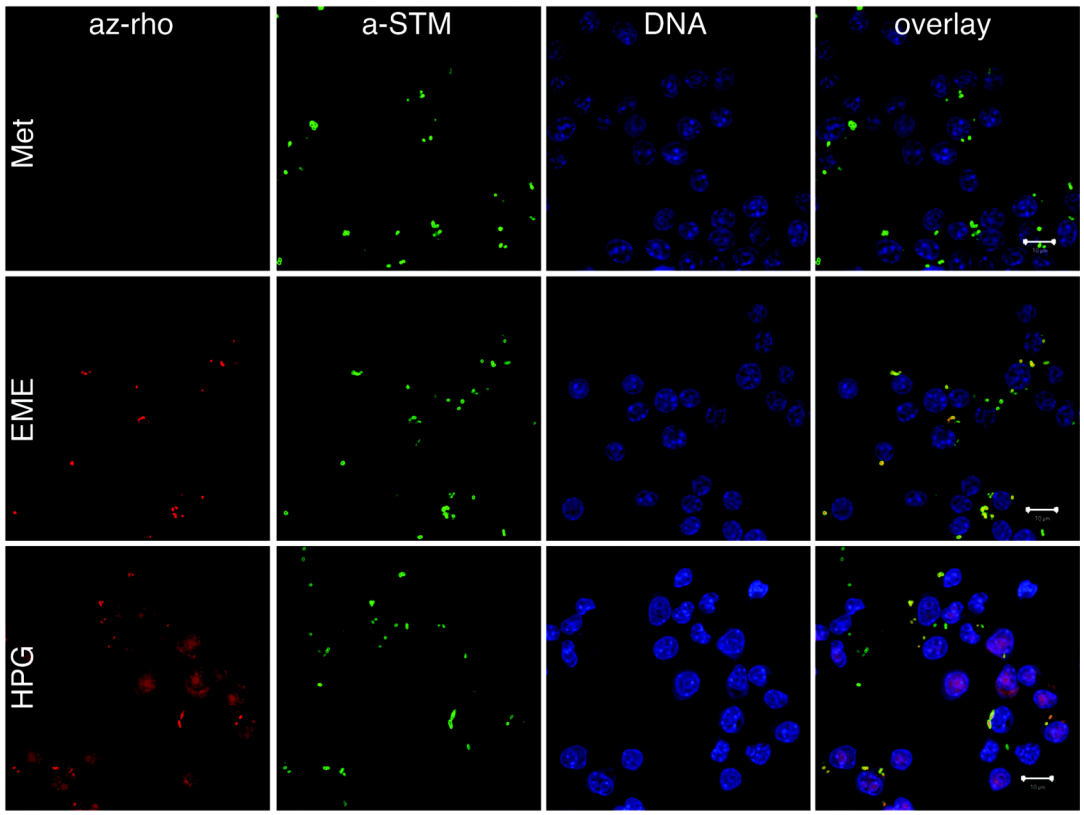


Figure 30 EME serves as an amino acid reporter for wild type *S. typhimurium* that is co-translationally incorporated instead of methionine. a) *S. typhimurium* IR715 was incubated with increasing concentrations of methionine (Met) and EME in M9 minimal medium. b) *S. typhimurium* IR715 was incubated with EME in the presence of tetracycline (Tet) in M9 minimal medium. az-rho: azido-rhodamine, CB: coomassie blue.

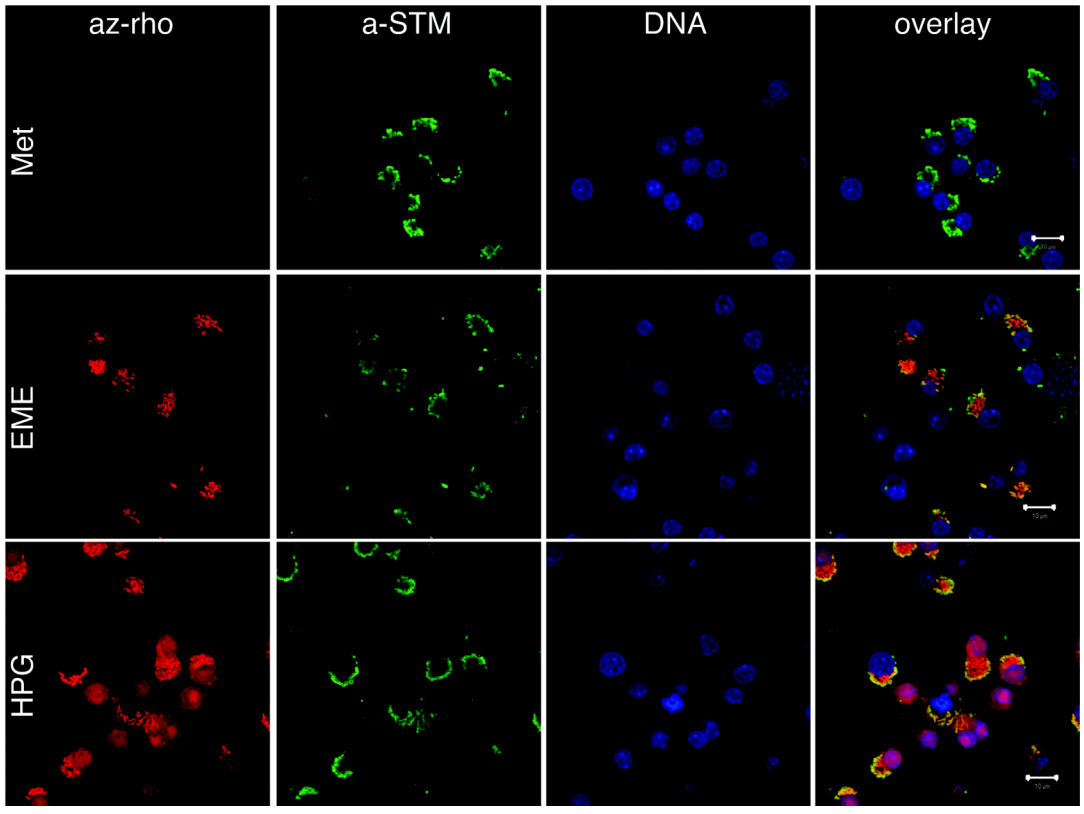
To test this hypothesis we conducted infection experiments with wild-type *S. typhimurium* and RAW264.7 macrophages. Before and 10 hours post-infection we pulse labeled with 1 mM Met, 1 mM EME or 1 mM homopropargylglycine (HPG) for 1 hour. HPG is an unselective amino acid reporter for methionine that is incorporated into mammalian and bacterial proteomes. The infected cells were fixed, reacted with az-rho by CuAAC and further prepared for analysis by fluorescence microscopy. EME was selectively incorporated into bacterial cells shortly after infection (0 hours) or 10 hours post-infection, as was indicated by the co-localization of the az-rho signal with the anti-*S. typhimurium* antibody staining and the absence of az-rho signal from the mammalian cytosol. In contrast, HPG was readily incorporated into bacterial and mammalian cells, as

Figure 31 EME allows selective labeling of wild-type *S. typhimurium* during infection of mammalian cells and visualization by fluorescence microscopy. RAW264.7 macrophages were infected with wild-type *S. typhimurium* and pulse labeled with Met, HPG, or EME 0 hours after infection (a) or 10 hours (b) post-infection. az-rho: azido-rhodamine (red), a-STM: anti-*S. typhimurium* antibody staining (green), DNA: DNA staining with TO-PRO-3, Met: methionine, HPG: homopropargylglycine.

a)


az-rho
a-STM
DNA
overlay
Met
EME
HPG

b)


az-rho
a-STM
DNA
overlay
Met
EME
HPG

was expected. No az-rho signal was observed in the Met labeled samples. These data show that EME possesses intrinsic selectively for incorporation into *S. typhimurium* and exclusion from mammalian proteomes.

For a more detailed analysis of the labeled proteomes we also investigated EME incorporation into *S. typhimurium* during infection of macrophages by SDS-PAGE and in-gel fluorescence scanning. For this purpose macrophages were infected with wild-type *S. typhimurium* 14028 and pulse-labeled with 1 mM EME 16 hours post-infection for 1 hour. Since we have learned previously that pulse-labeling in full growth medium can reduce the incorporation of the orthogonal amino acid reporter into the mammalian proteome, we tested EME incorporation in full growth medium as well as Met-depleted growth medium. While we could clearly detect EME incorporation in the presence of *S. typhimurium*, we also observed EME labeling of the mammalian proteome, which was not diminished by the presence of Met in the growth medium. We attribute the discrepancy between the in-gel and the fluorescence microscopy data, to the better dynamic range of the in-gel analysis. Even though EME does most likely not provide sufficient selectivity for the bacterial proteome for all applications (in particular mass spectrometry based proteomics), it provides a starting point for further developments.

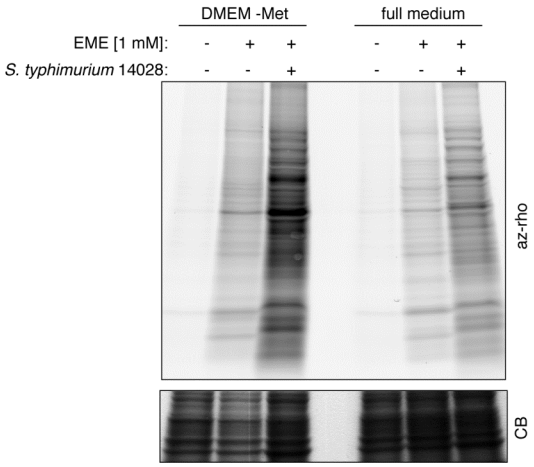


Figure 32 Full medium does not reduce EME incorporation into the mammalian proteome. RAW264.7 macrophages were infected with wild-type *S. typhimurium* and pulse-labeled with EME 16 hours post-infection. Cells were lysed, reacted with az-rho by CuAAC, and analyzed by SDS-PAGE and in-gel fluorescence scanning. az-rho: azido-rhodamine, CB: coomassie blue, DMEM-Met: Dulbecco's Modified Eagle Medium without methionine, full medium: full DMEM with 10% fetal bovine serum.

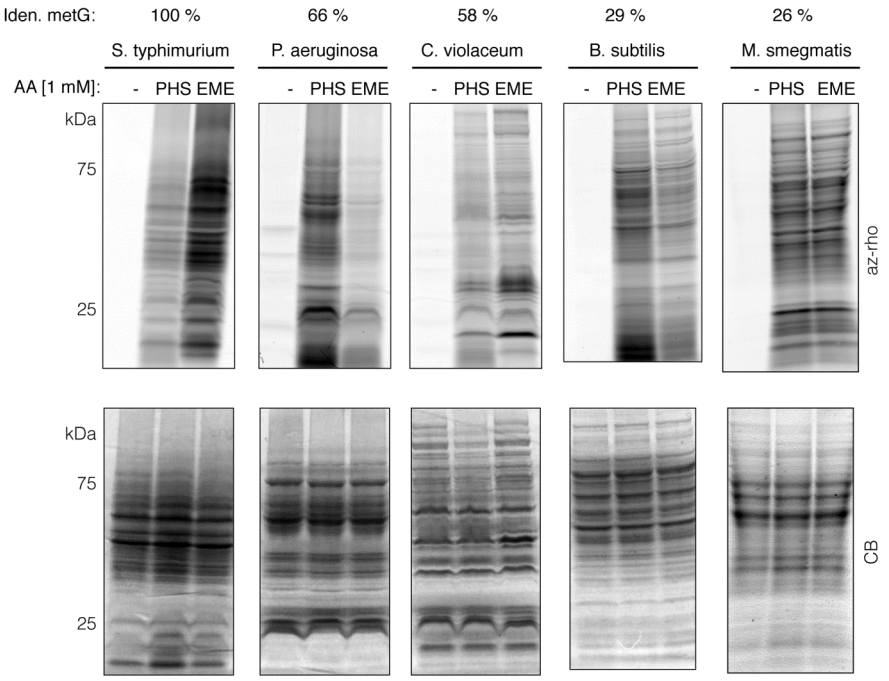


Figure 33 PHS and EME are incorporated into various bacterial species. *S. typhimurium*, *P. aeruginosa, C. violaceum, B. subtilis,* and *M. smegmatis* were pulse-labeled with PHS or EME in M9 minimal medium. Cell were lysed, reacted with az-rho by CuAAC, and analyzed by SDS-PAGE and in-gel fluorescence scanning. Iden. metG: percentage identity between individual metG genes and the metG gene of *S. typhimurium*. az-rho: azido-rhodamine, CB: coomassie blue, AA: amino acid, EME: : (S)-2-amino-4-(prop-2-yn-1-ylthio)butanoic acid, PHS: (S)-2-amino-4-(prop-2-yn-1-yloxy)butanoic acid.

To assess the general possibility to use these kinds of amino acid reporters in other bacterial species, which might benefit more from an intrinsically selective amino acid reporter than *S. typhimurium*, we tested EME and PHS incorporation in a variety of other bacterial species: *P. aeruginosa, C. violaceum, B. subtilis,* and *M. smegmatis*. All of the tested species incorporated EME and PHS, yet none of them as efficiently as *S. typhimurium*. The observed fluorescence intensity (incorporation level) did not scale with the percentage identity of the corresponding metG genes with the *S. typhimurium* metG

gene (*P. aeruginosa* 66 %, *C. violaceum* 58 %, *B. subtilis* 29 %, *M. smegmatis* 26 %). In addition to MetRS activity, differences in amino acid reporter uptake and metabolization could also cause the observed differences in incorporation levels. These data emphasizes the requirement for empiric determination of amino acid reporter acceptance and also the possibility to use amino acid reporters like PHS or EME as orthogonal amino acid reporters.

Although EME was not completely excluded from the mammalian proteome we wanted to experimentally address the observed differences in EME incorporation between *S. typhimurium* and mammalian cells. We hypothesized that the difference is largely due to discrepant MetRS specificities. Over-expression of *E. coli* MetRS might therefore rescue the mammalian EME incorporation deficiency, provided the bacterial enzyme is able to properly interact with the mammalian tRNA. Indeed, over-expression of *E. coli* MetRS in 293T cells improved EME incorporation, while HPG incorporation (a known mammalian and bacterial amino acid reporter for Met) was not elevated to the same extent. Over-expression of *E. coli* MetRS improved EME labeling almost to the same level as the proper mammalian amino acid reporter HPG. From this preliminary experiment we concluded that the reduced EME incorporation rate in mammalian cells could be most likely attributed to differences in MetRS activity, rather than amino acid reporter uptake or metabolization, parameters that should not have been changed by *E. coli* MetRS heterologous expression.

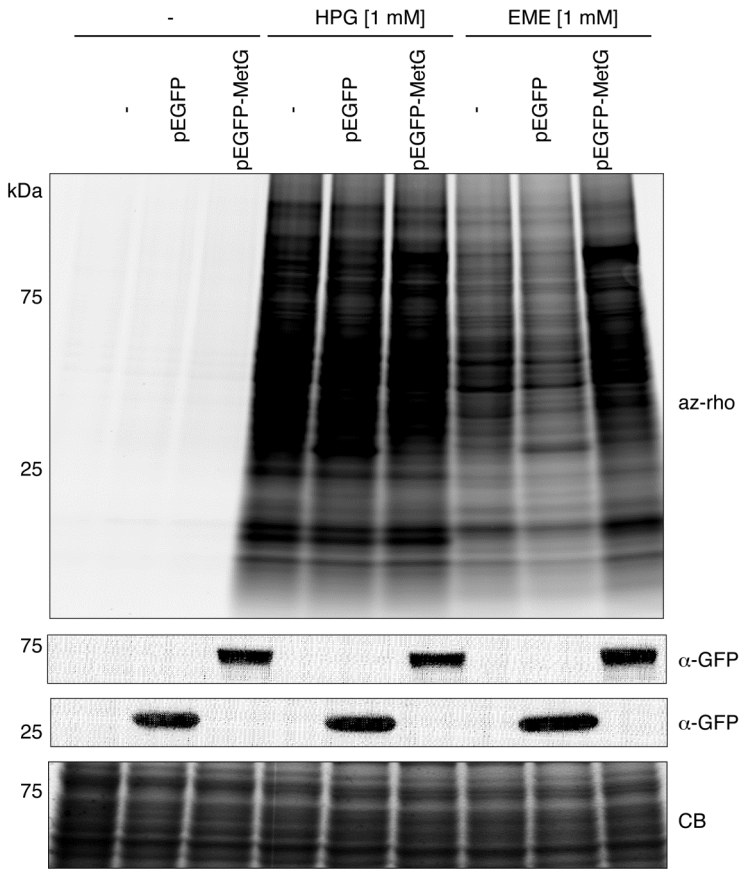


Figure 34 Over-expression bacterial methionyl-tRNA-synthetase (MetG) facilitates increased EME incorporation into the mammalian proteome. 293T cells were transiently transfected with a pEGFP-MetG expression construct generating a MetG-GFP protein fusion and pulse-labeled with HPG or EME. Cells were lysed, reacted with az-rho by CuAAC, and analyzed by SDS-PAGE and in-gel fluorescence scanning. HPG: homopropargylglycine, EME: (S)-2-amino-4-(prop-2-yn-1-ylthio)butanoic acid, α-GFP: anti-GFP Western Blot analysis.

4.4 Materials and methods

4.4.1 Synthesis of (*S*)-2-amino-4-(prop-2-yn-1-yloxy)butanoic acid (PHS)

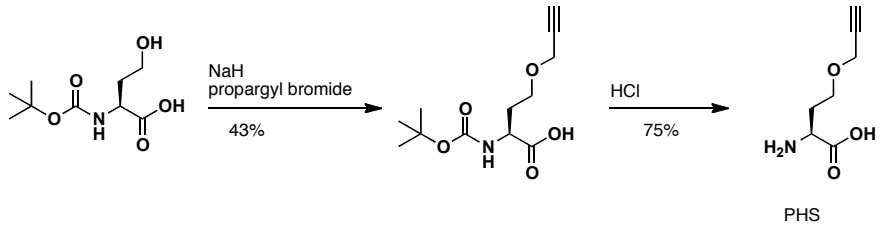


Figure 35 Synthesis of (*S*)-2-amino-4-(prop-2-yn-1-yloxy)butanoic acid (PHS).

A 100 ml flame dried round bottom flask was charged with 25 ml anhydrous dimethylformamide under argon atmosphere. 2-((*tert*-butoxycarbonyl)amino)-4-hydroxybutanoic acid (500 mg, 2.28 mmol) was added to the reaction vessel under stirring and the reaction mix was cooled to 0 °C. Sodium hydride (137 mg, 5.70 mmol) was added to the reaction mix under stirring. After H_2 evolution had ceased propargyl bromide (339 mg, 2.85 mmol) was added under stirring and the reaction was allowed to warm up to room temperature and was stirred for 2 h. The reaction was quenched with water and the solvents were evaporated under reduced pressure. The resulting material was resuspended in water and washed with ethyl acetate and diethylether. The aqueous solution was acidified with sodium bisulfate (1 M) and extracted with 3 portions of ethyl acetate. The organic fractions were pooled and dried over magnesium sulfate, filtered and evaporated under reduced pressure. The resulting material was purified by column chromatography over silica gel using 3% methanol in dichloromethane to yield (S)-2-((tert-butoxycarbonyl)amino)-4-(prop-2-yn-1-yloxy)butanoic acid (43%). 1H NMR (600 MHz, $CDCl_3$): 5.37-5.36 (d, 1H), 4.31 (s, 1H), 4.08-4.01 (m, 2H), 3.58-3.53 (m, 2H), 2.36 (s, 1H), 2.07-1.95 (m, 2H). ^{13}C NMR (600 MHz, $CDCl_3$): 176.45, 155.70, 80.05, 79.3, 74.75, 66.27, 58.17, 51.46, 31.59, 28.29. Theoretical mass: 257.13, mass observed: $[M+Na]^+$ 280.08, $[M+2Na-H]^+$ 302.08.

(S)-2-((tert-butoxycarbonyl)amino)-4-(prop-2-yn-1-yloxy)butanoic acid (271 mg, 1.05 mmol) was dissolved in3 ml concentrated HCl and stirred for 1.5 hours at room temperature. The reaction mix was diluted with 15 ml water and purified over Dowex 50WX-100 to yield the final product (*S*)-2-amino-4-(prop-2-yn-1-yloxy)butanoic acid (PHS) (75%). ^{1}H NMR (600 MHz, D_2O): 4.09 (s, 2H), 3.75-3.73 (m, 1H), 3.67-3.61 (m, 2H), 2.75 (s, 1H), 2.12-1.96 (m, 2H). ^{13}C NMR (600 MHz, D_2O): 173.94, 79.07, 75.76, 66.82, 57.90, 53.44, 29.65. Theoretical mass: 157.0739, mass observed: $[M+H]^+$ 158.2096.

4.4.2 Synthesis of (*S*)-2-amino-4-(prop-2-yn-1-ylthio)butanoic acid (EME)

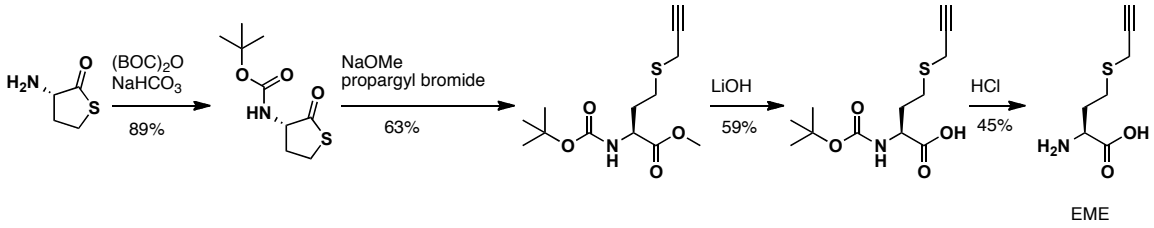


Figure 36 Synthesis of (*S*)-2-amino-4-(prop-2-yn-1-ylthio)butanoic acid (EME).

A flame dried 50 ml round bottom flask equipped with a condenser was charged with 10 ml anhydrous methanol and sodium methoxide (26 mg, 4.83 mmol) was added under argon atmosphere and dissolved under stirring at room temperature. (S)-N-*tert*-butoxycarbonyl-homocysteine thiolactone (100 mg, 4.6 mmol) was added and the reaction mixture was incubated for 10 min at room temperature under stirring. Propargyl bromide (66 mg, 5.52 mmol) was added and the reaction mixture was refluxed for 3 h. The reaction mixture was cooled to ambient temperature and quenched with water. The

solvents were evaporated under reduced pressure. The resulting material was resuspended in ethyl acetate and washed with concentrated sodium bicarbonate, 1 M sodium bisulfate and brine. The organic fraction was dried over magnesium sulfate, filtered and evaporated under reduced pressure. The resulting material was purified by column chromatography over silica gel with 15% ethyl acetate in hexanes to yield (S)-methyl 2-((tert-butoxycarbonyl)amino)-4-(prop-2-yn-1-ylthio)butanoate (63 %). 1H NMR (600 MHz, $CDCl_3$): 5.10-5.09 (m, 1H), 4.42-4.41 (m, 1H), 3.73 (s, 3H), 3.23 (d, 2H), 2.73-2.70 (t, 2H) 2.21 (s, 1H), 2.15-2.12 (m, 1H), 1.95-1.91 (m, 1H), 1.42 (s, 9H). ^{13}C NMR (600 MHz, $CDCl_3$): 172.71, 155.31, 80.11, 79.57, 71.30, 52.66, 52.49, 32.14, 29.72, 28.33, 27.21, 19.10.

(S)-methyl 2-((tert-butoxycarbonyl)amino)-4-(prop-2-yn-1-ylthio)butanoate (83 mg) was dissolved in 8 ml water and acetonitrile (1:1) and stirred at room temperature. Lithium hydroxide (6.92 mg) was dissolved in water and acetonitrile (1:1) and slowly added to the stirred reaction mixture. The reaction mixture was stirred at room temperature for 90 min. The organic solvent was evaporated under reduced pressure and the remaining aqueous phase was acidified with 1 M sodium bisulfate. The aqueous phase was extracted three times with ethyl acetate and the organic fractions were pooled, dried over magnesium sulfate, filtered and evaporated under reduced pressure. The resulting material was purified by column chromatography over silica gel with 5% methanol in dichloromethane to 10% methanol in dichloromethane to yield (S)-2-((tert-butoxycarbonyl)amino)-4-(prop-2-yn-1-ylthio)butanoic acid (59%). ^{1}H NMR (600 MHz, CD_3OD): 4.20-4.13 (m, 1H), 3.28 (m, 2H), 2.76-2.67 (m, 2H), 2.56 (s, 1H), 2.10-1.88 (m, 2H), 1.42 (s, 9H). ^{13}C

NMR (600 MHz, CD_3OD): 171.62, 156.75, 79.45, 79.11, 70.94, 60.19, 52.76, 31.02, 27.37, 27.25, 17.96, 13.13. Theoretical mass: 273.10, observed mass: $[M+Na]^+$ 296.08, $[M+2Na-H]^+$ 318.08.

(S)-2-((tert-butoxycarbonyl)amino)-4-(prop-2-yn-1-ylthio)butanoic acid was dissolved in 3 ml concentrated HCl and stirred for 1.5 hours at room temperature. The reaction mix was diluted with 15 ml water and purified over Dowex 50WX-100 to yield the final product (*S*)-2-amino-4-(prop-2-yn-1-ylthio)butanoic acid (EME) (45%). ^1H NMR (600 MHz, D_2O): 3.76-3.73 (t, 1H, J = 6.24), 3.44-3.37 (m, 2H), 2.87-2.84 (m, 2H), 2.69-2.68 (m, 1H), 2.32-2.15 (m, 2H). ^{13}C NMR (600 MHz, D_2O): 174.06, 80.34, 72.18, 53.82, 29.84, 26.52, 18.03. Theoretical mass: 173.0510, observed mass: $[M+H]^+$ 174.1730.

4.4.3 Synthesis of (*R*)-2-amino-3-(but-3-yn-1-ylthio)propanoic acid (BCY)

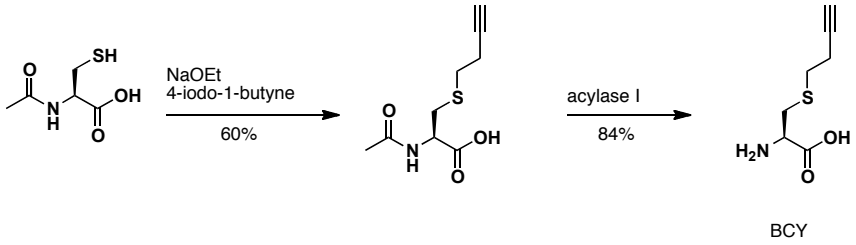


Figure 37 Synthesis of (*R*)-2-amino-3-(but-3-yn-1-ylthio)propanoic acid (BCY).

A 50 ml flame dried round bottom flask equipped with a condenser was charged with 12 ml anhydrous ethanol under argon atmosphere. N-acetyl-cysteine (212 mg, 1.30 mmol) was added to the reaction vessel under stirring at room temperature. Sodium ethoxide (195 mg, 2.86 mmol) was added to the reaction mix from a 21 wt. % solution of sodium ethoxide in ethanol and the reaction mix was stirred for 10 min at room temperature. Iodo

butyne (281 mg, 1.56 mmol) was added and the reaction mix was refluxed for 3 h. Since the starting material was not completely consumed (TLC) another 0.6 eq. of iodobutyne was added and the reaction mix was refluxed for 1h, after which starting material was completely consumed (TLC). The reaction was quenched with water and the solvents were evaporated under reduced pressure. The resulting material was purified by column chromatography over silica with 10% methanol, 0.5% acetic acid in ethyl acetate to yield (R)-2-acetamido-3-(but-3-yn-1-ylthio)propanoic acid (60%). 1H NMR (600 MHz, CD3OD): 4.61-4.59 (m, 1H), 3.11-3.08 (dd, 1H, 4.8, 13.8), 2.91-2.87 (dd, 1H, 7.8, 13.8), 2.73-2.71 (t, 2H), 2.48-2.45 (m, 2H), 2.30 (s, 1H), 2.01 (s, 3H). 13C NMR (600 MHz, CD3OD): 170.73, 170.36, 80.35, 67.76, 50.74, 50.74, 31.57, 29.30, 19.45, 17.65. Theoretical mass: 215.06, observed mass: $[M+Na]^+$ 238.00, $[M+2Na-H]^+$ 260.00.

(R)-2-acetamido-3-(but-3-yn-1-ylthio)propanoic acid (127 mg) was dissolved in 7 ml water and the pH was adjusted to 7.8 pH with 1 M NH_4OH and 1 M HCl. 1.5 mg of acylase I was added and the reaction mix was carefully stirred over night at 37 °C. The reaction mix was cooled to ambient temperature and 15 mg activated charcoal was added. The reaction mix was filtered through a glasswool, sand, Celite plug, which was rinsed with water. The aqueous solution was extracted with ethyl acetate and the observed precipitate in the aqueous layer was removed by filtration. The aqueous solvent was evaporated at 40 °C under reduced pressure and the resulting material was dissolved in 1 M HCl and purified over Dowex 50WX-100 to yield the final product (*R*)-2-amino-3-(but-3-yn-1-ylthio)propanoic acid (BCY) (84%). ^1H NMR (600 MHz, D_2O): 3.99-3.97 (m, 1H), 3.04-3.01 (dd, 1H, 4.2, 15), 2.91-2.88 (dd, 1H, 7.2, 15), 2.56-2.54 (m, 2H), 2.33-

2.31 (m, 2H), 2.17 (s, 1H). ^{13}C NMR (600 MHz, D_2O): 171.11, 83.16, 70.43, 52.49, 31.29, 30.08, 18.66. Theoretical mass: 173.0510, observed mass $[M+H]^+$ 174.2300.

4.4.4 Synthesis of (*R*)-2-amino-3-(prop-2-yn-1-ylthio)propanoic acid (PCY)

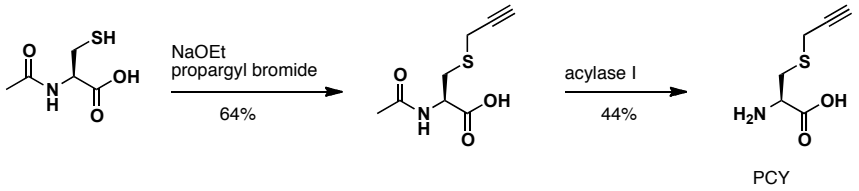


Figure 38 Synthesis of (R)-2-amino-3-(prop-2-yn-1-ylthio)propanoic acid (PCY).

A 50 ml flame dried round bottom flask was charged with 6 ml anhydrous ethanol under argon atmosphere. N-acetyl-cysteine (300 mg, 1 eq.) was added and dissolved and sodium ethoxide in ethanol (21%) (273 mg, 2.2 eq.) was added at room temperature. The reaction mix was incubated for 10 min at room temperature. Propargyl bromide (262 mg, 1.2 eq.) was added and the reaction mix was refluxed for 2.5 hours. The reaction mix was cooled to ambient temperature and quenched with water. The organic solvent was evaporated under reduced pressure and the reaction mix was acidified with 1 M HCl. The reaction mix was extracted 3 times with ethyl acetate and the pooled organic fractions were dried over magnesium sulfate and evaporated under reduced pressure. The resulting material was purified by column chromatography on silica gel with 5% methanol in dichloro methane and 0.5% acetic acid to yield the final product (R)-2-acetamido-3-(prop-2-yn-1-ylthio)propanoic acid (64%). ^1H NMR (600 MHz, CD_3OD): 6.78-6.71 (m, 1H), 4.82-4.80 (m, 1H), 3.32-3.29 (m, 1H), 3.26-3.22 (m, 2H), 3.11-3.08 (m, 1H), 2.30

(m, 1H), 2.06 (s, 1H). ^{13}C NMR (600 MHz, CD_3OD): 172.32, 171.99, 79.04, 71.53, 51.76, 32.51, 21.10, 18.60. Theoretical mass: 201.05, observed mass: $[M+H]^+$ 202.10.

(R)-2-amino-3-(prop-2-yn-1-ylthio)propanoic acid (PCY) was prepared from (R)-2-acetamido-3-(prop-2-yn-1-ylthio)propanoic acid in the same way as BCY (44 %). ^{1}H NMR (600 MHz, D_2O): 3.93-3.91 (dd, J = 4.2, 3.6, 1H), 3.30-3.29 (m, 1H), 2.59 (m, 1H). ^{13}C NMR (600 MHz, D_2O): 172.59, 79.98, 72.62, 53.27, 31.79, 18.76.

4.4.5 Amino acid reporter labeling of bacteria in liquid culture

See 2.4.6

4.4.6 Infection of RAW264.7 macrophages and pulse-labeling with amino acid reporter

See 2.4.5

4.4.7 Analysis of amino acid reporter labeling by fluorescence microscopy

See 3.4.6

4.4.8 Analysis of amino acid reporter labeling by SDS-PAGE and in-gel florescence scanning

See 3.4.5

4.4.9 Transfection of HEK293T cells with bacterial methionyl-tRNA synthetase and pulse-labeling with EME

E. coli MetG (methionyl-tRNA-synthetase) was cloned into pEGFP-N1 (pEGFP-MetG) by amplifying metG from pWSK29_MetG with the primers: GAGCTCATGACTCA AGTCGCGAAGAAAATTCTGG and GGTACCAGTTTCACCTGATGACCCGGTTT AGC, by PCR and cloning the resulting fragment into pEGFP-N1 with SacI and KpnI.

This cloning strategy, however, introduced a frame shift in the GFP protein fusion, which was corrected by site-directed mutagenesis using the QuikChange Site-Directed Mutagenesis Kit and the primers: CTAAACCGGGTCATCAGGTGAAA TGGTACCGCGGGC and GCCCGCGGTACCATTTCACCTGATGACCCGGTTTAG.

pEGFP-N1 was transfected into 293T cells using the X-tremeGENE (Roche) transfection reagent, following the supplied protocol with a ratio of 2:1 between DNA and transfection reagent. MetG expression was analyzed by epifluorescence visualization of the GFP fluorescence and by Western Blot analysis (Living Colors anti-GFP antibody JL-8, Clonetech).

Transfected cells were pulse-labeled with 1 mM EME or 1 mM HPG 24 hours post-transfection in fresh Dulbecco's Modified Eagle Medium (DMEM) growth medium without methionine.

5 Exploring orthogonal amino acid reporter labeling of *S. typhimurium* effector proteins

5.1 Abstract

Intracellular pathogens are able to live and replicate within a mammalian cell without clearance of the hosts immune system. A very specialized set of virulence factors enables them to survive in this environment and escape from bactericidal mechanisms the host cell imposes on them. Various virulence factors are secreted into the host cytosol (effector proteins) were they modify host functions and pathways. Both the identification of these factors and the analysis of their function are central to the understanding of the pathogen-host-interaction. Previous proteomic approaches to detect effector proteins have been extremely difficult, due to the complexity of the mixed pathogen host proteome and low abundance of secreted bacterial effector proteins. This section discusses our attempts to use orthogonal amino acid reporters for the proteomic detection of *S. typhimurium* effector proteins during infection.

5.2 Introduction

A highly specialized subset of symbiotic or pathogenic bacteria has evolved to invade and replicate within different host cell types. These bacteria are taken up either by phagocytic cell populations or actively invade non-phagocytic cells, such as epithelial cells. The first step in both processes - adherence to the plasma membrane - is mediated by various adherence structures. These structures comprise a plethora of bacterial proteins, ranging from rather simple structured adhesins, like fibronectin binding proteins, to multimeric complexes such as pili and fimbria (Pizarro-Cerdá & Cossart 2006). In the second step, the actual invagination takes place, which can be either

achieved by interaction with cognate membrane bound host receptors, or by remodeling mammalian signaling pathways from within the cell. The latter requires the concerted expression, activation, and delivery of effector proteins into the mammalian host cell by means of protein secretion systems (Zhou & Galán 2001; DiGiuseppe Champion & J. S. Cox 2007; Coburn et al. 2007). Once in the mammalian cytosol, these effector proteins can modulate crucial host pathways - often by mimicking mammalian protein activities - leading to the uptake of the bacterial cell (Bhavsar et al. 2007). Upon cell entry the bacterium finds itself encompassed by vesicular membrane structures, derived from the host membranes. Bacteria such as *S. typhymurium* or *M. tuberculosis* remain within these vesicular structures, replicate, and remodel them to their benefit. Other bacteria however, such as *L. monocytogenes*, egress from these compartments and adopt a cytosolic lifestyle. In both cases, the bacterium will adjust to the given environment as well as modulate it to guarantee its survival (Bhavsar et al. 2007; Coombes et al. 2004).

Each step during the above outlined infection process requires the expression and activity of specific bacterial proteins - concerted in time and space. Indeed bacterial protein expression changes upon superficial contact with mammalian host cells, e.g. in *S. typhimurium* (B B Finlay et al. 1989). The application of microarray technologies to dissect changes in gene expression during various stages of the intracellular life cycle bears valuable information about necessary virulence factors (Coombes et al. 2004; Rohde et al. 2007). However, analysis of gene expression on the transcript level has its inherent limitations, especially if taken as a surrogate for protein levels, since protein levels and activity – the general effectors of biological function – are determined by many other parameters in addition to mRNA levels (J. Cox & Mann 2007). Furthermore

the analysis of transcript level does not provide information about protein secretion activity.

Recent data in *S. typhimurium* emphasizes the role of posttranscriptional regulation in bacteria and shows poor correlation of transcript and protein levels (Ansong et al. 2009). The advent of modern mass spectrometry based proteomic technologies gave biological researcher excess to an immediate representation of proteome composition and hence cellular function (Aebersold & Mann 2003). Several studies explored the proteomic composition of pathogens under established culture conditions (Rosenkrands et al. 2000; C. G. Zhang et al. 2005). While these studies give a general list of bacterial proteins, they reveal only limited information about necessary virulence factors and dynamic proteome changes. Therefore, several *in vivo* and *in vitro* studies aimed at the analysis of the bacterial proteome under relevant infection conditions. In traditional *in vitro* experiments bacteria are cultivated under conditions that mimic intracellular conditions, such as high pH and low magnesium, which are known to induce the expression and secretion of certain virulence factors. This approach has been widely used for the identification of secreted effector proteins (Kodama et al. 2008; Yahr et al. 1997). Even though these experiments contribute to the understanding of bacterial virulence they do not truly reflect the native situation within the host cell. More elaborate experiments have tried to measure bacterial protein expression in cell culture infection experiments or *in vivo*. A recent study investigated *S. typhimurium* proteome changes upon infection of murine macrophages either expressing or lacking the host resistance regulator Nramp1 (natural resistance-associated macrophage protein 1) (Shi et al. 2006). Another study used GFP-expressing *S. typhimurium* to infect mice (Becker et al. 2006). This allowed for

flow cytometric sorting of *S. typhimurium* containing cells isolated from organs of infected mice and subsequent proteomic analysis. The overwhelming amount of host proteins impedes the analysis of less abundant bacterial proteins (Rodland et al. 2008) and none of the above-mentioned studies was able to detect secreted effector proteins in the context of the host cell.

Salmonellae are Gram-negative enterobacteria with a broad host range, in which they can cause either systemic disease in the form of enteric (typhoid) fever, or gastroenteritis (Ohl & S I Miller 2001). *S. typhimurium* in particular infects cattle, pigs, chickens, and humans. In these hosts it usually causes a self-limiting gastroenteritis. However, in certain mouse strains it causes a typhoid fever-like systemic disease, the susceptibility to which is greatly determined by Nramp1 (Jones & Falkow 1996; Plant & Glynn 1979). This mouse model has been widely used for the analysis of systemic *Salmonella* infection (Ohl & S I Miller 2001; Fields et al. 1986).

Salmonella spp. possesses a number of virulence factors, such as virulence plasmids, toxins, fimbriae, and importantly *Salmonella* pathogenicity islands (SPI) (van Asten & van Dijk 2005). SPIs are thought to be acquired through horizontal gene transfer between species, indicated by an anomalous GC content and the presence of fragments of transposable elements (Groisman & Ochman 1996). Even though more than five SPIs have been described in the literature, the focus has been greatly on SPI-1 and SPI-2 (Hansen-Wester & M Hensel 2001; Schlumberger & Hardt 2006). Both SPIs encode, amongst other proteins, for a type III secretion system (TTSS), which is a multi-component secretion machinery for the purpose of protein effector delivery into the host cell cytosol (Mueller et al. 2008). So far, more than 30 effector proteins have been

identified, whose secretion depends on the activity of these two SPI encoded secretion systems (McGhie et al. 2009). Conventionally, SPI-1 effectors and SPI-2 effectors are associated with very distinct functions. SPI-1 effectors are necessary for the invasion of non-phagocytic cells and the establishment of the *Salmonella* containing vacuole (SCV), which is a vesicular structure that harbors *S. typhimurium* upon cellular invasion. In the typhoid fever model SPI-1 TTSS (TTSS1) deficient *S. typhimurium* show attenuated infectivity when administered orally, but not when given intravenously (Ohl & S I Miller 2001). Furthermore, TTSS1 is implicated in the molecular etiology of *S. typhimurium* induced enteritis. These results indicate the importance of TTSS1 in the interaction and penetration of the intestinal epithelium.

Later during the infection process SPI-2 effectors promote intracellular replication and systemic dissemination. It has been shown that SPI-2 TTSS (TTSS2) is necessary for systemic infection in the mouse model and for replication in tissue culture. In epithelial cells and macrophages, TTSS2 is responsible for the formation of membrane structures called *Salmonella*-induced filaments (Sif) – Sif protrude from SCVs and are possibly involved in vesicular trafficking (Haraga et al. 2008). However, contrary to the current model, recent data suggest an unappreciated interaction between SPI-1 and SPI-2 effectors, as classical SPI-1 effectors like SopB have been implicated in late SCV positioning functions (Wasylnka et al. 2008). Despite the above-mentioned focus on TTSS1/2, other secretion systems have been discovered in *S. typhimurium* and have been implicated in its virulence. One of them is the sci-encoded (*Salmonella enterica* centisome 7 genomic island) type VI secretion system, which when ablated attenuates infectivity in various model systems (Filloux et al. 2008). The list of secreted effector

proteins is constantly growing and it is likely that more effector proteins are to be discovered, especially considering the vast number of non-annotated ORFs. Even tough some progress has been made in predicting bacterial effector proteins using bioinformatics (Y. Wang et al. 2011; McDermott et al. 2011), it is still desirable to develop a robust proteomics strategy that would allow global analysis of effector production and secretion in the infected host, since some effectors might always elude the bioinformatics prediction and an empiric determination and analysis of the dynamics of this concerted process is invaluable for its complete understanding. All previous proteomic studies of effector proteins or bacterial infection have failed to achieve this goal.

The orthogonal amino acid reporters presented in the preceding chapters (AOA, ANL, PEP) could help to isolate secreted bacterial effector proteins from the mammalian cytosol or membrane fraction and therefore could allow the proteomic detection of secreted bacterial effector proteins. In the following we present our efforts to use orthogonal amino acid reporters for this purpose and describe the current limitations of the orthogonal amino acid system.

5.3 Results

Initially we wanted to test whether the orthogonal amino acids AOA and ANL can be generally incorporated into secreted bacterial effector proteins. To address this question we investigated the incorporation of AOA and ANL into secreted bacterial effector proteins in liquid culture. *S. typhimurium* SPI-1 or SPI-2 dependent protein secretion can be induced in liquid culture (Ellermeier & Slauch 2003; Yu et al. 2010). *S. typhimurium* pWSK29_MetG$_{PLL}$ IR715 was grown in full LB medium over night and

diluted into fresh LB medium. After 2 or 3 hours of growth 1 mM ANL, AOA, or Met was added to the growth medium until 4 hours of growth were completed. The culture supernatant was separated from the bacteria and the containing proteins were precipitated with trichloroacetic acid. The precipitated proteins were reacted with az-rho or alk-rho by CuAAC and analyzed by SDS-PAGE and in-gel fluorescence analysis (Figure 39). AOA and ANL were incorporated into the secreted effector proteins (identity of these proteins was determined under similar conditions by other members of the Hang laboratory). Interestingly, earlier addition of the orthogonal amino acid reporter (2 h) impaired protein secretion in comparison to the methionine control and the 1 h sample. This speaks for a detrimental effect on protein secretion upon longer incubation times with AOA or ANL. Both amino acid reporters showed this effect. The smaller effect observed by fluorescence could be due to an increased labeling intensity in the 2 h sample compared to the 1 h sample, which is able to balance out some of the signal loss due to impaired protein secretion. Repeating the same experiment with *S. typhimurium* 14028 pWSK29_MetG$_{PLL}$ and *S. typhimurium* 14028 ΔSPI-1 pWSK29_MetG$_{PLL}$, a TTSS-1 inactive mutant, pulse-labeling for 1 hour, confirmed that most of the labeled protein species indeed are secreted SPI-1 effector proteins (Figure 40). So, generally AOA or ANL-labeled effector proteins can be secreted through the TTSS.

