



TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Chemie

Lehrstuhl für Organische Chemie II

A chemical proteomic platform to elucidate infection-mediated AMPylation of the human proteome

Dissertation zur Erlangung des akademischen Grades eines Doktors

der Naturwissenschaften von

Theresa Rauh

München 2021



TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Chemie

Lehrstuhl für Organische Chemie II

A chemical proteomic platform to elucidate infection-mediated AMPylation of the human proteome

Theresa Rauh

Vollständiger Abdruck der von der Fakultät der Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Franz Hagn
Prüfer der Dissertation: 1. Prof. Dr. Stephan A. Sieber
2. Prof. Dr. Kathrin Lang

Diese Dissertation wurde am 04.03.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 21.04.2021 angenommen.

„Was du für den Gipfel hältst, ist nur eine Stufe“

Seneca

Danksagung

Am Anfang der Doktorarbeit hatte ich das Gefühl am Fuße eines großen Berges zu stehen, welcher nur darauf wartet über die nächsten Jahre hinweg erklommen zu werden. Es eröffneten sich viele unterschiedliche Wege die Wanderung anzugehen, und oft war es nicht abzusehen, wobei es sich um eine Abkürzung, und wo um einen Umweg handelte. Ohne genaue Karte startete ich die spannende Expedition, und möchte mich an dieser Stelle zuerst bei Prof. Dr. Stephan A. Sieber bedanken, welcher mir die nötige Ausrüstung für solch eine Unternehmung zur Seite gestellt hat. Auch war er immer eine hilfreiche Unterstützung, wenn die Orientierung nicht mehr ganz da und der Weg nicht mehr zu sehen war. Als nächstes geht mein Dank an Pavel, durch den ich zur AMPylierung gekommen bin, erst während meiner Masterarbeit, und anschließend mit meinen Infektionsassays. Danke dir, dass du mir stets mit Rat und Tat zur Seite warst, wenn der Weg auf den ersten Blick als unbegebar erschien. Danke auch für die Diskussionen und jeden ermutigenden Zuspruch. Deine Ruhe und dein Optimismus waren immer ansteckend!

Wenn man sich auf solch eine lange Reise begibt, sollte eigentlich alles an notwendiger Ausrüstung dabei sein. Mangelt es aber doch mal an der ein oder anderen Stelle, so kann man sich glücklicherweise mit Kooperationen behelfen. An dieser Stelle möchte ich mich besonders bei Dr. Sophie Brameyer und Prof. Dr. Kirsten Jung von der LMU für die gute Kommunikation und erfolgreiche Zusammenarbeit bedanken. Ohne euch hätte ich mein Paper nicht so schnell auf den Weg gebracht. Ein weiteres Dankeschön in diesem Zusammenhang geht an Dr. Benjamin Winkeljann und Clara Lettl. Vielen Dank auch euch für eine erfolgreiche Zusammenarbeit, welche uns alle ein Stück vorangebracht hat.

Wie bei allen Expeditions-Berichten bedarf es auch bei dieser Arbeit am Ende einer kritischen Hinterfragung. Deshalb danke ich ganz herzlich Pavel, Jan, Stuy und Isabel für die schnelle und gewissenhafte Korrektur dieser Arbeit. Danke, dass ihr euch die Zeit dafür genommen habt.

Wer allein wandert kommt vielleicht abschnittsweise schneller voran. In der Gesellschaft von Freunden wandert es sich jedoch viel gelassener, und Langeweile hat keine Chance. So wird der Berg Meter um Meter erklommen, ohne dass die Anstrengung bemerkt wird. An dieser Stelle möchte ich mich bei allen bedanken, die mit mir gewandert sind. Ihr seid super Kollegen und Freunde, und habt die Zeit am AK erst du der gemacht, die sie war! Ein großes Dankeschön geht an Babsi und Lisa für die allmorgentlichen Runden Tee, die immer viel zu schnell wieder vorbei waren. Vielen Dank für all unsere Gespräche und die schöne Zeit auf dem Tollwood, beim Spazieren gehen oder beim Wandern. Sobald Corona vorbei ist, gibt es hoffentlich wieder

mehr davon! Ein weiteres Riesendankeschön geht an alle, die mit mir ein ums andere Mal den inneren Schweinehund erfolgreich bekämpft haben. Ohne euch hätte ich es weder um 5.40 Uhr aus dem Bett geschafft um pünktlich um sieben in Fröttmaning zu stehen, wäre keinen Gewittern davon geradelt noch wäre ich freiwillig so lange im Kreis gerannt, bis mir die Zunge zum Boden hängt. Ob bei den Fitness Stunden bei Peter Weiß, den Inliner-Runden, dem Radeln an die Uni oder den morgendlichen Boulder-Sessions, vielen lieben Dank an Caro, Babsi, Volker, Thomas, Patta, Zani, Stuy und Till. Es war immer eine super Zeit mit euch, die nicht selten in Muskelkater geendet hat, auch wenn man manchmal nur einer Staubwolke hinterhergejagt ist.

Gegenseitige Unterstützung ist beim Wandern sehr wichtig. Darum geht ein weiteres, großes Dankeschön an alle Leidensgenossen in Labor C, Martin, Angela, Jan und ThAngela, die auch cool geblieben sind, wenn der -80er mal wieder aus dem letzten Loch gepfiffen hat. Danke für die super Unterstützung und Arbeitsatmosphäre im Lab, und danke für den nicht enden wollenden Süßigkeiten-Vorrat! Die sauren Würmer haben mich durch den ein oder anderen Tag gebracht. Danke auch an Seppi, der mich schon durch die Höhen und Tiefen des Masters begleitet hat, und der, wie prophezeit, nach der Masterarbeit doch noch den Weg an den AK gefunden hat. Nur was Sonntagswanderungen betrifft ist dir nicht immer zu trauen ;)

Des Weiteren geht ein riesengroßes Dankeschön an alle weiteren Weggefährten des AK Sieber und AK Hacker: Alex, Davide, Didi, Ines, Isabel, Johnny, Konstantin, Markus, Martin K, Micha, Michi, Nina, Robert, Stephan and Thomas M. Ich möchte mich bei euch allen für feuchtfröhliche Grillabende oder Kaffee-Küchen-Abende, Ausflüge an den Echinger See, die lustige Zeit im Kleinwalsertal und die super Arbeitsatmosphäre bedanken. Bei jeder Frage oder Problemstellung findet man bei uns stets Hilfe und den passenden Ansprechpartner, was nicht selbstverständlich ist. Das gleiche gilt für das ehemalige Expeditionsteam, namentlich Dóri, Frani, Hacki, Christian, Kyu, Anja, Philipp, Markus, Vadim, Anni, Matthias und Lena.

Bei einer langen Wanderung sind Durchhaltevermögen und Gesellschaft allerdings nicht alles, auch der organisatorische und technische Support ist nicht zu unterschätzen. An dieser Stelle möchte ich mich besonders bei Christina, Mona und Katja B. bedanken. Auch bedanke ich mich bei Katja G. und unseren Azubis Linda, Alina, Marie, Isabella und Tanja. Vielen Dank für euren täglichen Einsatz, ohne euch würde im Labor nichts mehr laufen. Dir liebe Christina wünsche ich für deine Zukunft alles Gute, du wirst am AK vermisst werden.

Es ist hilfreich und zeitsparend, wenn nicht jede Wegbiegung einzeln abgegangen und erkundet werden muss, sondern wenn hier Helfer zur Hand sind. Deshalb bedanke ich mich sehr bei

meinen ehemaligen Forschungspraktikanten und Bacheloranten Christopher, Anita, Laura, Niklas, Sara und Monika. Vielen Dank euch für die Unterstützung bei der Synthese (trotz penetranten Gestanks) und der Auswertung etlicher NMR-Daten.

Auch beim Wandern kann es Höhen und Tiefen geben, wobei es von Vorteil ist, einen starken Rückhalt zu haben. Ich bedanke mich zuletzt an dieser Stelle bei meiner Familie und bei Thomas. Danke für ALLES! Danke für eure permanente Unterstützung, für's wieder aufhelfen, wenn ich mal ausgerutscht bin, für eure Motivation, für's Da-Sein und für's Zuhören – auch wenn ihr euch vermutlich manchmal fragt, was ich eigentlich den ganzen Tag in diesem Labor so mache. Ein großes Dankeschön geht auch an Mareike und Martin. Mit euch macht Wandern einfach am meisten Spaß! Danke auch allen anderen Freunden aus der Heimat oder vom Studium! Ich bin froh, dass ich euch alle habe!

Die letzten Höhenmeter dieser Wanderung sind die steilsten. Es bedarf einiges an Durchhaltevermögen, um die Wochen des finalen Diss-Schreibens am Schreibtisch zu meistern. Ist das kleine Plateau kurz unter den Gipfel erreicht, fehlt nur noch ein kurzer Anstieg zum erfolgreichen Abschluss. Am Ende ist der Gipfel endlich da, mit einer grandiosen Aussicht. Mit einem Lächeln denke ich an den Aufstieg zurück. Es war eine super schöne Zeit, die doch ganz schnell vorbeigegangen ist. Jetzt ist auch deutlich zu erkennen, was während des Aufstiegs von Wald und Wolken teilweise verborgen war: der Gipfel ist als Teil eines ausgebauten Höhenweges noch lange nicht das Ende. Von hier ergeben sich wieder eine Fülle an neuen Möglichkeiten, man muss sich nur für einen Weg entscheiden.

Table of contents

Abstract.....	III
Zusammenfassung	V
Introductory Remarks	VII
List of tables	IX
List of figures	XI
List of abbreviations	XIII
1. Introduction	1
1.1. Attenuating pathogenic bacteria – a possible entry point	2
1.2. Post-translational modifications – a leverage point for pathogens	3
1.3. AMPylation – the revival of a post-translational modification	5
1.3.1. ATase domains	6
1.3.2. Fic domains	7
1.4. AMPylators in metazoans	9
1.5. Bacterial Fic domain-containing proteins act intra- and extracellular	9
1.6. <i>V. parahaemolyticus</i> VopS – an AMPylator as tool to promote infection	12
1.7. A range of tools to detect AMPylated proteins.....	14
1.8. MS-based proteomic profiling	17
1.9. Scope of this work	20
2. Results and discussions	24
2.1. Optimization of pro-N6pA labeling conditions	25
2.2. Intrinsically AMPylated proteins in HeLa cells	26
2.3. Identification and optimization of suitable infection conditions	28
2.4. Construction and validation of the inactive <i>V. parahaemolyticus</i> mutant VopS-H348A	33
2.5. Assessment of different infection and quantification strategies	34
2.5.1. Protein enrichment experiments according to approach A combined with LFQ	34
2.5.2. Protein enrichment experiments according to approach B combined with DML ...	36
2.5.3. Protein enrichment / full proteome analysis according to approach A combined with SILAC labeling	38
2.5.4. Protein enrichment experiments according to approach B combined with LFQ	42
2.5.5. Protein enrichment experiments according to approach C combined with LFQ	48
2.6. Human host AMPylation targets of other pathogenic bacteria	50
2.7. Screening of other human cell lines for VopS-mediated AMPylation targets	52
2.8. Validation of Rho GTPases as VopS targets	55

3. Conclusion and outlook	59
4. Experimental section.....	64
4.1. Buffers	64
4.2. Cell culture	64
4.2.1. General cell culture procedures	65
4.2.2. Thawing and cryopreservation of cells	65
4.2.3. Cellular maintenance	66
4.3. Bacterial growth	66
4.3.1. Strains and media	66
4.3.2. Bacterial maintenance	68
4.4. Proteomic labeling strategies.....	68
4.4.1. Probe treatment	68
4.4.2. Analytical <i>in situ</i> labeling	69
4.4.3. Preparative <i>in situ</i> labeling.....	70
4.5. MS/MS measurements and data analysis	75
4.6. Biochemical assays.....	78
4.6.1. CFU assays.....	78
4.6.2. AMPylation assay	78
4.6.3. Intact protein MS measurement (IPMS)	79
4.6.4. SDS-PAGE of recombinantly AMPylated proteins.....	79
4.6.5. Site-ID of recombinant proteins.....	80
5. References	81
6. Appendix	95
7. Licenses	103

Abstract

Currently, we all are struggling with the fast spread of a new pandemic, COVID-19, which is believed to be one of the greatest threats to human health in modern times. Effective mRNA vaccines have presented as a solution to this disease, whose development and market approval took place within one year. Regarding another major threat to humanity, the antibiotic crisis, the opposite picture is emerging. We face a strong rise of multidrug-resistant bacteria. The emergence of such strains is favored by bacterial transfer of resistance genes, which also affect clinically important pathogenic strains, including *Staphylococcus aureus*, *Pseudomonas aeruginosa* and others. As the development of new antibiotics is unprofitable, there is a lack thereof. Nevertheless, to stay on top of the pathogenic strains, it is essential to understand their virulence strategies and the cellular consequences caused by their infection or invasion. New treatment options that only deprive bacteria their pathogenicity and do not kill them represent a promising alternative to antibiotics.

In this thesis, the focus is on the human pathogenic bacterium *Vibrio parahaemolyticus*. The Gram-negative bacterium was observed to apply its T3SS1 (type III secretion system) effector protein VopS to AMPylate human host proteins once it is docked to the human cell. VopS belongs to the larger protein family of Fic domain-containing proteins. These proteins are highly conserved throughout all domains of life and thus present a valuable starting point for the exploration of bacterial infection strategies and possible intervention points.

In this work, an assay combining a suitable infection strategy with high resolution mass spectrometry to track VopS mediated AMPylation targets in the human proteome under physiological conditions was established. For this purpose, human cells were treated with the cell permeable adenosine derivative probe **pro-N6pA** and infected with bacteria afterward. By comparing different MS quantification methods including label-free quantification, SILAC or dimethyl-labeling, and by testing different infection strategies the optimal workflow was developed. The comparison of human cells infected with the wild type bacterium *versus* cells infected with a mutated strain expressing an inactive AMPylator followed by LFQ analysis revealed seven Rho GTPases as VopS targets. Through the usage of a TEV-cleavable linker, the AMPylation sites of three Rho GTPases could be mapped to conserved threonine residues.

Further studies involving other human cell lines showed, that VopS mediated AMPylation is cell type specific and that it varies between cell types originating from different tissues.

By transferring the approach to other human pathogens including *P. aeruginosa* strain PAO1, several human proteins emerged as potential targets of the corresponding bacterial AMPylator. While further investigation in this area is required, the established infection assay proves to be a valuable tool to gain deeper insights into bacterial infection mechanisms. Further studies should focus on tissue or cell type specific AMPylation targets on the one hand, and on the other hand explore the potential of additional bacterial AMPylators. With more than 72 000 proteins belonging to the Fic domain-containing protein family, some of them are very likely to perform similar tasks to VopS and harm the human cell during infection.

Overall, this work introduces a proteomic profiling strategy that enables the identification of bacterial AMPylation targets in the human proteome under physiological conditions. Furthermore, this method also has the potential of a broader applicability regarding the exploration of other bacterially induced post-translational modifications.

Zusammenfassung

Derzeit ringt die Menschheit mit der Ausbreitung einer neuen Pandemie, COVID-19, die als eine der größten Bedrohungen der Neuzeit gilt. Als Mittel gegen diese Krankheit wurden wirksame mRNA-Impfstoffe vorgestellt, deren Entwicklung und Marktzulassung innerhalb eines Jahres erfolgten. Bei einer anderen großen Bedrohung der Menschheit, der Antibiotika-Krise, zeichnet sich jedoch ein gegenteiliges Bild ab. Es findet eine starke Zunahme an multiresistenten Bakterienstämmen statt. Die Entstehung solcher Stämme wird durch die bakterielle Übertragung von Resistenzgenen begünstigt, die auch klinisch wichtige pathogene Stämme nicht verschonen, wie zum Beispiel *Staphylococcus aureus*, *Pseudomonas aeruginosa* und andere. Da die Entwicklung neuer Antibiotika nicht profitabel ist, wird Forschung in diesem Bereich vielerorts nicht weiterverfolgt. Um den pathogenen Stämmen jedoch weiterhin einen Schritt voraus zu sein ist es unerlässlich, ihre Virulenzstrategien und deren zelluläre Folgen zu verstehen. Neue Behandlungsmöglichkeiten, die den Bakterien nur ihre Pathogenität nehmen, sie jedoch nicht abtöten, stellen eine vielversprechende Alternative zu Antibiotika dar.

Diese Arbeit legt den Fokus auf das humanpathogene Bakterium *Vibrio parahaemolyticus*, welches mittels seines T3SS1 (Typ III Sekretionssystem) Effektorproteins VopS menschliche Wirtsproteine AMPyliert, sobald es an die menschliche Zelle andockt ist. VopS gehört zu der größeren Proteinfamilie der Fic-Proteine. Dies sind Proteine, welche die konservierte Fic-Sequenz beinhalten. Sie sind in allen Domänen des Lebens vorhanden und stellen somit einen wertvollen Ausgangspunkt für die Erforschung bakterieller Infektionsstrategien und daraus resultierender Beeinflussungsmöglichkeiten dar.

In dieser Arbeit wurde ein Assay etabliert, der eine geeignete Infektionsstrategie mit hochauflösender Massenspektrometrie kombiniert, um AMPylierungsziele von VopS im menschlichen Proteom unter physiologischen Bedingungen zu identifizieren. Zu diesem Zweck wurden menschliche Zellen mit der zellpermeablen Sonde **pro-N6pA** behandelt, welche ein Adenosinderivat darstellt, und anschließend mit Bakterien infiziert. Durch den Vergleich verschiedener MS-Quantifizierungsmethoden wie der markierungsfreien Quantifizierung, SILAC oder der Dimethyl-Markierung, sowie durch die Anwendung verschiedener Infektionsstrategien wurde ein optimaler Arbeitsablauf entwickelt. Dieser ermöglichte die Identifikation von sieben Rho-GTPasen als VopS-Targets. Durch die Verwendung eines TEV-

spaltbaren Linkers konnten konservierte Threonin-Reste von drei Rho-GTPasen als AMPylierungsstellen identifiziert werden. Weitere Studien mit anderen humanen Zelllinien zeigten, dass die VopS-vermittelte AMPylierung zelltypspezifisch ist und dass sie zwischen Zelltypen aus verschiedenen Geweben variiert.

Die Übertragung dieses Ansatzes auf andere Humanpathogene, einschließlich des *P. aeruginosa*-Stammes PAO1, identifizierte mehrere menschliche Proteine als potenzielle Ziele des entsprechenden bakteriellen AMPylators. Während weitere Untersuchungen in diesem Bereich erforderlich sind, erweist sich der etablierte Infektionsassay als wertvolles Werkzeug, um tiefere Einblicke in bakterielle Infektionsmechanismen zu gewinnen. Weitere Studien sollten sich zum einen auf gewebe- oder zelltypspezifische AMPylierungsziele konzentrieren und zum anderen das Potenzial weiterer bakterieller AMPylatoren erforschen. Bei mehr als 72 000 Mitgliedern der Fic-Proteinfamilie ist es sehr wahrscheinlich, dass einige von ihnen ähnliche Aufgaben wie VopS erfüllen und die menschliche Zelle während der Infektion schädigen.

Insgesamt stellt diese Arbeit eine proteomische Profilierungsstrategie vor, die die Identifizierung von bakteriellen AMPylierungszielen im menschlichen Proteom unter physiologischen Bedingungen ermöglicht. Darüber hinaus hat diese Methode auch das Potenzial einer breiteren Anwendbarkeit im Hinblick auf die Erforschung anderer bakteriell induzierter posttranslationaler Modifikationen.

Introductory Remarks

This dissertation was completed between January 2018 and February 2021 under the supervision of Prof. Dr. Stephan A. Sieber at the Chair of Organic Chemistry II at the Technical University of Munich.

Parts of this thesis have been published in:

Journal publication:

T. Rauh, S. Brameyer, P. Kielkowski, K. Jung, S. A. Sieber, “MS-Based *in Situ* Proteomics Reveals AMPylation of Host Proteins during Bacterial Infection”, *ACS Infect. Dis.* **2020**, *6*, 3277–3289.

Figures from the publication were adapted for this thesis and are reprinted with permission from <https://pubs.acs.org/doi/10.1021/acsinfecdis.0c00740>, Copyright (2020) American Chemical Society

Conference presentations:

T. Rauh, S. Brameyer, P. Kielkowski, K. Jung, S. A. Sieber, “Establishing a Proteomic Profiling Approach to Detect AMPylation During Bacterial Infection *in situ*”, AMPylation Plus conference, 2020

T. Rauh, S. Brameyer, P. Kielkowski, K. Jung, S. A. Sieber, “Establishing a Proteomic Profiling Approach to Detect AMPylation During Bacterial Infection *in situ*”, Chemical Biology EMBO Workshop, 2020

Publication not highlighted in this thesis:

B. Winkeljann, M. G. Bauer, M. Marczynski, T. Rauh, S. A. Sieber, O. Lieleg, “Covalent Mucin Coatings Form Stable Anti-Biofouling Layers on a Broad Range of Medical Polymer Materials”, *Adv. Mater. Interfaces* **2020**, *7*, 1902069.

Contributions:

Sophie Brameyer, [Ludwig Maximilian University of Munich (LMU)] designed and constructed the *V. parahaemolyticus* strain RIMD 2210633 VopS mutant H348A as well as the in frame deletion mutants of *E. coli* CFT073, c4136::Km and c4409::Km.

List of tables

Table 1: Significantly enriched proteins resulting from a LFQ proteome profiling experiment in HeLa cells treated with 100 μ M pro-N6pA or 0.1 vol% DMSO	27
Table 2: Enrichment data of VopS targets CDC42, RHOA and RAC1 for protein enrichment experiments using approach A in combination with LFQ analysis.....	35
Table 3: Enrichment data of VopS targets CDC42, RHOA and RAC1 for protein enrichment experiments using approach B in combination with DML	37
Table 4: Enrichment data of CDC42, RHOA and RAC1 infection experiments with <i>V. parahaemolyticus</i> wild type and mutant VopS-H348A. Results for full proteome analysis and protein enrichment experiments	41
Table 5: Summary of all proteins that met the threshold criteria plus RHOG, CDC42 and RAC1 in the proteomic profiling experiment infecting HeLa cells with <i>V. parahaemolyticus</i> wild type according to approach B and analyzed via LFQ	43
Table 6: Enrichment data of initially identified VopS targets in the infection experiment with <i>V. parahaemolyticus</i> mutant VopS-H348A using approach B in combination with LFQ.....	45
Table 7: Overview of performed experiments combining different infection strategies with different quantification methods	47
Table 8: Enrichment data of all proteins that met the threshold criteria plus mistakenly predicted VopS targets in protein profiling experiments according to approach C	50
Table 9: Enrichment data of PLXNA1 plus the VopS target proteins CDC42, RHOA, RAC1 and RHOG of the SH-SY5Y cell infection experiment with <i>V. parahaemolyticus</i>	53
Table 10: Enrichment data of PLXNA1 and the VopS target proteins CDC42, RHOG, RHOA, RAC1, and RHOC of the HaCat cell infection experiment with <i>V. parahaemolyticus</i>	55
Table 11: Buffers used in this thesis.....	64
Table 12: Information about the used human cell lines.....	65
Table 13: Summary of bacterial strains including culture conditions and sources.....	67

List of figures

Figure 1: Summary of important post-translational modifications, divided into five different classes.	3
Figure 2: Protein AMPylation on a protein's serine, tyrosine or threonine side chain with ATP as precursor and the subsequent release of pyrophosphate.	5
Figure 3: Mechanism of AMPylation by a Fic domain.	7
Figure 4: Overview of selected virulence strategies of <i>V. parahaemolyticus</i> .	12
Figure 5: Cellular consequences of VopS mediated AMPylation.	14
Figure 6: Different chemical probes applied during <i>in vitro</i> and <i>in situ</i> experiments to identify AMPylated proteins.	16
Figure 7: Cellular uptake of pro-N6pA followed by probe activation.	17
Figure 8: Schematic overview of a standard proteomic profiling workflow.	18
Figure 9: Graphical illustration of pro-N6pA uptake followed by its cellular conversion to functional N6pATP .	24
Figure 10: SDS-PAGE analysis of time- and concentration-dependent labeling of HeLa cells.	26
Figure 11: Intrinsically AMPylated proteins in human cells and CFU/mL to OD ₆₀₀ ratio for <i>V. parahaemolyticus</i> wild type.	29
Figure 12: Optimized infection workflow of pro-N6pA treated HeLa cells with <i>V. parahaemolyticus</i> RIMD 2210633 infection.	30
Figure 13: HeLa cell infection with <i>V. parahaemolyticus</i> .	31
Figure 14: AMPylation of Rho GTPases does not immediately lead to HeLa cell rounding.	32
Figure 15: CFU/mL to OD ₆₀₀ ratio for <i>V. parahaemolyticus</i> VopS-H348A and gel-based analysis thereof.	33
Figure 16: HeLa cell infection with <i>V. parahaemolyticus</i> II.	37
Figure 17: Volcano plots of infection experiments of HeLa cells using approach A with <i>V. parahaemolyticus</i> wild type and SILAC labeling.	39
Figure 18: Volcano plots of infection experiments of HeLa cells using approach A with <i>V. parahaemolyticus</i> VopS-H348A and SILAC labeling.	40
Figure 19: Protein enrichment plots of HeLa cell infection with <i>V. parahaemolyticus</i> wild type or mutant VopS-H348A.	45
Figure 20: Applying wild type bacteria and mutants expressing inactive AMPylators yielded the largest number of VopS target proteins.	49
Figure 21: Results of infection experiments of the two human cell lines SH-SY5Y and HaCat infected with <i>V. parahaemolyticus</i> wild type and the mutant VopS-H348A.	54

Figure 22: Confirmation of RHOG as a target of VopS and investigation of VopS AMPylation activity toward RHOC and RHOG <i>in vitro</i> and <i>in situ</i>	57
Figure 23: Summary and conclusion of the project.	62
Figure S1: Volcano plot of the <i>in situ</i> infection experiment of HeLa cells with <i>V. parahaemolyticus</i> mutant VopS-H348A.	95
Figure S2: Identification of suitable infection conditions and enrichment data of infection experiments with the Gram-positive bacterium <i>S. aureus</i> USA300.....	96
Figure S3: Identification of suitable infection conditions and enrichment data of infection experiments with the Gram-negative bacterium <i>P. aeruginosa</i> PAO1 and its two transposon mutants expressing the inactive AMPylators PW3486 genotype PA1366-A11::ISpho-A/hah (PA1366) as well as PW2059 genotype PA0574-A11::ISlacZ/hah (PA0574).....	98
Figure S4: Identification of suitable infection conditions and enrichment data of infection experiments with the Gram-negative bacterium <i>E. coli</i> CFT073 and its two in frame deletion mutants c4136::Km and c4409::Km expressing inactive AMPylators.....	101

List of abbreviations

ABPP	activity-based protein profiling
ADP	adenosine diphosphate
ADPRT	adenosine diphosphate ribosyl transferase
AMP	adenosine monophosphate
Asn	asparagine
Asp	aspartic acid
ATase	adenylyl transferase
ATP	adenosine triphosphate
BCA	bicinchoninic acid
BSA	bovine serum albumin
cDNA	complementary deoxyribonucleic acid
CFU	colony forming units
CID	collision induced dissociation
CTP	cytidine triphosphate
CuAAC	copper(I)-catalyzed azide-alkyne cycloaddition
Da	dalton
dd	double distilled
DMEM	Dulbecco's modified Eagle's medium
DML	dimethyl labeling
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	Exempli gratia
ER	endoplasmic reticulum
ESI	electro spray ionization
<i>et. al.</i>	et alii
ETD	electron-transfer dissociation
FA	formic acid
FBS	fetal bovine serum
FDR	false discovery rate
FP	full proteome analysis
GDP	guanosine diphosphate
Glu	glutamic acid
GTP	guanosine triphosphate
h	hours
HCD	higher-energy collision-induced dissociation
His	histidine
HR-MS	high resolution mass spectrometry
IAA	iodoacetamide
IC ₅₀	concentration of 50% inhibition
ICAT	isotope-coded affinity tags
IPMS	intact-protein mass spectrometry
iTRAQ	isobaric tags for relative and absolute quantification
kDa	kilodalton
LB	lysogeny broth
LC	liquid chromatography
LC-MS	liquid chromatography coupled to mass spectrometry
LFQ	label-free quantification

LMU	Ludwig Maximilian University
LUCA	last universal common ancestor
Lys	lysine
LysC	endoproteinase lys-C
MeOH	methanol
µg	micrograms
min	minutes
mL	milliliter
mM	millimolar
µM	micromolar
MOI	multiplicity of infection
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
m/z	mass to charge
NaCl	sodium chloride
NAPPA	nucleic acid programmable protein array
NMP	nucleotide monophosphate
OD	optical density
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
PBS	phosphate buffered saline
PE	protein enrichment
pro-N6pA	phosphoramidate probe
PTM	post-translational modification
pH	power of hydrogen
ppm	parts per million
RhN ₃	rhodamine azide
RNA	ribonucleic acid
rpm	revolutions per minute
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
SDS-PAGE	sodium dodecyl sulfate – polyacrylamide gel electrophoresis
Ser	serine
SILAC	stable isotope labeling with amino acids in ecell culture
STRING	search tool for the retrieval of interacting genes/proteins
T3SS	type III secretion system
T6SS	type VI secretion system
TBTA	tris(benzyltriazolylmethyl)amine
TCEP	tris (2-carboxyethyl)phosphine
TEAB	triethylammonium bicarbonate buffer
TFA	trifluoro acetic acid
Thr	threonine
TMT	tandem mass tag
Tyr	tyrosine
UDP	uridine diphosphate
UMP	uridine monophosphate
UPR	unfolded protein response
UTP	uridine triphosphate
<i>V. parahaemolyticus</i>	<i>Vibrio parahaemolyticus</i>
wt	wild type

1. Introduction

1. Introduction

Previous generations regularly suffered from deadly infectious diseases. Infections, that are straightforward to treat nowadays. In the late nineteenth century, Louis Pasteur developed his germ theory of diseases. With contributions from Robert Koch and other scientists, finally the knowledge was spread that there is a connection between bacteria and the infection characteristics that can be observed.¹ This marked the beginning of antibiotics research. First, Salvarsan was identified as an effective reagent against syphilis by Paul Ehrlich, then penicillin was discovered by Alexander Fleming.¹ Since then, a multitude of different antibiotics have been identified.^{1,2} However, due to selective pressure, the occurrence of antibiotic-resistant bacteria was usually reported not long after the corresponding market launch.² Nowadays, more than 70% of pathogens are resistant to at least one antibiotic. Low profit margin and high research costs further prevent pharmaceutical companies from developing new antibiotics,² and as a result, we are approaching a likely post-antibiotic era.³ Therefore, new strategies are needed to mitigate pathogenic bacteria.

1.1. Attenuating pathogenic bacteria – a possible entry point

Bacteria have developed resistance to every available antibiotic on the market.⁴ While efforts have been made to address this, such as alternative therapies like nanomedicine or *in silico* methods, these also lack in efficiency. In order to progress, deeper knowledge about bacterial virulence or pathogenicity factors is required.⁵ Post-translational modifications (PTMs) are of greatest biological relevance, e.g. in cellular physiology.⁶ Bacterial pathogens deploy these modifications as their tool to interact with the host's signaling pathways, thereby creating advantages for themselves.^{6,7} It is therefore necessary to gain a deeper understanding of how pathogens exploit PTMs. Maybe then, these modifications offer a new angle, which can be used to attenuate multi-resistant bacteria and to deprive their pathogenicity. However, PTMs are difficult to detect in the cellular context because of their dynamic nature. Moreover, the available protein pool is not uniformly modified.⁴

Proteomics combined with mass spectrometry (MS) analysis provide a valuable tool to analyze intricate protein samples like cell lysates or whole tissues.⁴ In addition, it is a suitable method to study PTMs, as different enrichment strategies have been established for them. Proteomics thus makes it possible to study a pathogen's use of post-translational modifications in more detail, thereby revealing new ways to combat them.

1.2. Post-translational modifications – a leverage point for pathogens

The diversity of the human proteome greatly exceeds the number of coding genes in the human genome.⁸ This fact cannot be explained only by alternative splicing or by the transcription machinery removing non-coding introns during RNA splicing.^{8,9} One of the main drivers for this complexity in protein formation are PTMs.⁸ Starting with a small set of possible PTMs back at the time of our last universal common ancestor (LUCA),¹⁰ the number of known PTMs today reaches more than 300.⁶ These modifications allow cells to directly respond to changes in their surrounding environment.¹⁰ They range from covalent additions of small chemical groups (e.g. phosphorylation or acetylation) to bigger and more complex molecules or proteins (e.g. lipidation or ubiquitylation). In addition, they also include proteolysis, which affects the maturation, location or degradation of proteins (Figure 1).^{6,8} PTMs have been observed in prokaryotic as well as eukaryotic cells⁶ where they activate or inhibit the enzymatic activity of key proteins. Alternatively, they are also able to influence protein interactions as well as the location of the protein.^{7,10}

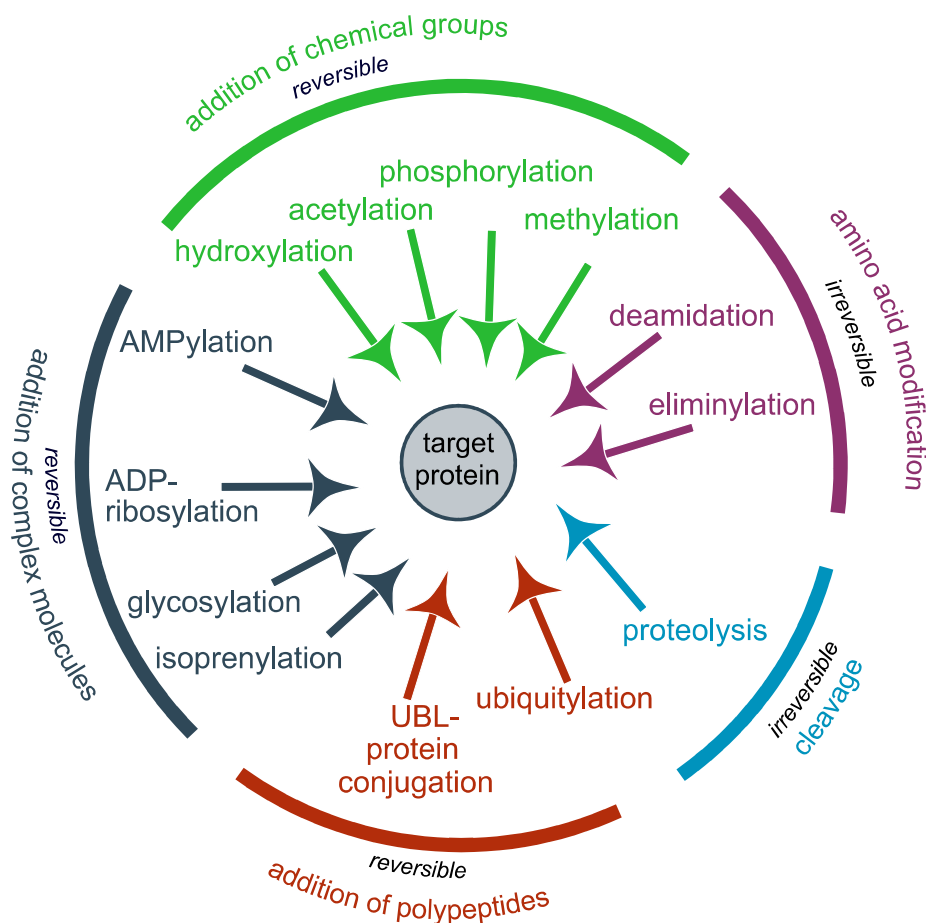


Figure 1: Summary of important post-translational modifications, divided into five different classes. They range from small changes like hydroxylation to more severe changes like ubiquitylation.⁶ Adapted from Ribet *et al.*⁶

Since post-translational modifications represent key players in cell signaling and cellular physiology, it is not surprising that pathogens apply these tools to promote their infection process and ensure their survival.^{6,11} Reported in 1969 for the first time, Collier *et al.* observed diphtheria toxin, produced by the pathogen *Corynebacterium diphtheriae*, to transfer an adenosine diphosphate ribose moiety (ADP-ribosylate) to the mammalian Elongation Factor-2 (EF-2),^{7,11,12} preventing protein synthesis and thus resulting in cell death.^{7,11} Since then, many more bacterial strains have been identified using PTMs to circumvent the host's immune response or to promote their own replication.^{7,11} To do so, they either directly modify host proteins or they influence host signaling pathways by modifying crucial key players and thus impairing protein-protein interactions. Furthermore, they can inhibit or activate the host's own PTM machinery or they insert bacterial factors, which then are unknowingly modified by the host.⁶

The question of which host proteins are subject to bacterially induced post-translational modifications is difficult to predict and must be answered experimentally on a case-by-case basis. However, many bacterial effectors and toxins share common targets consisting of both proteins and signaling pathways which again simplifies this prediction.¹³ Besides the NF- κ B pathway or the mitogen activated protein kinase (MAPK) pathway, both with essential roles in the host immune response, Rho GTPases are preferentially targeted by bacterial PTMs.^{6,7} This class of proteins represents molecular switches, being inactive when bound to guanosine diphosphate (GDP) and active when bound to guanosine triphosphate (GTP).¹⁴ By influencing this activity cycle, bacterial pathogens get access to cellular signaling and metabolism pathways.⁸ Furthermore, Rho GTPases regulate the actin cytoskeleton and are involved in the innate and adaptive immune response.^{6,14} Bacterial pathogens have developed a variety of different ways to interfere with Rho GTPase signaling.^{6,7,14} Some examples are given below.

