Organization and functional complex formation within the biosynthetic machinery of glycopeptide antibiotics

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard-Karls-Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Dardan Beqaj

aus Studençan/Kosovo

Tübingen

2025

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	10.07.2025
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter/-in:	apl. Prof. Dr. Evi Stegmann
2. Berichterstatter/-in:	Prof. Dr. Harald Gross
3. Berichterstatter/-in:	Prof. Dr. Wolfgang Wohlleben
4. Berichterstatter/-in:	Prof. Dr. Andreas Bechthold

Erklärung

Hiermit erkläre ich, dass ich diese Schrift selbständig und nur mit den angegebenen Hilfsmitteln verfasst habe. Ferner sind alle Stellen, die im Wortlaut oder dem Sinn nach den Werken anderer Autoren entnommen sind, durch Angabe der Quellen kenntlich gemacht. Eine detaillierte Abgrenzung meiner eigenen Leistungen von den Beiträgen meiner Kooperationspartner habe ich im Anhang vorgenommen.

……..

Tübingen, den 05.05.2025

To my wife, Selvije and my children Henor and Juna.

Acknowledgments

I would like to express my deepest gratitude to everyone who has played a significant role in making this thesis possible. First and foremost, I extend my heartfelt thanks to my supervisor, Evi Stegmann, for her invaluable guidance, support, and encouragement throughout my entire PhD research and for the opportunity to work on this project. I feel truly fortunate to have had such outstanding supervision. Her mentorship over these past years shaped my perspective on science and strengthened my passion for microbiology.

I am also sincerely grateful to my second supervisor, Wolfgang Wohlleben, for his unwavering support and exceptional expertise. His insightful guidance has helped me to critically assess and strengthen my work. It has been a privilege to learn from him and benefit from his vast knowledge of natural products and the biology of actinomycetes. I want to thank also Harald Gross, for his willingness to write the second review for this thesis.

I was incredibly lucky to pursue this journey in the Stegmann Lab. It has provided an exceptional working environment and fun with all the people that I had the chance to work with. I extend my deepest thanks to Jens, Naybel, Alena, Demi, Elisa, Luca, Irina, Erik, Luca Puschmann, Anna, Alessa and Sigrid for being exceptional coworkers. To Franzi, Yvonne and Davide, thanks for all the breakfasts together with me, Jens and Alena. It helped so much to start the hard PhD days with such a great company. And thanks to all the people from the department of Microbial active compounds, for the great working environment.

This work wouldn’t be the same without the outstanding cooperation I had. I want to thank Samuel Wagner, for all the suggestions and helpful comments and driving me into the world of blue native PAGE. Thanks to Nicola, for the fantastic work we did together and her continuous support during all experiments. To Thilo Stehle and Nina for all the comments and help on docking domains and their insights into structural biology.

I would also like to thank all my former students and HiWis: Denijel, Jule, Maja, Andreas, Meike and Sylvia for their support, which contributed to finishing this work successfully.

A big thanks also to Andreas Kulik, who supported my work with his exceptional expertise in HPLC-MS. To Günther, for his introduction into the bacterial two-hybrid

system, critical comments during all my talks and suggestions to improve my research work. I want to thank also Aysun and Melanie, for their support in administrative tasks and being always ready for help. To Thomas Härtner, for supporting me and our lab for everything we needed during the working days.
This work wouldn’t be possible without the support of my whole family and friends. Thanks to my parents and siblings who always believed in me and having time to listen to me when I needed it and to my friends for being supportive all the time.
Lastly and most, I want to express my deepest gratitude to my wife Selvije, whose love and support for me can’t be expressed in words. She was always motivating me to do better and inspired me with her personality and courage. I want to give my last and greatest thanks to my kids, Henor and Juna, who gifted me the best feelings of being a father and experiencing unconditional love during this time. They taught me the meaning of why I am doing this.

To all of you, thank you for being such great people.

Table of Contents

A. Abbreviations

°C	degree Celsius
µM	Micromolar
A. balhimycina	*Amycolatopsis balhimycina*
AA	Amino acid
B2H	Bacterial two-hybrid assay
Bal	Balhimycin
bp	Base pair
cDD	C-terminal docking domain
DD	Docking domain
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
HPLC	High performance liquid chromatography
kb	Kilobase
m/z	Mass/charge ratio
MLP	MbtH-like protein
nDD	N-terminal docking domain
NRPS	Non-ribosomal peptide synthetase
PCR	Polymerase chain reaction
TID	TurboID
Van	Vancomycin
WT	Wildtype

B. Zusammenfassung

Glykopeptid-Antibiotika (GPAs) sind Naturstoffe, die von Aktinomyzeten produziert werden und sich basierend auf ihrer chemischen Diversität in fünf Typen (I-V) unterteilen. Aufgrund ihrer Wirksamkeit gegen multiresistente grampositive Bakterien sind sie von besonderem Interesse, mit Vancomycin und Teicoplanin in der klinischen Anwendung. Die Biosynthese von Balhimycin, einem Vancomycin-ähnlichen GPA, dient als Modellsystem, an dem zentrale Erkenntnisse zur GPA-Biosynthese gewonnen wurden. Obwohl die wesentlichen Schritte dieser Biosynthese aufgeklärt sind, ist das Zusammenspiel der beteiligten Enzyme mit den Exportmechanismen bislang nur unzureichend verstanden. Dieses Zusammenspiel stellt einen aktuellen Forschungsschwerpunkt dar, um die molekularen Grundlagen der GPA-Produktion besser zu verstehen.

Die vorliegende Arbeit widmet sich einer umfassenden Untersuchung der GPA-Biosynthese in *Amycolatopsis balhimycina*, mit besonderem Fokus auf der Kopplung von Biosynthese und Export während der Balhimycin-Produktion. In zwei miteinander verknüpften Ansätzen wurden zum einen die Funktion des ABC-Transporters Tba beim Export von Balhimycin sowie dessen Einfluss auf die Biosynthese analysiert. Zum anderen erfolgte eine detaillierte Charakterisierung der molekularen Interaktionen innerhalb der Biosynthese-Maschinerie.

Im ersten Teil der Arbeit wurden die Substratspezifität von Tba sowie seine funktionelle Beteiligung an Transportprozessen und der Organisation des Balhimycin-Biosynthesekomplexes analysiert. Nach Deletion des *tba*-Gens zeigte sich eine signifikante Reduktion der Balhimycin-Produktion, was sich in gesenkten intra- und extrazellulären Konzentrationen widerspiegelte. Die Wiedereinführung des nativen *tba*-Gens stellte die Produktion wieder her. Heterolog exprimierte Transporter-Gene aus anderen GPA-Typen konnten die Deletion nicht komplementieren. Zudem wurde festgestellt, dass Modifikationen wie Glykosylierung und Chlorierung die Exporteffizienz nicht beeinflussen, wobei die Glykosylierung dennoch für die antibakterielle Aktivität unverzichtbar bleibt. Durch MD-Simulation konnte bestätigt werden, dass die Substratspezifität vom Heptapeptid-Rückgrat des GPAs bestimmt wird und dessen Modifikationen wenig Einfluss darauf haben. Mittels „Proximity-dependent Labelling“ konnten Hinweise auf eine direkte Interaktion von Tba mit der Balhimycin Biosynthesemaschinerie gewonnen werden. Dies legt nahe, dass Tba nicht nur am Export von Balhimycin beteiligt ist, sondern möglicherweise auch eine strukturelle Rolle eines biosynthetischen Mikrokompartiments übernimmt.

Im zweiten Teil der Arbeit wurden Protein-Protein-Wechselwirkungen innerhalb der Balhimycin-Biosynthese untersucht. Im Zentrum standen dabei die nicht-ribosomalen Peptidsynthetasen (NRPS) BpsA-C, die das Heptapeptid-Rückgrat zusammensetzen, sowie

deren Modifikationsenzyme. Wir konnten spezifische Docking-Domänen definieren, die für die Interaktion zwischen den NRPS entscheidend sind. Funktionsanalysen durch Deletionsexperimente und in vitro-Bindungsstudien zeigten, dass diese Domänen eine zentrale Rolle für die Stabilität des NRPS-Komplexes spielen. Darüber hinaus zeigten sogenannte MbtH-like-Proteine (MLPs), die als Kofaktoren für die Adenylierungsdomänen von NRPS wirken, Interaktionen mit den einzelnen NRPS-Modulen. Gen-Deletionsanalysen zeigten, dass im Genom von *A. balhimycina* mehrere homologe MLP-Gene vorhanden sind, die sich in ihrer Funktion gegenseitig ersetzen können. Mittels AlphaFold wurden Proteinkomplexe der MLP-NRPS Interaktion vorhergesagt, welche die A-domänen als Bindungsstelle der MLP zeigte. Diese Modelle dienten zugleich der Bestimmung der Interaktionsfläche sowie der daran beteiligten Aminosäuren. Durch Mutationsanalysen konnten Aminosäuren identifiziert werden, welche für die Bildung der Interaktionsfläche zuständig sind.

Auch die Interaktionen zwischen den Modifikationsenzymen und der NRPS-Biosynthesemaschinerie wurden untersucht. Die Cytochrom-P450-Monooxygenasen OxyA und OxyC, die oxidative Verknüpfungen von Aminosäuren katalysieren, interagierten mit der X-Domäne der NRPS. Die X-Domäne ist als Recruiting-Domäne für die Oxygenasen charakterisiert worden. Für OxyB konnte hingegen keine entsprechende Interaktion nachgewiesen werden. Native PAGE-Analysen und „Proximity-dependent Labelling“-Experimente bestätigten darüber hinaus eine enge räumliche Anordnung der NRPS mit der Halogenase BhaA, die die Chlorierung von β-Hydroxytyrosin während der Peptidsynthese an der NRPS katalysiert.

Zusammenfassend zeigt diese Arbeit, dass die Biosynthese von Balhimycin ein hochgradig regulierter und räumlich organisierter Prozess ist, bei dem Export, Peptidzusammenbau und Modifikationen durch ein Netzwerk spezifischer Protein-Protein-Interaktionen miteinander gekoppelt sind. Dieses Modell liefert wertvolle neue Einblicke in die molekularen Mechanismen der GPA-Produktion und bildet zugleich eine Grundlage für die gezielte biotechnologische Optimierung von GPA-Biosynthesewegen.

C. Summary

Glycopeptide antibiotics (GPAs) are natural products produced by actinomycetes and are classified into five types according to their chemical diversity. Clinically utilized glycopeptide antibiotics, vancomycin and teicoplanin, are of particular interest due to their activity against multi-resistant Gram-positive bacteria. The biosynthesis of balhimycin, a vancomycin-type GPA, represents a model system that provided important insights into GPA biosynthesis. While the catalytic steps of the biosynthesis have been described, the interplay between the enzymes involved and the export mechanism is still poorly understood. This relation now marks a crucial research topic focused on elucidating the molecular basis of GPA biosynthesis.

The present work focuses on an in-depth examination of GPA biosynthesis in *Amycolatopsis balhimycina*, emphasizing the integration of biosynthesis and export during balhimycin production. The function of the ABC transporter Tba in balhimycin export and its impact on biosynthesis were examined by two interconnected approaches, along with a systematic characterization of the molecular interactions within the biosynthetic machinery.

The first part of the work investigated Tba's substrate specificity, its functional role in transport processes, and its impact on the organization of the balhimycin biosynthesis complex. After the deletion of the *tba* gene, a significant decrease in balhimycin production has been observed, as indicated by reduced intra- and extracellular levels. While the integration of the native *tba* gene recovered production, transporter genes expressed heterologously from different GPA types failed to complement the gene deletion. Modifications like glycosylation and chlorination were shown to have no impact on export efficiency, whereas glycosylation was crucial for antibacterial action. Molecular dynamics simulations confirmed that substrate selectivity is dictated by the heptapeptide backbone of the GPA, while modifications show no impact. Proximity-dependent labeling provided evidence for a putative interaction between Tba and the balhimycin biosynthetic machinery. This indicates that Tba may be involved not only in the export of balhimycin but also in having a functional role within a biosynthetic microcompartment.

The second part of this work investigated the protein-protein interactions involved in balhimycin biosynthesis. The emphasis was on the non-ribosomal peptide synthetases (NRPS) BpsA-C, responsible for assembling the heptapeptide backbone, together with associated modification enzymes. Distinct docking domains crucial for interactions between the NRPS modules have been identified. Functional analysis via deletion experiments and in vitro binding experiments demonstrated that these domains are crucial for the stability of the NRPS complex. Furthermore, MbtH-like proteins (MLPs), which serve as cofactors for the adenylation domains of NRPSs, have been shown to interact with NRPS modules. Gene deletions revealed several

homologous MLP genes are present in the *A. balhimycina* genome, capable of functionally compensating for each other. Using AlphaFold, protein complexes of the MLP-NRPS interactions were predicted, showing the A-domains as binding sites for the MLPs. These models allowed the identification of interaction interfaces and the key amino acids, which were further validated using mutational analysis and bacterial two-hybrid assays.

The interactions between the modification enzymes and the NRPS biosynthetic machinery were further investigated. The cytochrome P450 monooxygenases OxyA and OxyC, which facilitate oxidative crosslinking of amino acids, interacted with the X-domain of the NRPS, identified as a recruitment domain for these oxygenases. No related interaction could be found for OxyB. Native PAGE analysis and proximity-dependent labeling experiments corroborated the intimate spatial arrangement of the NRPSs with the halogenase BhaA, which catalyzes the chlorination of β-hydroxytyrosine during peptide synthesis on the NRPS.

The present work demonstrates that balhimycin biosynthesis is a strictly regulated and spatially organized process in which export, peptide assembly, and modifications are interlinked via a network of unique protein-protein interactions. This model offers significant new insights into the molecular mechanism of GPA production and provides a basis for the targeted biotechnological improvement of GPA biosynthetic pathways.

D. List of publications and manuscripts

Primary research articles:

1. “Unveiling the substrate specificity of the ABC transporter Tba and its role in glycopeptide biosynthesis”. *iScience*, 2025.