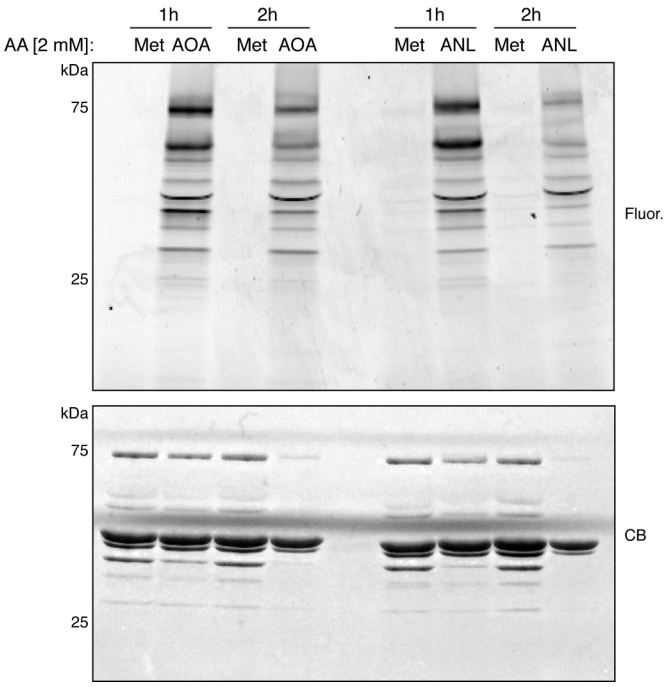


Figure 39 AOA and ANL are incorporated into secreted *S. typhimurium* proteome. SPI-1 dependent protein secretion of *S. typhimurium* IR715 pWSK_MetG$_{PLL}$ was induced in full LB growth medium. *S. typhimurium* IR715 pWSK_MetG$_{PLL}$ was pulse-labeled with AOA, ANL, or Met and the proteins in the culture supernatant were reacted with az-rho or alk-rho by CuAAC and analyzed by SDS-PAGE and in-gel fluorescence. Fluor.: az-rho or alk-rho fluorescence, CB: coomassie blue, AA: amino acid, az-rho:azido-rhodamine, alk-rho: alkyne-rhodamine.

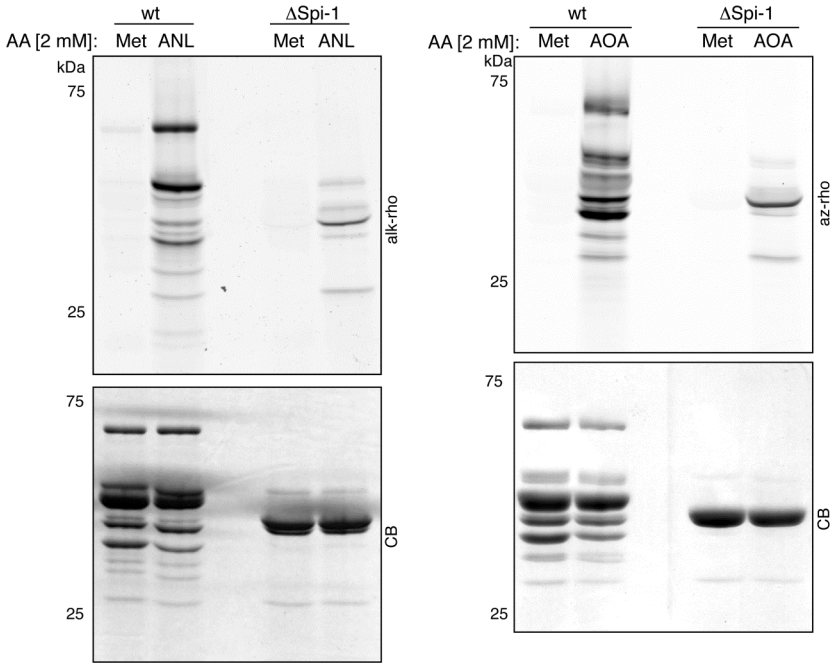


Figure 40 AOA and ANL are incorporated into SPI-1 effector proteins. SPI-1 dependent protein secretion of *S. typhimurium* IR715 pWSK_MetG$_{PLL}$ and *S. typhimurium* IR715 ΔSPI-1 pWSK_MetG$_{PLL}$ was induced in full LB growth medium. *S. typhimurium* was pulse-labeled with AOA, ANL, or Met and the proteins in the culture supernatant were reacted with az-rho or alk-rho by CuAAC and analyzed by SDS-PAGE and in-gel fluorescence. Fluor.: az-rho or alk-rho fluorescence, CB: coomassie blue, AA: amino acid, az-rho:azido-rhodamine, alk-rho: alkyne-rhodamine.

For previous proteomics experiments (see chapter 2) we pulse-labeled infected macrophages around 16 hours post-infection. Before moving into large-scale proteomics experiments for the detection of secreted bacterial effector proteins we wanted to assess whether known bacterial effector proteins are actually expressed at this time during infection. Eriksson et al published a microarray study profiling *S. typhimurium* gene expression during infection of macrophages and compared it to *S. typhimurium* grown in mammalian growth medium (Eriksson et al. 2003). Their data revealed that the transcription of numerous known effector proteins, amongst them SseJ and SifA, is up-regulated in comparison to the control population 12 hours post-infection (Figure 41).

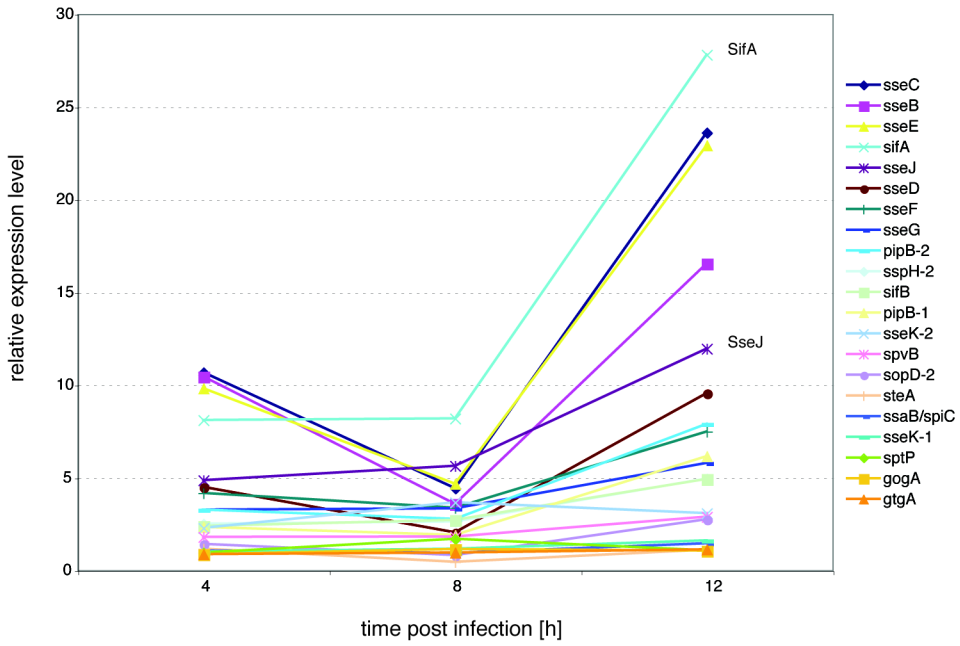


Figure 41 Relative expression level of selected known effector proteins in S. typhimurium during infection of macrophages. Relative expression level: relative expression compared to opsonized *S. typhimurium* in mammalian growth medium. (Data from Eriksson et al 2003).

SseJ, one of the most strongly up-regulated effector proteins in the microarray study by Eriksson and colleagues, is known to take part in the formation of *Salmonella*-induced filaments (Sifs) and as a SPI-2 regulated effector proteins plays an important role in the establishment of the intracellular infection (Lossi et al. 2008). We analyzed its expression dynamics during infection of RAW264.7 macrophages by fluorescence microscopy. Macrophages were infected with *S. typhimurium* IR715 pWSK29_SseJ, expressing a HA-tagged SseJ, regulated by the endogenous SseJ promoter. SseJ cannot be detected 2 hours post-infection, but can be detected after 7 hours and is most strongly expressed at the 20-hour time point.

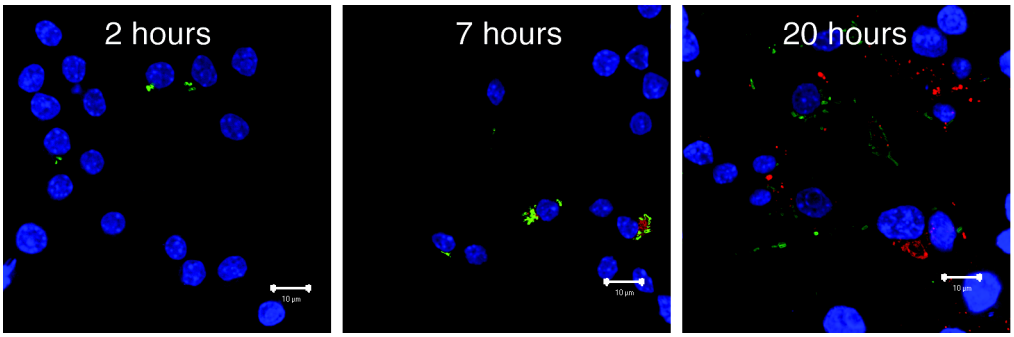


Figure 42 SseJ is expressed during infection of RAW264.7 macrophages. RAW264.7 cells were infected with S. typhimurium pWSK29_HASseJ and analyzed by fluorescence microscopy. Blue: DNA stain TO-PRO-3, green: anti-*Salmonella* antibody stain, red: anti-HA antibody stain. Bar indicated 10 μm.

As a final step towards establishing orthogonal amino acid incorporation into bacterial effector proteins during infection, in preparation of large-scale proteomics experiments, we analyzed the incorporation of the amino acid reporter EME (see chapter 4) into SseJ during infection of macrophages. Macrophages were infected with *S. typhimurium* IR715 pWSK29_SseJ and pulse labeled with EME 16 hours post-infection.

SseJ was immunoprecipitated, reacted with az-rho by CuAAC, and analyzed by SDS-PAGE and in-gel fluorescence scanning. EME was incorporated into SseJ during infection of mammalian cells (Figure 43). Based on the mild Brij detergent lysis we concluded that the immunoprecipitated SseJ population should be largely of mammalian cytosolic origin, hence secreted. These data demonstrate that amino acid reporters can be incorporated into SPI-1 and SPI-2 effector proteins in *S. typhimurium* and that this incorporation does not preclude effector secretion. Furthermore, known effector proteins, such as SseJ, are expressed between 7 and 20 hours of infection of macrophages, the time window we have been conducting our proteomics studies.

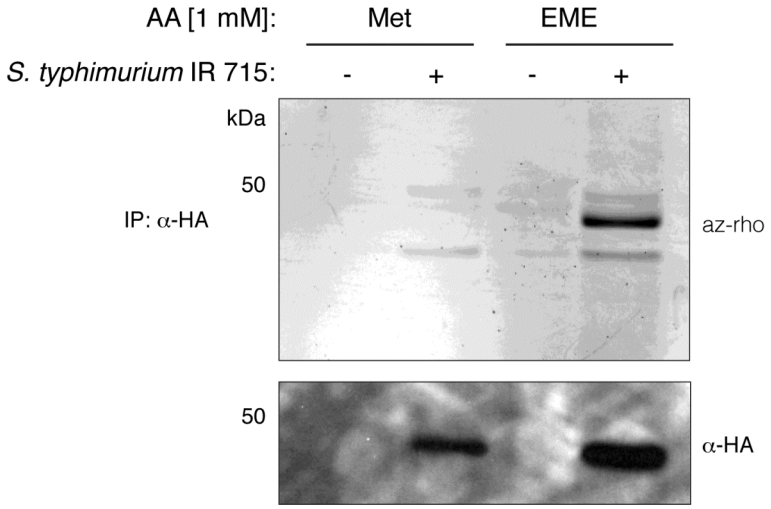


Figure 43 SseJ is labeled with the amino acid reporter EME during infection of RAW264.7 macrophages.

The identification of secreted effector proteins by mass spectrometry-based proteomics requires the separation of intact bacterial cells from the mammalian proteome. The mammalian cytosolic, membrane, and even nucleic fraction should contain any secreted bacterial effector proteins. It is therefore indispensable to lyse the infected

mammalian cell without also lysing the containing bacterial cell bodies since any contamination of the mammalian fraction with non-secreted bacterial proteins impedes the assignment of an identified bacterial protein to the secreted proteome fraction.

The literature is filled with a wide range of different selective lysis conditions that have been used for the analysis of secreted effector proteins. Initially we tested a standard fraction fractionation procedure based on an initial lysis with 0.2% saponin, followed by a 0.1% Triton X-100 lysis (Kuhle & Michael Hensel 2002). RAW264.7 cells were infected with *S. typhimurium* pWSK29_$MetG_{PLL}$ and pulse-labeled with AOA 16 hours post-infection for 1, 2, or 3 hours. Subsequently the infected cells were fractionated with 0.2% saponin and 0.1% Triton X-100 into a cytosolic (0.2% saponin), a membrane (0.1% Triton X-100), and a pellet fraction, which should contain the nucleus and the intact bacteria. The resulting proteome fractions were reacted with az-rho by CuAAC and analyzed by SDS-PAGE, in-gel fluorescence scanning, and Western blot. The S-tag epitope tag of $MetG_{PLL}$ expressed only in the bacterial cells served as a bacterial cytosolic marker. Fluorescence signal was detected in all three fractions. The mammalian cytosolic fraction differed qualitatively from the membrane and the pellet fraction, while the latter two were highly similar. $MetG_{PLL}$ was detected in all three fractions, indicating that a major portion of bacterial cells is lysed during cell lyses or prior to that, during infection. This data indicated that these lysis conditions were not selective enough and resulted in bacterial contaminations in the mammalian cytosol, provided post-labeling cell lysis was the reason for the contamination.

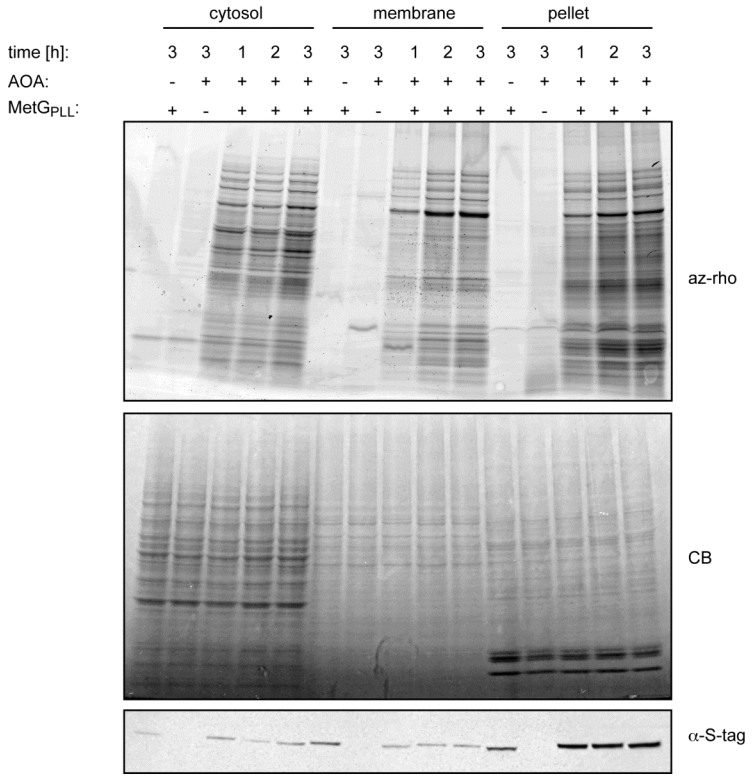


Figure 44 AOA can be detected in various proteome fractions. RAW264.7 cells were infected with *S. typhimurium* pWSK29_MetGPLL and pulse-labeled with 1 mM AOA for different times 16 hours post-infection. Cells were fractionated with 0.2% saponin (cytosol) and 0.1% Triton X-100 (membrane). az-rho: azido-rhodamine, CB: coomassie blue,

Due to the high background lysis observed in the first fractionation experiment (Figure 44), we investigated various lysis procedures for their selectivity for mammalian cells. Blotting for MetG in the isolated mammalian proteome fraction served as the read-out for bacterial co-lysis. Infected RAW264.7 cell were infected with *S. typhimurium* pWSK29_MetG$_{PLL}$ and lysed by different procedures 16 hours post-infection. Cells were either lysed by mechanical disruption in sucrose and imidazole (N1) or triethanolamine (N2) or by detergents (Triton X-100, saponin). Western blot analysis of MetG$_{PLL}$ as a

marker for bacterial integrity revealed the mechanical lysis (N1) as the most selective, while the traditional 0.1% Triton X-100 lysis seemed to be the least selective (Figure 45).

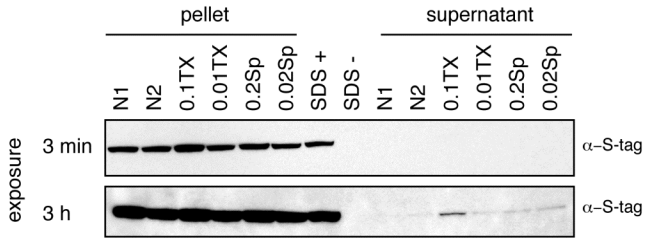


Figure 45 Analysis if differential lysis procedures. N1: mechanical lysis in sucrose and imidazol, N2: mechanical lysis in triethanolamine, 0.1TX: 0.1% Triton X-100, 0.01TX: 0.01% Triton X-100, 0.2Sp: 0.2% saponin, 0.02Sp: 0.02% saponin, α-S-tag: anti-S-tag ($MetG_{PLL}$) Western blot analysis.

Based on these results we conducted an infection and fractionation experiment. RAW264.7 cells were infected with *S. typhimurium* pWSK29_$MetG_{PLL}$ and pulse-labeled with AOA 16 hours post-infection for 2 hours. Subsequently the cells were lyzed by method N1, 0.01% Triton X-100, or 0.02% saponin. The resulting fractions were analyzed by CuAAC with az-rho, SDS-PAGE, and in-gel-fluorescence scanning as well as Western blot. Compared to the precious experiment (Figure 44) the improved selective fractionation conditions caused less bacterial lysis, since no $MetG_{PLL}$ signal was detected in the supernatant fractions (mammalian cytosol). In addition, the total fluorescence intensity in the supernatant fraction seemed to be reduced, indicating reduced levels of bacterial proteins in these fractions. This might speak for a reduced level of bacterial contaminants.

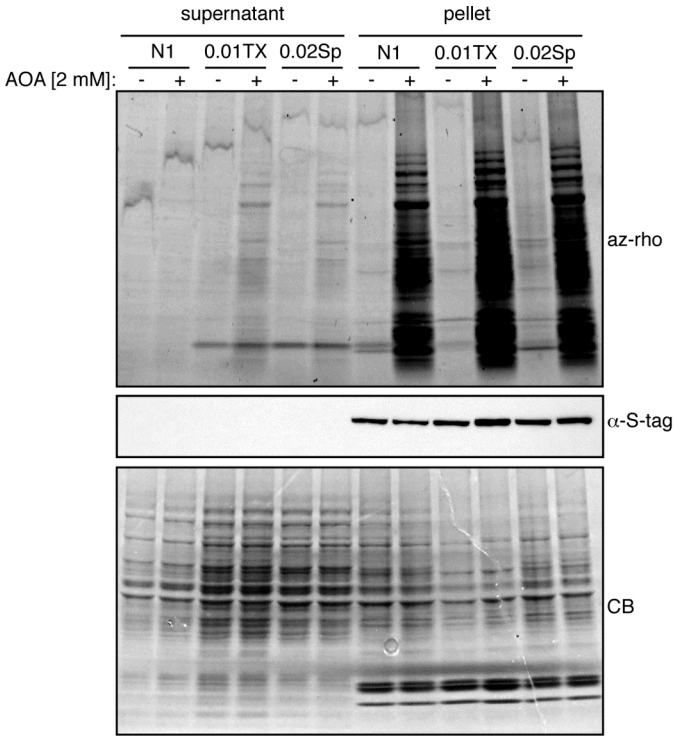


Figure 46 Improved fractionation conditions. RAW264.7 cells were infected with S. typhimurium pWSK29_MetGPLL and pulse-labeled with AOA for 2 hours 16 hours post-infection. N1: mechanical lysis in sucrose and imidazol, 0.01TX: 0.01% Triton X-100, 0.02Sp: 0.02 % saponin, az-rho: azido-rhodamine, CB: coomassie blue.

With these fractionation conditions in hand we decided to approach the first large-scale proteomics experiment. For the initial experiment we decided for the most selective fractionation condition N1, even tough this condition also provides the least fluorescence signal, which suggested very little protein retrieval by affinity enrichment of AOA-labeled proteins. Since the expected protein amount was very small we choose to use an in-solution digest protocol to minimize protein loss during the sample preparation for mass spectrometry (Wiśniewski et al. 2009). RAW264.7 cells were infected with *S. typhimurium* pWSK29_MetG$_{PLL}$ and pulse-labeled with AOA 16 hours post-infection for 2 hours. Cells were washed, harvested, and frozen in liquid nitrogen. Cells were lysed

and fractionated by the N1 procedure and the resulting protein fractions were precipitated by methanol/chloroform precipitation and resuspended in SDS buffer for CuAAC with azide-azo-biotin. Clicked protein samples were precipitated again and bound to streptavidin beads. The protein bound beads were washed and finally the selectively bound proteins were eluted with sodium dithionite. Protein reduction, alkylation, and digestion with Lys-C and trypsin was carried out on Microcon YM-10 filters (Wiśniewski et al. 2009). Resulting peptides were purified on StageTips and analyzed by LC-MS/MS.

Table 2 Identified proteins in the N1 supernatant fraction by Filter Aided Sample Preparation (FASP).

#	Identified Protein	Accession Number
1	mouse	IPI:IPI00319992.1
2	mouse	IPI:IPI00407130.4
3	mouse	IPI:IPI00886297.1
4	60 kDa chaperonin	P0A1D3|CH60_SALTY
5	Elongation factor G	P0A1H3|EFG_SALTY
6	Elongation factor Tu	P0A1H5|EFTU_SALTY
7	Spermidine/putrescine-binding periplasmic prot	P0A2C7|POTD_SALTY
8	Chaperone protein htpG	P58480|HTPG_SALTY
9	Enolase	P64076|ENO_SALTY
10	Chaperone protein dnaK	Q56073|DNAK_SALTY
11	Chaperone protein clpB	Q7CQ01|CLPB_SALTY
12	30S ribosomal subunit protein S1	Q7CQT9|Q7CQT9_SALTY
13	Pyruvate formate lyase I, induced anaerobicall	Q7CQU1|Q7CQU1_SALTY

Not surprisingly, given the low fluorescence intensity in the N1 supernatant fraction (Figure 46), the total number of proteins was very low (Table 2). Three proteins were mouse proteins, the remaining ten were *S. typhimurium* proteins, presenting an overall enrichment of *S. typhimurium* proteins similar what was observed in previous AOA labeling studies (Grammel et al. 2010). The identified *S. typhimurium* proteins did

not contain any known or previously described potential effector proteins and almost all of them were previously identified from intact bacteria infecting RAW264.7 macrophages (see chapter 1). We concluded that none of these proteins are in fact effector proteins.

Due to the low overall protein retrieval from the N1 supernatant fraction we turned to the Triton X-100 fractionation method, resulting in an increased fluorescence signal in the supernatant fraction (Figure 46). The RAW264.7 cell infection experiment was carried out as before, however cells were lysed in 0.1% Triton X-100. The resulting fractions were precipitated by methanol/chloroform precipitation and reacted with azido-azo-biotin for affinity enrichment. Since we expected higher protein recovery upon Triton X-100 lysis, compared to the N1 lysis, samples for LC-MS/MS analysis were prepared by in-gel digestion, providing a means of sample complexity reduction. Captured proteins were selectively eluted by sodium dithionite and separated by SDS-PAGE (Figure 47). SDS-PAGE analysis already indicated that the major protein components retrieved from the supernatant and the pellet fraction could be identical. LC-MS/MS and data analysis was carried out as previously (Grammel et al. 2010). 55 *S. typhimurium* proteins were identified in the supernatant fraction (Table 5), all of which were also detected in the bacterial pellet fraction, which comprised 140 identified *S. typhimurium* (Table 6). None of the identified proteins in the supernatant or pellet fraction was a known and characterized effector proteins (e.g. SifA, SseJ), which should be expressed 16 hours post-infection (Figure 41, Figure 42, Figure 43). In addition, non of the recently predicted putative effector proteins were among the identified proteins (Samudrala et al. 2009; Arnold et al. 2009). In conclusion, neither the highly selective N1 mechanical lysis nor

the harsher Triton X-100 lysis resulted in the proteomic identification of a known or putative effector protein. The reduction in general fluorescence recovery in the supernatant fraction only was accompanied by a reduction in dynamic range of bacterial protein detection, since the top protein hits in both, the N1 and the Triton X-100, lysis are the same. These proteins are highly abundant chaperones and translational components.

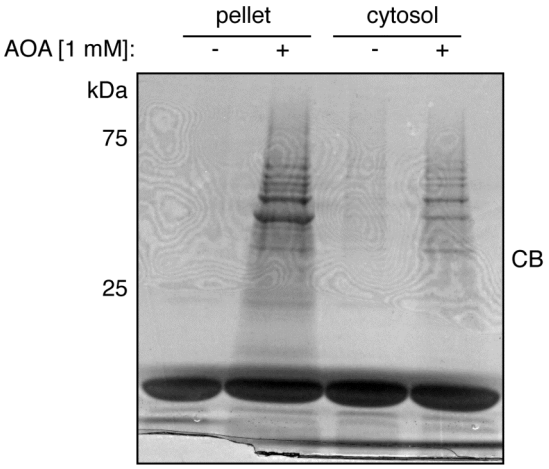


Figure 47 In-gel sample preparation enriched bacterial proteins from infected RAW264.7 cells, fractionated by 0.1% Triton X-100 lysis.

5.4 Materials and Methods

5.4.1 Labeling effector proteins in liquid culture with AOA or ANL

An overnight culture of *S. typhimurium* IR715 pWSK29_MetG$_{PLL}$, *S. typhimurium* 14028 pWSK29_MetG$_{PLL}$, or *S. typhimurium* 14028 ΔSPI-1 pWSK29_MetG$_{PLL}$ was centrifuged and resuspended in fresh LB medium with 100 μg/ml carbenicillin. For secretion assays the cultures were grown in 15 ml plastic culture tubes at 37 °C with aeration (2 ml culture volume). For amino acid labeling, 2 mM AOA, ANL, or Met was added either 2 or 3

hours afterwards. Bacteria were harvested by centrifugations, twice. To the culture supernatant trichloroacetic acid was added to a final concentration of 10% an the sample was incubated over night at 4 °C. The sample was centrifuged for 30 min at 20,000g at 4 °C. The resulting pellet was carefully rinsed with ice-cold acetone three times. The resulting pellets were resuspended in 20 μl 4% SDS, 50 mM TEA, 150 mM NaCl. To this solution 5 μl click-chemistry master mix was added (see 3.4.5). The samples were precipitated by methanol/chloroform precipitation after CuACC to remove all small molecules and finally completely loaded on a SDS-PAGE for separation and in-gel fluorescence scanning.

5.4.2 Cloning of SseJ into pWSK29

Initially a pWSK29_HA plasmid was generated by annealing two primers (TCGAGTTATGCATAATCAGGAACATCATAAGGATAG and AATTCTATCCTTATGATGTTCCTGATTATGCATAAC), digesting them with EcoRI and XhoI, and ligating them in pWSK29, also cut with EcoRI and XhoI. SseJ was amplified from *S. typhimurium* IR715 by PCR (GAATTCTTCAGTGGAATAATGATGAGCTATAAAAC and TCTAGAGATAGCAGTCAGATAATATGTACCAGGC). The resulting PCR product was digested with EcoRI and XbaI and ligated into pWSK29_HA, also cut with EcoRI ans XbaI.

5.4.3 Imaging of SseJ during infection by fluorescence microscopy

For infection see 3.4.4, for fluorescence microscopy see 3.4.6.

5.4.4 Pulse-labeling with EME and immunoprecipitation of SseJ during infection

Confluent 6 well dishes of RAW264.7 cells were infected with *S. typhimurium* IR715 pWSK29_HASseJ grown in LB medium to stationary phase at an MOI of 10 (see 3.4.4). 16 hours post-infection cells were washed with PBS and incubated with methionine free DMEM and 1 mM EME or 1 mM Met for 1 hour. Cells were washed with PBD and lysed in Brij-97 buffer (M. M. Zhang et al. 2010). Immunoprecipitation, CuAAC on beads, and analysis was carried out as reported previously (M. M. Zhang et al. 2010).

5.4.5 Fractionation of infected RAW264.7 macrophages

For detergent based lysis, infected cells were resuspended in ice-cold 50 mM TEA buffer pH 7.4 containing 1 mM PMSF, 1 RocheComplete protease inhibitor tablet per 25 ml, and different concentrations of detergent (0.1% Triton X-100, 0.01% Triton X-100, 0.2% saponin, or 0.02% saponin). Cells were lysed on ice for 10 min and centrifuged at 20,000 g for 5min at 4°C to yield the supernatant and pellet fraction.

For mechanical lysis, cells were resuspended in either 250 mM sucrose, 3 mM imidazole, 0.5 mM EDTA, 1 mM PMSF, and 1 RocheComplete protease inhibitor tablet per 25 ml (N1) or 50 mM TEA pH 7.4, 1 mM PMSF, and 1 RocheComplete protease inhibitor tablet per 25 ml (N2). Cells were pushed trough a 21.5 gauge needle 20 times. Lysates were centrifuged at 20,000 g for 5min at 4 °C to yield the supernatant and pellet fraction.

5.4.6 Sample preparation for LC-MS/MS analysis

Sample preparation by in-gel digestion, LC-MS/MS and data analysis was carried out as in 2.4.11, 2.4.12, 2.4.13, 2.4.14, and 2.4.15.

5.4.7 Filter aided sample preparation (FASP)

Sample preparation by filter aided sample preparation (FASP) was strictly carried out as published previously (Wiśniewski et al. 2009).

6 A chemical reporter for protein AMPylation[4]

6.1 Abstract

Protein AMPylation is an emerging posttranslational modification, which plays key roles in bacterial pathogenesis and cell biology. Enzymes with AMPylation activity, referred to as AMPylators, have been identified in several bacterial pathogens and eukaryotes. To facilitate the study of this unique modification, we developed an alkynyl chemical reporter for detection and identification of protein AMPylation substrates. Covalent functionalization of AMPylation substrates with the alkynyl reporter in lieu of adenylyl 5'-monophosphate (AMP) allows their subsequent bioorthogonal labeling with azide-fluorescent dyes or affinity enrichment tags. We show that this chemical reporter is transferred by a range of AMPylators onto their cognate protein substrates and allows rapid detection and identification of AMPylated substrates.

6.2 Introduction

Protein AMPylation refers to the covalent modification of protein side chain hydroxyl groups with adenylyl 5'-monophosphate (AMP) through a phosphodiester bond. This posttranslational modification is installed by the enzymatic transfer of AMP from adenosine-5'-triphosphate (ATP) to the substrate hydroxyl group. This catalytic activity was initially described for *E. coli* glutamine synthetase (GS) adenylyl transferase, which tightly regulates GS activity (Kingdon et al. 1967; Shapiro & Stadtman 1970). Additional proteins with AMPylation activity (AMPylators) have been identified that contain either

---

[4] This work was carried out in collaboration with Phi Luong (UT Southwestern) and Kim Orth (UT Southwestern). Phi Luong conducted all cloning and expression experiments.

the filamentation induced by cAMP (fic) domain or the adenylyl transferase (ATase) domain, which confer AMPylation activity (Woolery et al. 2010; Lisa N Kinch et al. 2009). Many of these characterized AMPylators serve as bacterial virulence factors that are secreted into the mammalian host cell during infection. There, they AMPylate mammalian host proteins to alter their function for the benefit of the pathogen. In particular, secreted bacterial AMPylators, such as VopS (*Vibrio parahaemolyticus*), IbpA (*Legionella pneumophila*), and DrrA (*Legionella pneumophila*) have been shown to target mammalian small GTPases, like RhoA, Rac1, Cdc42, and Rab1 (Melanie L Yarbrough, Yan Li, et al. 2009; Worby et al. 2009; Müller et al. 2010). Another bacterial effector protein has been identified that selectively deAMPylates a GTPase, indicating the dynamic nature of this modification (Tan & Luo 2011; Neunuebel et al. 2011). The AMPylation of small GTPases, on threonine or tyrosine residues, interferes with their proper function, either by sterically blocking the interaction with downstream signaling components or GTPase activating proteins (GAPs). Interestingly, fic and ATase domains are not limited to bacterial effector proteins, but have been identified in archaea and eukaryotes as well. Most eukaryotic genomes appear to contain a fic domain protein. AMPylation activity has been observed for the human protein HYPE and the Drosophila protein dFic (Lisa N Kinch et al. 2009; Worby et al. 2009). The widespread presence of these domains suggests a ubiquitous role for protein AMPylation as a regulated and reversible posttranslational modification. While radioactive ATP, targeted mass spectrometry, and specific antibodies, can be used to detect AMPylated substrates, more general and efficient analytical tools are still needed for the unbiased identification of new AMPylated substrates and the analysis of their regulation. We therefore developed

an alkynyl chemical reporter for bioorthogonal detection, enrichment and identification of AMPylated proteins (Figure 48).

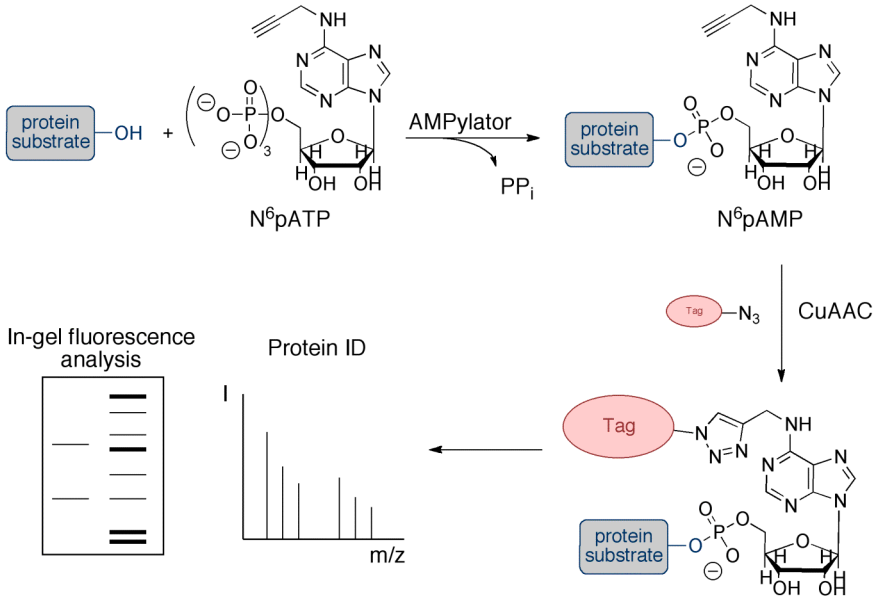


Figure 48 N^6pATP – N^6-propargyl adenosine-5'-triphosphate; N^6pAMP – N^6-propargyl adenosine-5'-monophosphat; PP_i – pyrophosphate; CuAAC – Cu(I)-catalyzed azide-alkyne cycloaddition; Tag – rhodamine fluorescence dye or cleavable biotin enrichment tag .

6.3 Results

Alkynyl chemical reporters allow Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) of labeled substrates with azide-functionalized detection and enrichment reagents. This technology has facilitated the analysis of various posttranslational modifications and nucleic acid biogenesis (Sletten & Bertozzi 2009). Recent structural studies of AMPylators and previous studies of fluorescent AMP analogs, suggested that a modification of the N^6 position of the adenine ring could be tolerated (Chock et al. 1973; Rhee et al. 1981; Xiao et al. 2010; Luong et al. 2010a; Palanivelu et al. 2011). Thus, we

synthesized the ATP analog N^6-propargyl adenosine-5'-triphosphate (N^6pATP) as a potential chemical reporter for AMPylation.

To assess the activity of N^6pATP as a possible chemical reporter for AMPylation, we used a well-established in vitro system, based on the recently identified AMPylator VopS and its cognate mammalian target Cdc42 (M. L. Yarbrough, Y. Li, L. N. Kinch, et al. 2009). Cdc42(Q61L), subsequently referred to as Cdc42, was incubated with VopS in the presence of increasing concentrations of N^6pATP under previously reported in vitro AMPylation conditions (Luong et al. 2010b). N^6pAMP transfer was analyzed by CuAAC with azido-rhodamine (az-rho) dye and in-gel fluorescence scanning (Charron et al. 2009). Increasing concentrations of N^6pATP yielded dose-dependent increase in fluorescence labeling of Cdc42 (Figure 49). The chemical reporter proved to be transferred only in the presence of VopS and appeared to be highly selective for the native substrate Cdc42, as judged by the lack of concomitant BSA labeling (Figure 49).

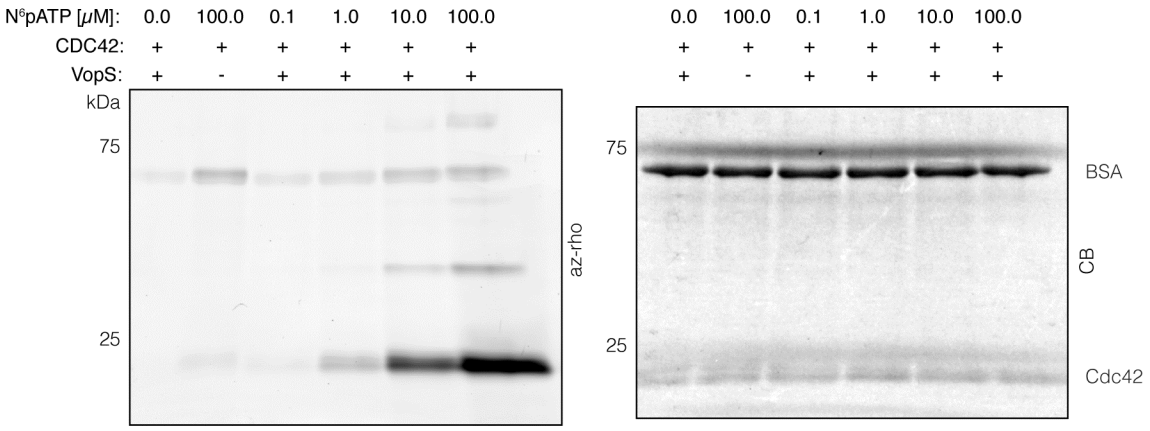


Figure 49 VopS labeled its native substrate with the chemical reporter N^6pATP. Cdc42 (0.5 μg) was incubated with VopS (1 ng) with increasing concentrations of N^6pATP for 1 hour at 30°C in reaction buffer and analyzed by CuAAC and in-gel fluorescence. az-rho: azido-rhodamine, CB: coomassie blue.

To investigate the selectivity of N^6pATP, we made use of substrate and AMPylator mutants. When VopS was incubated with the T35A mutant of Cdc42, in which the target threonine residue of VopS was mutated to an alanine, no transfer was observed (Figure 50b) (M. L. Yarbrough, Y. Li, L. N. Kinch, et al. 2009). In addition, a catalytic inactive mutant of VopS, H348A, failed to transfer the chemical reporter (Figure 50b) (M. L. Yarbrough, Y. Li, L. N. Kinch, et al. 2009). The dependence on this previously described catalytic histidine suggests that N^6pAMP is transferred by the same catalytic mechanism as the native AMP group. These data emphasize the distinct selectivity of N^6pATP, which was further demonstrated by competitive inhibition with ATP (Figure 50c). N^6pATP appears to be a general cofactor for AMPylation, since it was also utilized by VopS to modify its alternative substrates RhoA and Rac1 (Figure 50d). It should be noted that labeling of Cdc42(T35S) by VopS was also observed (Figure 50e), supporting the notion that AMPylation may occur on serine residues. Moreover, it served as a chemical reporter for another fic domain AMPylator, IbpA (Fic2, Figure 51a). Like VopS, Fic2 modifies Cdc42, however, it transfers AMP onto tyrosine residue Y32 (Worby et al. 2009) (Figure 51b). We also tested the ATase domain AMPylator DrrA for its ability to utilize N^6pATP. Indeed, DrrA accepted N^6pATP as cofactor and labeled the reported protein substrate Rab1 (Figure 51c) (Worby et al. 2009). These experiments demonstrate the versatile nature of N^6pATP as a chemical reporter for all known AMPylator families, namely fic domain and ATase domain AMPylators. N^6pATP labeling appears to be independent of the target protein substrate and could also allow the identification of serine-modified AMPylation substrates.

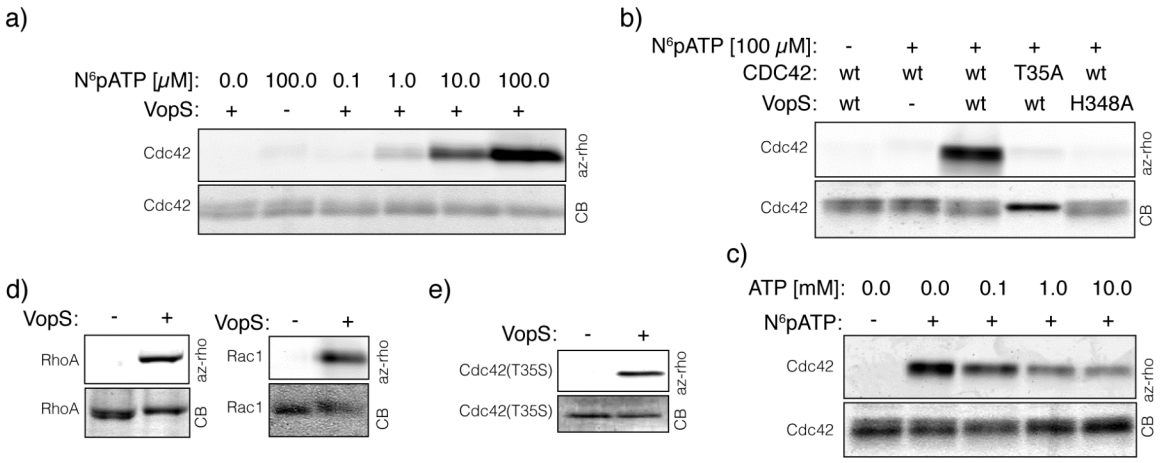


Figure 50 In vitro analysis of VopS activity, using N^6-propargyl adenosine-5'-triphosphate (N^6pATP) as a chemical reporter, by click-chemistry and in-gel fluorescence scanning. All AMPylation reactions were carried out in 15 μl total volume for 1 h at 30°C . A) Cdc42 (0.5 μg) was incubated with VopS (1 ng) and increasing concentrations of N^6pATP. B) Cdc42 (0.5 μg) or Cdc42(T35A) (0.5 μg) was incubated with VopS (1 ng) or VopS(H348A) (1 ng) and N^6pATP. C) Analysis of Cdc42 labeling by VopS, using N^6-propargyl-adenosine-5'-triphosphate (N^6pATP) as a chemical reporter, with increasing concentrations of adenosine-5'-triphosphate, by click-chemistry and in-gel fluorescence scanning. VopS (1 ng) was incubated with Cdc42 (0.5 μg), N^6pATP (100 μM) and increasing concentrations of ATP in a total volume of 15 μl for 1 h at 30°C. D) RhoA (0.5 μg) was incubated with VopS (10 ng). Rac1 (0.5 μg) was incubated with VopS (10 ng) in the presence of N^6pATP (100 μM) in a total volume of 15 μl for 1 h at 30°C. E) Cdc42(T35S) (0.5 μg) was incubated with VopS (10 ng). az-rho: azido-rhodamine fluorescence; CB: coomassie blue.