The *Yersinia enterocolitica* outer protein T (YopT) possesses proteolytic activity toward Rho GTPases, cleaving off their isoprenyl anchor. Consequently, the proteins are no longer membrane bound and released into the cytosol. This leads to cytoskeletal disruption due to impaired cellular function.^{15,16} Toxins of the human pathogens *Clostridioides difficile* (TcdA and TcdB)^{17,18} and *Clostridium sordellii* (tcsL and tcsH)^{19,20} glycosylate conserved threonine (Thr) residues of Rho GTPases using uridine diphosphate (UDP) glucose as a donor molecule. This results in several consequences, including the prevention of the interaction of Rho proteins with their effectors.²¹ Besides inhibiting protein signaling, bacterial pathogens also explore the activation of Rho GTPases. Cytotoxic necrotizing factors (CNFs) from *Escherichia coli*

permanently activate Rho proteins by deamidation of conserved glutamine residues, thus affecting the cytoskeletal integrity.^{22,23} CNF-like proteins from *Yersinia pseudotuberculosis* (CNFy) or from *Vibrio parahaemolyticus* and *Vibrio cholerae* (VopC) operate in a similar way.^{24,25} As a further consequence of Rho protein deamidation by *E. coli* CNFs, the host polyubiquitinates these Rho GTPases, subjecting them to proteasomal degradation.²⁶ Additionally, ADP-ribosylation is also used by pathogens to impair Rho GTPase function. This strategy is employed, for example, by the *Clostridium botulinum* C3 exoenzyme or the *Staphylococcus aureus* epidermal cell differentiation inhibitor (EDIN) toxins.^{27,28} AMPylation, which is closely related to ADP-ribosylation, was observed to be used by pathogens to affect GTPase signaling as another PTM.^{29,30}

1.3. AMPylation – the revival of a post-translational modification

In the 1960s, *Earl Stadtman and coworkers* discovered AMPylation for the first time.³¹ They reported this PTM as an aspect of controlling glutamine synthetase activity.³² Post-translational modifications often use ribonucleotides, as they are highly abundant with cellular ATP concentrations up to 10 millimolar (mM).³³ Furthermore, they offer four bonds with high energy as well as electrophilic centers. In the case of ATP, this allows a nucleophilic attack at the γ -phosphate resulting in protein phosphorylation or at the β -phosphate resulting in pyrophosphorylation.³⁴ As a third possibility, ATP donates an adenosine 5'-monophosphate (AMP) moiety to a protein's serine (Ser), tyrosine (Tyr) or threonine (Thr) side chain (Figure 2).³⁵ After this exergonic process, the released pyrophosphate is enzymatically hydrolyzed, shifting the equilibrium toward the product side.³⁴

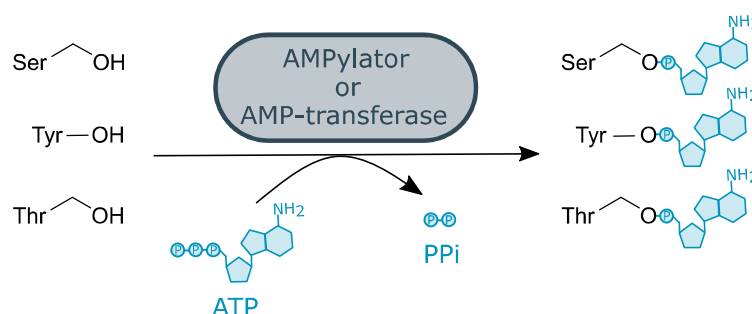


Figure 2: Protein AMPylation on a protein's serine, tyrosine or threonine side chain with ATP as precursor and the subsequent release of pyrophosphate.

After its initial discovery, AMPylation was infrequently reported on in the literature until 2009, when the two pathogens *Vibrio parahaemolyticus* and *Histophilus somni* were found to AMPylate host proteins during their infection process.^{29,30} With this rediscovery of AMPylation as molecular mechanism, special attention was also paid to the drivers of AMPylation, the so-called AMPylators. These proteins are classified into two different groups, depending on their structural properties and their enzymatic mechanism.³⁴ The majority of AMPylators contains either an ATase (adenylyl transferase) domain or a Fic-domain (filamentation induced by cyclic AMP).^{31,36} However, just recently Sreelatha *et al.* observed the pseudokinase and protein adenylyl transferase (SelO) to AMPylate proteins responsible for redox homeostasis.³⁷ However, it has not yet been conclusively determined whether the physiological function of SelO is AMPylation or rather UMPylation (transfer of uridine monophosphate).³⁸

1.3.1. ATase domains

Several proteins have been assigned to the family of ATase domain-containing proteins.³⁴ This protein family belongs to the group of nucleotidyl transferases and is characterized by a unique three-stranded β -sheet.^{34,35} Additionally, all of them share the conserved motif G-X₁₁-D-X-D where the aspartate residues are essential for magnesium ion coordination.³⁹ Although there are various representatives of this protein family, AMPylation activity has been observed in only two of them.^{34,35} One of them, as previously mentioned, is the *E. coli* glutamine synthetase ATase (GS-ATase).^{31,32} The second ATase domain-containing protein with AMPylation activity is an effector protein of *Legionella pneumophila*, a human pathogen causing Legionnaires' disease. The *N*-terminal region of the multifunctional virulence effector protein DrrA, also known as SidM, shares high resemblances with the *C*-terminal region of the GS-ATase.³⁹ Thus, it was not surprising to identify DrrA as AMPylator. This effector protein AMPylates the small GTPase Rab1b, keeping it in its active state, while conversely Rab1b stimulates the AMPylator by allosteric binding.^{40,41} However, DrrA also accepts cytidine triphosphate (CTP) and GTP as nucleophile precursors instead of ATP, though with lower catalytic activities.³⁹

As this twists the ATP substrate, both steric and electrostatic consequences prevent the Fic domain from attacking the α -phosphate.^{45,48} Depending on the nature of the inhibitory α -helix, most Fic proteins can be assigned to three distinct classes.⁴² Nevertheless, some Fic proteins, mainly bacterial pathogenicity factors, are structurally diverse and cannot be assigned to any of the three classes.³⁵ They are lacking regulatory mechanisms, probably due to missing intrinsic targets.⁴²

Class I Fic domains

These Fic proteins are characterized by the fact, that the regulatory α -helix is not positioned on the protein itself, but on a separate protein.⁴⁵ Based on these findings, the class I Fic proteins, which are invariably found in bacteria, are referred to as toxins or FicT and their counterparts, the inhibitory proteins, as antitoxins, FicA.^{42,49} FicTA family members as for example VbhTA of *Bartonella schoenbuchensis* are involved in the AMPylation of DNA gyrase GyrB and topoisomerase IV ParE.⁵⁰ As a result, class I Fic toxins significantly disrupt bacterial DNA topology and influence bacterial reproduction.^{50,51}

Class II Fic domains

Class II Fic proteins account for roughly 80% of all Fic proteins.⁴² The inhibitory α -helix is located *N*-terminal to the Fic domain in their case.⁴⁵ Proteins of this class are either single proteins or part of multidomain proteins.⁴⁵ Since they also contain various other domains, class II proteins can be assigned to a variety of subclasses.³⁵ Thus, the members of this group have very different protein properties, which complicates the identification of possible substrates or biological functions.³⁵ Nevertheless, this group includes several characterized proteins, including the human orthologue FICD.³⁰

Class III Fic domains

The last class of Fic proteins, class III, is characterized by the intramolecular inhibitory α -helix positioned at the *C*-terminus of the protein.⁴⁵ This class consists of bacterial single-domain proteins.³⁵ Therefore, it is assumed that these proteins are regulated by autoregulatory mechanisms, e.g., auto-AMPylation.⁵²

Other Fic proteins

As mentioned above, several Fic domain-containing proteins cannot be assigned to any of the three classes.³⁵ These Fic proteins lack an inhibitory α -helix, since their targets are not located inside the own cell.⁴² Examples for this protein class are the effector proteins VopS from *V. parahaemolyticus* and IbpA from *H. somni*.^{29,30}

1.4. AMPylators in metazoans

Most Fic domain-containing proteins are present in bacteria. However, there are also representatives in metazoans.⁴² Although mammalian AMPylators are not of great importance in this work, they provide the inevitable background for experiments in human cells. To date, only one orthologue of a Fic domain-containing protein was identified and characterized in several eukaryotic species.³⁵ This conserved protein has a class II Fic-domain.⁴⁵ The human representative is called FICD or HYPE.³⁰ Initial studies showed that the protein is located at the endoplasmic reticulum (ER) membrane, where it AMPylates as well as deAMPylates the ER chaperone HSPA5.^{53–55} In this way, it allows cellular response to the amount of unfolded proteins by influencing the unfolded protein response (UPR).⁵⁵ Additionally, alpha-synuclein (SNCA), a protein which is connected to Parkinson's disease, was identified as a novel FICD target *in vitro*.⁵⁶ A recent study also uncovers the involvement of FICD in human neurogenesis.⁵⁷ In general, several FICD targets have been observed. For most of them, the cellular consequences have not yet been fully elucidated, but are the target of ongoing research.^{58–60}

The human FICD orthologue in fruit flies (*Drosophila melanogaster*), called Fic, was identified as an essential protein for fly vision. Individuals bearing a Fic knockout mutant were blind, but viable and fertile. Thus, Fic is assumed to play an important role during the distribution of the visual neurotransmitter histamine.⁶¹ The ER chaperone HSPA5, along with other heat shock proteins, also proved to be a target of Fic in *D. melanogaster*.^{62,63}

1.5. Bacterial Fic domain-containing proteins act intra- and extracellular

As previously mentioned, there are more than 72 000 fic-domain containing proteins listed in the InterPro protein families database (IPR003812, 08.01.2021), the majority of them in bacteria.⁴³ These bacterial proteins can be further divided into two subclasses: AMPylators

necessary for maintaining intrinsic homeostasis and AMPylators promoting bacterial virulence.³⁵

Fic domain-containing proteins ensure bacterial homeostasis

This group of bacterial AMPylators includes the first ever discovered AMPylator, the GS-ATase (GlnE).^{31,32,35} This protein, which has both AMPylation and de-AMPylation functionality, AMPylates the protein glutamine synthetase, thereby regulating it.^{35,64} In its AMPylated version, which mainly occurs during high nitrogen levels, glutamine synthetase is inactive.⁶⁴ As a result, it does not catalyze the formation of L-glutamine from glutamate and ammonia.³⁵ However, when nitrogen levels decrease, the GS-ATase activates the glutamine synthetase again by deAMPylation it.³⁵ This mechanism is subject to a second, regulatory circuit involving the signal-transducing protein trimer PII. Dependent on the UMPylation state of PII, it either increases the AMPylation or deAMPylation activity of GlnE.³⁵

Another class of proteins ensuring cell homeostasis are the previously mentioned class I Fic toxins. Not only the bacterium *Bartonella schoenbuchensis* and other Gram-negative bacteria AMPylate the DNA gyrase GyrB and topoisomerase IV ParE,⁵⁰ but also *Yersinia pseudotuberculosis* and the Gram-positive bacterium *Staphylococcus aureus* AMPylate their corresponding counterparts.⁶⁵ To date, FicT AMPylation activity is thought to trigger the bacterial SOS response through cell filamentation and also to affect chromosome regulation and cell separation.^{50,65}

Fic-domain containing proteins drive bacterial virulence

Compared to intrinsic bacterial AMPylation, more is known about AMPylators being active in a host-pathogen setting. Several of these proteins were characterized in more detail,^{35,42,45,65,66} and three of them are detailed below.

Immunoglobulin-binding protein A (IbpA) from *Histophilus somni*

In addition to VopS of *V. parahaemolyticus*, IbpA of *H. somni* was discovered as one of the first AMPylators to modify host proteins.^{29,30} *H. somni* is a Gram-negative bacterium, which causes respiratory disease in livestock.³⁰ Its effector protein IbpA, which is translocated into the host cell by a two-partner secretion system, contains *N*-terminal adhesion domains as well as

two C-terminal Fic domains.³⁰ These Fic domains account for the bacterial cytotoxicity by AMPylating Tyr residues of the host Rho GTPases RAC1, RHOA and CDC42 in their active and their inactive states.⁶⁷ In this way, the downstream signaling of Rho GTPases is inhibited.^{30,67} Recently, an innovative *in vitro* approach suggests that Fe²⁺ homeostasis as well as endocytic receptors may also play an important role during *H. somni* infection, since it identifies various membrane-bound receptors as putative IbpA targets.⁵⁸

Beps – *Bartonella* effector proteins

Members of the *Bartonella* family, which are usually transmitted by blood-sucking arthropods, are characterized by a common infection strategy.⁶⁸ By applying a type IV secretion system, they translocate an armada of effector proteins (BepS) into the host cytosol.⁶⁸ Most of these effectors are equipped with a Fic domain.⁶⁹ BepA from *Bartonella henselae* targets, among others, two yet unidentified host proteins in HeLa cell lysate as well as the breast cancer anti-estrogen resistance protein 1 (BCAR1) with unknown consequences.^{58,70} Another *Bartonella* effector protein with AMPylation activity is Bep2 from *Bartonella rochalimae*. This protein was found to target the intermediate filamenting protein vimentin in the lysate of J774 mouse macrophages, again with unknown consequences.^{68,71} BepC as a third example also contains the conserved Fic motif, with the exception of one amino acid. The acidic residue (either Asp or Glu) is replaced by a Lys residue.⁷² The activity of this protein is therefore not AMPylation. Its Fic domain rather attracts the Rho guanine nucleotide exchange factor GEF-H1 while being located at the plasma membrane. This leads to RHOA/ROCK pathway activation and stress fiber formation.⁷²

Phosphocholine transferase (AnkX) from *Legionella pneumophila*

Similar to *Bartonella* members, *L. pneumophila* invades host cells by translocating effector proteins into the host cytosol *via* a type IV secretion system.⁷³ These proteins construct the *Legionella* containing vacuole (LCV), a membrane compartment for bacterial replication.⁴² Besides the already mentioned ATase domain-containing AMPylator of *L. pneumophila* DrrA,⁴² the bacterium also contains a protein bearing a Fic domain called AnkX. This effector protein phosphocholinates the two GTPases RAB1 and RAB35, probably inactivating them by disrupting their signaling processes.⁷⁴ However, the consequences of the phosphocholination of Rab family members are distinct from the effects of AMPylation by DrrA. In this way,

L. pneumophila is able to fine-tune the activity of RAB1 function by controlling its interaction partners.^{74,75}

1.6. *V. parahaemolyticus* VopS – an AMPylator as tool to promote infection

Protein adenylyl transferase VopS is an effector protein of the Gram-negative bacterium *V. parahaemolyticus*.⁷⁶ The majority of strains of this marine organism are virulent causing acute gastroenteritis when eating uncooked seafood.⁷⁷ Several of them have been sequenced, including strain RimD 2210633, which shows the most common serotype O3:K6.⁷⁶ *V. parahaemolyticus* was suspected to act exclusively as extracellular bacterium during its infection process. Still, further studies unveiled, that the bacterium also invades epithelial cells, where it replicates prior to escaping into the cytoplasm.⁷⁸ Nevertheless, its main infection approach is to attach to the host cell *via* multivalent adhesion molecules (MAMs).⁷⁶ Once in close proximity to the host cell, *V. parahaemolyticus* releases its toxins (thermostable direct hemolysin *tdh* and TDH-related hemolysin *trh*) and uses its two type III secretion systems (T3SSs) to translocate effector proteins into the host's cytosol (Figure 4).⁷⁷ Not depicted in this figure are the *V. parahaemolyticus* two type VI secretion systems (T6SSs), whose functions are only partially known.⁷⁹

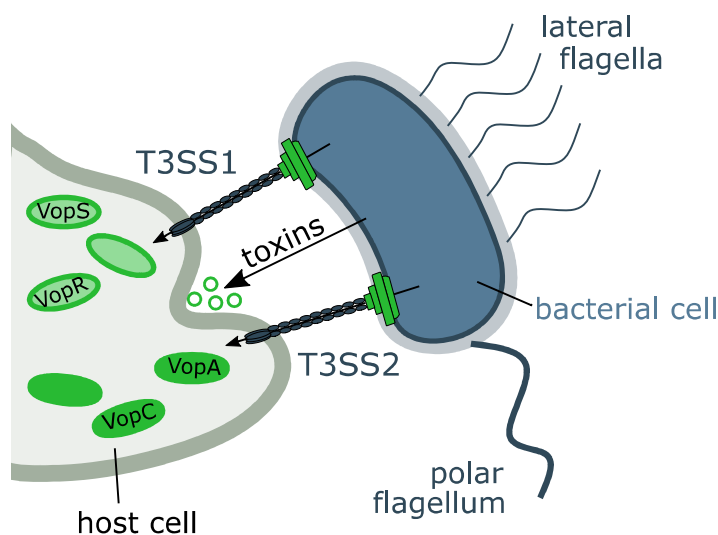


Figure 4: Overview of selected virulence strategies of *V. parahaemolyticus*. The Gram-negative bacterium possesses two type III secretion systems with T3SS1 translocating VopS. Furthermore, *V. parahaemolyticus* also releases different toxins thus further harming the host cell. Adapted from Rauh *et al.*⁸⁰

Type III secretion system 2 (T3SS2) is closely associated with *V. parahaemolyticus* virulence. Strains without this system have no pandemic potential. The system injects the effector proteins VopA/P, VopL, VopT as well as VopC into the host's cytosol.⁷⁶ As a short reminder, VopC activates Rho GTPases by deamidating them, thus influencing actin cytoskeletal integrity.⁷⁷ In contrast to T3SS2, T3SS1 is expressed in all *V. parahaemolyticus* strains.⁷⁷ This secretion system translocates the effector proteins VopQ, VopR, VPA0450 and VopS inside the host cells.⁷⁶ Initially, VopS (VP1686) was suspected to initiate apoptosis in macrophages by preventing the activation of the NF- κ B pathway without the involvement of toll-like receptors.⁸¹ Later on, the effector protein was identified as a Fic domain-containing protein without an inhibitory α -helix.^{29,45} VopS AMPylates a conserved Thr in the switch I region of the small Rho GTPases RAC1, RHOA and CDC42 with a catalytic constant k_{cat} of 18 s^{-1} .^{29,82} Besides these three confirmed targets, several more putative targets have been identified.^{83,84} Among them are other small Rho GTPases⁸⁴ and proteins like the ATP-dependent 6-phosphofructokinase (PFKP),⁸³ the Nucleoside diphosphate kinase A (NME1)⁸³ or the Endoplasmic reticulum Golgi intermediate compartment protein 2 (ERGIC2).⁸⁴ Additionally, *in vitro* nucleotide monophosphate (NMP) modification experiments also revealed GTP and CTP as putative nucleotide sources for VopS.⁶⁷

VopS mediated AMPylation of RHOA, RAC1 and CDC42 has far-reaching consequences.^{85–87} Overall, the modification inhibits downstream signaling of the Rho GTPases while also targeting the actin cytoskeleton.⁸⁷ First of all, VopS decreases cytokine production and inhibits caspase-1 activation by a hitherto unknown mechanism.^{85,87} On the other hand, its modification of the switch I region in Rho GTPases leads to inflammasome activation by the pattern recognition receptor pyrin.⁸⁶ More effects of VopS AMPylation were studied in great detail by Woolery *et al.*⁸⁷ Like many other bacterial pathogens targeting Rho GTPases, VopS also negatively impacts the NF- κ B and the MAPK signaling pathways due to their essential roles in the host innate immune response.^{7,87} Moreover, it also inhibits GTPase binding to the 1-activated kinase (PAK) or influences the interaction of RAC1 with an E3 ligase (Figure 5).⁸⁷

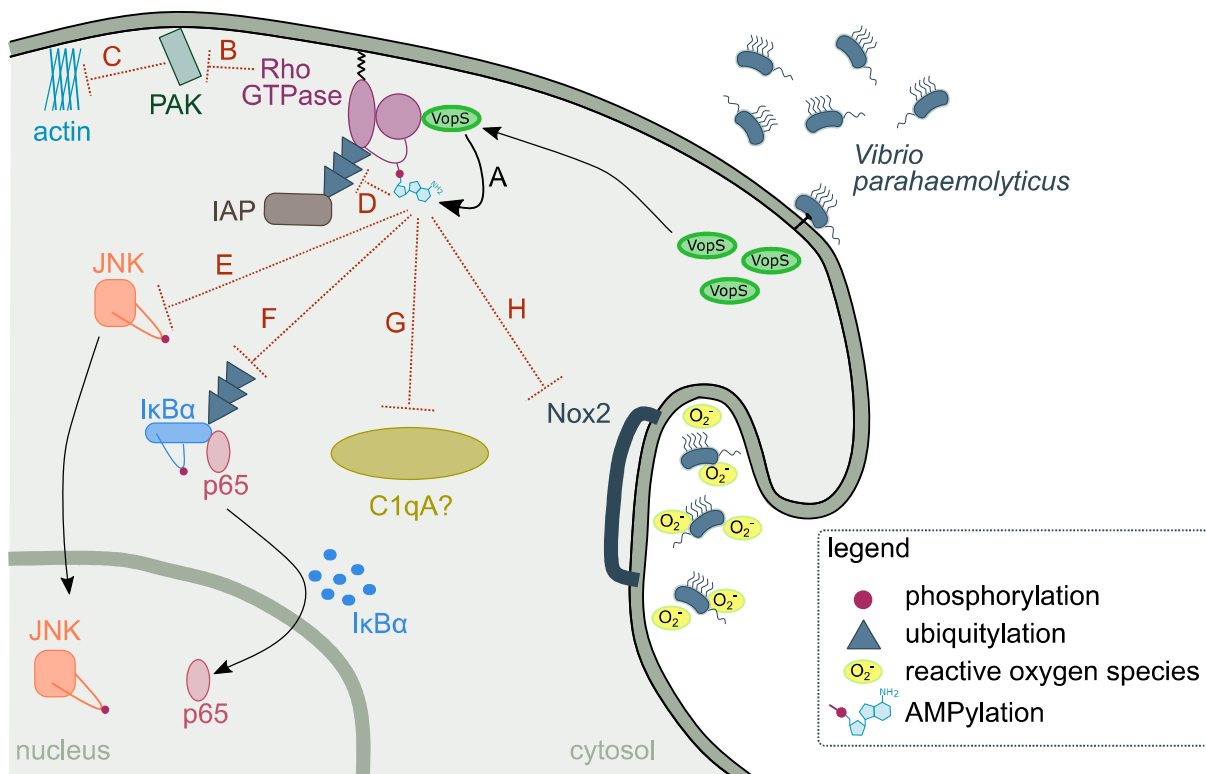


Figure 5: Cellular consequences of VopS mediated AMPylation. (A) VopS AMPylates Rho GTPases on the switch-1 loop, thus inhibiting a myriad of cellular processes. (B, C) First, it prevents GTPase binding to PAK and thereby influences cytoskeletal integrity. (D) It further inhibits the interaction between RAC1 and the E3 ligase (E, F) and decreases phosphorylation levels on the mitogen-activated protein kinases JNK and Erk and the NFκB inhibitor IκBα entailing more consequences. (G) Moreover, it also inhibits RAC1 binding to the complement C1q subcomponent subunit A C1QA. (H) Lastly, the AMPylated Rho GTPases influence the Nox2 complex and thereby the levels of reactive oxygen species.⁸⁷ Adapted from Woolery *et al.*⁸⁷

1.7. A range of tools to detect AMPylated proteins

Since the discovery that pathogens use AMPylation to promote infection, several methods have been pursued to decipher putative targets. Among the first experiments were *in vitro* AMPylation assays^{39,67} or the labeling of cell lysates with radiolabeled ATP. For the latter, the radiolabel (³²P) has to be positioned at the α-phosphate.^{29,70} However, a proteome-wide screening for AMPylation targets proved to be challenging with these approaches.³⁵ A different strategy uses antibodies to detect AMPylated amino acid residues, more precisely antibodies toward AMPylated-threonine or AMPylated-tyrosine.^{88,89} As antibodies enable the accurate detection as well as specific enrichment of modified proteins, they have the potential to detect yet unknown AMPylation targets.⁹⁰ However, the two originally available antibodies lag behind expectations since they either need high protein concentrations or show certain cross-

reactivities. Three recently developed monoclonal anti-AMP antibodies promise to overcome this limitation.⁹¹ They are able to recognize AMPylated residues independent of the protein backbone under native as well as denatured conditions.⁹¹

Alternative approaches equip AMP with various chemical handles, thus establishing a broad spectrum of probes. One of the first molecules to study AMPylation was propargylated ATP (**N6pATP**, Figure 6 A), since the attachment of a small, chemical group at the N⁶ position of adenine did not hinder its incorporation.⁹² Combined with click-chemistry and further analysis *via* mass spectrometry (MS), this probe allowed the enrichment of modified proteins in their native form from cell lysates.⁹³ However, this probe is not suitable to investigate AMPylation under physiological conditions in whole cells. The highly charged phosphate groups prevent cellular uptake as the molecule is too polar to cross the nonpolar lipid bilayers.⁵⁷ Also ATP derivatives equipped with an alkyne tag at position 2 of the adenine ring (**2eATP**, Figure 6 B) or equipped with a fluorophore (**Fl-ATP**, Figure 6 C) were used to study AMPylation.^{83,94} As these two probes have similar features as **N6pATP**, they share the same advantages and disadvantages.⁹³ Moreover, MS analysis is an aspect that can be improved to facilitate the identification of AMPylation targets. Li *et al.* used the mass increase upon AMPylation of 329 Da as well as its characteristic fragmentation patterns upon collision induced dissociation (CID).⁹⁵ However, the disadvantages of this method are the necessity to search a detailed sequence database followed by a manual reconciliation of all results.³⁵ Further method development showed, that CID as well as higher energy collision dissociation (HCD) fragmentation of AMPylated peptides yielded in losses of different chemical groups, whereas electron transfer dissociation (ETD) fragmentation left the AMP moiety unimpaired.⁹⁶ Thus, spectral interpretation was facilitated, but still required critical reconciliation.⁹⁶ Another approach implemented stable isotope-labeled ATP to detect AMPylated peptides in cell lysates.⁷¹ By introducing and searching for an AMP-reporter cluster, Pielek *et al.* identified vimentin as an AMPylation target of *B. rochalimae*'s effector protein Bep2.⁷¹ However, this method still does not enable the investigation of AMPylation under physiological conditions. Furthermore, handling radioactive probes poses a certain cytotoxicity risk.⁹³

A completely different approach was taken by Yu *et al.* They screened the human proteome for AMPylation targets of IbpA and VopS using an approach called nucleic acid programmable protein array (NAPPA).⁸⁴ In short, they attached purified human plasmid complementary deoxyribonucleic acid (cDNA) onto the array surface, where it was subsequently transcribed and translated yielding the recombinant protein of interest, equipped with a suitable tag.⁹³ After

removal of remaining DNA, proteins were incubated with $\text{GTP}\gamma\text{S}$, **N6pATP** and AMPylator. For the final step, the arrays were washed, proteins were labeled with a fluorescent dye and the readout was performed with a microarray scanner.⁹³ Using this procedure, several unknown putative targets of VopS and IbpA were identified.⁸⁴ This methodology provides a valuable tool for further studies. As a disadvantage, however, it should be noted that recombinant proteins may be misfolded or lack post-translational modifications. These are present under physiological conditions and can also influence the experimental result.⁹³

The latest approach by Gulen *et al.* used co-substrate-mediated covalent capture, a method that does not compete with endogenous ATP resources.⁵⁸ They applied binary probes composed of recombinant Fic proteins containing a strategically introduced cysteine (Cys) residue and nucleotide derivatives with thiol reactivity (thiol-reactive nucleotide derivatives, TReNDs).⁵⁸ This way, they detected further targets of the AMPylators IbpA, BepA and FICD in HEK293T cell lysate.⁵⁸ However, as with previous methods, this approach does not reflect physiological conditions. Thus, caution is advised when regarding the obtained results.

A different kind of approach was introduced by the Sieber lab. Kielkowski *et al.* introduced two cell-permeable pronucleotide probes (Figure 6 D, E).^{57,97}

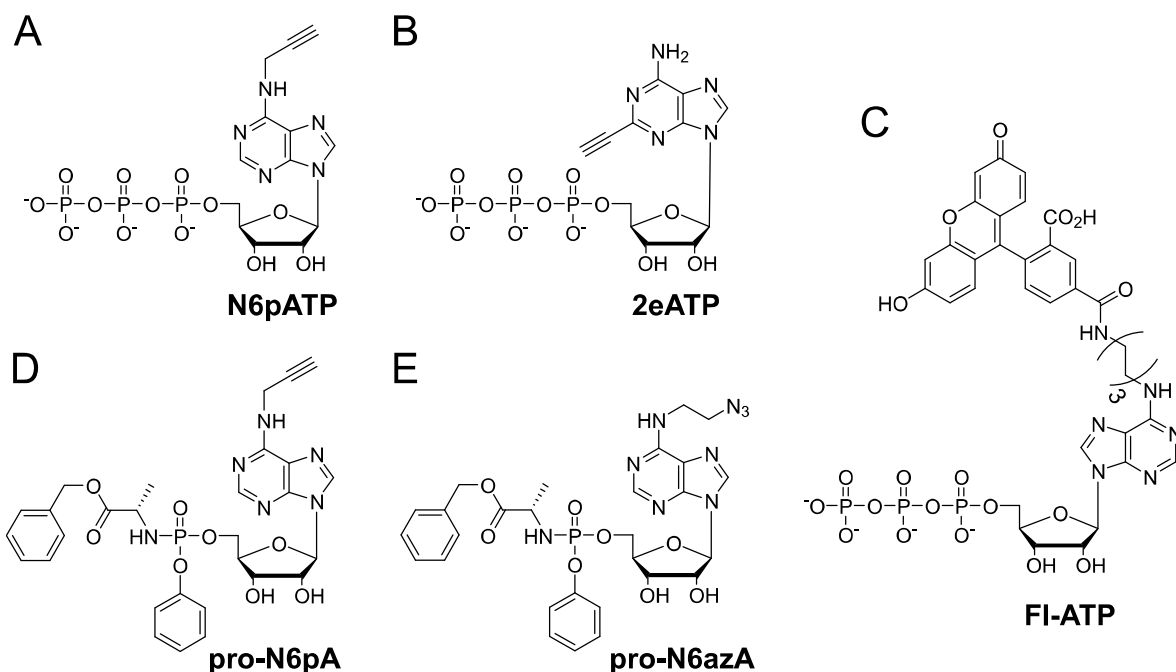


Figure 6: Different chemical probes applied during *in vitro* and *in situ* experiments to identify AMPylated proteins. The charged probes **N6pATP**, **2eATP** and **FI-ATP** (A-C) were used during *in vitro* experiments, while **pro-N6pA** and **pro-N6azA** (D, E) are applicable for *in situ* experiments.

N^6 -propargyl adenosine phosphoramidate pronucleotide (**pro-N6pA**) allowed probing of AMPylated proteins during neurogenesis⁵⁷ and N^6 -(2-azidoethyl)adenosine phosphoramidate (**pro-N6azA**) enabled live-cell imaging in different cell lines.⁹⁷ Both probes use the ProTide prodrug technology, which improves cellular uptake by masking charged groups.⁹⁸ Furthermore, structural considerations concerning kinase recognition are redundant since this first phosphorylation step is skipped (Figure 7).⁵⁷ These probes thus present a valuable tool for future studies to decipher novel AMPylation targets and elucidate implications of this post-translational modification.^{57,97}

1.8. MS-based proteomic profiling

As mentioned earlier, proteome diversity is partly caused by post-translational modifications.⁸ When examining these modifications, methods like genomics and transcriptomics reach their limit.^{99,100} Also protein-protein interactions or dynamic processes cannot be tracked this way.⁹⁹ As a result, proteomics in combination with MS analysis emerged as next generation technology. This technique, however, is highly dependent on previous genomic achievements, as the basis for the analysis of proteomic data are sequenced proteomes.¹⁰¹ Proteomic techniques, as they are applied nowadays, are strongly dependent on MS instrument equipment, resolution and performance as well as on MS techniques.⁹⁹

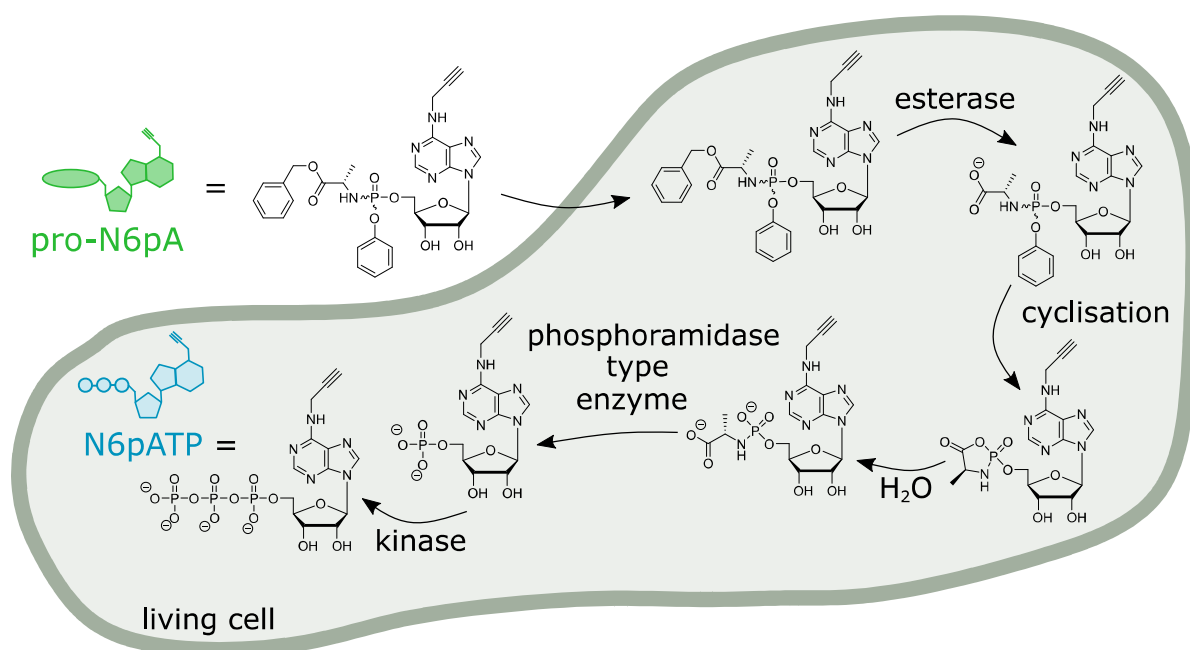


Figure 7: Cellular uptake of **pro-N6pA** followed by probe activation. Consecutive transformation steps remove the masking phosphoramidate moiety resulting in a triple phosphorylated ATP probe.^{57,98}

The predominantly used proteomic strategy is called bottom-up proteomics or shotgun proteomics.^{102,103} An example of this type of strategy is known as Activity-Based Protein Profiling (ABPP), a technique established by the Cravatt lab^{104–106} as well as the Boygo lab.^{107,108} A typical workflow for this methodology starts optionally with probe treatment of tissue, cells or cell lysates. Afterward, if possible, proteins are enriched or separated to overcome their sometimes low expression levels as well as lower ionization efficiency and sub-stoichiometry.⁹⁹ A widely used example for such an enrichment strategy is the reaction of labeled proteins with biotin followed by protein pull-out with bead-bound avidin.¹⁰⁹ Once enriched, the protease trypsin is applied to digest proteins resulting in proteolytic peptides, preferably yielding sequences comprised of 5 to 20 amino acids.¹⁰⁰ These peptides are further processed before being subjected to liquid chromatography (LC)-MS/MS based analysis (Figure 8).⁹⁹

However, proteomics as we know it nowadays would hardly be possible without the discovery of the so-called click chemistry. Thereby, an azide moiety reacts with a terminal alkyne resulting in a 1,2,3-triazole.⁹⁹ The initially heat catalyzed reaction was discovered by Huisgen,¹¹⁰ and further developed by Sharpless¹¹¹ and Meldal¹¹² into a Cu(I) catalyzed method which can take place at room temperature (copper(I)-catalyzed azide-alkyne cycloaddition CuAAC). Hereafter, reaction conditions were further fine-tuned for usage in complex samples such as protein lysates or tissues.¹⁰⁶

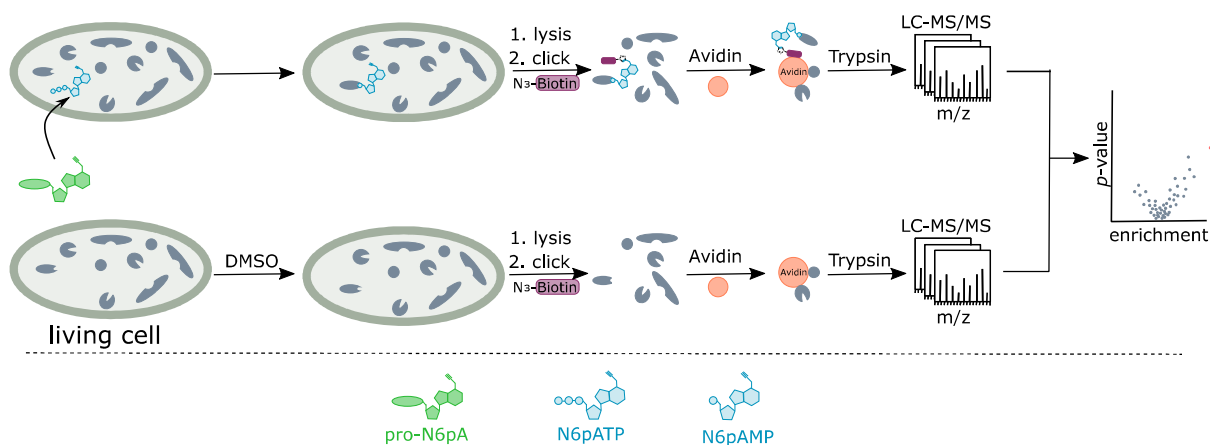


Figure 8: Schematic overview of a standard proteomic profiling workflow using an alkyne-tagged probe and a LC-MS/MS based readout. Adapted from Rauh *et al.*⁸⁰

Besides experimental setup, quantitative analysis of MS data also largely contributes to the success of a proteomic experiment. Quantification is either performed by label-free quantification (LFQ) strategies or by stable isotope labeling.¹¹³ LFQ represents a low cost technique, which can be applied to all kinds of biological material, including clinical isolates. In contrast to various labeling techniques, it allows for broader proteome coverage, even though this is associated with lower precision.¹¹³ Quantification during LFQ measurements is achieved by either comparing peptide intensities or by connecting the spectral count for a protein/peptide to its quantity in the sample.¹¹⁴ In order to compare results, it is important to ensure that protein expression levels are the same in all samples and to perform LFQ measurements sequentially using the exact same LC-MS/MS system.¹¹³ This method allows the comparison of any number of conditions with each other.¹¹⁵ Software such as MaxQuant^{115,116} greatly simplified LFQ data evaluation and contributed to the success of this method.