Nicola Gericke*, **Dardan Beqaj***, Thales Kronenberger, Andreas Kulik, Athina Gavriilidou, Mirita Franz-Wachtel, Ulrich Schoppmeier, Theresa Harbig, Johanna Rapp, Iwan Grin, Nadine Ziemert, Hannes Link, Kay Nieselt, Boris Macek, Wolfgang Wohlleben, Evi Stegmann, Samuel Wagner.

*Shared first authors and equal contribution

In the text, referred to as publication 1

This thesis also contains the following unpublished manuscript:

1. Molecular interplay of the glycopeptide antibiotics biosynthetic machinery
Dardan Beqaj*, Nina Pfahler, Samuel Wagner, Thilo Stehle, Mirita Franz-Wachtel, Boris Macek, Wolfgang Wohlleben, Evi Stegmann#

*First author

In the text, referred to as manuscript 1

Publications I contributed to but not included in this thesis:

1. MIBiG 4.0: advancing biosynthetic gene cluster curation through global collaboration, Nucleic Acids Research, Volume 53, Issue D1, 6 January 2025, Pages D678–D690,

Mitja M Zdouc, Kai Blin, Nico L L Louwen,…..,Christine Beemelmanns, **Dardan Beqaj**, Tim Berger,….., Tilmann Weber, Marnix H Medema

1. Introduction

1.1 Natural products as a source of antibiotics

Natural products (NPs), derived from microorganisms and plants, are bioactive compounds characterized by an exceptional degree of chemical diversity and structural complexity. These compounds play crucial ecological roles, serving as defense agents, signaling molecules, and metabolic regulators within ecosystems. They have been a prolific source for the development of pharmaceutical agents, with an estimated 60-70% of current drugs either derived from or inspired by natural products. This immense potential has fueled the ongoing search for new NPs, particularly in the battle against antibiotic resistance and diseases where treatment options remain limited [1].

NPs with antibiotic activity can originate from various biosynthetic pathways and thus are divided into different classes, including polyketides (PKs), ribosomally synthesized and post-translationally modified peptides (RiPPs), and non-ribosomal peptides (NRPs). Polyketides encompass antibiotics such as erythromycin, which is notable for its modular biosynthetic pathways that permit a wide range of structural modifications derived from Coenzyme A (CoA) activated acyl group precursors such as acetyl-CoA or malonyl-CoA [2]. RiPPs, as their name suggests, are peptides derived from the primary route of polypeptide biosynthesis, the ribosomes, followed by intricate post-translational modifications, with nisin being one of the most prominent examples exhibiting antibiotic activity [3].

NRPs are particularly significant, as they are synthesized by large, multi-enzyme complexes known as non-ribosomal peptide synthetases (NRPSs). Their ability to utilize not only proteinogenic amino acids but also unusual non-proteinogenic amino acids as precursors enables simultaneous peptide biosynthesis alongside the ribosomes, resulting in the generation of chemically diverse peptides with varying properties. NRPSs produce a wide array of complex NPs, including the first-generation glycopeptide antibiotics (GPAs) vancomycin and teicoplanin, which are invaluable in clinical settings due to their potent antibiotic activity. Telavancin, dalbavancin, and oritavancin are the latest FDA-approved GPAs, representing the second generation of semi-synthetically modified GPAs [4,5].

1.2 The genus *Amycolatopsis* as a promising source of glycopeptide antibiotics

The order Actinomycetes, which includes the genera *Amycolatopsis* and *Streptomyces*, represents a highly promising source of natural products and potential new drug candidates with antibacterial activity [6]. These organisms grow through mycelial development, typically differentiating into two types: substrate mycelium and aerial mycelium. The aerial mycelium is crucial for bacterial dispersal, as it produces spores. Actinomycetes are characterized by a high guanine-cytosine (G+C) content, typically ranging from 60% to 70% [7].

The genus *Amycolatopsis* is particularly noteworthy due to its genetic potential and established ability to produce NPs with various applications, including antibiotics. Over the past decades, *Amycolatopsis* species have been recognized as prolific producers of GPAs, which have been extensively studied and characterized. The first isolated GPA, vancomycin, is produced by A. orientalis and, even after 60 years of use, remains one of the last-resort antibiotics for treating infections caused by gram-positive bacteria, such as *Clostridium difficile* and methicillin-resistant *Staphylococcus aureus* (MRSA) [5,8,9]. Further examples include balhimycin from *A. balhimycina* [10], ristomycin from *A. japonicum* [11], avoparcin from *A. coloradensis* [12] and keratinimicin/keratinicyclin from *A. keratiniphila* [13]. Beyond GPAs the biosynthetic capabilities of *Amycolatopsis* species exhibit remarkable biosynthetic versatility, producing a range of chelators such as [*S,S*]-EDDS [14,15], albachelin [16], amychelin [17], ECO-0501 [18] and rifamycin [19]. This diverse biosynthetic repertoire highlights *Amycolatopsis* as a promising source for the discovery of novel NPs with potential biotechnological and pharmaceutical applications.

1.3 Glycopeptide antibiotics

1.3.1 Chemical diversity

GPAs are characterized by a heptapeptide backbone composed of both proteinogenic and non-proteinogenic amino acids (AAs), designated as AA1-7 (Figure 1). This heptapeptide backbone acquires a rigid, cup-shaped structure through cross-links formed by aromatic side chains. Furthermore, the backbone can be modified with halogens, methyl groups, sulfur, acyl chains, and glycosyl moieties [20,21].

The structural characteristics of the classical GPAs serve as the primary criterion for their classification, leading to their current grouping into four main types (Type I-IV) [22] (Figure 1). However, recent studies have suggested reclassifying Type V GPAs as GPA-related peptides

(GRPs) due to significant differences in the phylogenetic relationships, structure and mechanism of action compared to Types I-IV [23–25]. Consequently, Type V GPAs will not be further discussed in this thesis.

Distinct structural variations in the heptapeptide backbone define the different GPA types (I-IV). In Type I GPAs, such as vancomycin and balhimycin, aliphatic AAs occupy positions 1 and 3, whereas Types II (avoparcin), III (ristocetin A), and IV (teicoplanin) contain aromatic residues in these positions. Additionally, the cross-linking patterns in GPAs differ by type: Types I and II feature a ring system comprising the C-*O*-D, D-*O*-E, and A-B aryl rings, while Types III and IV include an additional F-*O*-G ring (Figure 1) [20,22]. Given the high complexity of GPA structures, their biosynthesis necessitates the coordinated action of multiple enzymes to ensure the precise incorporation of precursor amino acids and the execution of subsequent modifications.

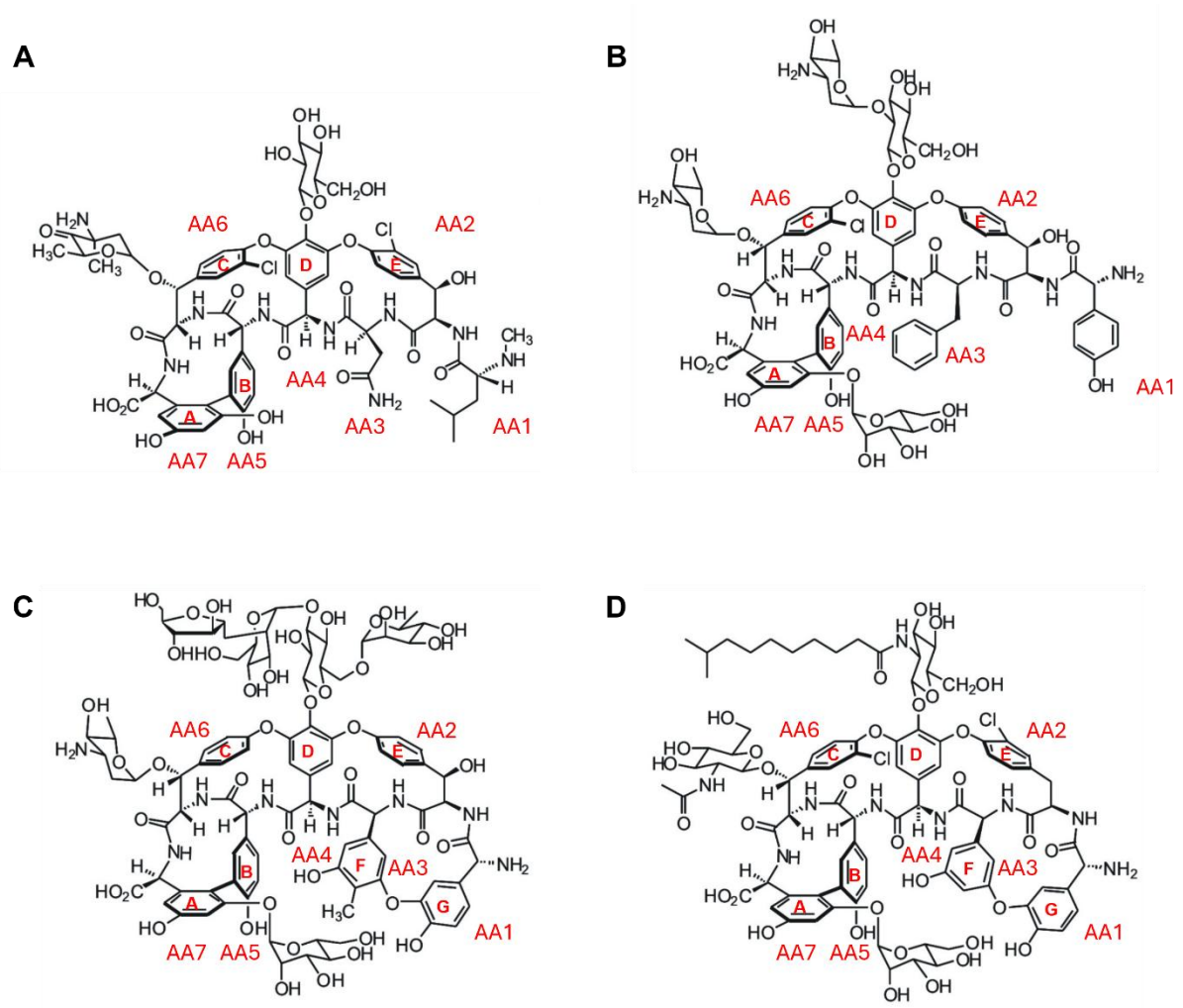


Figure 1: Representatives of GPAs from types I-IV. Shown are structures of (A) balhimycin as type I GPA, (B) actinoidin as type II, (C) ristomycin as type III, and (D) teicoplanin as type IV.

1.3.2 Genetic organization of biosynthetic genes

The genes encoding enzymes involved in the GPA pathways are organized into biosynthetic gene clusters (BGC), typically harboring genes responsible for precursor biosynthesis, backbone assembly, modifications of the backbone, transport, and resistance mechanisms, ensuring the coordinated production of the final bioactive compound.

Although GPAs were discovered earlier, the first reported GPA BGC was that of chloroeremomycin (*cep*), a Type I GPA [26]. Following the identification of the *cep* BGC, several additional GPA BGCs were published in the subsequent years, including those responsible for the production of balhimycin (*bal*) [27], vancomycin (*vps*) [28], teicoplanin (*tei*) [29,30], pekiskomycin (*pek*) [31], A40926 (*dbv*) [32] and A47934 (*sta*) [33].

Among the most extensively studied BGCs is the *bal* BGC (Figure 2), whose encoded enzymatic machinery serves as a model system for this thesis. The *bal* BGC encompasses nearly all the essential genes required for biosynthesis including those for precursor formation. This encompasses the *dpgABCD* genes for (S)-3,5-dihydroxyphenylglycine (Dpg) [34,35], the *hmaS* and *hmo* genes for 4-hydroxyphenylglycine (Hpg) [36], and the genes *bpsD*, *oxyD*, and *bhp* for biosynthesis of β-hydroxytyrosine (Bht) [37]. In addition, pdh_{sec} and $dahp_{sec}$ encode isoenzymes involved in primary metabolism, playing a critical role in bypassing feedback regulation of key precursors derived from the shikimate pathway [38,39].
Furthermore, the gene *tba* encodes an ATP-binding cassette transporter (ABC transporter) that is exclusively responsible for the export of balhimycin [40]. The resistance mechanism is mediated by the gene products of the *vanHAX* operon, which, interestingly, is not located within the *bal* BGC but elsewhere in the genome [41]. However, the *bal* BGC does contain *vanRS*-like genes that encode a two-component system, typically responsible for regulating the expression of *vanHAX.* In contrast, in *A. balhimycina*, this system does not appear to fulfill this regulatory role [42]. Additionally, the gene $vanY_{ab}$, which encodes a carboxypeptidase, has been shown to mediate resistance by cleaving D-Ala residues from the UDP-linked peptidoglycan precursor diacetyl-αγ-L-diaminobutryl-D-Ala-D-Ala [41].

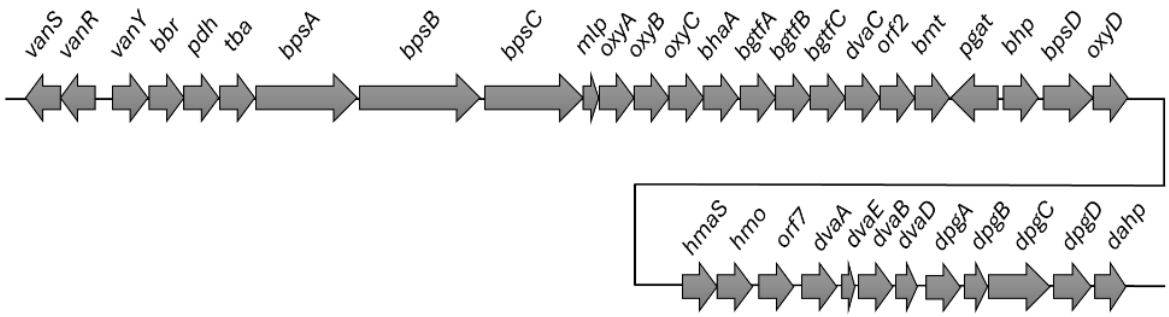


Figure 2: The **balhimycin BGC.** Shown are all genes located in the balhimycin BGC. BGC contains nearly all biosynthetic genes as well as resistance determinants, regulatory genes and transporter: a two component system (*vanRS*), a carboxypeptidase (*vanY*), an StrR-like transcriptional regulator (*bbr*), a prephenate dehydrogenase (*pdh*), an ABC transporter (*tba*), the non-ribosomal peptide synthetases (*bspA-C*), an MbtH-like protein (*mlp*), the P450 monooxygenases (*oxyA-C*), a halogenase (*bhaA*), the glycosyltransferases (*bgtfABC*), a putative deacetylase (*orf2*), the N-methyltransferase (*bmt*), a phenylglycine aminotransferase (*pgat*), the genes involved in the biosynthesis of β-hydroxytyrosine (*bhp, bpsD, oxyD*), the genes involved in the biosynthesis of hydroxyphenylgylcine (*hmaS, hmo*), a putative Na^+/H^+ antiporter (*orf7*), the genes involved in the biosynthesis of the sugar moieties (*dvaA-D*), the genes involved in the biosynthesis of dihydroxyphenylglycine (*dpgA-D*) and the DAHP synthase (*dahp*).