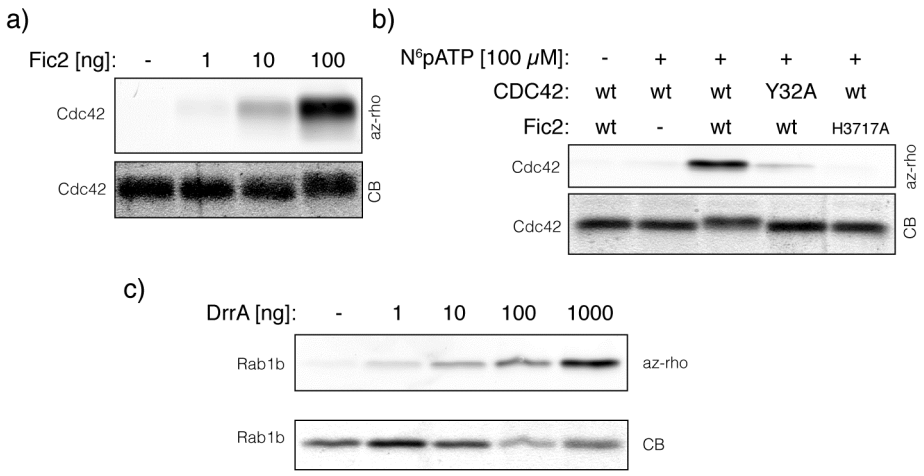


Figure 51 N^6pATP is used by all known AMPylation enzyme families. a) Cdc42 (0.5 μg) was incubated with increasing amounts of Fic2 (IbpA) in the presence of N^6pATP in reaction buffer. b) Cdc42 (0.5 µg) or Cdc42(Y32A) (0.5 µg) was incubated with Fic2 (10 ng) or Fic2(H3717A) (10 ng) and N^6pATP in a total volume of 15 µl for 1 h at 30°C. c) Increasing amounts of DrrA were incubated with Rab1a (0.5 µg) and N^6pATP (100 µM) in a total volume of 15 µl for 1 h at 30°C. az-rho: azido-rhodamine, CB: coomassie blue.

We next analyzed the ability of VopS, Fic2, and DrrA to modify their substrates with N^6pATP in mammalian cell lysates. It has been reported that addition of purified VopS or Fic2 to HeLa cell lysates in the presence of ^{32}P-a-ATP results in the labeling of one distinct band in the molecular weight range of small GTPases, as assessed by autoradiography (Melanie L Yarbrough et al. 2009; Worby et al. 2009). We incubated purified VopS, Fic2, and DrrA with Triton X-100 lysed HeLa cell lysates and N^6pATP (100 μM). As previously observed for ^{32}P-a-ATP labeled cell lysates, VopS and Fic2 labeled one distinct protein or protein population at the corresponding molecular weight of small GTPases (Figure 52a, b). Addition of DrrA also resulted in the modification of one distinct protein or protein population at ~21 kDa (Figure 52c), although more recombinant DrrA was needed to observe similar levels of labeling. In the absence of recombinant enzyme, the reporter did not significantly label the cell lysate. All three enzymes displayed auto-AMPylation activity as evidenced by a specific fluorescent signal at the corresponding molecular weight (Figure 52a, b, c). These results corroborate previous findings, underscoring the distinctive selectivity of these effector proteins for small host GTPases (Melanie L Yarbrough et al. 2009; Worby et al. 2009).

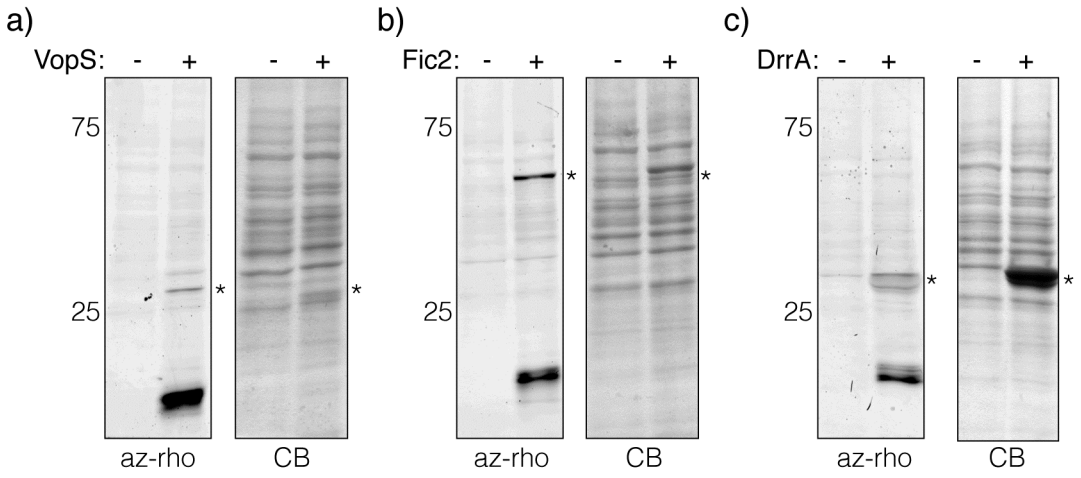
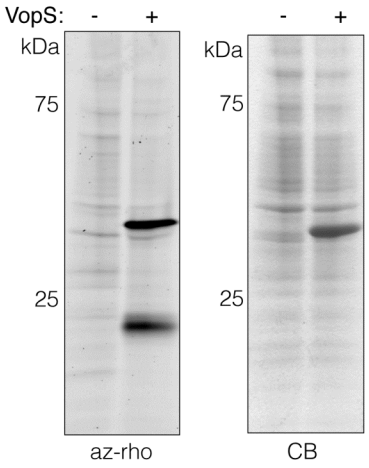


Figure 52 Detection of endogenous AMPylation substrates in cell lysates. AMPylation reactions (a-c) were carried out in 15 μl for 1 h at 30°C. a) VopS (100 ng) was added to HeLa cell lysate (10 mg) in the presence of N^6pATP (100 μM). b) Fic2 (300 ng) was added to HeLa cell lysate (10 mg) in the presence of N^6pATP (100 μM). c) DrrA (4 μg) was added to HeLa cell lysate (10 mg) in the presence of N^6pATP (100 μM).

To explore the utility of N^6pATP for the unbiased identification of AMPylated substrates, we incubated HeLa cell lysate (1 mg) with N^6pATP (100 μM) in the presence or absence of VopS (88μg). The bulk of the samples was then reacted with a cleavable biotin enrichment tag (*ortho*-hydroxy-azidoethoxy-azo-biotin) and incubated with Streptavidin beads for affinity capture of labeled proteins (Y.-Y. Yang et al. 2010). Captured proteins were digested on beads and the resulting peptides were collected and subjected to protein identification by tandem mass spectrometry. For data analysis, all protein hits of the negative control were subtracted from the VopS containing sample, which revealed two final protein identifications: VopS and Cdc42, with VopS peptides representing the majority of the identified peptides (Figure 53b, c). The dominant VopS fraction correlates well with the observed fluorescence profiling of the same samples (Figure 53a). This data demonstrates that N^6pATP can be used to identify substrates of individual AMPylators in complex cellular protein samples and confirms Cdc42 as an endogenous VopS target.

a)



b)

protein	# TS	# UP
VopS	60	20
Cdc42	3	3

c)

Cdc42 peptide	XCorr	DCn
NVFDEAILAALEPPEPK	3.14	0.279
NVFDEAILAALEPPEPKK	3.20	0.404
QKPITPETAEK	2.22	0.249

Figure 53 Proteomic analysis of endogenous VopS substrates in Hela cell lysates confirms Cdc42. a) 1 mg HeLa cell lysates was incubated with VopS (88 μg) and N^6pATP (100 μM). b) Proteomic analysis of VopS substrates in HeLa cell lysate identified two proteins above background, VopS and Cdc42. c) Identified Cdc42 peptides. # TS – number of total spectra, # UP – number of unique peptides, XCorr – SEQUEST Xcorr score, DCn – SEQUEST DCn score, az-rho - azido-rhodamine fluorescence; CB, coomassie blue; asterisk indicates auto-AMPylation and recombinant AMPylator.

6.4 Materials and Methods

6.4.1 General methods and materials

All chemical reagents were purchased from Sigma-Aldrich. Analytical and preparative HPLC was conducted on a Waters 1525 Binary HPLC Pump system with a XBridge Prep C18 5 μm OBD, 19 x 150 mm, column or a Xbridge C18 5μm, 4.6 x 150 mm, analytical column. Compounds were eluted with a linear gradient of 100% acetonitrile and 20 mM triethylammonium bicarbonate, 2% acetonitrile, adjusted to pH 7.4 with acetic acid. ^{31}P NMR analysis was done on a Bruker Avance DPX400 and ^{1}H analysis on a Bruker Avance DMX600 system. HRMS analysis was conducted by the Proteomics Resource Center of The Rockefeller University on a LTQ-Orbitrap (Thermo Scientific). Azido-rhodamine (az-rho) and TBTA (tris((1-benzyl-1H-1,2,3-triazol-4-yl)methyl) amine) was generated by the Hang laboratory as previously described (Charron et al. 2009).

6.4.2 Chemical synthesis

6.4.2.1 6-propargyl-adenosine

6-propargyl-adenosine was synthesized as previously reported (Jiang et al. 2009). Briefly, 6-Choropurine nucleoside (1 eq.), calcium carbonate (2 eq.) and propargylamine (5 eq.) were added to absolute ethanol and refluxed under argon for 12 hours. The reaction mixture was filtered and the filtrate was kept at -20 °C over night. The white precipitate was collected by filtration, washed with ether and dried under vacuum to afford the pure product. ^{1}H NMR (DMSO): δ 8.36 (s, 1H), 8.24 (s, 1H), 8.20 (br, 1H), 5.86 (d, 1H, J = 6.1 Hz), 5.4 (d, 1H, J = 6.2 Hz), 5.28 (dd, 1H, J = 4.7, 2.3 Hz), 5.14 (d, 1H, J = 4.7 Hz), 4.56 (dd, 1H, J = 6.0, 5.2 Hz), 4.21 (br, 2H), 4.11 (dd, 1H, J = 4.7, 3.3 Hz), 3.92 (dd, 1H, J = 3.5, 3.4 Hz), 3.63 (m, 1H), 3.51 (m, 1H).

6.4.2.2 N^6-propargyl-adenosine-5'-triphosphate

N^6-propargyl-adenosine-5'-triphosphate was synthesized using the Yoshikawa (Yoshikawa et al. 1967) reaction for selective phosphorylation of the 5'-hydroxyl group, generally following the method of Ludwig (Ludwig 1981). Trimethyl phosphate was dried for 18 h over molecular sieves (3A) and under argon atmosphere. 0.5 ml trimethyl phosphate was transferred into a flame dried round bottom flask and 6-propargyl-adenosine (61 mg, 0.2 mmol) was added. Phosphoryl chloride (39.8 mg, 0.26 mmol) was added to the stirred suspension with a syringe and the reaction mix was stirred for 1.5 h at 0 °C until it cleared. To this solution tributylammonium pyrophosphate (549 mg, 1 mmol), in 2 ml anhydrous dimethylformamide, was added and stirred for 20 min at 0 °C. The reaction mix was further stirred for 5 min at room temperature, before 2 ml of triethylammonium bicarbonate (1M) was added. Solvent was evaporated under reduced pressure and half of the resulting material was purified by reverse phase HPLC using a gradient of 20 mM triethylammonium bicarbonate (pH 7.4, acetic acid, 2% acetonitrile) and 100 % acetonitrile. Fractions that contained the desired product were lyophilized to afford 54 mg of N^6-propargyl-adenosine-5'-triphosphate as a triethylammonium salt (54% extrapolated yield). ^{31}P NMR (D_2O): 1.74 (s), -9.63 (m), -10.27 (m); ^{1}H NMR (D2O): δ 8.55 (m, 1H), 8.34 (m, 1H), 6.16 (m, 1H), 4.56 (m, 1H), 4.40 (m, 3H), 4.30 (m, 1H), 4.22 (m, 1H), 4.07 (m, 1H), 2.63 (s, 1H). HRMS calcd. for $C_{13}H_{18}N_5O_{13}P_3$ [$(M+H)^+$] 546.0187, obsd. 546.0247.

6.4.3 Cloning and expression

VopS (31-387) pGEXTEV and His_6-Cdc42-(1-179)-(Q61L) pPROEXHTa were expressed and purified as described previously (Luong et al. 2010). GST-Fic2 pGEXTEV, GST-Fic2-(H3717A) pGEXTEV, Rab1a (Canis) pet28a, DrrA (1-339) pet28a, RacV12 pet28a, RhoApGEX-KG, His_6-Cdc42-(1-179)-(Q61L, T35S) pPROEXHTa, His_6-Cdc42-(1-179)-(Q61L, T35A) pPROEXHTa and His_6-Cdc42-(1-179)-(Q61L, Y32A) pPROEXHTa were expressed and transformed in BL21 (DE3) cells. Protein expression was performed overnight at room temperature with 0.4 mM IPTG. Protein pellets were purified using glutathione-agarose (Sigma) or immobilized metal ion affinity chromatography (Sigma). Proteins were buffer exchanged in 20 mM Tris pH 8.0, 100 mM NaCl, 10% glycerol, 1 mM DTT with an Amicon 10K concentrator and stored at -80 °C. Protein concentrations were assayed by the Bradford method (Bio-Rad) and purity by SDS-PAGE analysis.

6.4.4 AMPylation reactions

In vitro AMPylation reactions were carried out at 30 °C for 1 hour, if not indicated otherwise, in a total volume of 15 μl. For *in vitro* AMPylation reactions with purified substrates, protein substrates (0.5 μg) and AMPylation enzymes (VopS 1-10 ng, Fic2 1-100 ng, DrrA 1-1000 ng) were incubated with N^6pATP (100 μM, unless indicated otherwise) in 20 mM Hepes pH 7.4, 100 mM NaCl, 5 mM $MgCl_2$, 0.1 mg/ml BSA and 1mM DTT. For the ATP competition experiment, increasing concentrations of ATP from a dilution series was added to the assay mix.

For *in vitro* AMPylation reactions with cell lysates, AMPylation enzymes (100 ng VopS, 300 ng Fic2, 4 μg DrrA) were incubated with 10 μg cell lysate in 20 mM Hepes pH 7.4, 100 mM NaCl, 5 mM $MgCl_2$, and 1mM DTT.

HeLa cells were grown in DMEM medium supplemented with 10% fetal bovine serum. HeLa cell lysates were generated from 1 P15 dish of 80% confluent HeLa cells. Cells were harvested by trypsinization, collected by centrifugation, and washed once in PBS. The cell pellet was lysed with 1 ml 20 mM Tris, pH 8.0, 150 mM NaCl, 0.5% Triton X-100, 1 mM EDTA and Roche complete EDTA-free protease inhibitors (1 tablet/25 ml) for 20 min on ice. Cell lysate was cleared by centrifugation at 23,000 g, 4 °C, for 15 min. Protein concentration was determined using the BCA assay. The resulting protein concentration was 5 μg/μl.

6.4.5 Click-chemistry and in-gel fluorescence analysis

The AMPylation reaction was stopped by methanol/chloroform precipitation. To 15 μl assay volume, 85 μl water, 400 μl methanol and 100 μl chloroform was added and samples were vortexed; 300 μl water was added, samples were vortexed again and centrifuged at 23,000 g, at room temperature, for 5 min. The upper phase was carefully discarded and 1 ml ice-cold acetone was added. Samples were vortexed and centrifuged at 23,000g, at 4 °C, for 10 min. The supernatant was discarded and pellets were dried for 5 min. Subsequently, pellets were dissolved in 15 μl 4% SDS, 150 mM NaCl, 50 mM triethanolamine, pH 7.5 (4ST).

For in-gel fluorescence detection of N^6pATP modified substrate proteins, resolubilized protein pellets from AMPylation reactions were reacted with azido-rhodamine (az-rho). For this purpose a click-chemistry master mix was prepared with 7.5 μl 4ST, 0.25 μl 10

mM az-rho, 0.5 μl 50 mM TCEP (tris(2-carboxyethyl)phosphine), 1.25 μl 2 mM TBTA (tris((1-benzyl-1H-1,2,3-triazol-4-yl)methyl) amine) in 1:4 DMSO:n-butanol and 0.5 μl $CuSO_4$ per individual sample. 10 μl click-chemistry master mix was added to each individual protein sample to yield a total volume of 25 μl with 0.1 mM az-rho, 1 mM TCEP, 0.1 mM TBTA and 1 mM $CuSO_4$. Click-chemistry was carried out for 1 hour at room temperature. Click-chemistry was stopped by addition of 8.3 μl 4x Laemmli buffer and 1.6 μl 2-mercaptoethanol. Samples were incubated at 95 °C for 5 min and run for analysis on a 4-20% Tris-HCl Criterion™ Precast Gel (Biorad).

Protein gels were rinsed with deionized water and incubated for multiple hours in 50% water, 40% methanol, 10% acetic acid at 4 °C. Before in-gel fluorescence analysis protein gels were transferred to deionized water and incubated for at least 30 min at room temperature. Protein gels were scanned on a Amersham Bioscience Typhoon 9400 variable mode imager (excitation 532 nm, 580 nm filter, 30 nm band-pass). All image adjustments were done on the entire depicted image for all samples equally in Adobe® Photoshop® CS3. Subsequent to fluorescence scanning, protein gels were stained with GelCode Blue Stain Reagent (Thermo) or Coomassie Brilliant Blue (BioRad). Both stains are indicated as coomassie blue (CB) in all figures.

6.4.6 AMPylation reaction and sample preparation for proteomic analysis

For large-scale proteomic analysis, 1 mg of HeLa cell lysate (preparation see above) was incubated with 100 μM N^6pATP in the presence or absence of VopS (88 μg) for 1.5 h at 30 °C in a total volume of 1.5 ml of 20 mM Hepes pH 7.4, 100 mM NaCl, 5 mM $MgCl_2$, and 1mM DTT. The reaction mix was precipitated by methanol chloroform precipitation, adding 6 ml methanol, 1.5 ml chloroform, and 4.5 ml water in a 15 ml falcon tube.

Samples was vigorously vortexed and centrifuged for 30 min at 5200g at 4 °C. The upper phase was discarded and 10ml ice-cold acetone was added and mixed. Samples were centrifuged for 30 min at 5200g at 4 °C and the supernatant was discarded and the pellets were dried on air.

Protein pellets were resuspended in 400 μl 4 % SDS in PBS. A click-chemistry master mix was prepared with 21 μl of a 10 mM solution of *ortho*-hydroxy-azidoethoxy-azo-biotin (compound 5 in Yang et al. 2010), 42 μl of a 50 mM TCEP (tris(2-carboxyethyl)posphine; freshly prepared in water) solution, 105 μl of a 2 mM TBTA (tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine; in 4:1 n-butanol and DMSO) solution, 42 μl of a 50 mM $CuSO_4$ solution, and 1050 μl of 4 % SDS in PBS. 600 μl of the master mix was added to each sample and each sample was incubated for 1.5 h at room temperature. Click-chemistry was ended by methanol/chloroform precipitation (4 ml methanol, 1 ml chloroform, 3 ml water). The resulting pellet was rinsed 3 times with ice-cold methanol and vigorously vortexed each time. The pellets were dried on air. The resulting protein pellets were dissolved in 1 ml 6M urea/2M thiourea in 10 mM Hepes pH 8.0. Proteins were reduced by addition of DTT to 1mM (from 500 mM DTT stock) for 45 min at room temperature. Then, proteins were alkylated by addition of iodoacetamide to 5.5 mM from a 500 mM stock. Samples were incubated for 30 min at room temperature in the dark. High-capacity streptavidin beads (Thermo) were equilibrated with 6M urea/2M thiourea in 10 mM Hepes pH 8.0. The reduced and alkylated protein samples were directly loaded onto the equilibrated beads and incubated on a nutator mixer for 1.5 h at room temperature. The beads were washed 3 times with 6M urea/2M thiourea in 10 mM Hepes pH 8.0, 3 times with 0.2% SDS in PBS, 3 times with PBS, and 1 time in 4.5

M urea/1.5 M thiourea in 10 mM Hepes pH 8.0. Beads were drained and 200 μl 4.5 M urea/1.5 M thiourea in 10 mM Hepes pH 8.0 was added with 2 μg Lys-C. Samples were incubated in shaker (1000 rpm) for 4 h at room temperature. Samples were diluted by addition of 600 μl 10 mM Hepes pH 8.0 and 2 μl trypsin was added to each sample (2 μl/5 μl in 50 mM acetic acid stock). Samples were incubated over night at 1000 rpm at room temperature. Beads were spun down (2000g for 1 min) and the supernatant was collected for mass spectrometry analysis. The supernatant was then purified by Zip-Tip Pipette tips (Millipore).

6.4.7 Mass spectrometry

LC-MS analysis was performed at the Rockefeller Proteomics Resource Center, with a Dionex 3000 nano-HPLC coupled to an LTQ-Orbitrap ion trap mass spectrometer (ThermoFisher). Peptides were pressure loaded onto a home-made 75 μm diameter, 15 cm C18 reverse phase column and separated with a gradient running from 95% buffer A (HPLC water with 0.1% formic acid) and 5% buffer B (HPLC grade acetonitrile with 0.1% formic acid) to 55% B over 30 min, next ramping to 95% B over 10 min and holding 95% B for 10 min. One full MS scan (300-2000 MW) was followed by 3 data dependent scans of the most intense ions with dynamic exclusion enabled. The spray voltage was set to 1.94 kV and the flow rate through the column was set to 0.25 μL/min.

6.4.8 Data analysis

For protein identification all generated .raw files were searched using the SQUEST search algorithm under BioWorks 3.3.1 SP1. All searches were performed against a indexed database containing the international protein index (IPI) human database v3.86, the sequence of VopS, and the sequence of streptavidin. The database was indexed in

BioWorks 3.3.1 with carbamidomethylation as static modification, trypsin as proteolytic enzyme, 'fully enzymatic' as enzyme limit, 250-7000 as MW range, and 2 missed cleavage sites. SEQUEST searches were conducted with monoisotopic precursor and fragment masses, 50.00 ppm peptide tolerance, 1.0 AMU fragment ion tolerance, and B and Y ion series calculated. Carbamidomethylation was selected as static modification, methionine/tryptophan oxidation as well as asparagine/glutamine deamindation was selected as variable modification. The generated unified search files .srf were opened in Scaffold 3 (Proteome Software) and further analyzed. All searches were filtered in Scaffold 2 for 99.9% protein probability, 2 unique peptides, and 95% peptide probability. Data was exported into Excel and further analyzed. All identifications with hits in the control sample were discarded, as well as all identified contaminations (e.g. keratins).

7 Chemical reporters for RNA synthesis and mRNA polyadenylation[5]

7.1 Abstract

Modern array and sequencing technologies have tremendously improved our capabilities for transcriptome analysis. In parallel metabolic labeling approaches have been taken to label newly synthesized mRNAs with chemical reporters, which allow analysis of mRNA synthesis and degradation. In this chapter we present a new adenosine based chemical reporter for RNA synthesis that is incorporated into newly transcribed mRNA and into mRNA polyadenylation tails.

7.2 Introduction

Understanding the mechanisms of gene expression requires a thorough knowledge of the balance between the biogenesis of an mRNA and its degradation. While advances in transcriptome profiling methods have enabled the large-scale analysis of RNA steady-state levels, monitoring the synthesis, processing and decay rates of specific transcripts remains technically challenging. The application of nucleoside analogs such as 4-thiouridine (4-sU) has recently enabled metabolic labeling and enrichment of nascent transcripts to monitor RNA dynamics in mammalian cells (Rabani et al. 2011). Thiol-selective labeling allows the capture and enrichment of RNA but is incompatible with cellular imaging due to competing free thiols in cells. Alternatively, metabolic labeling with 5-ethynyluridine (EU) is compatible with fluorescence imaging of nascent

[5] This work was carried out in collaboration with Nicholas K. Conrad (UT Southwestern). Nicholas K. Conrad conducted all experiments pertaining to Figure 55, Figure 56, and Figure 57.

transcripts in cells, tissues and model organisms following copper (I)-catalyzed azide-alkyne cycloaddition (CuAAC or "click chemistry") with azide-functionalized fluorescent dyes (Jao & Salic 2008). Here we report N^6-propargyl adenosine (N^6pA; Figure 54a) is efficiently incorporated into RNA in mammalian cells and can be used for CuAAC-mediated fluorescence imaging and affinity enrichment of nascent transcripts. We also demonstrate how "clickable" nucleosides can be employed to monitor polyadenylate [poly(A)] tail dynamics in cells.

7.3 Results

To determine whether N^6pA labels cellular RNA, we analyzed metabolic incorporation of N^6pA into mammalian cells by fluorescence microscopy. HeLa cells were incubated in complete medium containing a range of N^6pA concentrations (1.0 μM - 1 mM). After 12 hours, the cells were washed, permeabilized, fixed, reacted with azido-rhodamine via CuAAC and visualized by fluorescence microscopy (Figure 54b). The majority of N^6pA-specific fluorescence signal localized to distinct substructures of the nucleus, most likely nucleoli, in accordance with the dominant transcriptional activity of RNA polymerase I (pol I). Incorporation of N^6pA was observed with as little as 10 μM of N^6pA and appeared to reach its maximum between 10 μM and 100 μM. To distinguish between DNA and RNA labeling, we examined N^6pA incorporation in the presence of hydroxyurea, a ribonucleotide reductase inhibitor, or actinomycin D, a transcription inhibitor (Figure 54c). While the addition of hydroxyurea had no effect, actinomycin D completely abolished fluorescence, demonstrating that N^6pA is incorporated into RNA, but not DNA. We conclude that, similar to EU, N^6pA is efficiently incorporated into rRNA by pol I in cells.

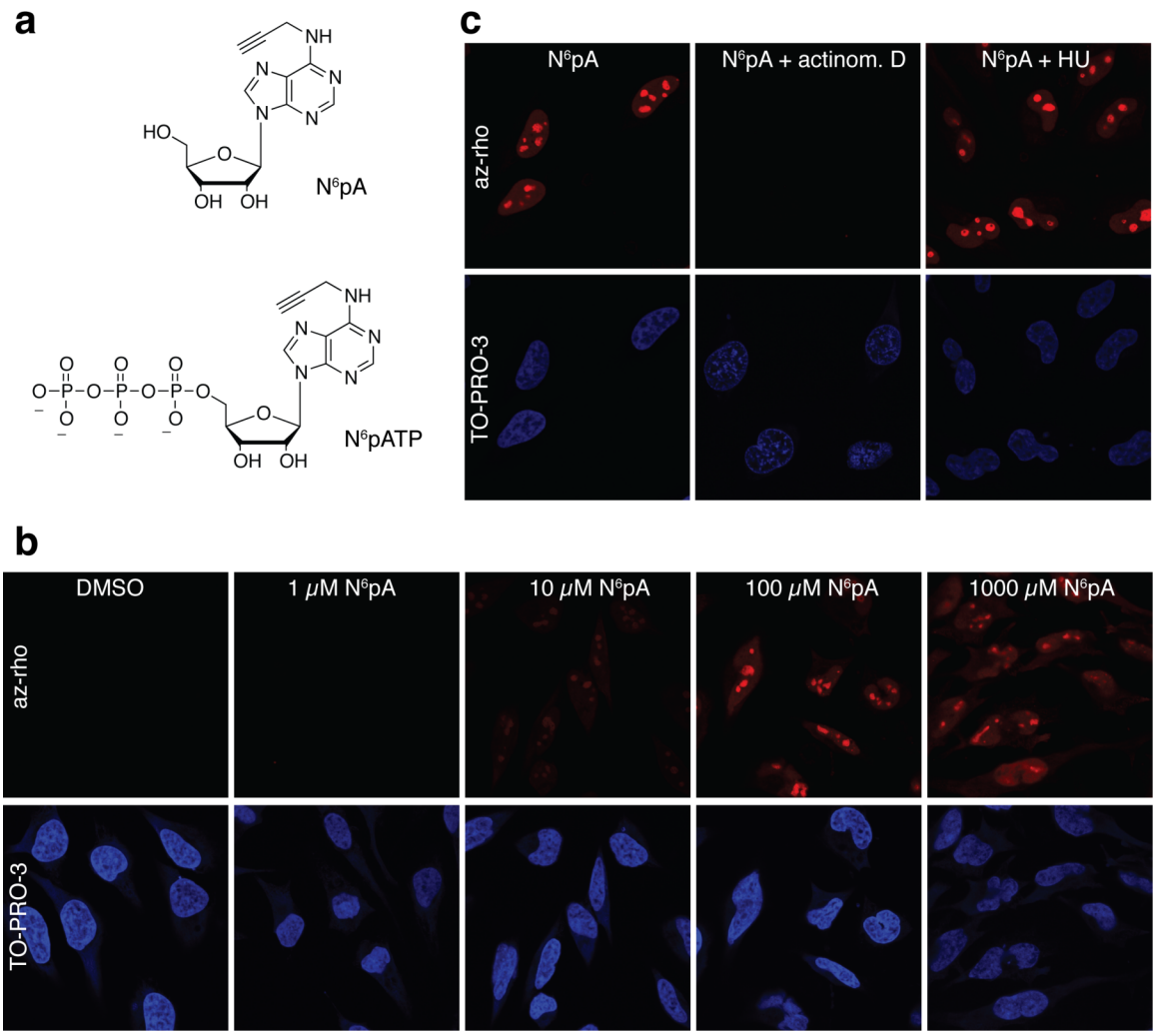


Figure 54 N^6pA is incorporated by pol I in cells. a) Structure of N^6pA and N^6pATP. b) Cells were incubated for 12 hours in media containing the indicated concentrations of N^6pA. After fixation and permeabilization, rhodoamine-azide was conjugated via click chemistry. Top panels are rhodamine signal and the bottom panels are nuclei counterstained with TO-PRO-3. c) Cells were incubated with 1 mM N^6pA in the presence of actinomycin D or hydroxyurea as indicated; fluorescent assay performed as in b).

In contrast to EU, N^6pA is a potential substrate for incorporation into poly(A) tails by cellular poly(A) polymerases. To test this, we performed *in vitro* polyadenylation reactions with HeLa nuclear extracts, a uniformly ^{32}P-labeled RNA substrate and N^6-propargyl adenosine-5´-triphosphate (N^6pATP; Figure 54a) (Grammel et al. 2011). Incubation of the substrate in nuclear extract under polyadenylation conditions leads to the addition of a poly(A) tail of ~150-250 nt, as previously observed (Wilusz & Shenk 1988) (Figure 55a, lanes 1 and 2). Inclusion of N^6pATP in the reaction at various concentrations (Figure 55a, lanes 3-5) had no apparent effect on polyadenylation efficiency, as judged by the intensity and lengths of the polyadenylated products (Figure 55a, lanes 2-5). However, when ATP was replaced completely with N^6pATP, no poly(A) extension was observed (Figure 55a, lane 6). To show incorporation of N^6pA into poly(A) tails, polyadenylation products were biotinylated using CuAAC with biotin-azide and captured on streptavidin beads. Polyadenylated RNAs generated in the presence of N^6pATP were retained by streptavidin beads, whereas those generated in its absence were not (Figure 55a, compare lanes 9-11 to lane 8). In principle, incorporation of only one N^6pATP and its subsequent biotinylation should be sufficient for binding. However, the recovery of the polyadenylated substrates correlated with higher N^6pATP:ATP ratios. This may indicate that N^6pATP is not utilized by poly(A) polymerase as efficiently as ATP or that any one of the downstream steps (i.e. biotinylation, binding or retention on the column) requires more than one modified residue for maximal recovery. Interestingly, while N^6pATP alone did not support processive polyadenylation, this substrate was recovered more efficiently than the control (Figure 55a, lane 12 versus lane 7). It is therefore likely that a few N^6pATP molecules were incorporated into the substrate, but

the tail could not be further extended. Perhaps natural adenosines in the poly(A) tail are required to promote processive polyadenylation or ATP may be necessary for the activity of putative ATP-dependent enzymes (e.g. helicases) involved in polyadenylation in extract. In either case, these data clearly show that N^6pA is utilized by a poly(A) polymerase *in vitro*.

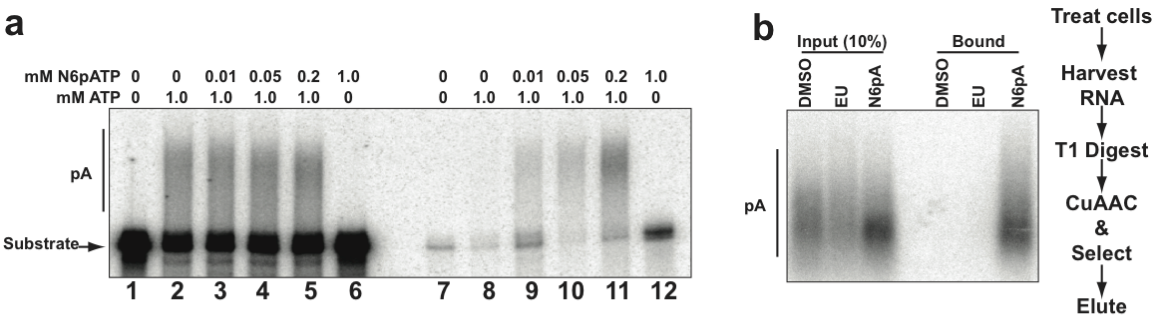


Figure 55 N^6pA is incorporated into poly(A) tails in vitro and in vivo. a) In vitro polyadenylation assays with a ^{32}P-labeled "pre-cleaved" substrate in HeLa nuclear extract (Wilusz & Shenk 1988). The substrate is a 247-nt in vitro transcribed RNA derived from the 3´ end of PANΔ79 RNA (Conrad et al. 2007) that contains the AAUAAA polyadenylation signal and terminates at the natural cleavage site. As a result, polyadenylation is uncoupled from the endonucleolytic cleavage step of 3´-end formation. The concentrations of N^6pATP and ATP in the reaction are shown above each lane. Lanes 1-6 contain the products of the polyadenylation reaction (20% of the total reaction), while lanes 7-12 are the streptavidin-bound transcripts. b) N^6pA is incorporated into poly(A) tails in vivo. 10% of the RNase T1-treated RNA (Input) or 100% of the streptavidin-bound RNA was analyzed by northern blot using an oligo dT probe to detect poly(A) tails. The experimental flow is shown to the right.

Next we tested whether N^6pA is incorporated into poly(A) tails in mammalian cells. To do this, we labeled HEK293 cells with 200 μM EU or N^6pA for 12 hr, harvested cells, and extracted RNA. To degrade total RNA but leave poly(A) tails intact, we treated

the RNA with RNase T1, a G-specific endonuclease, prior to biotinylation and selection. Heterogeneous poly(A) tails are observed in each of the input samples (Figure 55b). However, only the poly(A) tails from cells labeled with N^6pA bound the streptavidin column (Figure 55b), while the EU-treated or control (DMSO) poly(A) tails did not. These results demonstrate that N^6pA is incorporated into poly(A) tails by at least one poly(A) polymerase in cells. While several poly(A) polymerases are expressed in cells, our results are consistent with the conclusion that we are monitoring incorporation by the "canonical" poly(A) polymerase, PAPα (PAPOLA) (Eckmann et al. 2011). PAPα is responsible for the polyadenylation of cellular mRNAs that contain a typical AAUAAA poly(A) signal. Both our in vitro substrate (Figure 55a) and the majority of the "bulk" poly(A) tails (Figure 55b) fall into this category, so it seems likely that this is the poly(A) polymerase responsible for the activity we are monitoring. Further studies are necessary to determine whether the "noncanonical" poly(A) polymerases involved in a variety of alternative polyadenylation pathways (Schmidt & Norbury 2010) can also utilize N^6pA.

We next examined whether pol II and pol III utilize N^6pA. For pol II, we needed to experimentally distinguish between pol II and PAPα activity in cells. To do so, we tested whether the histone mRNAs, which are non-polyadenylated pol II transcripts (Marzluff et al. 2008), were labeled with N^6pA. Total RNA from N^6pA-treated cells was biotinylated, selected, and recovered transcripts were analyzed by semi-quantitative RT-PCR (Figure 56, top). H2A mRNA was efficiently recovered from N^6pA-treated cells as well as from positive control (EU) cells. In contrast, minimal H2A mRNA was observed in the negative control (DMSO). Similar results were obtained for histone HIST3H2BB mRNA (data not shown). Additionally, the pol III-transcribed 7SK RNA was detected

among the selected transcripts from EU and N^6pA-treated cells (Figure 56, bottom). From these results and those presented above, we conclude that N^6pA is incorporated into RNA by pol I, II, III and PAPα.

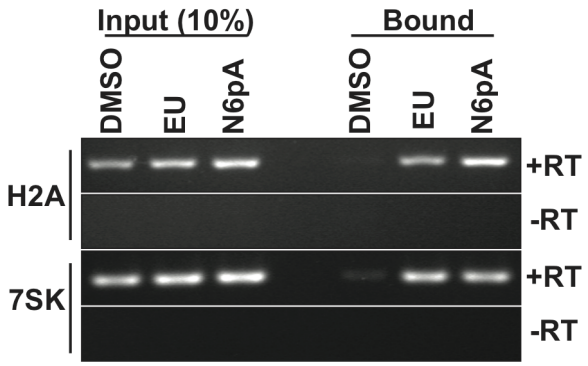


Figure 56 N^6pA is incorporated into pol II and pol III transcripts in vivo. RNA harvested from cells treated with DMSO, 200 μM EU, or 200 μM N^6pA for 12 hr was subject to click chemistry with biotin-azide and selected on a streptavidin column. 10% of the input and 100% of the bound RNA was analyzed by semi-quantitative RT-PCR using primers specific for the histone H2A mRNA or the 7SK small noncoding RNA. Negative controls lacking reverse transcriptase are also shown (-RT).

The general pathway for destruction of cytoplasmic mRNAs begins with a slow distributive deadenylation of the transcript, which is followed by nearly coincident decapping, processive deadenylation, and transcript decay (Parker & Song 2004; Yamashita et al. 2005). In most cases, distributive deadenylation has been observed using specific reporter transcripts driven from inducible promoters, but has not been demonstrated at the bulk level. Changes in bulk poly(A) lengths have been limited to steady-state observations primarily due to the lack of simple approaches for global pulse-chase analysis of poly(A) tail lengths. While steady-state analysis can be informative

(Apponi et al. 2010), interpreting poly(A) tail length differences in a pool of transcripts that contains both newly synthesized transcripts and older transcripts can be difficult. We therefore developed a pulse-chase strategy using EU or N^6pA to study bulk poly(A) tail dynamics in cells. HEK293 cells were pulse-labeled with 200 μM EU or N^6pA for 2 hr, washed, and chased with media lacking modified nucleoside. Cytoplasmic RNA was extracted from cells at given time intervals following the chase. Metabolically labeled transcripts were biotinylated by CuAAC, captured on streptavidin columns, and digested with RNase T1. Northern blot analysis of the resulting poly(A) tails showed that after a 2-hr pulse (t=0), the cytoplasmic poly(A) tails were ~150-200 nt long (Figure 57). Importantly, the tails shortened over time and decreased in their abundance, consistent with the model that most cytoplasmic mRNAs are gradually deadenylated prior to their destruction. This assay has the advantage of observing RNAs transcribed from endogenous promoters in the absence of general transcription inhibitors, which often have pleiotropic effects (Loflin et al. 1999). The ability to detect bulk poly(A) tails after a short labeling pulse allows us to examine poly(A) dynamics of transcripts that were synthesized within a 2-hr window, thereby limiting the heterogeneity of poly(A) tail length that arises over time. We conclude that alkyne-nucleosides and click chemistry can be employed to examine poly(A) tail dynamics in cells. In principle, this approach could be extended to examine poly(A) tail dynamics of specific cellular mRNAs using previously described RT-PCR techniques to detect poly(A) tail length (Sallés et al. 1999).

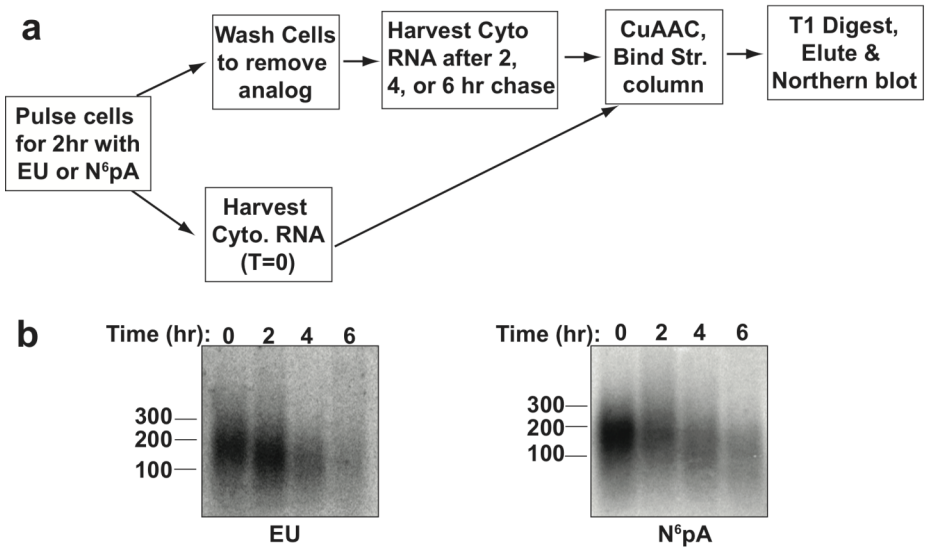


Figure 57 Pulse-chase assay to analyze poly(A) tail dynamics. a) Experimental flow. Cells were labeled for two hours with EU or N^6pA and chased for 2, 4, or 6 hr prior to cytoplasmic RNA extraction. Poly(A) tails of captured transcripts were then analyzed by northern blot using an oligo dT probe. b) Data from an in vivo poly(A) tail dynamics assay.

In summary, we report N^6pA is incorporated by all three cellular RNA polymerases (pol I, pol II, pol III) in mammalian cells. Moreover, N^6pA is added to poly(A) tails *in vitro* and *in vivo*, demonstrating that this chemical reporter is utilized by poly(A) polymerase. Furthermore, we showed that clickable nucleosides can be employed to monitor cellular poly(A) tail dynamics. These alkyne-adenosine reporters should provide new tools to monitor RNA metabolism *in vitro* and *in vivo* using bioorthogonal chemistry (Boyce & Bertozzi 2011) and Raman microscopy (Yamakoshi et al. 2011).

7.4 Materials and methods

7.4.1 Metabolic labeling of HeLa cells and visualization by fluorescence microscopy

For metabolic labeling experiments in mammalian cell culture, 5 x 10^4 HeLa cells were split into 12 well plates on glass cover slips 24 hours before the experiment. HeLa cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) with 10 % fetal bovine serum (FBS). Cells were washed once with phosphate-buffered saline (PBS) and then incubated with N^6pA in culture media for 12 hours at 37 °C at 5% CO_2. Actinomycin D (Sigma Aldrich) or hydroxyurea (Fisher Scientific) was added to the growth medium together with N^6pA at 2 μM or 10 mM, respectively.

After metabolic labeling, the cells were washed three times with PBS and subsequently fixed and permeabilized with 3.7% formaldehyde in PBS with 0.2% Triton X-100 for 30 min at room temperature. Then, cells were washed three times for 5 min each with PBS.