Isotope labeling-based approaches harness the property, that an isotopically labeled peptide shows identical behavior when subjected to LC-MS/MS analysis but has a label-dependent mass shift when comparing differently labeled samples.^{113,114} Isotopic labels can be introduced in different ways. Stable isotope labeling of amino acids in cell culture (SILAC) is a method to metabolically label different cell lines.¹¹⁷ Furthermore, labels can also be introduced chemically to already digested proteins. Examples thereof are isotope-coded affinity tags (ICAT),¹¹⁸ the triplex labeling method dimethyl labeling (DML),^{119,120} or multiplex labeling methods like isobaric tags for relative and absolute quantification (iTRAQ)¹²¹ or tandem mass tags (TMT).¹²²

During SILAC labeling, cells are grown in media containing only isotopically labeled arginine and lysine.¹¹⁷ Thus, the minimum of one labeled amino acid is incorporated into every tryptic peptide.¹¹⁴ To ensure sufficient labeling efficiency, cells should be passaged more than 5 times in the corresponding media prior to experimental usage.¹¹⁷ This way, up to three differently labeled conditions can be compared [light (unlabeled), medium ($^{13}\text{C}_6$) and heavy ($^{13}\text{C}_6$ $^{15}\text{N}_4$) tagged amino acids].¹¹⁴ The benefit of this technology is the combination of differently labeled samples at an early stage. As a result, fewer samples need to be processed. Furthermore, variations due to a slightly different sample handling are prevented.¹²³

Similar to SILAC labeling, the dimethyl labeling workflow also allows sample combination. However, this only takes place after digestion, thus at a later stage than during SILAC labeling.¹²⁴ This method is fast, and its labeling reagents (formaldehyde and sodium cyanoborohydride) are inexpensive to purchase.^{124,125} Since SILAC is mainly used for cell

culture systems, dimethyl labeling represents a good alternative for peptide quantification when examining tissues or the like.¹²⁴

Since PTMs have the power to radically change protein structure and function, they are important features which need to be explored in more detail.¹⁰⁰ Being already hard to detect in standard biochemistry experiments, they also pose a challenge to shotgun proteomics experiments.⁹⁹ PTMs can occur on a broad range of amino acids.⁸ They are dynamic modifications and their degree of endogenously modified residues varies depending on environmental conditions.^{57,99} Furthermore, as already mentioned previously, not all PTMs represent stable modifications during the MS ionization process.⁹⁶ On the other hand, bulky PTMs or PTMs such as AMPylation or ADP-ribosylation that introduce a negative charge can affect the ionization properties of peptides, making their identification more difficult.⁹⁹ To partially overcome these limitations, PTM selective enrichment strategies should be sought. Therefore, it is advantageous that enzymes catalyzing PTMs accept minor modifications of the structure to be transferred, like small handles necessary for click chemistry.⁹⁹

1.9. Scope of this work

Post-translational modifications are crucial switches for distorting and fine-tuning protein structure and function. Mammalian cells depend on the faultless functioning of this machinery. However, an increasing number of pathogens have been found to exploit PTMs in order to control host cellular signaling pathways and promote their own infection process.¹¹ Recent studies have shown that several pathogens harness AMPylation (mainly of Rho GTPases) to avoid, amongst other things, recognition by the host immune system.^{35,42} One of these pathogens is the Gram-negative bacterium *V. parahaemolyticus*, the leading cause for acute gastroenteritis due to uncooked seafood.^{76,77} As Fic domain-containing proteins are highly conserved throughout all kingdoms of life, they provide a valuable basis for the development of strategies to battle pathogenic bacteria.⁴² Several diverse methods to detect protein targets of pathogenic AMPylators have been explored.^{35,90} So far, most of them are only suitable for *in vitro* experiments due to poor cell permeability. Thus, they do not investigate AMPylation in its natural environment, meaning under physiological conditions. Consequently, false positive results are possible and identified targets need to be carefully validated. Additionally, these methods also have other disadvantages such as limited selectivity or handling of harmful compounds.^{29,88,89}

The scope of this thesis was to establish a proteomic strategy, which allows the *in situ* identification of AMPylated host proteins due to bacterial infection. The well characterized human HeLa cell line and the previously established cell permeable pronucleotide probe **pro-N6pA** served as basis.⁵⁷ Furthermore, *V. parahaemolyticus* was chosen as a characterized model organism.

Once a suitable probe concentration and optimized labeling times were established, proteomic profiling experiments were performed. By comparing dimethyl labeling, SILAC and LFQ quantification-based methods, the best proteomic technique to track AMPylated proteins in an infection setting was identified. In addition, various experimental setups were tested to ensure maximum reliably enriched proteins and to help distinguish between intrinsically AMPylated proteins and bacterial effector-mediated AMPylation. The inactive VopS-H348A mutant was constructed for this purpose in cooperation with Sophie Brameyer and its effects during HeLa cell infection were assessed. To further validate putative VopS targets, *in vitro* AMPylation assays and site-ID experiments using a TEV-cleavable linker were performed. After successful establishment of bacterial infection and subsequent proteomic profiling, other bacterial pathogens were also investigated for their ability to AMPylate host proteins. Therefore, mutants expressing inactive AMPylators were exploited. Apart from this, further studies in other human cell lines were performed to get an overview of other putative VopS AMPylation targets.

2. Results and discussions

2. Results and discussions

Post-translational modifications play an important role as signal regulators at the host-pathogen interface.⁷ So far, most methods investigating PTMs are only applicable during *in vitro* experiments. The drawback to these experiments is, that PTMs can depend on a variety of factors that are difficult to replicate in experiments with *in vitro* settings, such as correct protein folding or preceding PTMs (see section 1.7). Therefore, it is of great importance to establish a system to study post-translational modifications under physiological conditions, which is the goal of this work.

The previously developed probe **pro-N6pA** was chosen as the starting point for establishing such a workflow, in addition to the well-characterized human HeLa cell line and the sequenced model organism *V. parahaemolyticus*.^{57,126} **Pro-N6pA** is converted to **N6pATP** by metabolic transformation *via* a series of different enzymes once it is taken up by the cell (see Figure 7). Upon equipment with a triphosphate, **N6pATP** is further available as donor of an AMP group in the AMPylation reaction. Thus, it enables the incorporation of an alkyne tag into target proteins which can be subsequently selectively enriched (Figure 9).

Pro-N6pA was synthesized as previously published by Kielkowski *et al.*⁵⁷ In addition to probe synthesis, the effect of **pro-N6pA** on cell viability and proliferation was investigated in this study using an MTT (3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2*H*-tetrazolium bromide) assay.⁵⁷ **Pro-N6pA** was found to be non-cytotoxic, with an IC₅₀ value of 300 μ M in HeLa cells after incubation for 24 h.⁵⁷

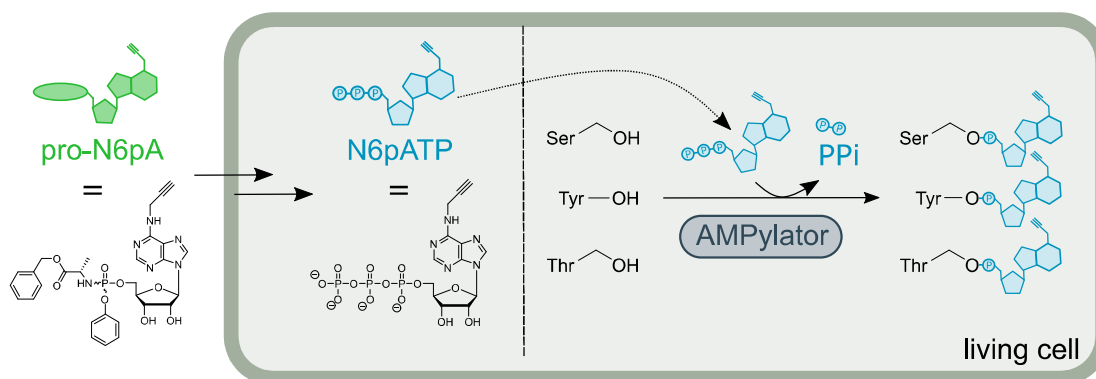


Figure 9: Graphical illustration of **pro-N6pA** uptake followed by its cellular conversion to functional **N6pATP**, which can further be exploited as AMP donor for the AMPylation reaction. A terminal alkyne tag is thus introduced into all AMPylated proteins, allowing modification-specific protein enrichment. Adapted from Rauh *et al.*⁸⁰

However, even with a suitable probe in hand, the detection of post-translationally modified proteins remains a challenge. On the one hand, **pro-N6pA** competes with significant endogenous ATP levels up to 10 mM.³³ On the other hand, cells experience varying levels of endogenous AMPylation, as the modification enables an adaptation to environmental conditions, as seen for the interaction between FICD and HSPA5 (see section 1.4) or between GS-ATase (GlnE) and the protein glutamine synthetase (see section 1.5).⁵⁷ **Pro-N6pA** therefore requires a sufficiently strong labeling efficiency to counteract these circumstances.

2.1. Optimization of **pro-N6pA** labeling conditions

Since the labeling conditions of **pro-N6pA** have not been thoroughly optimized in previous studies, time- and concentration-dependent labeling experiments were performed. Gel-based analysis is the method of choice for this purpose, since SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) allows the investigation of several labeling conditions in parallel and requires only a small amount of time compared to LC-MS/MS-based analysis.

During optimization, HeLa cells were either treated with 100 μ M **pro-N6pA** for different periods of time (Figure 10 A) or incubated with varying **pro-N6pA** concentrations for 16 h (Figure 10 B), as performed in previous studies.⁵⁷ After probe incubation for the designated time period, or with the desired concentrations, HeLa cells were lysed, protein concentrations were adjusted, and lysate containing alkyne-tagged proteins was further clicked to rhodamine-azide (further details given in section 4.4.2). Finally, the gel was visualized with a fluorescence gel scanner revealing several bands that represent protein targets. With respect to time-dependent labeling, increasing incubation times correlated with increasing band intensities. However, no significant changes were visible when comparing incubation times of 8 h to incubation times of 16 h. As anticipated, HeLa cells retained their normal morphological phenotype after treatment with 100 μ M **pro-N6pA** for 16 h, therefore overnight incubation was chosen for all further experiments for convenience. Concentration-dependent labeling yielded increasing band intensities up to 100 μ M **pro-N6pA**. As this concentration yielded intense labeling, probe concentration was set to 100 μ M **pro-N6pA** for all further experiments.

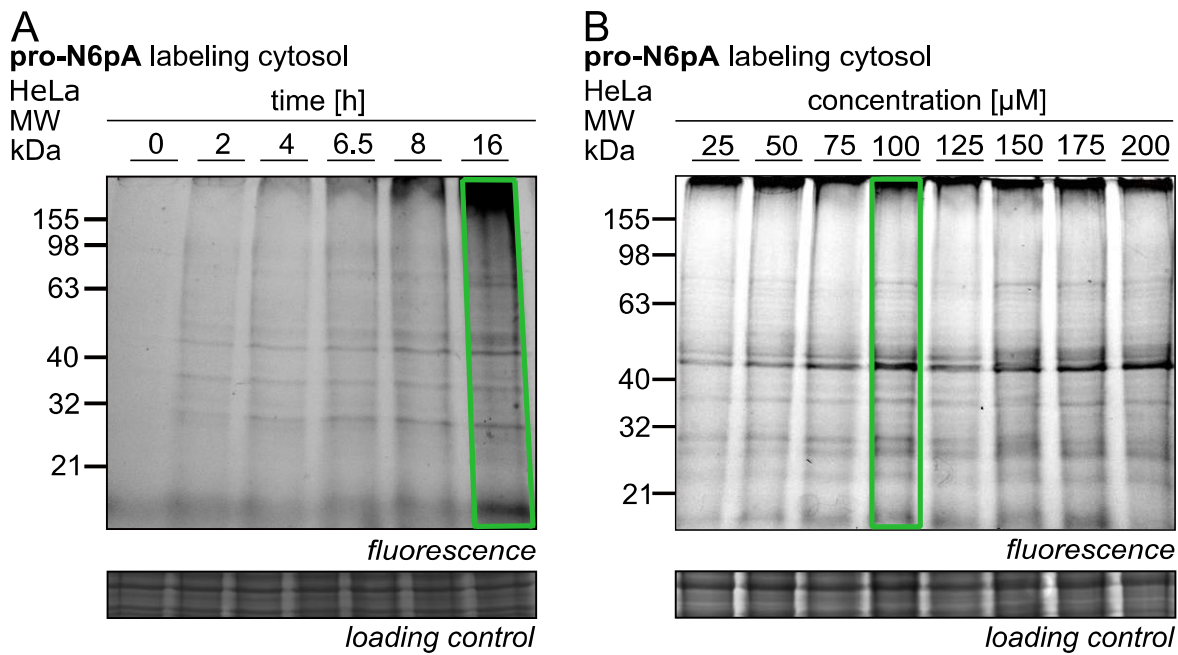


Figure 10: SDS-PAGE analysis of time- and concentration-dependent labeling of HeLa cells. Time-dependent labeling (A) (100 μM **pro-N6pA**) revealed an incubation time of 16 h as the best labeling option while concentration-dependent labeling (B) (16 h) identified a concentration of 100 μM **pro-N6pA** as most suitable. Adapted from Rauh *et al.*⁸⁰

2.2. Intrinsically AMPylated proteins in HeLa cells

With optimized conditions in hand for labeling AMPylated proteins, identification by LFQ MS-based experiments was the next focus. One consideration was, that this experimental workflow can identify intrinsically AMPylated proteins as well as those AMPylated during infection. Therefore, intrinsically AMPylated proteins should be collated in a reference list, which can be compared to those AMPylated after bacterial infection. Proteins on this reference list should not necessarily be excluded as potential targets of effector mediated AMPylation, but they should be considered with caution.

For the proteomic profiling experiment, HeLa cells were incubated with either 100 μM **pro-N6pA** or dimethyl sulfoxide (DMSO, 0.1 vol%) as a vehicle control for 16 h, in 12 replicates per condition. Afterward, cells were lysed and a CuAAC click reaction with biotin-azide was performed. Modified proteins were further enriched using bead-bound avidin, reduced with dithiothreitol (DTT) and alkylated with iodoacetamide (IAA). Subsequent overnight trypsin digest resulted in peptides for LC-MS/MS analysis (further details given in section 4.4.3). Measured raw data were processed using MaxQuant (version 1.6.2.10),¹¹⁶ which identified

measured peptides. A two-sample Student's *t*-test was applied to compare protein intensities in different samples and the results were illustrated using Perseus (version 1.6.2.3).

Thresholds of \log_2 enrichment > 1 and significance $p < 0.05$ were applied to delineate significant protein enrichment. After filtering to remove contaminants as well as proteins identified by reverse sequences or by a solely modified amino acid, 1240 protein groups remained, of which 25 met the threshold criteria (Table 1, Figure 11 A, p. 29).

Table 1: Significantly enriched proteins (\log_2 enrichment > 1 , p -value < 0.05) resulting from a LFQ proteome profiling experiment in HeLa cells (associated with Figure 11 A), which have been treated with 100 μ M **pro-N6pA** or 0.1 vol% DMSO as the vehicle control for 16 h.

gene name	protein name	\log_2 fold change	$-\log_{10}$ p-value	coverage [%]	previously identified?
ACP2	lysosomal acid phosphatase	4.25	17.81	38.9	✓
LAMP1	lysosome-associated membrane glycoprotein 1	3.90	7.95	17.7	✗
CTSA	lysosomal protective protein	3.57	8.62	17.9	✓
TPP1	tripeptidyl-peptidase 1	3.33	7.52	17.4	✓
CTSB	cathepsin B	3.32	10.42	21.2	✓
SCPEP1	retinoid-inducible serine carboxypeptidase	3.24	8.21	27.9	✓
ABHD6	monoacylglycerol lipase ABHD6	3.07	9.62	26.1	✓
PNPLA4	patatin-like phospholipase domain-containing protein 4	2.89	6.98	32.4	✓
SLC38A2	sodium-coupled neutral amino acid transporter 2	2.44	6.08	17	✗
PSAP	prosaposin	2.40	9.43	12.8	✗
PLD3	phospholipase D3	2.40	8.85	14.3	✗
CTSD	cathepsin D	2.24	8.62	27.9	✗
CTSZ	cathepsin Z	2.23	7.73	20.1	✓
PPME1	protein phosphatase methylesterase 1	2.20	5.55	59.1	✓
PPT1	palmitoyl-protein thioesterase 1	2.06	6.81	26.9	✗
APOB	apolipoprotein B-100	2.01	6.30	1.9	✗
PRCP	lysosomal Pro-X carboxypeptidase	1.63	7.04	12.1	✗
FDFT1	squalene synthase	1.60	5.56	36	✓
PLBD2	putative phospholipase B-like 2	1.46	6.00	10.2	✗
SCARB2	lysosome membrane protein 2	1.45	11.71	16	✗

MARCKS	myristoylated alanine-rich C-kinase substrate	1.32	5.67	37.7	×
SQSTM1	sequestosome-1	1.31	12.61	51.1	✓
PFKP	ATP-dependent 6-phosphofructokinase, platelet type	1.11	6.93	49.6	✓
HEXB	beta-hexosaminidase subunit beta	1.02	3.23	40.5	✓
FTH1	ferritin heavy chain	1.01	6.91	36.1	×
HSPA5	endoplasmic reticulum chaperone BiP	0.67	5.13	60.6	✓

Surprisingly, the well-known FICD target protein HSPA5 was not significantly enriched, with a \log_2 fold change value of only 0.67. But as HSPA5 is a known target of FICD, it was also included in the reference list.^{53–55} Thus, the final list comprises 26 entries. 14 of these proteins were already identified in previous studies investigating AMPylated proteins in HeLa cells under physiological conditions with two different phosphoramidate probes.^{57,97} It should be kept in mind that these studies also identified several other proteins, that were not present in the current study for unknown reason. It is noteworthy that some of them appeared in several of the later infection experiments, for example STOM or various cathepsins (see Figure 16 B, 19 A).

2.3. Identification and optimization of suitable infection conditions

To enable comparison of independently infected replicates or samples, it is necessary to fix the ratio of bacteria to human cells. To implement this, a measure of the number of viable bacteria within a bacterial culture is needed. The colony forming unit (CFU) value indicates the number of viable bacteria within a corresponding culture. It needs to be determined individually for each growth condition of an organism. This determination is performed *via* a CFU assay, which establishes a correlation between the optical density at 600 nm of a culture (OD_{600} value) and the CFU count per mL.¹²⁷ To obtain reliable results, this assay was performed at least twice in an identical manner. An overnight culture of *V. parahaemolyticus* was diluted 1:100 into fresh LB media (with a total of 3% NaCl) and cultivated (200 rpm, 30 °C) following standard growth conditions.¹²⁸ At defined time points, the OD_{600} value was measured followed by serial dilutions and inoculation on LB agar plates. Plates containing between 12 and 120 colonies were counted and included in the calculation. For *V. parahaemolyticus* wild type, an exponential trend line represented the obtained data most accurately (Figure 11 B).

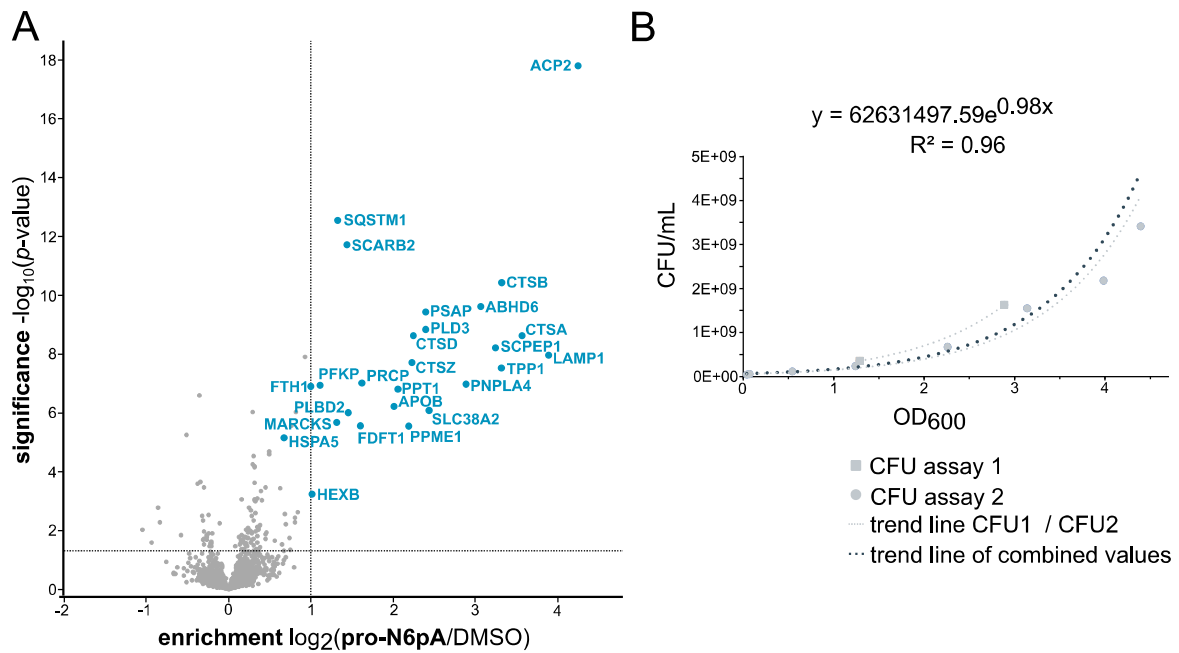


Figure 11: Intrinsically AMPylated proteins in human cells and CFU/mL to OD₆₀₀ ratio for *V. parahaemolyticus* wild type. (A) Volcano plot of the *in situ* LFQ reference experiment in HeLa cells treated with 100 μM **pro-N6pA** or 0.1 vol% DMSO (n = 12, after data processing: 1240 proteins). Proteins meeting the threshold criteria (log₂ enrichment > 1; p-value < 0.05) as well as the ER chaperone HSPA5 are highlighted in blue. (B) Correlation between OD₆₀₀ and the CFU count per mL of *V. parahaemolyticus* wild type when growing the bacteria in a 100 mL culture flask without baffles. The two independently performed CFU assays are accentuated in light gray, and data points can be assigned to the respective experiment based on their external shape. For each individual experiment, a trend line was calculated using an exponential fit. The trend line in dark gray was calculated by combining all measured data. The experimentally determined correlation between OD₆₀₀ and CFU/mL is indicated. Adapted from Rauh *et al.*⁸⁰

With the calculated correlation, different samples can be infected with identical multiplicities of infection (MOIs). For initial infection studies, HeLa cells were seeded into 6-cm dishes, including two dishes for cell counting, and treated with 100 μM **pro-N6pA** or DMSO (0.1 vol%) overnight (16 h). Cells grown in the additional two dishes were counted before the start of the experiment as they are the basis for the required CFU calculation. Next, the OD₆₀₀ value of the freshly inoculated *V. parahaemolyticus* day culture was measured and the required culture volume for a fixed stoichiometry between human and bacterial cells [bacterial multiplicity of infection (MOI)] was calculated. Following previous studies, MOIs of 10 or 20 were selected, representing a tenfold or twentyfold excess of bacteria over human cells, respectively.²⁹ Several experimental runs were required until two slightly different and suitable workflows were established (Figure 12).

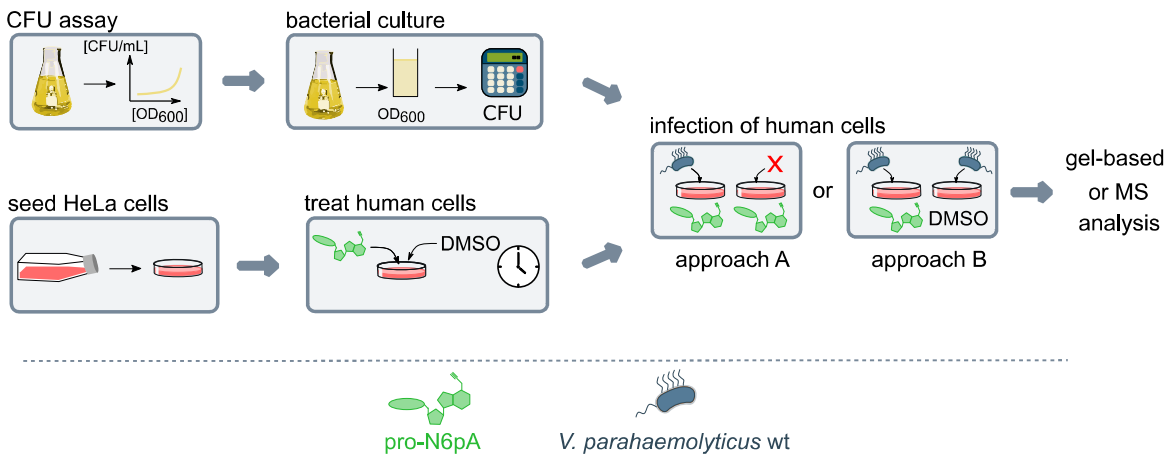


Figure 12: Optimized infection workflow of **pro-N6pA** treated HeLa cells with *V. parahaemolyticus* RIMD 2210633 infection, graphically illustrating the two slightly different approaches A and B. The individual working steps are shown, which consist of the performance of an CFU assay, bacterial growth and probe treatment of HeLa cells followed by bacterial infection. The successful infection workflow is followed by a proteomic profiling workflow. Adapted from Rauh *et al.*⁸⁰

Problems encountered included that samples with previously adjusted protein concentrations showed varying band intensities when analyzed *via* Coomassie stained SDS-PAGE. This problem occurred due to uneven cell lysis caused by problems in removing the remaining bacteria after successful infection. Optimization of the lysis buffer and stringent washing and centrifugation steps solved this problem. This was an important step, as the presence of too many bacterial proteins in the samples also interferes with the final LC-MS/MS based analysis of human proteins by reducing the quantity of their measured peptides. Another hurdle was the reduction of washing steps of bacterial pellets prior to the infection. These also had to be reduced to a minimum to exclude losses of bacteria remaining in the supernatant after centrifugation. Notably, to ensure sufficient probe availability, bacteria were resuspended in medium containing 10 μ M **pro-N6pA** before being added to HeLa cells. A valuable indicator for assessing the progress of *V. parahaemolyticus* infection was the HeLa cell phenotype.

The morphology of HeLa cells drastically changes after *V. parahaemolyticus* wild type infection (Figure 13 A 1,2).¹²⁹ This is because AMPylation of the small Rho GTPases affects the actin cytoskeleton, resulting in cell rounding.²⁹ Once the infection workflow was successfully established, infections with different bacterial MOIs (10 and 20) were performed for different periods of time (90 min and 150 min, Figure 13 B). After infection with tenfold excess of bacteria for 1.5 h, labeling of proteins different to those identified as intrinsically AMPylated (Figure 12 B, first two lanes) was already observed. Neither an increase in infection times nor an increase in the number of bacteria resulted in stronger labeling intensities. On this

basis, the parameters for all further experiments with *V. parahaemolyticus* infection were set to HeLa cell treatment with **pro-N6pA** for 16 h prior to infection with *V. parahaemolyticus* MOI of 10 for 90 minutes. A repetition of the infection experiment with the parameters above revealed two bands located around 23 kDa, which correspond to the size of human Rho GTPases (Figure 13 C). The absence of other bands does not preclude further AMPylation targets of VopS, since protein expression levels affect visibility on a gel.

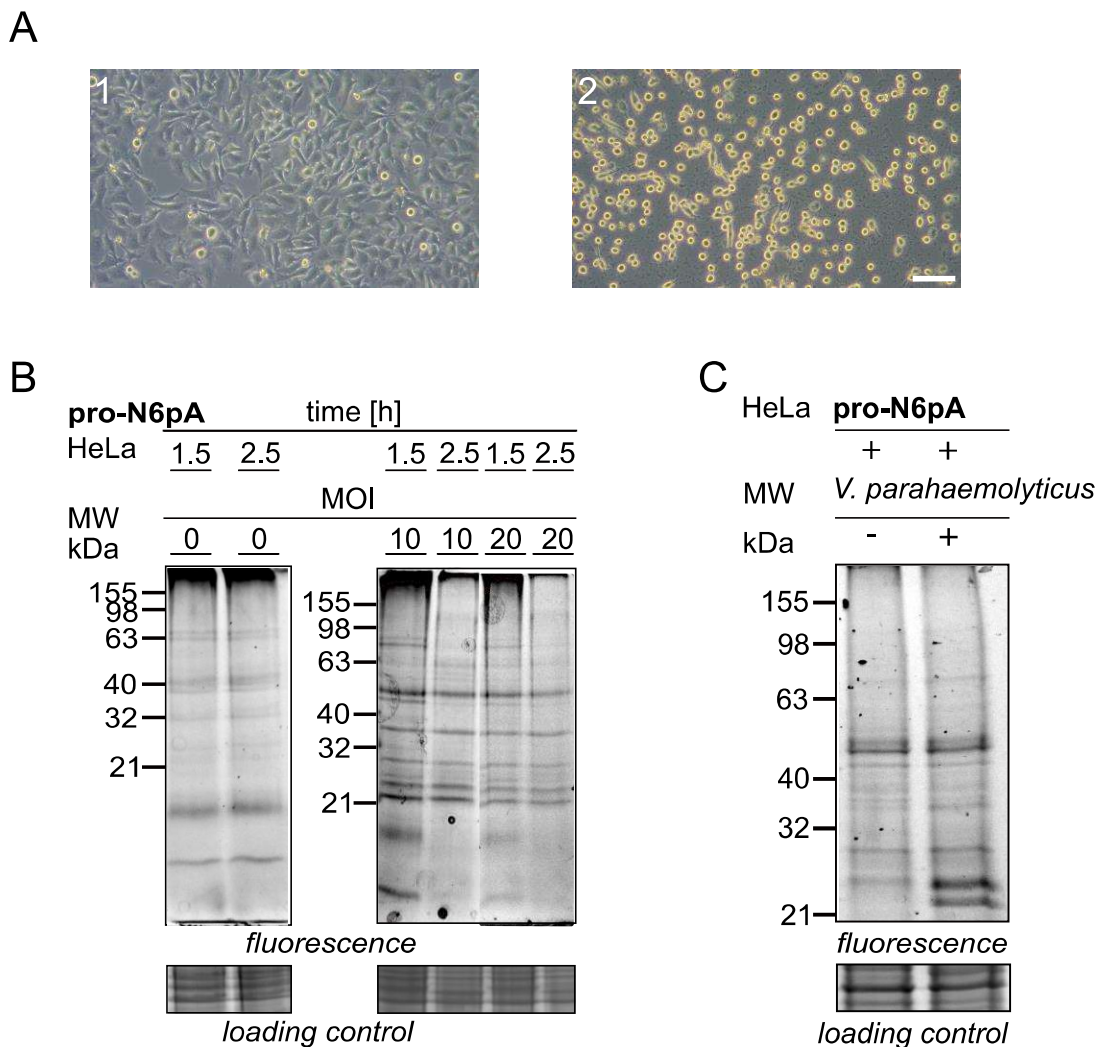


Figure 13: HeLa cell infection with *V. parahaemolyticus*. **(A)** Phenotypic changes of HeLa cells after bacterial infection. (1) HeLa cells, 100 μ M **pro-N6pA** (16 h), no bacteria or (2) HeLa cells, 100 μ M **pro-N6pA** (16 h), infected with *V. parahaemolyticus* wild type (MOI = 10 for 92 min). The scale bar represents 100 μ m. **(B)** SDS-PAGE analysis of HeLa cell infection with different MOIs (10 or 20) of *V. parahaemolyticus* showed that an infection with tenfold excess of bacteria for 1.5 h already resulted in the appearance of several specific protein bands. Increasing infection times or more bacteria did not yield stronger protein bands. **(C)** Gel-based analysis of uninfected HeLa cell lysate as well as HeLa cell lysate previously infected with *V. parahaemolyticus*. Both conditions were treated with 100 μ M **pro-N6pA**. Adapted from Rauh *et al.*⁸⁰

Enrichment of labeled proteins by affinity purification as well as analysis by the more sensitive and quantitative LC-MS/MS was next desired, both of which could facilitate identification of low abundant targeted proteins.

To further optimize infection conditions, the progress of protein AMPylation during bacterial infection was monitored. Because AMPylation is a highly dynamic process, the phenotypic cell rounding (Figure 13 A) could be a lagging indicator that does not reflect the actual status of AMPylation. There might be a temporal discrepancy between the post-translational modification and the resulting phenotypic change. To validate this hypothesis, **pro-N6pA** (100 μ M, 16 h) treated HeLa cells were infected with *V. parahaemolyticus* wild type (MOI 10). Every ten minutes, the infection of one sample was interrupted. Human cells were separated from remaining bacteria and clicked to rhodamine-azide after cell lysis. Surprisingly, the two characteristic 23 kDa bands were visible after infection times between 20 and 40 minutes (Figure 14 A). As suspected, phenotypic changes in HeLa cells lagged significantly and became apparent only after 60 to 90 minutes (Figure 14 B). With respect to AMPylation, this means that the resulting molecular dysregulation does not immediately lead to cell damage.

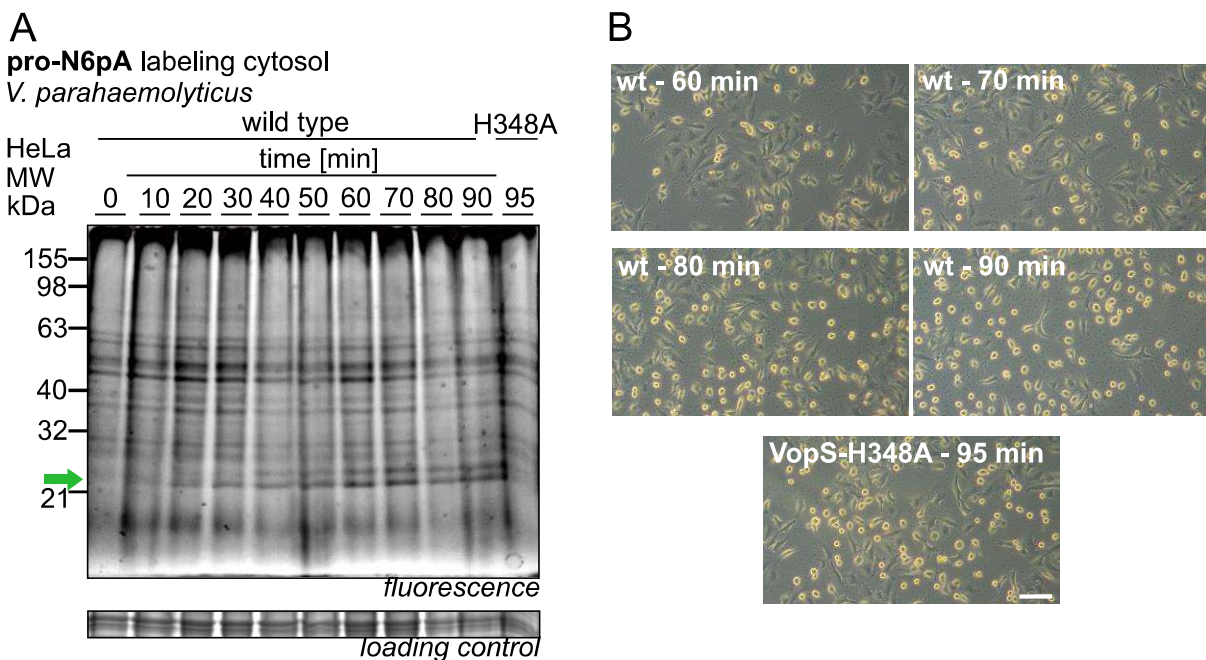


Figure 14: AMPylation of Rho GTPases does not immediately lead to HeLa cell rounding. **(A)** Labeling of AMPylated proteins (100 μ M **pro-N6pA**) in time increments of 10 min with increasing infection times up to 90 min (*V. parahaemolyticus* wt) and infection with the VopS-H348A mutant for 95 min as a control. **(B)** Microscope imaging of HeLa cells infected with *V. parahaemolyticus* wild type for different periods of time as indicated. As a control, the last picture represents infection with *V. parahaemolyticus* mutant VopS-H348A for 95 min (see section 2.4). The scale bar represents 100 μ m. Adapted from Rauh *et al.*⁸⁰

2.4. Construction and validation of the inactive *V. parahaemolyticus* mutant VopS-H348A

To be able to control or validate results obtained in future experiments, the inactive *V. parahaemolyticus* VopS mutant H348A (from now on referred to as mutant VopS-H348A) was constructed in cooperation with Sophie Brameyer (LMU). In bacteria expressing this VopS mutant, the catalytically active histidine was replaced by an alanine residue. Due to its methyl side chain, alanine can no longer form the required interactions with the ATP molecule, which precludes AMPylation reactions (compare Figure 3).²⁹ To validate the constructed mutant, the CFU to OD₆₀₀ correlation was determined first (Figure 15 A). Interestingly, the mutant VopS-H348A showed increased growth compared to the wild type strain under the same conditions. Next, HeLa cells were infected with *V. parahaemolyticus* expressing VopS-H348A according to the previously established procedure. The resulting HeLa cell lysate was clicked to the fluorescent dye and an SDS-PAGE was performed (Figure 15 B).