The biosynthetic core machinery consists of three NRPSs encoded by *bpsA*, *bpsB*, and *bpsC*. Downstream of these genes lies *mbtH*, which encodes an MbtH-like protein that presumably enhances the functionality of the NRPS, as shown for other NRPS pathways [43].

Beyond the core NRPS genes, BGC includes all genes required for modification enzymes. Among them are three P450 monooxygenases, *oxyA*, *oxyB*, and *oxyC*, which catalyze essential aryl-ether bond formations within the heptapeptide backbone (10.1002/cbic.200400328). Another key modification is the chlorination of Bht at positions two and six, catalyzed by the FAD-dependent halogenase *bhaA* [44].

Additionally, *bmt* encodes a methyltransferase responsible for modifying the backbone with a methyl group [45,46]. The BGC also contains *bgtfA*, *bgtfB*, and *bgtfC*, which encode glycosyltransferases that attach glucose and dehydrovancosamine moieties to the balhimycin aglycone [47,48]. Furthermore, the *dvaABCDE* genes play a key role in the biosynthesis of these sugar moieties [27,49].

1.4 Biosynthesis of balhimycin as a model of GPA biosynthesis

1.4.1 Precursor biosynthesis

Since the discovery of the balhimycin BGC, the functions of its encoded enzymes have been extensively studied through genetic manipulation, biochemical assays, and analytical chemical analyses. This has made it one of the best-characterized GPA BGCs to date [46].

The enzymatic catalysis performed by these enzymes is highly specific and begins with the supply of precursor amino acids (AA). These include the proteinogenic AA L-leucine at peptide position 1 (AA1) and asparagine at AA3 as well as the non-proteinogenic amino acids Bht at AA2 and AA6, Hpg at AA4 and AA5, and Dpg at AA7 (Figure 3).
The biosynthesis of Bht is is mediated by a standalone NRPS BpsD. This enzyme contains two catalytic domains: the adenylation (A) domain, which activates tyrosine and the peptidyl carrier protein (PCP) domain. The enzyme OxyD hydroxylates tyrosine, which is subsequently cleaved from BpsD by the hydrolase activity of Bhp [37].

The biosynthesis of Hpg involves a series of enzymatic steps. The process begins with prephenate, which is converted into p-hydroxyphenylpyruvate by the dehydrogenase Pdh. This intermediate is then utilized by HmaS, a 4-hydroxymandelate synthase, to produce 4-hydroxymandelate, which is further oxidized by HmO. The final step is catalyzed by the aminotransferase Pgat/Hpgt, which transfers an amino group from tyrosine to yield Hpg [36]. The synthesis of Dpg requires the coordinated action of DpgA-D. It begins with malonyl-CoA as the starter units, which are converted into 3,5-dihydroxyphenylacetyl-CoA through the enzymatic activities of DpgA, DpgB, and DpgD. Subsequently, DpgC catalyzes the conversion of 3,5-dihydroxyphenylglyoxylate into (S)-3,5-dihydroxyphenylglycine via the transamination by Pgat, again using tyrosine as an amino group donor [35,50].

1.4.2 Linking the amino acids into a heptapeptide backbone

Proteinogenic and non-proteinogenic amino acids are needed for the generation of the heptapeptide backbone, a process facilitated by NRPSs. These megaenzymes are composed of multiple modules, each harboring catalytic domains with distinct functions [51]. Each module incorporates a specific amino acid into the growing peptide chain. Importantly, the module arrangement determines the final peptide sequence [52–54]. Thus, identifying the substrate specificity of each module allows for structural predictions of the resulting compound. Since

balhimycin is a heptapeptide, its biosynthesis involves three NRPSs comprising seven modules, with modules 1-3 in BpsA, modules 4-6 in BpsB and module 7 in BpsC (Figure 3).

Each module contains catalytic domains responsible for amino acid incorporation. These include A-domains, that are responsible for the selection and activation of amino acids via adenylation, consuming ATP to generate aminoacyl adenylates and releasing inorganic pyrophosphate (PPi)[55]. This activation is particularly important for the generation of highly reactive substrates that are transferred to the thiol group of the phosphopantetheine (PPant) prosthetic group of the peptidyl carrier protein (PCP)-domain [56]. The PCP-domains shuttle amino acids and peptides between catalytic domains. This process is only possible after posttranslational modification of the PCP-domain by a phosphopantethenyl transferase (PPTase), which transfers a PPant moiety to a conserved serine residue on PCP helix-2. This modification generates a reactive thiol group, allowing a temporary loading of amino acids, while also providing structural flexibility for reaching catalytic sites within the NRPS [56,57]. Condensation domains (C-domains) are required for the peptide bond formation between the amino acids [58].

Additionally, a module may contain an epimerization domain (E-domain) which alters the stereochemistry of amino acids. by converting the C-terminal residues of PCP-bound peptides from L to D [59]. The thioesterase domain (TE-domains) functions as a final catalytic step, releasing the peptide chain from the NRPS [60]. In the balhimycin NRPSs, the TE domain is in the last module of BpsC. The reaction is catalyzed in two steps, starting with the transfer of the peptidyl group from the donor PCP to an activated serine residue of the active site in the TE-domain, forming an O-acyl-enzyme intermediate. This triggers a nucleophilic attack that results in peptide hydrolysis and release from the NRPS [61]. Finally, the X-domain, a unique domain found only in GPAs, functions as a recruitment platform for P450 monooxygenases. It facilitates the cyclization cascade and enables a sequential interaction with the monooxygenases and PCP-bound peptide substrate [62–65].

The precise order of action of these domains dictates the specific assembly of the balhimycin peptide backbone, which is further modified by a set of enzymes during its synthesis while bound to the NRPSs and after its release.

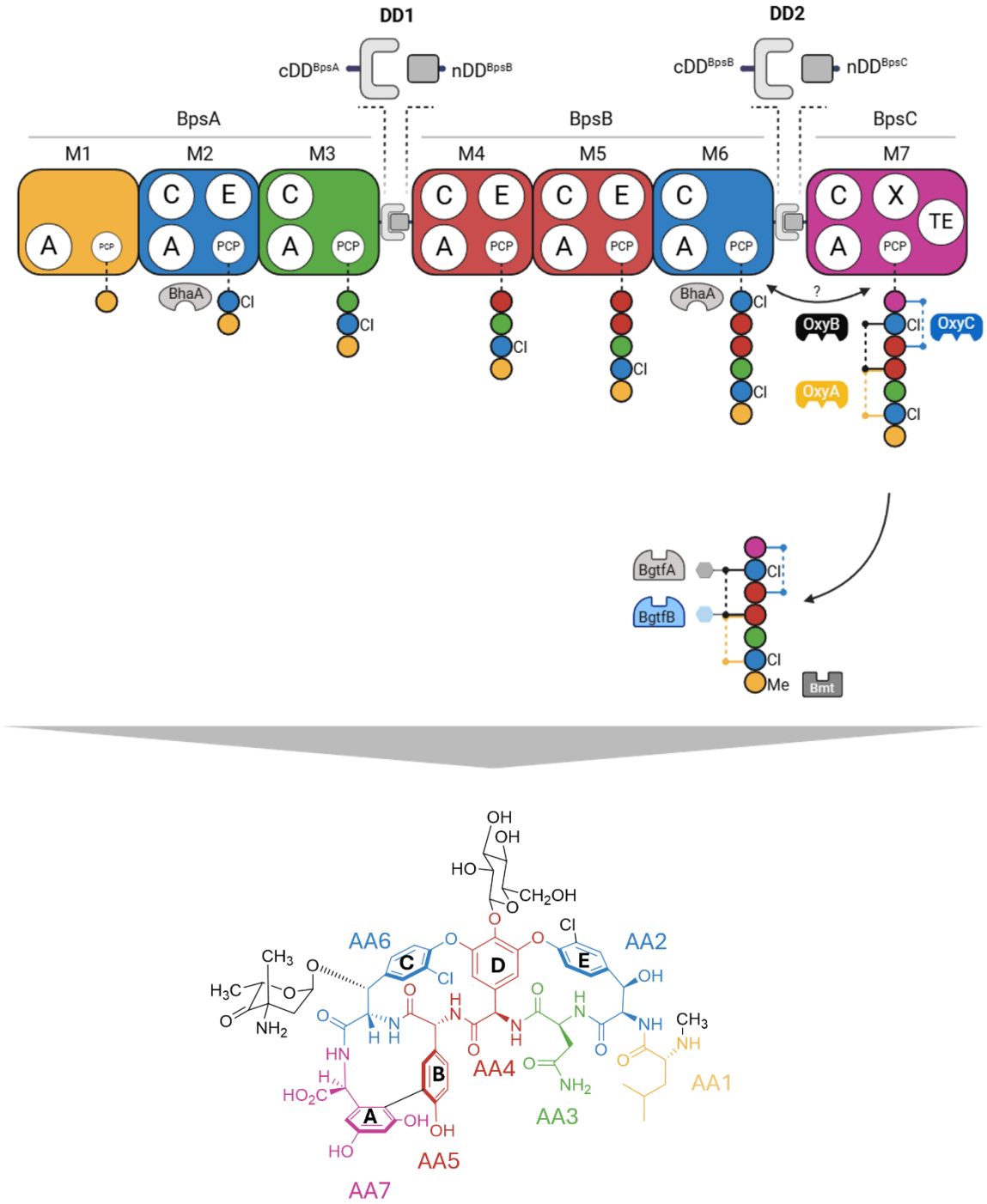


Figure 3: Schematic representation of balhimycin biosynthesis. The biosynthesis is catalyzed by the NRPSs BpsA-C. Colors indicate modules (M1-M7) and the corresponding amino acids incorporated into a growing peptide chain. In yellow L-leucine, in blue β-hydroxytyrosine, in green asparagine, in red 4-hydroxyphenylglycine and in pink 3,5-dihydroxyphenylglycine. The halogenase (BhaA) chlorinates amino acids (AA) at positions AA2 and AA6. The P450 monooxygenases (OxyA-C) catalyze crosslinking reactions, and glycosyltransferases (BgtfA and BgtfB) attach sugar moieties. The N-methyltransferase (Bmt) attaches a methyl group to the amino acid at position AA1. The balhimycin structure indicates the colored amino acids (AA1-AA7) and the lettering of the aromatic side chains contributing to the C-*O*-D, D-*O*-E and A-B crosslinks.

1.4.3 Modification of the balhimycin heptapeptide backbone

The final structure of balhimycin features numerous modifications of the heptapeptide backbone, indicating the action of various modification enzymes.
These modifications include chlorination, cyclisation, methylation and glycosylation. Chlorination occurs at AA2 and AA6, where β-hydroxytyrosine is incorporated. The attachment of the chlorine is catalyzed by BhaA, an FAD-dependent halogenase [44], and is suggested to be essential for balhimycin dimerization, an event crucial for the mechanism of action [66].

The cross-linking steps of the linear balhimycin backbone are key reactions leading to its final cup-shaped structure, a conformation required for GPA binding to its target. In the balhimycin biosynthesis, the three P450 monooxygenases OxyABC catalyze specific cross-links in a defined order. First, OxyB catalyzes the C-O-D aryl-ether bond between Bht at AA2 and Hpg at AA4. Subsequently, OxyA catalyzes the D-O-E aryl-ether bond between Hpg at AA4 and Bht at AA6. The A-B bond between Hpg at AA5 and Dpg AA7 is then catalyzed by OxyC [65].
Another modification in balhimycin is the methylation of L-leucine (AA1) which is predicted to be mediated by the BGC-encoded N-methyltransferase Bmt [45].
In addition to these mentioned modifications, the attachment of sugar moieties completes the balhimycin structure. The BGC encodes three glycosyltransferases, BgtfABC, each responsible for attaching a specific sugar moiety in a defined order. BgtfB first adds a glucose moiety to AA4, followed by BgtfA, which attaches an oxo-vancosamine moiety to AA6. Finally, BgtfC extends the structure by adding a further oxo-vancosamine moiety to the attached glucose [46]. However, this last step appears to occur infrequently under the conditions used.

1.4.4 The export as a final step of biosynthesis

GPA BGCs typically encode ABC transporters, which are dedicated to exporting the produced GPA rather than being involved in the resistance.
To date, the only functionally characterized transporter within GPA biosynthesis is the Tba from balhimycin. Gene deletion experiments demonstrated that in the absence of Tba balhimycin export is abolished [40].

1.4.5 Temporal cascade of the modification reactions

The biosynthesis of GPAs involves a series of precisely coordinated enzymatic modifications that contribute to their structural complexity and biological activity. Many of these modifications occur while the heptapeptide backbone is still tethered to the NRPS machinery.

Studies have shown that halogenation is one of the earliest tailoring steps. *In vivo* experiments using an engineered balhimycin NRPS system, which terminates after the incorporation of Bht at position AA2, resulted in halogenated dipeptides. This indicates that the FAD-dependent halogenase BhaA catalyzes the chlorination of AA2 while the substrate remains bound to the PCP domain of module 2 [67]. A second chlorination event occurs at AA6, which is also a Bht, likely following a similar mechanism during NRPS elongation.