To detect incorporation of N^6pA, fixed cells were reacted with an azide-rhodamine (az-rho) fluorescence dye. For this purpose, click chemistry master mix (300 μl) was added to each well for a 1 h incubation period at room temperature. The master mix was prepared in PBS by addition of azido-rhodamine (Charron et al. 2009) (N-(9-(2-(4-(6-azidohexanoyl)piperazine-1-carbonyl)phenyl)-6-(diethylamino)-3H-xanthen-3-ylidene)-N-ethylethanaminium chloride, from a 10 mM stock solution in DMSO) to a final concentration of 20 μM, TCEP (tris(2-carboxyethyl)phosphine, from a 50 mM freshly prepared aqueous solution) to a final concentration of 1 mM, TBTA (tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine, from a 2 mM solution in 4:1 n-butanol:DMSO) to a final

concentration of 100 μM, and $CuSO_4$ (from a 50 mM freshly prepared aqueous solution) to a final concentration of 1 mM. The components were added in the order listed and the mix was vigorously vortexed after the addition of each component. After click chemistry, the cells were washed three times for 5 min each. Cells were incubated with TO-PRO-3 DNA dye (1:1000 in PBS) for 30 min at room temperature. Cells were washed three times for 5 min each. Cover slips were mounted on glass slides in ProLong Gold Mounting medium (Invitrogen). The images were taken on an upright LSM510 laser scanning confocal microscope (Zeiss) equipped with a Krypton/Argon laser with 488 and 568 lines and a HeNe laser with a 633 line, using a 100x oil immersion objective.

7.4.2 CuAAC reaction and streptavidin selection of alkyne-modified RNA

Modified transcripts were selected using the Click-iT Nascent RNA Capture Kit (Invitrogen) with several alterations to the manufacturer's protocol. Primarily, we scaled down the procedure to cost-effectively accommodate the number of samples required for pulse-chase analyses. Additionally, the elution and detection schemes were experiment-specific. A typical click reaction contained from 100 ng - 3 μg RNA, 10 μl of Click-iT EU Buffer, 2 mM $CuSO_4$, 250 nM biotin-azide, 0.5 μl of Click-iT reaction additive 1, and 0.5 μl of Click-iT reaction additive 2 in a total volume of 20 μl. As per the manufacturer's instructions, the Click-iT reaction additives were added last; additive 2 was added 3 min after additive 1. Generally, we used the higher range of RNA concentrations (~1-3 μg) from cells treated for shorter times with ethynyluridine (EU, Invitrogen) or N^6pA. The reaction proceeded for 30 min, and was mixed tapping the tube every ~10 min. Subsequently, 20 μl of 10M NH_4OAc, 15 μg of GlycoBlue (Ambion), and 350 μl of ethanol were added and the RNA was precipitated overnight at -20 °C.

Dynabeads streptavidin beads were washed three times in an equal volume of Wash Buffer 2; 10 µl of slurry was used per click reaction. After the final wash, the beads were resuspended in 20 µl of Wash Buffer 2. Precipitated RNA from the click reaction was washed with 70% ethanol and resuspended in 5 µl H_2O. 25 µl of 2X binding buffer supplemented with 10 U RNase inhibitor (Promega) was mixed with the RNA suspension and heated to 65 °C for 5 min prior to the addition of 20 µl bead suspension. The resulting 50 µl binding reactions were incubated at room temperature for 60 min with gentle vortexing every ~10 min to keep beads from settling. The beads were subsequently washed three times with 100 µl of Click-iT reaction wash buffer 1 and then three times with 100 µl of Click-iT reaction wash buffer 2. Selected RNAs were eluted in an experiment-specific manner.

7.4.3 In vitro polyadenylation

In vitro polyadenylation reactions were performed using pre-cleaved substrates based on previously described protocols (Wilusz & Shenk 1988). The gel-purified, uniformly ^{32}P-labeled substrate contains the 3´-most 247 nt of PANΔ79 RNA (Conrad et al. 2007) including the AAUAAA hexamer and ends at the natural poly(A) addition site. Substrates were generated by in vitro transcription with bacteriophage T7 RNA polymerase using general laboratory protocols. The template was made by PCR amplification of PANΔ79 with primers NC137 (5´-cgtaatacgactcactatagggcagctgggagcgctctttcaatg-3´) and NC138 (5´-tatggattaaacattgacctttatttatgttgtaagttgc-3´). Excluding contributions from nuclear extract, each 25 µl polyadenylation reaction contained ~25 fmol of substrate, 20 mM creatine phosphate, 1 mM DTT, 2% (v/v) polyvinyl alcohol, 1 mM $MgCl_2$, 20 µg/ml tRNA, 20 U RNase Inhibitor, and 45% HeLa nuclear extract. The nuclear extract

contained 20 mM HEPES (pH 7.9), 20% glycerol, 50 mM KCl, 0.2 mM EDTA, and 0.5 mM DTT (Dignam et al. 1983). ATP and N^6pATP was added as indicated in Figure 2. After 60 min at 30 °C, reactions were stopped by the addition of 200 µl of 0.1 mg/ml proteinase K, 0.1% (v/v) SDS, 15 µg GlycoBlue, 20 µg torula yeast RNA (Sigma), and incubated for an additional 30 min at 37 °C. RNAs were phenol-chloroform extracted and ethanol precipitated using standard laboratory techniques.

After precipitation, the entire reaction was subjected to click chemistry and streptavidin selection as described above. After washing, the bound transcripts were eluted in two steps. First, the beads were heated to 90 °C in 50 µl of a solution of 98% formamide and 10 mM EDTA. Second, after removal of the eluted formamide solution, the beads were stripped by incubation for 60 min at 50 °C in 250 µl of a solution containing 0.1% SDS, 0.1 mg/ml proteinase K, 20 mM Tris-HCl (pH 6.8), and 22.5 µg of GlycoBlue. The formamide and proteinase K elutions were combined, 30 µl of 3M NaOAc was added, and the samples were phenol-chloroform extracted and ethanol precipitated using standard laboratory techniques. Input and eluted RNAs were run on 6% urea-PAGE, the gel was dried, and substrate was detected by Phosphorimager (Molecular Dynamics).

7.4.4 Metabolic labeling of poly(A) tails in vivo

To determine whether N^6pA was incorporated into poly(A) tails in vivo, HEK293 were grown in DMEM supplemented with 10% fetal bovine serum, 1X penicillin-streptomycin (Sigma), 2 mM L-glutamate, and 200 µM of EU, N^6pA, or DMSO as a control. After overnight incubation, a slightly lower confluence of cells treated with N^6pA was observed, suggesting a minor growth defect compared to the DMSO control. Total RNA

was extracted using Tri-Reagent (Molecular Research Center) following the manufacturers protocol and treated with RQ1 DNase (Promega). One μg total RNA was digested in 10 μl reactions containing 10 mM Tris-HCl (pH 7.5), 200 U RNase T1 (Ambion), 20 U RNase Inhibitor for 30 min at 37 °C. This treatment is sufficient to completely digest the RNA as a 5-fold range of RNase T1 concentrations yielded similar digestion products (data not shown). After digestion, the reaction was stopped by the addition of 190 μl of 300 mM NaOAc, 0.1 mg/ml proteinase K, 0.25% SDS, 2 mM EDTA, 20 mM Tris-HCl (pH 7.0) for 30 min at 37 °C. The RNA was then phenol-chloroform extracted and ethanol precipitated using standard lab techniques. 10% of the RNA was set aside as "Input" samples and the remaining RNA was selected using click chemistry and streptavidin as described above. Elution protocol was the same as used for the in vitro poly(A) selection.

7.4.5 Metabolic labeling and detection of histone mRNAs and 7SK RNA

Labeling and extraction of RNA was performed as described for the in vivo poly(A) labeling assay. One μg of total RNA was used for click chemistry and streptavidin selection as described above. After selection and washing, the selected RNA was reverse transcribed on the beads. To the washed pellets, we added 25 μl of a solution containing 100 ng random hexamers (Sigma) and 10 mM solution of all four dNTPs. The mix was heated for 5 min at 65 °C, briefly chilled on ice. To this solution, 8 μl of 5X first strand buffer (Invitrogen), 4 μl of 100 mM DTT and 40 U RNase Inhibitor were added for a total of 38 μl. Half of this mix was added to 1 μl (200 U) of M-MuLV reverse transcriptase (New England Biolabs) while the other half was mixed with 1 μl H_2O as a "no-RT" control. Reverse transcription proceeded at 25 °C for 15 min and then at 42 °C

for 1 hr. After the reverse transcription, 1 U of RNase H (Promega) and 5 ng RNase A (Sigma) was added to the reaction for 30 min at 37 °C. The reaction was stopped by the addition of 180 μl of a solution containing 300 mM NaOAc, 0.25% SDS, 2 mM EDTA, 20 mM Tris-HCl (pH 7.0) and 15 μg GlycoBlue. The cDNA was then phenol-chloroform extracted and ethanol precipitated using standard lab techniques. PCR was performed using 1/100th of the cDNA as template in a standard 20 μl PCR reaction with primers: NC996 (5´ gtctggtcgtggcaaacaag-3´) and NC997 (5´-tgcgcgtcttcttgttgtcc-3´) for histone H2A, NC998 (5´ atgccagacccgtccaaatc-3´) and NC999 (5´-tggtggagcgcttgttgtag-3´) for histone HIST3H2BB and ID15 (5´- taagagctcggatgtgagggcgatctg -3´) and ID16 (5´-cgaattcggagcggtgagggaggaag -3´) for 7SK RNA.

7.4.6 In vivo poly(A) tail dynamics assay

HEK293 cells were grown to ~80% confluency in 6-well plates and 200 μM EU, 200 μM N^6pA, or DMSO was added directly to the media. After 2 hr, the cells were washed twice with PBS containing calcium and magnesium (Sigma); cells were harvested at time intervals shown in Figure 4. To collect cytoplasmic RNA, cells were resuspended in Buffer I (0.32 M sucrose, 3mM $CaCl_2$, 2 mM $MgCl_2$, 0.1 mM EDTA, 10 mM Tris-HCl (pH 8.0), 1 mM DTT, 0.04 U/μl RNase Inhibitor, 0.5% Triton X-100), incubated on ice for 5 min, centrifuged at 500 x g for 3 min at 4 °C, and RNA in the supernatant was extracted using TriReagent followed by an additional phenol-chloroform extraction. One μg of cytoplasmic RNA was used for click chemistry and streptavidin selection as described above. After binding and washing, the beads were treated with 30 μl RNase T1 solution (20 mM Tris-HCl (pH 6.8), 8U RNase Inhibitor, 150 U RNase T1) for 15 min at

37°. Subsequently, RNA was eluted as for the *in vitro* poly(A) selection assay and poly(A) tails were detected by northern blotting.

7.4.7 Northern blot analysis

To examine bulk poly(A) tails by northern blot, the RNA was run on 1.8% formaldehyde-agarose gels using standard laboratory techniques and transferred to Zeta-Probe nylon membranes (Bio-Rad). An RNA marker (0.1-2kb; Invitrogen) was run as a size standard. The blots were probed 3hr to overnight at 42 °C in ExpressHybe (Clontech) with a 5´ end-labeled oligonucleotide dT_{40} probe. The membranes were washed twice in 2X saline-sodium citrate buffer, 0.1% SDS, twice in 1X SSC, 0.1% SDS at room temperature and detected by Phosphorimager.

8 Discussion and future outlook

In the past decade bioorthogonal chemistries and the resulting idea of chemical reporters has effectively shaped the modern field of chemical biology. The strategy of using highly selective non-perturbing, or at least selectively perturbing, chemistry to investigate basic biological phenomena, develop new pharmacophores, and identify drug targets was truly invigorated or maybe even completely new defined by a set of new chemical transformations. This thesis work set out to contribute to this still growing field by investigating the possibilities chemical reporters provide to analyze pathogen-host interactions. While I recognize the current limitations of these approaches, some of the discoveries I have made during my thesis work have strong potential to contribute to the current line of investigation and transcend the topic of bacterial pathogenesis. Most notably, various other laboratories are already applying the chemical reporters I have developed for protein AMPylation and RNA synthesis to identify new AMPylation substrates and investigate mRNA polyadenylation. Other explorations, as for example the serendipitous discovery of a somewhat intrinsically cell-selective amino acid reporter (i.e. EME), on the other hand, are only first steps towards more valuable research tools.

8.1 Development of cell-selective amino acid reporters for the investigation of bacterial protein synthesis and secretion during infection

Bacterial pathogens dramatically change their gene expression pattern during the course of an infection (Eriksson et al. 2003; Shi et al. 2006). It is believed that many genes that are selectively expressed during infection, contribute to disease. The identification of these genes or proteins has been a major interest in the field of bacterial pathogenesis, especially in face of more rapidly emerging multi-resistant strains and the growing need

for new antibiotics. Many bacterial pathogens, such as *S. typhimurium,* secrete effector proteins into the mammalian host cell, which are important factors of infectivity, through elaborate and complex secretion apparatuses. These secretion systems and their substrates are of major interest for the development of new antibiotics. Proteomics has the theoretic potential to identify bacterial proteins, selectively expressed or secreted during infection and their respective posttranslational modifications. Proteomic interrogation of the mammalian cytosol should reveal secreted bacterial effector proteins. Modern mass spectrometry-based proteomics also has the potential to allow a quantitative and dynamic analysis of these processes (J. Cox & Mann 2007). However, the confounding problem of the direct proteomic interrogation of the bacterial proteome during infection is the abundance and complexity of the mammalian proteome, which, by today's dynamic range standards of mass spectrometry, makes robust detection of the bacterial proteome, especially of the secreted fraction, very difficult or even impossible (C. G. Zhang et al. 2005).

To address this issue we developed a cell-selective labeling strategy, based on previous work of the Tirell group (Tanrikulu et al. 2009; Ngo et al. 2009) that allows selective labeling of the *S. typhimurium* proteome during infection with alkyne (AOA) or azide (ANL) bearing amino acids (Grammel et al. 2010). These amino acids are co-translationally incorporated into the bacterial proteome upon expression of a mutant methionyl-tRNA synthetase (MetRS) and replace methionine (Met) throughout the bacterial proteome. The incorporated alkyne or azide groups serve chemical reporters that allow visualization or enrichment of newly synthesized bacterial proteins. However, this approach bears two intrinsic limitations; first, only cells in which the described MetRS

mutant is active can be targeted with AOA or ANL, second only proteins that possess Met residues can be labeled.

The first limitation will likely preclude the application in many more distantly related (from *E. coli*) bacteria, even though sometimes striking cross reactivity has been observed (Ripmaster et al. 1995). It will be interesting to see, whether these *E. coli* derived MetRS mutants will be active in eukaryotic cells. This thesis presented some experimental evidence that this is the case. The translation of the *E. coli* MetRS/AOA system to eukaryotes could allow labeling of individual cell populations in multi-cellular organisms (e.g. *C. elegans*) in the future. Towards this goal, it has been recently shown that multi-cellular organisms, such as the zebrafish, can be labeled with methionine amino acid reporters that do not depend on mutant MetRS activity (Hinz et al. 2012). If the inhere used *E. coli* MetRS mutants do not possess activity in other model systems, similar approaches of protein engineering, as the one used for the identification of the current *E. coli* MetRS mutants, could be applied to the corresponding MetRS (Dieterich et al. 2006).

The second limitation was addressed in this thesis by developing *para*-ethynylphenylalanine (PEP) as a cell-selective amino acid reporter (Grammel et al. 2012). Since phenylalanine (Phe) is present at higher frequencies in many proteomes than Met, the replacement of Phe with PEP could provide better coverage of the proteome. Indeed we could detect better overall incorporation of PEP compared to AOA. A combination of AOA and PEP provided even further improved labeling intensity. Unfortunately, these results did not translate to the infection setting. Suboptimal PEP incorporation upon dual bicistronic expression of MetRS and phenylalanyl-tRNA-

synthetase (PheRS) should be overcome by generating a non-bicistronic expression system for PheRS. In general, however, this proofs that the strategy that was initially applied to ANL and *E. coli* MetRS could be also applied to other aminoacyl-tRNA-synthetases (aaRS). In fact, Met is one of the least frequent amino acids, while some other amino acids, such as alanine, can make up 10% of a proteome or more (http://www.pasteur.fr/~tekaia/aafreq.html) (Tekaia et al. 2002; Tekaia & Yeramian 2006). The engineering of an orthogonal amino acid reporter aaRS pair of one of these aaRSs could dramatically increase labeling efficiencies.

While exploring alternative alkynyl amino acid reporters of Met we discovered that EME has a certain intrinsic selectivity for *S. typhimurium* over mammalian cells. Whereas its selectivity is currently not sufficient for proteomics studies, it provides a mode for bacterial visualization during infection. In a first attempt to improve the selectivity by extending the alkyl chain by one methylene unit, our laboratory tried to improve the selectivity for *S. typhimurium*. However, this amino acid is still incorporated into the mammalian proteome. Nevertheless, these observations at least demonstrated that the differential intrinsic affinity of aaRSs of different organisms for an amino acid reporter could be exploited to achieve cell-selective labeling.

Our proteomics studies revealed that cell-selective (orthogonal) amino acid reporters can be used to enrich bacterial proteins from infected mammalian cells and identify them by mass spectrometry. Yet, even though this method provides a facile way to separate these two proteomes the dynamic range of amino acid reporter incorporation is still not completely represented by the dynamic range of the subsequent enrichment and mass spectrometry detection. Many of the identified bacterial proteins are proteins of

high abundance, such as components of the translational machinery. This became particularly clear when looking at the mammalian cytosolic fractions, which should contain secreted bacterial effector proteins. Whereas we could show that the SPI-2 effector protein SseJ is labeled with amino acid reporter during infection and effector proteins can be generally secreted by the type three secretion system (TTSS) after they have incorporated amino acid reporters, we could not detect any effector proteins in the mammalian cytosolic fraction by mass spectrometry. In these fractions we only detected the most abundant bacterial proteins also detected in the pellet fraction, even with the most selective lysis conditions.

The aim of enriching and detecting secreted effector proteins from infected mammalian cells by mass spectrometry is pushing the boundaries of the theoretical detection limit of current mass spectrometry based proteomics work-flows. Estimates for secreted bacterial effector proteins range from a few hundred to a few thousand molecules per cells, in some cases this number might even be lower (Winnen et al. 2008). Other data suggest that in a typical cell the abundance of proteins ranges from a few dozen to more then 10^6 molecules (Ghaemmaghami et al. 2003). This dynamic range has to be compared with the current dynamic range of modern mass spectrometers, that is the range of the least abundant and the most abundant molecule detected in a single experiment, which is currently 4 to 5 orders of magnitude under optimal conditions (de Godoy et al. 2008). The dynamic range of a particular mass spectrometry run is however dependent on countless factors that all have to be optimized (Andrews et al. 2011). These numbers show, that even under optimal conditions, the robust and reproducible detection of secreted bacterial effector proteins is a daunting task. The complexity of the current

workflow, that consists of the infection experiment, amino acid reporter pulse-labeling, lysis, CuAAC, affinity enrichment, proteomics sample preparation, mass spectrometry, leaves plenty of room for improvements. However, quantitative analysis of the bottlenecks in the future will be indispensable, before addressing the current shortcomings.

8.2 A chemical reporter for protein AMPylation

Protein AMPylation is emerging as an important posttranslational modification (PTM), especially in the context of pathogen-host interactions. Genome data suggest that protein AMPylation could be a ubiquitous PTM in prokaryotes as well as eukaryotes. To determine its prevalence in different organisms we have developed an alkynyl chemical reporter (N6pATP) for protein AMPylation that can be used by all known classes of AMPylation enzymes and allows unbiased detection of substrates in cell lysates at endogenous levels (Grammel et al. 2011). We demonstrated its utility by detecting the known VopS substrate Cdc42 in Hela cell lysates by mass spectrometry. Addition of recombinant AMPylation enzyme and N6pATP to cell lysates allows the enzyme specific identification of substrates, an approach that our lab has also applied to other PTMs, such as acetylation (Y.-Y. Yang et al. 2011). A general draw back of cofactor analogs as N6pATP is its charged character and the resulting lack of cell permeability. Caged cofactor analogs could circumvent this issue in the future (Ellis-Davies 2007). Caging charged phosphate groups with esters results in a neutral net charge and enhances cell permeability (Schultz et al. 1993). Such caging strategies could allow metabolic labeling with cofactor analogs in the future. In addition to the current application in cell lysates, N6pATP can be used for enzyme specific substrate identification with protein arrays. Our

collaborators are currently establishing a screening assay for AMPylation substrates using N6pATP and a nucleic acid programmable protein array (NAPPA)(Ramachandran et al. 2008). This assay provides the opportunity to screen thousands of proteins in high-throughput as potential AMPylation substrates using a combination of NAPPA and CuAAC for substrate detection by fluorescence. Preliminary data has show that this approach can be used to detect VopS substrates (RhoA, Rac3, Cdc42) and auto AMPylation (HYPE) in array format. This approach is independent of the individual expression patterns of cell lines and its reduced complexity might allow the detection of substrates that would go undetected in cell lysates.

8.2.1 An adenosine based chemical reporter

Numerous alkynyl chemical reporters for nucleic acid biosynthesis have been introduced in recent years (Salic & Mitchison 2008; Jao & Salic 2008; Neef & Luedtke 2011; Guan et al. 2011). This thesis work has shown that N^6-propargyl-adenosine (N^6pA) also serves as a chemical reporter for RNA biosynthesis, which is metabolically incorporated by RNA polymerase I, II, and III (pol I, II, or III). In contrast to previous chemical reporters of RNA biosynthesis, N^6pA is also incorporated into polyadenylation (polyA) tails of mRNA molecules. In the absence of transcriptional activity of pol I, II, and III polyA modification of mRNA could be selectively studied with N^6pA. However, the activated triphosphate N^6pATP does not seem to be an adequate substrate for polyA polymerase *in vitro*.

As is indicated by its incorporation into RNA, N^6pA is intracellularly transformed into its triphosphate N^6pATP, most likely, by the purine ribonucleoside salvage pathway. In addition to nucleic acid biosynthesis, N^6pATP could also be used as a co-factor/co-

substrate in other transformations. As we have seen, isolated AMPylation enzymes *in vitro* as well as in cell lysates use N^6pATP to modify their substrates. Hence, intracellularly formed N^6pATP might serve as a chemical reporter for potential AMPylation enzymes. Indeed, in addition to the labeling of RNA, we have also observed labeling of the proteome in mammalian and bacterial cells, when incubated with N^6pA (data not shown). However, these observed protein modifications could also be of different origin, e.g. ADP-ribosylation. Provided the intracellularly formed N^6pATP is integrated into nicotinamide adenine dinucleotide (NAD), the co-substrate for ADP-ribosyl transferases and poly ADP-ribose polymerases, by the enzyme NAD pyrophosphorylase, N^6pA could potentially serve as chemical reporter for protein ADP-ribosylation.

This shows that N^6pA is a facile chemical reporter for RNA biosynthesis and polyA mRNA modification, but potentially also for protein posttranslational modifications. This is particularly interesting due to the limited applicability (lack of cell-permeability) of the cofactor based chemical reporters N^6pATP and alkyne-modified NAD (Du et al. 2009; Jiang et al. 2010), for protein AMPylation and ADP-ribosylation, respectively.

9 Appendix

9.1 Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC)

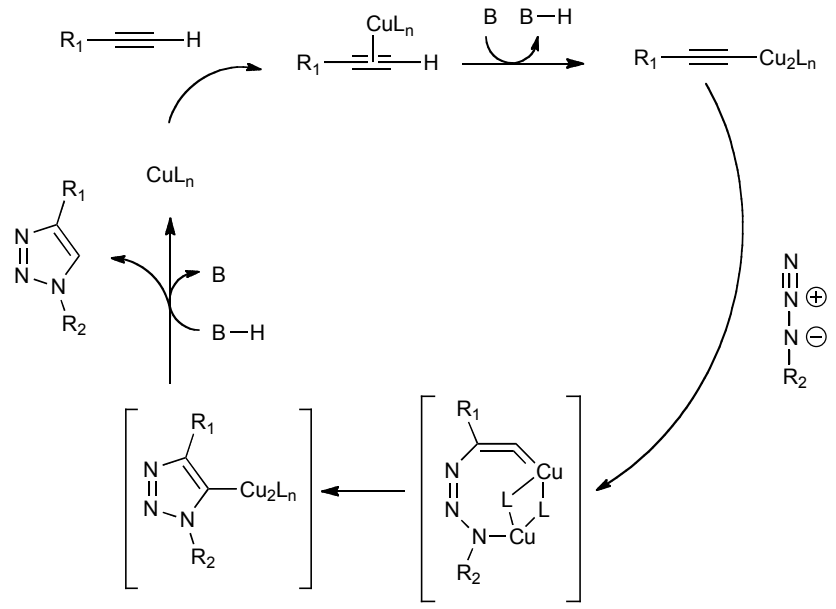


Figure 58 Current stepwise model of the Cu(I)-catalyzed azide-alkyne cycloaddition (Himo et al. 2005; Bock et al. 2006; Rodionov et al. 2005; Rostovtsev et al. 2002b).

9.2 Proteomics data

Table 3 *S. typhimurium* proteins identified from infected RAW264.7 cells (red – only detected from infected RAW264.7 cells and not from in vitro labeled *S. typhimurium*). # SC: number of spectral counts.

#	UniProt	Name	# SC
1	P0A1D3	60 kDa chaperonin	763
2	Q56073	Chaperone protein dnaK	271
3	P0A1H5	Elongation factor Tu	261
4	P0A1P0	Glyceraldehyde-3-phosphate dehydrogenase	152
5	P64076	Enolase	132
6	P0A1E3	Cysteine synthase A	122
7	P0A1H3	Elongation factor G	109
8	P58480	Chaperone protein htpG	74
9	Q8ZNN4	Methionyl-tRNA synthetase	71
10	Q7CQ01	Chaperone protein clpB	69
11	Q7DAP2	Putative carbonic anhydrase	57
12	P14146	Virulence transcriptional regulatory protein p	56
13	Q7CR42	Putative thiol-alkyl hydroperoxide reductase	51
14	Q7CPX8	Fur regulated Salmonella iron transporter	50
15	P64052	Elongation factor Ts	49
17	Q9L6N1	5-methyltetrahydropteroyltriglutamate--homocys	46
16	P0A1D5	10 kDa chaperonin	46
18	Q8ZP03	Mannose-specific enzyme IIAB	45
19	Q8ZRT0	Dihydrolipoyl dehydrogenase	44
20	Q7CPE2	ATP synthase subunit beta	43
22	P02936	Outer membrane protein A	41
23	P0A7W4	30S ribosomal protein S5	41
21	O30911	ATP-dependent hsl protease ATP-binding subunit	41
24	P26976	Non-specific acid phosphatase	41
25	P77983	Pyruvate kinase I	41
27	P62405	50S ribosomal protein L5	40
26	P0A251	Alkyl hydroperoxide reductase subunit C	40
29	Q8ZRT2	Pyruvate dehydrogenase, decarboxylase componen	39
28	Q7CQV9	DNA protection during starvation protein	39
30	Q7CPE1	ATP synthase subunit alpha	36
31	Q93DX3	Putative LysR family transcriptional regulator	34
32	P06173	DNA-directed RNA polymerase subunit beta	32
33	P66932	Trigger factor	31
34	Q7CPK0	sn-glycerol-3-phosphate-binding periplasmic pr	31
36	P65702	Phosphoglycerate kinase	30
35	P16657	Enoyl-[acyl-carrier-protein] reductase [NADH]	30
37	P0A7Z7	DNA-directed RNA polymerase subunit alpha	29
38	Q8ZRU0	Cell division protein ftsZ	29
39	P25077	Malate dehydrogenase	28
40	P60446	50S ribosomal protein L3	28
41	Q7CQT9	30S ribosomal subunit protein S1	26

#	UniProt	Name	# SC
42	Q9XCS1	Fur regulated Salmonella iron transporter	26
44	Q7CQU1	Pyruvate formate lyase I, induced anaerobicall	25
43	P0A7K0	50S ribosomal protein L11	25
45	P0A2R4	DNA-directed RNA polymerase subunit beta'	24
46	Q8ZN59	Inosine-5'-monophosphate dehydrogenase	24
47	P0A1S2	DNA-binding protein H-NS	22
48	P66541	30S ribosomal protein S2	21
49	P67093	Universal stress protein G	21
50	P0A2A3	50S ribosomal protein L1	20
51	Q8ZRB9	ATP-dependent protease La	20
52	Q8ZQS2	2,3-bisphosphoglycerate-dependent phosphoglyce	19
54	Q8ZRT1	Pyruvate dehydrogenase	18
55	Q93DX1	Putative acetyl-CoA hydrolase	18
53	P0A7V6	30S ribosomal protein S3	18
58	Q7CQN4	Major outer membrane lipoprotein 1	17
56	P0A295	Transcription termination factor rho	17
57	Q7CPU6	Fructose-bisphosphate aldolase	17
63	Q8ZP45	Iron-dependent alcohol dehydrogenase of the mu	16
60	P28354	Lysyl-tRNA synthetase	16
59	P0A2B3	30S ribosomal protein S7	16
61	P33321	Translation initiation factor IF-3	16
62	P66593	30S ribosomal protein S6	16
64	Q8ZPZ0	Isocitrate dehydrogenase [NADP]	16
69	Q8ZLM1	30S ribosomal protein S13	15
65	P0A249	Phosphoenolpyruvate-protein phosphotransferase	15
66	P0A2C9	3-oxoacyl-[acyl-carrier-protein] reductase	15
67	P26219	CDP-diacylglycerol pyrophosphatase	15
68	Q7CPZ4	Protein grpE	15
73	P65977	Protein recA	14
70	P0A297	50S ribosomal protein L10	14
72	P26982	Protease do	14
71	P14846	Carbamoyl-phosphate synthase large chain	14
74	Q8ZQE3	ATP-binding subunit of serine protease	14
78	Q7CPL5	50S ribosomal protein L14	13
77	P60726	50S ribosomal protein L4	13
75	P0A2E1	Serine hydroxymethyltransferase	13
76	P19480	Alkyl hydroperoxide reductase subunit F	13
79	Q8ZS18	Threonine synthase	13
80	P37430	Transcription elongation protein nusA	12
81	P61179	50S ribosomal protein L22	12
82	Q8ZLD5	Putative phosphatase	12
83	Q8ZRC0	ATP-dependent Clp protease ATP-binding subunit	12
84	O54297	30S ribosomal protein S4	11
85	P0A283	Glucose-specific phosphotransferase enzyme IIA	11
86	P67904	30S ribosomal protein S10	11
87	Q8ZK71	Peptide methionine sulfoxide reductase msrA	11
88	Q9F2E0	GTP-binding elongation factor family protein	11
92	Q7CPF1	Small heat shock protein ibpA	10

#	UniProt	Name	# SC
93	Q8ZLT3	Polyribonucleotide nucleotidyltransferase	10
90	P0A1Z2	Chaperone protein skp	10
89	O84944	Secretion system effector	10
91	Q7CP63	Response regulator	10
94	Q8ZRJ5	Aminoacyl-histidine dipeptidase	10
98	Q08016	Ornithine carbamoyltransferase	9
102	Q8ZK80	50S ribosomal protein L9	9
95	P0A1P6	Glutamine synthetase	9
101	Q7CQV2	Arginine 3rd transport system	9
100	Q7CQU5	Thioredoxin reductase	9
96	P0A2B1	30S ribosomal protein S20	9
97	P63411	Acetate kinase	9
99	Q7CPJ3	Putative inner membrane lipoprotein	9
103	Q8ZN40	Cysteine desulfurase	9
110	Q7CPL7	50S ribosomal protein L17	8
112	Q8ZKI4	Glucose-6-phosphate isomerase	8
104	O68883	Citrate synthase	8
108	P64222	Aminomethyltransferase	8
105	P27237	Oligopeptidase A	8
106	P63557	Acetylglutamate kinase	8
107	P63923	Phosphopentomutase	8
109	Q7CP96	Component of modulator for protease specific f	8
111	Q7CR87	Chaperone surA	8
113	Q8ZRP0	Outer membrane protein assembly factor yaeT	8
114	Q93GL9	Conjugative transfer: surface exclusion	8
117	P53636	Superoxide dismutase [Cu-Zn] 1	7
119	P66869	Succinyl-CoA ligase [ADP-forming] subunit beta	7
122	Q8ZMM8	Effector protein pipB2	7
123	Q8ZND6	Phosphotransacetylase	7
115	P0A1S4	DNA-binding protein stpA	7
125	Q8ZPG8	Putative hydrolase	7
116	P40732	Acetylornithine/succinyldiaminopimelate aminot	7
118	P65882	Adenylosuccinate synthetase	7
120	Q8ZKP7	Triosephosphate isomerase	7
121	Q8ZL96	Glycyl-tRNA synthetase beta subunit	7
124	Q8ZNW0	Pyruvate kinase II	7
126	Q8ZRB8	Peptidyl prolyl isomerase	7
127	Q8ZRP4	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N	7
132	P41033	Phosphoenolpyruvate carboxykinase [ATP]	6
137	Q8ZM63	Transketolase 1 isozyme	6
133	Q7CQM5	Secretion system regulator: transcriptonal act	6
134	Q8ZKW4	Aspartate--ammonia ligase	6
135	Q8ZL03	Small heat shock protein ibpB	6
136	Q8ZLT4	Cysteine sulfinate desulfinase	6
131	P37411	DNA gyrase subunit A	6
128	P0A1F6	Uridine phosphorylase	6
129	P0A7X0	30S ribosomal protein S8	6
130	P14062	6-phosphogluconate dehydrogenase, decarboxylat	6

#	UniProt	Name	# SC
138	Q8ZM69	D-3-phosphoglycerate dehydrogenase	6
139	P0A1D7	ATP-dependent Clp protease proteolytic subunit	5
140	P0A2M5	Uracil phosphoribosyltransferase	5
147	Q7CQ97	3-oxoacyl-[acyl-carrier-protein] synthase I	5
141	P43019	Superoxide dismutase [Mn]	5
142	P66409	30S ribosomal protein S14	5
143	P66738	Ribosome-recycling factor	5
144	P66764	S-adenosylmethionine synthetase	5
145	Q7CPC8	Putative enzyme	5
146	Q7CPY9	Transcriptional repressor of emrAB operon	5
148	Q8XFG8	Peptidyl-prolyl cis-trans isomerase	5
149	Q8ZKN5	Cystathionine gamma-synthase	5
150	Q8ZMF5	Sulfate adenylyltransferase subunit 1	5
151	Q8ZPT7	Putative enzyme	5
152	Q9ZF31	Translation initiation factor IF-2	5
154	P15111	NADP-specific glutamate dehydrogenase	4
157	P65692	6-phosphofructokinase	4
158	P66313	50S ribosomal protein L6	4
163	Q8ZKA4	Component of modulator for protease specific f	4
153	P0A263	Outer membrane protein C	4
167	Q8ZP57	Putative inner membrane protein	4
155	P26978	Bifunctional purine biosynthesis protein purH	4
156	P60428	50S ribosomal protein L2	4
159	Q7CPE7	Regulatory gene for high affinity phosphate up	4
160	Q7CPH6	Putative rhodanese-related sulfurtransferases	4
161	Q7CPM0	Acetyl CoA carboxylase	4
162	Q7CQ36	3-oxoacyl-[acyl-carrier-protein] synthase I	4
164	Q8ZKN4	Aspartokinase II	4
165	Q8ZLT0	Argininosuccinate synthase	4
166	Q8ZNV2	Aspartyl-tRNA synthetase	4
175	Q04873	UDP-glucose 6-dehydrogenase	3
176	Q8ZLX5	Ribosomal RNA large subunit methyltransferase	3
168	O33921	Glucose-1-phosphatase	3
172	P64046	Elongation factor P-like protein	3
177	Q8ZPC9	Gamma-aminobutyraldehyde dehydrogenase	3
179	Q9ZFV8	NADP-dependent malic enzyme	3
173	P64394	Integration host factor subunit beta	3
174	P67557	Glucans biosynthesis protein G	3
169	P0A2I3	DNA gyrase subunit B	3
170	P0AA19	Transcriptional regulatory protein ompR	3
171	P15434	Phenylalanyl-tRNA synthetase beta chain	3
178	Q8ZQ76	Aminopeptidase N	3
182	Q8ZPQ3	Putative transport protein	2
181	P65215	2-dehydro-3-deoxyphosphooctonate aldolase	2
180	P14845	Carbamoyl-phosphate synthase small chain	2
183	Q8ZQT5	Protein tolB	2
184	Q8ZRT7	Protein translocase subunit secA	2
185	Q9X523	Periplasmic dipeptidase	2

Table 4 *S. typhimurium* proteins identified from *in vitro* culture conditions in MgM medium. # SC: numer of spectral counts.

#	UniProt	Name	# SC
1	Q8ZQE3	ATP-binding subunit of serine protease	138
2	P64076	Enolase	136
3	Q7CPE2	ATP synthase subunit beta	105
4	P66932	Trigger factor	96
5	Q8ZRT0	Dihydrolipoyl dehydrogenase	95
6	Q7CPA1	Aspartate ammonia-lyase	89
7	Q8ZRU0	Cell division protein ftsZ	83
8	Q8ZM07	Putative alcohol dehydrogenase	76
9	Q8ZQS2	2,3-bisphosphoglycerate-dependent phosphoglyce	73
10	Q8ZRC0	ATP-dependent Clp protease ATP-binding subunit	73
11	Q8ZKP3	Glycerol kinase	70
12	P14146	Virulence transcriptional regulatory protein p	68
13	Q8ZM88	Putative periplasmic protein	66
14	P40732	Acetylornithine/succinyldiaminopimelate aminot	63
15	P63557	Acetylglutamate kinase	61
17	Q7CP63	Response regulator	60
16	P00929	Tryptophan synthase alpha chain	60
18	Q8ZJZ3	Putative NAD-dependent aldehyde dehydrogenase	55
19	Q8ZRB9	ATP-dependent protease La	54
20	P77983	Pyruvate kinase I	52
21	Q7CR62	Prolyl-tRNA synthetase	52
22	P0AA19	Transcriptional regulatory protein ompR	49
23	Q8ZMB8	Amino-acid acetyltransferase	48
24	P27237	Oligopeptidase A	47
25	Q7CQ00	Ribosome associated factor	46
26	P38039	Sulfite reductase [NADPH] flavoprotein alpha-c	45
27	Q7DAP2	Putative carbonic anhydrase	45
28	Q7CQ10	HTH-type transcriptional regulator iscR	44
29	Q8ZRJ5	Aminoacyl-histidine dipeptidase	44
30	Q8ZMF5	Sulfate adenylyltransferase subunit 1	43
31	P0A2E5	RNA polymerase sigma factor rpoS	42
32	Q8ZQT9	Cytochrome d terminal oxidase, polypeptide sub	42
33	P0A1F6	Uridine phosphorylase	41
34	P58661	Aspartate aminotransferase	40
35	P06202	Periplasmic oligopeptide-binding protein	39
36	Q8ZQW9	Phosphoglucomutase	38
38	P37430	Transcription elongation protein nusA	36
37	P15434	Phenylalanyl-tRNA synthetase beta chain	36
40	Q8ZJU9	Putative transport protein	35
39	Q7CQP3	Transcriptional activator of ntrL gene	35
41	Q8ZLZ4	Outer membrane channel	33
42	Q7CPE7	Regulatory gene for high affinity phosphate up	32
43	Q8ZPM0	Putative oxidoreductase	32
44	Q8ZQ40	Flavoprotein wrbA	32
45	P00905	Anthranilate synthase component II	31

#	UniProt	Name	# SC
47	P16657	Enoyl-[acyl-carrier-protein] reductase [NADH]	31
48	Q8ZPM7	Glutathionine S-transferase	31
46	P0A2C9	3-oxoacyl-[acyl-carrier-protein] reductase	31
52	Q8ZL96	Glycyl-tRNA synthetase beta subunit	30
51	Q7CQ36	3-oxoacyl-[acyl-carrier-protein] synthase I	30
50	P0A2F4	Superoxide dismutase [Fe]	30
53	Q8ZQL1	D-alanyl-D-alanine carboxypeptidase penicilli	30
49	P0A1V4	Adenylate kinase	30
55	Q8ZLT0	Argininosuccinate synthase	29
54	Q7CPU6	Fructose-bisphosphate aldolase	29
56	Q8ZPZ0	Isocitrate dehydrogenase [NADP]	29
57	O54297	30S ribosomal protein S4	28
58	P0A283	Glucose-specific phosphotransferase enzyme IIA	28
59	Q8ZNW8	Putative inner membrane protein	28
60	Q8ZP21	Putative homolog of glutamic dehydrogenase	28
61	Q8ZRP4	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N	28
66	Q8ZLL6	Peptidyl-prolyl cis-trans isomerase	27
62	P09384	Chemotaxis protein cheA	27
67	Q8ZP60	Phage shock protein	27
63	P33321	Translation initiation factor IF-3	27
64	P67366	UPF0227 protein ycfP	27
65	Q7CR64	UPF0325 protein yaeH	27
69	P66409	30S ribosomal protein S14	26
68	P0A249	Phosphoenolpyruvate-protein phosphotransferase	26
70	Q8ZP10	Cell division inhibitor	26
72	Q8ZLX9	Putative methyl-accepting chemotaxis protein	25
71	P66764	S-adenosylmethionine synthetase	25
77	Q7CP68	Hyperosmotically inducible periplasmic protein	24
79	Q9F2E0	GTP-binding elongation factor family protein	24
78	Q8ZL72	Aldehyde dehydrogenase B	24
76	P20753	Peptidyl-prolyl cis-trans isomerase A	24
75	P0A2T6	Catabolite gene activator	24
80	Q9ZF31	Translation initiation factor IF-2	24
74	P0A2E1	Serine hydroxymethyltransferase	24
73	P00910	Tryptophan biosynthesis protein trpCF	24
81	P0AA02	Transcription antitermination protein nusG	23
82	P22037	Protein mgtC	23
83	Q8ZNV2	Aspartyl-tRNA synthetase	23
86	Q8ZK81	30S ribosomal protein S18	22
85	P40860	Sulfate/thiosulfate import ATP-binding protein	22
84	P02910	Histidine-binding periplasmic protein	22
87	Q8ZP52	Aconitate hydratase 1	22
88	Q8ZQT5	Protein tolB	22
89	Q8ZQU2	Succinate dehydrogenase iron-sulfur subunit	22
93	P26264	3'(2'),5'-bisphosphate nucleotidase cysQ	21
96	Q8ZKB4	Fumarate reductase	21
90	P0A273	LexA repressor	21
97	Q8ZKL9	Acetylornithine deacetylase	21

#	UniProt	Name	# SC
91	P0A2B3	30S ribosomal protein S7	21
94	P60659	Biosynthetic arginine decarboxylase	21
98	Q8ZM76	Glycine dehydrogenase [decarboxylating]	21
95	P69936	NADP-dependent L-serine/L-allo-threonine dehyd	21
99	Q8ZPV9	Putative enzyme related to aldose 1-epimerase	21
92	P14846	Carbamoyl-phosphate synthase large chain	21
100	O30911	ATP-dependent hsl protease ATP-binding subunit	20
103	Q7CPI5	Putative outer membrane lipoprotein	20
104	Q8ZLQ5	Stringent starvation protein A	20
105	Q8ZNC8	Putative sugar nucleotide epimerase	20
101	P55900	Phosphoserine aminotransferase	20
102	P66738	Ribosome-recycling factor	20
106	Q8ZS18	Threonine synthase	20
108	P61179	50S ribosomal protein L22	19
110	Q8ZNW0	Pyruvate kinase II	19
109	Q7CQW0	Glutamine high-affinity transporter	19
107	P0A251	Alkyl hydroperoxide reductase subunit C	19
111	P65921	CTP synthase	18
112	Q8ZN40	Cysteine desulfurase	18
113	Q8ZQ76	Aminopeptidase N	18
114	Q8ZR95	Copper-exporting P-type ATPase A	18
115	Q8ZS19	Aspartokinase I	18
117	P0A295	Transcription termination factor rho	17
122	Q8ZK29	Probable cytosol aminopeptidase	17
123	Q8ZKF6	Acetyl-coenzyme A synthetase	17
124	Q8ZKP7	Triosephosphate isomerase	17
118	P0A7V6	30S ribosomal protein S3	17
120	P66692	Ribose-5-phosphate isomerase A	17
119	P10370	Histidinol dehydrogenase	17
125	Q8ZP02	Mannose-specific enzyme IID	17
116	P00898	Anthranilate synthase component 1	17
126	Q8ZPS9	Threonyl-tRNA synthetase	17
121	P67270	UPF0135 protein ybgI	17
138	Q9L6M9	Putative carboxymethylenebutenolidase	16
131	P66073	50S ribosomal protein L15	16
134	Q8ZLR6	Sigma cross-reacting protein 27A	16
129	P17845	Sulfite reductase [NADPH] hemoprotein beta-com	16
137	Q9F5L2	Putative serine/threonine protein kinase	16
130	P26398	Lipopolysaccharide biosynthesis protein rfbH	16
127	P06179	Flagellin	16
139	Q9ZNM2	FliY	16
135	Q8ZNZ7	Putative IclR family transcriptional repressor	16
136	Q8ZPS3	Phosphoenolpyruvate synthase	16
132	Q7CQT1	Putative cytoplasmic protein	16
133	Q8XFG8	Peptidyl-prolyl cis-trans isomerase	16
128	P0A1G5	DnaK suppressor protein	16
143	P26219	CDP-diacylglycerol pyrophosphatase	15
147	Q8ZL68	Putative glutathione S-transferase	15