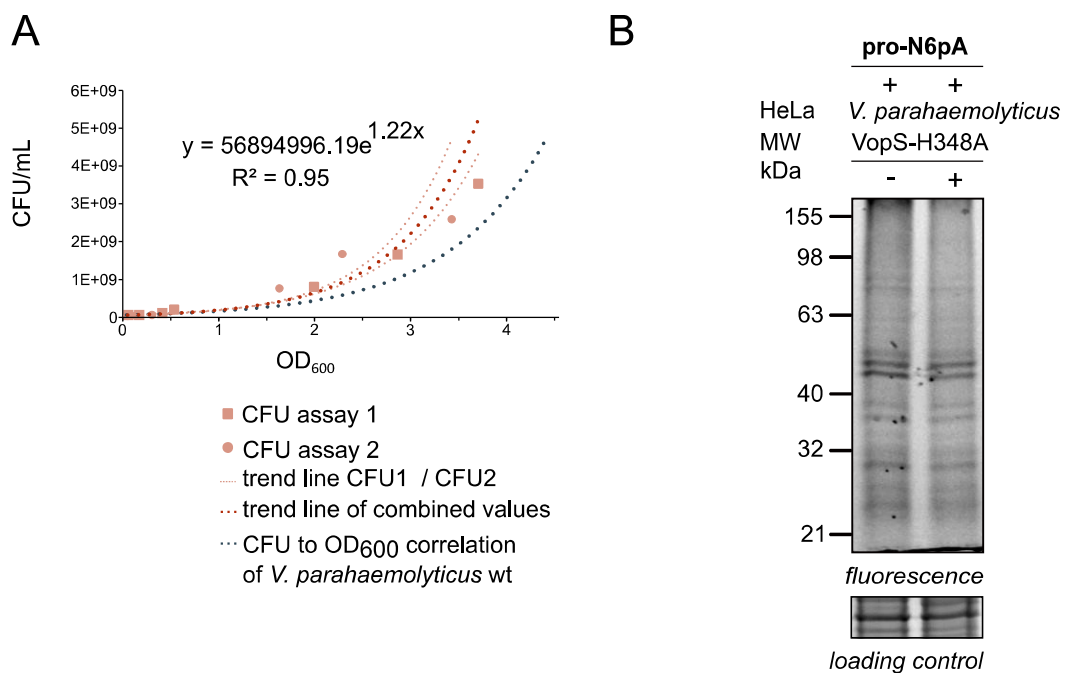


Figure 15: (A) Correlation between OD₆₀₀ and the CFU count per mL of *V. parahaemolyticus* VopS-H348A when growing the bacteria in a 100 mL culture flask without baffles. The two independently performed CFU assays are accentuated in light red, and data points can be assigned to the respective experiment based on their external shape. For each individual experiment, a trend line was calculated using an exponential fit. The trend line in dark red was calculated by combining all measured data. The experimentally determined correlation between OD₆₀₀ and CFU/mL is indicated. In addition, the correlation between CFU/mL and OD₆₀₀ for *V. parahaemolyticus* wild type is indicated in gray. (B) Gel-based analysis of uninfected HeLa cell lysate as well as HeLa cell lysate previously infected with *V. parahaemolyticus* VopS-H348A. Both conditions were treated with 100 μM pro-N6pA. Adapted from Rauh *et al.*⁸⁰

As expected, this infection did not cause the appearance of the characteristic protein bands around 23 kDa, indicating an inability of the mutant to AMPylate. In line with these results, VopS-H348A also abrogated the observed cell rounding in HeLa cells (Figure 14 B).

2.5. Assessment of different infection and quantification strategies

Since some VopS targets (RHOA, RAC1, CDC42)²⁹ are already known, these proteins could be monitored to assess the suitability of this workflow to identify unknown protein targets of the *V. parahaemolyticus* effector. For this purpose, mass spectrometry was used as a monitoring tool to evaluate the applied method.

2.5.1. Protein enrichment experiments according to approach A combined with LFQ

The first proteome profiling experiment combined a protein enrichment strategy using LFQ quantification with probe treatment of both conditions, but bacterial infection of only one (referred to as approach A, for illustration see Figure 12). Briefly, HeLa cells were seeded in 6-cm or 10-cm dishes and all dishes were treated with **pro-N6pA** for 16 h. Then, half of the dishes were infected with *V. parahaemolyticus* wild type, and the other half were treated the same except for the addition of bacteria. Finally, modified proteins were enriched by clicking them to azide-PEG₃-biotin followed by avidin bead pull-down. This experiment was designed such that host proteins AMPylated by *V. parahaemolyticus* should appear significantly enriched compared to the non-infected control, with intrinsically AMPylated proteins being instead equally enriched in both conditions. However, the disadvantage of this methodology is that samples infected with *V. parahaemolyticus* may experience dysregulation of protein expression due to bacterial infection. This then leads to non-specific accumulation of these proteins in the enrichment samples. As expected, the volcano plot showed hardly any protein enriched, that is intrinsically AMPylated by HeLa cells (Figure 16 A, p. 37). As the volcano plot showed an overall broad distribution of the identified proteins, the threshold was adjusted for this experiment to a more stringent log₂ enrichment > 2, while the *p*-value < 0.05 remained the same. Interestingly, CDC42 and RHOA as known VopS targets were significantly enriched, with RAC1 also enriched, but with a smaller log₂ fold change of 1.48 (Table 2). Several other proteins also met the threshold criteria (compare Figure 16 A) and could therefore represent possible, previously unknown targets of VopS.

Table 2: Enrichment data of VopS targets CDC42, RHOA and RAC1 for protein enrichment experiments using approach A in combination with LFQ analysis (Figure 16 A, Figure S1).

V. parahaemolyticus wild type, protein enrichment, approach A, LFQ

gene name	log ₂ fold change	-log ₁₀ <i>p</i> -value	coverage [%]	significantly enriched?
CDC42	2.50	5.14	42.9	✓
RHOA	2.25	3.05	39.3	✓
RAC1	1.48	2.69	45.3	×

V. parahaemolyticus mutant VopS-H348A, protein enrichment, approach A, LFQ

gene name	log ₂ fold change	-log ₁₀ <i>p</i> -value	coverage [%]	significantly enriched?
CDC42	0.39	0.79	37.2	×
RHOA	0.02	0.01	38.3	×
RAC1	1.08	3.40	31.2	×

Given these initial promising results, the *V. parahaemolyticus* VopS-H348A was used in the same proteomic profiling workflow. The corresponding infection time was set to the same time as for wild type infection (90 minutes). The resulting volcano plot confirmed the basic functionality of this proteome profiling approach, since now RHOA and CDC42 showed a log₂ fold enrichment < 0.5 (Table 2, Figure S1, section 6.). Only RAC1 did not show such a radical decrease in enrichment. However, since very large numbers of proteins were dysregulated after infection, another approach which takes these changes into account would enable clearer delineation of VopS targets. An alternative profiling approach facilitating the identification of bacterially AMPylated proteins was therefore desired next.

A final evaluation of the methodology was performed, this time using the Gram-positive pathogenic strain *Staphylococcus aureus* USA300. *S. aureus* is an opportunistic pathogen, which causes several human diseases and is present in human nares. Approximately 20% of the healthy population are permanently infected with the bacterium, while 60% are temporarily colonized.¹³⁰ The strain USA300 contains one AMPylator, a Fic-domain containing protein (A0A0H2XIM6, gene SAUSA300_1560), which belongs to the class II Fic proteins and possesses the conserved Fic motif.⁴⁵ Its function is yet to be elucidated. After CFU assay performance (Figure S2 A), initial gel-based analyses revealed an infection time for 2 h with *S. aureus* MOI of 20 as best infection conditions (data not shown). In contrast to previous results, the corresponding plot showed few differences between infected and non-infected

samples, and very few enriched proteins at all (threshold \log_2 enrichment > 1 ; Figure S2 B). Since no Gram-positive bacterial strain is yet known to AMPylate host proteins upon infection, this may indicate that this mechanism is a characteristic feature of Gram-negative strains.

2.5.2. Protein enrichment experiments according to approach B combined with DML

The previous method using a non-infected control was successful in terms of probe enrichment of RHOA and CDC42, but the resulting volcano plots showed a diffusely enriched protein cloud from which the monitored target proteins did not stand out. Therefore, a protein or peptide labeling strategy was chosen as this promises a more reliable quantification of identified peptides and proteins by comparing the same peptide in all samples, just with different labels. Since dimethyl labeling (DML) is experimentally faster to implement than SILAC labeling, this method was the first choice.^{119,131} Additionally, the diffuse volcano plot shape could be the result of bacterially induced proteome changes, so for further experiments all samples were infected with bacteria and only half treated with **pro-N6pA** (referred to as approach B, for illustration see Figure 12).

For DML, infection experiments can be performed according to standardized workflows. To introduce the corresponding labels, digested peptides were treated with solutions containing either “heavy” or “light” labels (45 mM sodium phosphate, 30 mM NaBD₃CN and 0.2 % ¹³C-formaldehyde-*d*₂ for “heavy” labeling or 45 mM sodium phosphate, 30 mM NaBH₃CN and 0.2 % formaldehyde for “light” labeling). Complementary conditions were then combined, *i.e.*, probe-treated, and “heavy” labeled peptides with DMSO-treated “light” labeled peptides and *vice versa*. Despite its high reliability, DML did not provide results as good as the previous proteomic profiling strategy.¹²⁵ Even with the enabled MaxQuant software functions of match between run and re-quantify, the volcano plot was only comprised of 272 proteins after filtering (Figure 16 B). This is a very low number compared to a typical protein count of over 1000. To qualitatively and quantitatively specify the peptide abundance in the samples during DML, the labeled peptides are compared with their counterparts in terms of the MS intensities occurring in the samples.¹³² With the approach used here, theoretically only proteins with incorporated **pro-N6pA** are enriched as their attached biotin tag adheres to the avidin beads. Since **pro-N6pA** is only present in half of the samples, AMPylated proteins are also enriched only in these samples.

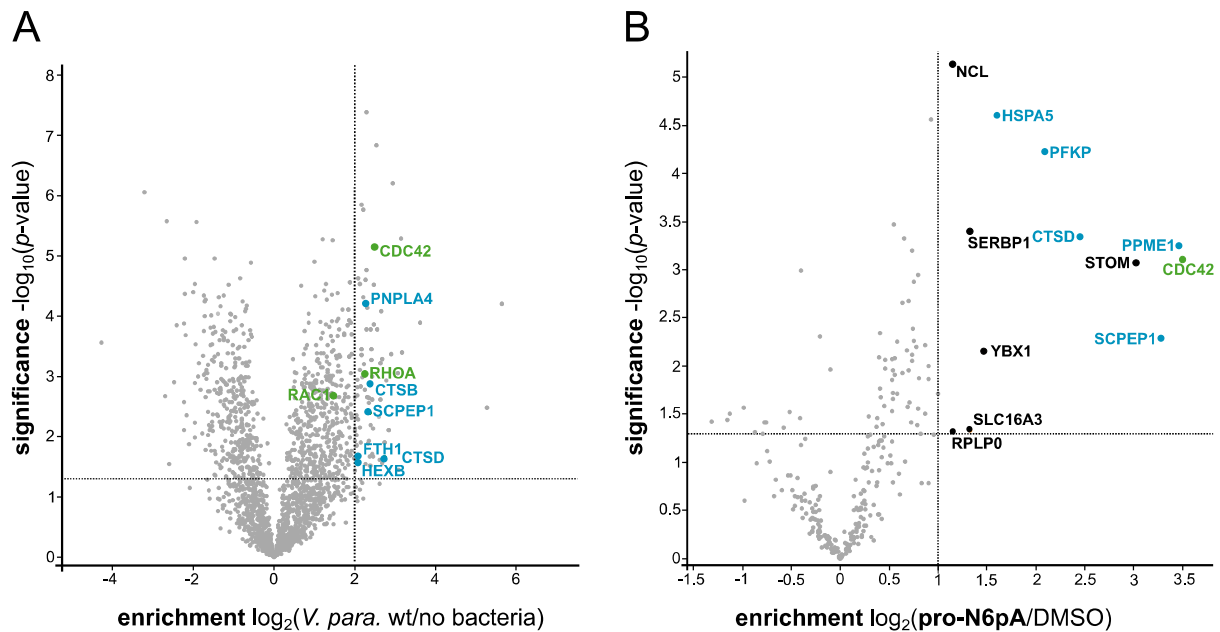


Figure 16: HeLa cell infection with *V. parahaemolyticus*. **(A)** Volcano plot of the *V. parahaemolyticus* wild type *in situ* infection experiment applying LFQ and approach A (100 μM pro-N6pA all samples, only half of them infected with bacteria). The experiment was performed with 5 replicates, and after data processing 2465 proteins remained. Threshold criteria were set to \log_2 enrichment > 2 and p -value < 0.05 . Intrinsically AMPylated proteins are highlighted in blue, while VopS target proteins are marked in green. **(B)** Volcano plot of proteomic profiling experiment infecting HeLa cells with *V. parahaemolyticus* wild type according to approach B (all infected, only half of them treated with probe) combined with DML ($n = 5$, after data processing: 272 proteins). Intrinsically AMPylated proteins are highlighted in blue, while VopS target proteins are marked in light green. Possible new targets of VopS are marked in black.

Thus, the counterparts of the peptides in the DMSO-treated samples are not bound to the avidin beads, *i.e.*, they are not enriched, which makes quantification impossible. Nevertheless, the volcano plot showed some proteins being part of the intrinsically AMPylated reference list and furthermore CDC42 as VopS target (Table 3).

Table 3: Enrichment data of VopS targets CDC42, RHOA and RAC1 for protein enrichment experiments using approach B in combination with DML; *N/A* = not available (Figure 16 B).

V. parahaemolyticus wild type, protein enrichment, approach B, DML

gene name	\log_2 fold change	$-\log_{10} p\text{-value}$	coverage [%]	significantly enriched?
CDC42	3.51	3.11	37.2	✓
RHOA	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>
RAC1	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>

2.5.3. Protein enrichment / full proteome analysis according to approach A combined with SILAC labeling

As DML did not provide the desired result, further experiments were performed using an alternative labeling strategy, SILAC labeling. By passaging HeLa cells at least six times in corresponding media (SILAC-DMEM with 10 % dialyzed FBS, 2 mM L-glutamine, and 214 μM [$^{13}\text{C}_6$, $^{15}\text{N}_4$] L-arginine HCl (Arg10), 419 μM [$^{13}\text{C}_6$, $^{15}\text{N}_2$] L-lysine 2 HCl (Lys8) for “heavy” cells or with 214 μM [$^{13}\text{C}_6$] L-arginine HCl (Arg6), 419 μM [4,4,5,5-D4] L-lysine 2 HCl (Lys4) for “light” cells), they exclusively incorporate the labeled amino acids. Thus, SILAC represents a robust method allowing more precise quantification as DML.¹¹⁷ Since peptide quantification in SILAC labeling experiments functions as for DML, this time all samples were treated with **pro-N6pA** (approach A) to ensure protein enrichment. To account for proteome changes due to the unilateral infection with bacteria, SILAC lysate was divided into two parts and a proteomic enrichment experiment as well as a full proteome analysis were performed. In this way, significant protein enrichment can be attributed to either actual enrichment or to enrichment caused by protein overexpression due to invading bacteria.

For SILAC labeling, two cell populations were similarly passaged, one in medium supplemented with “heavy” labeled amino acids and one with “light” labeled ones. The infection workflow was performed according to the standardized procedure such that “heavy” and “light” labeled cells were used for approximately equal numbers of infected and uninfected samples. After successful infection, the protein concentration of the lysate was determined. The advantage of SILAC labeling is, that samples can be combined directly after the bacterial infection. Thus, 250 μg of protein of a “heavy” labeled and infected sample were combined with 250 μg of protein of a “light” labeled and uninfected sample and *vice versa*. The same procedure was used for the full proteome analysis, except that in this case 100 μg protein per sample were considered sufficient. The resulting volcano plot of the full proteome analysis (\log_2 enrichment < -1 ; \log_2 enrichment > 1) indicated several human proteins up- or down-regulated (Figure 17 A). A STRING database analysis (v.11.0) of these proteins, looking for high confidence interactions, revealed no association between them. Importantly, RHOA, RAC1, and CDC42 were neither up- nor down-regulated. This means that their expression levels are not affected by the presence of *V. parahaemolyticus*. Since all three proteins showed strong significant enrichment in the second volcano plot (Figure 17 B), this is clearly due to their AMPylation by VopS.

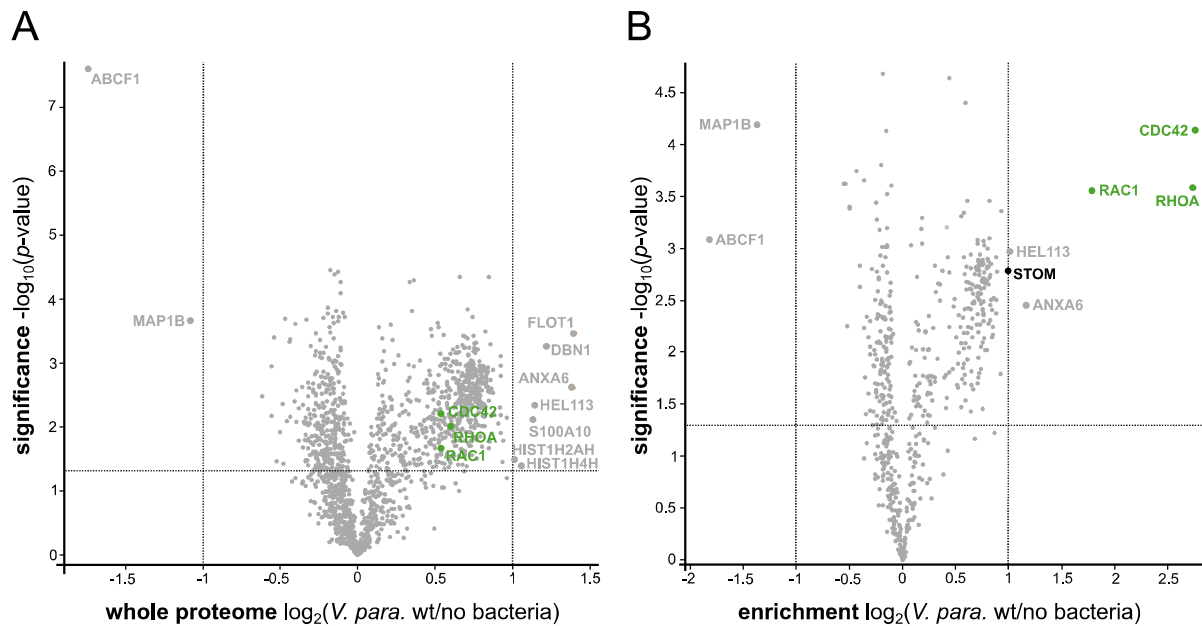


Figure 17: Volcano plots of infection experiments of HeLa cells using approach A with *V. parahaemolyticus* wild type and SILAC labeling. (A) Whole proteome analysis ($n = 6$, after data processing: 1538 proteins) and (B) proteomic profiling experiment ($n = 6$, after data processing: 572 proteins) both treated with $100 \mu\text{M}$ **pro-N6pA**. For each experiment, only one condition was treated with *V. parahaemolyticus* wild type MOI 10 for 90 min. Protein targets of VopS are highlighted in green, while proteins that are simultaneously enriched or reduced in both experiments are highlighted in gray. Possible new targets of VopS are marked in black. Adapted from Rauh *et al.*⁸⁰

Additionally, HEL113 (vimentin), STOM (stomatatin) and ANXA6 (annexin A6) were detected as significantly enriched. However, as HEL113 and ANXA6 were also enriched in the whole proteome analysis, they are presumably overexpressed due to the bacterial infection. This lowers the likelihood that these proteins are targets of VopS. STOM on the other hand had no elevated expression levels in the whole proteome analysis. However, because this protein exhibited increased \log_2 enrichment levels when HeLa cells were treated solely with **pro-N6pA**, it most likely represents an intrinsically AMPylated protein.⁹⁷ Nevertheless, the strategy combining approach A with SILAC labeling represents as the most promising method so far.

To further validate this strategy, the same experiment was performed with the VopS-H348A mutant (Figure 18). Full proteome analysis once more showed that RAC1, RHOA and CDC42 levels did not change upon bacterial infection. Proteomic enrichment also did not identify these proteins as significantly enriched. More precisely, only RHOA was detected in this sample set. CDC42 and RAC1 were filtered out during data processing because no "heavy/light" ratio could be calculated for them in three and four samples, respectively. The re-enrichment of STOM in this infection experiment confirms the previous assumption that it is not a VopS target.

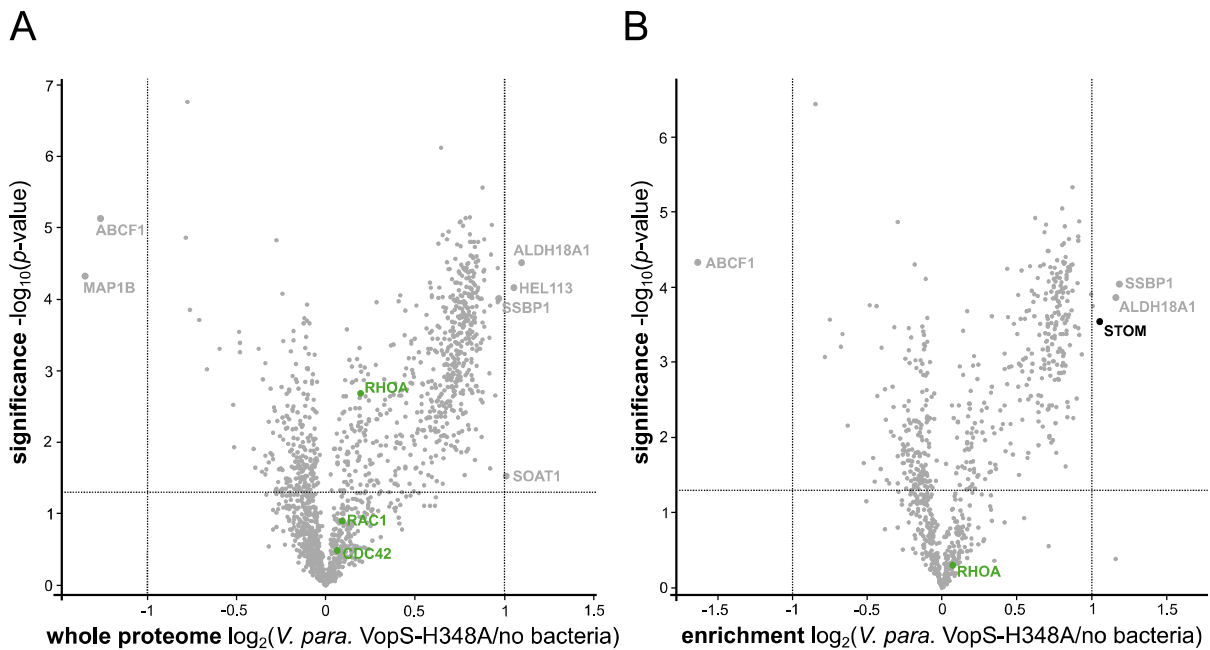


Figure 18: Volcano plots of infection experiments of HeLa cells using approach A with *V. parahaemolyticus* VopS-H348A and SILAC labeling. (A) Whole proteome analysis (n = 6, after data processing: 1487 proteins) and (B) proteomic profiling experiment (n = 6, after data processing: 784 proteins) both treated with 100 μ M **pro-N6pA**. For each experiment, only one condition was treated with *V. parahaemolyticus* VopS-H348A MOI 10 for 90 min. Protein targets of VopS are highlighted in green, while proteins that are simultaneously enriched or reduced in both experiments are highlighted in gray. Proteins enriched only in the proteome profiling experiment are shown in black.

The other two proteins enriched in this plot, SSBP1 and ALDH18A1 were proteins that have been identified as significantly overexpressed due to bacterial infection (Figure 18 A). The enrichment data regarding VopS target proteins in SILAC labeling experiments are summarized in Table 4.

To further validate the ability of this method to decipher unknown host AMPylation targets of bacterial pathogens, this profiling strategy was tested in the Gram-negative bacterium *Pseudomonas aeruginosa*. This pathogen is on the World Health Organization's top priority list of pathogens requiring new treatment options.¹³³ *P. aeruginosa* not only poses a threat to patients with cystic fibrosis, but it is also a leading cause of community-acquired and nosocomial-acquired pneumonia. Thereby, a considerable proportion of infections are caused by multidrug-resistant strains.¹³⁴ The strain *P. aeruginosa* PAO1 contains two Fic proteins on genes PA1366 and PA0574 (PA1366: Q9I3X8 and PA0574: Q9I5W2). The protein PA1366 is a promising candidate because, like VopS and IbpA, it has no regulatory α -helix. But unlike them, it has no conserved Fic motif.⁴⁵

Table 4: Enrichment data of CDC42, RHOA and RAC1 infection experiments with *V. parahaemolyticus* wild type and mutant VopS-H348A. Results for full proteome analysis and protein enrichment experiments (each n = 6) using approach A in combination with SILAC labeling; *N/A* = not available.

***V. parahaemolyticus* wild type, full proteome, approach A, SILAC**

gene name	log ₂ fold change	-log ₁₀ <i>p</i> -value	coverage [%]	significantly enriched?
CDC42	0.54	2.20	26.2	×
RHOA	0.60	2.00	47.7	×
RAC1	0.54	1.67	39.9	×

***V. parahaemolyticus* wild type, protein enrichment, approach A, SILAC**

gene name	log ₂ fold change	-log ₁₀ <i>p</i> -value	coverage [%]	significantly enriched?
CDC42	2.76	4.14	42.9	✓
RHOA	2.75	3.58	27.7	✓
RAC1	1.78	3.55	36.5	✓

***V. parahaemolyticus* mutant VopS-H348A, full proteome, approach A, SILAC**

gene name	log ₂ fold change	-log ₁₀ <i>p</i> -value	coverage [%]	significantly enriched?
CDC42	0.06	0.49	26.2	×
RHOA	0.19	2.69	52.3	×
RAC1	0.09	0.90	39.9	×

***V. parahaemolyticus* mutant VopS-H348A, protein enrichment, approach A, SILAC**

gene name	log ₂ fold change	-log ₁₀ <i>p</i> -value	coverage [%]	significantly enriched?
CDC42	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>
RHOA	0.07	0.29	26.4	×
RAC1	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>

The second Fic protein, PA0574, contains this conserved motif and in addition an inhibitory α -helix at the *N*-terminus (class II).⁴⁵ Once again, a CFU assay was performed prior to infection experiments (Figure S3 A). Initial gel-based analyses revealed an infection time of 5 h with *P. aeruginosa* MOI of 10 as best infection conditions (data not shown). During infection, *P. aeruginosa*, like *V. parahaemolyticus*, causes cell rounding of human host cells (Figure S3 B). Comparing the enriched proteins from the whole proteome analysis with the enriched

proteins from the proteome profiling experiment, it is striking that most of them overlap (Figure S3 C, D). These proteins therefore do not represent reliable hits as putative AMPylated proteins. Heterogeneous nuclear ribonucleoproteins C1/C2 (HNRNPC) is the only protein significantly enriched solely in the second data set. However, this most probably is only because it was not identified in the whole proteome analysis. Overall, this experiment gave no evidence that human proteins are AMPylated by the two bacterial Fic proteins of *P. aeruginosa* PAO1.

To assess once more the ability of Gram-positive bacteria to AMPylate host proteins, SILAC full proteome analysis as well as SILAC enrichment experiments were also performed with *S. aureus* USA300. As short reminder, the previous strategy of *S. aureus* infection followed by protein enrichment and LFQ analysis resulted in hardly any enriched proteins. This was again observed in the SILAC labeling experiments (Figure S2 C, D). For this reason, the focus for further experiments was set exclusively on Gram-negative bacteria.

In summary, SILAC labeling proved to be a robust methodology to evaluate the ability of pathogenic bacteria to AMPylate human host proteins. However, it required the full proteome analysis as well as protein enrichments experiments being performed in parallel. Since the methodology entails a smaller number of samples due to the combination of “heavy” and “light” labeled conditions, this does not limit its utility. In addition to RHOA, RAC1 and CDC42, several more small Rho GTPases as well as other proteins have been identified as putative VopS targets in previous *in vitro* experiments.^{83,84,93} None of them have been enriched in the performed *in situ* experiments so far. Therefore, these proteins are either mistakenly predicted as VopS targets or that there is still room for improvement regarding the experimental setup and analysis strategy.

2.5.4. Protein enrichment experiments according to approach B combined with LFQ

As a next strategy, approach B was combined with a quantification method that does not require a peptide to be identified in all samples in order to quantify it. Therefore, all samples were infected with *V. parahaemolyticus* wild type MOI 10, whereby one half was treated with 100 μ M **pro-N6pA**, the other with 0.1 vol% DMSO. This methodological setup eliminates uncertainties due to changes on the proteome level caused by infecting or invading bacteria. Furthermore, LFQ provides accurate sample analysis even if some peptides are not identified in all samples during MS measurements.¹¹⁵ As a minor drawback, the intrinsically AMPylated host proteins also emerge as significantly enriched in the resulting volcano plot. Thus,

intrinsically AMPylated proteins which have not been identified as those might mistakenly be assumed as VopS targets, representing false positives.

In addition to some of the proteins on the reference list, the plot surprisingly showed not only the known VopS targets RHOA, RAC1, and CDC42, but also the two putative targets RHOG and RHOC near the cutoff line (Figure 19 A, p. 45).⁸⁴ Due to the tighter protein distribution of the plot, the threshold was again set to \log_2 fold change > 1 . Once more, several other proteins met the threshold criteria (Table 5). They could be either unknown targets of VopS, or intrinsically AMPylated by HeLa cells and were not identified in previous experiments. However, these enriched proteins do not overlap with the enriched proteins from previous experiments, making it less likely that they are VopS targets.

Table 5: Summary of all proteins that met the threshold criteria plus RHOG, CDC42 and RAC1 in the proteomic profiling experiment infecting HeLa cells with *V. parahaemolyticus* wild type according to approach B and analyzed via LFQ; green = VopS protein targets, blue = intrinsically AMPylated proteins, dark green = possible unknown VopS targets.

***V. parahaemolyticus* wild type, protein enrichment, approach B, LFQ**

gene name	protein name	\log_2 fold enrichment	$-\log_{10}$ p-value	coverage [%]
RHOA	Transforming protein RhoA	1.77	5.40	48.2
RHOC	Rho-related GTP-binding protein RhoC	1.28	2.61	41.5
RHOG	Rho-related GTP-binding protein RhoG	0.94	3.91	34.6
CDC42	Cell division control protein 42 homolog	0.93	2.00	34.7
RAC1	Ras-related C3 botulinum toxin substrate 1	0.76	3.28	45.9
CTSA	Lysosomal protective protein	5.26	8.69	20.2
ABHD6	Monoacylglycerol lipase ABHD6	4.49	9.72	32
ACP2	Lysosomal acid phosphatase	3.88	7.82	35.5
LAMP1	Lysosome-associated membrane glycoprotein 1	3.85	5.11	8.6
SCPEP1	Retinoid-inducible serine carboxypeptidase	3.49	4.81	23.5
PNPLA4	Patatin-like phospholipase domain-containing protein 4	3.32	5.88	30.4
CTSB	Cathepsin B	3.17	4.55	20.1
APOB	Apolipoprotein B-100	3.04	4.08	1.6
TPP1	Tripeptidyl-peptidase 1	2.58	2.93	11.9
PLBD2	Putative phospholipase B-like 2	2.14	2.26	7
CTSZ	Cathepsin Z	2.13	2.94	10.6
CTSD	Cathepsin D	1.92	2.81	18.3
SCARB2	Lysosome membrane protein 2	1.85	5.08	24.1
SLC38A2	Sodium-coupled neutral amino acid transporter 2	1.33	5.69	17
FDFT1	Squalene synthase	1.28	6.97	45.6
SQSTM1	Sequestosome-1	1.24	5.63	48

FTH1	Ferritin heavy chain	1.13	4.41	61.7
PPME1	Protein phosphatase methylesterase 1	1.12	6.16	32.1
HSPA5	Endoplasmic reticulum chaperone BiP	1.12	4.84	55.8
ACSS2	Acetyl-coenzyme A synthetase	3.01	4.74	10.5
AK4	Adenylate kinase 4, mitochondrial	2.36	2.08	26
CTSL	Cathepsin L1	2.32	3.74	18.5
DPP7	Dipeptidyl peptidase 2	2.21	3.75	10.9
STC2	Stanniocalcin-2	2.16	3.26	13.8
SLC38A1	Sodium-coupled neutral amino acid transporter 1	1.73	2.14	5.3
FAM129A	Protein Niban	1.71	3.82	4.1
LGMN	Legumain	1.49	1.40	17.8
PCSK9	Proprotein convertase subtilisin/kexin type 9	1.34	4.15	12.6
SCARB1	Scavenger receptor class B member 1	1.08	2.65	6.7
ABHD10	Mycophenolic acid acyl-glucuronide esterase, mitochondrial	1.02	2.87	35

Previously, Yu *et al.* screened the human proteome for AMPylation targets of VopS via a NAPPA array (for overview see section 1.7).⁸⁴ Eight Rho GTPases produced a positive fluorescence signal in this assay, which confirmed them as AMPylation targets.⁸⁴ By aligning the sequences of these Rho GTPases, the group identified a conserved motif which is suspected to contain the AMPylation site. They then searched for this motif in all human proteins and identified seven others carrying this amino acid sequence. RHOC was among the hits of the initial NAPPA screen, whereas RHOG was also identified to contain the conserved motif.⁸⁴ For an unknown reason, RHOG was not identified as VopS hit in the NAPPA array. Besides these proteins, they also identified or predicted RAC2, RAC3, RHOB, RND3, RHOD, RHOJ, RHOQ, RND1, RND2 and ERGIC2 as VopS targets, none of which were found in the current volcano plot.⁸⁴

To validate the current approach, it was repeated with the *V. parahaemolyticus* mutant VopS-H348A once more. Statistical analysis of the obtained raw data confirmed the functionality of the approach by revealing a drastically diminished enrichment with a \log_2 fold change < 0 for all AMPylation targets identified in the wild type plot (Figure 19 B). Only RHOC was not found in this experiment (Table 6).

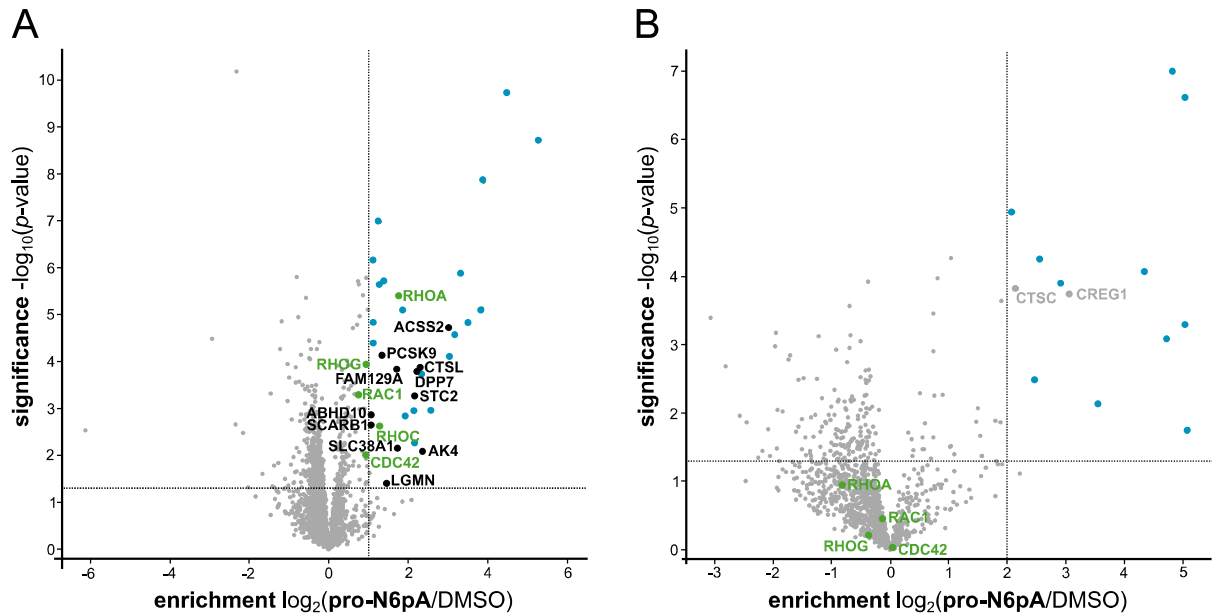


Figure 19: Protein enrichment plots of HeLa cell infection with *V. parahaemolyticus* wild type (**A**; $n = 5$; after data processing: 2766 proteins) or mutant VopS-H348A (**B**; $n = 5$; after data processing: 1099 proteins). All samples were infected with bacteria while only half of them were treated with 100 μM **pro-N6pA** (approach B). Analysis was performed *via* LFQ. Cutoff lines were determined as \log_2 fold enrichment > 1 and a p -value < 0.05 . Intrinsically AMPylated proteins are highlighted in blue. For clarity, their naming has been omitted. VopS targets are shown in light green while black presents possible new VopS targets. Proteins in gray met the threshold criteria but are excluded as VopS targets. Partially adapted from Rauh *et al.*⁸⁰

Most of the significantly enriched proteins are part of the reference list containing intrinsically AMPylated proteins. Proteins not included in this list but still significantly enriched could either be unknown to be AMPylated by a human AMPylator, or they are background binders.

Table 6: Enrichment data of initially identified VopS targets in the infection experiment with *V. parahaemolyticus* mutant VopS-H348A using approach B in combination with LFQ; *N/A* = not available.