Subsequent oxidative cross-linking steps, catalyzed by the P450 monooxygenases OxyB, OxyA, and OxyC, also take place while the peptide remains NRPS-bound. Evidence suggests that the first D-O-E cross-link is introduced at the hexapeptide level, initiating a defined sequence [46,68]. OxyB forms the C-*O*-D aryl-ether bond, followed by OxyA introducing the D-*O*-E linkage, and finally OxyC catalyzing the A-B bond after the incorporation of Dpg (AA7) by module 7 [63,69]. Methylation and glycosylation occur after the peptide is released from the NRPS complex [47,48].

Altogether, these findings highlight the sequential and coordinated nature of GPA maturation. However, it remains unclear at which precise stage of NRPS-mediated peptide biosynthesis each individual modification takes place and how the timing of these reactions is mechanistically regulated.

1.5 The role of protein-protein interactions and their implications in glycopeptide biosynthesis

The biosynthesis of GPAs requires precise coordination between the modules and domains of the NRPSs, as the precursor amino acids must be activated, linked, and transferred from module to module. This process follows the collinearity rule [51,52]. This correct sequence of catalysis depends not only on cooperation between adjacent domains within the same NRPS but also between domains located on different NRPS. In addition, modification enzymes acting in trans are involved. This complex interplay is based on specific structural features of the domains and the overall architecture of the NRPSs, indicating their high structural flexibility as a prerequisite for proper, efficient catalytic function.

Communication between NRPSs is facilitated by specific domains known as communication-mediating domains (COM-domains) or docking domains (DDs). These domains are located at the N- and C-termini of the NRPSs and were first described in the tyrocidine synthetases TycA-C [70]. They can be defined as mediators of protein–protein interactions between NRPSs, thereby enabling substrate shuttling between modules located on different NRPSs, an essential feature for maintaining the integrity of the biosynthetic assembly line [71–74]. Since their

discovery, DD have been characterized and grouped into two types specific to NRPS systems: the Helix-Hand Motif COM-domains [71] and the Three α-Helix Bundle DDs, also known as PAX domains [75,76]. Additionally, a third group of β-hairpin DDs has been identified, which are present in both NRPS and hybrid PKS–NRPS systems [77]. In GPA NRPSs, including BpsA-C of balhimycin, homologous sequences have been identified (Figure 3), but their function and structure roles remain unexplored.

Building on the importance of inter-domain and inter-protein communication in NRPS systems, it becomes evident that catalytic function is not solely determined by the activity of individual domains. Instead, structural flexibility plays a crucial role, enabling domains to adopt alternative conformations necessary for interactions with other NRPS domains or with auxiliary modification enzymes. A prominent example of such a domain is the PCP domain. It acts as a central domain, essential not only for the transfer of substrates between catalytic domains but also for presentation of the bound peptide to modification enzymes. This substrate transfer relies on precise interactions with both up- and downstream domains, as demonstrated in structural and biochemical studies involving A-domain [78,79], C-domains [80] and TE-domains [81].

On the other hand, these interactions with accessory proteins can also play a critical role in NRPSs functionality. MbtH-like proteins (MLPs), for instance, have been identified as essential activators of A-domains forming protein-protein complexes with NRPSs. While their exact mechanism of action remains incompletely understood their importance in NRPS-mediated biosynthesis have demonstrated in several cases [82–84].

An exclusive domain, present only in GPA NRPSs is the X-domain. Due to its homology with C-domains, its function was initially unclear. However, structural and functional analyses revealed that the X-domain serves as a recruitment platform for cytochrome P450 monooxygenases, which are responsible for the oxidative cross-linking of GPA heptapeptides [46,62].

The PCP-domain is also central to this interaction. Studies on the PCP7 and the monooxygenase P450sky of the cyclic depsipeptide skyllamycin showed that P450 monooxygenases can recognize and bind PCP-domains to selectively oxidize their desired substrates. Despite high sequence similarity among PCP domains, even subtle variations in their tertiary structure can significantly influence recognition and binding specificity [85].

1.6 Spatial restriction of cellular processes

Recent findings increasingly support the notion that bacterial metabolic processes do not occur in an unorganized intracellular environment. Instead, they can be spatially restricted and compartmentalized within the cell, creating protected microenvironments tailored to specific biochemical reactions. This compartmentalization offers several advantages, including localized concentration of enzymes and substrates, improved metabolic efficiency, and enhanced regulatory control over cellular processes [86].

Compartmentalization in bacteria can occur both in the cytoplasm and in or associated to the cellular membrane. Although the cytoplasm was long considered a homogenous mixture of molecules and reactions, early studies on *Phormidium uncinatum* suggested the presence of distinct granular structures. Structural analyses in *Synechocystis* sp. PCC 6803 provided the first X-ray crystallographic evidence of carboxysomes, protein-based microcompartments involved in carbon fixation [87,88].

Bacteria are now known to utilize a variety of microcompartments to spatially organize metabolic functions. These include membrane-associated microcompartments i.e. chromatophores, anammoxosomes, and magnetosomes, as well as membrane free microcompartments like carboxysomes and metabolosomes. Additionally, bacteria employ phase-separated compartments, such as nucleolus-like condensates, to dynamically regulate biochemical processes [86]. Importantly, this subcellular organization is not restricted to primary metabolism. Previous studies have shown that enzymes involved in the biosynthesis of natural products also utilize spatial organization, often through membrane association and protein complex formation, to ensure efficient metabolite production. This is shown in a study of cereulide NRPS machinery and its cognate ABC transporter CesCD in the producer *Bacillus cereus,* where the NRPS machinery co-localizes with its cognate ABC transporter CesCD. Deletion or inactivation of the ABC transporter disrupts cereulide production, pointing out that the interaction between them is crucial for continuous biosynthesis [89]. Similarly, investigation on the lantibiotic nisin, showed that its biosynthesis is facilitated by a putative membrane-associated lanthionine synthetase complex, composed of the dehydratase NisB, the cyclase NisC, and the ABC transporter NisT. Co-localization of these enzymes further supports the concept of a spatially organized biosynthetic assembly line [90,91].

Altogether, these findings illustrate that the biosynthesis of natural products is embedded within a highly organized cellular framework, where spatial proximity of enzymes and transport

systems enables coordinated catalysis, efficient substrate channeling, and robust production of secondary metabolites.

2 Aim of this work

The biosynthesis of balhimycin is characterized by a high degree of enzymatic complexity involving numerous catalytic steps and tightly regulated interactions. Over the past decades, extensive studies have shown that this biosynthetic process is governed not only by metabolic flux but also by multiple layers of regulation, including transcriptional control and feedback mechanisms. Furthermore, it has been revealed that the structural features of NRPSs significantly impact the specificity of reactions, likely determining the timing and spatial organization of events and thereby adding an additional layer of regulation.

Recent structural and biochemical data increasingly highlight that NRPSs are not rigid, isolated catalysts. Instead, they are dynamic, highly flexible proteins that function through transient and coordinated interactions with other biosynthetic and accessory proteins. This has led to a paradigm shift, recognizing that NRPSs are components of larger, modular protein complexes rather than stand-alone enzymes.

Despite this progress, the interaction network of NRPSs involved in the biosynthesis of glycopeptide antibiotics (GPAs) remains poorly understood and necessitates systematic investigation to unveil the protein-protein interactions that facilitate the orchestration of the GPA multi-enzymatic machinery.

This thesis aims to elucidate the protein interactions and spatial organization of the balhimycin biosynthetic machinery. To achieve this, various approaches have been employed, including *in vivo* engineering of NRPSs, proteomics, bacterial two-hybrid assays, proximity-dependent biotin labeling, and artificial intelligence-based predictions of protein-protein interaction and complexes.

Understanding the specific interaction requirements of biosynthetic enzymes is essential for the rational design and reprogramming of the NRPS pathway to generate novel GPA derivatives. Additionally, insights into the spatial organization of biosynthetic machinery can be leveraged to optimize the biotechnological production of GPAs, for example, by engineering cells to promote the complex formation of these machineries or enhance the stability and efficiency of the biosynthetic assemblies.

By using these strategies, the interactions between the balhimycin NRPSs BpsA-C as well as with modification enzymes BhaA, OxyA-C, and MLPs were analyzed, and their specific interaction deciphered. Further, the export mediated by Tba, was explored to elucidate its specificity and its role on the balhimycin biosynthetic machinery.

3 Publication 1: Unveiling the substrate specificity of the ABC transporter Tba and its role in the glycopeptide biosynthesis

Nicola Gericke*, **Dardan Beqaj*,** Thales Kronenberger, Andreas Kulik, Athina Gavriilidou, Mirita Franz-Wachtel, Ulrich Schoppmeier, Theresa Harbig, Johanna Rapp, Iwan Grin, Nadine Ziemert, Hannes Link, Kay Nieselt, Boris Macek, Wolfgang Wohlleben, Evi Stegmann[#] and Samuel Wagner[#]. *iScience. 2025 Mar 3;28(4):112135.*
https://doi.org/10.1016/j.isci.2025.112135
**Dardan Beqaj and Nicola Gericke contributed equally to this publication*

3.1 Contributions

All experiments in this publication were planned and designed by Dardan Beqaj, Nicola Gericke, Evi Stegmann, Wolfgang Wohlleben and Samuel Wagner.

Nicola Gericke performed the *in silico* analysis of transporter sequences and phylogenetic analysis resulting in Figure 1, Figure S1, Figure S2, Figure S3, and Figure S12 and analyzed AlphaFold2 models of all transporters manually and contributed to Figure 3B. Athina Gavriilidou and Nadine Ziemert identified the glycopeptide antibiotic (GPA) biosynthetic gene clusters (BGCs) and provided the transporter sequences, contributing to Figure 1 and Figure S3. Iwan Grin performed pairwise comparison of transporter sequences and automated AlphaFold2 usage, contributing to Figure 3B. Athina Gavriilidou performed hierarchical clustering of the transporter sequences, contributing to Figure S3. Thales Kronenberger performed all MD simulations, resulting in Figure 3 and Figure S7.
Dardan Beqaj and Nicola Gericke constructed all plasmids and *Amycolatopsis balhimycina* strains, performed all *in vivo* experiments, and analyzed all corresponding data, resulting in Figure 2, Figure 4, Figure S4, Figure S5, Figure S6, and Figure S8. Andreas Kulik performed all HPLC-MS measurements of balhimycin, contributing to Figure 2, Figure 4, Figure S4, Figure S6, Figure S8, and Figure S11. Ulrich Schoppmeier did statistical analysis of the *in vivo* balhimycin export assay, contributing to Figure 2 and Figure 4.
Dardan Beqaj and Nicola Gericke performed the transcriptomics experiment and final analysis, resulting in Figure 5A and Figure S9A. Raw data generation and primary analysis for transcriptomics were performed by Kay Nieselt and Theresa Harbig, contributing to Figure S9A.
Dardan Beqaj and Nicola Gericke performed the metabolomics experiment and final analysis, resulting in Figure 5B and Figure S9B. Raw data generation and primary analysis for metabolomics were performed by Hannes Link and Johanna Rapp, contributing to Figure S9B.

Dardan Beqaj and Nicola Gericke performed the proteomics experiment and final analysis, resulting in Figure 6 and Figure S11. Raw data generation and primary analysis for proteomics were performed by Mirita Franz-Wachtel and Boris Macek.
Dardan Beqaj and Nicola Gericke participated in the development of the study, drafting the manuscript, and writing of the paper.

3.2 Summary of results

The production of GPAs in *A. balhimycina* is a highly regulated process that relies on efficient biosynthesis and export mechanisms. The ABC transporter Tba has been identified as an exporter responsible for translocating balhimycin out of the cell [40]. The primary goal of this study was to investigate the substrate specificity of Tba and examine correlations between the phylogenetic relationships of GPA-associated ABC transporters and the type of GPA they export. A particular focus was placed on identifying the substrate-binding pocket responsible for GPA recognition and translocation.

While previous reports have suggested that deletion of *tba* leads to intracellular balhimycin accumulation, our study revealed a drastic reduction in balhimycin levels upon deletion or inactivation of *tba*. This unexpected finding raised the possibility that Tba may not solely function as an exporter but could also be involved in the regulation of balhimycin biosynthesis. This led to the hypothesis that Tba might interact with the biosynthetic machinery, prompting further investigation into a potential mechanistic link between GPA export and biosynthesis.

A comparative genomic analysis of GPA BGCs across actinomycetes revealed that each BGC encodes a single putative ABC transporter. These transporters are predicted to form homodimers, each monomer comprising six transmembrane helices containing transmembrane domain (TMD) and characteristic nucleotide-binding domain (NBD) motifs (Publication 1: Figure 1; Table S1; Figure S1). Sequence analysis showed over 50% identity among transporters, with type I-IV GPA transporters displaying higher similarity to each other than to type V (glycopeptide antibiotic related peptides (GRP)) transporters (Publication 1: Figure S2).

Phylogenetic analysis revealed clustering of transporters according to GPA type, suggesting that substrate specificity is predominantly determined by the core peptide backbone rather than modifications. Evolutionarily, transporters appear to follow the divergence of their respective BGCs (Publication 1: Figure 1; Figure S3).

The functional relevance of Tba was investigated using an *A. balhimycina Δtba* mutant, in which intra- and extracellular balhimycin levels were analyzed. HPLC-MS analysis revealed a drastic reduction in both export and overall accumulation of balhimycin, indicating impaired biosynthesis. Complementation with the native *tba* gene fully restored export, confirming the transporter's essential role in balhimycin biosynthesis (Publication 1: Figure 2a–d). Complementation with the native *tba* gene restored export, validating its role (Publication 1: Figure 2a-d).

To assess substrate specificity among GPA transporters, the *A. balhimycina Δtba* mutant was complemented with the ABC transporter genes from the vancomycin (*tva*) and ristomycin (*tri*) BGCs. Vancomycin is a type I, while ristomycin represents type III GPA, differing primarily in their peptide backbone structures and the attached sugar residues (Publication 1: Figure 3A). Complementation with *tva* efficiently restored balhimycin export, whereas *tri* did not, indicating a clear substrate preference based on the peptide backbone rather than glycosylation pattern (Publication 1: Figure 2a-d). Membrane localization assays using SDS and BN-PAGE confirmed correct insertion and expression of the heterologous transporters, ruling out misfolding or localization artifacts as a cause for the observed export differences (Publication 1: Figure 4e).