#	UniProt	Name	# SC
148	Q8ZLI7	Fe/S biogenesis protein nfuA	15
145	Q7CPU8	Putative periplasmic immunogenic protein	15
140	P0A1N0	Surface presentation of antigens protein spaK	15
146	Q7CPZ4	Protein grpE	15
141	P0A1W4	GTP-binding protein lepA	15
142	P0A2D1	Oxidoreductase ucpA	15
144	P37426	Ribonucleoside-diphosphate reductase 1 subunit	15
158	Q8ZJV7	Purine nucleoside phosphorylase deoD-type	14
159	Q8ZJV8	Deoxyribose-phosphate aldolase	14
154	P65748	Inorganic pyrophosphatase	14
155	Q7CP94	Putative exoribonuclease	14
151	P0A2F6	Single-stranded DNA-binding protein	14
160	Q8ZKL6	Argininosuccinate lyase	14
161	Q8ZLQ1	Serine endoprotease	14
150	P0A2E3	RNA polymerase sigma factor rpoD	14
156	Q7CPY9	Transcriptional repressor of emrAB operon	14
162	Q8ZN59	Inosine-5'-monophosphate dehydrogenase	14
152	P14062	6-phosphogluconate dehydrogenase, decarboxylat	14
149	P02941	Methyl-accepting chemotaxis protein II	14
163	Q8ZNW2	Glucose-6-phosphate 1-dehydrogenase	14
164	Q8ZNW4	2-keto-3-deoxygluconate 6-phosphate aldolase	14
157	Q7CR25	Morphogene putative regulator of murein genes	14
153	P15875	2-isopropylmalate synthase	14
171	Q7CPD8	UPF0438 protein yifE	13
170	P64036	Elongation factor P	13
169	P63365	Phosphate import ATP-binding protein pstB	13
172	Q7CPL6	50S ribosomal protein L18	13
174	Q8ZM69	D-3-phosphoglycerate dehydrogenase	13
173	Q7CQ04	Putative tRNA/rRNA methyltransferase	13
175	Q8ZNE6	Putative chemotaxis signal transduction protei	13
165	P0A1Q0	Alpha,alpha-trehalose-phosphate synthase [UDP-	13
166	P0A2K1	Tryptophan synthase beta chain	13
167	P37439	PTS system glucose-specific EIICB component	13
176	Q8ZR22	Periplasmic disulfide isomerase, thiol-disulph	13
168	P60217	5'-methylthioadenosine/S-adenosylhomocysteine	13
189	Q8ZKR6	Response reguator in two-component regulatory	12
182	P43019	Superoxide dismutase [Mn]	12
181	P37419	Putative 8-amino-7-oxononanoate synthase/2-ami	12
184	P67904	30S ribosomal protein S10	12
186	Q7CPQ3	Putative ABC superfamily transport protein	12
187	Q7CPU7	Putative membrane protein	12
190	Q8ZML4	Putative inner membrane protein	12
188	Q7CQ54	Putative acyltransferase	12
177	P0A1R2	Imidazole glycerol phosphate synthase subunit	12
183	P60317	ProP effector	12
178	P0A2I1	DNA topoisomerase 1	12
191	Q8ZPD4	Putative NADP-dependent oxidoreductase	12
179	P0A2S0	Leucine-responsive regulatory protein	12

#	UniProt	Name	# SC
192	Q8ZRB1	Glycoprotein/polysaccharide metabolism	12
185	Q56062	Methylisocitrate lyase	12
193	Q8ZRP0	Outer membrane protein assembly factor yaeT	12
180	P14845	Carbamoyl-phosphate synthase small chain	12
195	P0A2A3	50S ribosomal protein L1	11
201	P63923	Phosphopentomutase	11
207	Q8ZJZ8	Isoaspartyl dipeptidase	11
203	Q7CP78	Putative translation initiation inhibitor	11
212	Q9F173	DNA polymerase I	11
197	P0A2H9	Thiol:disulfide interchange protein dsbA	11
208	Q8ZKX9	Putative oxidoreductase	11
196	P0A2A5	50S ribosomal protein L28	11
199	P50335	Protein yhjJ	11
202	P69066	Chaperone protein sicA	11
209	Q8ZN71	Dihydrodipicolinate synthase	11
194	P06185	Outer membrane protease E	11
210	Q8ZNF3	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate	11
200	P58663	Capsular synthesis regulator component B	11
211	Q8ZNM0	Cytidine deaminase	11
204	Q7CQE8	Putative GTP-binding protein	11
205	Q7CQK9	Putative glutaredoxin protein	11
206	Q7CQP4	Putative oxidoreductase	11
198	P0A6B1	Acyl carrier protein	11
220	Q8ZJU7	Transcriptional regulator	10
215	P60829	UPF0307 protein yjgA	10
214	P0A2I3	DNA gyrase subunit B	10
221	Q8ZL67	Mannitol-1-phosphate 5-dehydrogenase	10
222	Q8ZL95	Glycyl-tRNA synthetase alpha subunit	10
217	Q7CPN9	Putative ABC superfamily transport protein	10
223	Q8ZM41	Putative cytoplasmic protein	10
213	P0A1Y6	NADH-quinone oxidoreductase subunit C/D	10
224	Q8ZNN5	Putative ATP-binding protein	10
216	Q04866	Chain length determinant protein	10
225	Q8ZNS0	AMP nucleosidase	10
218	Q7CQF9	Putative cytoplasmic protein	10
227	Q9X523	Periplasmic dipeptidase	10
226	Q8ZPE8	NAD-dependent malic enzyme	10
219	Q7CR43	Response regulator in two-component regulatory	10
230	P0A299	50S ribosomal protein L7/L12	9
238	Q8ZK71	Peptide methionine sulfoxide reductase msrA	9
239	Q8ZKI5	Aspartokinase	9
240	Q8ZKM0	Phosphoenolpyruvate carboxylase	9
241	Q8ZKN4	Aspartokinase II	9
242	Q8ZKX1	Glucosamine--fructose-6-phosphate aminotransfe	9
234	Q02755	Methyl-accepting chemotaxis citrate transducer	9
233	P77980	3-dehydroquinate synthase	9
235	Q7CPR8	Putative SH3 domain protein	9
236	Q7CPR9	Bifunctional protein hldE	9

#	UniProt	Name	# SC
228	O68838	Glutamate--cysteine ligase	9
243	Q8ZN45	Putative sulfurtransferase	9
244	Q8ZN60	GMP synthase [glutamine-hydrolyzing]	9
245	Q8ZND4	Putative transketolase	9
231	P16328	Flagellar hook-associated protein 2	9
246	Q8ZNZ5	Putative GAF domain-containing protein	9
247	Q8ZP37	Nitrate reductase 1, alpha subunit	9
248	Q8ZPL8	Putative periplasmic protein	9
229	O84944	Secretion system effector	9
249	Q8ZPT7	Putative enzyme	9
253	Q9L328	Catalase	9
250	Q8ZQ14	3-oxoacyl-[acyl-carrier-protein] synthase II	9
232	P58696	Asparaginyl-tRNA synthetase	9
252	Q935S7	Chromosome partition protein mukB	9
237	Q7CQY3	Transcriptional repressor of iron-responsive g	9
251	Q8ZQX5	Glutaminyl-tRNA synthetase	9
259	P26978	Bifunctional purine biosynthesis protein purH	8
256	P0AA28	Thioredoxin-1	8
264	P65882	Adenylosuccinate synthetase	8
271	Q8ZKN5	Cystathionine gamma-synthase	8
268	Q7CPH6	Putative rhodanese-related sulfurtransferases	8
257	P17215	Leu/Ile/Val/Thr-binding protein	8
272	Q8ZLK3	Membrane protein	8
266	P67347	UPF0190 protein yedY	8
267	P68684	30S ribosomal protein S21	8
273	Q8ZM21	Putative glutathione S-transferase	8
274	Q8ZME2	(P)ppGpp synthetase I	8
255	P0A2A9	30S ribosomal protein S16	8
269	Q7CQ05	Autonomous glycyl radical cofactor	8
270	Q7CQ15	Putative inner membrane protein	8
262	P60802	Erythronate-4-phosphate dehydrogenase	8
263	P63411	Acetate kinase	8
254	O52325	Bifunctional polymyxin resistance protein arnA	8
260	P37592	Outer membrane porin protein ompD	8
261	P40676	Transcriptional regulator slyA	8
275	Q8ZPN7	N-ethylmaleimide reductase	8
276	Q8ZPP1	Riboflavin synthase, alpha chain	8
265	P67341	UPF0189 protein ymdB	8
277	Q8ZQL2	Putative glutathione S-transferase	8
278	Q8ZQX7	Glucosamine-6-phosphate deaminase	8
258	P19480	Alkyl hydroperoxide reductase subunit F	8
279	Q8ZR92	Putative thioredoxin protein	8
280	Q8ZRA7	RND family, acridine efflux pump	8
281	Q8ZRJ9	Homolog of sapA	8
282	Q8ZRT7	Protein translocase subunit secA	8
286	P0A1R6	DNA-binding protein HU-alpha	7
300	Q8ZJY6	Putative cobalamin synthesis protein	7
301	Q8ZK31	Valyl-tRNA synthetase	7

#	UniProt	Name	# SC
290	P26976	Non-specific acid phosphatase	7
294	Q7CPI1	PTS family, mannitol-specific enzyme IIABC com	7
302	Q8ZL99	Putative periplasmic or exported protein	7
303	Q8ZLD5	Putative phosphatase	7
292	P67686	Universal stress protein B	7
295	Q7CPJ6	Putative ATPase involved in cell division	7
289	P0A2P2	Tryptophanyl-tRNA synthetase	7
296	Q7CPM0	Acetyl CoA carboxylase	7
297	Q7CPM2	Putative oxidoreductase	7
304	Q8ZLQ6	N-acetylneuraminate lyase	7
305	Q8ZLR3	Glutamate synthase, small subunit	7
306	Q8ZM16	Hydrogenase-2, large subunit	7
293	Q56026	Cell invasion protein sipD	7
307	Q8ZN42	Chaperone protein hscA	7
288	P0A2K3	Glutamyl-tRNA synthetase	7
308	Q8ZND5	Putative transketolase	7
287	P0A1Y4	NADH-quinone oxidoreductase subunit G	7
284	P06614	HTH-type transcriptional regulator cysB	7
309	Q8ZPA8	Alcohol dehydrogenase class III	7
310	Q8ZPL7	Fumarate hydratase class II	7
298	Q7CQK7	Tyrosyl-tRNA synthetase	7
283	O85139	3-oxoacyl-[acyl-carrier-protein] synthase 3	7
311	Q8ZQF0	Pyruvate dehydrogenase/oxidase FAD and thiamin	7
312	Q8ZQX6	Sugar Specific PTS family, n-acetylglucosamine	7
299	Q7CQY5	Putative phosphatase in N-acetylglucosamine me	7
313	Q8ZQZ2	Putative cytoplasmic protein	7
291	P63223	Phosphoheptose isomerase	7
314	Q8ZRT5	GMP reductase	7
285	P0A1G7	Chaperone protein dnaJ	7
339	Q9L9I6	Regulator of sigma D	6
324	Q56121	Peptide chain release factor 3	6
317	P0A2L1	UPF0047 protein yjbQ	6
323	P74861	Aromatic-amino-acid aminotransferase	6
329	Q8ZLK4	D-ribulose-5-phosphate 3-epimerase	6
330	Q8ZLR4	Glutamate synthase, large subunit	6
321	P63343	Cell division protease ftsH	6
325	Q7CPQ6	Putative periplasmic protein	6
319	P58537	Inositol-1-monophosphatase	6
326	Q7CQ27	Putative iron-dependent peroxidase	6
327	Q7CQ39	Putative aspartate-semialdehyde dehydrogenase	6
328	Q7CQ71	50S ribosomal protein L25	6
322	P65363	Cell division topological specificity factor	6
331	Q8ZPC4	Putative outer membrane lipoprotein	6
316	P0A1A9	Exodeoxyribonuclease III	6
332	Q8ZPY6	Putative molecular chaperone	6
315	O30916	Inositol phosphate phosphatase sopB	6
333	Q8ZQC4	Cytidylate kinase	6
334	Q8ZR27	Putative glycerol dehydrogenase	6

#	UniProt	Name	# SC
320	P58688	Bifunctional protein folD	6
335	Q8ZR86	HTH-type transcriptional repressor allR	6
336	Q8ZRA6	Acridine efflux pump	6
337	Q8ZRB2	Acyl-CoA thioesterase II	6
338	Q8ZRB8	Peptidyl prolyl isomerase	6
318	P37412	3-isopropylmalate dehydrogenase	6
380	Q9L6L5	Fatty acid oxidation complex subunit alpha	5
363	Q8ZJY1	Putative PTS permease	5
364	Q8ZK42	Putative periplasmic protein	5
354	P66593	30S ribosomal protein S6	5
348	P36936	Peptidase E	5
356	P67651	Regulator of ribonuclease activity A	5
365	Q8ZL52	L-threonine 3-dehydrogenase	5
359	Q7CPJ3	Putative inner membrane lipoprotein	5
341	P05415	Glucose-1-phosphate adenylyltransferase	5
366	Q8ZLJ1	Putative RNase R	5
367	Q8ZLJ4	33 kDa chaperonin	5
345	P0A7X0	30S ribosomal protein S8	5
368	Q8ZLN3	Putative ferripyochelin binding protein	5
369	Q8ZLP6	Putative outer membrane protein	5
353	P66431	30S ribosomal protein S15	5
349	P58580	Glutathione synthetase	5
370	Q8ZMD2	Putative nucleotide binding	5
381	Q9XCS1	Fur regulated Salmonella iron transporter	5
358	P74881	Phosphoribosylformylglycinamidine synthase	5
350	P58671	4-hydroxy-3-methylbut-2-en-1-yl diphosphate sy	5
340	O52765	Histidyl-tRNA synthetase	5
347	P33901	NADH-quinone oxidoreductase subunit F	5
371	Q8ZNM5	Putative oxidoreductase	5
352	P63905	Deoxycytidine triphosphate deaminase	5
357	P74871	Arginyl-tRNA synthetase	5
360	Q7CQG1	Putative translation factor	5
361	Q7CQG2	Putative oxoacyl-(Acyl carrier protein) reduct	5
372	Q8ZP85	Putative Smr domain protein	5
373	Q8ZPD7	Putative cytoplasmic protein	5
342	P06204	Dihydroorotase	5
355	P67563	Seryl-tRNA synthetase	5
362	Q8XG72	Arginine transport system	5
374	Q8ZQQ4	UvrABC system protein B	5
375	Q8ZR26	Putative aminotransferase	5
376	Q8ZRC1	Putative lipoprotein	5
351	P63603	Shikimate kinase 2	5
343	P0A277	Xanthine phosphoribosyltransferase	5
346	P32200	Acyl-[acyl-carrier-protein]--UDP-N-acetylgluco	5
377	Q8ZRS0	Carbonic anhydrase	5
378	Q8ZRU8	S-adenosyl-L-methionine-dependent methyltransf	5
344	P0A2B1	30S ribosomal protein S20	5
379	Q93GL9	Conjugative transfer: surface exclusion	5

#	UniProt	Name	# SC
391	P40810	Dihydroxy-acid dehydratase	4
403	Q8ZJX0	Putative cytoplasmic protein	4
404	Q8ZK11	Putative ATPase involved in DNA repair	4
405	Q8ZK78	Putative cell envelope opacity-associated prot	4
397	Q7CP96	Component of modulator for protease specific f	4
406	Q8ZKB1	Phosphatidylserine decarboxylase proenzyme	4
407	Q8ZKE2	Fumarase B	4
408	Q8ZKH9	Glycerol-3-phosphate acyltransferase	4
394	P66008	Soluble pyridine nucleotide transhydrogenase	4
398	Q7CPC8	Putative enzyme	4
409	Q8ZL36	Putative inner membrane protein	4
399	Q7CPK0	sn-glycerol-3-phosphate-binding periplasmic pr	4
385	P0A1F8	Aspartate-semialdehyde dehydrogenase	4
392	P60428	50S ribosomal protein L2	4
382	O54296	30S ribosomal protein S11	4
383	P0A1B3	Arginine repressor	4
410	Q8ZLQ2	Putative periplasmic protein	4
387	P0AA13	50S ribosomal protein L13	4
411	Q8ZLR0	Putative cytosine deaminase	4
400	Q7CPQ5	Putative intracellular proteinase	4
388	P17853	Phosphoadenosine phosphosulfate reductase	4
396	Q56019	Cell invasion protein sipB	4
384	P0A1B5	Phospho-2-dehydro-3-deoxyheptonate aldolase, T	4
412	Q8ZMW8	Putative lipoprotein	4
413	Q8ZMX7	ATP-dependent RNA helicase	4
414	Q8ZN19	Pyridoxine 5'-phosphate synthase	4
386	P0A9Z4	Nitrogen regulatory protein P-II 1	4
393	P65889	Phosphoribosylaminoimidazole-succinocarboxamid	4
415	Q8ZN82	Transketolase 2 isozyme	4
416	Q8ZNC2	Amidophosphoribosyltransferase	4
390	P37604	D-alanyl-D-alanine carboxypeptidase dacD	4
425	Q9RM66	Universal stress protein C	4
417	Q8ZP20	Periplasmic trehalase	4
418	Q8ZP34	Response regulator in two-component regulatory	4
419	Q8ZP42	Formyltetrahydrofolate hydrolase	4
420	Q8ZPF0	Putative alpha amylase	4
395	P67197	Putative phosphotransferase ydiA	4
421	Q8ZQR7	Molybdate transporter	4
389	P22715	UDP-glucose 4-epimerase	4
401	Q7CQX4	2-oxoglutarate dehydrogenase	4
422	Q8ZRD1	1-deoxy-D-xylulose-5-phosphate synthase	4
402	Q7CR87	Chaperone surA	4
423	Q93GQ7	Plasmid partition protein B	4
424	Q93GS8	DNA replication	4
440	Q8ZK91	Putative hydrolase of the alpha/beta superfami	3
430	P51066	Isocitrate lyase	3
434	P67912	ADP-L-glycero-D-manno-heptose-6-epimerase	3
435	Q7CPP7	50S ribosomal protein L21	3

#	UniProt	Name	# SC
436	Q7CPS3	Putative cytoplasmic protein	3
441	Q8ZM45	Nucleoside-triphosphatase rdgB	3
437	Q7CPZ7	4.5S-RNP protein	3
426	P0A1W2	Signal peptidase I	3
431	P63423	Acetyltransferase ypeA	3
442	Q8ZN89	DNA ligase	3
427	P0A1Y8	NADH-quinone oxidoreductase subunit E	3
432	P64209	GTP cyclohydrolase 1	3
429	P26397	CDP-glucose 4,6-dehydratase	3
443	Q8ZP18	Putative resistance protein MccF	3
438	Q7CQF3	Thymidine kinase	3
444	Q8ZP75	Putative cytoplasmic protein	3
445	Q8ZPF1	Putative glycosyl hydrolase	3
449	Q9Z5Z2	Oxygen-insensitive NADPH nitroreductase	3
446	Q8ZQY0	Asparagine synthetase B	3
439	Q7CQY7	Putative phosphate starvation-inducible protei	3
433	P65792	Glutamate 5-kinase	3
447	Q8ZRM7	2,5-diketo-D-gluconic acid reductase B	3
428	P21267	Glutamate-1-semialdehyde 2,1-aminomutase	3
448	Q8ZRW0	LPS-assembly protein lptD	3
450	P0A2Q4	HTH-type transcriptional regulator metR	2
461	Q7CP84	Putative peptide maturation protein	2
462	Q7CP91	Peptidyl-prolyl cis-trans isomerase	2
454	P37413	Homoserine O-succinyltransferase	2
453	P17750	Catalase-peroxidase	2
457	P41789	Nitrogen regulation protein NR(I)	2
465	Q8ZKV4	Protein for a late step of protoheme IX synthe	2
466	Q8ZL61	L-lactate dehydrogenase [cytochrome]	2
472	Q93IN1	Cyclic di-GMP-binding protein	2
467	Q8ZLD4	Glutathione oxidoreductase	2
463	Q7CPJ7	RNA polymerase sigma factor	2
458	P65510	Glutathione-regulated potassium-efflux system	2
451	P0A7S6	30S ribosomal protein S12	2
460	Q56020	Cell invasion protein sipC	2
455	P37431	3-demethylubiquinone-9 3-methyltransferase	2
452	P10372	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino	2
456	P37600	Thiosulfate reductase	2
459	Q06399	Uncharacterized lipoprotein yedD	2
468	Q8ZPC0	Acetyl transferase	2
469	Q8ZPU3	PTS family sugar specific enzyme III for cello	2
470	Q8ZR02	D-alanyl-D-alanine carboxypeptidase, penicilli	2
471	Q8ZR16	RNase I	2
464	Q7CR52	Sigma factor-binding protein crl	2

Table 5 Identified *S. typhimurium* proteins in the Triton X-100 supernatant fraction upon in-gel digestion.

UniProt	Name	# SC
P0A1H5	Elongation factor Tu	66
Q56073	Chaperone protein dnaK	64
P58480	Chaperone protein htpG	34
Q7CQ01	Chaperone protein clpB	33
P0A1H3	Elongation factor G	15
P64052	Elongation factor Ts	14
P0A249	Phosphoenolpyruvate-protein phosphotransferase	14
P0A1P0	Glyceraldehyde-3-phosphate dehydrogenase	13
O30911	ATP-dependent hsl protease ATP-binding subunit	13
P64076	Enolase	12
P0A1D5	10 kDa chaperonin	12
Q7CPK0	sn-glycerol-3-phosphate-binding periplasmic pr	11
Q7CR42	Putative thiol-alkyl hydroperoxide reductase	9
Q7CQ36	3-oxoacyl-[acyl-carrier-protein] synthase I	9
Q8ZS18	Threonine synthase	8
P65702	Phosphoglycerate kinase	8
P26976	Non-specific acid phosphatase	8
P14062	6-phosphogluconate dehydrogenase, decarboxylat	8
Q8ZPZ0	Isocitrate dehydrogenase [NADP]	7
Q8ZP49	UPF0259 membrane protein yciC	7
Q8ZL96	Glycyl-tRNA synthetase beta subunit	7
P0A7Z7	DNA-directed RNA polymerase subunit alpha	7
Q8ZRU0	Cell division protein ftsZ	6
Q8ZRS8	Aconitate hydratase 2	6
Q7CPE1	ATP synthase subunit alpha	6
P14146	Virulence transcriptional regulatory protein p	6
P0A2C9	3-oxoacyl-[acyl-carrier-protein] reductase	6
Q8ZM63	Transketolase 1 isozyme	5
Q7CR87	Chaperone surA	5
Q7CQ11	NifU homolog	5
P77983	Pyruvate kinase I	5
P66738	Ribosome-recycling factor	5
P65977	Protein recA	5
P40732	Acetylornithine/succinyldiaminopimelate aminot	5
P0A1E3	Cysteine synthase A	5
Q9F2E0	GTP-binding elongation factor family protein	4
Q8ZNV2	Aspartyl-tRNA synthetase	4
Q7CPU6	Fructose-bisphosphate aldolase	4
P66764	S-adenosylmethionine synthetase	4
P64281	Transcription elongation factor greA	4
P25077	Malate dehydrogenase	4
P0A2E1	Serine hydroxymethyltransferase	4
Q7CPE2	ATP synthase subunit beta	3
P37430	Transcription elongation protein nusA	3

UniProt	Name	# SC
P0A7W4	30S ribosomal protein S5	3
P0A7K0	50S ribosomal protein L11	3
O85140	Malonyl CoA-acyl carrier protein transacylase	3
Q9L6N1	5-methyltetrahydropteroyltriglutamate--homocys	2
Q8ZQT5	Protein tolB	2
Q8ZQS4	Putative ABC transport protein	2
Q8ZPG8	Putative hydrolase	2
Q7CQ22	Putative PerM family permease	2
P65692	6-phosphofructokinase	2
P14846	Carbamoyl-phosphate synthase large chain	2
P0A9X6	Rod shape-determining protein mreB	2

Table 6 Identified *S. typhimurium* proteins in the Triton X-100 pellet fraction upon in-gel digestion.

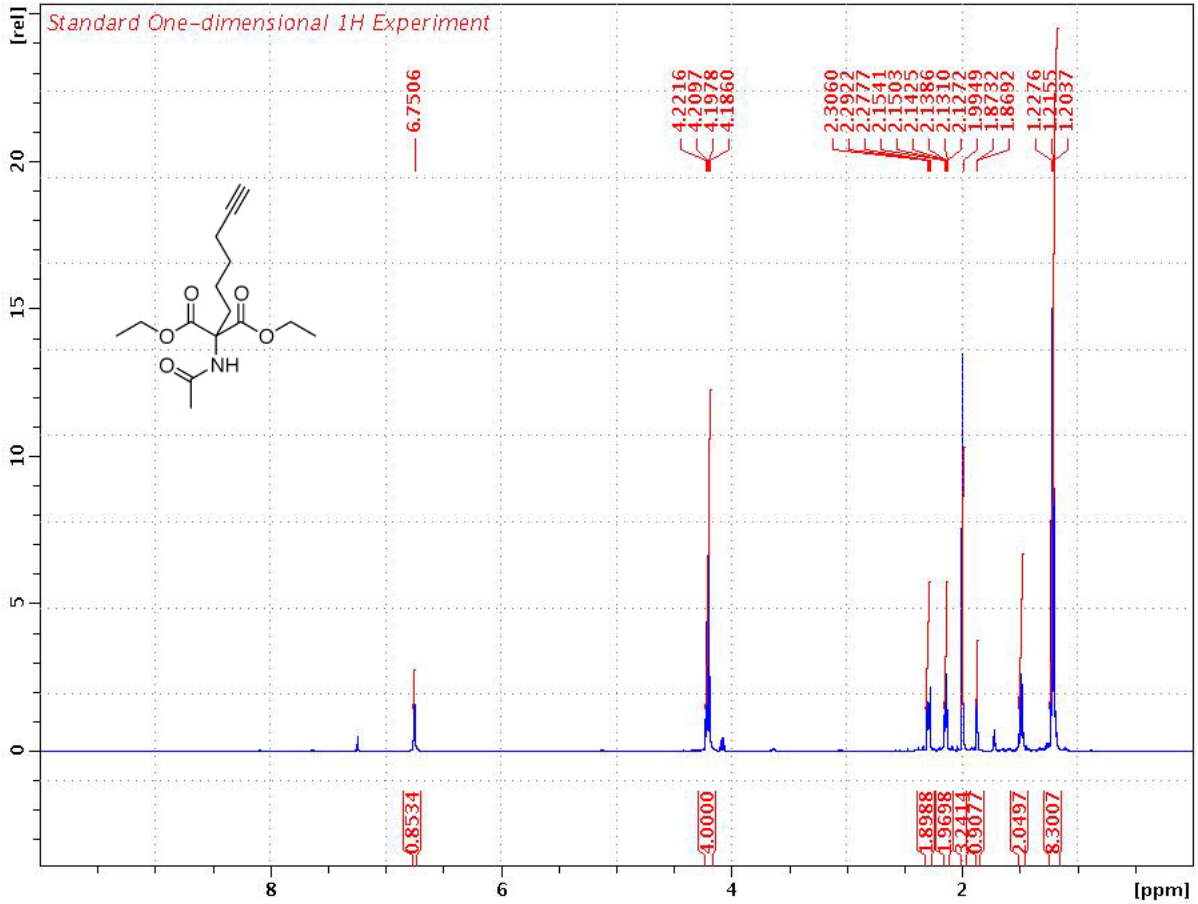
UniProt	Name	#SC
Q56073	Chaperone protein dnaK	143
Q7CQ01	Chaperone protein clpB	107
P06173	DNA-directed RNA polymerase subunit beta	91
P0A2R4	DNA-directed RNA polymerase subunit beta'	83
P0A1H3	Elongation factor G	76
P0A1H5	Elongation factor Tu	73
P58480	Chaperone protein htpG	65
Q8ZRT1	Pyruvate dehydrogenase	57
Q8ZRT2	Pyruvate dehydrogenase, decarboxylase componen	39
Q8ZQT9	Cytochrome d terminal oxidase, polypeptide sub	37
P0A1P0	Glyceraldehyde-3-phosphate dehydrogenase	34
P64076	Enolase	32
Q8ZNN4	Methionyl-tRNA synthetase	31
P0A1S2	DNA-binding protein H-NS	27
P63343	Cell division protease ftsH	27
Q7DAP2	Putative carbonic anhydrase	27
P0A7W4	30S ribosomal protein S5	26
P02936	Outer membrane protein A	25
Q8ZP45	Iron-dependent alcohol dehydrogenase of the mu	25
P64052	Elongation factor Ts	23
Q7CQT9	30S ribosomal subunit protein S1	23
P0A1D5	10 kDa chaperonin	22
P0A7K0	50S ribosomal protein L11	22
P14846	Carbamoyl-phosphate synthase large chain	21
P0A7V6	30S ribosomal protein S3	19
Q8ZLT3	Polyribonucleotide nucleotidyltransferase	19
P0A1S4	DNA-binding protein stpA	18
P0A2A3	50S ribosomal protein L1	18
Q8ZND6	Phosphotransacetylase	18
P0A7Z7	DNA-directed RNA polymerase subunit alpha	17
Q7CQU1	Pyruvate formate lyase I, induced anaerobicall	17
Q8ZRB9	ATP-dependent protease La	17
O30911	ATP-dependent hsl protease ATP-binding subunit	16
P40732	Acetylornithine/succinyldiaminopimelate aminot	16
P60446	50S ribosomal protein L3	16
P66541	30S ribosomal protein S2	16
P66593	30S ribosomal protein S6	16
Q9L6N1	5-methyltetrahydropteroyltriglutamate--homocys	16
P26976	Non-specific acid phosphatase	15
P65702	Phosphoglycerate kinase	15
Q7CR42	Putative thiol-alkyl hydroperoxide reductase	14
Q7CPE2	ATP synthase subunit beta	13
Q8ZL96	Glycyl-tRNA synthetase beta subunit	13
Q8ZRS8	Aconitate hydratase 2	13
P37411	DNA gyrase subunit A	12

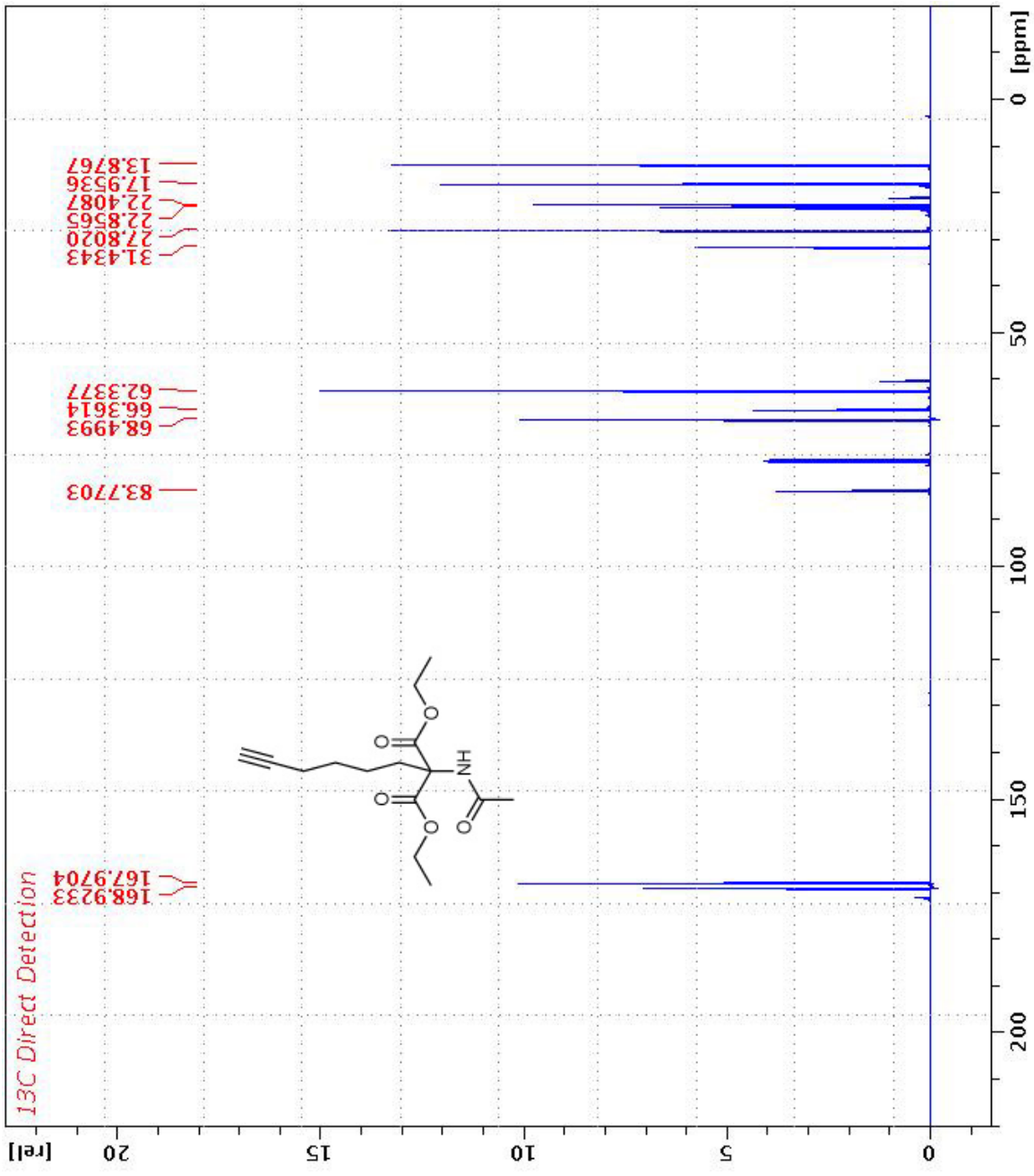
Q7CPK0	sn-glycerol-3-phosphate-binding periplasmic pr	12
Q7CPL7	50S ribosomal protein L17	12
Q7CQ36	3-oxoacyl-[acyl-carrier-protein] synthase I	12
Q9ZF31	Translation initiation factor IF-2	12
P0AA13	50S ribosomal protein L13	11
P14146	Virulence transcriptional regulatory protein p	11
Q8ZNV2	Aspartyl-tRNA synthetase	11
Q8ZP03	Mannose-specific enzyme IIAB	11
P0A1G7	Chaperone protein dnaJ	10
P26219	CDP-diacylglycerol pyrophosphatase	10
P60726	50S ribosomal protein L4	10
P65977	Protein recA	10
Q8ZK80	50S ribosomal protein L9	10
Q8ZM76	Glycine dehydrogenase [decarboxylating]	10
P0A297	50S ribosomal protein L10	9
P0A2B3	30S ribosomal protein S7	9
P26353	Flavohemoprotein	9
P62405	50S ribosomal protein L5	9
Q7CQV9	DNA protection during starvation protein	9
Q8ZM63	Transketolase 1 isozyme	9
Q8ZQE3	ATP-binding subunit of serine protease	9
P0A1E3	Cysteine synthase A	8
P0A9X6	Rod shape-determining protein mreB	8
P15111	NADP-specific glutamate dehydrogenase	8
P22036	Magnesium-transporting ATPase, P-type 1	8
Q7CPE4	ATP synthase subunit b	8
Q7CR87	Chaperone surA	8
Q9F2E0	GTP-binding elongation factor family protein	8
Q9ZFV8	NADP-dependent malic enzyme	8
P0A1X0	Outer membrane lipoprotein slyB	7
P0A295	Transcription termination factor rho	7
P37586	Methionine synthase	7
P60428	50S ribosomal protein L2	7
Q7CP96	Component of modulator for protease specific f	7
Q8ZK31	Valyl-tRNA synthetase	7
Q8ZQ17	RNase E	7
Q8ZRT7	Protein translocase subunit secA	7
Q8ZRZ0	Isoleucyl-tRNA synthetase	7
P0A251	Alkyl hydroperoxide reductase subunit C	6
P15434	Phenylalanyl-tRNA synthetase beta chain	6
P25077	Malate dehydrogenase	6
P65882	Adenylosuccinate synthetase	6
P66313	50S ribosomal protein L6	6
P74881	Phosphoribosylformylglycinamidine synthase	6
Q8ZKA4	Component of modulator for protease specific f	6
Q8ZPZ0	Isocitrate dehydrogenase [NADP]	6
Q8ZQZ6	Leucyl-tRNA synthetase	6
Q8ZRB8	Peptidyl prolyl isomerase	6
Q8ZRC0	ATP-dependent Clp protease ATP-binding subunit	6

O54297	30S ribosomal protein S4	5
P0A2C9	3-oxoacyl-[acyl-carrier-protein] reductase	5
P0A2E3	RNA polymerase sigma factor rpoD	5
P65692	6-phosphofructokinase	5
P66932	Trigger factor	5
Q8ZKM0	Phosphoenolpyruvate carboxylase	5
Q8ZLM1	30S ribosomal protein S13	5
Q8ZN40	Cysteine desulfurase	5
Q8ZRP0	Outer membrane protein assembly factor yaeT	5
Q935S7	Chromosome partition protein mukB	5
Q93DX3	Putative LysR family transcriptional regulator	5
P0A1Y6	NADH-quinone oxidoreductase subunit C/D	4
P16657	Enoyl-[acyl-carrier-protein] reductase [NADH]	4
P63923	Phosphopentomutase	4
Q7CPE1	ATP synthase subunit alpha	4
Q7CQU5	Thioredoxin reductase	4
Q7CQV8	Outer membrane protease, receptor for phage OX	4
Q7CQW9	Tol protein required for outer membrane integr	4
Q8ZLK3	Membrane protein	4
Q8ZM69	D-3-phosphoglycerate dehydrogenase	4
Q8ZPG8	Putative hydrolase	4
Q9F173	DNA polymerase I	4
Q9XCS1	Fur regulated Salmonella iron transporter	4
O84944	Secretion system effector	3
P0A2E1	Serine hydroxymethyltransferase	3
P0AA19	Transcriptional regulatory protein ompR	3
P65215	2-dehydro-3-deoxyphosphooctonate aldolase	3
P66764	S-adenosylmethionine synthetase	3
Q8ZKN4	Aspartokinase II	3
Q8ZLJ1	Putative RNase R	3
Q8ZLT4	Cysteine sulfinate desulfinase	3
Q8ZMM7	VirK-like protein	3
Q8ZPS3	Phosphoenolpyruvate synthase	3
Q8ZQ40	Flavoprotein wrbA	3
Q8ZR37	Enterobactin synthetase, component F	3
Q8ZRT0	Dihydrolipoyl dehydrogenase	3
Q8ZS19	Aspartokinase I	3
Q93DX1	Putative acetyl-CoA hydrolase	3
P0A1V6	Ribose-phosphate pyrophosphokinase	2
P0A2F6	Single-stranded DNA-binding protein	2
Q7CPL4	50S ribosomal protein L16	2
Q7CPL6	50S ribosomal protein L18	2
Q7CQ11	NifU homolog	2
Q7CQY7	Putative phosphate starvation-inducible protei	2
Q8ZLT0	Argininosuccinate synthase	2
Q8ZRB4	Putative ABC superfamily	2
Q93GL9	Conjugative transfer: surface exclusion	2