V. parahaemolyticus mutant VopS-H348A, protein enrichment, approach B, LFQ

gene name	\log_2 fold change	$-\log_{10} p$ -value	coverage [%]	significantly enriched?
CDC42	0.06	0.04	37.20	×
RHOA	-0.81	0.96	38.30	×
RAC1	-0.13	0.45	31.20	×
RHOC	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>
RHOG	-0.37	0.21	33.50	×

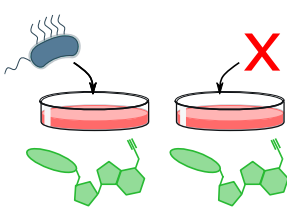
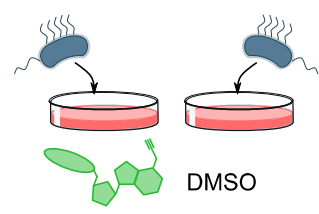
To check again if the *P. aeruginosa* PAO1 Fic domain-containing proteins have AMPylation targets in the human proteome, HeLa cells were infected with this pathogen. After sample processing and analysis, the plot revealed 17 significantly enriched proteins (Figure S3 E), of which 10 can be assigned as intrinsically AMPylated proteins. Of the remaining seven proteins, three were identified in a previous infection experiment using the same strategy but *V. parahaemolyticus* wild type instead of *P. aeruginosa*. Therefore, it is unlikely that these proteins do represent AMPylation targets of *P. aeruginosa*. They are more likely artefacts, which are caused by the infection workflow itself. It is possible, that bacterial infection of human cells leads to protein overexpression as part of a stress response. Since **pro-N6pA** is only present in one half of the samples, this may lead to an accumulation of the dysregulated proteins there. Four residual proteins have been enriched (PSAT1, PSPH, SLC7A1, MRPL1), that may be yet unknown AMPylation targets of this pathogen. However, none of these proteins had been identified in the previous infection experiment using SILAC labeling, and only the profile plot of PSAT1 looked promising (Figure S3 F). In this case, corresponding peptides were measured in all samples while for PSPH, SLC7A1 and MRPL1, no peptides were measured in several samples (mainly in DMSO treated ones), leading to missing value imputation when analyzing the data. Since the log₂ LFQ intensities for the **pro-N6pA** or DMSO-treated samples for these three proteins are not different from each other, their enrichment is probably the result of imputed numbers with low value. If no peptides of a protein are measured in a sample by LC-MS/MS, no LFQ intensity value for this protein can be given in the data tables. These missing values are further imputed during the data analysis by estimating a low LFQ intensity referring to the whole matrix. Further experiments would be necessary to confirm this assumption (see section 2.6).

As initial experiments did not show any promising results with Gram-positive bacteria, the Gram-negative strain *Escherichia coli* CFT073 was used to delve further into putative human AMPylation targets. This uropathogenic strain is among the major causes of urinary tract infections.¹³⁵ It is completely sequenced and its genome encodes for two putative Fic proteins, on genes c4409 and c4136 (A0A0H2VER3 and A0A0H2VBH5). Gene c4409 encodes a Fic domain-containing protein with conserved Fic motif and a class II inhibitory α -helix while the second one is a class I Fic enzyme that does not contain the conserved Fic motif.⁴⁵ CFU assays were also performed for this bacterium (Figure S4 A). Infections with this strain also caused slight cell rounding, however with a different phenotypic appearance than *V. parahaemolyticus* wild type infections (Figure S4 B). Initial gel-based analyses revealed an infection time of 2.5 h with *E. coli* MOI of 10 as the best infection conditions (data not shown). The resulting volcano

plot revealed 44 significantly enriched proteins, 20 of which are on the intrinsically AMPylated reference list (Figure S4 C). Some of the remaining 23 hits have been already observed in previous experiments with other pathogens, suggesting they are also intrinsically AMPylated proteins or background binders. Of the other enriched proteins, none had a literature known connection to AMPylation. For these proteins, further experiments are required to confirm them as AMPylation targets (see section 2.6).

As short summary, so far the three quantification methods LFQ, DML and SILAC were applied.^{115,119,131} They were combined with different proteomic methods, namely full proteome analysis as well as a protein enrichment strategy. Furthermore, two different infection approaches were tested (Table 7). In approach A, both conditions were treated with 100 μ M **pro-N6pA**, but only one of them with the pathogenic bacterium. For approach B, both conditions were treated with bacterium but only one was pre-incubated with 100 μ M **pro-N6pA**. The other was pre-incubated with 0.1 vol% DMSO.

Table 7: Overview of performed experiments combining different infection strategies with different quantification methods. The type of the performed proteomic experiment is indicated (PE = protein enrichment; FP = full proteome analysis). “---” indicates that this combination was not applied.

infection approach \ quantification method	 approach A	 approach B
LFQ	PE	PE
DML	---	PE
SILAC	PE/FP	---

2.5.5. Protein enrichment experiments according to approach C combined with LFQ

So far, proteomic profiling combined with the established infection workflow allowed the detection of AMPylated host proteins upon bacterial infection. Strategies in which only half of the samples were infected required additional full proteomic analysis to account for changes in the human proteome caused by the infectious bacteria. In contrast, strategies that infected all samples but treated only half of them with **pro-N6pA** had the disadvantage of showing all AMPylated proteins enriched, regardless of the AMPylator responsible. A third strategy to profile AMPylated infected both conditions with pathogenic bacteria, as it is desirable to avoid performing a whole proteome analysis in addition to each enrichment experiment. Thereby, one condition was infected with the wild type bacterium, while the other condition was infected with a mutant version of the corresponding bacterial strain, expressing an inactive AMPylator. Furthermore, both conditions were treated with **pro-N6pA** as the absence of intrinsically AMPylated proteins enables a simpler evaluation of the enriched proteins (referred to as approach C, Figure 20 A). At first, this strategy was performed using *V. parahaemolyticus* wild type and its inactive mutant VopS-H348A. This method presented as most optimal, since it not only allowed the detection of the previously discovered AMPylation targets (RHOA, RAC1, CDC42, RHOC, and RHOG), but also identified RAC3 and RHOB as targets of VopS under physiological conditions (Figure 20 B).⁸⁴ Both of these two proteins have been listed by Yu *et al.* as potential VopS targets.⁸⁴ Altogether, seven different host proteins have been identified as AMPylation targets of VopS, all of them small Rho GTPases (Figure 20 C). The method further allowed to refute putative targets of VopS. The protein ERGIC2 was among the AMPylated proteins predicted by Yu *et al.*⁸⁴ Also Lewallen *et al.* observed several proteins as VopS targets, among them PFKP, NME1 and NAGK.⁸³ The current method did not confirm these four proteins as VopS targets under physiological conditions (Table 8). Since *in vitro* assays are not capable of reproducing all conditions prevailing in the cell, they are likely to lead to false positives. This might be the case for these proteins, as the corresponding profile plot clearly indicates no protein enrichment (Figure 20 D). An alternative explanation for the missing enrichment is incorporation problems, as **pro-N6pA** is slightly bulkier than AMP itself. Although VopS accepts **pro-N6pA** as substrate, (compare Figure 22 C) steric hindrance may occur in the target proteins due to the additional alkyne group. Furthermore, influencing the results by imputing missing values can also be ruled out.

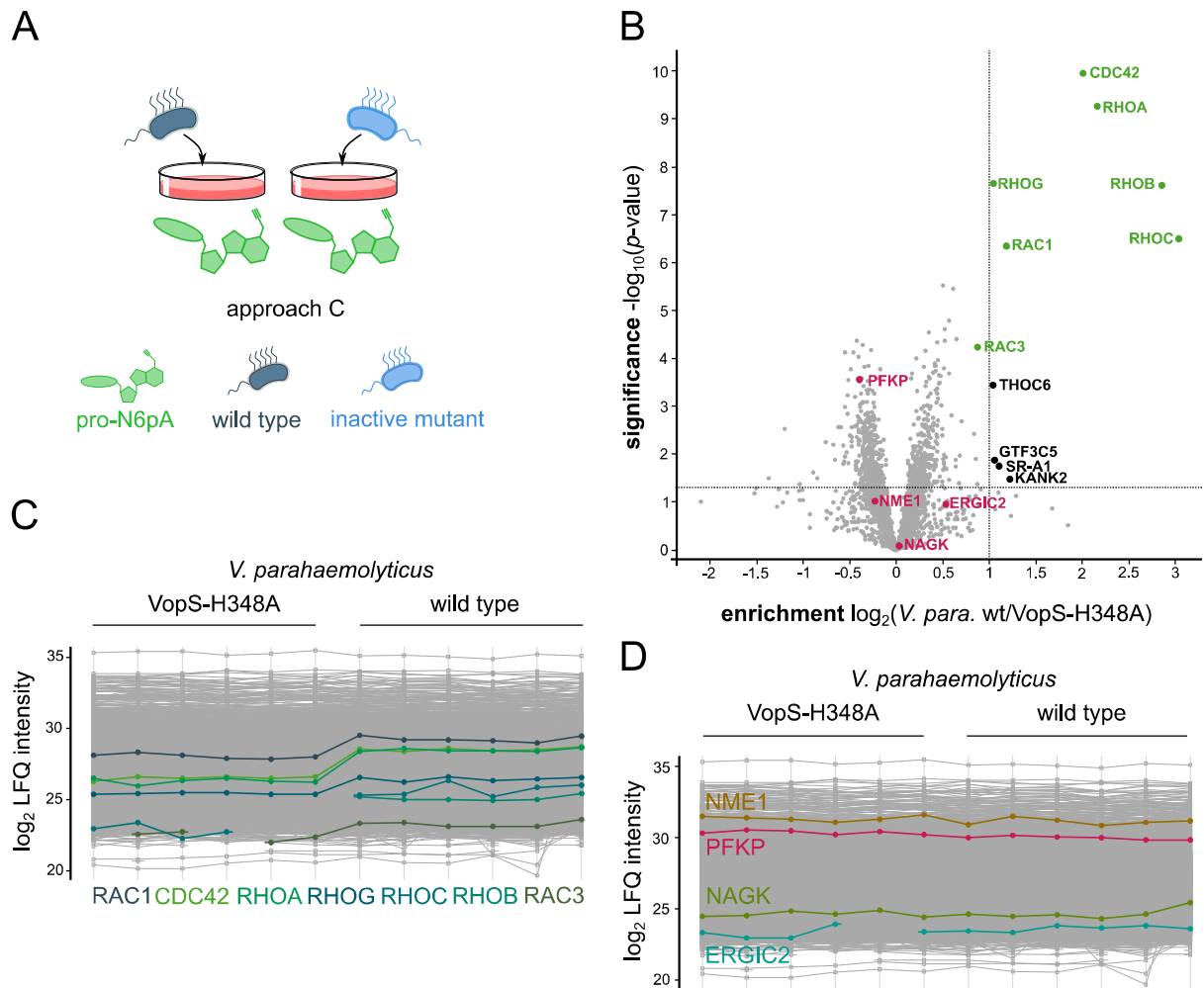


Figure 20: Applying wild type bacteria and mutants expressing inactive AMPylators yielded the largest number of VopS target proteins. (A) Graphic illustration of infection approach C. (B) Protein enrichment plot of HeLa cell infection with *V. parahaemolyticus* wild type or mutant VopS-H348A ($n = 6$; after data processing: 2491 proteins). All samples were treated with $100 \mu\text{M}$ pro-N6pA. Analysis was performed via LFQ. Cutoff lines were determined as \log_2 fold enrichment > 1 and a p -value < 0.05 . VopS protein targets are highlighted in light green, while putative new VopS targets are marked in black. Proteins which were previously predicted as VopS targets but could not be confirmed are shown in pink. (C) Profile plot of \log_2 intensities of VopS protein targets RAC1, CDC42, RHOA, RHOG, RHOC, RHOB and RAC3 (in descending order). (D) Profile plot of \log_2 intensities of previously identified VopS protein targets, which could not be confirmed in this experiment. Adapted from Rauh *et al.*⁸⁰

In addition to the assumed and confirmed VopS targets, four more proteins met the threshold line (Table 8), namely KANK2, SR-A1, GTF3C5 and THOC6. So far, they are not known to be related to any AMPylation event. However, all four profile plots did not represent convincing data, due to many missing and later-on imputed LFQ intensity values (data not shown).

Table 8: Enrichment data of all proteins that met the threshold criteria plus mistakenly predicted VopS targets in protein profiling experiments treating both conditions with probe, one condition with wild type bacteria and one with a mutant expressing an inactive AMPylator (approach C); green = confirmed VopS targets, pink = VopS targets that could not be confirmed, dark green = possible new VopS targets.

V. parahaemolyticus wild type vs. mutant VopS-H348A, protein enrichment, approach C, LFQ

gene name	log ₂ fold change	-log ₁₀ p-value	coverage [%]	significantly enriched?
RHOC	3.03	6.51	47.3	✓
RHOB	2.85	7.60	20.9	✓
RHOA	2.16	9.25	60.1	✓
CDC42	2.01	9.96	42.9	✓
RAC1	1.20	6.32	45.9	✓
RHOG	1.05	7.65	35.1	✓
RAC3	0.88	4.23	27.6	×
ERGIC2	0.54	0.97	13	×
NAGK	0.04	0.09	20.1	×
NME1	-0.22	1.02	74	×
PFKP	-0.40	3.52	45.9	×
KANK2	1.22	1.48	6.3	✓
SR-A1	1.11	1.79	6.9	✓
GTF3C5	1.07	1.91	13.7	✓
THOC6	1.03	3.43	20.5	✓

2.6. Human host AMPylation targets of other pathogenic bacteria

Since the optimized profiling strategy could theoretically be applied to any organism, further bacterial pathogens were screened for their ability to AMPylate human host proteins. Recent experiments with *V. parahaemolyticus* have shown that infection experiments comparing AMPylation induced by wild type bacteria with that of the inactive mutant, revealed more hits in the human proteome than, for example, the SILAC experiments. For this reason, previously used strains were tested again, even though they did not produce promising results in initial experiments. For this purpose, transposon mutants of the two putative Fic proteins of *P. aeruginosa* strain PAO1 (PW3486 genotype PA1366-A11::ISphoA /hah, referred to as mutant PA1366 as well as PW20 59 genotype PA0574-A11::ISlacZ/hah, referred to as mutant PA0574) were obtained from the University of Washington Manoil Lab PAO1 transposon mutant library.¹³⁶ Furthermore, deletion mutants of the putative Fic proteins of *E. coli* strain CFT073 (CFT073Δc4136::Km and CFT073Δc4409::Km) were constructed in cooperation with Sophie Brameyer (LMU).

Prior to the infection experiments, CFU assays for each mutant (PA1366, PA0574, CFT073 Δ c4136::Km and CFT073 Δ c4409::Km) were performed (Figures S3 G, H and Figures S4 D, E). At first, proteins enriched during *P. aeruginosa* strain PAO1 infection were compared to proteins enriched during infections with its mutants PA1366 and PA0574. As short reminder, mutant PA1366 looked most promising due to its missing inhibitory α -helix.⁴⁵ HeLa cells infected with one of the mutants at a time still showed the phenotypic cell rounding mentioned above for the wild type strain (Figure S3 B). This is an indication that the observed cell rounding by this pathogen, in contrast to *V. parahaemolyticus* infection, is not caused by the Fic proteins. This is consistent with the literature, since *P. aeruginosa* also possesses a T3SS that translocates, among others, the two effector proteins ExoS and ExoT.¹³⁷ These two effector proteins share 76 % sequence identity and have both GTPase activating protein (GAP) and ADP ribosyl transferase (ADPRT) activity. Like *V. parahaemolyticus*, they target small GTPases thus affecting the host cell cytoskeleton and resulting in cell rounding.¹³⁷ Infection experiments were again performed for 5 h. The volcano plots of these two infection experiments each showed not more than a handful of enriched proteins (Figure S3 I, K). However, none of these proteins represented convincing enrichment results regarding their profile plots. All of them showed similar LFQ intensity values in wild type as well as mutant treated samples. This fact combined with several missing LFQ intensity values led to protein enrichment due to intensive missing value imputation, although limits for protein exclusion have been set to the presence of at least 80% valid values in at least one condition. The enriched proteins also did not overlap with previous results obtained from SILAC labeling or the combination of approach B with LFQ. In addition, PSAT1, which appeared to be an actual enriched protein in previous experiments, could not be confirmed with the latest measurements as it had log₂ enrichment values around zero. For infection experiments with mutant PA0574 this again confirms recent literature, stating that this protein might have other functions than AMPylation.⁶⁵ Due to the lack of coherence between experiments and the absence of clear hits, infections with *P. aeruginosa* were not pursued any further.

Next, enriched proteins during *E. coli* strain CFT073 infection were compared to proteins enriched during infections with its deletion mutants CFT073 Δ c4136::Km and CFT073 Δ -c4409::Km. Again, there was no phenotypic difference between infection with the wild type strain or a deletion mutant (Figure S4 F, G). This suggests that neither gene c4136 nor c4409 is involved in this process. After infecting HeLa cells with either wild type or one of the mutants for 2.5 h with an MOI of 10, the resulting volcano plots revealed more significantly enriched proteins than during *P. aeruginosa* infection (Figure S3 I, K). However, similar to infections

with *P. aeruginosa*, there was no overlap with significantly enriched proteins from previous experiments. Furthermore, the corresponding profile plots again revealed consistent LFQ intensities for wild type and mutant treated samples with several missing values. As a consequence, the focus was further shifted toward infection experiments with *V. parahaemolyticus* in other human cell lines and target validation of the identified targets in HeLa cells.

2.7. Screening of other human cell lines for VopS-mediated AMPylation targets

Several of the small GTPases predicted as VopS targets have not been identified under physiological conditions yet.⁸⁴ One possible reason why these proteins did not appear in the latest *V. parahaemolyticus* infection experiment are their low expression levels in HeLa cells according to the ProteomicsDB (<https://www.proteomicsdb.org/>) and the Human Protein Atlas (<http://www.proteinatlas.org>).^{138,139} According to the same sources, human cell lines with higher expression levels of these proteins are HaCat and SH-SY5Y cells. HaCat cells are human skin/keratinocytes cells while SH-SY5Y cells are derived from human bone marrow with neuroblastoma.^{140,141} Notably, these cell types have no known association with *V. parahaemolyticus* infection to date.

Infecting SH-SY5Y cells with the optimized and fine-tuned infection strategy (approach C with the proteomic profiling workflow and LFQ analysis) and *V. parahaemolyticus* wild type versus its inactive mutant VopS-H348A yielded mixed results. As expected, also SH-SY5Y cells take on a round shape after successful infection (Figure 21 A1). However, it was surprising that only the previously identified Rho GTPases RHOA, RAC1, CDC42 and RHOG could be confirmed as VopS targets (Figure 21 B). Peptides of other Rho GTPases have not been detected during LC-MS/MS analysis. Interestingly, another protein showed strong enrichment (Table 9). Plexin-A1 (PLXNA1) is an integral membrane protein displaying GAP activity toward RAP1 GTPases, thus fitting into current knowledge about VopS targeting GTPases.¹⁴² RAP1 is a key player during neuronal development, and its inhibition impedes axon formation during cerebral cortex and hippocampal development in *in vivo* experiments.¹⁴² Moreover, also the profile plot of PLXNA1 strongly suggests this protein as true enrichment target. LFQ intensity values throughout the replicates are consistent with higher values for *V. parahaemolyticus* wild type treated samples. In addition, values for all replicates have been measured, eliminating influences due to missing value imputation.

Table 9: Enrichment data of PLXNA1 plus the VopS target proteins CDC42, RHOA, RAC1 and RHOG of the SH-SY5Y cell infection experiment with *V. parahaemolyticus* wild type and mutant VopS-H348A using approach C in combination with LFQ.

***V. parahaemolyticus* wild type vs. mutant VopS-H348A, protein enrichment, approach C, LFQ, SH-SY5Y**

gene name	log ₂ fold change	-log ₁₀ <i>p</i> -value	coverage [%]	significantly enriched?
PLXNA1	4.41	7.96	3.9	✓
CDC42	1.16	5.23	42.9	✓
RHOA	1.02	3.31	37.8	✓
RAC1	0.85	4.90	39.9	×
RHOG	0.83	3.24	23.6	×

While PLXNA1 has not been found during infections with the inactive mutant VopS-H348A, it was measured with LFQ intensities of about 26 in all samples infected with the *V. parahaemolyticus* wild type (Figure 21 D). However, re-evaluation of data produced by infecting HeLa cells with the latest strategy did not yield any results, as PLXNA1 was not found at all during these measurements.

As third human cell line, HaCat cells were infected with *V. parahaemolyticus* wild type and mutant VopS-H348A. Unexpectedly, this cell line did not show such pronounced cell rounding after 90 min infection with *V. parahaemolyticus* (Figure 21 A2). In the resulting volcano plot, only VopS target CDC42 has been significantly enriched (Figure 21 C, E) which might explain the limited cell rounding. In addition, the human protein SYNGR2 was also enriched. However, the available log₂ LFQ intensities suggest that this is due to imputed numbers with low values (data not shown). For all samples infected with *V. parahaemolyticus* wild type, LFQ intensities can be indicated for SYNGR2. However, corresponding peptides were only measured in two VopS-H48A treated samples, showing LFQ intensities in the same range as wild type treated samples. Thus, SYNGR2 enrichment is probably due to the matrix-based imputation of low LFQ values. Of the other known VopS targets, only RHOG approached the threshold to be considered significantly enriched. Since at least two unique peptides of RHOC, RHOA, RAC1, and PLXNA1 were identified during LC-MS/MS analysis, they are also shown in the plot (Table 10). However, all of them were far from being significantly enriched. As associated peptides were either identified in all samples or missing identifications mainly affected the negative control, this excluded a misalignment in the plot due to falsely imputed values (Figure 21 E). No peptides were measured for RHOB and RAC3.

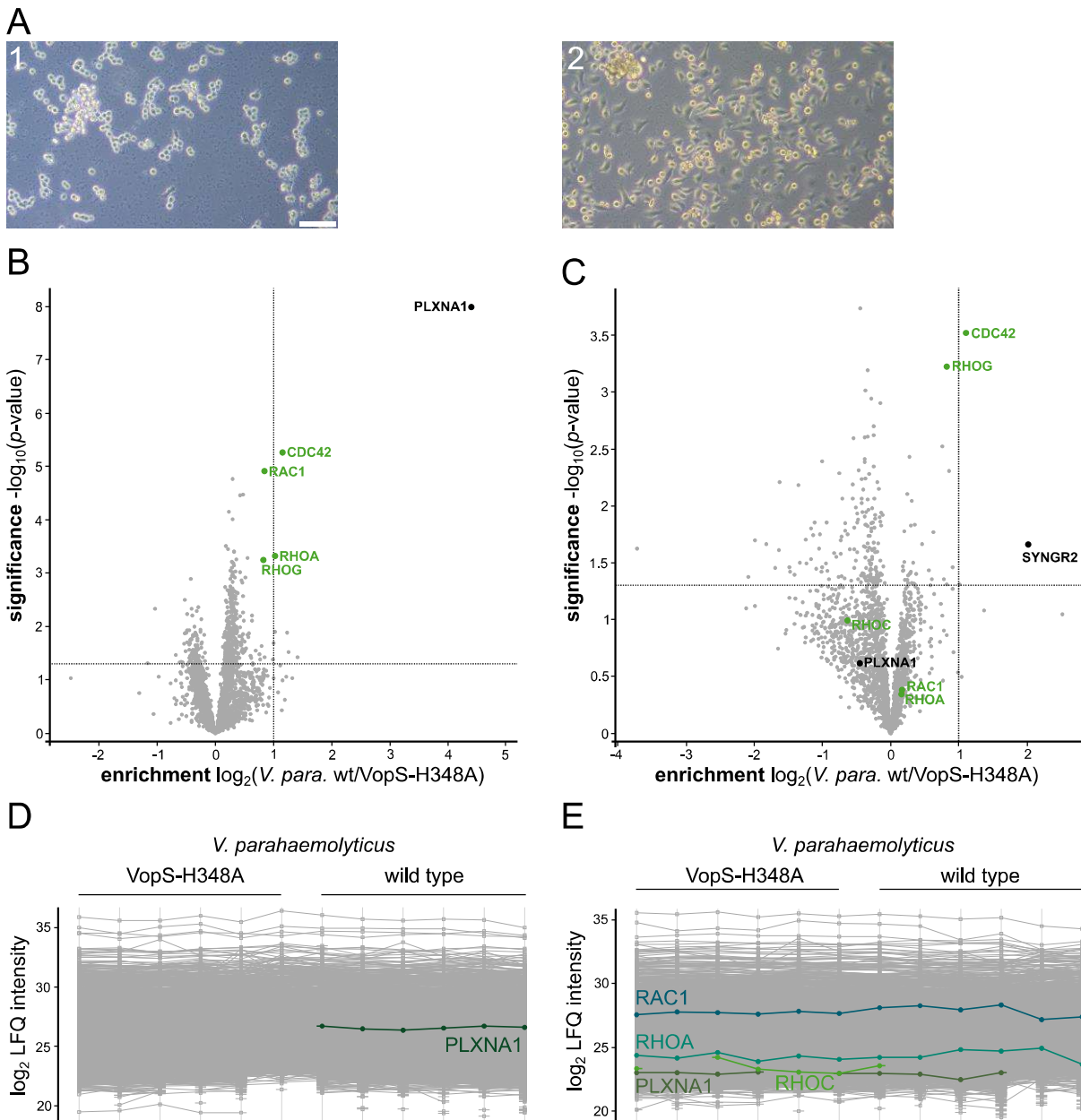


Figure 21: Results of infection experiments of the two human cell lines SH-SY5Y and HaCat infected with *V. parahaemolyticus* wild type and the mutant VopS-H348A using approach C combined with LFQ analysis. (A) Phenotypic appearances of SH-SY5Y cells (1, 84 min infection) or HaCat cells (2, 86 min infection) after infection with *V. parahaemolyticus* wild type. (B, C) Volcano plots of SH-SY5Y (B) or HaCat (C) cell infection with *V. parahaemolyticus* wild type or mutant VopS-H348A (B: n = 6; after data processing: 2135 proteins; C: n = 6, after data processing: 2265 proteins). All samples were pre-treated with 100 μ M **pro-N6pA** for 16 h. Analysis was performed *via* LFQ. Cutoff lines were determined as \log_2 fold enrichment > 1 and a p -value < 0.05. VopS protein targets are highlighted in light green, while putative new VopS targets are marked in black. (D, E) Profile plots of \log_2 LFQ intensities of PLXNA1 during SH-SY5Y cell infection (D) or of the VopS protein targets RAC1, RHOA, PLXNA1 and RHOC during HaCat cell infection (E). LFQ intensities of the respective proteins are highlighted in the indicated color.

Table 10: Enrichment data of PLXNA1 and the VopS target proteins CDC42, RHOG, RHOA, RAC1, and RHOC of the HaCat cell infection experiment with *V. parahaemolyticus* wild type and mutant VopS-H348A using approach C in combination with LFQ.

***V. parahaemolyticus* wild type vs. mutant VopS-H348A, protein enrichment, approach C, LFQ, HaCat**

gene name	log ₂ fold change	-log ₁₀ <i>p</i> -value	coverage [%]	significantly enriched?
CDC42	1.11	3.52	42.9	✓
RHOG	0.82	3.23	48.7	×
RHOA	0.18	0.36	43.5	×
RAC1	0.17	0.38	45.9	×
PLXNA1	-0.45	0.62	5	×
RHOC	-0.64	0.98	38.3	×

The lack of enrichment of Rho GTPases in HaCat cells after infection was surprising, since enrichment was observed for HeLa as well as SH-SY5Y cells. Possible explanations are tissue-specific or rather cell line-specific circumstances, as HaCat cells are the only non-cancerogenic cell line used in this study.^{126,140,141} Non-cancerogenic cells strongly deviate from cancerogenic cell lines concerning cell proliferation and cell differentiation, which might cause impaired **pro-N6pA** labeling capacities *e.g.* by slower conversion of **pro-N6pA** to **N6pATP**.¹⁴³ Moreover, HaCat cells are known to release ATP in the surrounding media, which might lead to lower **pro-N6pA** concentrations inside the cell.¹⁴⁴ Another possibility is, that HaCat cells themselves are difficult to access for VopS. Nevertheless, this is a completely new topic that will require further investigations.

2.8. Validation of Rho GTPases as VopS targets

So far, several Rho GTPases have been identified as AMPylation targets of VopS since they were enriched in previous proteome profiling experiments. To validate these proteins, other biochemical methods were used. To directly confirm AMPylation by VopS, recombinant RHOG was incubated with recombinant VopS missing the first 30 amino acids (VopS Δ 30, still active but with deleted signal sequence, kindly provided by the Itzen Lab)²⁹ in the presence of ATP or **N6pATP**. In parallel, the same approach was repeated with recombinant CDC42 (also kindly provided by the Itzen Lab) as positive control. As a negative control, the recombinant protein was incubated either without AMP source or without AMPylator. When using ATP as

AMP donor, successful protein modification was evaluated by intact-protein mass-spectrometry (IPMS). In case of **N6pATP**, the attachment of the AMP moiety could also be confirmed by clicking the sample to rhodamine-azide followed by a fluorescent SDS-PAGE. IPMS measurements of the recombinant proteins alone revealed their exact molecular mass with 21 003 Da for CDC42 and 25 247 Da for RHOG (Figure 22 A). Upon successful AMPylation, a mass increase of 329 Da is expected in case of ATP, and 367 Da in case of **N6pATP** (Figure 22 B). IPMS measurements of both proteins, CDC42 and RHOG, confirmed them as *in vitro* VopS substrates and the measurements showed that ATP represents a suitable substrate for the AMP transfer by VopS. Labeling of RHOG and VopS by **N6pATP** was also observed by in gel fluorescence after click reaction of the probe alkyne to rhodamine-azide (Figure 22 C). This is consistent with previous findings of VopS auto-AMPylation and that small chemical handles at the N6 position of the adenine are tolerated.^{44,92}

In order to identify the site of AMPylation, the *in vitro* AMPylated recombinant RHOG was subjected to reduction with DTT, followed by alkylation with IAA and overnight trypsin digestion. The resulting peptides were evaluated by LC-MS/MS analysis on a Q Exactive instrument. Although this mass spectrometer is not able to fragment peptides *via* ETD fragmentation, which helps to prevent loss of different chemical groups of the AMP moiety, modified peptides were nevertheless detected due to their high abundance [close to 100% conversion as observed during IPMS measurements (Figure 22 A)]. Final data analysis with MaxQuant was performed by searching for the attachment of an AMP moiety to Ser, Tyr or Thr side chains. The results revealed Thr 35 as the RHOG AMPylation site, as predicted by Yu *et al.*⁸⁴

To further investigate the AMPylation site of RHOG under physiological conditions, the established infection workflow was exploited. After successful HeLa cell treatment with **pro-N6pA** and infection with *V. parahaemolyticus*, the treated proteome was clicked to an azide-TEV-cleavable-biotin linker.¹⁴⁵ This way, modified proteins were enriched due to the biotin avidin interaction and the unbound proteins were washed out. Next, enriched proteins were digested, and unmodified peptides were removed. Finally, modified peptides bound to the beads were released by treatment with AcTEV protease. As there was only a limited amount of modified peptide this time, MS measurements were performed on the Orbitrap Fusion. HCD fragmentation was used for most of the measurement as previously reported for AMPylated proteins.⁵⁷

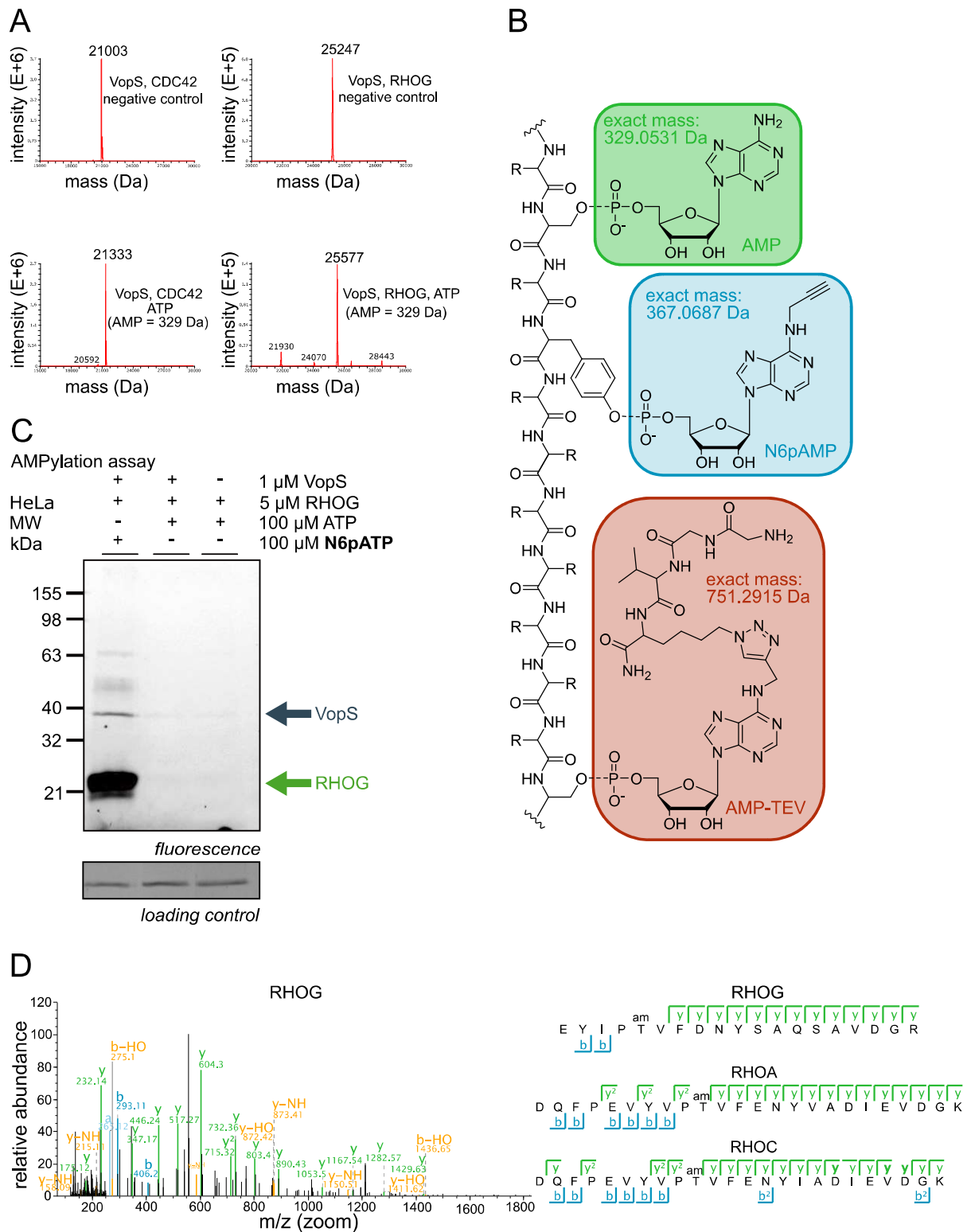


Figure 22: Confirmation of RHOG as a target of VopS and investigation of VopS AMPylation activity toward RHOC and RHOG *in vitro* and *in situ*. **(A)** IPMS measurements without ETD fragmentation of AMPylated as well as unmodified recombinant CDC42 and RHOG. Both proteins have one AMPylation site. **(B)** Expected mass increase once an AMP (329.0531 Da in green) or a N6pAMP moiety (367,0687 Da in blue) is attached to a hydroxyl group containing amino acid. The structure of an AMP moiety clicked to the remaining TEV sequence after cleavage is shown in red along with the resulting mass increase (751,2915 Da). **(C)** Gel-based analysis of recombinant RHOG AMPylated by VopS *in vitro* confirms **N6pATP** as suitable substrate for the AMPylation reaction. **(D)** *V. parahaemolyticus* wild type infection of HeLa cells revealed that the

VopS mediated AMPylation of human Rho GTPases occurred on a conserved Thr residue (RHOG: Thr35; RHOA: Thr37; RHOC: Thr37). Partially adapted from Rauh *et al.*⁸⁰

The instrument was constantly searching for specific reporter ions (m/z 558.3, 279.65, or 186.77 – adenine-TEV, single, double, or triple charged; m/z of 672.33, 336.67, or 224.78 – adenosine-TEV, single, double, or triple charged; or m/z of 770.31, 385.66, or 257.44 – phosphoadenosine-TEV, single, double, or triple charged). Once one of them has been identified, the fragmentation method was switched to ETD. Thus, the modified peptide could be fragmented with the AMP modification still in place, allowing determination of the modification site. It is not only because of a low proportion of endogenously modified proteins, that the task of identifying AMPylation sites remains challenging, despite this optimized MS method. Certainly, not all modified proteins are equipped with an alkyne tag, since **pro-N6pA** competes with endogenously available ATP.⁵⁷ Nevertheless, by analyzing the obtained raw data and searching for a modification consisting of AMP clicked to the TEV linker (Figure 22 B), the AMPylation sites of RHOA, RHOG and RHOC could be identified (Figure 22 D). As expected, the RHOA AMPylation site was assigned to Thr 37.²⁹ The AMPylation sites for RHOG and RHOC were assigned to Thr 35 and Thr 37, respectively, as predicted by Yu *et al.*⁸⁴

3. Conclusion and outlook

3. Conclusion and outlook

AMPylation as post-translational modification was discovered more than 50 years ago, but its impact has been recognized only in recent years. Therefore, we are only at the beginning of recognizing the scope of this modification, identifying modified proteins as well as investigating consequences of it in more detail. The discovery that pathogens, among them *Vibrio parahaemolyticus*, mediate AMPylation of human host proteins by their effector proteins was an important step. Research in this area is becoming increasingly important as the number of multidrug-resistant bacteria increases. Therefore, new strategies are needed to combat these bacteria as well as ways, to deprive their pathogenicity.