Further evidence supporting peptide backbone recognition was obtained through molecular docking and molecular dynamics (MD) simulations using vancomycin and balhimycin as model substrates. These simulations positioned type I GPAs, vancomycin and balhimycin, within a predicted binding cavity at the center of the inward-open TMD core, revealing stable interactions between amino acid residues at positions 4-7 of the GPAs and key transporter residues, including F269, R310, G309, and G313. Notably, no significant interactions were observed with residues 1-3, which are typically used for GPA classification. R1 sugar moieties remained cytoplasm-facing during simulations, suggesting a minor role in substrate binding and possibly influencing solubility instead (Publication 1: Figure 3). These results reinforce the hypothesis that GPA transporters primarily recognize the peptide backbone.

Experimental validation confirmed this hypothesis. Site-directed mutagenesis of key TMD residues of Tba revealed that substitution of T316 (located at the base of the substrate binding pocket) significantly reduced balhimycin export, whereas mutations in Q309 and R310 had no detectable effect. This suggests a stabilizing rather than direct substrate-binding role for T316. (Publication 1: Figure 4a).

Additional experiments explored the influence of glycosylation and chlorination on balhimycin export efficiency. Deletion of the glycosyltransferase gene *bgtfB* in *A. balhimycina* resulted in the production of a non-glycosylated heptapeptide that was exported at ~5.5-fold higher levels) than the glycosylated balhimycin (Publication 1: Figure 4d). Despite the increased export, this sugar-free derivative exhibited reduced antibacterial activity against *Bacillus subtilis*, highlighting the importance of glycosylation for the mode of action (Publication 1: Figure 4e). Conversely, deletion of the halogenase gene *bhaA*, which blocks chlorination, had no detectable effect on export efficiency or bioactivity (Publication 1: Figure 4d-e). A double

mutant lacking both *bgtfB* and *bhaA* exhibited even higher export levels than the single mutants, though the resulting product lacked antibacterial activity (Publication 1: Figure 4d-e). These findings reinforce that Tba-mediated GPA transport is largely independent of glycosylation and chlorination. Instead, they support the model in which substrate recognition by the ABC transporter is primarily governed by as peptide backbone of the GPA.
Contrary to prior findings, deletion or inactivation of *tba* in *A. balhimycina* did not result in the intracellular accumulation of balhimycin. Instead, balhimycin biosynthesis was nearly abolished. A catalytically inactive mutant (Tba^{E545Q}) failed to export balhimycin and showed negligible production both intra- and extracellularly (Publication 1: Figure 2a-d). These findings suggest that active transport is essential for balhimycin biosynthesis, potentially through a feedback regulation mechanism.

To investigate whether this regulation occurs at the transcriptional level, transcriptome analysis of *A. balhimycina Δtba* and *A. balhimycina* wild type was performed, revealing no significant downregulation of the BGC expression in *A. balhimycina Δtba* when compared to the wildtype. In contrast, a control mutant lacking the known BGC regulator Bbr, *A. balhimycina Δbbr* [92] exhibited strong downregulation, confirming that Tba does not act as a transcriptional regulator of the biosynthetic genes (Publication 1: Figure 5a). Metabolomic analysis also ruled out precursor accumulation as a potential inhibitory mechanism, as the levels of key amino acids and intermediates remained unchanged between wildtype and *Δtba* strains (Publication 1: Figure 5b).

To assess potential physical interaction between Tba and the biosynthetic machinery, a proximity-dependent biotinylation assay using TurboID (TID) was employed. A recombinant *A. balhimycina* strain producing a Tba^{TID} fusion protein restored balhimycin export, validating the functionality of the fusion construct. Comparative proteomic analysis of biotinylated proteins revealed a significant enrichment of balhimycin biosynthetic enzymes, including NRPSs (BpsB, BpsC), modification enzymes (OxyC, BhaA, BgtfAB, Bmt), and enzymes involved in precursor supply (Pgat, BpsD, OxyD, HmaS, DvaB, DvaC) (Publication 1: Figure 6c-d). These results suggest that Tba forms part of a biosynthetic microenvironment, potentially facilitating metabolite channeling or assembly line stabilization during balhimycin production.

3.3 Paper 1

Unveiling the substrate specificity of the ABC transporter Tba and its role in glycopeptide biosynthesis

Author list:
Nicola Gericke[1,11], Dardan Beqaj[2,11], Thales Kronenberger[1,3,4], Andreas Kulik[2], Athina Gavriilidou[5], Mirita Franz-Wachtel[6], Ulrich Schoppmeier[1,7], Theresa Harbig[8], Johanna Rapp[7,9], Iwan Grin[1], Nadine Ziemert[3,5], Hannes Link[7,9], Kay Nieselt[8], Boris Macek[6], Wolfgang Wohlleben[3,7,10], Evi Stegmann[2,3,7,12*], Samuel Wagner[1,3,7,*]

Author affiliations and footnotes:
1 Cellular and Molecular Microbiology, Interfaculty Institute of Microbiology and Infection Medicine (IMIT), University of Tübingen, Elfriede-Aulhorn-Str. 6, 72076 Tübingen, Germany

2 Microbial active compounds, Interfaculty Institute of Microbiology and Infection Medicine (IMIT), University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

3 Partner-Site: DZIF Tübingen, Elfriede-Aulhorn-Str. 6/Auf der Morgenstelle 28, 72076 Tübingen,Germany

4 School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, Yliopistonrinne 3, 70211 Kuopio, Finland.

5 Translational Genome Mining for Natural Products, Interfaculty Institute of Microbiology and Infection Medicine Tübingen (IMIT), University of Tübingen, Auf der Morgenstelle 24, 72076 Tübingen, Germany

6 Proteome Center Tübingen, Institute of Cell Biology, University of Tübingen, Auf der Morgenstelle 15, 72076 Tübingen, Germany

7 Excellence Cluster "Controlling Microbes to Fight Infections" (CMFI), University of Tübingen, 72076 Tübingen, Germany

8 Interfaculty Institute for Bioinformatics and Medical Informatics (IBMI), University of Tübingen, Sand 14, 72076 Tübingen, Germany.

9 Bacterial Metabolomics, Interfaculty Institute of Microbiology and Infection Medicine (IMIT), University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

10 Microbiology/Biotechnology, Interfaculty Institute of Microbiology and Infection Medicine (IMIT), University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

11 These authors contributed equally

12 Lead contact

* Corresponding authors
evi.stegmann@uni-tuebingen.de (E.S.), samuel.wagner@med.uni-tuebingen.de (S.W.)

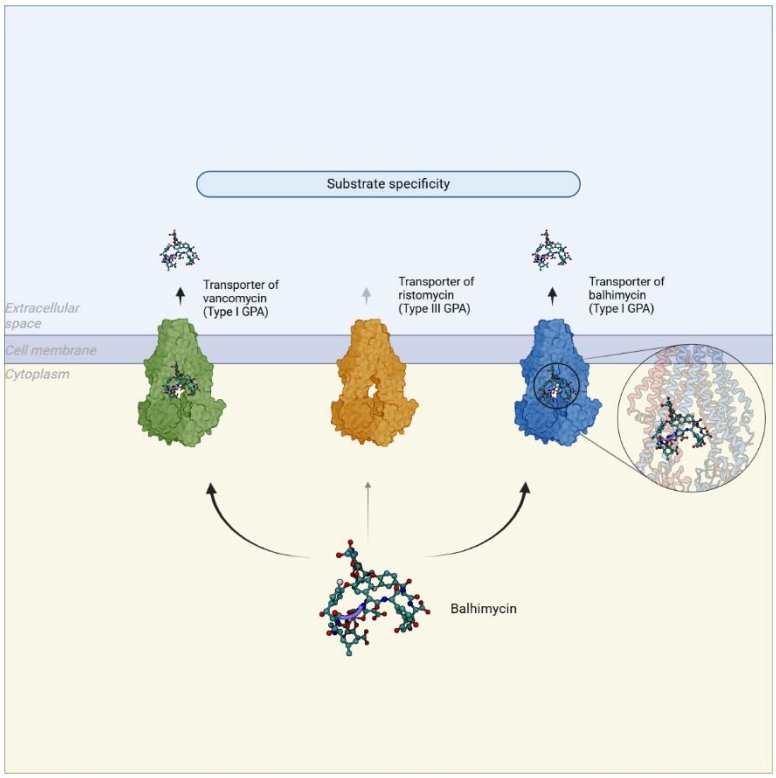
SUMMARY

Glycopeptide antibiotics (GPA) such as vancomycin are essential last-resort antibiotics produced by actinomycetes. Their biosynthesis is encoded within biosynthetic gene clusters, also harboring genes for regulation, and transport. Diverse types of GPAs have been characterized that differ in peptide backbone composition and modification patterns. However, little is known about the ATP-binding cassette (ABC) transporters facilitating GPA export. Employing a multifaceted approach, we investigated the substrate specificity of GPA ABC-transporters towards the type-I GPA balhimycin. Phylogenetic analysis suggested and trans-complementation experiments confirmed that balhimycin is exported only by the related type I GPA transporters Tba and Tva (transporter of vancomycin). Molecular dynamics simulations and mutagenesis experiments showed that Tba exhibits specificity towards the peptide backbone rather than the modifications. Unexpectedly, deletion or functional inactivation of Tba halted balhimycin biosynthesis. Combined with proximity biotinylation experiments, this suggested that the interaction of the active transporter with the biosynthetic machinery is required for biosynthesis.

Keywords

Glycopeptide antibiotics (GPA), ABC transporters, Natural products, Actinomycetes, NRPS

Graphical abstract



Introduction

Glycopeptide antibiotics (GPAs) belong to a diverse class of bioactive compounds synthesized by various actinomycetes. Notable examples include vancomycin and teicoplanin, which are clinically used to treat infections caused by multidrug-resistant gram-positive bacterial pathogens[1].

Based on their backbone composition, GPAs have been classified into five structural subclasses. Type I-IV GPAs are commonly referred to as “true” GPAs[2]. They share a similar backbone composition, featuring type-specific amino acids (AAs) at positions AA1 and AA3 with either aliphatic (I) or aromatic (II, III, IV) side chains. However, the modification pattern of the backbone varies between GPAs within the same type[3]. Type V GPAs have recently been classified as glycopeptide-related peptides (GRPs) due to significant differences in structural, phylogenetic, and functional features[4]. GPA backbones are synthesized by non-ribosomal peptide synthetases (NRPS)[5] from both proteinogenic and non-proteinogenic amino acids. During peptide assembly the aromatic side chains are halogenated and cross-linked by dedicated enzymes, resulting in the antibiotic's characteristic multicyclic and cup-shaped structure[6–8]. Following peptide assembly, additional diversifications such as glycosylation, acylation, methylation, and sulfonation occur[9]. The enzymes responsible for these modifications are typically encoded within biosynthetic gene clusters (BGCs).

The export of GPAs represents the last step in their biosynthesis process. To date, only one transporter associated with GPAs, the transporter of the type I (vancomycin type) GPA balhimycin (Tba) in *Amycolatopsis balhimycina* has been functionally characterized[10]. This ATP-binding cassette (ABC) transporter specifically mediates the export of balhimycin and is not involved in self resistance[10,11]. ABC transporters form a large superfamily of integral membrane proteins and are known to facilitate ATP-driven transport of various substrates across the cytoplasmic membrane. However, they may fulfill other roles in addition to transport[12]. ABC transporters can be associated with accessory domains[13], play a vital role in the biosynthesis of their substrates[14,15], or be involved in producer’s self-immunity[16,17]. The latter is particularly relevant to ABC transporters within BGCs encoding compounds with antimicrobial activity.

Considering the paucity of data on how the processes of biosynthesis and export of GPAs are coordinated, we performed a comprehensive investigation of the specificity and functional roles of GPA-associated transporters, with a particular emphasis on the ABC transporter of balhimycin, Tba. Phylogenetic analysis, molecular dynamics simulation (MD simulation), and transport assays revealed that the specificity of Tba towards GPAs of the same type as balhimycin is dictated by the peptide backbone rather than the sugar modifications. Unexpectedly, deletion or functional inactivation of Tba did not lead to intracellular accumulation of balhimycin, suggesting a regulatory role of the transporter in balhimycin biosynthesis. This notion was supported by the observation that the transporter is in close proximity to the biosynthetic machinery.

Results

Bioinformatic analysis and phylogenetic reconstitution of GPA ABC transporters suggest type-associated specificity

Members of the ABC transporter family share a similar architecture consisting of two cytoplasmic nucleotide binding domains (NBD) responsible for ATP binding and hydrolysis, coupled with two transmembrane domains (TMD) forming the substrate translocation pathway at their mirror axis[18]. Based on these features, we searched for putative ABC transporter genes in 89 GPA and GRP BGCs from producers of various families such as *Pseudonocardiaceae*, *Streptosporangiaceae*, *Micromonosporaceae*, *Streptomycetaceae*, and *Nocardiaceae.* Our analysis identified a single gene encoding a putative transporter in each BGC (Table S1).

Analysis of transmembrane topology based on prediction of the membrane integration propensity of segments of the protein sequence[19], revealed six hydrophobic regions in the N-terminal half of all scanned proteins (exemplary in Figure S1). Furthermore, typical motifs for functional NBDs -Walker A, Walker B, and ABC signature motif (C-loop) as well as A-, Q-, D-loop, and H-switch[18]- were identified in the C-terminal half of every protein. These findings strongly suggest that the GPA BGCs encode a homodimeric ABC transporter with six transmembrane (TM) helices per TMD, wherein each half-transporter consists of one TMD and one NBD, similar to Tba[10]. All transporters shared an amino acid sequence identity of more than 50% (Figure S2). Notably, sequence identities between transporters from putative type I-IV GPA BGCs were substantially higher than those to the transporters of GRP BGCs (Figure S2). Transporters with the highest similarity to Tba (88% sequence identity) were identified in *Amycolatopsis orientalis* DSM 40400, *Amycolatopsis keratiniphila* HCCB10007, and *Amycolatopsis regifaucium* DSM45072. These strains harbor BGCs for vancomycin (the first two)[20,21] and decaplanin[22], both of which, like balhimycin, are type I GPAs.