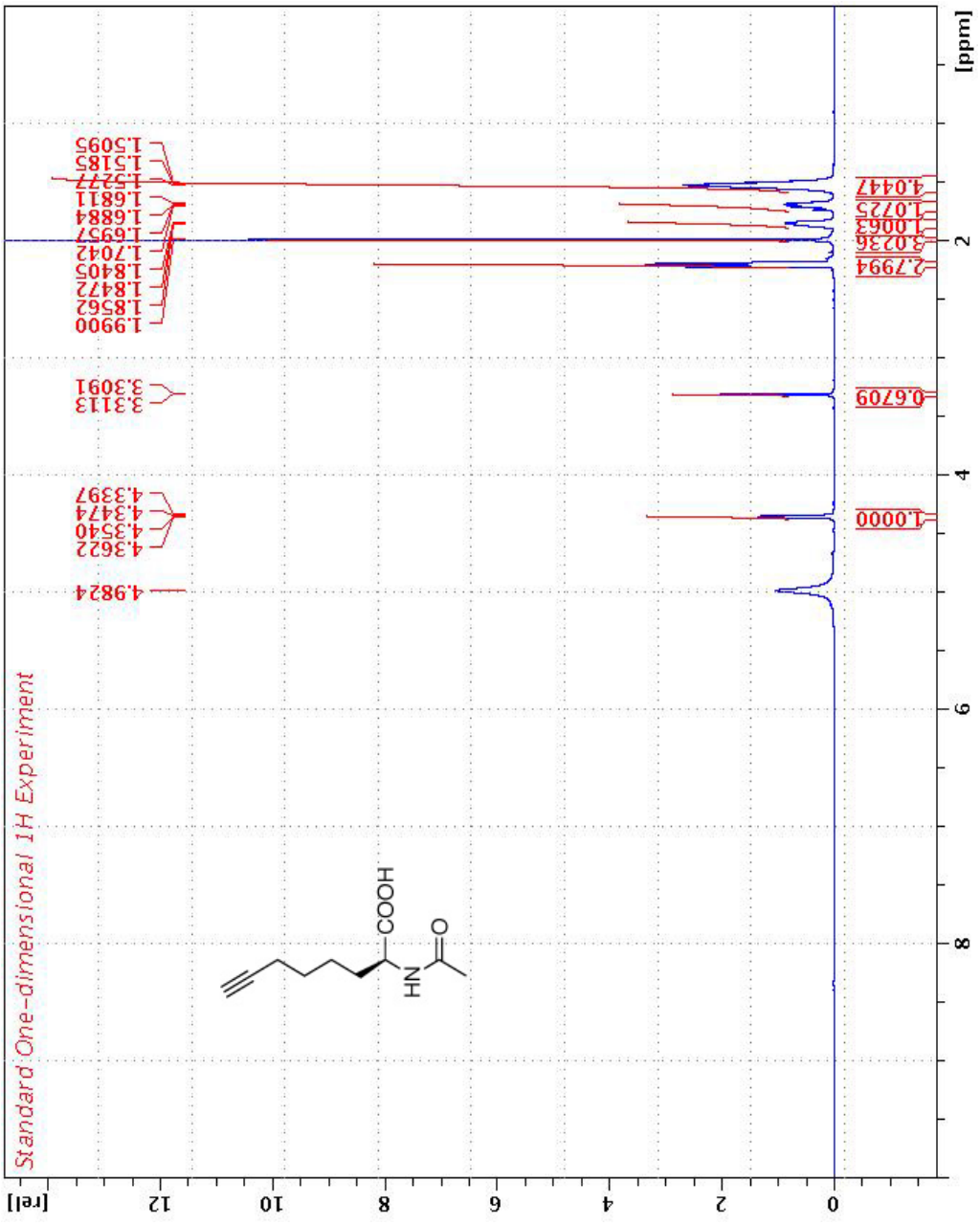
9.3 Chemical characterization data

9.3.1 Diethyl 2-acetamido-2-(hex-5-yn-1-yl)malonate

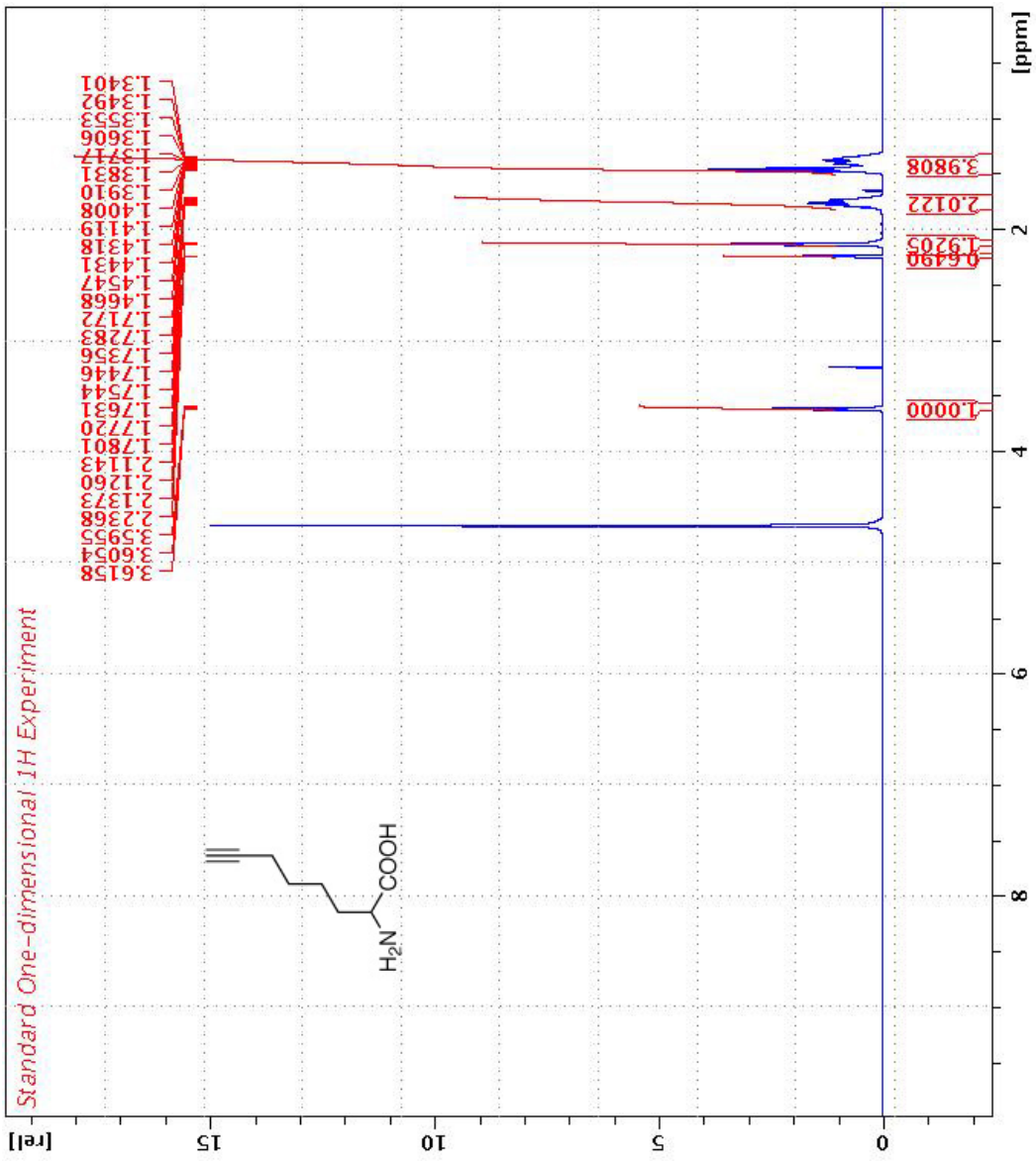


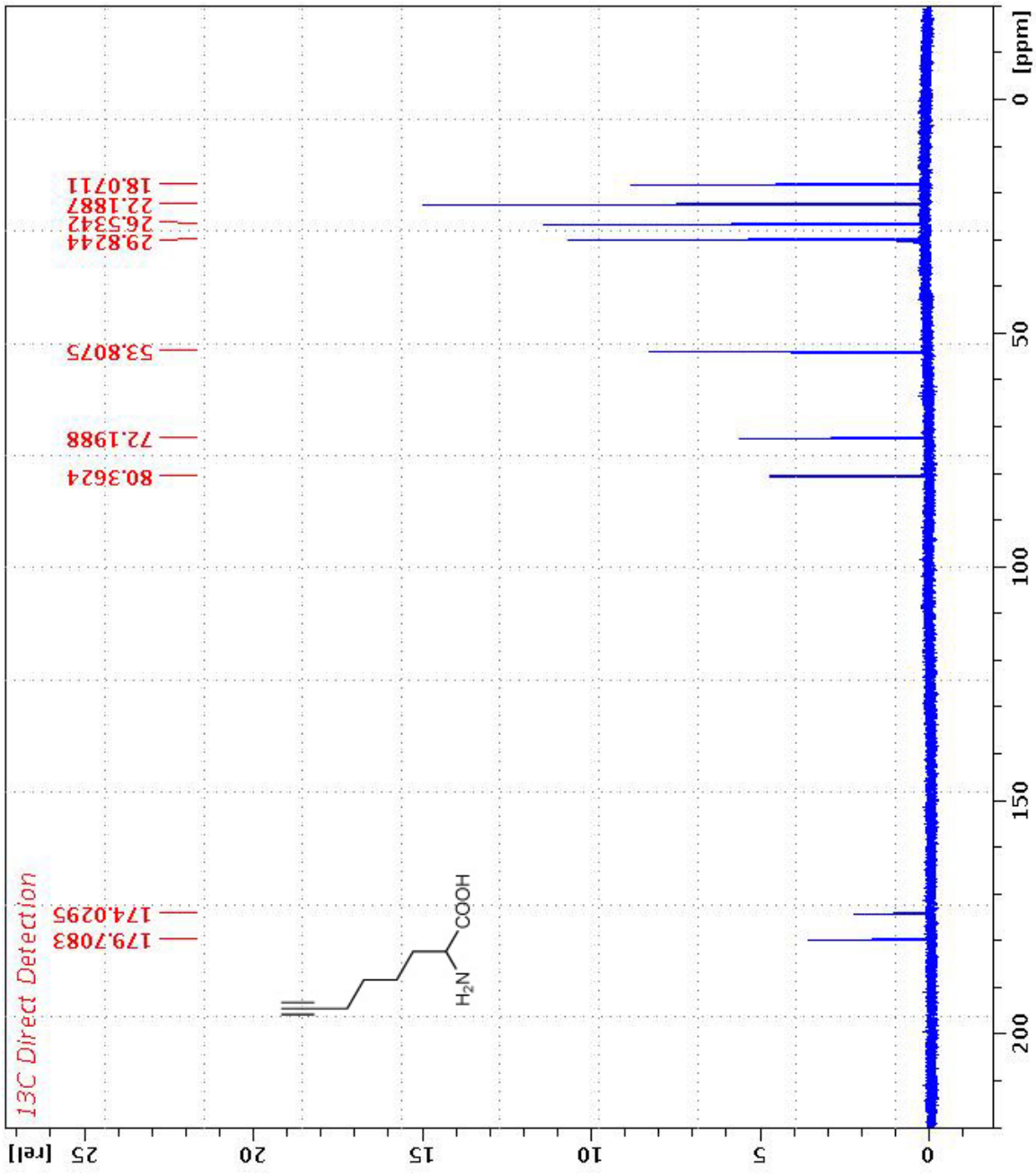

13C Direct Detection
168.9233
167.9704
83.7703
68.4993
66.3614
62.3377
31.4343
27.8020
22.8565
22.4087
17.9536
13.8767
[rel]
20
15
10
5
0
[ppm]
200
150
100
50
0
O
O
O
O
NH
O

9.3.2 2-acetamidooct-7-ynoic acid

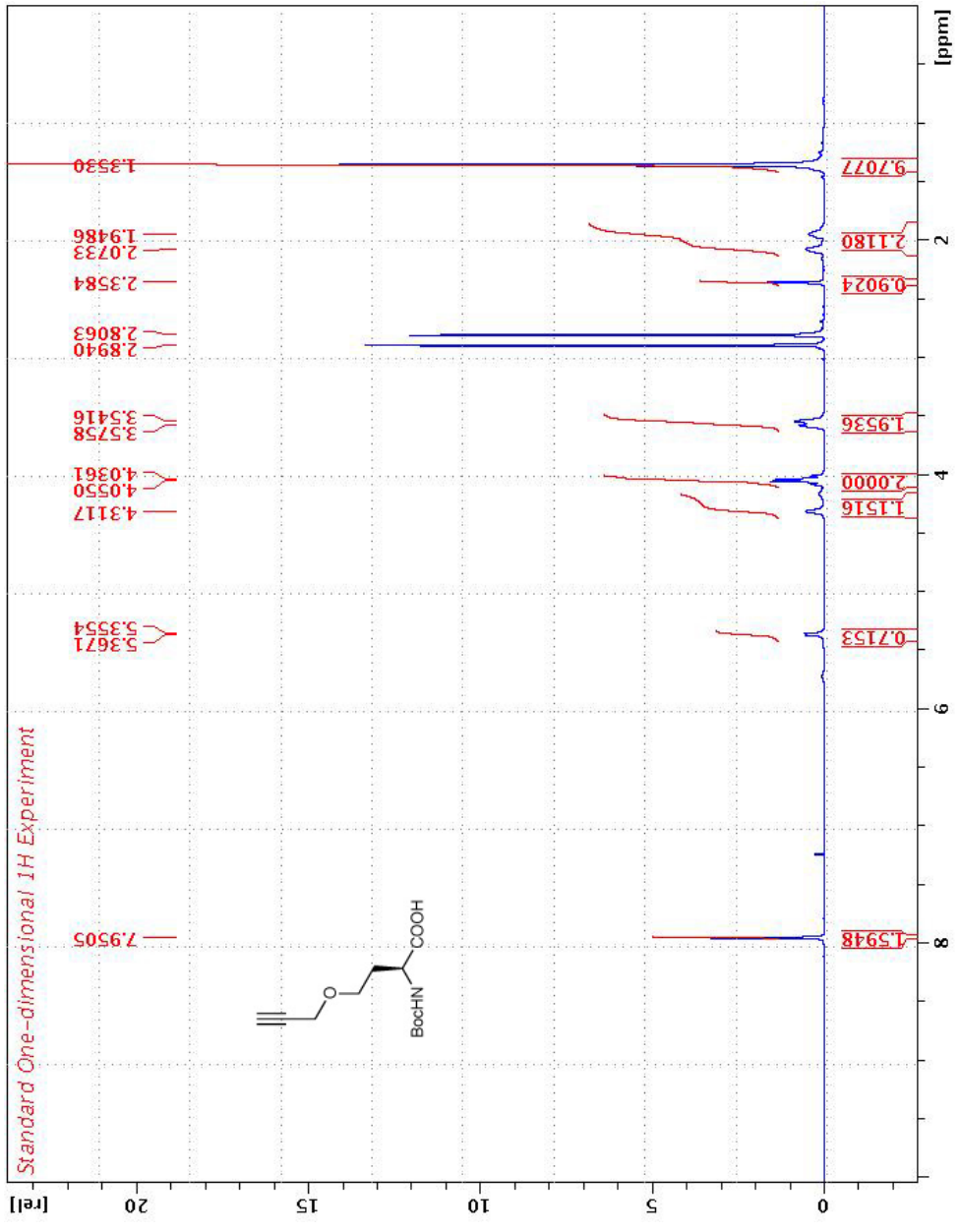


9.3.3 2-amino-7-octynoic acid (AOA)

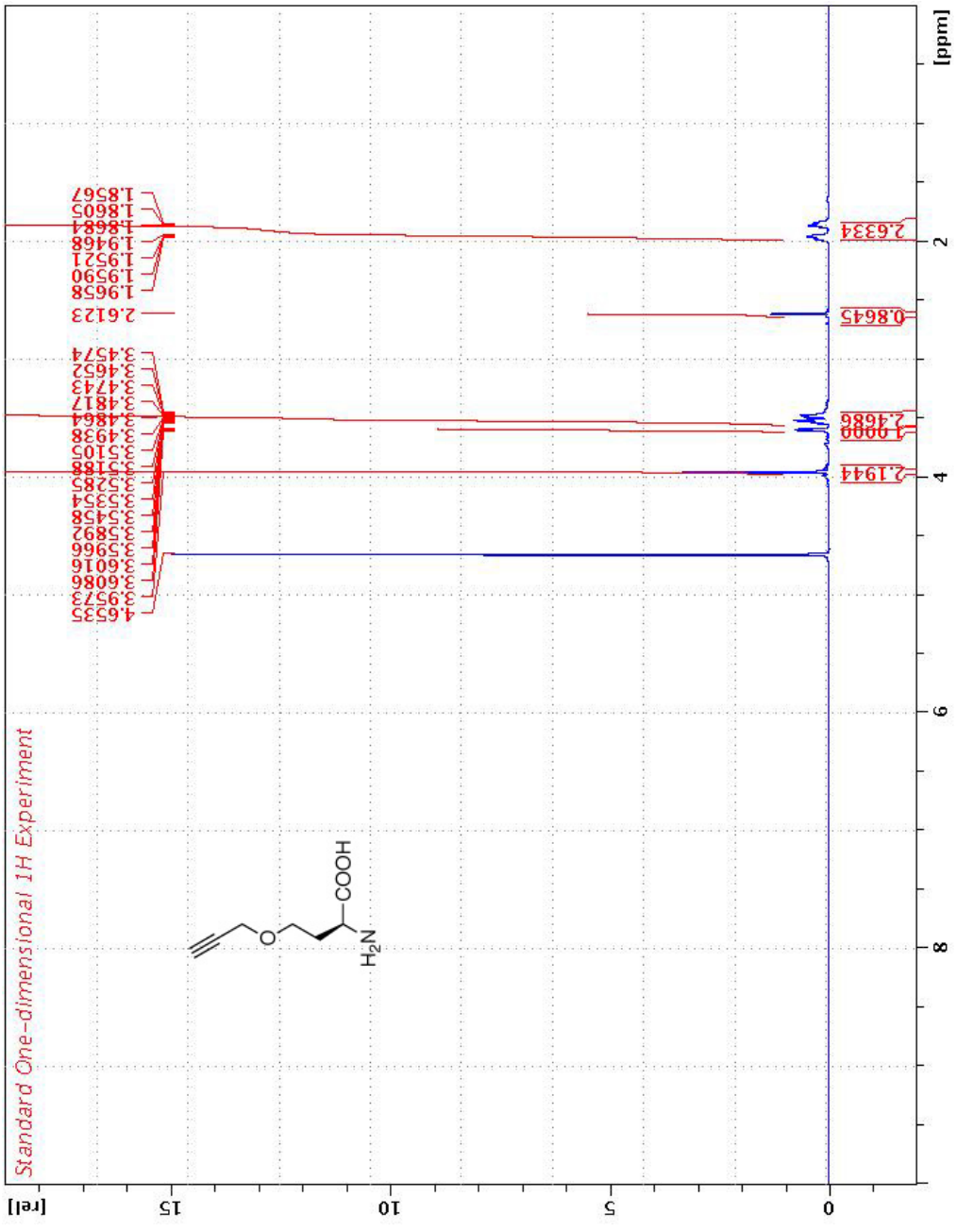


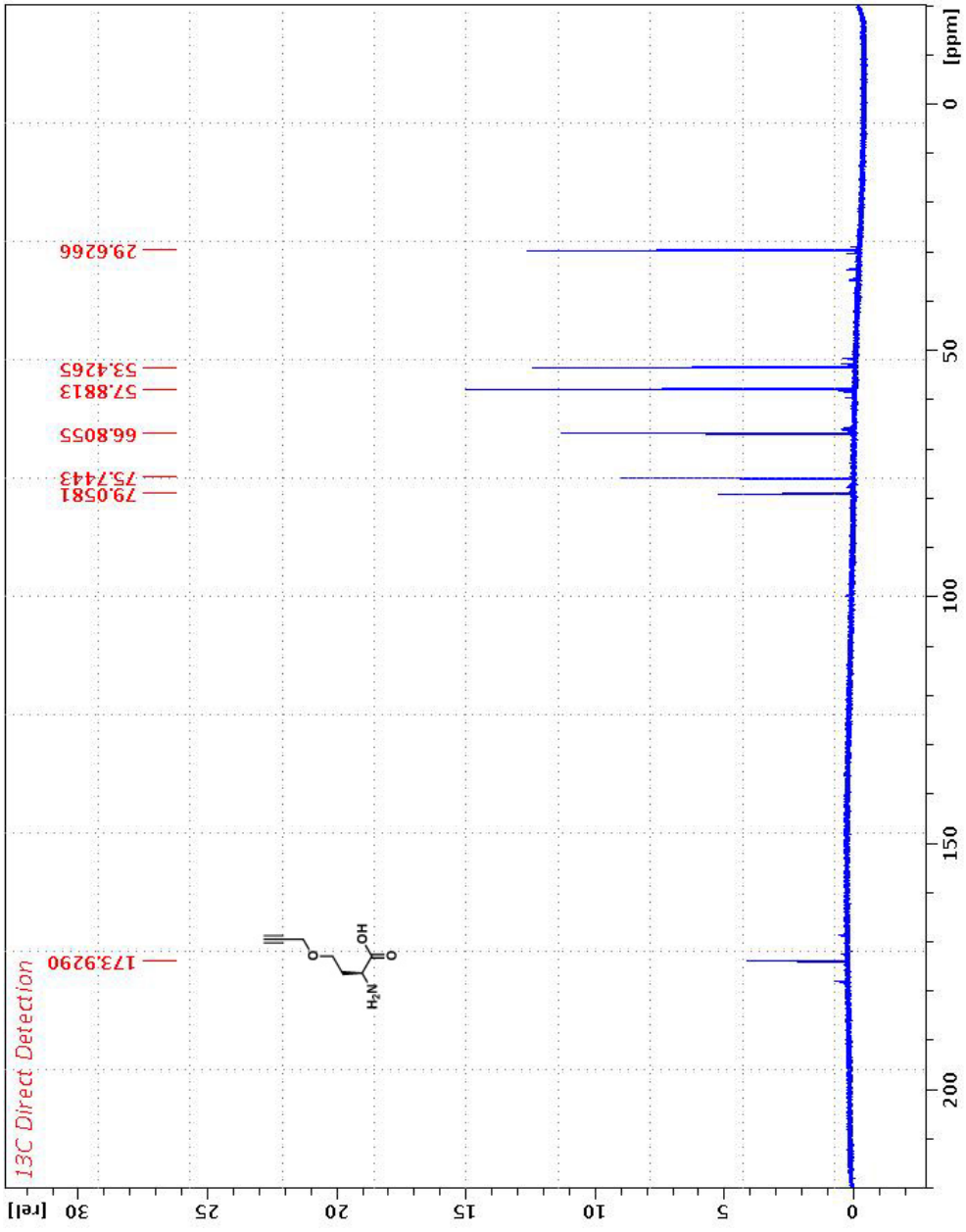

13C Direct Detection
179.7083
174.0295
53.8075
72.1988
80.3624
29.8244
26.5342
22.1887
18.0711
COOH
H2N
[rel]
25
20
15
10
5
0
[ppm]
200
150
100
50
0

9.3.4 (S)-2-((tert-butoxycarbonyl)amino)-4-(prop-2-yn-1-yloxy)butanoic acid

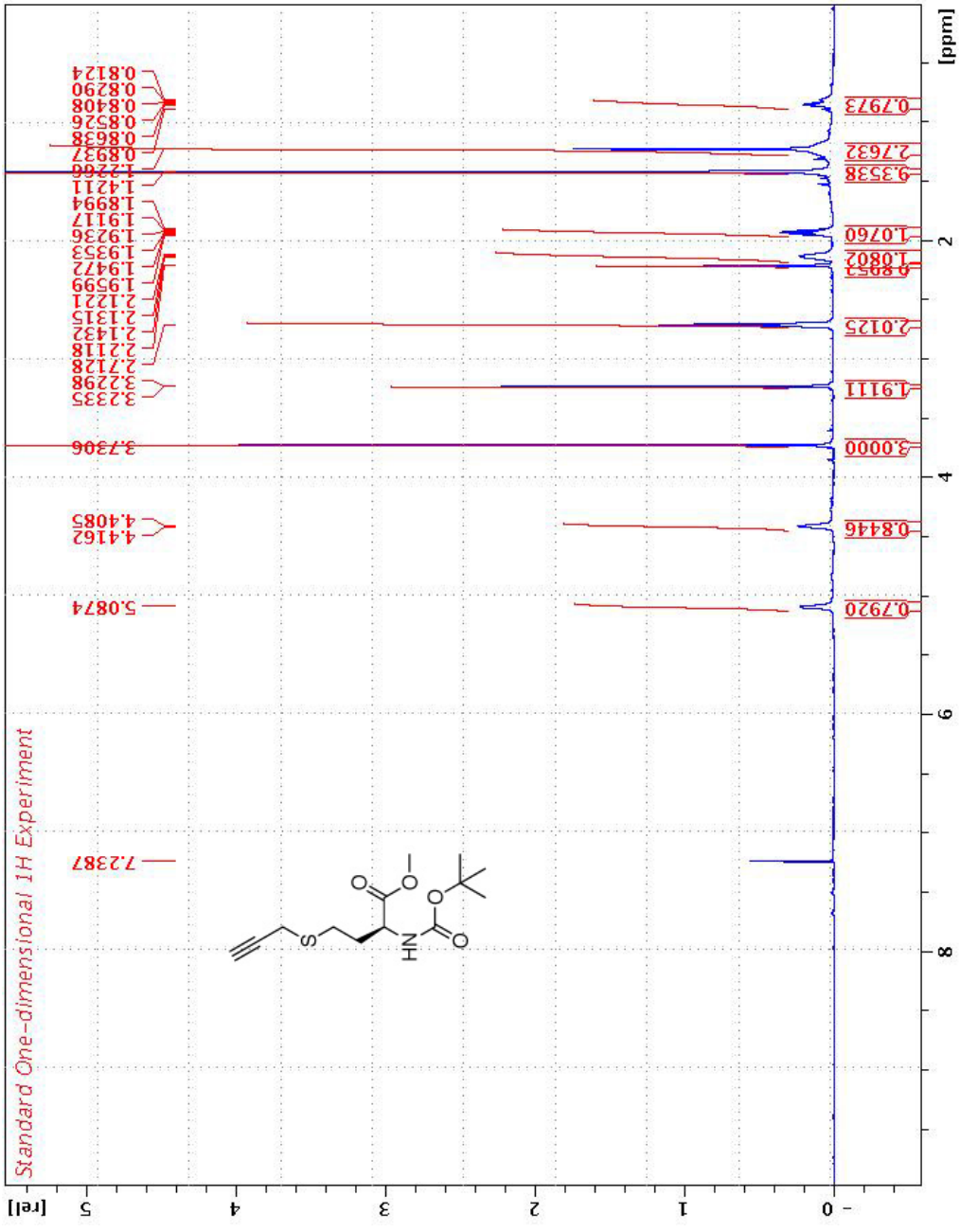


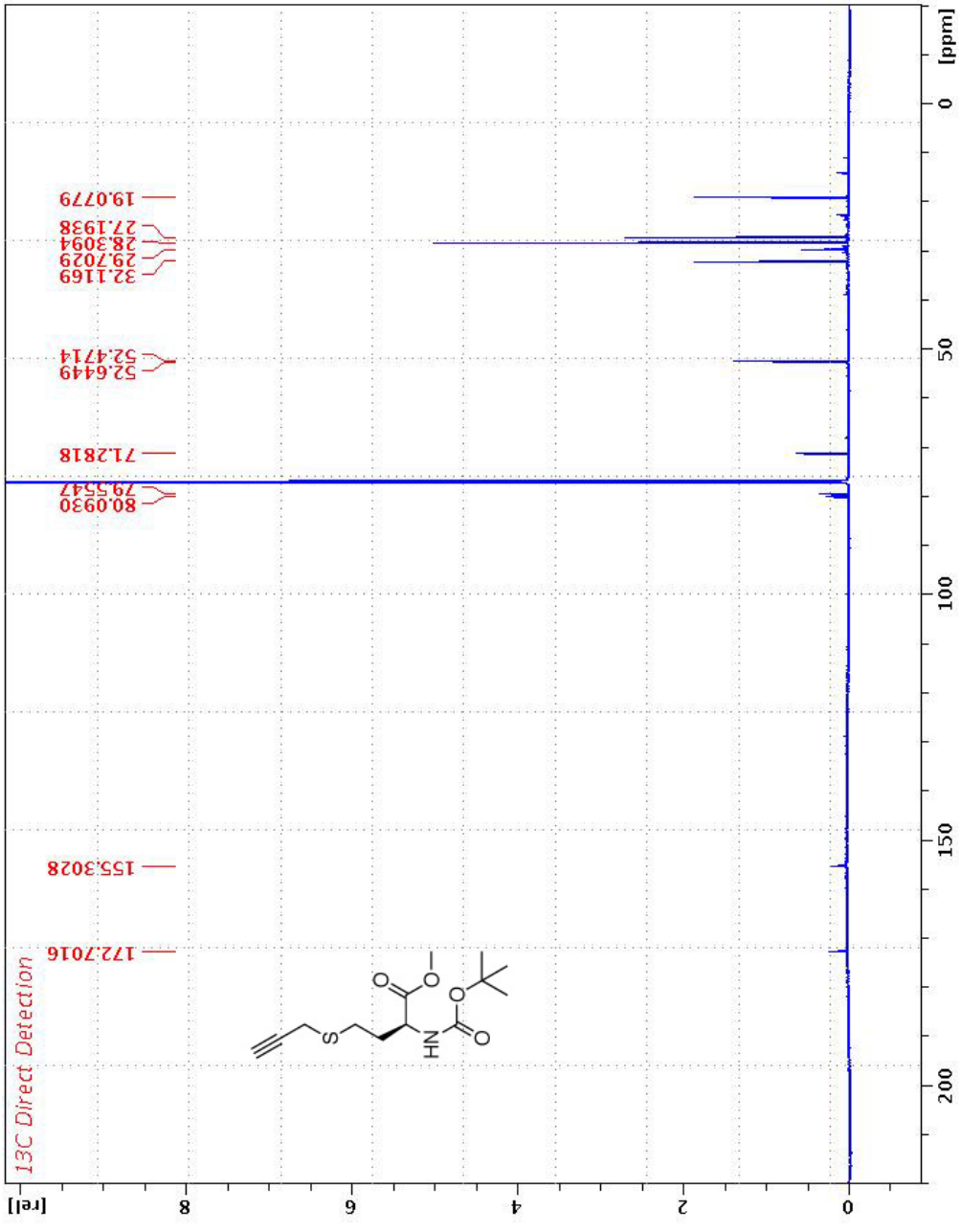
9.3.5 (*S*)-2-amino-4-(prop-2-yn-1-yloxy)butanoic acid (PHS)



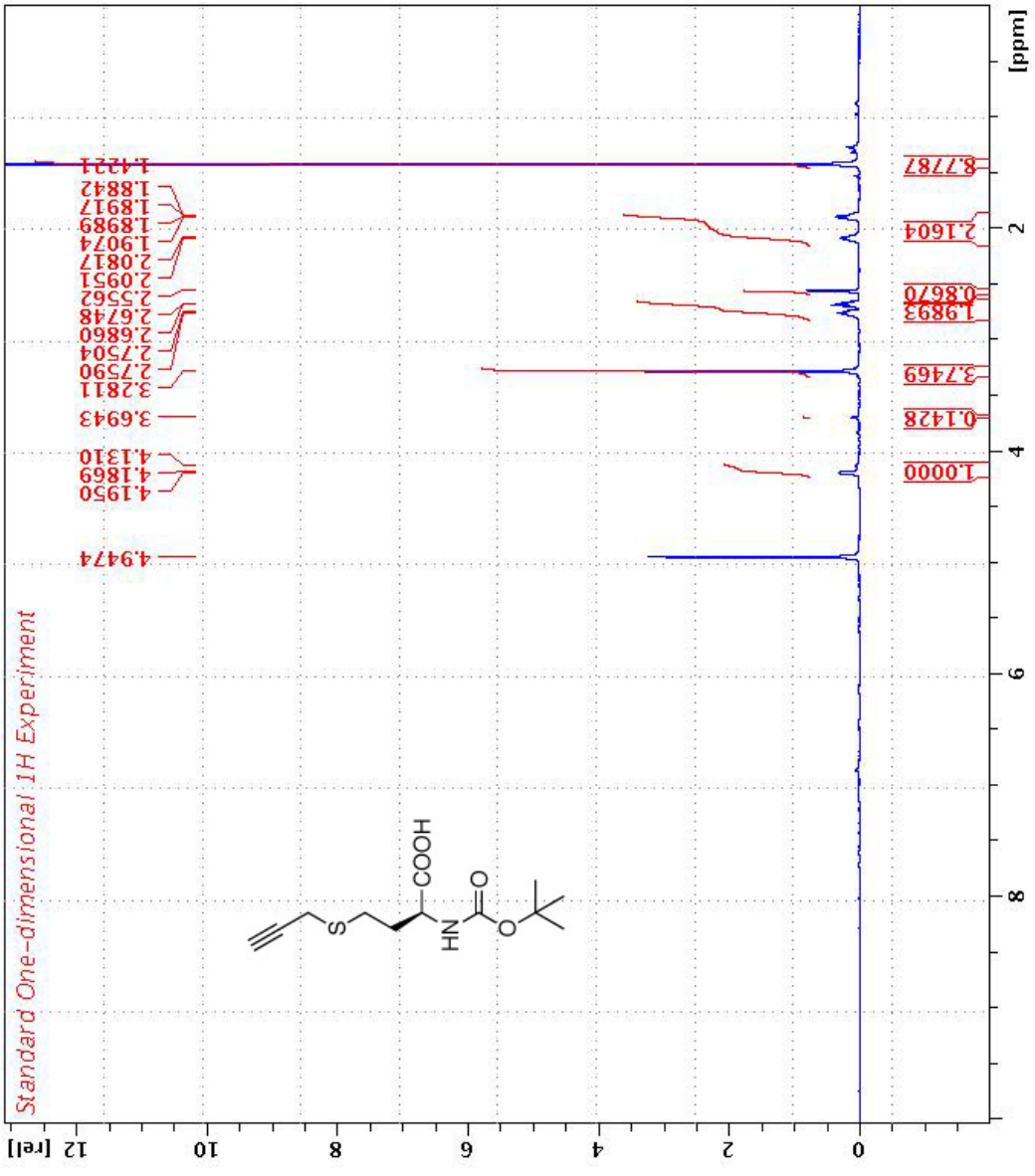

13C Direct Detection
[rel]
[ppm]
173.9290
79.0581
75.7443
66.8055
57.8813
53.4265
29.6266
HO
O
O
H2N

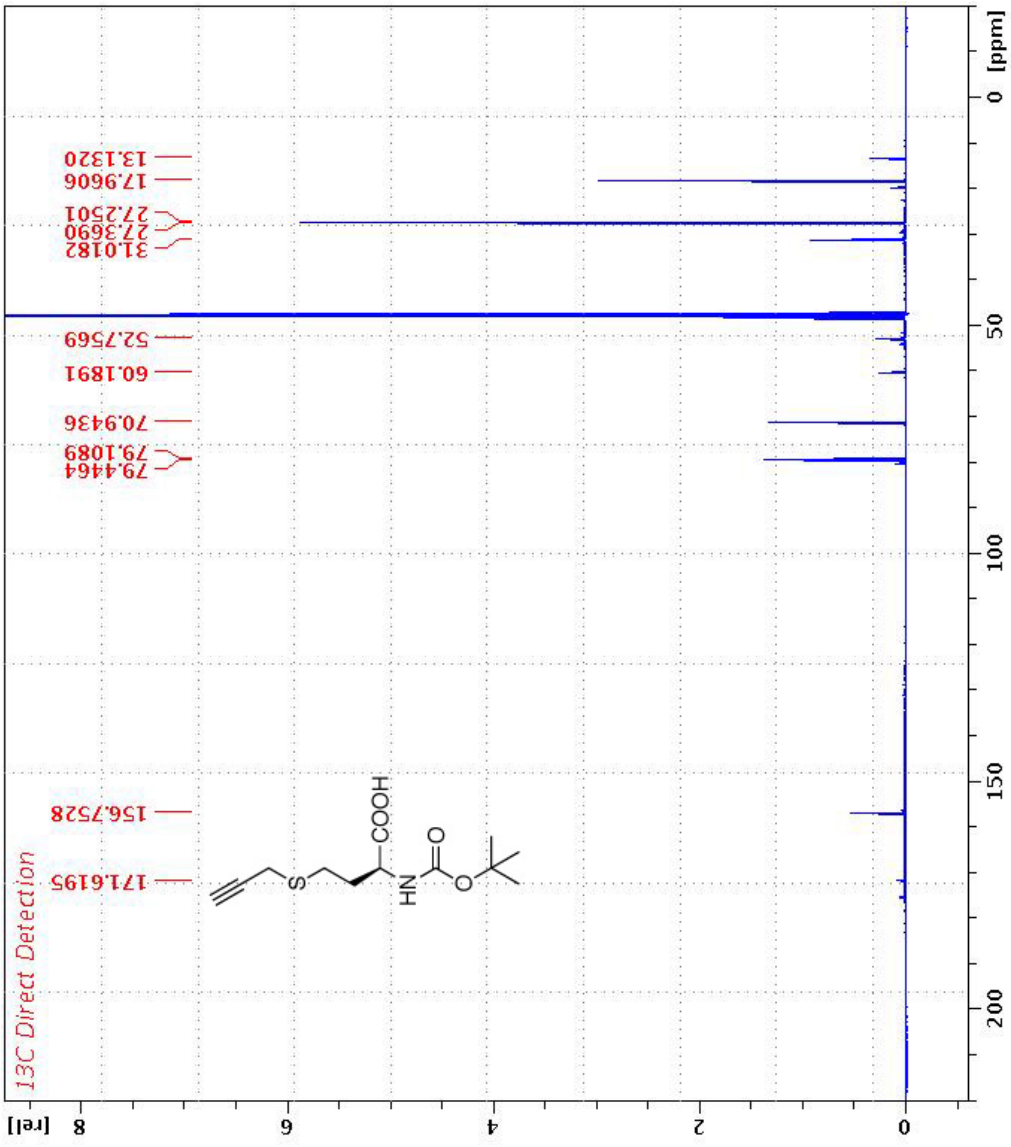
9.3.6 (S)-methyl 2-((tert-butoxycarbonyl)amino)-4-(prop-2-yn-1-ylthio)butanoate



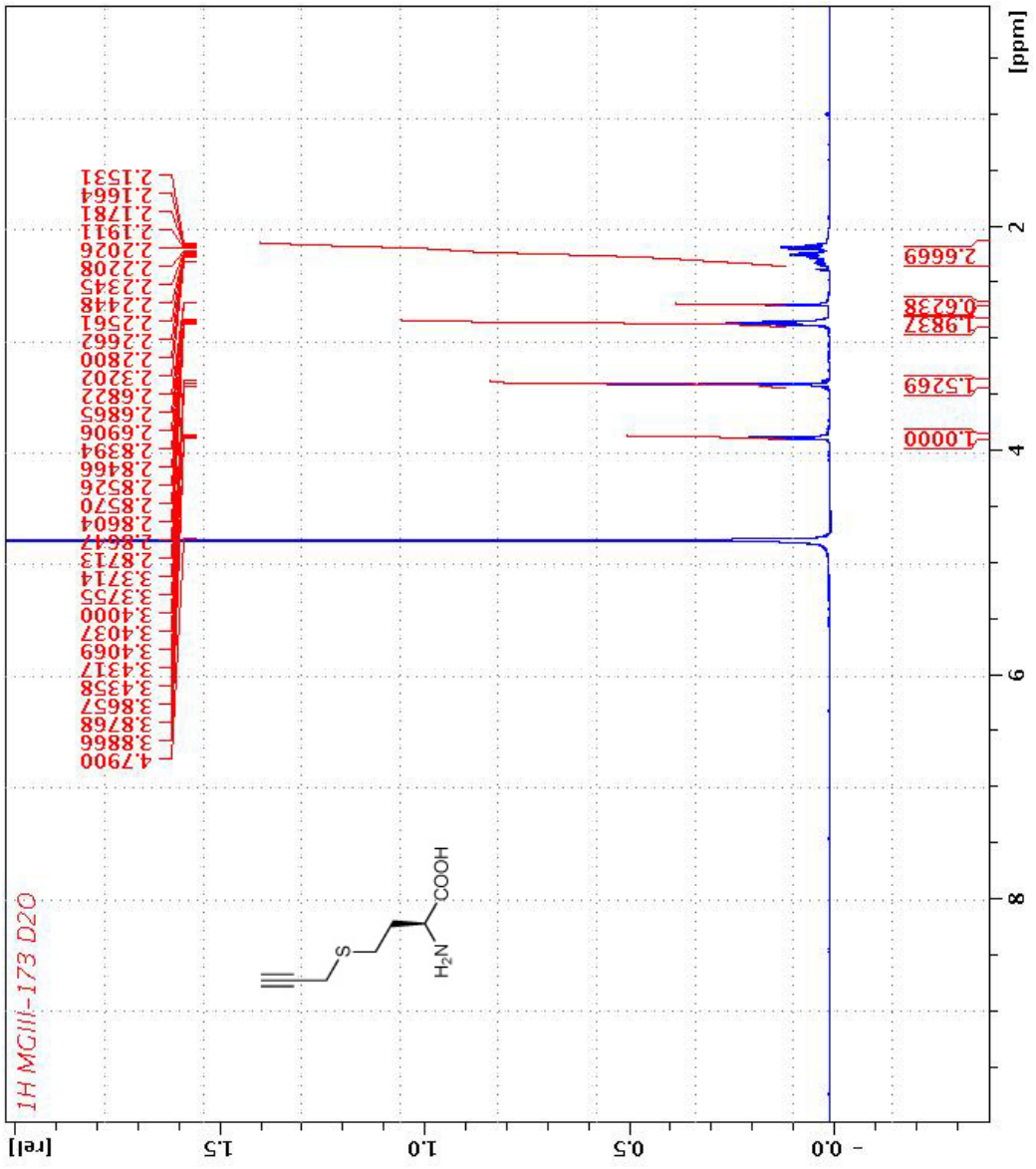

13C Direct Detection
172.7016
155.3028
80.0930
79.5547
71.2818
52.6449
52.4714
32.1169
29.7029
28.3094
27.1938
19.0779
[rel]
[ppm]

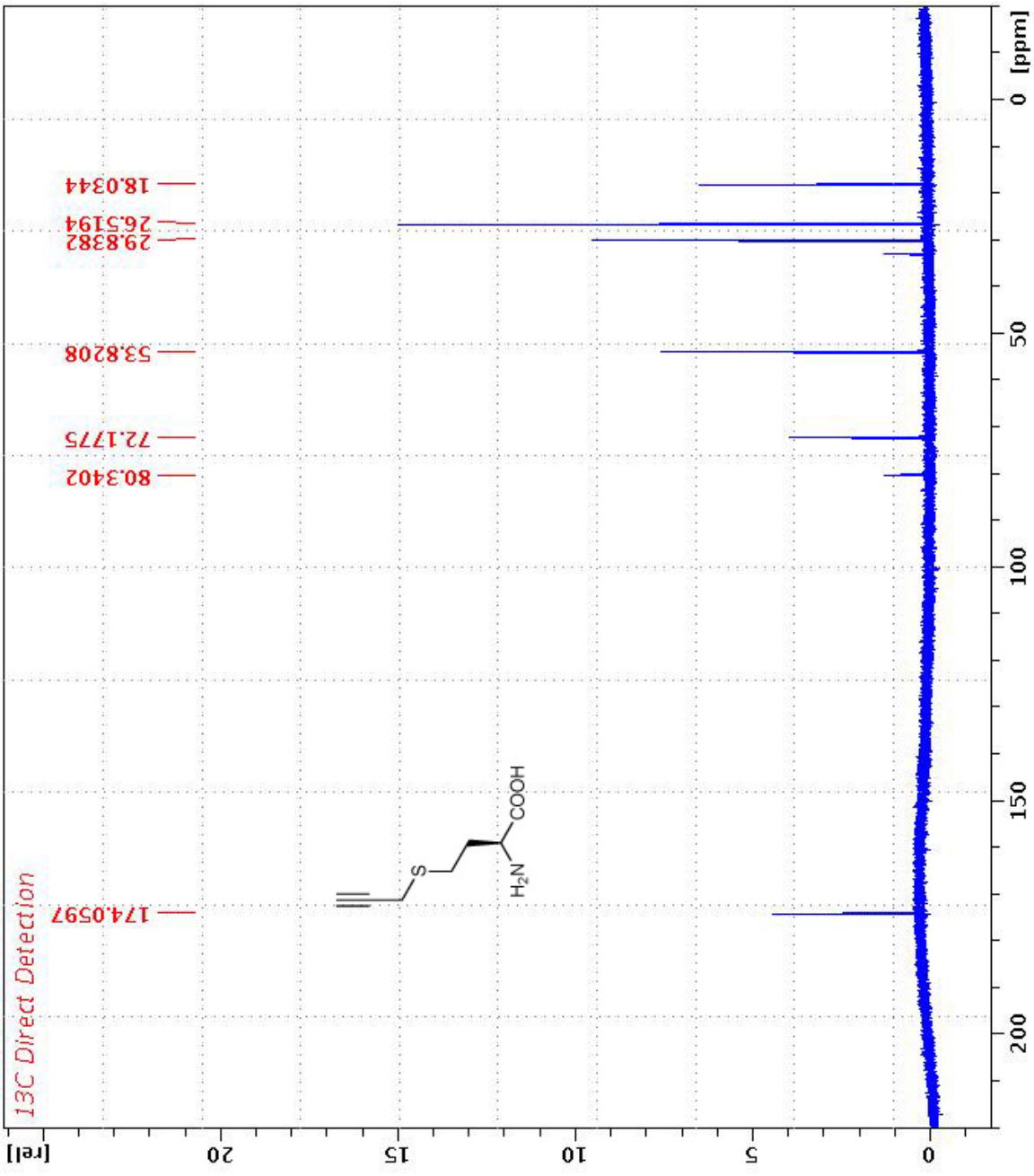
9.3.7 (S)-2-((tert-butoxycarbonyl)amino)-4-(prop-2-yn-1-ylthio)butanoic acid