In vitro experiments may not succeed in correctly mimicking all the influences present in the live cell. For this reason, their results may deviate from actual circumstances in that not all AMPylation targets are identified or false positives are included. In this thesis, a method to identify AMPylation targets of VopS in the human proteome under physiological conditions was successfully developed. To do so, the cell permeability of the previously introduced pronucleotide probe **pro-N6pA** was exploited.⁵⁷ In combination with the sequenced *Vibrio parahaemolyticus* strain RIMD 2210633 and the well-characterized human HeLa cell line, different infection workflows, followed by different proteomic quantification methods, were evaluated for their potential to identify targets of VopS.^{76,126}

The initial approach applied LFQ to samples that were all treated with **pro-N6pA**, but only half of them infected with *V. parahaemolyticus* wild type. Although the approach managed to confirm RHOA and CDC42 as VopS target, and its functionality was further confirmed with the *V. parahaemolyticus* mutant VopS-H348A, it did not prove suitable. The approach did not consider possible influences of the infecting bacteria on the human proteome. The second approach explores quantification *via* dimethyl labeling of sample that were all infected with *V. parahaemolyticus* but only half of them treated with **pro-N6pA**. This way, undetected changes to the human proteome were avoided. However, as this quantification method is based on the comparison of differently labeled peptides, it failed to quantify most of them due to missing protein enrichment for DMSO treated samples. As a third approach, SILAC labeling was combined with complete treatment of the samples with **pro-N6pA** and infection of only half of them. To include possible changes on the proteome level this time, a full proteome analysis was performed in parallel to the enrichment experiments. This approach confirmed RHOA, RAC1 and CDC42 as VopS targets while showing that their enrichment is not caused

by cellular upregulation. When infecting HeLa cells with *P. aeruginosa* PAO1 and *S. aureus* USA 300, this approach did not yield any new results. It only suggested that Gram-positive bacteria might not AMPylate host proteins, as *S. aureus* infection hardly reveal any enriched proteins. Strategy four combined LFQ quantification with samples that were all infected with *V. parahaemolyticus* but only half of them treated with **pro-N6pA**. This approach turned out to be promising, as it not only identified the already confirmed VopS targets RHOA, RAC1 and CDC42, but furthermore also revealed RHOG and RHOC as VopS targets under physiological conditions. The approach was once more validated using the inactive *V. parahaemolyticus* mutant VopS-H348A, which caused significant less enrichment of the VopS target proteins. Moreover, applying this approach to HeLa cell infection with *P. aeruginosa* PAO1 or *E. coli* CFT073 revealed several potential AMPylation targets.

The final approach combined all the previously identified advantages. It used LFQ as quantification method and treated all samples with **pro-N6pA** to increase protein enrichment. In addition, all samples were infected with *V. parahaemolyticus*, either with the wild type or the inactive VopS-H348A mutant, to exclude bacterial induced changes on the proteome level. In addition to the already identified VopS targets under physiological conditions, this method was also able to detect RHOB and RAC3. When performing this strategy with *P. aeruginosa* PAO1 and *E. coli* CFT073 wild type infection and the corresponding inactive mutants, a lack of consistency throughout the proteomic data prevented the identification of novel AMPylation targets.

To check if some VopS targets remained undetected in previous experiments due to a possibly low expression level in HeLa cells, the human HaCat and SH-SY5Y cell lines were subjected to this final approach. These infection experiments indicated that not all VopS targets have been identified to date, as they show a tissue- or cell line-specific enrichment pattern that differs from that in HeLa cells. This is a topic that needs further attention in the future.

To finally validate the newly identified VopS targets, and thus also validating the established workflow, *in vitro* and *in situ* AMPylation assays as well as site-ID experiments were performed. IPMS measurements confirmed the *in vitro* AMPylation of RHOG by VopS, while the AMPylation sites of RHOA (Thr37), RHOG (Thr35) and RHOC (Thr37) were mapped *in situ* by clicking the infected and probe treated proteome to a TEV-cleavable linker.

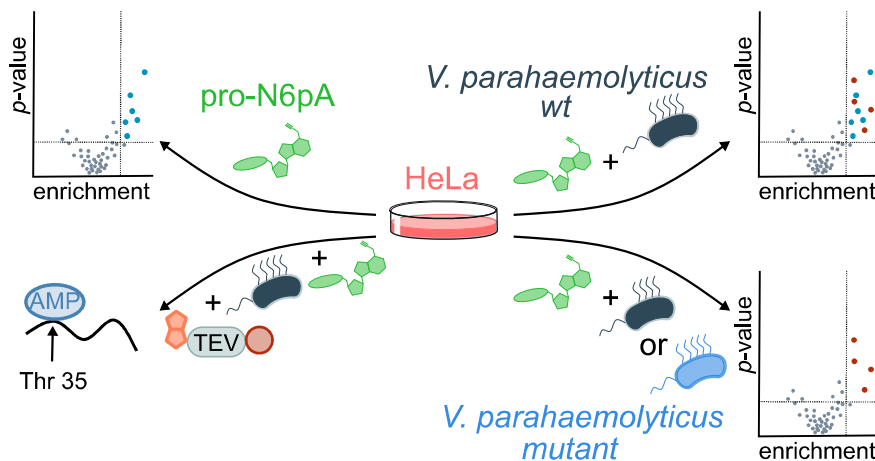


Figure 23: Summary and conclusion of the project. **Pro-N6pA** was applied to map intrinsically AMPylated human proteins and to establish an infection strategy revealing *in situ* host protein AMPylation targets of the *V. parahaemolyticus* effector protein VopS. As best infection strategy proved the treatment of all cell samples with **pro-N6pA** (100 μ M) and their infection with either the wild type bacterial strain or an mutant strain expressing an inactive AMPylator. This way, several Rho GTPases were confirmed as *in situ* AMPylation targets of VopS, and their AMPylation site was identified. Adapted from Rauh *et al.*⁸⁰

To date, it is known that VopS belongs to a larger group of effector proteins, which exploit AMPylation during the infection process. It is necessary to identify more proteins that proceed with such a mechanism and to elucidate their AMPylation targets under physiological conditions. With the established approach in hand, this goal is within reach. Furthermore, by examining several cell lines of different origin, a holistic picture of possible targets can be achieved. What complicates the whole matter, however, are bacterial proteins that contain a Fic domain but do not catalyze AMPylation. By replacing **pro-N6pA** with other pronucleotide probes, the established workflow can be adapted to identify protein targets of other post-translational modifications introduced during pathogenic infections.

4. Experimental section

4. Experimental section

For all proteomic experiments, if not stated otherwise, MS-grade solvents were used. All given percentages represent percentage by volume (v/v) for liquid substances and percentage by weight (w/v) for all solids. The majority of the experimental procedures were adapted from Rauh *et al.*⁸⁰ (<https://pubs.acs.org/doi/10.1021/acsinfecdis.0c00740>)

4.1. Buffers

Table 11: Buffers used in this thesis.

buffer	ingredients
AMPylation buffer (10x)	200 mM Hepes pH 7.5 1 M NaCl 50 mM MgCl ₂ 1 mg/mL BSA 10 mM DTT
lysis buffer	PBS pH 7.4 1 % NP40 1 % sodium deoxycholate 1 tablet protease inhibitor (cOmplete, Mini, EDTA-free protease inhibitor cocktail, Roche)
peptide elution buffer	80% acetonitrile 0.5% formic acid in LC-MS grade H ₂ O
protein digestion buffer	3.9 M urea 1.1 M thiourea 20 mM Hepes, pH 7.5
SDS loading buffer (2x)	63 mM Tris-HCL 10% glycerol 139 mM sodium dodecyl sulfate (SDS) 0.0025% bromophenol blue 5% 2-mercaptoethanol

4.2. Cell culture

Human epithelioid cervix carcinoma HeLa cells were purchased from Sigma-Aldrich (93021013_1VL). Human epithelioid and keratinocyte HaCat cells were obtained from the group of Markus Gerhardt at the Technical University of Munich and human epithelioid neuroblastoma SH-SY5Y cells were purchased from Sigma-Aldrich (94030304-1VL).

4.2.1. General cell culture procedures

Cells were maintained in the corresponding media (Table 12) at 37 °C in a humidified 5% CO₂ atmosphere. Cell culture work was performed in a laminar flow bench (Herasafe KS, Thermo Scientific, USA) equipped with a vacuum pump (BVC21, Vacuubrand, Germany) to remove media. Visual cell inspection was performed with a PrimoVert microscope (Zeiss, Germany), while cells were counted using a Neubauer improved chamber. The cells were routinely tested for mycoplasma contamination.

For SILAC experiments, HeLa cells were passaged at least six times in SILAC-DMEM (Sigma-Aldrich) supplemented with 10% dialyzed FBS and 2 mM L-glutamine as well as 214 µM [13C6, 15N4] L-arginine HCl (Arg10) and 419 µM [13C6, 15N2] L-lysine 2 HCl (Lys8) (Cambridge Isotope Laboratories) resulting in “heavy” cells or with 214 µM [13C6] L-arginine HCl (Arg6) and 419 µM [4,4,5,5-D4] L-lysine 2 HCl (Lys4)(Cambridge Isotope Laboratories) resulting in “light” cells instead.

Table 12: Information about the used human cell lines.

cell line	species	cell type	media for cultivation
HeLa	homo sapiens	cervix / adenocarcinoma	DMEM high glucose supplemented with 10% FBS and 2 mM L-glutamine
HaCat	homo sapiens	skin / keratinocytes	calcium-free DMEM high glucose supplemented with 10% FBS and 2 mM L-glutamine
SH-SY5Y	homo sapiens	bone marrow / neuroblastoma	DMEM / F-12 supplemented with 10% FBS, 2 mM L-glutamine and 1% MEM non-essential amino acids solution (100x)

4.2.2. Thawing and cryopreservation of cells

A water bath (MPC, Huber, Germany) at 37 °C was used to thaw cells as fast as possible which have been stored in cryotubes in a liquid nitrogen tank (XSS-36/6, VWR, Belgium). Subsequently, 9 mL of the corresponding prewarmed media were added and the cell pellet was resuspended. To remove remaining DMSO, cells were centrifuged at 600g for 5 min at r.t. (Centrifuge 5810, Eppendorf, Germany) and the supernatant was removed. The residual pellet

was resuspended in 10 mL of prewarmed medium, transferred to a T-75 flask and further incubated at 37 °C.

To preserve cells, they were pelletized at 600g for 5 min at r.t. and the supernatant was discarded. The cell pellet was resuspended in 2 mL of the respective cryo medium (70% DMEM, 20% FBS, 10% DMSO) before transferring it to cryo vials. This suspension was then frozen using an isopropanol chamber (Mr. Frosty, Thermo Scientific, USA) at -80 °C (Forma 900, Thermo Scientific) for one day. Afterward, vials were stored in a liquid nitrogen tank.

4.2.3. Cellular maintenance

To avoid growth restrictions, cells were regularly passaged. Therefore, they were washed with prewarmed PBS to remove remaining media and detached with 1 mL (T-75) or 2 mL (T-175) Accutase solution (Sigma-Aldrich). Digestion was performed for 5 to 10 min at 37 °C and traced *via* the microscope. After complete detachment, 4 mL (T-75) or 9 mL (T-175) of the corresponding medium were added to the cells. Depending on the previous cell density, the suspension was diluted 1:10 to 1:4 with prewarmed medium and cells were further cultivated at 37 °C.

4.3. Bacterial growth

4.3.1. Strains and media

The *Vibrio parahaemolyticus* strain RIMD 2210633 was a kind gift from Dr. Tetsuya Iida and Dr. Takeshi Honda (Research Institute for Microbial Diseases, Osaka University). The strain was routinely cultured in lysogeny broth (LB) medium (10 g/L casein peptone, 5 g/L NaCl, 5 g/L yeast extract, pH 7.5) + 3% NaCl at 30 °C with agitation at 200 rpm. In cooperation with Sophie Brameyer, the *V. parahaemolyticus* RIMD 2210633 mutant H348A in VopS was obtained by double homologous recombination using the suicide plasmid pNPTS138-R6KT-VopS-H348A. The procedure in more detail is documented in Rauh *et al.*⁸⁰

The strain *Pseudomonas aeruginosa* PAO1 was obtained from the Institute Pasteur in France and *Escherichia coli* CFT073 was obtained from the Urological Clinic in Munich (Dr. Giuseppe Magistro). The *P. aeruginosa* transposon mutants PW3486 genotype PA1366-A11::ISphoA/hah as well as PW2059 genotype PA0574-A11::ISlacZ/hah were obtained from the

University of Washington Manoil Lab PAO1 transposon mutant library.¹³⁶ Transposon mutants were plated onto pure LB agar plates from glycerol stock dilutions and grown for 24 h at 37 °C. Subsequently, colonies were picked for each transposon mutant and grown in overnight cultures containing 20 µg/mL tetracycline. *Pseudomonas aeruginosa* PAO1 wild type was routinely cultured in LB medium, whereas the transposon mutants were cultivated in LB medium containing 5 µg/mL tetracycline.

The *E. coli* in frame deletion mutants c4136::Km and c4409::Km were constructed by Red/ET recombination in cooperation with Sophie Brameyer as described by Rauh *et al.*⁸⁰ The *E. coli* CFT073 deletion mutants c4136::Km and c4409::Km were grown in LB medium supplemented with 50 µg/mL kanamycin, whereas the *E. coli* wild type was cultured in pure LB medium. All of them were grown at 37 °C with agitation at 200 rpm.

S. aureus strain USA300 was purchased from the American Type Culture Collection (ATCC BAA-1556) and routinely cultured in pure B medium (10 g/L casein peptone, 5 g/L NaCl, 5 g/L yeast extract, 1 g/L K₂HPO₄, pH 7.5).

Table 13: Summary of bacterial strains used in this thesis including their culture conditions and their sources.

bacterial strain	growth condition	source
<i>Vibrio parahaemolyticus</i> strain RIMD 2210633 wild type	30 °C, 200 rpm, LB-medium containing 3% NaCl	Dr. Tetsuya Iida and Dr. Takeshi Honda (Research Institute for Microbial Diseases, Osaka University)
<i>Vibrio parahaemolyticus</i> strain RIMD 2210633 VopS-H348A	30 °C, 200 rpm, LB-medium containing 3% NaCl	constructed from the wild type strain in cooperation with Sophie Brameyer
<i>Pseudomonas aeruginosa</i> PAO1 wild type	37 °C, 200 rpm, LB-medium	Institute Pasteur in France
<i>Pseudomonas aeruginosa</i> PW2059 genotype PA0574-A11::ISlacZ/hah	37 °C, 200 rpm, LB-medium plus 5 µg/mL tetracycline	University of Washington Manoil Lab PAO1 transposon mutant library
<i>Pseudomonas aeruginosa</i> PW3486 genotype PA1366-A11::ISphoA/hah	37 °C, 200 rpm, LB-medium plus 5 µg/mL tetracycline	University of Washington Manoil Lab PAO1 transposon mutant library
<i>Escherichia coli</i> CFT073 wild type	37 °C, 200 rpm, LB-medium	Urological Clinic in Munich (Dr. Giuseppe Magistro)

<i>Escherichia coli</i> CFT073 c4136::Km	37 °C, 200 rpm, LB-medium plus 50 µg/mL kanamycin	constructed from the wild type strain in cooperation with Sophie Brameyer
<i>Escherichia coli</i> CFT073 c4409::Km	37 °C, 200 rpm, LB-medium plus 50 µg/mL kanamycin	constructed from the wild type strain in cooperation with Sophie Brameyer
<i>Staphylococcus aureus</i> USA300 wild type	37 °C, 200 rpm, B-medium	ATCC BAA-1556

4.3.2. Bacterial maintenance

For overnight cultures, 5 mL of the appropriate medium were added to a sterile culture tube and inoculated with 5 µL of the respective glycerol stock. Solutions were supplemented in case of the *P. aeruginosa* transposon mutants with 20 µg/mL tetracycline and for the *E. coli* deletion strains with 50 µg/mL kanamycin. The culture was then grown at the appropriate temperature (*V. parahaemolyticus* 30 °C, all other strains 37 °C) with 200 rpm for 16 h. To exclude undesired contaminations, a sterile control (medium only) was always included.

For glycerol stock preparation, 2 mL of overnight culture were pelletized (6 min, 6000g, 4 °C). The supernatant was removed, and the bacterial pellet resuspended in 500 µL of the corresponding medium (including antibiotics) and 500 µL glycerol. The bacterial solutions were further aliquoted and frozen at -80 °C.

For day cultures, sterile culture flasks (100 mL or 250 mL, with or without baffles) were filled to 20% with the corresponding medium. Antibiotics were added if necessary, and the culture was inoculated 1:100 with overnight culture.

4.4. Proteomic labeling strategies

4.4.1. Probe treatment

Labeling of cells was performed as previously described.⁵⁷ Briefly, HeLa cells were either seeded in 6 cm dishes, 10 cm dishes, or 15 cm dishes and grown to 80–90% confluency. Then, the old culture medium was removed and new medium containing 100 µM **pro-N6pA** (stock 100 mM in DMSO) or 0.1% DMSO were added. Cell dishes were statically incubated at 37 °C and 5% CO₂ for 16 h before further treatment.

4.4.2. Analytical *in situ* labeling

Without bacterial infection

HeLa cells were seeded in 6 cm dishes and treated with **pro-N6pA** (stock solution 100 mM in DMSO) or DMSO with various concentrations and for different periods of time. Labeling was performed as previously mentioned (4.4.1. probe treatment). Then, medium containing probe or DMSO was removed, and cells were washed twice with ice-cold PBS (2 x 1 mL). Next, cells were lysed by adding 150 μ L lysis buffer and scrapping the cells off the plate. The lysis was performed at 4 °C for at least 15 min while rotating the samples. The insoluble fraction has been separated from the soluble fraction (15 min, 4 °C, 13 000 rpm) before the protein concentration of the latter has been determined by BCA assay (Roti Quant, Roth). As next step, protein concentration was adjusted to equal protein amounts and samples were filled up to a total volume of 200 μ L using 0.2% SDS in PBS. Click reactions were performed as described previously¹⁴⁶ for 1.5 h at 25 °C and shaking at 450 rpm using 0.096 mM rhodamine-azide (10 mM stock in DMSO, Rh-N3, base click), 0.96 mM TCEP (100 mM stock in ddH₂O, Roth), 0.096 mM TBTA ligand (83.5 mM stock in DMSO, TCI), and 0.96 mM CuSO₄ (50 mM stock in ddH₂O). Finally, the proteins were precipitated by adding 1 mL acetone and storing the samples in upright position at -20 °C overnight. As next step, the proteins were harvested by centrifugation (15 min, 4 °C, 13 000 rpm) and the pelletized proteins were resuspended in 100 μ L 2x SDS loading buffer (63 mM Tris-HCl, 10% glycerol, 139 mM sodium dodecyl sulfate (SDS), 0.0025% bromophenol blue, 5% 2-mercaptoethanol) by sonication (10% intensity, 10 s) before heating the sample for 5 min at 95 °C. Each sample was thoroughly mixed prior to applying 40 μ L per gel-lane on an SDS-PAGE gel (12 pockets, 12.5% acrylamide). A Fujifilm LAS 4000 luminescent image analyzer equipped with a Fujinon VRF43LMD3 lens and a 575DF20 filter (both Fujifilm) were used to record the fluorescence image of the gel.

Including bacterial infection

For each analytical infection assay, HeLa cells were seeded in as much 6 cm dishes as needed for the planned number of samples plus two. All dishes were treated with **pro-N6pA** as described previously (4.4.1. probe treatment). In parallel to the probe treatment, an overnight culture of the desired bacterial strain and the respective medium was inoculated. The next morning, a fresh day culture of the desired bacterial strain was inoculated under the same

conditions used to determine the CFU to OD₆₀₀ correlation. Next, the number of HeLa cells in the two additionally seeded dishes was counted. Therefore, cells were washed with 1 mL PBS, detached with 0.6 µL Accutase solution (Sigma-Aldrich), taken up in 1.4 mL DMEM and diluted 1:1 with 0.5% trypan blue solution. Subsequently, cells were counted using a Neubauer improved cell counting chamber. As soon as the HeLa cells, which were seeded for the infection assay, were exposed to probe treatment for 16 h, the OD₆₀₀ of the previously inoculated bacterial day culture (already grown for at least 2.5 h) was determined. Having the number of HeLa cells per dish as well as the OD₆₀₀ value of the bacterial culture in hand, the required volume of bacterial culture needed for a certain multiplicity of infection was calculated and harvested (10 min, 4 °C, 6000g). The bacterial pellet was resuspended in DMEM solely supplemented with 2 mM L-glutamine and 10 µM **pro-N6pA**. In the meantime, HeLa cells were washed with 1 mL PBS. After that, the infection was initiated by adding 2 mL of DMEM containing probe and bacteria with the calculated MOI. Infected cells were incubated under static conditions at 37 °C and 5% CO₂. At various time points, images of each dish were acquired at 10x magnification using the Zeiss microscope Primovert equipped with a Zeiss AxioCam ERc 5s. Then, cells were scraped off and pelletized for 10 min at 4 °C and 750g. The cells were washed with 1 mL PBS and resuspended in 150 µL lysis buffer. Further sample processing, containing click reaction, protein precipitation and in-gel analysis were performed analogously to analytical *in situ* labeling without bacterial infection.

4.4.3. Preparative *in situ* labeling

Label-free quantification (LFQ) without bacterial infection

As described above, cells (6 cm dishes) were treated with **pro-N6pA** probe and DMSO for 16 h. Afterward, cells were washed twice with 1.5 mL ice-cold PBS, while scratching them off during the second washing step. Next, intact cells were pelletized (15 min, 4 °C, 750g) and resuspended in 150 µL ice-cold lysis buffer. The lysis was performed at 4 °C for at least 15 min while rotating the samples. The insoluble fraction has been separated from the soluble fraction (15 min, 4 °C, 13 000 rpm) before determining the protein concentration of the latter by BCA assay (Roti Quant, Roth). As next step, protein concentration was adjusted to equal protein amounts and samples were filled up to a total volume of 200 µL with PBS. Click chemistry was performed as described in analytical *in situ* labeling using 0.096 mM azide-PEG3-biotin conjugate instead of rhodamine-azide. Finally, the proteins were precipitated by adding at least 4-fold volume excess of acetone and storing the samples at -20 °C overnight. As next step, the

proteins were harvested by centrifugation (15 min, 4 °C, 13 000 rpm) and washed twice with ice-cold MeOH. Therefore, 1 mL of MeOH was added to the protein pellet which was resuspended by sonication (10% intensity, 10 s, Sonopuls HD 2070 ultrasonic rod, Bandelin electronic GmbH) and harvested again as described previously followed by the enrichment procedure. Prior to the enrichment, 50 µL of avidin-agarose beads per sample were washed trice with 1 mL 0.2% SDS in PBS. The protein pellets were resuspended in 0.2% SDS in PBS by sonication (10% intensity, 10 s) before separating the remaining insoluble part by centrifugation (5 min, 25 °C, max. speed). After that, the supernatant was transferred to the avidin-agarose beads and incubated under rotation for 1.5 h at 25 °C. Subsequently, the beads were washed with 0.2% SDS in PBS trice, twice with 6 M urea solution and finally three times with PBS (after each washing step beads have been collected by centrifugation for 3 min and 400g). The washed beads were resuspended in 200 µL digestion buffer (3.9 M urea, 1.1 M thiourea in 20 mM Hepes, pH 7.5) and the enriched proteins were first reduced (1 mM DTT, 45 min, 25 °C) and second alkylated (5.5 mM IAA, 30 min, 25 °C). Last, the alkylation reaction was quenched by adding 4 mM DTT (30 min, 25 °C). Proteins were predigested with 1 µL LysC (0.5 mg/mL, Wako) for 2 h at 25 °C, diluted in 600 µL 50 mM triethylammonium bicarbonate buffer (TEAB, 1 M) and digested with 1.5 µL trypsin (0.5 mg/mL, Promega) for 15 h at 37 °C. Then, trypsin digestion was stopped by adding 1% formic acid (FA) and in the following, peptides were desalted with 50 mg Sep-Pak C18 cartridges (Waters Corp.). Therefore, the C18 material of each column was equilibrated with 1 mL acetonitrile, 1 mL elution buffer (80% acetonitrile, 0.5% FA in H₂O) and 3 mL wash buffer (0.5% FA in H₂O). As soon as the peptides were loaded on the cartridges, they were washed with 3 mL wash buffer and eluted (2x 250 µL elution buffer). Afterward, peptides were freeze-dried using a speedvac centrifuge and reconstituted in 30 µL 1% FA. Finally, they were sonicated for 10 min, filtered through 0.22 µm Ultrafree-MC centrifugal filters (Merck, UFC30GVNB) and stored at -20 °C until MS/MS measurement. All enrichment experiments were measured on the Q Exactive instrument.

Label-free quantification (LFQ) including bacterial infection

For the label-free quantification-based infection assays, HeLa cells were seeded either in 6 cm dishes (infection with *V. parahaemolyticus* or *S. aureus*) or in 10 cm dishes (infection with *V. parahaemolyticus* wild type vs VopS-H348A, infections with *P. aeruginosa* or *E. coli*). The further procedure was performed the same way as described for analytical *in situ* labeling including bacterial infection with some changes. In brief, the two dishes thought for cell

counting were treated with **pro-N6pA**, whereas the other dishes were either treated with 100 μM **pro-N6pA** or 0.1% DMSO, depending on the experimental setup. An overnight culture of the bacterial strain was inoculated, with which a fresh day culture was started the next day. The two additional dishes of HeLa cells were counted (for 10 cm dish: washed with 2 mL PBS, added 1 mL Accutase, resuspended in 3 mL DMEM, diluted 1:1 with trypan blue), the bacterial OD_{600} value was measured and the respective volume for the desired MOI was calculated. The bacteria were harvested and resuspended in DMEM (plus 2 mM L-glutamine) and either 10 μM **pro-N6pA** or 0.01% DMSO. To each 10 cm dish, 7 mL of this suspension were added (2 mL to 6 cm dish) before further incubation at 37 °C and 5% CO_2 for a previously defined period. Once the desired infection time was reached, images of each dish were taken and the human cells were scraped off, washed in PBS, and lysed in 250 μL lysis buffer. Further sample handling was performed according to the previously described preparative *in situ* labeling procedure using LFQ analysis. Equal protein amounts were adjusted prior to performing the click reaction with azide-PEG3-biotin conjugate. The proteins were precipitated overnight, pelletized and washed twice with ice-cold MeOH. The pellet was resuspended in 1 mL 0.2% SDS in PBS and added to 50 μL prewashed (3x 1 mL 0.2% SDS in PBS) avidin-agarose beads. After incubation for 1.5 h at 25 °C, the beads were washed (3x 1 mL 0.2% SDS in PBS, 2x 1 mL 6 M urea, 3x 1 mL PBS) and the proteins reduced and alkylated (1 mM DTT, 45 min, 25 °C; 5.5 mM IAA, 30 min, 25 °C). The alkylation reaction was stopped (4 mM DTT, 30 min, 25 °C) and the proteins were predigested [1 μL LysC (0.5 mg/mL), 2 h, 25 °C] prior to being digested overnight (600 μL 50 mM TEAB, 1.5 μL trypsin (0.5 mg/mL), 15 h, 37 °C). Digestion was stopped (1% FA) and peptides were desalted as described previously. These peptides were freeze-dried and reconstituted in 30 μL 1% FA for MS/MS measurements on the Q Exactive instrument.

Dimethyl labeling (DML) including bacterial infection

HeLa cells were seeded in 10 cm dishes (two additional for counting) and treated with either 100 μM **pro-N6pA** or DMSO as previously described. Further procedure was conducted as described for preparative *in situ* labeling with infection and LFQ analysis with the feature that all dishes were treated with bacteria (MOI 10). After successful digestion, the “heavy” and “light” dimethyl-labeling was performed as on-column procedure as previously reported.¹³² Therefore, 50 mg Sep-Pak C18 cartridges (Waters Corp.) were washed twice with 1 mL acetonitrile, once with 1 mL elution buffer (80% acetonitrile, 0.5% FA in H_2O) and thrice with

1 mL wash buffer (0.5% FA in H₂O). Then, samples were loaded and washed with wash buffer (5x 1 mL). As next step, samples were labeled with the respective on-column dimethyl-labeling solution (5x 1 mL, for “heavy”: 45 mM sodium phosphate pH 7.5, 30 mM NaBD₃CN, 0.2% ¹³C-formaldehyde-d₂; for “light”: 45 mM sodium phosphate pH 7.5, 30 mM NaBH₃CN, 0.2% formaldehyde) in a way, that each condition (probe, MOI 10 and DMSO, MOI 10) consisted of the same number of “light” and “heavy” labeled samples. After labeling, samples were washed twice with 1 mL wash buffer before being eluted with 3x 250 µL elution buffer. Complementary samples were mixed (“light” label, probe treated with “heavy” label, DMSO treated and *vice versa*), freeze-dried and reconstituted in 30 µL 1% FA. They were sonicated for 10 min and further filtered through 0.22 µm Ultrafree-MC® centrifugal filters (Merck, UFC30GVNB) for MS/MS measurements, which were performed on the Q Exactive instrument.

SILAC labeling including bacterial infection - enrichment

Both, “heavy” and “light” labeled HeLa cells were seeded in 10 cm dishes (2 additional dishes per label) and treated with 100 µM **proN6pA**. Further procedure was performed according to preparative *in situ* labeling including infection and analysis *via* LFQ with minor modifications. Half of the dishes of “heavy” and “light” labeled HeLa cells were treated with bacteria, the rest without bacteria (MOI 0). Bacterial pellets were resuspended in SILAC-DMEM supplemented with 2 mM L-glutamine, 10 µM **pro-N6pA** and “heavy” or “light” lysine and arginine. After cell lysis, a BCA assay was performed, and the resulting cell lysate was divided into two portions for further analysis: protein enrichment (250 µg) and full proteome analysis (100 µg). For the enrichment, “light” lysate being treated with bacteria was combined with “heavy” lysate treated without bacteria and *vice versa*. All combined samples were adjusted to a total volume of 250 µL with 0.2% SDS in PBS

SILAC labeling including bacterial infection – whole proteome analysis

For the whole proteome analysis, the same cell lysate was used as for the SILAC enrichment experiments. After successfully determining the lysate’s protein concentration by BCA assay, 100 µg protein of the “heavy” lysate treated with bacteria was combined with 100 µg of the “light” lysate treated without bacteria and *vice versa*. The volume of the combined samples was adjusted to 250 µL using 0.2% SDS in PBS, and 0.96 mM CuSO₄ were added to each sample. They were further incubated for 1.5 h at 25 °C prior to being precipitated overnight using

1.1 mL acetone. Protein pellets were harvested by centrifugation (15 min, 4 °C, 13 000 rpm) and washed with ice-cold MeOH. Therefore, 1 mL of MeOH was added to the protein pellet which was resuspended by sonication (10% intensity, 10 s) and harvested again as described previously. The MeOH washing step was repeated once again, and the resulting protein pellet was dissolved in 200 µL digestion buffer (3.9 M urea, 1.1 M thiourea in 20 mM Hepes, pH 7.5). Reduction and alkylation of disulfides as well as further protein digestion was performed as described for preparative *in situ* labeling containing infection and LFQ analysis. The protein digestion was stopped by adding 1% FA and peptides were desalted with Sep-Pak C18 cartridges (Waters Corp.) as previously described. As last step, peptides were freeze-dried, reconstituted in 1% FA in H₂O to a final concentration of 2 µg/µL and filtered. Whole proteome samples were measured on the Orbitrap Fusion with a sample injection volume of 4 µL.

***In situ* site identification of AMPylated proteins during infection using a TEV-cleavable linker**

After performing an infection assay of HeLa cells seeded in 15 cm dishes as described before with *V. parahaemolyticus* wild type having **pro-N6pA** and bacteria (MOI 10) as positive sample and DMSO treated cells infected with bacteria (MOI 10) as negative control, cell lysate was adjusted to a protein concentration of 6 mg and a total volume of 2 mL. Click reaction was performed with all three samples (2x positive, 1x negative) for 1.5 h at 25 °C (0.096 mM azide-TEV-biotin (10 mM stock in DMSO), 0.51 mM TCEP (53 mM stock in ddH₂O), 0.1 mM TBTA (83.5 mM stock in DMSO), and 0.96 mM CuSO₄ (stock 50 mM in ddH₂O)).¹⁴⁵ Subsequently, proteins were precipitated with at least 4-fold excess of acetone in an upright position overnight at -20 °C. MeOH washing of the pellets as well as avidin-agarose bead enrichment and on-beads digest was performed as described for preparative *in situ* labeling with infection and LFQ analysis with minor deviations. Namely, 100 µL of beads slurry were used per sample and the digestion buffer only contained urea (no thiourea). After digestion, the beads were transferred onto membrane filters [Ultrafree-MC centrifugal filters (Merck, UFC30GVNB)] which were washed beforehand (500 µL 1% FA in H₂O, 1 min with 1000g). The flow through was further processed according to the standard enrichment protocol. The beads were washed with H₂O (2x 50 µL), PBS (3x 600 µL) and again H₂O (3x 600 µL). They were further resuspended in 150 µL TEV buffer (141 µL H₂O, 7.5 µL 20x TEV buffer (Invitrogen), 1.5 µL 100 µM DTT in H₂O) and transferred into 1.5 mL low-bind Eppendorf tubes. Next, they were centrifuged (2 min, 25 °C, 400g), the supernatant was removed, and the remaining beads once again resuspended

in 150 μL TEV buffer. 50 U of AcTEV protease (5 μL , 10 U/ μL , Invitrogen) were added to each sample. TEV digestion was conducted overnight with agitation at 29 $^{\circ}\text{C}$ prior to transferring the beads onto a new membrane filter. As next step, the beads were spun down (1 min, 1000g) and washed with H_2O (2x 50 μL). The flow through was collected and acidified with 4 μL FA. Samples were desalted on stage tips (self-made pipet tips containing double C18 layer, Empore disc C18, 47MM, Agilent Technologies) which were equilibrated and washed before [70 μL MeOH, 70 μL elution buffer (80% acetonitrile, 0.5% FA in H_2O) and 3x 70 μL 0.5% FA in H_2O ; 1 min with 1000g]. The peptides were transferred onto the membranes, spun down (5 min, 25 $^{\circ}\text{C}$, 1000g) and washed thrice with 70 μL 0.5% FA in H_2O . Finally, the modified peptides were eluted (2x 30 μL elution buffer) and lyophilized prior to being reconstituted in 30 μL 1% FA and filtered. Samples were measured on the Fusion instrument.

4.5. MS/MS measurements and data analysis

Q Exactive

Enrichment samples were analyzed on a Q Exactive Plus instrument (Thermo Fisher) coupled to an UltiMate 3000 nano-HPLC (Dionex) equipped with an Acclaim C18 PepMap100 75 μm ID x 2 cm trap column (Thermo Fisher) and a 25 cm Aurora Series emitter column (25 cm x 75 μm ID, 1.6 μm FSC C18) (Ionoptics) in an EASY-spray setting. Both columns were heated to 40 $^{\circ}\text{C}$ during the measurement process. For analysis, 4 μL of each peptide sample were injected. The samples were first loaded on the trap column with 0.1% trifluoroacetic acid (TFA) applying a flow rate of 5 $\mu\text{L}/\text{min}$. Second, the samples were transferred to the separation column with a flow rate of 0.4 $\mu\text{L}/\text{min}$ where peptides were separated with a 152 min gradient (buffer A: H_2O with 0.1% FA, buffer B: acetonitrile with 0.1% FA). The gradient consisted of the following steps: holding buffer B at 5% for 7 min, further increasing buffer B to 22% during 105 min, to 32% the next 10 min and to 90% within 10 min. Once buffer B reached a concentration of 90%, this solvent ratio was maintained for another 10 min before decreasing it to 5% in 0.1 min, at which level it stayed until the end of the run. Peptides were ionized at a capillary temperature of 275 $^{\circ}\text{C}$ and the instrument was operated in a Top12 data dependent mode. For full scan acquisition, the Orbitrap mass analyzer was set to a resolution of $R = 140\,000$, an automatic gain control (AGC) target of 3×10^6 , and a maximal injection time of 80 ms in a scan range of 300–1500 m/z . Precursors having a charge state of >1 , a minimum AGC target of 1×10^3 and intensities higher than 1×10^4 were selected for fragmentation. Peptide fragments were generated by HCD (higher-energy collisional dissociation) with a

normalized collision energy of 27% and recorded in the Orbitrap at a resolution of $R = 17\,500$. Moreover, the AGC target was set to 1×10^5 with a maximum injection time of 100 ms. Dynamic exclusion duration was set to 60 s and isolation was performed in the quadrupole using a window of 1.6 m/z.

Orbitrap Fusion

Whole proteome samples were measured on an Orbitrap Fusion instrument (Thermo Fisher Scientific Inc.) being equipped with an UltiMate 3000 nano HPLC system (Dionex) and the following trap and separation columns (heated to 50 °C, both Thermo Fisher): Acclaim C18 PepMap 100 75 μm ID x 2 cm and Acclaim PepMap RSLC C18 75 μm ID x 50 cm. Samples were measured in an EASY-spray setting and loaded on the trap column with a flow rate of 5 $\mu\text{L}/\text{min}$ using 0.1% TFA. Subsequently, samples were separated using a flow rate of 0.3 $\mu\text{L}/\text{min}$ and a 152 min gradient (buffer A: H_2O with 0.1% FA, buffer B: acetonitrile with 0.1% FA, gradient: starting with 5% buffer B, holding it for 7 min, the next 105 min increasing buffer B from 5% to 22%, from 22% to 32% buffer B in 10 min, then to 90% buffer B in 10 min and holding it there for 10 min, decreasing it to buffer B in 0.1 min and holding it there until the end of the run). The instrument was operated in a 3 s top speed data dependent mode. Peptides were ionized with a capillary temperature of 275 °C and MS full scans were performed with a resolution of $R = 120\,000$ in the Orbitrap. Further settings were an AGC ion target value of 2×10^5 and a scan range between 300 to 1500 m/z having a maximal injection time of 50 s. For MS2 scans, precursors having charge states between 2 and 7 and intensities higher than 5×10^3 were further selected for fragmentation. The fragmentation was conducted with an HCD collision energy of 30% and the isolation was performed in the quadrupole using a window of 1.6 m/z. The AGC target was set to 1×10^4 , the maximum injection time to 35 ms and the function “inject ions for all available parallelizable time” was enabled. Furthermore, the dynamic exclusion time was set to 60 s with 10 ppm low and high mass tolerance. The site-ID experiments applying the TEV cleavable linker were analyzed on the Orbitrap Fusion with slightly modified parameters. After HCD fragmentation, peptides having the following targeted masses of m/z 558.3, 279.65, or 186.77 (adenine-TEV, single, double, or triple charged), m/z of 672.33, 336.67, or 224.78 (adenosine-TEV, single, double, or triple charged) or m/z of 770.31, 385.66, or 257.44 (phosphoadenosine-TEV, single, double, or triple charged) were further selected for electron-transfer dissociation (ETD) fragmentation while prioritizing the

highest charge states. ETD fragmentation scans were acquired in the Orbitrap with a resolution of $R = 30\,000$, a maximum injection time of 40 ms and an AGC target value of 5×10^4 .