To determine the substrate spectrum of the identified transporters, we investigated their evolutionary relationship with GPAs. Using all 89 transporter sequences, we generated a phylogenetic tree employing the Maximum Likelihood algorithm (Figure S3), with an unrelated ABC transporter (Abc30) from *A. balhimycina* serving as an outgroup. By correlating the transporters with the GPA known GPA/GRP types I-V[4] potentially encoded by the BGCs that also encode the respective transporter, we have categorized them according to these, creating a classification based on their substrates (in colors, Figure S3). Type V (GRP) transporters form a distinct cluster separate from all type I-IV transporters in the phylogenetic tree, indicating an earlier evolutionary divergence and significant differences in their putative substrates. Consequently, our focus was exclusively on type I-IV GPA transporters. Consistent with the high sequence identity, phylogenetic analysis of only type I-IV GPAs did not reveal distinct and widely separated clades, but a rough grouping of transporters associated with the export of GPAs of the same type, indicating an underlying specificity for them (in colors, Figure 1). Transporters exporting GPAs of the same type but with different modification patterns tended to cluster closely together (Figure 1). For example, balhimycin, decaplanin, and vancomycin, which share a type I backbone but vary in glycosylation patterns, formed a distinct clade (Bootstrap >81.6%). These findings indicate that the evolutionary trajectory of the transporters follows that of the BGCs, as these also build clades based on GPA type[23]. They also suggest that the specificity of the transporters may be associated with the peptide backbone rather than its modifications.

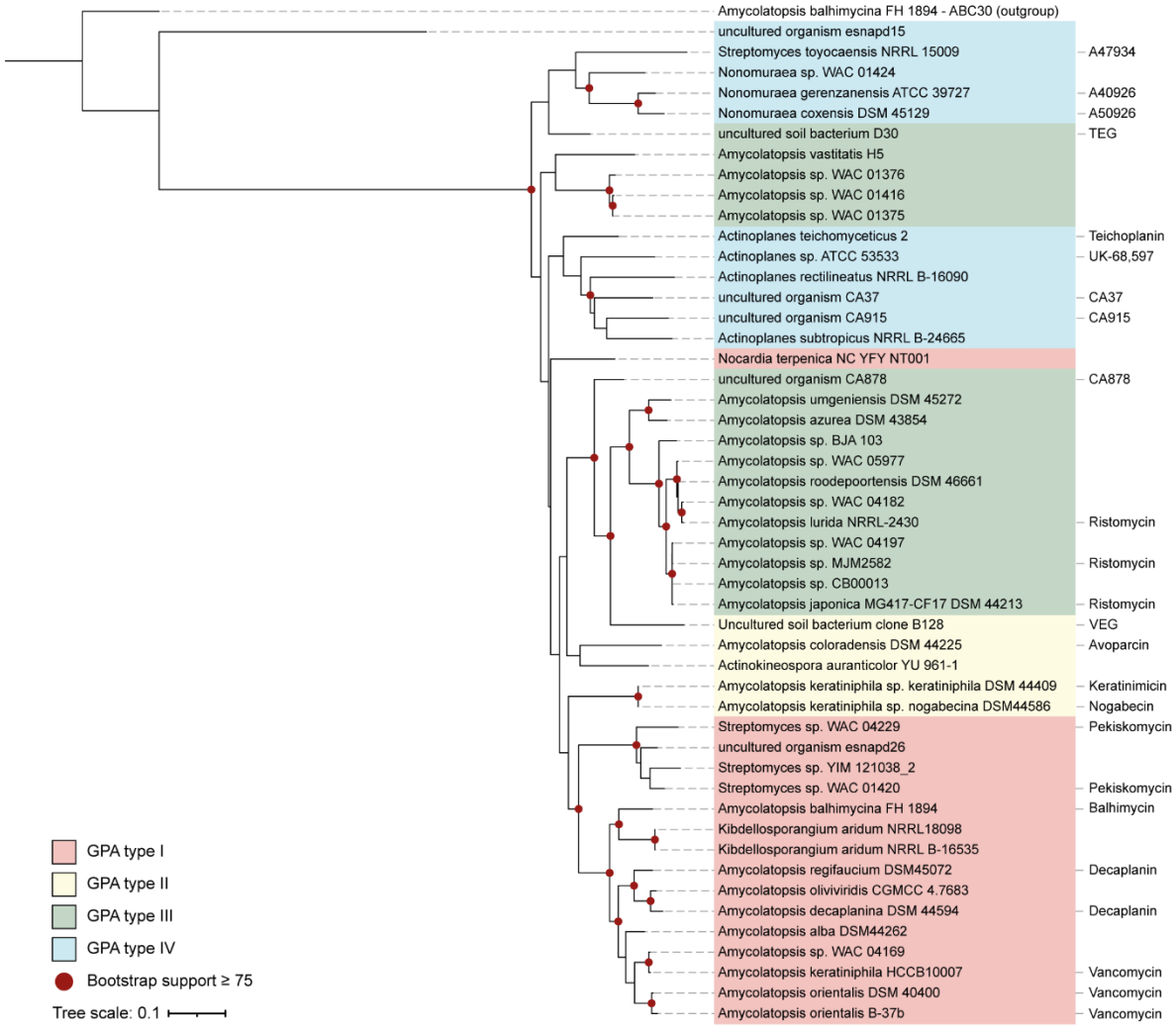


Figure 1. Evolutionary relationship within the group of GPA ABC transporters. Phylogenetic tree (Maximum Likelihood (LG+F+G) algorithm with 1,000 bootstrap repetitions) of ABC transporters encoded in BGCs of GPAs. Nodes with bootstrap support ≥75% are marked by red dots on the branches. The tree was rooted using a putative ABC transporter encoded in an uncharacterized BGC in *A. balhimycina* (Abc30). The strain from which the BGC and thus the associated ABC transporter originates is indicated in the tree. (Predicted) GPA types are labeled accordingly: type I (red), type II (yellow), type III (green), type IV (blue). Known and predicted GPAs are mentioned behind the branches.

GPA-associated ABC transporters demonstrate specificity in exporting their native substrates

In silico analyses and phylogenetic reconstitution of putative ABC transporters of GPAs suggested a significant degree of similarity among them. To empirically confirm their substrate specificity, we developed an *in vivo* system based on *trans* complementation of an *A. balhimycina* Δ*tba* mutant. To this end, we constructed the *A. balhimycina* Δ*tba* deletion and evaluated balhimycin export using high-performance liquid chromatography-mass spectrometry (HPLC-MS). Specifically, we quantified balhimycin in culture supernatants and mycelium extracts obtained from a 96-hour culture of *A. balhimycina* wildtype and *A. balhimycina* Δ*tba* under balhimycin-producing conditions. The HPLC-MS results confirmed the presence of balhimycin in both strains, with prominent peaks corresponding to fully

glycosylated (*m/z* 1587.49 $[M+H]^+$), double-glycosylated (*m/z* 1446.41 $[M+H]^+$), and mono-glycosylated (*m/z* 1305.34 $[M+H]^+$) derivatives in each sample (Figure 2a,b; Figure S4). Since accurate protein concentrations of the produced transporter could not be measured due to purification constraints, we relied on HPLC-MS results for quantification, normalizing to dry cellular weight for better comparability (Figure 2c,d).

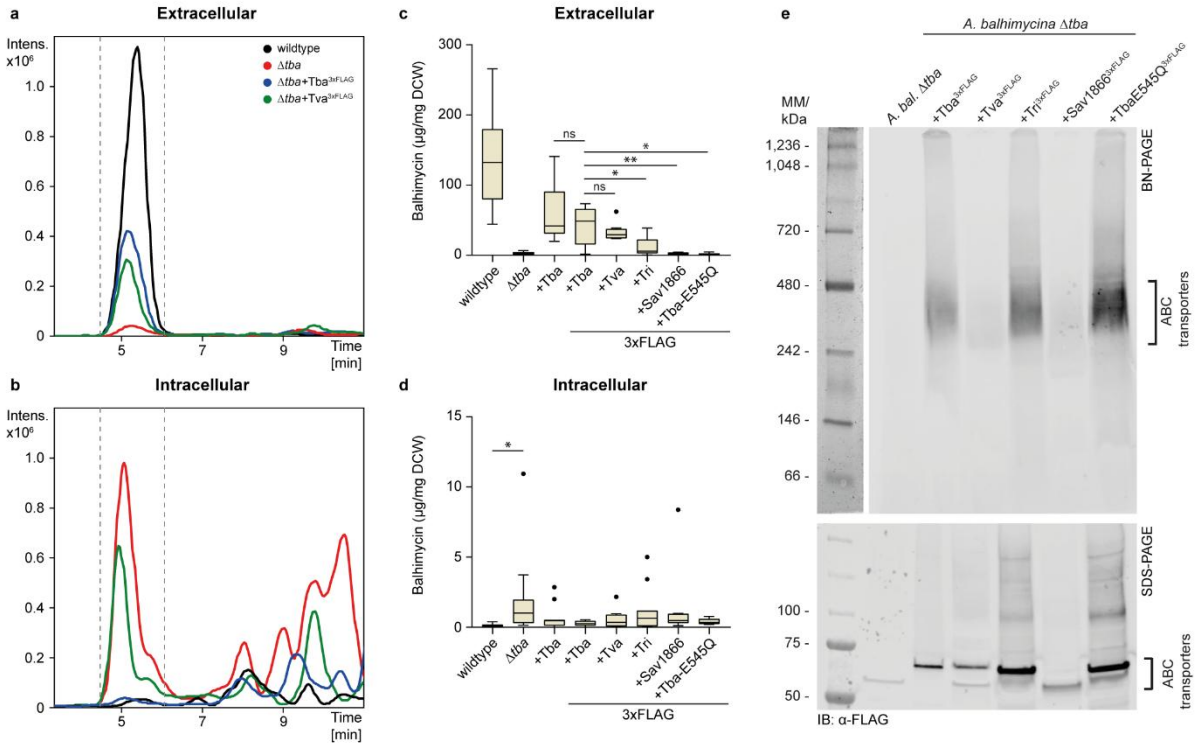


Figure 2. Analysis of complementation efficiency of *A. balhimycina* Δ*tba* using HPLC-MS detection of balhimycin. a-b: Raw (not normalized) HPLC-MS chromatograms of the culture supernatants (**a**) and mycelium extracts (**b**) of *A. balhimycina* wildtype (black), Δ*tba* (red), *Δtba+tba*^{3xFLAG} (blue) and *Δtba+tva*^{3xFLAG} (green). The peaks (retention time 5-5.5 min) represent the extracted ion chromatogram (EIC) of the protonated balhimycin mass *m/z* 1446.41 $[M+H]^+$ (positive mode, smoothed 10.63-10.76 GA) (Figure S4). **c-d**: Quantification of extracellular (**c**) and intracellular (**d**) levels of balhimycin in *A. balhimycina* and respective mutant strains based on HPLC-MS data. Values were normalized to the dry cell weight (DCW). All data are shown in boxplot representation. Statistical outliers are displayed as dots. Statistically significant differences are highlighted with asterisks ($p<0.05$ (*); $p<0.01$ (**)) ("Wilcoxon signed-rank test" and "Benjamini-Hochberg Procedure") (Table S2). **e:** Immunoblotting after BN- (top) and SDS-page (bottom) of solubilized crude membranes of *A. balhimycina* Δ*tba* and respective mutant strains. Detection with 3xFLAG specific antibodies

The extracellular amount of balhimycin in *A. balhimycina Δtba* was significantly reduced in comparison to *A. balhimycina* wildtype, by a factor of 50 (Figure 2c), in accordance with previous results[10]. Unexpectedly, the level of balhimycin that accumulated intracellularly in *A. balhimycina Δtba* was approximately 70-fold lower than the level of balhimycin exported by *A. balhimycina* wildtype, indicating impaired biosynthesis. To confirm that the reduction of balhimycin in *A. balhimycina* Δ*tba* resulted solely from the deletion of *tba*, we complemented the mutant with the native *tba* gene and with *tba*^{3xFLAG} under the control of a strong constitutive promoter (*ermE*p*)[24]. Each gene was integrated into the genome at an ectopic locus, specifically the *ϕC31* attachment site. The expression of *tba* with a 3xFLAG epitope tag enabled us to verify the production and dimerization of the transporter via immunoblot analysis. Both complementations restored balhimycin export function in *A. balhimycina* Δ*tba* (Figure 2c),

although not to wildtype levels, possibly due to the integration of the gene at a non-native locus. Statistical analysis revealed no significant difference between the amount of balhimycin exported by Tba or Tba^{3xFLAG} (Table S4), suggesting that the C-terminal epitope tag did not affect the export function of Tba.

To investigate the substrate specificity of Tba, we complemented *A. balhimycina* Δ*tba* with genes coding for putative GPA transporters and a non-GPA transporter. We used transporters encoded in the BGC of vancomycin (Tva) from *A. keratiniphila* HCCB10007 and ristomycin (Tri) from *A. japonica* MG417-CF17 DSM 44213, both with a 3xFLAG epitope tag. Vancomycin (type I) differs from balhimycin only in the glycosylation pattern (Figure 3a), while ristomycin (type III) differs in the amino acid composition of the backbone and has a higher degree of glycosylation compared to balhimycin (Figure S5). All complemented mutants were compared with *A. balhimycina* Δ*tba*+*tba*^{3xFLAG} instead of the wildtype, ensuring a consistent starting point for analysis. Our results showed significant differences in the export of balhimycin by the ABC transporters used. Tva^{3xFLAG} exhibited a similar export to Tba^{3xFLAG}, with no significant difference ($p>0.05$). In contrast, Tri^{3xFLAG} exported significantly less ($p=0.03$) balhimycin than Tba^{3xFLAG}. Intracellularly, we observed only a slight accumulation of balhimycin in presence of Tva^{3xFLAG} or Tri^{3xFLAG} compared to Tba^{3xFLAG}, but similar to *A. balhimycina* Δ*tba*, it was only detectable at very low levels (<1% of wildtype production). For complementation with a gene encoding a non-GPA ABC transporter, we selected the unrelated but well characterized multidrug transporter Sav1866 from *Staphylococcus aureus* [25], which shares only 33% sequence identity with Tba. In our assay, no export by Sav1866^{3xFLAG} was detected mirroring the findings from *A. balhimycina* Δ*tba*. Additionally, no significant intracellular accumulation was detectable.