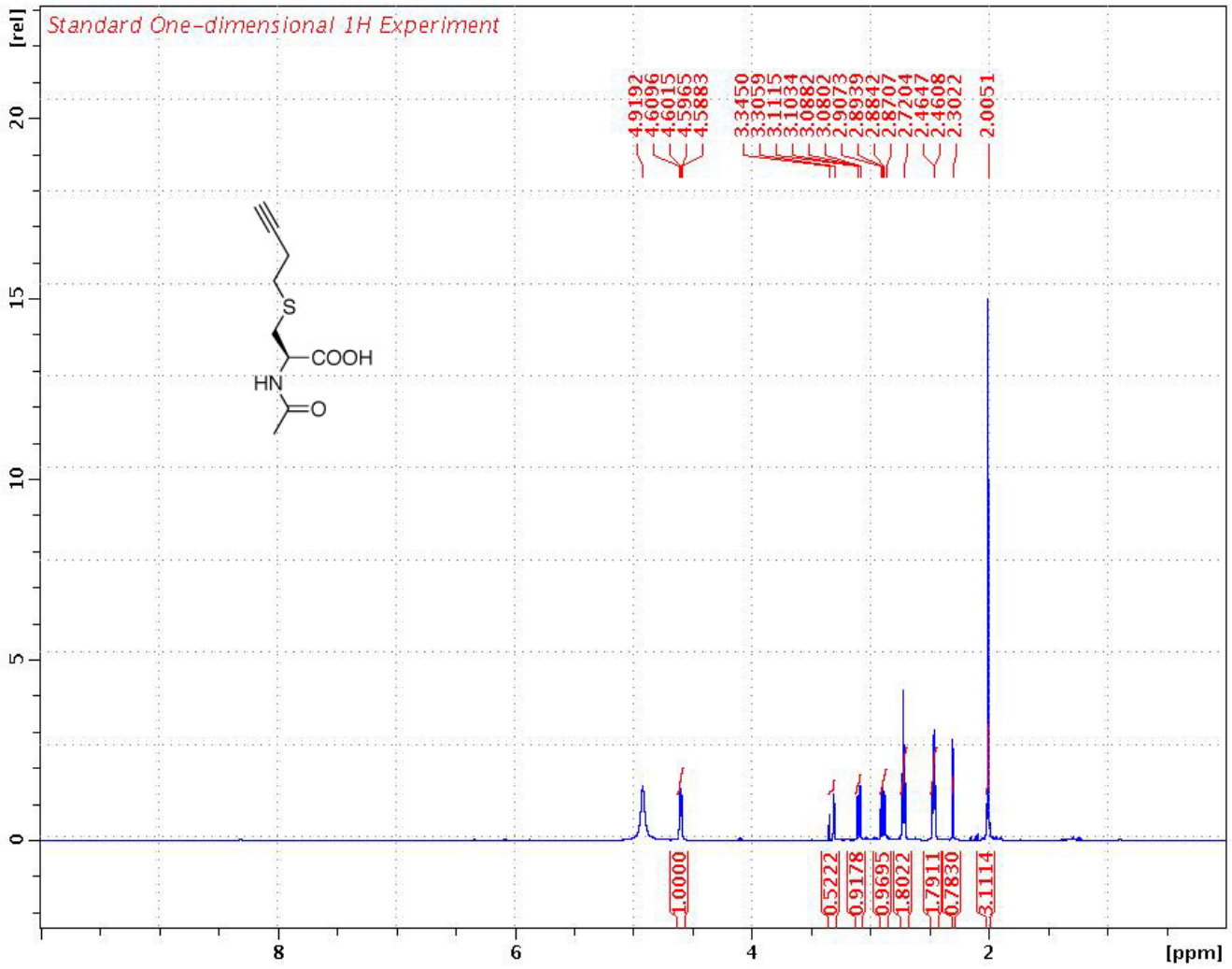

13C Direct Detection
[rel]
0
2
4
6
8
[ppm]
0
50
100
150
200
156.7528
171.6195
79.4464
79.1089
70.9436
60.1891
52.7569
31.0182
27.3690
27.2501
17.9606
13.1320
COOH
NH
O
O
S

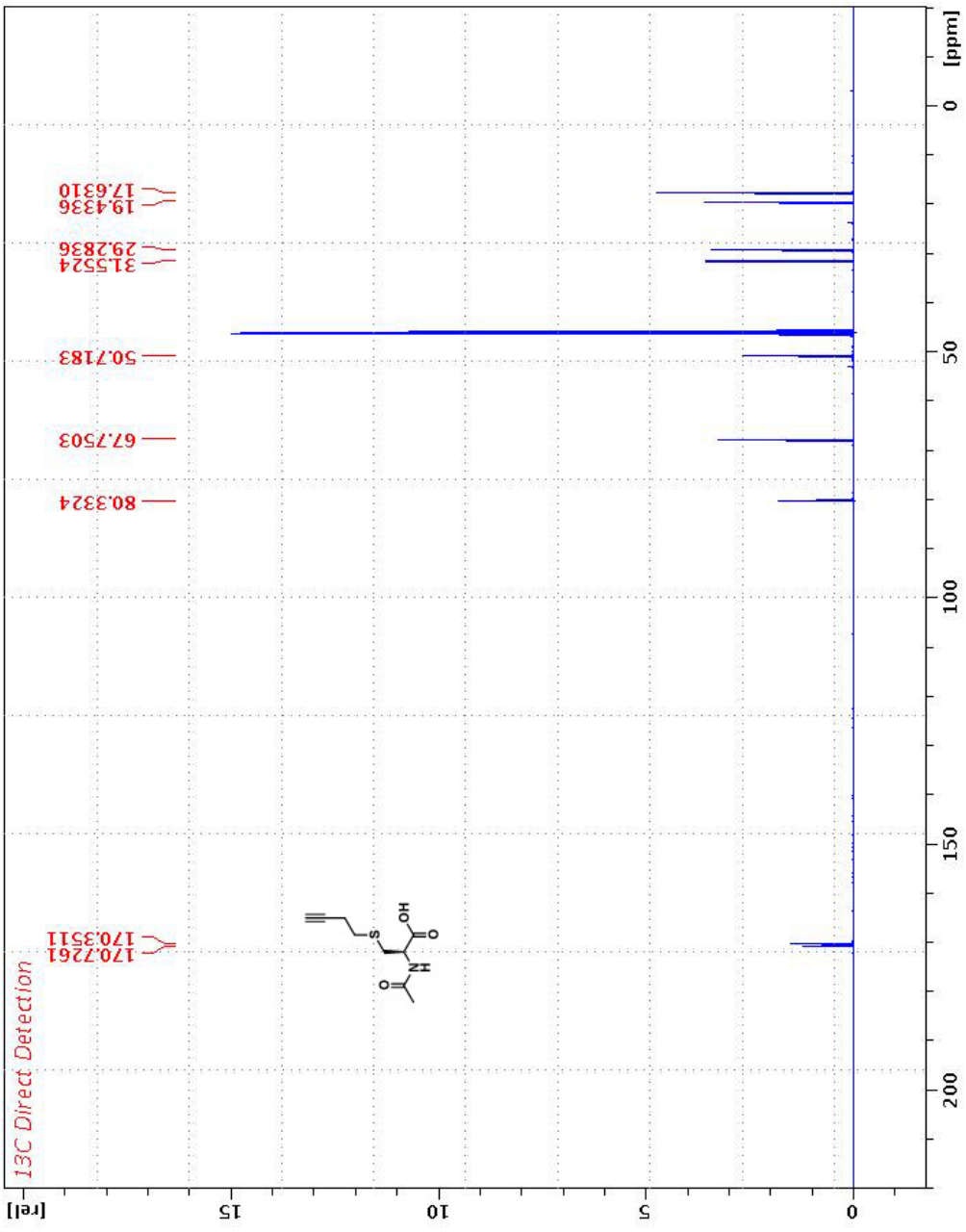
9.3.8 (*S*)-2-amino-4-(prop-2-yn-1-ylthio)butanoic acid (EME)



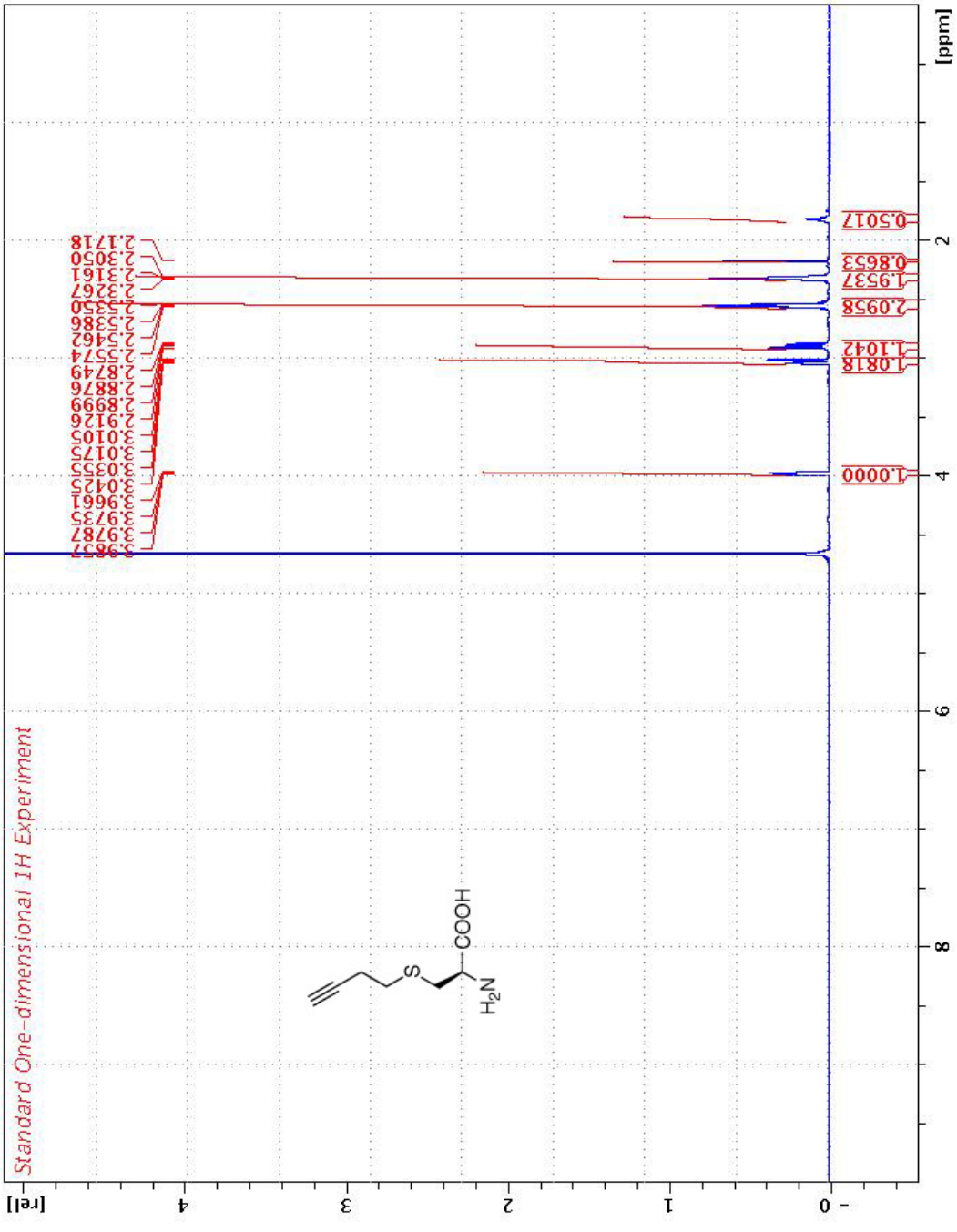

13C Direct Detection
174.0597
80.3402
72.1775
53.8208
29.8382
26.5194
18.0344
COOH
H_2N
S
[rel]
0
5
10
15
20
[ppm]
0
50
100
150
200

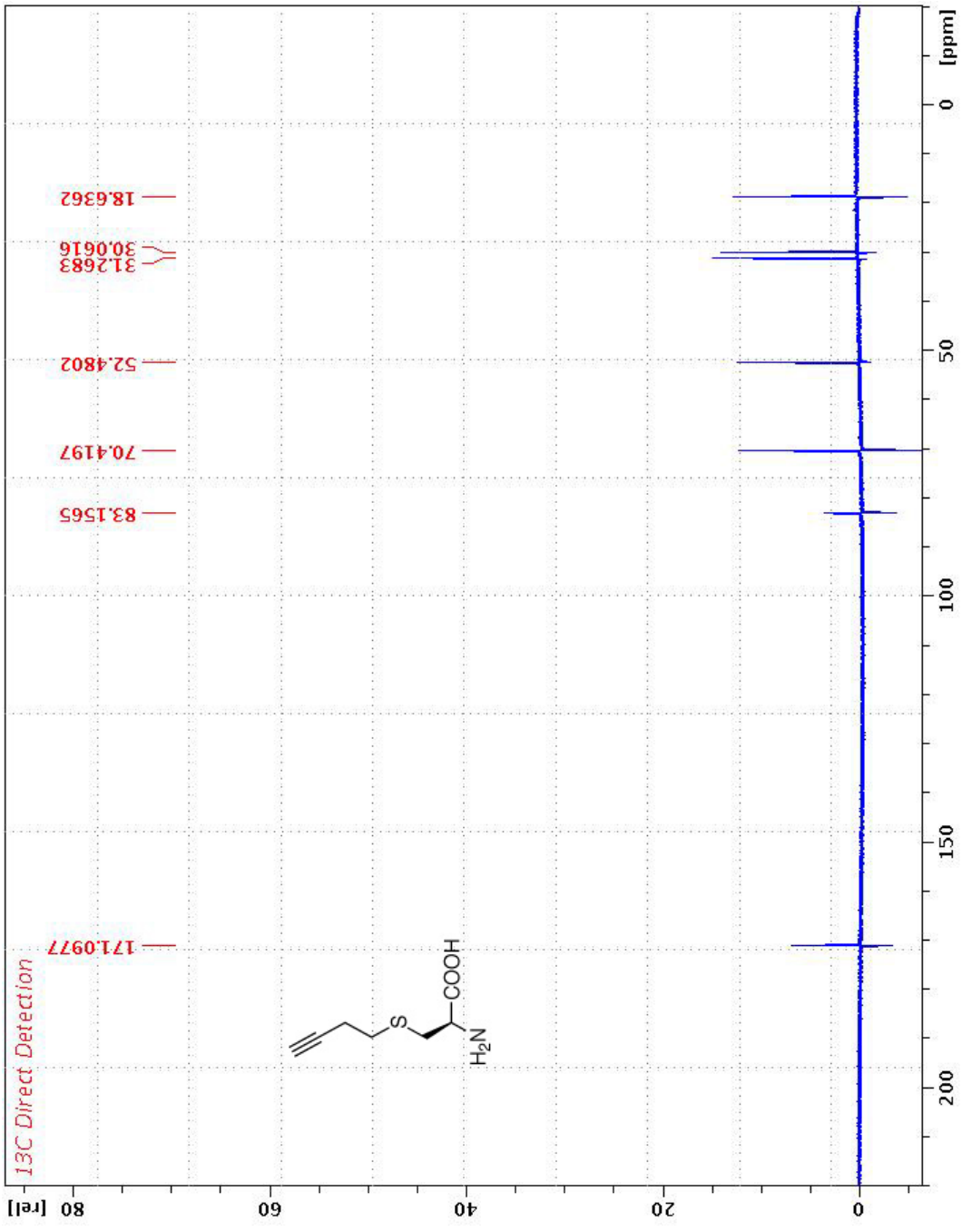
9.3.9 (R)-2-acetamido-3-(but-3-yn-1-ylthio)propanoic acid



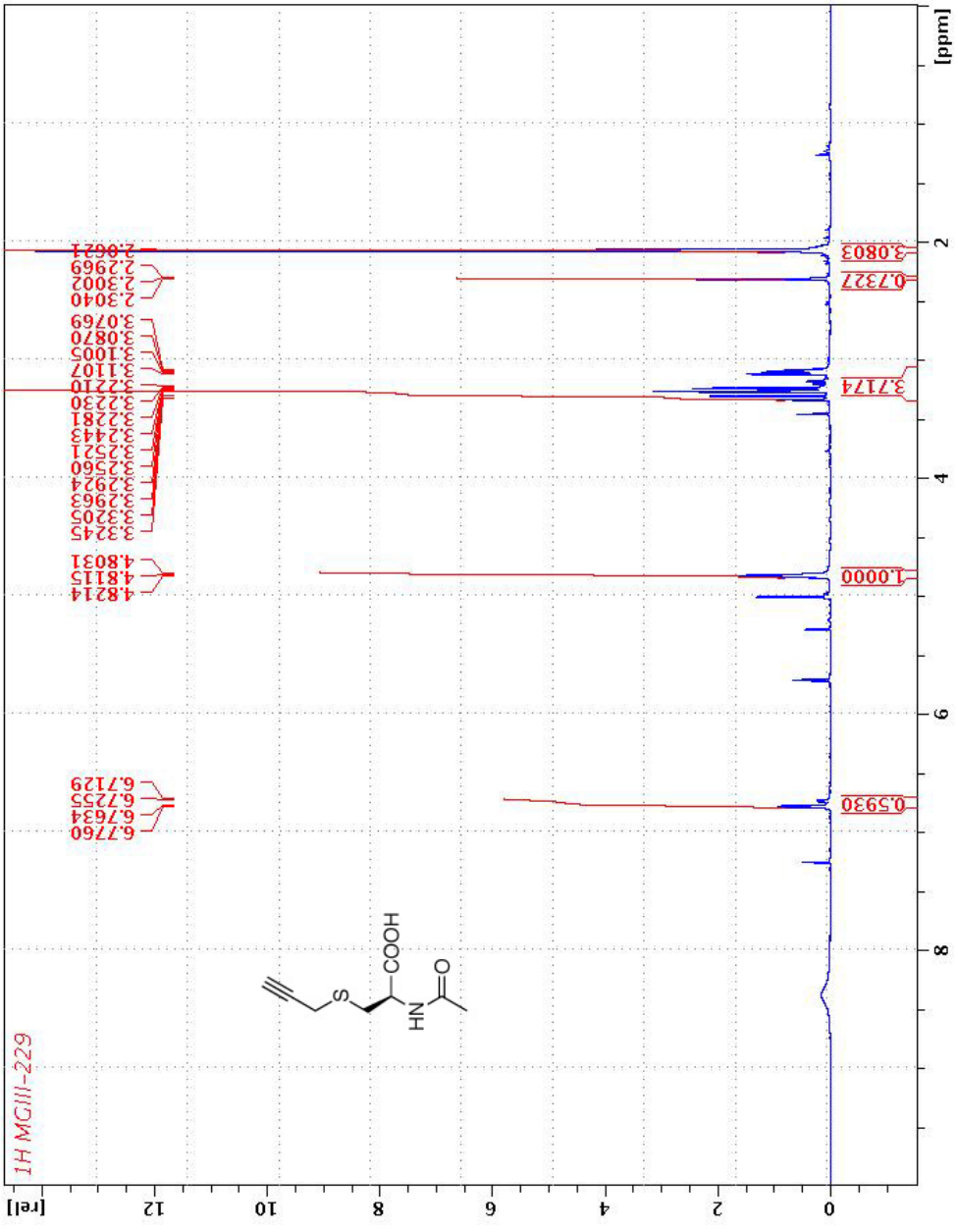

13C Direct Detection
170.7261
170.3511
80.3324
67.7503
50.7183
31.5524
29.2836
19.4336
17.6310
[rel]
[ppm]

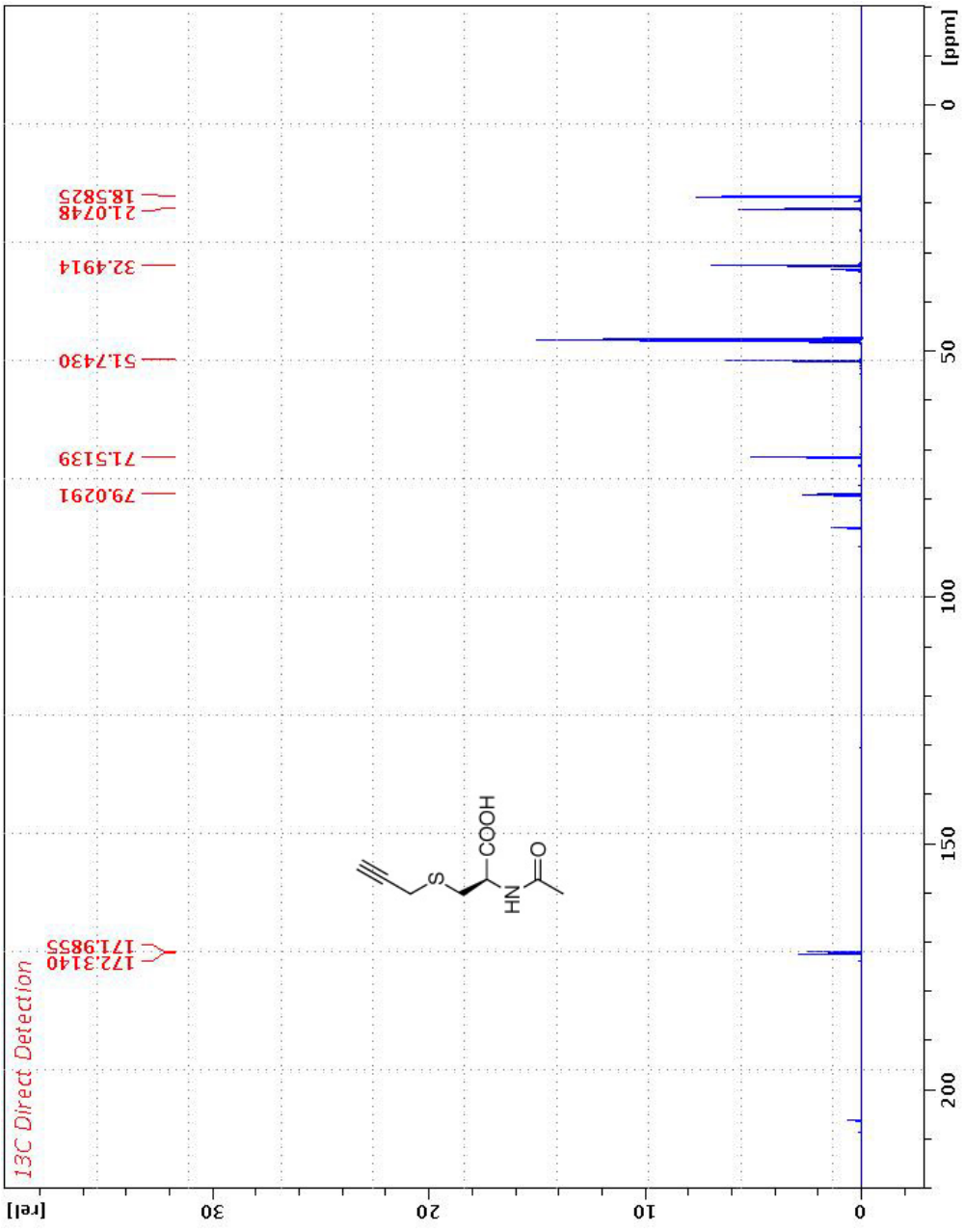
9.3.10 (R)-2-amino-3-(but-3-yn-1-ylthio)propanoic acid (BCY)



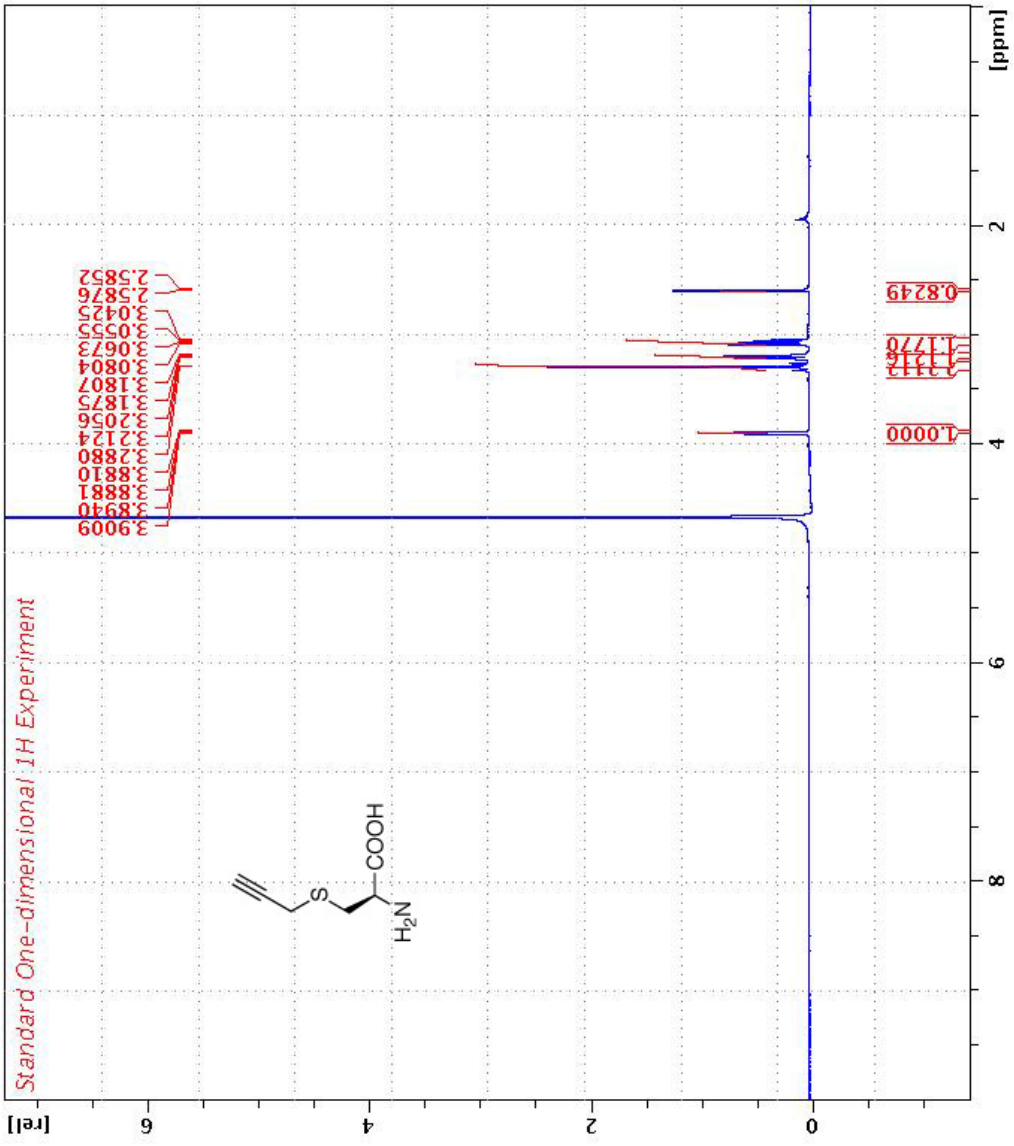

13C Direct Detection
[rel]
80
60
40
20
0
[ppm]
0
50
100
150
200
171.0977
83.1565
70.4197
52.4802
31.2683
30.0616
18.6362
COOH
S
H2N

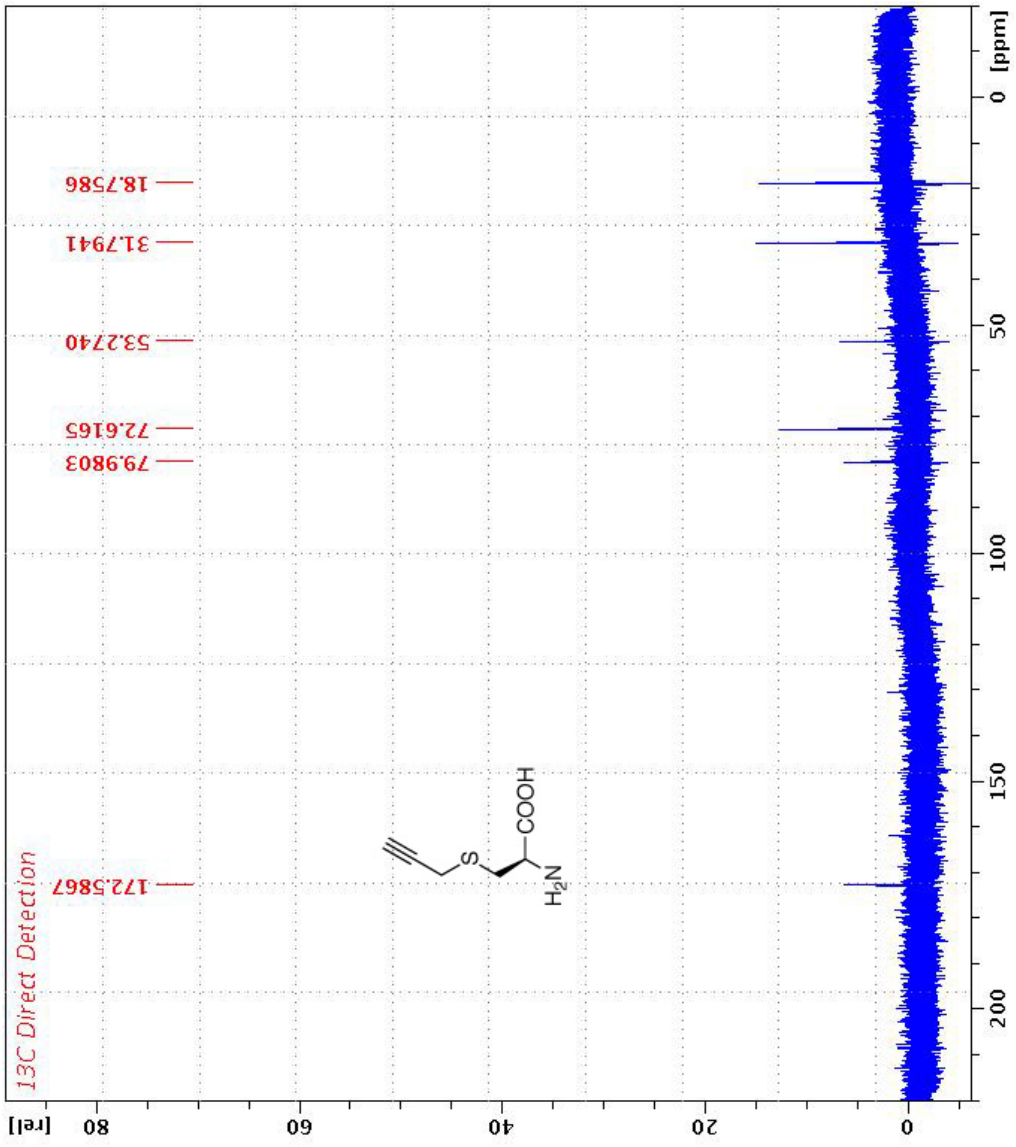
9.3.11 (R)-2-acetamido-3-(prop-2-yn-1-ylthio)propanoic acid



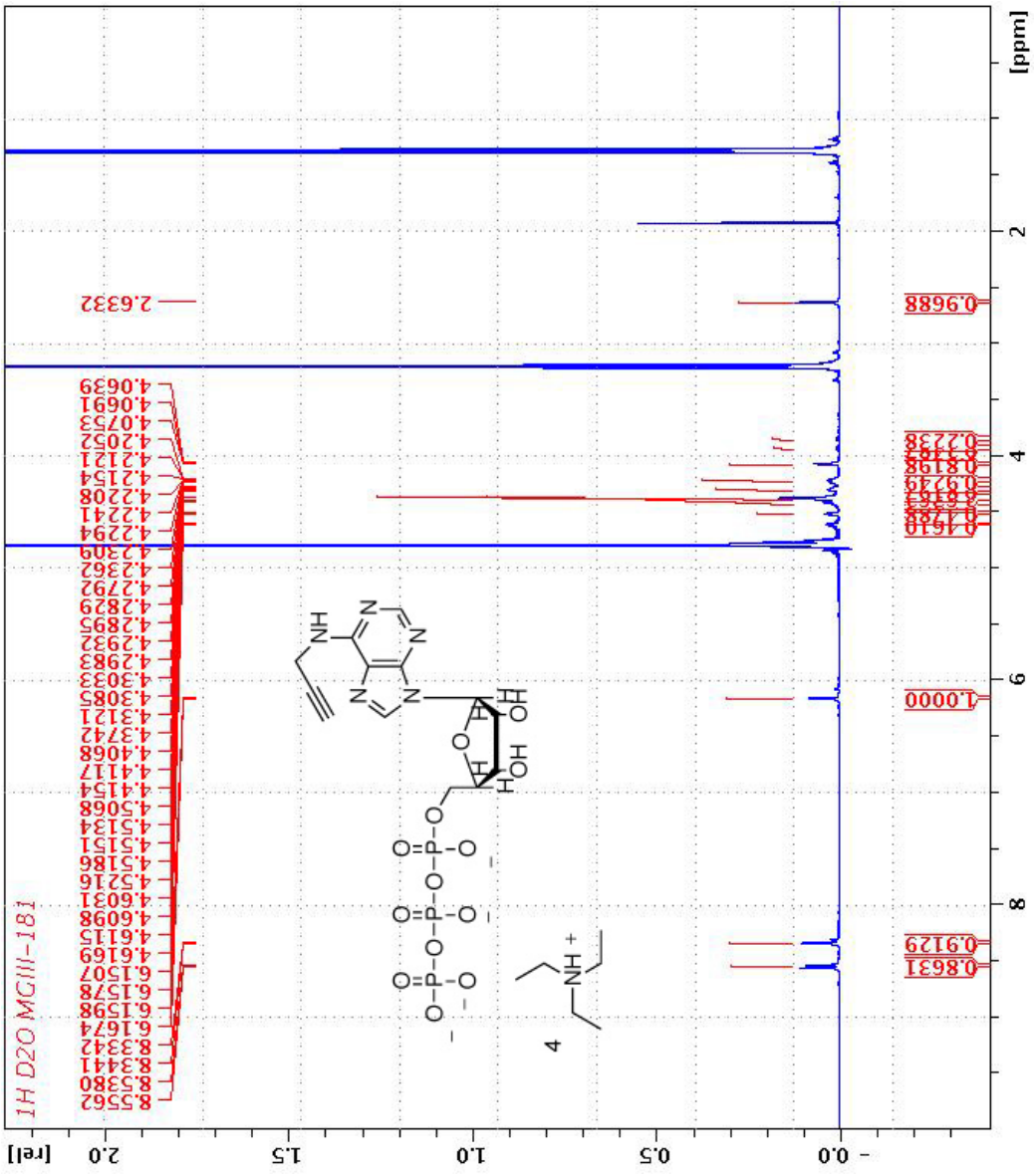

13C Direct Detection
172.3140
171.9855
79.0291
71.5139
51.7430
32.4914
21.0748
18.5825
COOH
HN
O
S
[rel]
0
10
20
30
[ppm]
0
50
100
150
200

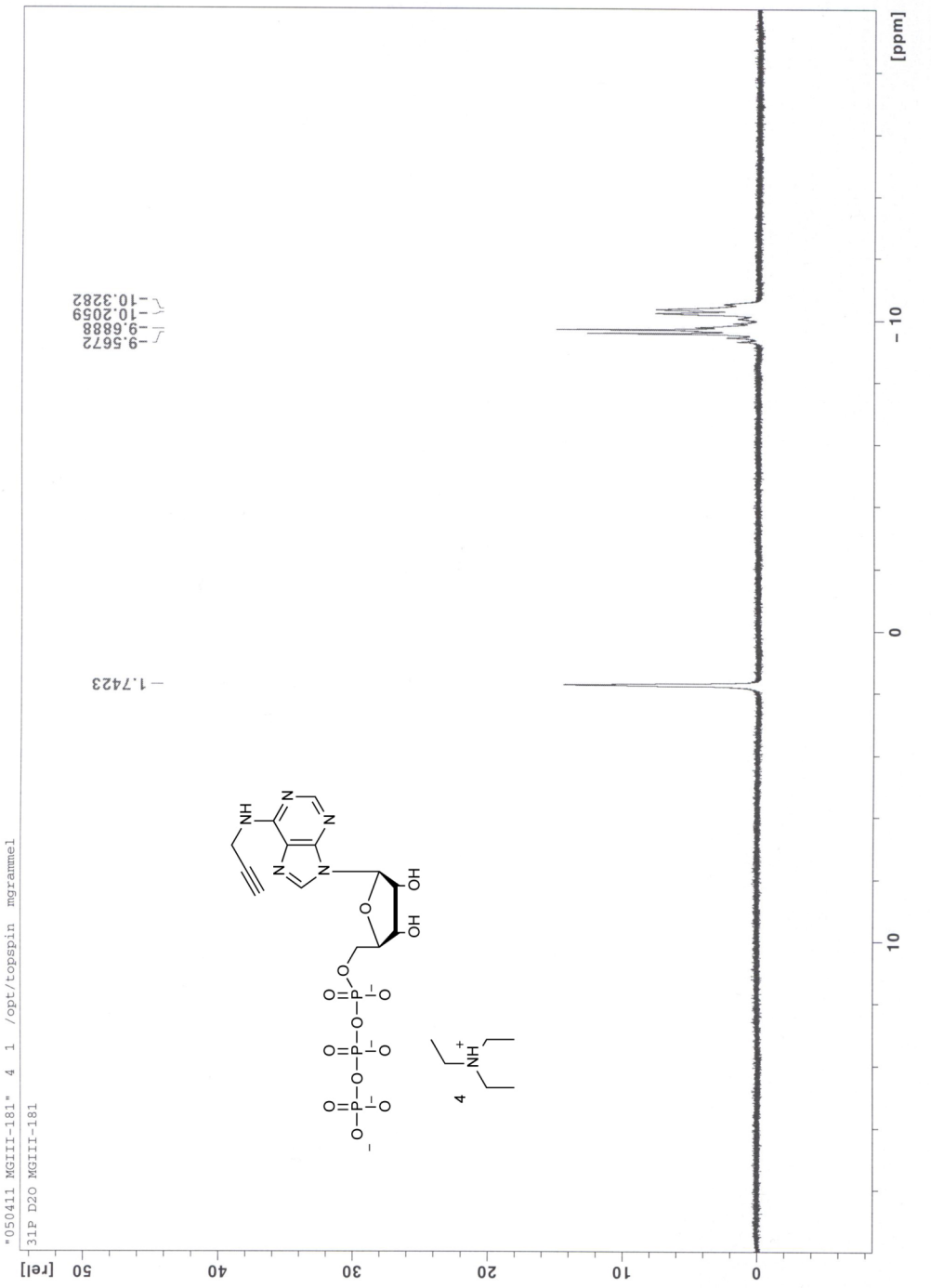
9.3.12 (R)-2-amino-3-(prop-2-yn-1-ylthio)propanoic acid (PCY)

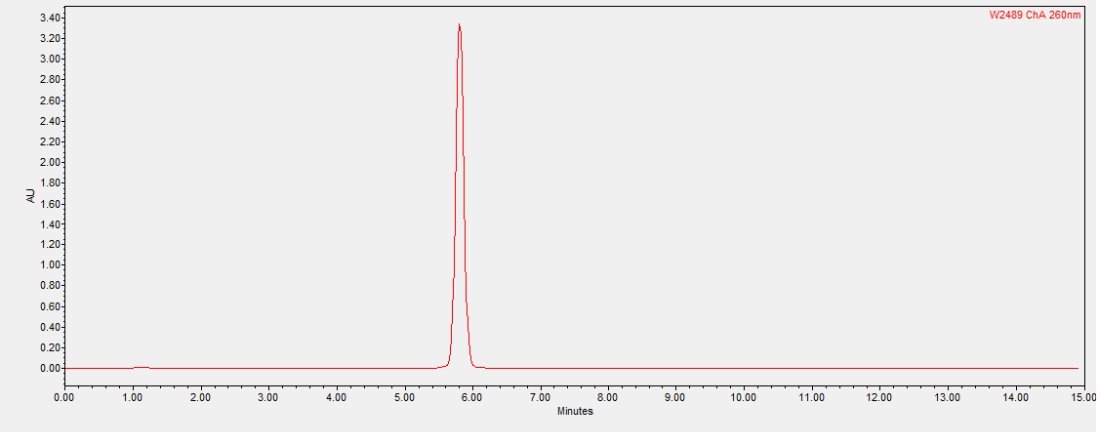



13C Direct Detection
172.5867
79.9803
72.6165
53.2740
31.7941
18.7586
COOH
S
H2N
[rel]
0
20
40
60
80
0
50
100
150
200
[ppm]

9.3.13 N^6-propargyl-adenosine-5'-triphosphate




"050411 MGIII-181" 4 1 /opt/topspin mgrammel
31P D2O MGIII-181
-1.7423
-9.5672
-9.6888
-10.2059
-10.3282
4
[rel]
[ppm]



Analytical HPLC trace for N^6-propargyl-adenosine-5'-triphosphate. Product was eluted with a linear gradient, using 20 mM triethylammonium bicarbonate (pH 7.4, acetic acid, 2% acetonitrile) and 100% acetonitrile and the absorbance was measured at 260 nm.

10 References

Aebersold, R. & Mann, M., 2003. Mass spectrometry-based proteomics. *Nature*, 422(6928), pp.198–207.

Allen, J.J. et al., 2007. A semisynthetic epitope for kinase substrates. *Nature Methods*, 4(6), pp.511–516.

Allen, J.J., Lazerwith, S.E. & Shokat, Kevan M, 2005. Bio-orthogonal affinity purification of direct kinase substrates. *Journal of the American Chemical Society*, 127(15), pp.5288–5289.

Anderson, C.T., Wallace, I.S. & Somerville, C.R., 2012. Metabolic click-labeling with a fucose analog reveals pectin delivery, architecture, and dynamics in Arabidopsis cell walls. *Proceedings of the National Academy of Sciences of the United States of America*, 109(4), pp.1329–1334.

Andrews, G.L. et al., 2011. Improving proteome coverage on a LTQ-Orbitrap using design of experiments. *Journal of the American Society for Mass Spectrometry*, 22(4), pp.773–783.

Ansong, C. et al., 2009. Global systems-level analysis of Hfq and SmpB deletion mutants in Salmonella: implications for virulence and global protein translation. *PloS One*, 4(3), p.e4809.

Apponi, L.H. et al., 2010. Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Human Molecular Genetics*, 19(6), pp.1058–1065.

Arnold, R. et al., 2009. Sequence-based prediction of type III secreted proteins. *PLoS Pathogens*, 5(4), p.e1000376.

van Asten, A.J.A.M. & van Dijk, J.E., 2005. Distribution of "classic" virulence factors among Salmonella spp. *FEMS Immunology and Medical Microbiology*, 44(3), pp.251–259.

Babu, M.M. et al., 2006. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *Journal of Bacteriology*, 188(8), pp.2761–2773.

Baskin, J.M. et al., 2010. Visualizing enveloping layer glycans during zebrafish early embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 107(23), pp.10360–10365.

Beatty, K.E. et al., 2005. Selective dye-labeling of newly synthesized proteins in bacterial cells. *Journal of the American Chemical Society*, 127(41), pp.14150–14151.

Beatty, K.E. & Tirrell, D.A., 2008. Two-color labeling of temporally defined protein populations in mammalian cells. *Bioorganic & Medicinal Chemistry Letters*, 18(22), pp.5995–5999.

Becker, D. et al., 2006. Robust Salmonella metabolism limits possibilities for new antimicrobials. *Nature*, 440(7082), pp.303–307.

Bedford, M.T. & Clarke, S.G., 2009. Protein arginine methylation in mammals: who, what, and why. *Molecular Cell*, 33(1), pp.1–13.

Berndt, N., Hamilton, A.D. & Sebti, S.M., 2011. Targeting protein prenylation for cancer therapy. *Nature Reviews. Cancer*, 11(11), pp.775–791.

Bertozzi, C R & Kiessling, L.L., 2001. Chemical glycobiology. *Science (New York, N.Y.)*, 291(5512), pp.2357–2364.

Bertozzi, Carolyn R, 2011. A decade of bioorthogonal chemistry. *Accounts of Chemical Research*, 44(9), pp.651–653.

Bhavsar, A.P., Guttman, J.A. & Finlay, B Brett, 2007. Manipulation of host-cell pathways by bacterial pathogens. *Nature*, 449(7164), pp.827–834.

Binda, O. et al., 2011. A chemical method for labeling lysine methyltransferase substrates. *Chembiochem: A European Journal of Chemical Biology*, 12(2), pp.330–334.

Blethrow, J.D. et al., 2008. Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates. *Proceedings of the National Academy of Sciences of the United States of America*, 105(5), pp.1442–1447.

Bock, V.D., Hiemstra, H. & van Maarseveen, J.H., 2006. CuI-Catalyzed Alkyne-Azide "Click" Cycloadditions from a Mechanistic and Synthetic Perspective. *European Journal of Organic Chemistry*, 2006(1), pp.51–68.

Boyce, M. & Bertozzi, Carolyn R, 2011. Bringing chemistry to life. *Nature Methods*, 8(8), pp.638–642.

Carlson, H.K. et al., 2010. Use of a semisynthetic epitope to probe histidine kinase activity and regulation. *Analytical Biochemistry*, 397(2), pp.139–143.

Champney, W.S. & Jensen, R.A., 1970. Molecular events in the growth inhibition of Bacillus subtilis by D-tyrosine. *Journal of Bacteriology*, 104(1), pp.107–116.

Chang, P.V. et al., 2009. Metabolic labeling of sialic acids in living animals with alkynyl sugars. *Angewandte Chemie (International Ed. in English)*, 48(22), pp.4030–4033.

Charron, G. et al., 2011. Alkynyl-farnesol reporters for detection of protein S-prenylation in cells. *Molecular bioSystems*, 7(1), pp.67–73.

Charron, G. et al., 2009a. Robust fluorescent detection of protein fatty-acylation with chemical reporters. *Journal of the American Chemical Society*, 131(13), pp.4967–4975.

Charron, G. et al., 2009b. Robust fluorescent detection of protein fatty-acylation with chemical reporters. *Journal of the American Chemical Society*, 131(13), pp.4967–4975.