Bioinformatics and statistics

For peptide and protein identification, MS data were processed with MaxQuant (version 1.6.2.10) having Andromeda as search engine.¹⁴⁷ Searches were performed against the UniProt database for Homo sapiens (taxon identifier: 9606, canonical version, reviewed and unreviewed proteome, not older than three months prior to MS measurements). For infection assays, all proteins in the UniProt database of the respective bacteria (*Vibrio parahaemolyticus* serotype O3:K6, strain RIMD 2210633, taxon identifier: 223926; *Pseudomonas aeruginosa* PAO1, taxon identifier: 208964; *Escherichia coli* CFT073, taxon identifier: 199310; and *Staphylococcus aureus* USA300, taxon identifier: 367830; canonical versions, reviewed and unreviewed proteomes, not older than three months prior to MS measurements) were added to the MaxQuant contaminants file. As search parameters, mostly default settings were applied (trypsin/P as digest enzyme, max. 2 missed cleavages, oxidation (M) and acetylation (protein N-term) as variable modifications, carbamidomethylation (C) as fixed modification, label-free quantification LFQ with a minimal ratio count of 2, min peptide length 7, 20 ppm for precursor mass tolerance (FTMS MS/MS match tolerance) and 0.5 Da for fragment mass tolerance (ITMS MS/MS match tolerance). Furthermore, proteins were identified using PSM FDR 0.01, protein FDR 0.01, min peptides 2, min razor + unique peptides 2, min unique peptides 2, razor protein FDR enabled, second peptides enabled. Additionally, the match between run function with its settings match time window of 0.7 min and alignment time window of 20 min was also enabled. Multiplicity was set to 2 for DML and SILAC experiments. Then, parameter groups were defined and in the setting group specific parameters, labels were selected. For samples containing treated “heavy” lysate and untreated “light” lysate, DimethLys8 and DimethNter8 (DML) or Arg10 and Lys8 (SILAC) were selected for the first and DimethLys0 and DimethNter0 (DML) or Arg6 and Lys4 (SILAC) for the latter. For samples containing treated “light” lysate and untreated “heavy” lysate, DimethLys8 and DimethNter8 (DML) or Arg10 and Lys8 (SILAC) were selected for the first and DimethLys0 and DimethNter0 (DML) or Arg6 and Lys4 (SILAC) for the latter. The requantify option was enabled for DML and SILAC samples. Peptides of the site-ID experiments were searched for the fixed carbamidomethyl (C) modification (57.02146) as well as for the variable modification of the “light” TEV linker coupled to AMP (751.2915) on serine, tyrosine, or threonine residues. Furthermore, the options

min peptides, min razor + unique peptides as well as min unique peptides were used with default settings (1 – 1 – 0). For further statistical data analysis, the Perseus software (version 1.6.2.3) was used. Therefore, the protein groups table in the txt folder resulting from the MaxQuant analysis was uploaded into the program. LFQ intensities and DML or SILAC ratios were \log_2 transformed and putative contaminants as well as reverse hits were removed. Next, all samples were categorical annotated according to their treatment conditions. Then, the resulting matrix was filtered against 75% of valid values in at least one group. Missing values were imported from normal distribution (width 0.3, down shift 1.8, total matrix) and p -values were obtained by a two-sided two sample t-test over replicates with a Benjamini–Hochberg false discovery rate correction (FDR 0.05). Data were visualized by generating scatter plots [Student's t-test difference (treated/control) against $-\log$ student's t-test p -value (treated/control)]. For DML/SILAC experiments, rows were filtered for 75% of valid values in total and a one-sample Student's t-test was performed.

4.6. Biochemical assays

4.6.1. CFU assays

For each bacterial strain and each growth condition, the CFU (colony forming units) vs OD₆₀₀ assay was performed at least twice. Therefore, an overnight culture of the desired bacterial strain was diluted 1:100 into fresh media and cultivated under standard condition. At various time points, the OD₆₀₀ value was measured and 5 μ L of the day culture were further diluted. Each dilution was plated on pure agar plates of the corresponding medium in triplicates and incubated for 24 h at 37 °C. Each plate, having between 12 and 120 colonies per plated sample, was counted. Having various data points connecting each OD₆₀₀ value to a certain CFU amount, different mathematical functions were applied to represent the OD₆₀₀ to CFU ratio in the area of interest.

4.6.2. AMPylation assay

The *in vitro* AMPylation assay was performed as described previously.⁹² Briefly, the purified AMPylator VopS (AA 31–378, 1 μ M, kind gift of the Itzen lab) was incubated with either 100 μ M ATP or 100 μ M **N6pATP** and the AMPylation targets Cdc42 (AA 1–188, 50 μ M, kind gift of the Itzen lab) or RhoG (full protein, 5 μ M, Abcam) with 2 μ L 10x AMPylation buffer (1x

buffer: 20 mM Hepes pH 7.5, 100 mM NaCl, 5 mM MgCl₂, 0.1 mg/mL BSA and 1 mM DTT). Samples were diluted to a final volume of 20 μ L. In case of negative control experiments, the volume of the missing ingredient was compensated by water. The AMPylation reaction was performed for 1.5 h at 30 °C before being further analyzed.

4.6.3. Intact protein MS measurement (IPMS)

Intact protein MS of recombinant protein or proteins modified *via* the AMPylation assay were performed as described previously.¹⁴⁸ The samples were measured with an UltiMate 3000 HPLC system (Dionex) being equipped with a Massprep online desalting cartridge (Waters) and coupled to a Finnigan LTQ FT Ultra mass spectrometer (Thermo Fisher Scientific). Ions were generated by electrospray ionization (capillary temperature 275 °C, spray voltage 4.0 kV, tube lens 110 V, capillary voltage 48 V, sheath gas 60 arb, aux gas 10 arb, sweep gas 0.2 arb) and separated with a flow rate of 0.4 mL/min and the following gradient: (buffer A: 0.1% FA in H₂O, buffer B: 0.1% FA, 9.9% H₂O in acetonitrile, 1 min hold 6% buffer B, in 1.5 min increase buffer B to 95%, hold at 95% buffer B for 2 min, decrease to 6% buffer B in 0.2 min, stay at 6% buffer B for further 0.3 min). The instrumented was operated with a resolution of $R = 200\,000$ and a mass range of m/z from 600 to 2000 while acquiring full scans. The ProMass software for Xcalibur (Version 2.8) was used for deconvolution (input m/z range: 600–2000; output mass range: 15 000–30 000 Da).

4.6.4. SDS-PAGE of recombinantly AMPylated proteins

After successful *in vitro* AMPylation of recombinant proteins, the reaction volume was filled up with PBS to a total volume of 50 μ L. Subsequently, a click reaction was performed by adding 0.096 mM rhodamine-azide (10 mM stock in DMSO, Rh-N3, base click), 0.96 mM TCEP (100 mM stock in ddH₂O, Roth), 0.096 mM TBTA ligand (83.5 mM stock in DMSO, TCI), and 0.96 mM CuSO₄ (50 mM stock in ddH₂O). After incubation for 1.5 h at 25 °C and 450 rpm, 25 μ L of 2x Laemmli buffer (63 mM Tris HCl, 2% SDS, 10% glycerol, 0.0025% bromophenol blue, 5% β -mercaptoethanol) were added. The samples were thoroughly mixed before boiling them for 5 min at 95 °C. Afterward, a standard SDS-PAGE analysis was performed.

4.6.5. Site-ID of recombinant proteins

Once the successful AMPylation reaction was confirmed by IPMS, the samples were diluted in 150 μ L digestion buffer (3.9 M urea, 1.1 M thiourea in 20 mM Hepes, pH 7.5) and treated with DTT (1.2 mM, 1 h, 25 °C), IAA (6.5 mM, 1.5 h, 25 °C), and DTT (4.7 mM, 45 min, 25 °C). Next, the samples were predigested with LysC (1 μ L LysC, 0.5 mg/mL, 2 h, 25 °C) before diluting them in 600 μ L 50 mM TEAB and digesting them with trypsin (1.5 μ L, 0.5 mg/mL, 15 h, 37 °C). After digestion, samples were acidified (1% FA in H₂O) and desalted using stage-tips. The membranes were washed and equilibrated [70 μ L acetonitrile, 70 μ L elution buffer (80% acetonitrile, 0.5% FA in H₂O), 3x 70 μ L 0.5% FA in H₂O] before being loaded with samples. Afterward, membranes were washed again (3x 70 μ L 0.5% FA in H₂O) and peptides were eluted with elution buffer (2x 30 μ L). The freeze-dried peptides were reconstituted in 25 μ L 1% FA in H₂O and filtered [0.22 μ m Ultrafree-MC centrifugal filters (Merck, UFC30GVNB)]. The samples (5 μ L) were injected on the Q Exactive instrument for further analysis.

5. References

- (1) Nicolaou, K. C., and Rigol, S. (2018) A brief history of antibiotics and select advances in their synthesis. *J Antibiot (Tokyo)* 71, 153–184.
- (2) Watkins, R. R., and Bonomo, R. A. (2016) Overview: Global and Local Impact of Antibiotic Resistance. *Infect Dis Clin North Am* 30, 313–322.
- (3) Fodor, A., Abate, B. A., Deák, P., Fodor, L., Gyenge, E., Klein, M. G., Koncz, Z., Muvevi, J., Ötvös, L., Székely, G., Vozik, D., and Makrai, L. (2020) Multidrug Resistance (MDR) and Collateral Sensitivity in Bacteria, with Special Attention to Genetic and Evolutionary Aspects and to the Perspectives of Antimicrobial Peptides-A Review. *Pathogens* 9.
- (4) Tsakou, F., Jersie-Christensen, R., Jenssen, H., and Mojsoska, B. (2020) The Role of Proteomics in Bacterial Response to Antibiotics. *Pharmaceuticals (Basel)* 13.
- (5) Tiwari, V. (2019) Post-translational modification of ESKAPE pathogens as a potential target in drug discovery. *Drug Discov Today* 24, 814–822.
- (6) Ribet, D., and Cossart, P. (2010) Post-translational modifications in host cells during bacterial infection. *FEBS Lett* 584, 2748–2758.
- (7) Ribet, D., and Cossart, P. (2010) Pathogen-mediated posttranslational modifications: A re-emerging field. *Cell* 143, 694–702.
- (8) Walsh, C. T., Garneau-Tsodikova, S., and Gatto, G. J. (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chem Int Ed Engl* 44, 7342–7372.
- (9) Berget, S. M., Moore, C., and Sharp, P. A. (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A* 74, 3171–3175.
- (10) Beltrao, P., Bork, P., Krogan, N. J., and van Noort, V. (2013) Evolution and functional cross-talk of protein post-translational modifications. *Mol Syst Biol* 9, 714.
- (11) Chambers, K. A., and Scheck, R. A. (2020) Bacterial virulence mediated by orthogonal post-translational modification. *Nat Chem Biol* 16, 1043–1051.
- (12) Collier, R. J., and Cole, H. A. (1969) Diphtheria toxin subunit active in vitro. *Science* 164, 1179–1181.
- (13) Salomon, D., and Orth, K. (2013) Lost after translation: post-translational modifications by bacterial type III effectors. *Curr Opin Microbiol* 16, 213–220.
- (14) Aktories, K. (2015) Rho-modifying bacterial protein toxins. *Pathog Dis* 73, ftv091.

- (15) Shao, F., Merritt, P. M., Bao, Z., Innes, R. W., and Dixon, J. E. (2002) A Yersinia Effector and a Pseudomonas Avirulence Protein Define a Family of Cysteine Proteases Functioning in Bacterial Pathogenesis. *Cell* 109, 575–588.
- (16) Shao, F., Vacratsis, P. O., Bao, Z., Bowers, K. E., Fierke, C. A., and Dixon, J. E. (2003) Biochemical characterization of the Yersinia YopT protease: cleavage site and recognition elements in Rho GTPases. *Proc Natl Acad Sci U S A*, 904–909.
- (17) Just, I., Wilm, M., Selzer, J., Rex, G., Eichel-Streiber, C. von, Mann, M., and Aktories, K. (1995) The enterotoxin from Clostridium difficile (ToxA) monoglucosylates the Rho proteins. *J Biol Chem* 270, 13932–13936.
- (18) Just, I., Selzer, J., Wilm, M., Eichel-Streiber, C. von, Mann, M., and Aktories, K. (1995) Glucosylation of Rho proteins by Clostridium difficile toxin B. *Nature* 375, 500–503.
- (19) Just, I., Selzer, J., Hofmann, F., Green, G. A., and Aktories, K. (1996) Inactivation of Ras by Clostridium sordellii lethal toxin-catalyzed glucosylation. *J Biol Chem* 271, 10149–10153.
- (20) Genth, H., Hofmann, F., Selzer, J., Rex, G., Aktories, K., and Just, I. (1996) Difference in protein substrate specificity between hemorrhagic toxin and lethal toxin from Clostridium sordellii. *Biochem Biophys Res Commun* 229, 370–374.
- (21) Sehr, P., Joseph, G., Genth, H., Just, I., Pick, E., and Aktories, K. (1998) Glucosylation and ADP ribosylation of rho proteins: effects on nucleotide binding, GTPase activity, and effector coupling. *Biochemistry* 37, 5296–5304.
- (22) Schmidt, G., Sehr, P., Wilm, M., Selzer, J., Mann, M., and Aktories, K. (1997) Gln 63 of Rho is deamidated by Escherichia coli cytotoxic necrotizing factor-1. *Nature* 387, 725–729.
- (23) Flatau, G., Lemichez, E., Gauthier, M., Chardin, P., Paris, S., Fiorentini, C., and Boquet, P. (1997) Toxin-induced activation of the G protein p21 Rho by deamidation of glutamine. *Nature* 387, 729–733.
- (24) Lockman, H. A., Gillespie, R. A., Baker, B. D., and Shakhnovich, E. (2002) Yersinia pseudotuberculosis produces a cytotoxic necrotizing factor. *Infect Immun* 70, 2708–2714.
- (25) Zhang, L., Krachler, A. M., Broberg, C. A., Li, Y., Mirzaei, H., Gilpin, C. J., and Orth, K. (2012) Type III effector VopC mediates invasion for Vibrio species. *Cell Rep*, 453–460.
- (26) Doye, A., Mettouchi, A., Bossis, G., Clément, R., Buisson-Touati, C., Flatau, G., Gagnoux, L., Piechaczyk, M., Boquet, P., and Lemichez, E. (2002) CNF1 Exploits the Ubiquitin-Proteasome Machinery to Restrict Rho GTPase Activation for Bacterial Host Cell Invasion. *Cell* 111, 553–564.

-
- (27) Aktories, K., Weller, U., and Chhatwal, G. S. (1987) Clostridium botulinum type C produces a novel ADP-ribosyltransferase distinct from botulinum C2 toxin. *FEBS Lett* 212, 109–113.
- (28) Sugai, M., Hashimoto, K., Kikuchi, A., Inoue, S., Okumura, H., Matsumoto, K., Goto, Y., Ohgai, H., Moriishi, K., and Syuto, B. (1992) Epidermal cell differentiation inhibitor ADP-ribosylates small GTP-binding proteins and induces hyperplasia of epidermis. *J Biol Chem* 267, 2600–2604.
- (29) Yarbrough, M. L., Li, Y., Kinch, L. N., Grishin, N. V., Ball, H. L., and Orth, K. (2009) AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling. *Science*, 269–272.
- (30) Worby, C. A., Mattoo, S., Kruger, R. P., Corbeil, L. B., Koller, A., Mendez, J. C., Zekarias, B., Lazar, C., and Dixon, J. E. (2009) The fic domain: regulation of cell signaling by adenylation. *Mol Cell* 34, 93–103.
- (31) Kingdon, H. S., Shapiro, B. M., and Stadtman, E. R. (1967) Regulation of glutamine synthetase. 8. ATP: glutamine synthetase adenylyltransferase, an enzyme that catalyzes alterations in the regulatory properties of glutamine synthetase. *Proc Natl Acad Sci U S A* 58, 1703–1710.
- (32) Stadtman, E. R., Shapiro, B. M., Kingdon, H. S., Woolfolk, C. A., and Hubbard, J. S. (1968) Cellular regulation of glutamine synthetase activity in Escherichia coli. *Advances in Enzyme Regulation* 6, 257–289.
- (33) Zimmerman, J. J., Saint André-von Arnim, A. von, and McLaughlin, J. (2011) Cellular Respiration. In *Pediatric Critical Care* pp 1058–1072, Elsevier.
- (34) Hedberg, C., and Itzen, A. (2015) Molecular perspectives on protein adenylation. *ACS Chem Biol*, 12–21.
- (35) Casey, A. K., and Orth, K. (2018) Enzymes Involved in AMPylation and deAMPylation. *Chem Rev*, 1199–1215.
- (36) Utsumi, R., Nakamoto, Y., Kawamukai, M., Himeno, M., and Komano, T. (1982) Involvement of cyclic AMP and its receptor protein in filamentation of an Escherichia coli fic mutant. *J Bacteriol* 151, 807–812.
- (37) Sreelatha, A., Yee, S. S., Lopez, V. A., Park, B. C., Kinch, L. N., Pilch, S., Servage, K. A., Zhang, J., Jiou, J., Karasiewicz-Urbańska, M., Łobocka, M., Grishin, N. V., Orth, K., Kucharczyk, R., Pawłowski, K., Tomchick, D. R., and Tagliabracci, V. S. (2018) Protein AMPylation by an Evolutionarily Conserved Pseudokinase. *Cell*, 809-821.e19.

- (38) Yang, Y., Yue, Y., Song, N., Li, C., Yuan, Z., Wang, Y., Ma, Y., Li, H., Zhang, F., Wang, W., Jia, H., Li, P., Li, X., Wang, Q., Ding, Z., Dong, H., Gu, L., and Li, B. (2020) The YdiU Domain Modulates Bacterial Stress Signaling through Mn²⁺-Dependent UMPylation. *Cell Rep* 32, 108161.
- (39) Müller, M. P., Peters, H., Blümer, J., Blankenfeldt, W., Goody, R. S., and Itzen, A. (2010) The Legionella effector protein DrrA AMPylates the membrane traffic regulator Rab1b. *Science*, 946–949.
- (40) Barthelmes, K., Ramcke, E., Kang, H.-S., Sattler, M., and Itzen, A. (2020) Conformational control of small GTPases by AMPylation. *Proc Natl Acad Sci U S A*, 5772–5781.
- (41) Du, J., Wrisberg, M.-K. von, Gulen, B., Stahl, M., Pett, C., Hedberg, C., Lang, K., Schneider, S., and Itzen, A. (2021) Rab1-AMPylation by Legionella DrrA is allosterically activated by Rab1. *Nat Commun* 12, 460.
- (42) Harms, A., Stanger, F. V., and Dehio, C. (2016) Biological Diversity and Molecular Plasticity of FIC Domain Proteins. *Annu Rev Microbiol*, 341–360.
- (43) Mitchell, A., Chang, H.-Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., McAnulla, C., McMenamin, C., Nuka, G., Pesseat, S., Sangrador-Vegas, A., Scheremetjew, M., Rato, C., Yong, S.-Y., Bateman, A., Punta, M., Attwood, T. K., Sigrist, C. J. A., Redaschi, N., Rivoire, C., Xenarios, I., Kahn, D., Guyot, D., Bork, P., Letunic, I., Gough, J., Oates, M., Haft, D., Huang, H., Natale, D. A., Wu, C. H., Orengo, C., Sillitoe, I., Mi, H., Thomas, P. D., and Finn, R. D. (2015) The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res*, D213-21.
- (44) Kinch, L. N., Yarbrough, M. L., Orth, K., and Grishin, N. V. (2009) Fido, a novel AMPylation domain common to fic, doc, and AvrB. *PLoS One*, e5818.
- (45) Engel, P., Goepfert, A., Stanger, F. V., Harms, A., Schmidt, A., Schirmer, T., and Dehio, C. (2012) Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. *Nature*, 107–110.
- (46) Garcia-Pino, A., Christensen-Dalsgaard, M., Wyns, L., Yarmolinsky, M., Magnuson, R. D., Gerdes, K., and Loris, R. (2008) Doc of prophage P1 is inhibited by its antitoxin partner Phd through fold complementation. *J Biol Chem*, 30821–30827.
- (47) Garcia-Pino, A., Zenkin, N., and Loris, R. (2014) The many faces of Fic: structural and functional aspects of Fic enzymes. *Trends Biochem Sci*, 121–129.
- (48) Goepfert, A., Stanger, F. V., Dehio, C., and Schirmer, T. (2013) Conserved inhibitory mechanism and competent ATP binding mode for adenylyltransferases with Fic fold. *PLoS One*, e64901.

-
- (49) Gerdes, K., Ed. (2013) *Prokaryotic Toxin-Antitoxins*, Springer Berlin Heidelberg, Berlin, Heidelberg.
- (50) Harms, A., Stanger, F. V., Scheu, P. D., Jong, I. G. de, Goepfert, A., Glatter, T., Gerdes, K., Schirmer, T., and Dehio, C. (2015) Adenylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology. *Cell Rep*, 1497–1507.
- (51) Goepfert, A., Harms, A., Schirmer, T., and Dehio, C. (2013) Type II Toxin-Antitoxin Loci: The fic Family. In *Prokaryotic Toxin-Antitoxins* (Gerdes, K., Ed.) pp 177–187, Springer Berlin Heidelberg, Berlin, Heidelberg.
- (52) Stanger, F. V., Burmann, B. M., Harms, A., Aragão, H., Mazur, A., Sharpe, T., Dehio, C., Hiller, S., and Schirmer, T. (2016) Intrinsic regulation of FIC-domain AMP-transferases by oligomerization and automodification. *Proc Natl Acad Sci U S A*, E529-37.
- (53) Bunney, T. D., Cole, A. R., Broncel, M., Esposito, D., Tate, E. W., and Katan, M. (2014) Crystal structure of the human, FIC-domain containing protein HYPE and implications for its functions. *Structure* 22, 1831–1843.
- (54) Sanyal, A., Chen, A. J., Nakayasu, E. S., Lazar, C. S., Zbornik, E. A., Worby, C. A., Koller, A., and Mattoo, S. (2015) A novel link between Fic (filamentation induced by cAMP)-mediated adenylation/AMPylation and the unfolded protein response. *J Biol Chem* 290, 8482–8499.
- (55) Perera, L. A., Rato, C., Yan, Y., Neidhardt, L., McLaughlin, S. H., Read, R. J., Preissler, S., and Ron, D. (2019) An oligomeric state-dependent switch in the ER enzyme FICD regulates AMPylation and deAMPylation of BiP. *EMBO J* 38, e102177.
- (56) Sanyal, A., Dutta, S., Camara, A., Chandran, A., Koller, A., Watson, B. G., Sengupta, R., Ysselstein, D., Montenegro, P., Cannon, J., Rochet, J.-C., and Mattoo, S. (2019) Alpha-Synuclein Is a Target of Fic-Mediated Adenylation/AMPylation: Possible Implications for Parkinson's Disease. *J Mol Biol* 431, 2266–2282.
- (57) Kielkowski, P., Buchsbaum, I. Y., Kirsch, V. C., Bach, N. C., Drukker, M., Cappello, S., and Sieber, S. A. (2020) FICD activity and AMPylation remodelling modulate human neurogenesis. *Nat Commun* 11, 517.
- (58) Gulen, B., Rosselin, M., Fauser, J., Albers, M. F., Pett, C., Krisp, C., Pogenberg, V., Schlüter, H., Hedberg, C., and Itzen, A. (2020) Identification of targets of AMPylating Fic enzymes by co-substrate-mediated covalent capture. *Nat Chem*, 732–739.
- (59) Truttman, M. C., Wu, Q., Stiegeler, S., Duarte, J. N., Ingram, J., and Ploegh, H. L. (2015) HypE-specific nanobodies as tools to modulate HypE-mediated target AMPylation. *J Biol Chem* 290, 9087–9100.

- (60) Broncel, M., Serwa, R. A., Bunney, T. D., Katan, M., and Tate, E. W. (2016) Global Profiling of Huntingtin-associated protein E (HYPE)-Mediated AMPylation through a Chemical Proteomic Approach. *Mol Cell Proteomics* 15, 715–725.
- (61) Rahman, M., Ham, H., Liu, X., Sugiura, Y., Orth, K., and Krämer, H. (2012) Visual neurotransmission in *Drosophila* requires expression of Fic in glial capitate projections. *Nat Neurosci* 15, 871–875.
- (62) Moehlman, A. T., Casey, A. K., Servage, K., Orth, K., and Krämer, H. (2018) Adaptation to constant light requires Fic-mediated AMPylation of BiP to protect against reversible photoreceptor degeneration. *Elife* 7.
- (63) Nitika, Porter, C. M., Truman, A. W., and Truttmann, M. C. (2020) Post-translational modifications of Hsp70 family proteins: Expanding the chaperone code. *J Biol Chem* 295, 10689–10708.
- (64) Carroll, P., Pashley, C. A., and Parish, T. (2008) Functional analysis of GlnE, an essential adenylyl transferase in *Mycobacterium tuberculosis*. *J Bacteriol*, 4894–4902.
- (65) Lu, C.-H., McCloskey, A., Chen, F.-R., Nakayasu, E. S., Zhang, L.-Q., and Luo, Z.-Q. (2020) Fic Proteins Inhibit the Activity of Topoisomerase IV by AMPylation in Diverse Bacteria. *Front Microbiol*, 2084.
- (66) Veyron, S., Peyroche, G., and Cherfils, J. (2018) FIC proteins: from bacteria to humans and back again. *Pathog Dis* 76.
- (67) Mattoo, S., Durrant, E., Chen, M. J., Xiao, J., Lazar, C. S., Manning, G., Dixon, J. E., and Worby, C. A. (2011) Comparative analysis of *Histophilus somni* immunoglobulin-binding protein A (IbpA) with other fic domain-containing enzymes reveals differences in substrate and nucleotide specificities. *J Biol Chem*, 32834–32842.
- (68) Siamer, S., and Dehio, C. (2015) New insights into the role of *Bartonella* effector proteins in pathogenesis. *Curr Opin Microbiol*, 80–85.
- (69) Engel, P., Salzburger, W., Liesch, M., Chang, C.-C., Maruyama, S., Lanz, C., Calteau, A., Lajus, A., Médigue, C., Schuster, S. C., and Dehio, C. (2011) Parallel evolution of a type IV secretion system in radiating lineages of the host-restricted bacterial pathogen *Bartonella*. *PLoS Genet*, e1001296.
- (70) Palanivelu, D. V., Goepfert, A., Meury, M., Guye, P., Dehio, C., and Schirmer, T. (2011) Fic domain-catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA. *Protein Sci*, 492–499.
- (71) Pieleis, K., Glatter, T., Harms, A., Schmidt, A., and Dehio, C. (2014) An experimental strategy for the identification of AMPylation targets from complex protein samples. *Proteomics*, 1048–1052.

-
- (72) Marlaire, S., and Dehio, C. (2020) *Bartonella effector protein C mediates actin stress fiber formation via recruitment of GEF-H1 to the plasma membrane.*
- (73) Roy, C. R., and Mukherjee, S. (2009) Bacterial FIC Proteins AMP Up Infection. *Sci Signal*, 14.
- (74) Mukherjee, S., Liu, X., Arasaki, K., McDonough, J., Galán, J. E., and Roy, C. R. (2011) Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature*, 103–106.
- (75) Hardiman, C. A., and Roy, C. R. (2014) AMPylation is critical for Rab1 localization to vacuoles containing *Legionella pneumophila*. *mBio*, e01035-13.
- (76) Broberg, C. A., Calder, T. J., and Orth, K. (2011) *Vibrio parahaemolyticus* cell biology and pathogenicity determinants. *Microbes Infect* 13, 992–1001.
- (77) Letchumanan, V., Chan, K.-G., and Lee, L.-H. (2014) *Vibrio parahaemolyticus*: a review on the pathogenesis, prevalence, and advance molecular identification techniques. *Front Microbiol* 5, 705.
- (78) Souza Santos, M. de, and Orth, K. (2014) Intracellular *Vibrio parahaemolyticus* escapes the vacuole and establishes a replicative niche in the cytosol of epithelial cells. *mBio* 5, e01506-14.
- (79) Ben-Yaakov, R., and Salomon, D. (2019) The regulatory network of *Vibrio parahaemolyticus* type VI secretion system 1. *Environ Microbiol* 21, 2248–2260.
- (80) Rauh, T., Brameyer, S., Kielkowski, P., Jung, K., and Sieber, S. A. (2020) MS-Based in Situ Proteomics Reveals AMPylation of Host Proteins during Bacterial Infection. *ACS Infect Dis* 6, 3277–3289.
- (81) Bhattacharjee, R. N., Park, K.-S., Kumagai, Y., Okada, K., Yamamoto, M., Uematsu, S., Matsui, K., Kumar, H., Kawai, T., Iida, T., Honda, T., Takeuchi, O., and Akira, S. (2006) VP1686, a *Vibrio* type III secretion protein, induces toll-like receptor-independent apoptosis in macrophage through NF- κ B inhibition. *J Biol Chem* 281, 36897–36904.
- (82) Luong, P., Kinch, L. N., Brautigam, C. A., Grishin, N. V., Tomchick, D. R., and Orth, K. (2010) Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *J Biol Chem* 285, 20155–20163.
- (83) Lewallen, D. M., Steckler, C. J., Knuckley, B., Chalmers, M. J., and Thompson, P. R. (2012) Probing adenylation: using a fluorescently labelled ATP probe to directly label and immunoprecipitate VopS substrates. *Mol Biosyst* 8, 1701–1706.

- (84) Yu, X., Woolery, A. R., Luong, P., Hao, Y. H., Grammel, M., Westcott, N., Park, J., Wang, J., Bian, X., Demirkan, G., Hang, H. C., Orth, K., and LaBaer, J. (2014) Copper-catalyzed azide-alkyne cycloaddition (click chemistry)-based detection of global pathogen-host AMPylation on self-assembled human protein microarrays. *Mol Cell Proteomics* 13, 3164–3176.
- (85) Higa, N., Toma, C., Koizumi, Y., Nakasone, N., Nohara, T., Masumoto, J., Kodama, T., Iida, T., and Suzuki, T. (2013) *Vibrio parahaemolyticus* effector proteins suppress inflammasome activation by interfering with host autophagy signaling. *PLoS Pathog* 9, e1003142.
- (86) Xu, H., Yang, J., Gao, W., Li, L., Li, P., Zhang, L., Gong, Y.-N., Peng, X., Xi, J. J., Chen, S., Wang, F., and Shao, F. (2014) Innate immune sensing of bacterial modifications of Rho GTPases by the Pyrin inflammasome. *Nature* 513, 237–241.
- (87) Woolery, A. R., Yu, X., LaBaer, J., and Orth, K. (2014) AMPylation of Rho GTPases subverts multiple host signaling processes. *J Biol Chem* 289, 32977–32988.
- (88) Smit, C., Blümer, J., Eerland, M. F., Albers, M. F., Müller, M. P., Goody, R. S., Itzen, A., and Hedberg, C. (2011) Efficient synthesis and applications of peptides containing adenylylated tyrosine residues. *Angew Chem Int Ed Engl* 50, 9200–9204.
- (89) Hao, Y.-H., Chuang, T., Ball, H. L., Luong, P., Li, Y., Flores-Saaib, R. D., and Orth, K. (2011) Characterization of a rabbit polyclonal antibody against threonine-AMPylation. *J Biotechnol* 151, 251–254.
- (90) Müller, M. P., Albers, M. F., Itzen, A., and Hedberg, C. (2014) Exploring adenylylation and phosphocholination as post-translational modifications. *Chembiochem* 15, 19–26.
- (91) Höpfner, D., Fauser, J., Kaspers, M. S., Pett, C., Hedberg, C., and Itzen, A. (2020) Monoclonal Anti-AMP Antibodies Are Sensitive and Valuable Tools for Detecting Patterns of AMPylation. *iScience* 23, 101800.
- (92) Grammel, M., Luong, P., Orth, K., and Hang, H. C. (2011) A chemical reporter for protein AMPylation. *J Am Chem Soc* 133, 17103–17105.
- (93) Yu, X., and LaBaer, J. (2015) High-throughput identification of proteins with AMPylation using self-assembled human protein (NAPPA) microarrays. *Nat Protoc* 10, 756–767.
- (94) Creech, C., Kanaujia, M., and Causey, C. P. (2015) Synthesis and evaluation of 2-ethynyladenosine-5'-triphosphate as a chemical reporter for protein AMPylation. *Org Biomol Chem* 13, 8550–8555.
- (95) Li, Y., Al-Eryani, R., Yarbrough, M. L., Orth, K., and Ball, H. L. (2011) Characterization of AMPylation on threonine, serine, and tyrosine using mass spectrometry. *J Am Soc Mass Spectrom* 22, 752–761.

- (96) Hansen, T., Albers, M., Hedberg, C., and Sickmann, A. (2013) Adenylylation, MS, and proteomics—Introducing a “new” modification to bottom-up proteomics. *Proteomics* 13, 955–963.
- (97) Kielkowski, P., Buchsbaum, I. Y., Becker, T., Bach, K., Cappello, S., and Sieber, S. A. (2020) A Pronucleotide Probe for Live-Cell Imaging of Protein AMPylation. *Chembiochem* 21, 1285–1287.
- (98) Mehellou, Y., Rattan, H. S., and Balzarini, J. (2018) The ProTide Prodrug Technology: From the Concept to the Clinic. *J Med Chem* 61, 2211–2226.
- (99) Parker, C. G., and Pratt, M. R. (2020) Click Chemistry in Proteomic Investigations. *Cell* 180, 605–632.
- (100) Timp, W., and Timp, G. (2020) Beyond mass spectrometry, the next step in proteomics. *Sci Adv* 6, eaax8978.
- (101) Cravatt, B. F., Wright, A. T., and Kozarich, J. W. (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* 77, 383–414.
- (102) Zhang, Y., Fonslow, B. R., Shan, B., Baek, M.-C., and Yates, J. R. (2013) Protein analysis by shotgun/bottom-up proteomics. *Chem Rev* 113, 2343–2394.
- (103) Wolters, D. A., Washburn, M. P., and Yates, J. R. (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* 73, 5683–5690.
- (104) Liu, Y., Patricelli, M. P., and Cravatt, B. F. (1999) Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci U S A* 96, 14694–14699.
- (105) Speers, A. E., and Cravatt, B. F. (2004) Chemical strategies for activity-based proteomics. *Chembiochem* 5, 41–47.
- (106) Speers, A. E., Adam, G. C., and Cravatt, B. F. (2003) Activity-based protein profiling in vivo using a copper(i)-catalyzed azide-alkyne 3 + 2 cycloaddition. *J Am Chem Soc* 125, 4686–4687.
- (107) Greenbaum, D., Medzihradszky, K. F., Burlingame, A., and Bogyo, M. (2000) Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chemistry & Biology* 7, 569–581.
- (108) Jeffery, D. A., and Bogyo, M. (2003) Chemical proteomics and its application to drug discovery. *Current Opinion in Biotechnology* 14, 87–95.
- (109) Hofmann, K., and Kiso, Y. (1976) An approach to the targeted attachment of peptides and proteins to solid supports. *Proc Natl Acad Sci U S A* 73, 3516–3518.
- (110) Huisgen, R. (1963) 1,3-Dipolar Cycloadditions. Past and Future. *Angew. Chem. Int. Ed. Engl.* 2, 565–598.

- (111) Rostovtsev, V. V., Green, L. G., Fokin, V. V., and Sharpless, K. B. (2002) A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective “ligation” of azides and terminal alkynes. *Angew Chem Int Ed Engl* 41, 2596–2599.
- (112) Tornøe, C. W., Christensen, C., and 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. *J Org Chem* 67, 3057–3064.
- (113) Schulze, W. X., and Usadel, B. (2010) Quantitation in mass-spectrometry-based proteomics. *Annu Rev Plant Biol* 61, 491–516.
- (114) Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* 389, 1017–1031.
- (115) Cox, J., Hein, M. Y., Lubner, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* 13, 2513–2526.
- (116) Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26, 1367–1372.
- (117) Ong, S.-E., and Mann, M. (2006) A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat Protoc* 1, 2650–2660.
- (118) Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 17, 994–999.
- (119) Boersema, P. J., Aye, T. T., van Veen, T. A. B., Heck, A. J. R., and Mohammed, S. (2008) Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates. *Proteomics* 8, 4624–4632.
- (120) Hsu, J.-L., Huang, S.-Y., Chow, N.-H., and Chen, S.-H. (2003) Stable-isotope dimethyl labeling for quantitative proteomics. *Anal Chem* 75, 6843–6852.
- (121) Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlett-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 3, 1154–1169.
- (122) Thompson, A., Schaefer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R. A. W., Mohammed, A. K. A., and Hamon, C. (2006) Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Anal Chem* 78, 4235.