The presence of the 3xFLAG epitope tag allowed us to verify proper production and insertion of the transporters into the membrane. This validation was crucial, particularly for transporters exhibiting minimal or no balhimycin export. All ABC transporters were detectable in the membrane fraction by immunodetection after SDS-PAGE (Figure 2e). Furthermore, BN-PAGE analysis following membrane solubilization with the nonionic detergent LMNG demonstrated that each transporter was detectable at the expected size (Figure 2e), indicating that neither aggregation nor incorrect dimerization significantly contributed to the reduced transport activity. However, the varying intensity of the bands indicated that the transporters differ in their accumulation levels in the membrane. For precise protein quantification, alternative methods should be considered in future studies to enable more accurate normalization and comparison of different transporters.

Altogether, the results of the *trans* complementation assay indicate that only the transporter gene from the BGC of a type I GPA, vancomycin, could complement *A. balhimycina Δtba*, whereas the transporter gene from the BGC of a type III GPA, ristomycin could not. Given that vancomycin and balhimycin differ solely in their glycosylation pattern, we speculated that sugars may not play a pivotal role in binding to the ABC transporter, a hypothesis that we set out to validate.

The specificity of the ABC transporter depends on the GPA backbone, rather than the sugar moieties

To further confirm the sugar independence and evaluate potential interaction between GPAs and ABCs, we applied a molecular modeling pipeline encompassing docking and long MD simulations (Figure 3b). We used AlphaFold2[26] to generate structural models of all 89 ABC transporters. The predicted structures were nearly identical with differences only in the unstructured N-terminal region. The models contained all the previously described structural

elements of type IV ABC transporters (Figure 3c; Figure S12)[27]. Models featured a short N-terminal cytosolic helix (elbow helix, (Figure 3c; Figure S12), a conserved TMD core comprising six TM helices, of which two perform a domain swap with those of the other half-transporter, and two coupling helices (*i.e.* CH1+2), responsible for the interaction of the TMD[27]. We docked the type I GPAs vancomycin and balhimycin (Figure 3a) within a predicted binding cavity in the center of the inward-open TMD core (Figure 3d). These proposed binding modes underwent MD simulations and had their protein-ligand interaction frequencies monitored along the trajectory. Representative structures from the simulations displayed conserved interactions between the GPAs' backbone and residues from the TMD core pocket, specifically, GPA amino acids (AA) at position 4-7 (Figure 3d,e) interacting with F269 (AA 6, numbering as in Figure 3a,f, R310 (AAs 5 and 7), G309, and G313 (AA 4). Interestingly, no significant differences were observed in the interaction pattern for AAs 1-3, which are primarily responsible for the classification of the GPAs. We highlight this as a limitation of our current binding model. Additionally, in the proposed binding mode the R1 sugar moieties of GPAs point to the cytoplasm (Figure 3d, e). Based on that, we suggest that the sugar would contribute more to the compound solubility than to the binding energy.

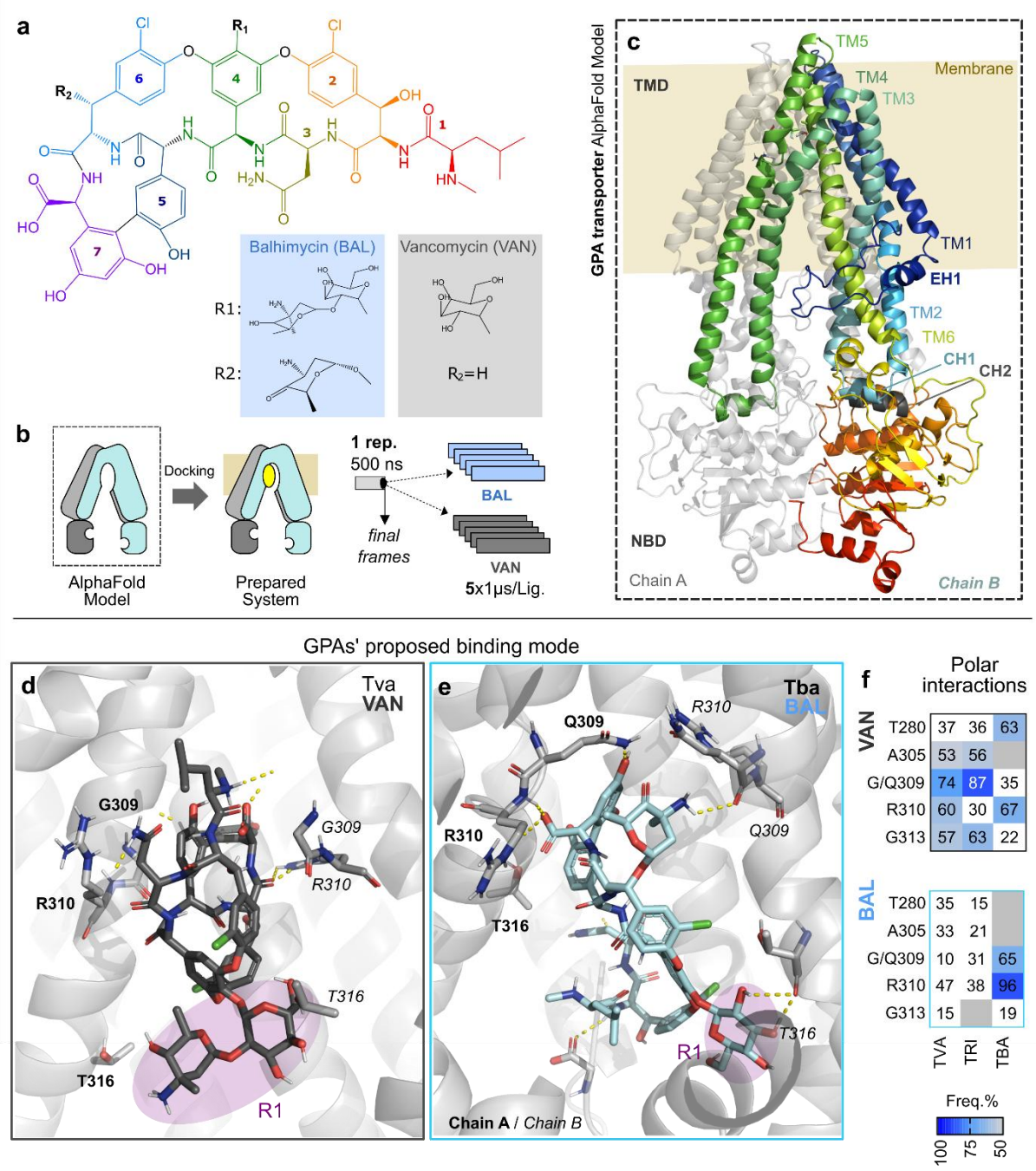


Figure 3. GPA binding mode on GPA transporter. a: 2D schematic representation of the GPAs exemplified by balhimycin and vancomycin, highlighting the individual portions and the insertion points for the sugar moieties. **b:** Overview of the modeling pipeline (described in Star Methods 13). **c:** 3D representation of a type IV GPA transporter, showing its six transmembrane (TM) helices (colored from dark blue to green) and the coupling (CH 1 and 2) and elbow (EH) helices. **d-e:** Proposed binding mode for relevant GPAs, generated from representative conformations from the MD simulation for the transporter Tva bound to vancomycin (**d**) and Tba (**e**) bound to balhimycin. R1 substituent is highlighted in purple. Hydrogen bonds are depicted as yellow dashed lines. **f:** heatmap representations of the hydrogen bond interaction frequencies in the analyzed trajectory. Numbering based on Tva sequence.

In vivo validation of binding interactions between Tba and balhimycin confirms binding of the peptide backbone

Our MD simulation suggested that the transporter's binding affinity towards the GPA substrates relies on interactions with the backbone. Notably, our model points to polar contacts between specific TMD residues and the GPA backbone, prompting us to validate the effect of mutations on transport efficiency.
To address this, we constructed mutants of *A. balhimycina* to probe the impact of Tba mutations on the transport process. Specifically, we introduced mutations in the TMD of the transporter Tba^{3xFLAG} at residues Q309 (G309 in Tva), R310 (conserved in all studied transporters), and T316 (conserved in GPA I-III transporters and located at the base of the substrate binding pocket), substituting them with either glycine or alanine (designated Q309G, R310A, and T316A, respectively). Complementation of *A. balhimycina Δtba* with these constructs resulted in three recombinant strains carrying the individual mutations. Interestingly, only the T316A substitution significantly decreased the export level compared to the transporter Tba^{3xFLAG}, emphasizing the critical role of T316 in the export process. Similar to a strain lacking Tba (Figure 2c,d), the *A. balhimycina* recombinant strain carrying the TbaT316A mutant slightly accumulated intracellular balhimycin, however approximately 85-fold lower than the exported level by the wildtype. In contrast, the Q309G and R310A substitutions neither reduced the extracellular balhimycin levels nor had an effect on intracellular levels (Figure 4a, b; Table S2; Figure S6a). Although T316 is not directly involved in balhimycin binding according to our model, its location on the cytoplasmic side of the putative substrate binding pocket suggests a role in pocket formation and stabilization. Importantly, we ruled out that the reduction in export levels was due to protein misfolding, as $TbaT316A^{3xFLAG}$ was detected at the correct size by SDS-PAGE and BN-PAGE, albeit at lower accumulation levels than the wildtype protein (Figure 4c).

Our MD simulations suggest that the R1 sugar moieties do not interfere with binding (Figure S7), prompting us to assess the effects of GPA modifications on transport efficiency. In particular, we aimed to determine if non-glycosylated and non-chlorinated balhimycin could be exported by Tba. To investigate this, we used an *A. balhimycina* mutant lacking the halogenase gene *bhaA*, responsible for chlorination of the β-hydroxytyrosine residues at AAs 2 and 6 of the balhimycin backbone[28] (Figure 3a), and a mutant lacking the glycosyltransferase gene *bgtfB*, responsible for attaching the first sugar moiety to the hydroxyphenylglycine residue at AA4 (R1, Figure 3a). The deletion of *bgtfB* results in the production of a non-glycosylated heptapeptide, as glycosylation occurs stepwise in a defined order. Missing the first glycosylation step prevents the attachment of subsequent sugars [29]. In addition, we included a double mutant lacking both genes, *bhaA* and *bgtfB*. By analyzing the supernatants via HPLC-MS we identified the peaks of the corresponding final products of each mutant and used the intensities for subsequent quantification. The HPLC-MS analysis of culture supernatants from *A. balhimycina* Δ*bhaA* showed peaks corresponding to masses of *m/z* 1237.5 $[M+H]^+$ and *m/z* 1377.49 $[M+H]^+$ (Figure S6b; Table S3), assignable to balhimycin without chlorine and with either one or two sugar moieties, respectively. In the culture supernatant of *A. balhimycina* Δ*bgtfB*, a peak with a corresponding mass of *m/z* 1142.28 $[M+H]^+$ (Figure S6b; Table S3) was identified as chlorinated balhimycin lacking all sugar moieties. The double mutant exhibited peaks with prominent masses of *m/z* 1074.36 $[M+H]^+$ and *m/z* 1088.38 $[M+H]^+$, corresponding to balhimycin without chlorine atoms and sugar moieties, and the same molecule with an additional methyl group, respectively (Figure S6b; Table S3). All data were confirmed by comparing the UV-spectra derived from the DAD detector with in-house databases (Figure S8).

Interestingly, the absence of sugar modifications on the balhimycin peptide backbone led to a significant increase in exported balhimycin derivatives, approximately 5.5-fold higher in *A. balhimycina* Δ*bgtfB* as compared to the glycosylated balhimycin in the wildtype (Figure 4d),

However, the biological activity decreased in the absence of the sugars, as we confirmed by exposing the indicator strain *Bacillus subtilis* DSM10 to the culture supernatants of the *A. balhimycina* mutants (Figure 4e). Moreover, we observed an even higher export of balhimycin derivatives in the *A. balhimycina* double mutant Δ*bhaA*Δ*bgtfB*. The biological activity of the balhimycin derivative produced by this mutant was totally abolished (Figure 4e), which confirms the importance of these modifications in the mode of action of balhimycin. In *A. balhimycina* Δ*bhaA* the balhimycin derivative lacking chlorine atoms appeared to be exported at the same levels as fully modified balhimycin in the wildtype strain (Figure 4d), and didn't exhibit any decrease in inhibitory activity. This indicates that chlorine atoms do not interfere with the transport process.