Chen, Y. et al., 2007. Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Molecular & Cellular Proteomics: MCP*, 6(5), pp.812–819.

Chenault, H.K., Dahmer, J. & Whitesides, G.M., 1989. Kinetic resolution of unnatural and rarely occurring amino acids: enantioselective hydrolysis of N-acyl amino acids catalyzed by acylase I. *Journal of the American Chemical Society*, 111(16), pp.6354–6364.

Chock, P.B. et al., 1973. Epsilon-adenylylated glutamine synthetase: an internal fluorescence probe for enzyme conformation. *Proceedings of the National Academy of Sciences of the United States of America*, 70(11), pp.3134–3138.

Choudhary, C. et al., 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science (New York, N.Y.)*, 325(5942), pp.834–840.

Choudhary, C. & Mann, M., 2010. Decoding signalling networks by mass spectrometry-based proteomics. *Nature Reviews. Molecular Cell Biology*, 11(6), pp.427–439.

Clark, P.M. et al., 2008. Direct in-gel fluorescence detection and cellular imaging of O-GlcNAc-modified proteins. *Journal of the American Chemical Society*, 130(35), pp.11576–11577.

Coburn, B., Sekirov, I. & Finlay, B Brett, 2007. Type III secretion systems and disease. *Clinical Microbiology Reviews*, 20(4), pp.535–549.

Conrad, N.K. et al., 2007. Mutational analysis of a viral RNA element that counteracts rapid RNA decay by interaction with the polyadenylate tail. *Proceedings of the National Academy of Sciences of the United States of America*, 104(25), pp.10412–10417.

Coombes, B.K., Hardwidge, P.R. & Finlay, B Brett, 2004. Interpreting the host-pathogen dialogue through microarrays. *Advances in Applied Microbiology*, 54, pp.291–331.

Cooper, G.M. & Shendure, J., 2011. Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nature Reviews. Genetics*, 12(9), pp.628–640.

Cox, J. & Mann, M., 2007. Is proteomics the new genomics? *Cell*, 130(3), pp.395–398.

Dale, J.A. & Mosher, H.S., 1973. Nuclear magnetic resonance enantiomer regents. Configurational correlations via nuclear magnetic resonance chemical shifts of diastereomeric mandelate, O-methylmandelate, and .alpha.-methoxy-.alpha.-trifluoromethylphenylacetate (MTPA) esters. *Journal of the American Chemical Society*, 95(2), pp.512–519.

Dalhoff, C. et al., 2006. Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases. *Nature Chemical Biology*, 2(1), pp.31–32.

DeGraw, A.J. et al., 2010. Evaluation of alkyne-modified isoprenoids as chemical reporters of protein prenylation. *Chemical Biology & Drug Design*, 76(6), pp.460–471.

Deal, R.B., Henikoff, J.G. & Henikoff, S., 2010. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science (New York, N.Y.)*, 328(5982), pp.1161–1164.

DiGiuseppe Champion, P.A. & Cox, J.S., 2007. Protein secretion systems in Mycobacteria. *Cellular Microbiology*, 9(6), pp.1376–1384.

Dieterich, D.C. et al., 2010. In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nature Neuroscience*, 13(7), pp.897–905.

Dieterich, D.C. et al., 2006. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proceedings of the National Academy of Sciences of the United States of America*, 103(25), pp.9482–9487.

Dignam, J.D., Lebovitz, R.M. & Roeder, R.G., 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Research*, 11(5), pp.1475–1489.

Dion, M.F. et al., 2007. Dynamics of replication-independent histone turnover in budding yeast. *Science (New York, N.Y.)*, 315(5817), pp.1405–1408.

Dixon, S.J. et al., 2009. Systematic mapping of genetic interaction networks. *Annual Review of Genetics*, 43, pp.601–625.

Du, J. et al., 2011. Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science (New York, N.Y.)*, 334(6057), pp.806–809.

Du, J., Jiang, H. & Lin, H., 2009. Investigating the ADP-ribosyltransferase activity of sirtuins with NAD analogues and 32P-NAD. *Biochemistry*, 48(13), pp.2878–2890.

Dumont, A. et al., 2012. Click-Mediated Labeling of Bacterial Membranes through Metabolic Modification of the Lipopolysaccharide Inner Core. *Angewandte Chemie (International Ed. in English)*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22323101 [Accessed March 8, 2012].

Eckmann, C.R., Rammelt, C. & Wahle, E., 2011. Control of poly(A) tail length. *Wiley Interdisciplinary Reviews. RNA*, 2(3), pp.348–361.

Ellermeier, C.D. & Slauch, J.M., 2003. RtsA and RtsB coordinately regulate expression of the invasion and flagellar genes in Salmonella enterica serovar Typhimurium. *Journal of Bacteriology*, 185(17), pp.5096–5108.

Ellis-Davies, G.C.R., 2007. Caged compounds: photorelease technology for control of cellular chemistry and physiology. *Nature Methods*, 4(8), pp.619–628.

Eriksson, S. et al., 2003. Unravelling the biology of macrophage infection by gene expression profiling of intracellular Salmonella enterica. *Molecular Microbiology*, 47(1), pp.103–118.

Farazi, T.A., Waksman, G. & Gordon, J.I., 2001. The biology and enzymology of protein N-myristoylation. *The Journal of Biological Chemistry*, 276(43), pp.39501–39504.

Fields, P.I. et al., 1986. Mutants of Salmonella typhimurium that cannot survive within the macrophage are avirulent. *Proceedings of the National Academy of Sciences of the United States of America*, 83(14), pp.5189–5193.

Filloux, A., Hachani, A. & Bleves, S., 2008. The bacterial type VI secretion machine: yet another player for protein transport across membranes. *Microbiology (Reading, England)*, 154(Pt 6), pp.1570–1583.

Finlay, B B, Heffron, F & Falkow, S., 1989. Epithelial cell surfaces induce Salmonella proteins required for bacterial adherence and invasion. *Science (New York, N.Y.)*, 243(4893), pp.940–943.

Friedman, M., 1999. Chemistry, nutrition, and microbiology of D-amino acids. *Journal of Agricultural and Food Chemistry*, 47(9), pp.3457–3479.

Gao, X. et al., 2011. Membrane targeting of palmitoylated Wnt and Hedgehog revealed by chemical probes. *FEBS Letters*, 585(15), pp.2501–2506.

Gauchet, C., Labadie, G.R. & Poulter, C.D., 2006. Regio- and chemoselective covalent immobilization of proteins through unnatural amino acids. *Journal of the American Chemical Society*, 128(29), pp.9274–9275.

Ghaemmaghami, S. et al., 2003. Global analysis of protein expression in yeast. *Nature*, 425(6959), pp.737–741.

Gkogkas, C., Sonenberg, N. & Costa-Mattioli, M., 2010. Translational control mechanisms in long-lasting synaptic plasticity and memory. *The Journal of Biological Chemistry*, 285(42), pp.31913–31917.

Goddard-Borger, E.D. & Stick, R.V., 2007. An efficient, inexpensive, and shelf-stable diazotransfer reagent: imidazole-1-sulfonyl azide hydrochloride. *Organic Letters*, 9(19), pp.3797–3800.

de Godoy, L.M.F. et al., 2008. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*, 455(7217), pp.1251–1254.

Goldberg, A.D., Allis, C.D. & Bernstein, E., 2007. Epigenetics: a landscape takes shape. *Cell*, 128(4), pp.635–638.

Grammel, M. et al., 2011. A chemical reporter for protein AMPylation. *Journal of the American Chemical Society*, 133(43), pp.17103–17105.

Grammel, M. et al., 2012. Cell-selective labeling of bacterial proteomes with an orthogonal phenylalanine amino acid reporter. *Chemical Communications (Cambridge, England)*, 48(10), pp.1473–1474.

Grammel, M., Conrad, N.K. & Hang, H.C., 2012. Chemical reporters for monitoring RNA synthesis and poly(A) tail dynamics. *Chembiochem*, April19.

Grammel, M., Zhang, Mingzi M & Hang, H.C., 2010. Orthogonal alkynyl amino acid reporter for selective labeling of bacterial proteomes during infection. *Angewandte Chemie (International Ed. in English)*, 49(34), pp.5970–5974.

Groisman, E.A. & Ochman, H., 1996. Pathogenicity islands: bacterial evolution in quantum leaps. *Cell*, 87(5), pp.791–794.

Guan, L. et al., 2011. Intracellular detection of cytosine incorporation in genomic DNA by using 5-ethynyl-2'-deoxycytidine. *Chembiochem: A European Journal of Chemical Biology*, 12(14), pp.2184–2190.

Guttman, M. & Rinn, J.L., 2012. Modular regulatory principles of large non-coding RNAs. *Nature*, 482(7385), pp.339–346.

Hancock, S.M. et al., 2010. Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *Journal of the American Chemical Society*, 132(42), pp.14819–14824.

Hang, H.C. et al., 2003. A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. *Proceedings of the National Academy of Sciences of the United States of America*, 100(25), pp.14846–14851.

Hang, H.C. et al., 2007. Chemical probes for the rapid detection of Fatty-acylated proteins in Mammalian cells. *Journal of the American Chemical Society*, 129(10), pp.2744–2745.

Hang, H.C. & Linder, M.E., 2011. Exploring protein lipidation with chemical biology. *Chemical Reviews*, 111(10), pp.6341–6358.

Hang, H.C., Wilson, J.P. & Charron, G., 2011. Bioorthogonal chemical reporters for analyzing protein lipidation and lipid trafficking. *Accounts of Chemical Research*, 44(9), pp.699–708.

Hannoush, R.N. & Arenas-Ramirez, N., 2009. Imaging the lipidome: omega-alkynyl fatty acids for detection and cellular visualization of lipid-modified proteins. *ACS Chemical Biology*, 4(7), pp.581–587.

Hansen-Wester, I. & Hensel, M, 2001. Salmonella pathogenicity islands encoding type III secretion systems. *Microbes and Infection / Institut Pasteur*, 3(7), pp.549–559.

Hanson, S.R. et al., 2007. Tailored glycoproteomics and glycan site mapping using saccharide-selective bioorthogonal probes. *Journal of the American Chemical Society*, 129(23), pp.7266–7267.

Hao, Y.-H. et al., 2011. Characterization of a rabbit polyclonal antibody against threonine-AMPylation. *Journal of Biotechnology*, 151(3), pp.251–254.

Haraga, A., Ohlson, M.B. & Miller, Samuel I, 2008. Salmonellae interplay with host cells. *Nature Reviews. Microbiology*, 6(1), pp.53–66.

Hart, G.W. et al., 2011. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annual Review of Biochemistry*, 80, pp.825–858.

Hassa, P.O. et al., 2006. Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? *Microbiology and Molecular Biology Reviews: MMBR*, 70(3), pp.789–829.

He, Y.-F. et al., 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science (New York, N.Y.)*, 333(6047), pp.1303–1307.

Heal, W.P. et al., 2011. Bioorthogonal chemical tagging of protein cholesterylation in living cells. *Chemical Communications*, 47(14), p.4081.

Heal, W.P. et al., 2008. N-Myristoyl transferase-mediated protein labelling in vivo. *Organic & Biomolecular Chemistry*, 6(13), p.2308.

Heiman, M. et al., 2008. A translational profiling approach for the molecular characterization of CNS cell types. *Cell*, 135(4), pp.738–748.

Henikoff, S., 2008. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nature Reviews. Genetics*, 9(1), pp.15–26.

Himo, F. et al., 2005. Copper(I)-Catalyzed Synthesis of Azoles. DFT Study Predicts Unprecedented Reactivity and Intermediates. *Journal of the American Chemical Society*, 127(1), pp.210–216.

Hinz, F.I. et al., 2012. Non-canonical amino acid labeling in vivo to visualize and affinity purify newly synthesized proteins in larval zebrafish. *ACS Chemical Neuroscience*, 3(1), pp.40–49.

Hottiger, M.O. et al., 2010. Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends in Biochemical Sciences*, 35(4), pp.208–219.

Huisgen, R., 1963. 1,3-Dipolar Cycloadditions. Past and Future. *Angewandte Chemie International Edition in English*, 2(10), pp.565–598.

Ibba, M., Kast, P. & Hennecke, H., 1994. Substrate specificity is determined by amino acid binding pocket size in Escherichia coli phenylalanyl-tRNA synthetase. *Biochemistry*, 33(23), pp.7107–7112.

Ingolia, N.T., Lareau, L.F. & Weissman, J.S., 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, 147(4), pp.789–802.

Islam, K. et al., 2011. Expanding cofactor repertoire of protein lysine methyltransferase for substrate labeling. *ACS Chemical Biology*, 6(7), pp.679–684.

Issartel, J.P., Koronakis, V. & Hughes, C., 1991. Activation of Escherichia coli prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. *Nature*, 351(6329), pp.759–761.

Ito, S. et al., 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science (New York, N.Y.)*, 333(6047), pp.1300–1303.

Jao, C.Y. & Salic, A., 2008. Exploring RNA transcription and turnover in vivo by using click chemistry. *Proceedings of the National Academy of Sciences of the United States of America*, 105(41), pp.15779–15784.

Jewett, J.C. & Bertozzi, Carolyn R, 2010. Cu-free click cycloaddition reactions in chemical biology. *Chemical Society Reviews*, 39(4), pp.1272–1279.

Jiang, H. et al., 2010. Clickable NAD analogues for labeling substrate proteins of poly(ADP-ribose) polymerases. *Journal of the American Chemical Society*, 132(27), pp.9363–9372.

Jiang, H. et al., 2009. Mechanism-based small molecule probes for labeling CD38 on live cells. *Journal of the American Chemical Society*, 131(5), pp.1658–1659.

Johnson, J.A. et al., 2010. Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Current Opinion in Chemical Biology*, 14(6), pp.774–780.

Jones, B.D. & Falkow, S., 1996. Salmonellosis: host immune responses and bacterial virulence determinants. *Annual Review of Immunology*, 14, pp.533–561.

Jung, H., O'Hare, C.M. & Holt, C.E., 2011. Translational regulation in growth cones. *Current Opinion in Genetics & Development*, 21(4), pp.458–464.

Kang, R. et al., 2008. Neural palmitoyl-proteomics reveals dynamic synaptic palmitoylation. *Nature*, 456(7224), pp.904–909.

Karnoub, A.E. & Weinberg, R.A., 2008. Ras oncogenes: split personalities. *Nature Reviews. Molecular Cell Biology*, 9(7), pp.517–531.

Khidekel, N. et al., 2003. A chemoenzymatic approach toward the rapid and sensitive detection of O-GlcNAc posttranslational modifications. *Journal of the American Chemical Society*, 125(52), pp.16162–16163.

Khidekel, N. et al., 2007. Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics. *Nature Chemical Biology*, 3(6), pp.339–348.

Kho, Y. et al., 2004. A tagging-via-substrate technology for detection and proteomics of farnesylated proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 101(34), pp.12479–12484.

Kiick, K.L. et al., 2002. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proceedings of the National Academy of Sciences of the United States of America*, 99(1), pp.19–24.

Kinch, Lisa N et al., 2009. Fido, a novel AMPylation domain common to fic, doc, and AvrB. *PloS One*, 4(6), p.e5818.

Kingdon, H.S., Shapiro, B.M. & Stadtman, E.R., 1967a. Regulation of glutamine synthetase. 8. ATP: glutamine synthetase adenylyltransferase, an enzyme that catalyzes alterations in the regulatory properties of glutamine synthetase. *Proceedings of the National Academy of Sciences of the United States of America*, 58(4), pp.1703–1710.

Kingdon, H.S., Shapiro, B.M. & Stadtman, E.R., 1967b. Regulation of glutamine synthetase. 8. ATP: glutamine synthetase adenylyltransferase, an enzyme that catalyzes alterations in the regulatory properties of glutamine synthetase. *Proceedings of the National Academy of Sciences of the United States of America*, 58(4), pp.1703–1710.

Kirshenbaum, K., Carrico, I.S. & Tirrell, D.A., 2002. Biosynthesis of proteins incorporating a versatile set of phenylalanine analogues. *Chembiochem: A European Journal of Chemical Biology*, 3(2-3), pp.235–237.

Kodama, T. et al., 2008. Identification of two translocon proteins of Vibrio parahaemolyticus type III secretion system 2. *Infection and Immunity*, 76(9), pp.4282–4289.

Koenigs, M.B., Richardson, E.A. & Dube, D.H., 2009. Metabolic profiling of Helicobacter pylori glycosylation. *Molecular bioSystems*, 5(9), pp.909–912.

Kouzarides, T., 2007. Chromatin modifications and their function. *Cell*, 128(4), pp.693–705.

Kuhle, V. & Hensel, Michael, 2002. SseF and SseG are translocated effectors of the type III secretion system of Salmonella pathogenicity island 2 that modulate aggregation of endosomal compartments. *Cellular Microbiology*, 4(12), pp.813–824.

Laughlin, S.T. et al., 2008a. In vivo imaging of membrane-associated glycans in developing zebrafish. *Science (New York, N.Y.)*, 320(5876), pp.664–667.

Laughlin, S.T. et al., 2008b. In vivo imaging of membrane-associated glycans in developing zebrafish. *Science (New York, N.Y.)*, 320(5876), pp.664–667.

Laughlin, S.T. & Bertozzi, Carolyn R, 2009a. Imaging the glycome. *Proceedings of the National Academy of Sciences of the United States of America*, 106(1), pp.12–17.

Laughlin, S.T. & Bertozzi, Carolyn R, 2009b. In vivo imaging of Caenorhabditis elegans glycans. *ACS Chemical Biology*, 4(12), pp.1068–1072.

Laughlin, S.T. & Bertozzi, Carolyn R, 2009c. In vivo imaging of Caenorhabditis elegans glycans. *ACS Chemical Biology*, 4(12), pp.1068–1072.

Leonard, S.E., Reddie, K.G. & Carroll, K.S., 2009. Mining the thiol proteome for sulfenic acid modifications reveals new targets for oxidation in cells. *ACS Chemical Biology*, 4(9), pp.783–799.

Li, Yan et al., 2011. Characterization of AMPylation on threonine, serine, and tyrosine using mass spectrometry. *Journal of the American Society for Mass Spectrometry*, 22(4), pp.752–761.

Lin, S. et al., 2011. Site-specific incorporation of photo-cross-linker and bioorthogonal amino acids into enteric bacterial pathogens. *Journal of the American Chemical Society*, 133(50), pp.20581–20587.

Linder, M.E. & Deschenes, R.J., 2007. Palmitoylation: policing protein stability and traffic. *Nature Reviews. Molecular Cell Biology*, 8(1), pp.74–84.

Link, A.J. et al., 2006. Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids. *Proceedings of the National Academy of Sciences of the United States of America*, 103(27), pp.10180–10185.

Link, A.J., Vink, M.K.S. & Tirrell, D.A., 2004. Presentation and detection of azide functionality in bacterial cell surface proteins. *Journal of the American Chemical Society*, 126(34), pp.10598–10602.

Liu, C.C. & Schultz, P.G., 2010. Adding new chemistries to the genetic code. *Annual Review of Biochemistry*, 79, pp.413–444.

Liu, F. et al., 2009. The engineering of bacteria bearing azido-pseudaminic acid-modified flagella. *Chembiochem: A European Journal of Chemical Biology*, 10(8), pp.1317–1320.

Liu, J. et al., 2012. Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin. *Proceedings of the National Academy of Sciences of the United States of America*, 109(2), pp.413–418.

Loflin, P.T. et al., 1999. Transcriptional pulsing approaches for analysis of mRNA turnover in mammalian cells. *Methods (San Diego, Calif.)*, 17(1), pp.11–20.

Lossi, N.S. et al., 2008. The Salmonella SPI-2 effector SseJ exhibits eukaryotic activator-dependent phospholipase A and glycerophospholipid: cholesterol acyltransferase activity. *Microbiology (Reading, England)*, 154(Pt 9), pp.2680–2688.

Luchansky, S.J., Goon, Scarlett & Bertozzi, Carolyn R, 2004. Expanding the diversity of unnatural cell-surface sialic acids. *Chembiochem: A European Journal of Chemical Biology*, 5(3), pp.371–374.

Ludwig, J., 1981. A new route to nucleoside 5'-triphosphates. *Acta Biochimica Et Biophysica; Academiae Scientiarum Hungaricae*, 16(3-4), pp.131–133.

Luong, P. et al., 2010a. Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *The Journal of Biological Chemistry*, 285(26), pp.20155–20163.

Luong, P. et al., 2010b. Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *The Journal of Biological Chemistry*, 285(26), pp.20155–20163.

López-Otín, C. & Hunter, T., 2010. The regulatory crosstalk between kinases and proteases in cancer. *Nature Reviews. Cancer*, 10(4), pp.278–292.

Macek, B., Mann, M. & Olsen, J.V., 2009. Global and site-specific quantitative phosphoproteomics: principles and applications. *Annual Review of Pharmacology and Toxicology*, 49, pp.199–221.

Martin, B.R. et al., 2011. Global profiling of dynamic protein palmitoylation. *Nature Methods*, 9(1), pp.84–89.

Martin, B.R. & Cravatt, B.F., 2009. Large-scale profiling of protein palmitoylation in mammalian cells. *Nature Methods*, 6(2), pp.135–138.

Martin, D.D.O. et al., 2008. Rapid detection, discovery, and identification of post-translationally myristoylated proteins during apoptosis using a bio-orthogonal azidomyristate analog. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, 22(3), pp.797–806.

Marzluff, W.F., Wagner, E.J. & Duronio, R.J., 2008. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nature Reviews. Genetics*, 9(11), pp.843–854.

McDermott, J.E. et al., 2011. Computational prediction of type III and IV secreted effectors in gram-negative bacteria. *Infection and Immunity*, 79(1), pp.23–32.

McGhie, E.J. et al., 2009. Salmonella takes control: effector-driven manipulation of the host. *Current Opinion in Microbiology*, 12(1), pp.117–124.

Mercer, A.C. et al., 2009. In vivo modification of native carrier protein domains. *Chembiochem: A European Journal of Chemical Biology*, 10(6), pp.1091–1100.

Meta, M. et al., 2006. Protein farnesyltransferase inhibitors and progeria. *Trends in Molecular Medicine*, 12(10), pp.480–487.

Milne, S.B. et al., 2010. Capture and release of alkyne-derivatized glycerophospholipids using cobalt chemistry. *Nature Chemical Biology*, 6(3), pp.205–207.

Mito, Y., Henikoff, J.G. & Henikoff, S., 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nature Genetics*, 37(10), pp.1090–1097.

Mueller, C.A., Broz, P. & Cornelis, G.R., 2008. The type III secretion system tip complex and translocon. *Molecular Microbiology*, 68(5), pp.1085–1095.

Mukherjee, S., Hao, Y.-H. & Orth, K., 2007. A newly discovered post-translational modification--the acetylation of serine and threonine residues. *Trends in Biochemical Sciences*, 32(5), pp.210–216.

Müller, M.P. et al., 2010a. The Legionella effector protein DrrA AMPylates the membrane traffic regulator Rab1b. *Science (New York, N.Y.)*, 329(5994), pp.946–949.

Müller, M.P. et al., 2010b. The Legionella effector protein DrrA AMPylates the membrane traffic regulator Rab1b. *Science (New York, N.Y.)*, 329(5994), pp.946–949.

Neef, A.B. & Luedtke, N.W., 2011. Dynamic metabolic labeling of DNA in vivo with arabinosyl nucleosides. *Proceedings of the National Academy of Sciences of the United States of America*, 108(51), pp.20404–20409.

Neef, A.B. & Schultz, Carsten, 2009. Selective fluorescence labeling of lipids in living cells. *Angewandte Chemie (International Ed. in English)*, 48(8), pp.1498–1500.

Nessen, M.A. et al., 2009. Selective enrichment of azide-containing peptides from complex mixtures. *Journal of Proteome Research*, 8(7), pp.3702–3711.

Neunuebel, M.R. et al., 2011. De-AMPylation of the small GTPase Rab1 by the pathogen Legionella pneumophila. *Science (New York, N.Y.)*, 333(6041), pp.453–456.

Ngo, J.T. et al., 2009. Cell-selective metabolic labeling of proteins. *Nature Chemical Biology*, 5(10), pp.715–717.

Ngo, J.T. & Tirrell, D.A., 2011. Noncanonical amino acids in the interrogation of cellular protein synthesis. *Accounts of Chemical Research*, 44(9), pp.677–685.

Nomura, D.K., Dix, M.M. & Cravatt, B.F., 2010. Activity-based protein profiling for biochemical pathway discovery in cancer. *Nature Reviews. Cancer*, 10(9), pp.630–638.

Nurse, P. & Hayles, J., 2011. The cell in an era of systems biology. *Cell*, 144(6), pp.850–854.

Ohl, M.E. & Miller, S I, 2001. Salmonella: a model for bacterial pathogenesis. *Annual Review of Medicine*, 52, pp.259–274.

Palanivelu, D.V. et al., 2011. Fic domain catalyzed adenylylation: Insight provided by the structural analysis of the type IV secretion system effector BepA. *Protein Science: A Publication of the Protein Society*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21213248.

Parker, R. & Song, H., 2004. The enzymes and control of eukaryotic mRNA turnover. *Nature Structural & Molecular Biology*, 11(2), pp.121–127.

Paulsen, C.E. et al., 2012. Peroxide-dependent sulfenylation of the EGFR catalytic site enhances kinase activity. *Nature Chemical Biology*, 8(1), pp.57–64.

Pawson, T. & Scott, J.D., 2005. Protein phosphorylation in signaling--50 years and counting. *Trends in Biochemical Sciences*, 30(6), pp.286–290.

Peters, W. et al., 2010. Enzymatic site-specific functionalization of protein methyltransferase substrates with alkynes for click labeling. *Angewandte Chemie (International Ed. in English)*, 49(30), pp.5170–5173.

Pezacki, J.P. et al., 2011. Chemical contrast for imaging living systems: molecular vibrations drive CARS microscopy. *Nature Chemical Biology*, 7(3), pp.137–145.

Pizarro-Cerdá, J. & Cossart, P., 2006. Bacterial adhesion and entry into host cells. *Cell*, 124(4), pp.715–727.

Plant, J. & Glynn, A.A., 1979. Locating salmonella resistance gene on mouse chromosome 1. *Clinical and Experimental Immunology*, 37(1), pp.1–6.

Polevoda, B. & Sherman, F., 2003. N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins. *Journal of Molecular Biology*, 325(4), pp.595–622.

Pratt, M.R. et al., 2004. Deconvoluting the functions of polypeptide N-alpha-acetylgalactosaminyltransferase family members by glycopeptide substrate profiling. *Chemistry & Biology*, 11(7), pp.1009–1016.

Prescher, J.A. & Bertozzi, Carolyn R, 2005. Chemistry in living systems. *Nature Chemical Biology*, 1(1), pp.13–21.

Prescher, J.A. & Bertozzi, Carolyn R, 2006. Chemical technologies for probing glycans. *Cell*, 126(5), pp.851–854.

Prescher, J.A., Dube, D.H. & Bertozzi, Carolyn R, 2004. Chemical remodelling of cell surfaces in living animals. *Nature*, 430(7002), pp.873–877.

Rabani, M. et al., 2011. Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nature Biotechnology*, 29(5), pp.436–442.

Rabuka, D. et al., 2006. A chemical reporter strategy to probe glycoprotein fucosylation. *Journal of the American Chemical Society*, 128(37), pp.12078–12079.

Ramachandran, N. et al., 2008. Next-generation high-density self-assembling functional protein arrays. *Nature Methods*, 5(6), pp.535–538.

Rangan, K.J. et al., 2010. Rapid visualization and large-scale profiling of bacterial lipoproteins with chemical reporters. *Journal of the American Chemical Society*, 132(31), pp.10628–10629.

Rathert, P. et al., 2008. Specificity of protein lysine methyltransferases and methods for detection of lysine methylation of non-histone proteins. *Molecular bioSystems*, 4(12), pp.1186–1190.

Resh, M.D., 2006. Trafficking and signaling by fatty-acylated and prenylated proteins. *Nature Chemical Biology*, 2(11), pp.584–590.

Rexach, J.E. et al., 2012. Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. *Nature Chemical Biology*, 8(3), pp.253–261.

Rexach, J.E. et al., 2010. Quantification of O-glycosylation stoichiometry and dynamics using resolvable mass tags. *Nature Chemical Biology*, 6(9), pp.645–651.

Rhee, S.G. et al., 1981. Fluorometric studies of aza-epsilon-adenylylated glutamine synthetase from Escherichia coli. *The Journal of Biological Chemistry*, 256(12), pp.6010–6016.

Ripmaster, T.L., Shiba, K. & Schimmel, P., 1995. Wide cross-species aminoacyl-tRNA synthetase replacement in vivo: yeast cytoplasmic alanine enzyme replaced by human polymyositis serum antigen. *Proceedings of the National Academy of Sciences of the United States of America*, 92(11), pp.4932–4936.

Rodionov, V.O., Fokin, V.V. & Finn, M.G., 2005. Mechanism of the ligand-free CuI-catalyzed azide-alkyne cycloaddition reaction. *Angewandte Chemie (International Ed. in English)*, 44(15), pp.2210–2215.

Rodland, K.D. et al., 2008. Use of high-throughput mass spectrometry to elucidate host-pathogen interactions in Salmonella. *Future Microbiology*, 3(6), pp.625–634.

Rohde, K.H., Abramovitch, R.B. & Russell, D.G., 2007. Mycobacterium tuberculosis invasion of macrophages: linking bacterial gene expression to environmental cues. *Cell Host & Microbe*, 2(5), pp.352–364.

Rosenkrands, I. et al., 2000. Towards the proteome of Mycobacterium tuberculosis. *Electrophoresis*, 21(17), pp.3740–3756.

Rostovtsev, V.V. et al., 2002a. A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angewandte Chemie (International Ed. in English)*, 41(14), pp.2596–2599.

Rostovtsev, V.V. et al., 2002b. A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angewandte Chemie (International Ed. in English)*, 41(14), pp.2596–2599.

Roth, A.F. et al., 2006. Global analysis of protein palmitoylation in yeast. *Cell*, 125(5), pp.1003–1013.

Salic, A. & Mitchison, T.J., 2008. A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 105(7), pp.2415–2420.

Sallés, F.J., Richards, W.G. & Strickland, S., 1999. Assaying the polyadenylation state of mRNAs. *Methods (San Diego, Calif.)*, 17(1), pp.38–45.

Samudrala, R., Heffron, Fred & McDermott, J.E., 2009. Accurate prediction of secreted substrates and identification of a conserved putative secretion signal for type III secretion systems. *PLoS Pathogens*, 5(4), p.e1000375.

Sankaran, K. & Wu, H.C., 1994. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *The Journal of Biological Chemistry*, 269(31), pp.19701–19706.

Sawa, M. et al., 2006. Glycoproteomic probes for fluorescent imaging of fucosylated glycans in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 103(33), pp.12371–12376.

Saxon, E. & Bertozzi, C R, 2000. Cell surface engineering by a modified Staudinger reaction. *Science (New York, N.Y.)*, 287(5460), pp.2007–2010.

Schlumberger, M.C. & Hardt, W.-D., 2006. Salmonella type III secretion effectors: pulling the host cell's strings. *Current Opinion in Microbiology*, 9(1), pp.46–54.

Schmidt, M.-J. & Norbury, C.J., 2010. Polyadenylation and beyond: emerging roles for noncanonical poly(A) polymerases. *Wiley Interdisciplinary Reviews. RNA*, 1(1), pp.142–151.

Schultz, C et al., 1993. Acetoxymethyl esters of phosphates, enhancement of the permeability and potency of cAMP. *The Journal of Biological Chemistry*, 268(9), pp.6316–6322.

Schwanhäusser, B. et al., 2011. Global quantification of mammalian gene expression control. *Nature*, 473(7347), pp.337–342.

Seet, B.T. et al., 2006. Reading protein modifications with interaction domains. *Nature Reviews. Molecular Cell Biology*, 7(7), pp.473–483.

Shapiro, B.M. & Stadtman, E.R., 1970. The regulation of glutamine synthesis in microorganisms. *Annual Review of Microbiology*, 24, pp.501–524.

Shi, L. et al., 2006. Proteomic analysis of Salmonella enterica serovar typhimurium isolated from RAW 264.7 macrophages: identification of a novel protein that contributes to the replication of serovar typhimurium inside macrophages. *The Journal of Biological Chemistry*, 281(39), pp.29131–29140.

Shogren-Knaak, M.A., Alaimo, P.J. & Shokat, K M, 2001. Recent advances in chemical approaches to the study of biological systems. *Annual Review of Cell and Developmental Biology*, 17, pp.405–433.

Sletten, E.M. & Bertozzi, Carolyn R, 2011. From mechanism to mouse: a tale of two bioorthogonal reactions. *Accounts of Chemical Research*, 44(9), pp.666–676.

Sletten, E.M. & Bertozzi, Carolyn R, 2009. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angewandte Chemie (International Ed. in English)*, 48(38), pp.6974–6998.

Song, C.-X. et al., 2011. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nature Biotechnology*, 29(1), pp.68–72.

Speers, A.E. & Cravatt, B.F., 2005. A tandem orthogonal proteolysis strategy for high-content chemical proteomics. *Journal of the American Chemical Society*, 127(28), pp.10018–10019.

Speers, A.E. & Cravatt, B.F., 2004. Profiling enzyme activities in vivo using click chemistry methods. *Chemistry & Biology*, 11(4), pp.535–546.

Sprung, R. et al., 2005. Tagging-via-substrate strategy for probing O-GlcNAc modified proteins. *Journal of Proteome Research*, 4(3), pp.950–957.

Szychowski, J. et al., 2010. Cleavable biotin probes for labeling of biomolecules via azide-alkyne cycloaddition. *Journal of the American Chemical Society*, 132(51), pp.18351–18360.

Tan, M. et al., 2011. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*, 146(6), pp.1016–1028.

Tan, Y. & Luo, Z.-Q., 2011. Legionella pneumophila SidD is a deAMPylase that modifies Rab1. *Nature*, 475(7357), pp.506–509.

Tanrikulu, I.C. et al., 2009. Discovery of Escherichia coli methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 106(36), pp.15285–15290.

Tarrant, M.K. & Cole, P.A., 2009. The chemical biology of protein phosphorylation. *Annual Review of Biochemistry*, 78, pp.797–825.

Tcherkezian, J. et al., 2010. Transmembrane receptor DCC associates with protein synthesis machinery and regulates translation. *Cell*, 141(4), pp.632–644.

Tekaia, F. & Yeramian, E., 2006. Evolution of proteomes: fundamental signatures and global trends in amino acid compositions. *BMC Genomics*, 7, p.307.

Tekaia, F., Yeramian, E. & Dujon, B., 2002. Amino acid composition of genomes, lifestyles of organisms, and evolutionary trends: a global picture with correspondence analysis. *Gene*, 297(1-2), pp.51–60.

Tornøe, C.W., Christensen, C. & Meldal, M., 2002. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(i)-catalyzed 1,3-dipolar cycloadditions of

terminal alkynes to azides. *The Journal of Organic Chemistry*, 67(9), pp.3057–3064.

Tsou, L.K., Zhang, Mingzi M & Hang, H.C., 2009. Clickable fluorescent dyes for multimodal bioorthogonal imaging. *Organic & Biomolecular Chemistry*, 7(24), pp.5055–5058.

Uttamapinant, C. et al., 2010. A fluorophore ligase for site-specific protein labeling inside living cells. *Proceedings of the National Academy of Sciences of the United States of America*, 107(24), pp.10914–10919.

Vander Heiden, M.G., Cantley, L.C. & Thompson, C.B., 2009. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science (New York, N.Y.)*, 324(5930), pp.1029–1033.

Varki, A., 1999. *Essentials of glycobiology*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press [Bethesda, Md.; National Center for Biotechnology Information]. Available at: http://www.usc.edu/hsc/nml/e-resources/info/essgly.html [Accessed March 18, 2012].

Vocadlo, D.J. et al., 2003. A chemical approach for identifying O-GlcNAc-modified proteins in cells. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16), pp.9116–9121.

Walsh, C.T., Garneau-Tsodikova, S. & Gatto, G.J., Jr, 2005. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angewandte Chemie (International Ed. in English)*, 44(45), pp.7342–7372.

Wan, J. et al., 2007. Palmitoylated proteins: purification and identification. *Nature Protocols*, 2(7), pp.1573–1584.

Wang, F., Robbins, S., et al., 2010. Genetic incorporation of unnatural amino acids into proteins in Mycobacterium tuberculosis. *PloS One*, 5(2), p.e9354.

Wang, L. & Schultz, P.G., 2004. Expanding the genetic code. *Angewandte Chemie (International Ed. in English)*, 44(1), pp.34–66.

Wang, Q., Zhang, Y., et al., 2010. Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science (New York, N.Y.)*, 327(5968), pp.1004–1007.

Wang, R., Zheng, W., et al., 2011. Labeling substrates of protein arginine methyltransferase with engineered enzymes and matched S-adenosyl-L-methionine analogues. *Journal of the American Chemical Society*, 133(20), pp.7648–7651.

Wang, Y., Zhang, Q., et al., 2011. High-accuracy prediction of bacterial type III secreted effectors based on position-specific amino acid composition profiles. *Bioinformatics (Oxford, England)*, 27(6), pp.777–784.

Wasylnka, J.A. et al., 2008. Role for myosin II in regulating positioning of Salmonella-containing vacuoles and intracellular replication. *Infection and Immunity*, 76(6), pp.2722–2735.

Wilson, J.P. et al., 2011. Proteomic analysis of fatty-acylated proteins in mammalian cells with chemical reporters reveals S-acylation of histone H3 variants. *Molecular & Cellular Proteomics: MCP*, 10(3), p.M110.001198.

Wilusz, J. & Shenk, T., 1988. A 64 kd nuclear protein binds to RNA segments that include the AAUAAA polyadenylation motif. *Cell*, 52(2), pp.221–228.

Winnen, B. et al., 2008. Hierarchical effector protein transport by the Salmonella Typhimurium SPI-1 type III secretion system. *PloS One*, 3(5), p.e2178.

Wiśniewski, J.R. et al., 2009. Universal sample preparation method for proteome analysis. *Nature Methods*, 6(5), pp.359–362.

Wollack, J.W. et al., 2009. A minimalist substrate for enzymatic peptide and protein conjugation. *Chembiochem: A European Journal of Chemical Biology*, 10(18), pp.2934–2943.

Woolery, A.R. et al., 2010. AMPylation: something old is new again. , 1, p.113.

Worby, C.A. et al., 2009a. The fic domain: regulation of cell signaling by adenylylation. *Molecular Cell*, 34(1), pp.93–103.

Worby, C.A. et al., 2009b. The fic domain: regulation of cell signaling by adenylylation. *Molecular Cell*, 34(1), pp.93–103.

Worby, C.A. et al., 2009c. The fic domain: regulation of cell signaling by adenylylation. *Molecular Cell*, 34(1), pp.93–103.

Xiao, J. et al., 2010. Structural basis of Fic-mediated adenylylation. *Nature Structural & Molecular Biology*, 17(8), pp.1004–1010.

Yahr, T.L. et al., 1997. Identification of type III secreted products of the Pseudomonas aeruginosa exoenzyme S regulon. *Journal of Bacteriology*, 179(22), pp.7165–7168.

Yamakoshi, H. et al., 2011. Imaging of EdU, an alkyne-tagged cell proliferation probe, by Raman microscopy. *Journal of the American Chemical Society*, 133(16), pp.6102–6105.

Yamashita, A. et al., 2005. Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nature Structural & Molecular Biology*, 12(12), pp.1054–1063.

Yang, J. et al., 2008. VFDB 2008 release: an enhanced web-based resource for comparative pathogenomics. *Nucleic Acids Research*, 36(Database issue), pp.D539–542.

Yang, X.-J. & Seto, E., 2008. Lysine acetylation: codified crosstalk with other posttranslational modifications. *Molecular Cell*, 31(4), pp.449–461.

Yang, Y.-Y., Grammel, M., et al., 2010. Comparative analysis of cleavable azobenzene-based affinity tags for bioorthogonal chemical proteomics. *Chemistry & Biology*, 17(11), pp.1212–1222.

Yang, Y.-Y. et al., 2010. Comparative analysis of cleavable azobenzene-based affinity tags for bioorthogonal chemical proteomics. *Chemistry & Biology*, 17(11), pp.1212–1222.

Yang, Y.-Y. et al., 2011. Identification of lysine acetyltransferase p300 substrates using 4-pentynoyl-coenzyme A and bioorthogonal proteomics. *Bioorganic & Medicinal Chemistry Letters*, 21(17), pp.4976–4979.

Yang, Y.-Y., Ascano, J.M. & Hang, H.C., 2010. Bioorthogonal chemical reporters for monitoring protein acetylation. *Journal of the American Chemical Society*, 132(11), pp.3640–3641.

Yao, J.Z. et al., 2012. Fluorophore Targeting to Cellular Proteins via Enzyme-Mediated Azide Ligation and Strain-Promoted Cycloaddition. *Journal of the American Chemical Society*, 134(8), pp.3720–3728.

Yap, M.C. et al., 2010. Rapid and selective detection of fatty acylated proteins using omega-alkynyl-fatty acids and click chemistry. *Journal of Lipid Research*, 51(6), pp.1566–1580.

Yarbrough, M. L., Li, Y., Kinch, L. N., et al., 2009. AMPylation of Rho GTPases by Vibrio VopS Disrupts Effector Binding and Downstream Signaling. *Science*, 323(5911), pp.269–272.

Yarbrough, Melanie L, Li, Yan, et al., 2009. AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling. *Science (New York, N.Y.)*, 323(5911), pp.269–272.

Yarbrough, Melanie L et al., 2009. AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling. *Science (New York, N.Y.)*, 323(5911), pp.269–272.

Yarema, K.J., Goon, S & Bertozzi, C R, 2001. Metabolic selection of glycosylation defects in human cells. *Nature Biotechnology*, 19(6), pp.553–558.

Yi, L. et al., 2010. A highly efficient strategy for modification of proteins at the C terminus. *Angewandte Chemie (International Ed. in English)*, 49(49), pp.9417–9421.

Yoshikawa, M., Kato, T. & Takenishi, T., 1967. A novel method for phosphorylation of nucleosides to 5'-nucleotides. *Tetrahedron Letters*, 50, pp.5065–5068.

Yount, J.S. et al., 2010. Palmitoylome profiling reveals S-palmitoylation–dependent antiviral activity of IFITM3. *Nature Chemical Biology*, 6(8), pp.610–614.

Yu, X.-J. et al., 2010. pH sensing by intracellular Salmonella induces effector translocation. *Science (New York, N.Y.)*, 328(5981), pp.1040–1043.

Zaro, B.W., Yang, Y.-Y., et al., 2011. Chemical reporters for fluorescent detection and identification of O-GlcNAc-modified proteins reveal glycosylation of the ubiquitin ligase NEDD4-1. *Proceedings of the National Academy of Sciences of the United States of America*, 108(20), pp.8146–8151.

Zaro, B.W., Bateman, L.A. & Pratt, M.R., 2011. Robust in-gel fluorescence detection of mucin-type O-linked glycosylation. *Bioorganic & Medicinal Chemistry Letters*, 21(17), pp.5062–5066.

Zhang, C., Kenski, D.M., et al., 2005. A second-site suppressor strategy for chemical genetic analysis of diverse protein kinases. *Nature Methods*, 2(6), pp.435–441.

Zhang, C.G., Chromy, B.A. & McCutchen-Maloney, S.L., 2005. Host-pathogen interactions: a proteomic view. *Expert Review of Proteomics*, 2(2), pp.187–202.

Zhang, M. M. et al., 2010. Tandem fluorescence imaging of dynamic S-acylation and protein turnover. *Proceedings of the National Academy of Sciences*, 107(19), pp.8627–8632.

Zhang, Mingzi M, in preparation

Zhang, Z. et al., 2011. Identification of lysine succinylation as a new post-translational modification. *Nature Chemical Biology*, 7(1), pp.58–63.

Zheng, T. et al., 2011. Tracking N-acetyllactosamine on cell-surface glycans in vivo. *Angewandte Chemie (International Ed. in English)*, 50(18), pp.4113–4118.

Zhou, D. & Galán, J., 2001. Salmonella entry into host cells: the work in concert of type III secreted effector proteins. *Microbes and Infection / Institut Pasteur*, 3(14-15), pp.1293–1298.