- (123) Wasinger, V. C., Zeng, M., and Yau, Y. (2013) Current status and advances in quantitative proteomic mass spectrometry. *Int J Proteomics* 2013, 180605.
- (124) Lau, H.-T., Suh, H. W., Golkowski, M., and Ong, S.-E. (2014) Comparing SILAC- and stable isotope dimethyl-labeling approaches for quantitative proteomics. *J Proteome Res* 13, 4164–4174.
- (125) Hsu, J.-L., and Chen, S.-H. (2016) Stable isotope dimethyl labelling for quantitative proteomics and beyond. *Philos Trans A Math Phys Eng Sci* 374.
- (126) Scherer, W. F., Syverton J. T., and Gey, G. O. (1953) Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. *J Exp Med* 97, 695–710.
- (127) Peñuelas-Urquides, K., Villarreal-Treviño, L., Silva-Ramírez, B., Rivadeneyra-Espinoza, L., Said-Fernández, S., and León, M. B. de (2013) Measuring of Mycobacterium tuberculosis growth. A correlation of the optical measurements with colony forming units. *Braz J Microbiol* 44, 287–289.
- (128) Whitaker, W. B., Parent, M. A., Naughton, L. M., Richards, G. P., Blumerman, S. L., and Boyd, E. F. (2010) Modulation of responses of *Vibrio parahaemolyticus* O3:K6 to pH and temperature stresses by growth at different salt concentrations. *Appl Environ Microbiol* 76, 4720–4729.
- (129) Fairman, K., and Jacobson, B. S. (1983) Unique morphology of HeLa cell attachment, spreading and detachment from microcarrier beads covalently coated with a specific and non-specific substratum. *Tissue Cell* 15, 167–180.
- (130) Jayakumar, J., Kumar, V. A., Biswas, L., and Biswas, R. (2020) Therapeutic applications of lysostaphin against *Staphylococcus aureus*. *J Appl Microbiol*.
- (131) Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1, 376–386.
- (132) Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A. J. R. (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc* 4, 484–494.
- (133) Shaw, E., and Wuest, W. M. (2020) Virulence attenuating combination therapy: a potential multi-target synergy approach to treat *Pseudomonas aeruginosa* infections in cystic fibrosis patients. *RSC Med Chem* 11, 358–369.
- (134) Yayan, J., Ghebremedhin, B., and Rasche, K. (2015) Antibiotic Resistance of *Pseudomonas aeruginosa* in Pneumonia at a Single University Hospital Center in Germany over a 10-Year Period. *PLoS One* 10, e0139836.

- (135) Luo, C., Hu, G.-Q., and Zhu, H. (2009) Genome reannotation of *Escherichia coli* CFT073 with new insights into virulence. *BMC Genomics* 10, 552.
- (136) Jacobs, M. A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R., Chun-Rong, L., Guenther, D., Bovee, D., Olson, M. V., and Manoil, C. (2003) Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 100, 14339–14344.
- (137) Hauser, A. R. (2009) The type III secretion system of *Pseudomonas aeruginosa*: infection by injection. *Nat Rev Microbiol* 7, 654–665.
- (138) Wilhelm, M., Schlegl, J., Hahne, H., Gholami, A. M., Lieberenz, M., Savitski, M. M., Ziegler, E., Butzmann, L., Gessulat, S., Marx, H., Mathieson, T., Lemeer, S., Schnatbaum, K., Reimer, U., Wenschuh, H., Mollenhauer, M., Slotta-Huspenina, J., Boese, J.-H., Bantscheff, M., Gerstmair, A., Faerber, F., and Kuster, B. (2014) Mass-spectrometry-based draft of the human proteome. *Nature* 509, 582–587.
- (139) Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., Alm, T., Asplund, A., Björk, L., Breckels, L. M., Bäckström, A., Danielsson, F., Fagerberg, L., Fall, J., Gatto, L., Gnann, C., Hober, S., Hjelmare, M., Johansson, F., Lee, S., Lindskog, C., Mulder, J., Mulvey, C. M., Nilsson, P., Oksvold, P., Rockberg, J., Schutten, R., Schwenk, J. M., Sivertsson, Å., Sjöstedt, E., Skogs, M., Stadler, C., Sullivan, D. P., Tegel, H., Winsnes, C., Zhang, C., Zwahlen, M., Mardinoglu, A., Pontén, F., Feilitzén, K. von, Lilley, K. S., Uhlén, M., and Lundberg, E. (2017) A subcellular map of the human proteome. *Science* 356.
- (140) Boukamp, P., Petrussevska, R. T., Breitkreutz, D., Hornung, J., Markham, A., and Fusenig, N. E. (1988) Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. *J Cell Biol* 106, 761–771.
- (141) Ross, R. A., Spengler, B. A., and Biedler, J. L. (1983) Coordinate morphological and biochemical interconversion of human neuroblastoma cells. *J Natl Cancer Inst* 71, 741–747.
- (142) Wang, N., Dhumale, P., Chiang, J., and Püschel, A. W. (2018) The Sema3A receptor Plexin-A1 suppresses supernumerary axons through Rap1 GTPases. *Sci Rep* 8, 15647.
- (143) 2.3 In-vitro-Eigenschaften transformierter Zellen (2010). In *Molekulare Onkologie* (Wagener, C., and Müller, O., Eds.), Georg Thieme Verlag, Stuttgart.
- (144) Burrell, H. E., Wlodarski, B., Foster, B. J., Buckley, K. A., Sharpe, G. R., Quayle, J. M., Simpson, A. W. M., and Gallagher, J. A. (2005) Human keratinocytes release ATP and utilize three mechanisms for nucleotide interconversion at the cell surface. *J Biol Chem* 280, 29667–29676.

-
- (145) Backus, K. M., Correia, B. E., Lum, K. M., Forli, S., Horning, B. D., González-Páez, G. E., Chatterjee, S., Lanning, B. R., Teijaro, J. R., Olson, A. J., Wolan, D. W., and Cravatt, B. F. (2016) Proteome-wide covalent ligand discovery in native biological systems. *Nature* 534, 570–574.
- (146) Kleiner, P., Heydenreuter, W., Stahl, M., Korotkov, V. S., and Sieber, S. A. (2017) A Whole Proteome Inventory of Background Photocrosslinker Binding. *Angew. Chem. Int. Ed. Engl.* 56, 1396–1401.
- (147) Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 10, 1794–1805.
- (148) Hoegl, A., Nodwell, M. B., Kirsch, V. C., Bach, N. C., Pfanzelt, M., Stahl, M., Schneider, S., and Sieber, S. A. (2018) Mining the cellular inventory of pyridoxal phosphate-dependent enzymes with functionalized cofactor mimics. *Nat Chem* 10, 1234–1245.

6. Appendix

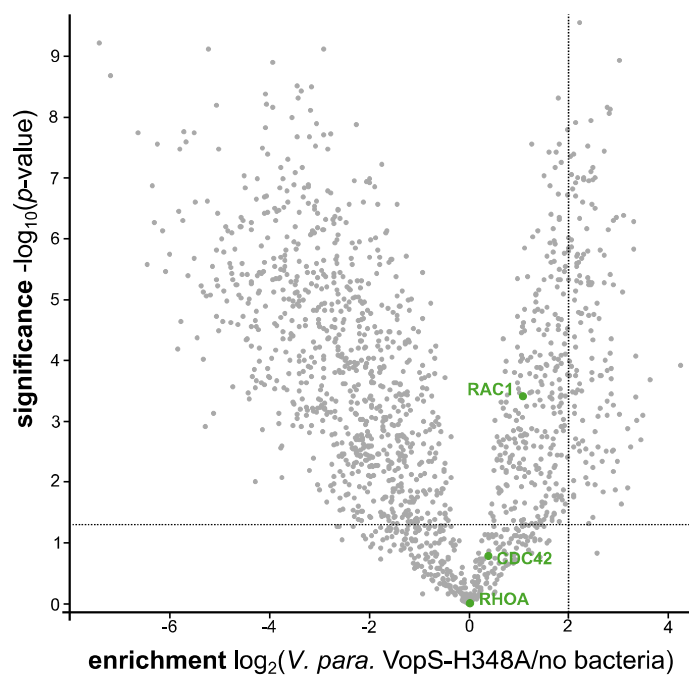


Figure S1: Volcano plot of the *in situ* infection experiment of HeLa cells with *V. parahaemolyticus* mutant VopS-H348A. All samples were treated with **pro-N6pA** (100 μ M) but only half of them infected with bacteria (approach A). MS analysis was performed *via* LFQ. The experiment was performed with 5 replicates, and after data processing 1548 proteins remained. Threshold criteria were set to \log_2 enrichment > 2 and p -value < 0.05 . The VopS target proteins RAC1, CDC42 and RHOA are marked in green.

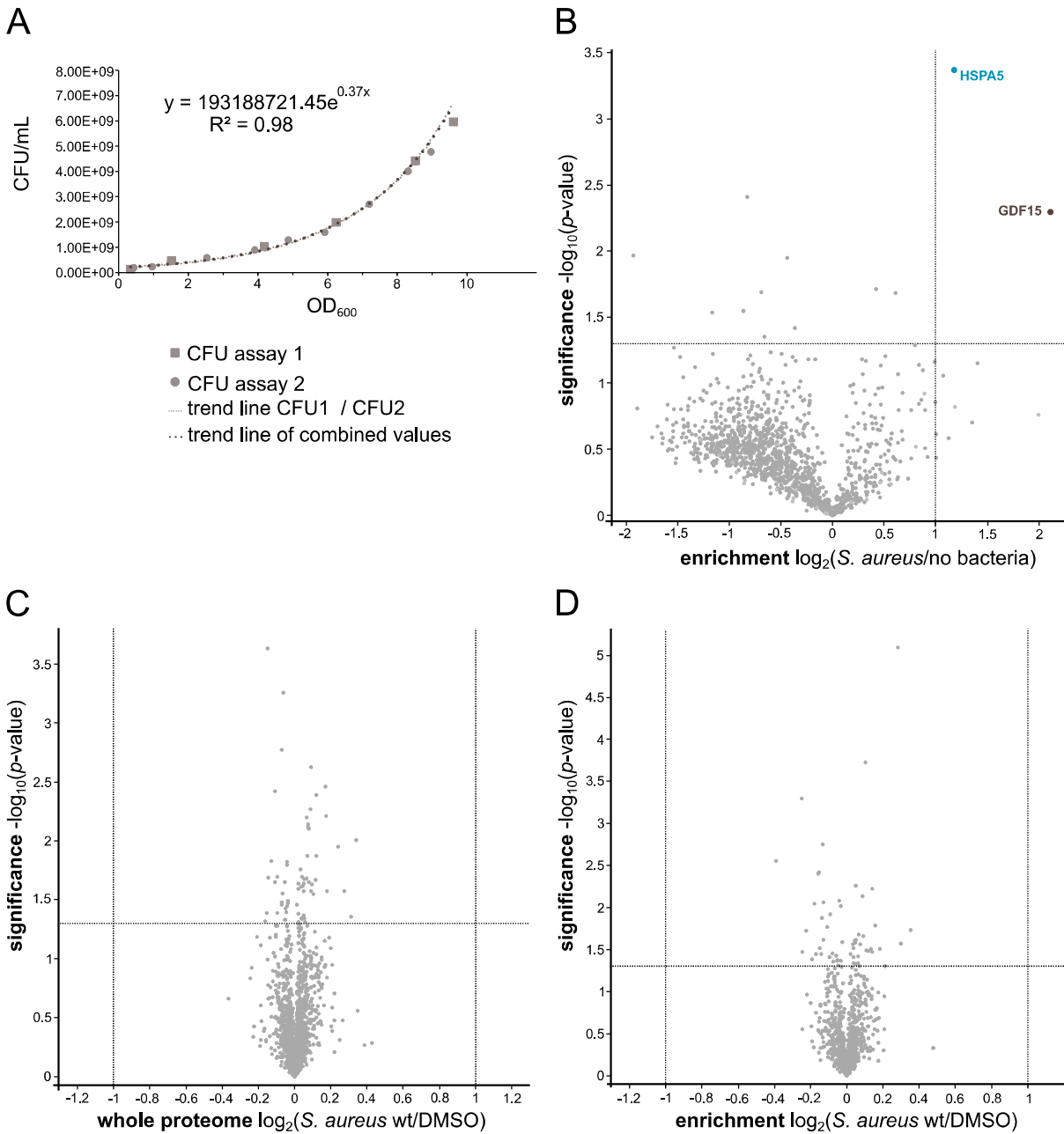
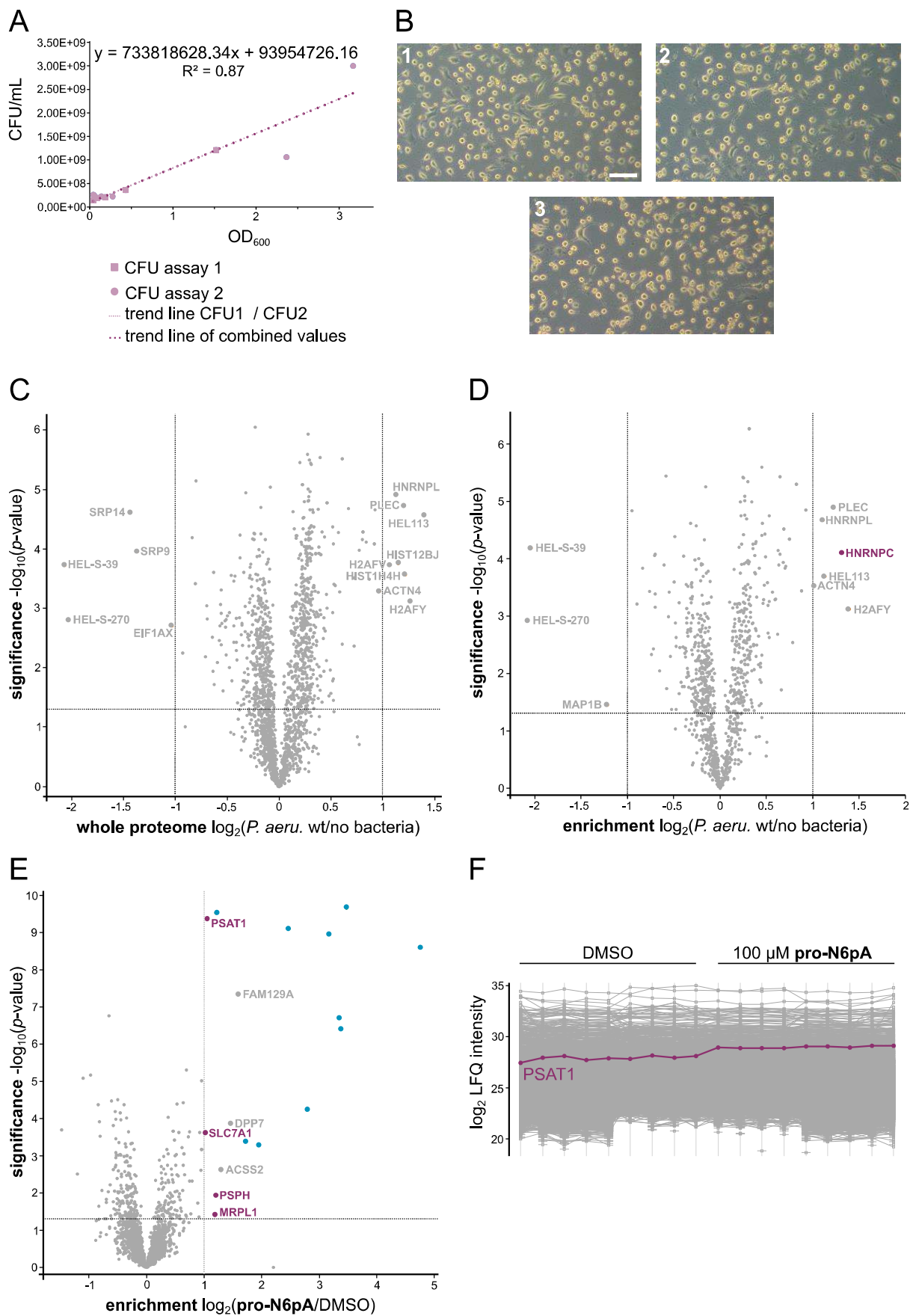


Figure S2: Identification of suitable infection conditions and enrichment data of infection experiments with the Gram-positive bacterium *S. aureus* USA300. (A) Correlation between OD₆₀₀ and the CFU count per mL of *S. aureus* USA300 when growing the bacteria in a 250 mL culture flask with baffles. The two independently performed CFU assays are accentuated in light brown, and data points can be assigned to the respective experiment based on their external shape. For each individual experiment, a trend line was calculated using an exponential fit. The trend line in dark brown was calculated by combining all measured data. The experimentally determined correlation between OD₆₀₀ and CFU/mL is indicated. (B) Volcano plot of HeLa cell infection with *S. aureus* USA300 (MOI 20), whereby all samples have been treated with **pro-N6pA** (100 μM) but only half of them with bacteria (approach A, n = 5, after data processing: 1350 proteins). Intrinsically AMPylated proteins are marked in blue, while possible human AMPylation targets are highlighted in brown. (C, D) Volcano plots of full proteome analysis (C, n = 6, after data processing: 1792 proteins) and protein enrichment experiment (D, n = 6, after data processing: 1157 proteins) of *S. aureus* USA300 infected HeLa cells with SILAC labeling. All samples were treated with 100 μM **pro-N6pA** but only half of them infected with bacteria (approach A).



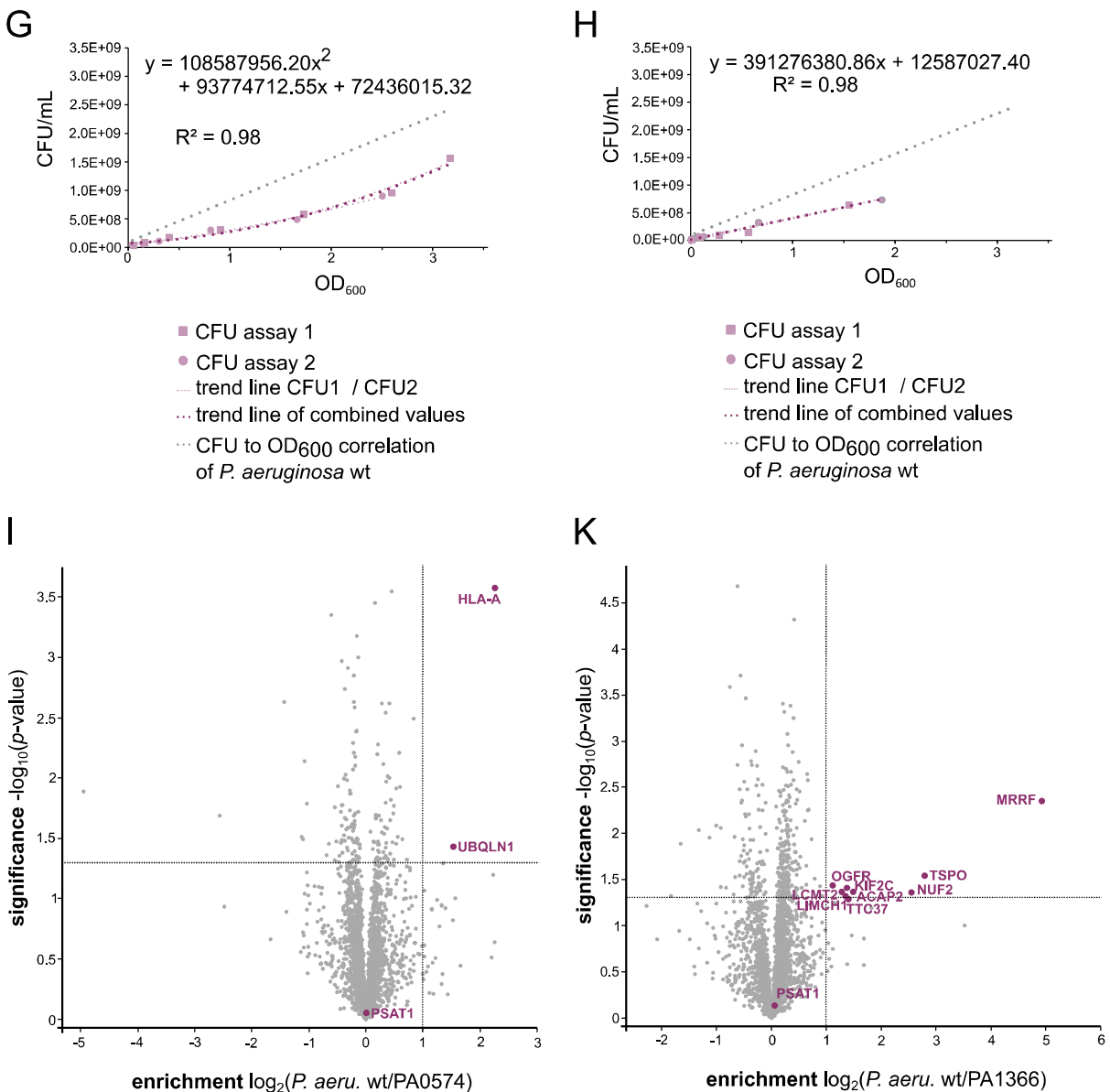
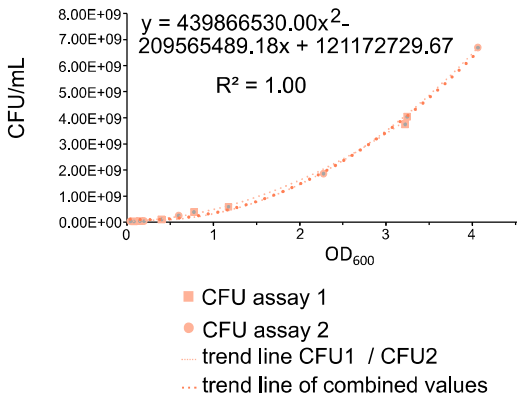


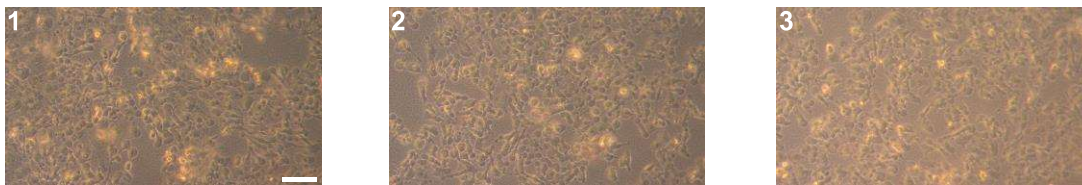
Figure S3: Identification of suitable infection conditions and enrichment data of infection experiments with the Gram-negative bacterium *P. aeruginosa* PAO1 and its two transposon mutants expressing the inactive AMPylators PW3486 genotype PA1366-A11::ISpho-A/hah (PA1366) as well as PW2059 genotype PA0574-A11::ISlacZ/hah (PA0574). (A, G, H) OD₆₀₀ to CFU/mL correlation of *P. aeruginosa* PAO1 (A), PA0574 (G) and PA1366 (H) when growing the bacteria in a 250 mL culture flask with baffles. The two independently performed CFU assays are accentuated in mauve, and data points can be assigned to the respective experiment based on their external shape. For each individual experiment, a trend line was calculated using a linear fit (wild type A; PA1366 H) or an exponential fit (PA0574 G), depending on what function reflected the measured data points best. The trend line in dark purple was calculated by combining all measured data. In the CFU assays of the transposon mutants (G, H), the fit of the wild type bacterium is shown in gray. The experimentally determined correlation between OD₆₀₀ and CFU/mL is indicated. (B) Phenotypic changes of HeLa cells after bacterial infection. (1) HeLa cells, 100 μM pro-N6pA (16 h), infected with *P. aeruginosa* PAO1 (MOI = 10 for 282 min). (2) HeLa cells, 100 μM pro-N6pA (16 h), infected with *P. aeruginosa* PA0574 (MOI = 10 for 281 min) or (3) HeLa cells, 100 μM pro-N6pA (16 h), *P. aeruginosa* PA1366 (MOI = 10 for 273 min). The scale bar represents 100 μm. (C, D) Volcano plots of full proteome analysis (C, n = 6, after data processing: 1755 proteins) and protein enrichment experiment (D, n = 6, after data processing: 949 proteins) of *P. aeruginosa* PAO1 infected HeLa cells with SILAC labeling. Proteins highlighted in gray are up- or down-regulated during full proteome analysis. Possible AMPylation targets are shown in mauve. All samples were treated with 100 μM pro-N6pA but only half of them

infected with bacteria (approach A). **(E, F)** Volcano plot and profile plot of HeLa cell infection with *P. aeruginosa* wild type (n = 9; after data processing: 2159 proteins). All samples were infected with bacteria while only half of them were treated with 100 μ M **pro-N6pA** (approach B). The analysis was performed by LFQ. **(E)** Intrinsically AMPylated proteins are highlighted in blue for a better overview, without indication of their name. Possible AMPylation targets are shown in mauve. Proteins that have been enriched in infection experiments with other bacterial strains and thus most likely represent background binders are shown in gray. **(F)** Profile plot of log₂ intensities of PSAT1. **(I, K)** Volcano plots of HeLa cell infection with *P. aeruginosa* wild type versus transposon mutant PA0574 **(I, n = 5, after data processing: 3236 proteins)** or PA1366 **(K, n = 5, after data processing: 3068 proteins)**. All samples were treated with 100 μ M **pro-N6pA** (approach C). Possible AMPylation targets are shown in mauve. Partially adapted from Rauh *et al.*⁸⁰

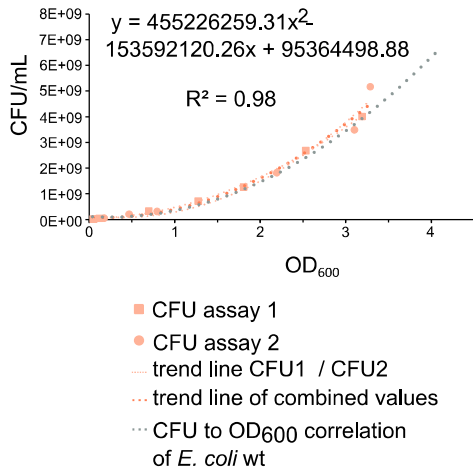
A



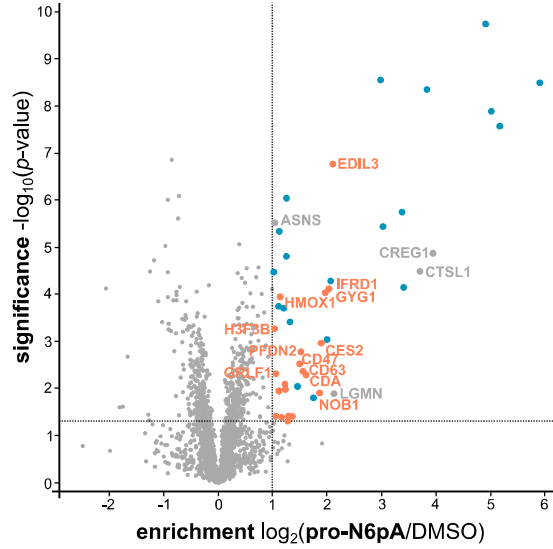
B



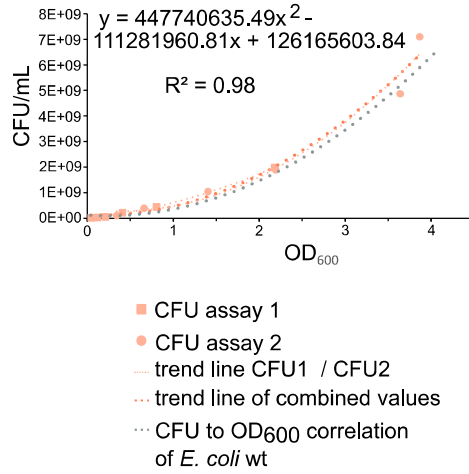
D



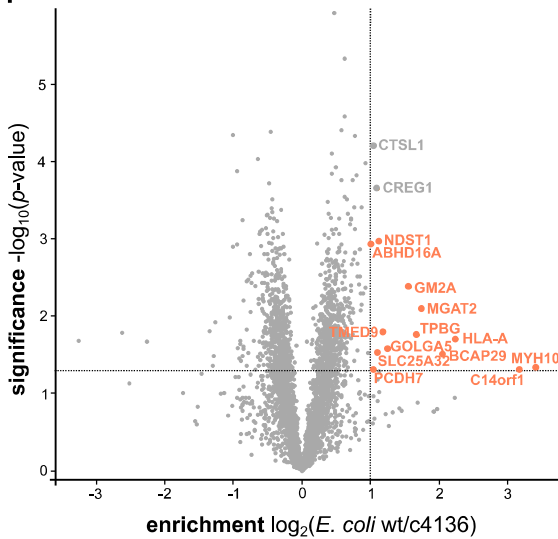
C



E



F



G

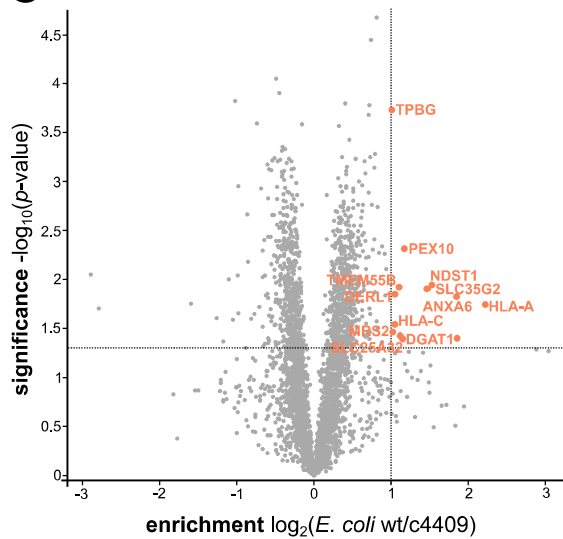


Figure S4: Identification of suitable infection conditions and enrichment data of infection experiments with the Gram-negative bacterium *E. coli* CFT073 and its two in frame deletion mutants c4136::Km and c4409::Km expressing inactive AMPylators. (A, D, E) Correlation between OD₆₀₀ and the CFU count per mL of *E. coli* CFT073 wild type (A), c4136::Km (D) and c4409::Km (E) when growing the bacteria in a 100 mL culture flask without baffles. The two independently performed CFU assays are accentuated in light orange, and data points can be assigned to the respective experiment based on their external shape. For each individual experiment, a trend line was calculated using a second degree polynomial fit. The trend line in dark orange was calculated by combining all measured data. In the CFU assays of in frame deletion mutants (D, E), the trend line of the wild type bacterium is shown in gray. The experimentally determined correlation between OD₆₀₀ and CFU/mL is indicated. (B) Phenotypic changes of HeLa cells after *E. coli* infection. (1) HeLa cells, 100 μM **pro-N6pA** (16 h), infected with *E. coli* CFT073 wild type (MOI = 10 for 142 min). (2) HeLa cells, 100 μM **pro-N6pA** (16 h), infected with *E. coli* c4136::Km (MOI = 10 for 142 min) or (3) HeLa cells, 100 μM **pro-N6pA** (16 h), *E. coli* c4409::Km (MOI = 10 for 144 min). The scale bar represents 100 μm. (C) Volcano plot of HeLa cell infection with *E. coli* wild type (n = 5; after data processing: 2506 proteins). All samples were infected with bacteria while only half of them were treated with 100 μM **pro-N6pA** (approach B). The analysis was performed by LFQ. Intrinsically AMPylated proteins are highlighted in blue. Possible AMPylation targets are shown in orange. Proteins that most likely represent background binders are shown in gray. (F, G) Volcano plots of HeLa cell infection with *E. coli* wild type versus in frame deletion mutants c4136::Km (F, n = 5, after data processing: 3372 proteins) or c4409::Km (K, n = 5, after data processing: 3375 proteins). All samples were treated with 100 μM **pro-N6pA** (approach C). Possible AMPylation targets are shown in orange. Partially adapted from Rauh *et al.*⁸⁰

7. Licenses

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Feb 16, 2021

This Agreement between Ms. Theresa Rauh ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number 5010640162134

License date Feb 16, 2021

Licensed Content Publisher John Wiley and Sons

Licensed Content Publication FEBS Letters

Licensed Content Title Post-translational modifications in host cells during bacterial infection

Licensed Content Author Pascale Cossart, David Ribet

Licensed Content Date May 20, 2010

Licensed Content Volume 584

Licensed Content Issue 13

Licensed Content Pages 11

Type of use Dissertation/Thesis

Requestor type University/Academic

Format Print and electronic

Portion Figure/table

Number of figures/tables 1

Will you be translating? No

Title An MS-based chemical proteomic strategy to elucidate bacterial manipulation of the human proteome by protein AMPylation

Institution name Technische Universität München

Expected presentation date May 2021

Order reference number 4982400470447

Portions Figure 1

Requestor Location Ms. Theresa Rauh
Lichtenbergstraße 4
Garching, 85748
Germany
Attn: Ms. Theresa Rauh

Publisher Tax ID EU826007151

Total 0.00 EUR

Terms and Conditions

Copyright Clearance Center RightsLink® Home ? Email Support Theresa Rauh

Molecular Perspectives on Protein Adenylation
Author: Christian Hedberg, Aymelt Itzen
Publication: ACS Chemical Biology
Publisher: American Chemical Society
Date: Jan 1, 2015
Copyright © 2015, American Chemical Society

Quick Price Estimate

This service provides permission for reuse only. If you do not have a copy of the portion you are using, you may copy and paste the content and reuse according to the terms of your agreement. Please be advised that obtaining the content you license is a separate transaction not involving RightsLink.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

Note: Individual Scheme and Structure reuse is free of charge and does not require a license. If the scheme or structure is identified as a "Figure" in the article, permission is required.

I would like to... reuse in a Thesis/Dissertation Format Print and Electronic

Requestor Type Non-profit Select your currency EUR - €

Portion Table/Figure/Micrograph Quick Price Click Quick Price

Number of Table/Figure/Micrographs 1

QUICK PRICE CONTINUE

To request permission for a type of use not listed, please contact the publisher directly.

Copyright Clearance Center RightsLink® Home ? Email Support Theresa Rauh

Molecular Perspectives on Protein Adenylation
Author: Christian Hedberg, Aymelt Itzen
Publication: ACS Chemical Biology
Publisher: American Chemical Society
Date: Jan 1, 2015
Copyright © 2015, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION), Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.
- If credit is given to another source for the material you requested, permission must be obtained from that source.

BACK CLOSE WINDOW

Order Number: 1097764
Order Date: 16 Feb 2021

Payment Information

Theresa Rauh theresa.rauh@tum.de Payment method: Invoice	Billing Address: Ms. Theresa Rauh Lichtenbergstraße 4 Garching, 85748 Germany +49 8928913676 theresa.rauh@tum.de	Customer Location: Ms. Theresa Rauh Lichtenbergstraße 4 Garching, 85748 Germany
--	---	--

Order Details

1. The Journal of biological chemistry Billing Status: Open
Article: AMPylation of Rho GTPases subverts multiple host signaling processes.
Order license ID 1097764-1
Order detail status Completed
ISSN 0021-9258
Type of use Republish in a thesis/dissertation
Publisher AMERICAN SOCIETY FOR BIOCHEMISTRY AND MOL...
Portion Chart/graph/table/figure 0,00 EUR
Republication Permission

LICENSED CONTENT

Publication Title	The Journal of biological chemistry	Rightsholder	American Soc for Biochemistry & Molecular Biology
Article Title	AMPylation of Rho GTPases subverts multiple host signaling processes.	Publication Type	Journal
Author/Editor	AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH, AMERICAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BI	Start Page	32977
Date	01/01/1905	End Page	32988
Language	English	Issue	47
Country	United States of America	Volume	289

REQUEST DETAILS

Portion Type	Chart/graph/table /figure	Distribution	Worldwide
Number of charts / graphs / tables / figures requested	1	Translation	Original language of publication
Format (select all that apply)	Print,Electronic	Copies for the disabled?	No
Who will republish the content?	Academic institution	Minor editing privileges?	Yes
Duration of Use	Life of current and all future editions	Incidental promotional use?	No
Lifetime Unit Quantity	Up to 499	Currency	EUR
Rights Requested	Main product		

NEW WORK DETAILS

Title	An MS-based chemical proteomic strategy to elucidate bacterial manipulation of the human proteome by protein AMPylation	Institution name	Technical University of Munich
Instructor name	Prof. Dr. Stephan A. Sieber	Expected presentation date	2021-03-31

ADDITIONAL DETAILS

The requesting person / organization to appear on the license
TUM - Theresa Rauh

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Figure 8	Title of the article/chapter the portion is from	AMPylation of Rho GTPases subverts multiple host signaling processes.
Editor of portion(s)	Woolery, Andrew R; Yu, Xiaobo; LaBaer, Joshua; Orth, Kim	Author of portion(s)	Woolery, Andrew R; Yu, Xiaobo; LaBaer, Joshua; Orth, Kim
Volume of serial or monograph	289	Issue, if republishing an article from a serial	47
Page or page range of portion	32977-32988	Publication date of portion	2014-11-21

MS-Based in Situ Proteomics Reveals AMPylation of Host Proteins during Bacterial Infection

Author: Theresa Rauh, Sophie Brameyer, Pavel Kielkowski, et al
Publication: ACS Infectious Diseases
Publisher: American Chemical Society
Date: Dec 1, 2020
 Copyright © 2020, American Chemical Society

Quick Price Estimate

This service provides permission for reuse only. If you do not have a copy of the portion you are using, you may copy and paste the content and reuse according to the terms of your agreement. Please be advised that obtaining the content you license is a separate transaction not involving RightsLink.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

Note: Individual Scheme and Structure reuse is free of charge and does not require a license. If the scheme or structure is identified as a "Figure" in the article, permission is required.

Permission for this particular request is granted for print and electronic formats, and translations, at no charge. Figures and tables may be modified. Appropriate credit should be given. Please print this page for your records and provide a copy to your publisher. Requests for up to 4 figures require only this record. Five or more figures will generate a printout of additional terms and conditions. Appropriate credit should read: "Reprinted with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.

If credit is given to another source for the material you requested, permission must be obtained from that source.

I would like to... **Format**

Requestor Type **Select your currency**

Portion **Quick Price**

Number of Table/Figure /Micrographs

QUICK PRICE **CONTINUE**

To request permission for a type of use not listed, please contact the publisher directly.

MS-Based in Situ Proteomics Reveals AMPylation of Host Proteins during Bacterial Infection

Author: Theresa Rauh, Sophie Brameyer, Pavel Kielkowski, et al
Publication: ACS Infectious Diseases
Publisher: American Chemical Society
Date: Dec 1, 2020
 Copyright © 2020, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK

CLOSE WINDOW