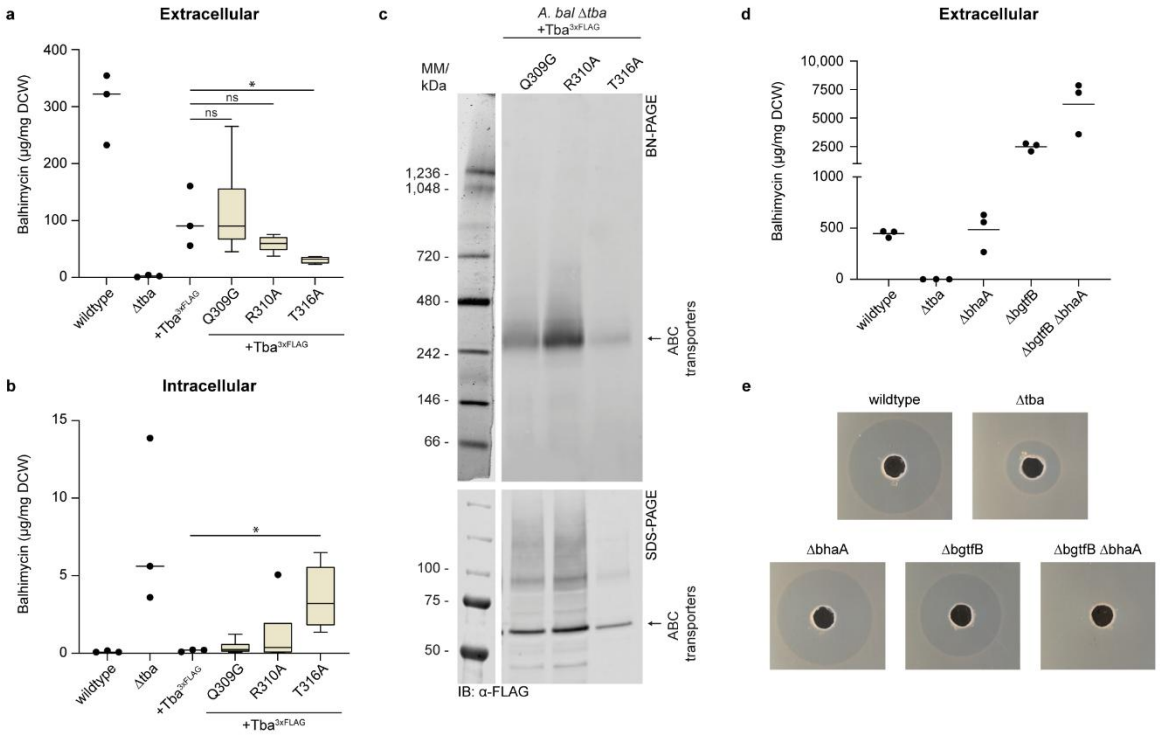


Figure 4. Analysis of transport specificity of the balhimycin transporter Tba. a-b: Quantification of extracellular (**a**) and intracellular (**b**) levels of balhimycin in *A. balhimycina* and respective mutant strains based on HPLC-MS data (Figure S6a). Values were normalized to the dry cell weight (DCW). Data are shown in boxplot representation or by individual data points (+median) if n≤3. Statistical outliers are displayed as dots as well. Statistically significant differences compared to +Tba^{3xFLAG} are highlighted with asterisks ($p<0.05$ (*)) ("Wilcoxon signed-rank test" and "Benjamini-Hochberg Procedure") (Table S2). **c**: Immunoblotting after BN- (up) and SDS-page (down) of solubilized crude membranes of *A. balhimycina* transporter mutant strains. Detection with 3xFLAG specific antibodies. **d**: Quantification of extracellular levels of balhimycin in *A. balhimycina* and respective mutant strains based on HPLC-MS data (Figure S6b). Values were normalized to the DCW. Data are shown by individual data points (+median). **e**: Bioassay on *B. subtilis* DSM10 indicator plates using culture supernatant of the indicated *A. balhimycina* strains.

The ABC transporter Tba interacts with the GPA biosynthetic machinery at the membrane

The low intracellular balhimycin levels in the *A. balhimycina Δtba* mutant and in the mutant strain producing the export-defective transporter TbaT316A^{3xFLAG} (Figure 2d and Figure 4b) were unexpected. We anticipated intracellular accumulation of balhimycin following transporter

deletion or inactivation, considering that *A. balhimycina* produces intrinsically resistant cell wall precursors. To confirm that active transport through Tba is required for the biosynthesis of balhimycin to occur, we constructed a mutant expressing TbaE545Q^{3xFLAG}, a variant that carries a point mutation in the NBD, leading to a catalytically inactive protein unable to hydrolyze ATP. In this mutant, we could only detect negligible amounts of balhimycin both intracellularly and extracellularly (Figure 2c,d). This shows that TbaE545Q^{3xFLAG} is indeed unable to export balhimycin. Furthermore, the mere presence of the transporter is not sufficient to promote balhimycin biosynthesis, which would lead to its intracellular accumulation. Thus, it appears that a regulatory mechanism responsible for the arrest of balhimycin biosynthesis is triggered by the lack of active transport. When not actively exported, balhimycin could negatively regulate the transcription of the biosynthetic genes, thereby decreasing biosynthesis. We conducted a differential expression analysis of the transcriptome of *A. balhimycina* wildtype and *A. balhimycina Δtba* to test this hypothesis (Figure 5a; Figure S9a). We included in our investigation the *A. balhimycina Δbbr* mutant lacking the StrR-like regulator gene *bbr*. Bbr is known to activate the transcription of the BGC genes [30]. No significant difference in BGC expression was observed between the wildtype and *A. balhimycina Δtba* after 48 hours of cultivation. However, a clear downregulation was apparent in *A. balhimycina Δbbr* (Figure 5a). This result shows that the absence of Tba does not affect the transcription of the genes in the BGC.

Since we observed no effect at the transcriptional level, we explored the possibility of biosynthesis inhibition due to the accumulation of precursors or intermediates of balhimycin biosynthesis acting as feedback inhibitors. By measuring the levels of amino acid and different balhimycin intermediates in *A. balhimycina* wildtype and *Δtba* at the same time point (48 h) using LC-MS/MS and flow-injection mass spectrometry, respectively, we observed neither a significant increase or decrease in relevant amino acids (leucine, asparagine, and tyrosine) involved in balhimycin biosynthesis (Figure. 5b), nor accumulation of intermediates (Figure S9b, Figure S10).

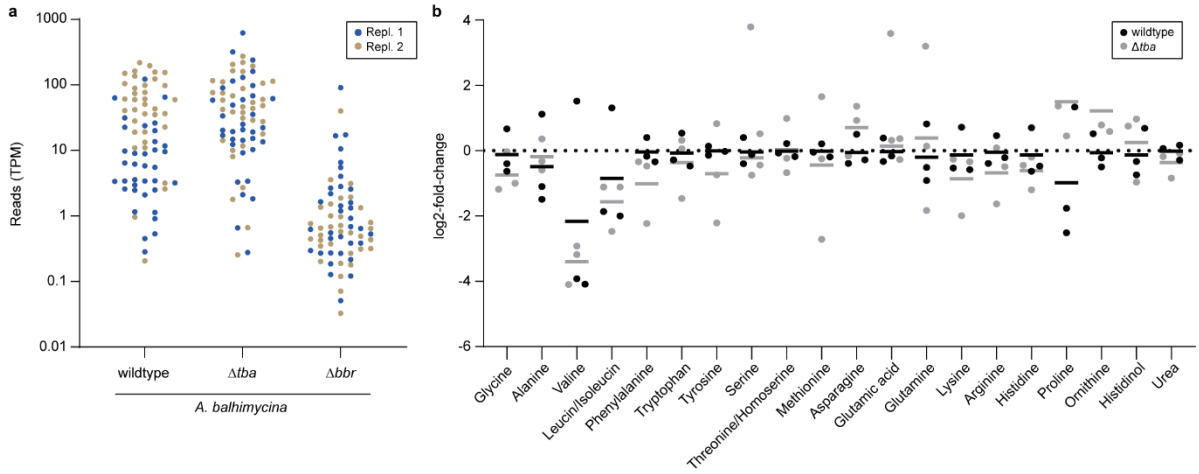


Figure 5. Transcriptome and metabolome analysis of *A. balhimycina Δtba*. a: Display of transcription of every gene within the balhimycin BGC in *A. balhimycina* wildtype, Δ*tba*, and Δ*bbr* in transcripts per million (TPM). Two individual replicates are shown in beige and blue, respectively. **b:** Amino acids levels after 48 h measured by LC-MS/MS and normalized to the dry cell weight (DCW). *A. balhimycina* wildtype (black) and Δ*tba* (gray) are displayed as log2 fold changes in comparison to a 13C internal standard.

Both transcriptome and metabolome analyses have remained inconclusive regarding the cause of the inhibition of balhimycin biosynthesis in the absence of the transporter. To explore the possibility that the transporter might interact with the biosynthetic machinery and thereby

exert a direct regulatory function, we sought to identify the proteins localized in the vicinity of the transporter. To address this question, we employed a proximity dependent biotinylation approach using the recently developed and promiscuous biotin ligase TurboID (TID)[31]. Fusing TID to any protein of interest results in short-range (~10 nm) biotinylation of primary amines on lysine residues of proximal proteins, thereby allowing the identification of the microenvironment of specific proteins[32] (Figure 6a). We successfully established this method in *A. balhimycina DSM 44591* by using Tba as a bait protein for proximity studies (Figure 6b). We generated a recombinant strain by introducing a Tba^{TID} fusion construct into *A. balhimycina Δtba* and confirmed that production of Tba^{TID} restored balhimycin export (Figure S11a). In order to detect biotinylation that is specific for Tba^{TID}, we constructed two control strains of *A. balhimycina*. One strain produces TID as a cytosolic protein, while the other strain produces a GPA-unrelated transporter of *A. balhimycina,* Abc30, fused to TID. To enrich and identify the proteins biotinylated by Tba^{TID} we performed pull-downs with streptavidin beads (Figure S11b), followed by HPLC-MS/MS analysis. This yielded 563 proteins after processing the data. To ensure robust statistical analysis, we excluded hits identified in only one replicate. Intriguingly, comparison between the Tba^{TID} and control strains revealed a clear enrichment of balhimycin biosynthetic enzymes. These included NRPSs (BpsB, BpsC), modification enzymes (OxyC, BhaA, BgtfAB, Bmt), and enzymes involved in precursor supply (Pgat, BpsD, OxyD, HmaS, DvaB, DvaC). The presence of these biosynthetic enzymes in close proximity to the Tba transporter suggests that an active transporter might "stimulate" the production of balhimycin via interaction with the biosynthetic machinery, a hypothesis that awaits confirmation in follow-up studies.

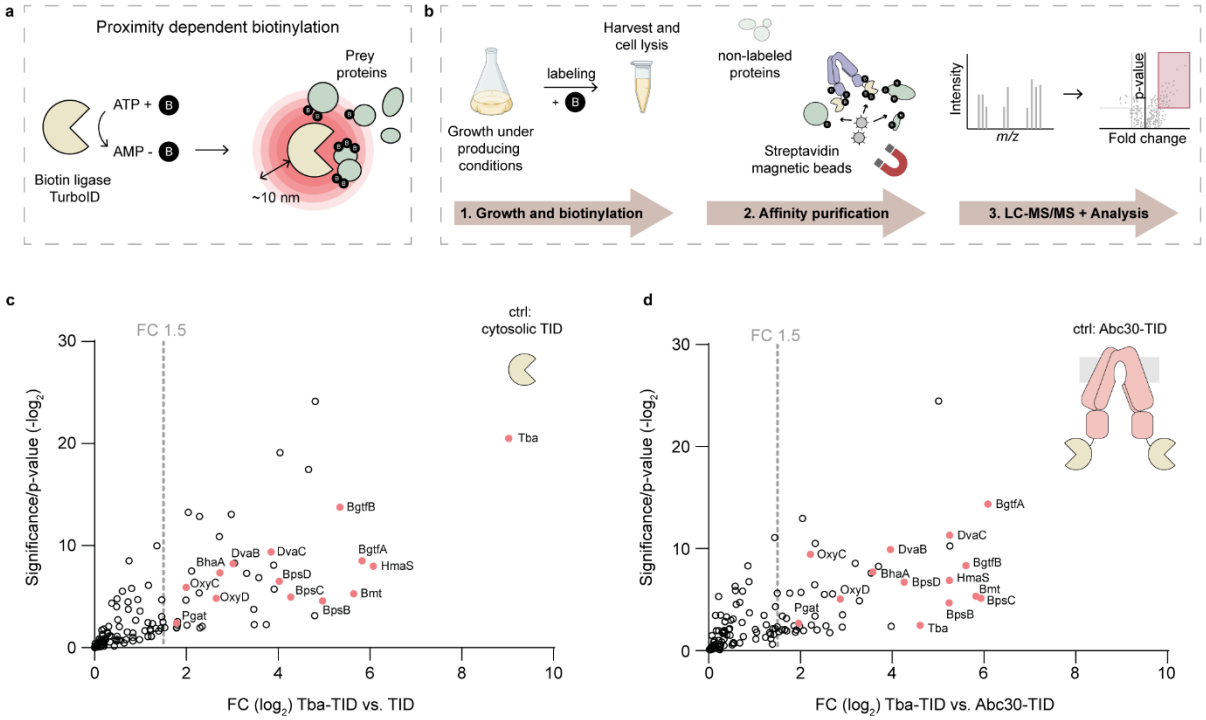


Figure 6. Identification of Tba's microenvironment by proximity dependent biotinylation. **a:** Schematic overview of the concept of proximity dependent biotinylation by the promiscuous biotin ligase TurboID. B: Biotin **b:** Schematic representation of the experimental procedure of proximity dependent biotinylation in *A. balhimycina*. **c-d:** Representation of all log2 fold change positive proteins detected by LC-MS/MS after proximity dependent biotinylation. Hits of Tba-TID compared to cytosolic TID (**c**) or membrane protein Abc30 (**d**). Proteins detected and encoded in the BGC of balhimycin are coloured in salmon. NRPSs (BpsB, BpsC), modification enzymes (OxyC, BhaA, BgtfAB, Bmt), and enzymes involved in precursor supply (Pgat, BpsD, OxyD, HmaS, DvaB, DvaC).

Discussion

Our in-depth study of GPA-related ABC transporters provides valuable new insights into their evolutionary relationships, substrate specificity, and regulatory functions in GPA biosynthesis. The phylogeny of these transporters is consistent with the evolutionary trajectory of their respective substrates, as proposed by Hansen et al. and Waglechner et al. It is suggested that type I GPAs evolved from type IV GPAs[23,33]. Consequently, our analysis implies that the putative ABC transporters of type IV GPAs are more ancestral than those of type I GPAs. Notably, transporters encoded by BGCs of the same GPA type often cluster together. This clustering suggests that their specificity is likely to be determined by differences in the amino acid composition of the GPA backbone rather than by variations in modifications, as GPAs of the same type can exhibit different modification patterns.

This hypothesis was further supported by MD-simulation and *in vivo* export analysis. Only the ABC transporters Tba and Tva, which are encoded in BGCs of type I GPAs, were found to export comparable quantities of balhimycin. In contrast, the putative ristomycin (type III GPA) ABC transporter Tri exported significantly less balhimycin. Our results suggest that the selectivity of GPA-associated ABC transporters is due to the chemical composition of the peptide backbone, since in ristomycin it consists exclusively of aromatic amino acids, whereas balhimycin has two aliphatic amino acids at position AA1 and AA3. We showed that Tba exports a non-glycosylated derivative of balhimycin at significantly higher levels than the fully glycosylated form. This suggests that the bulky and flexible sugar moieties, although crucial for biological activity, impedes efficient transport. There are examples of transporters in ribosomally synthesized and post-translationally modified peptides (RiPPs) produced by a wide range of bacteria that have been experimentally characterized. NisT, a transporter associated with the lantibiotic peptide nisin, is an example of a transporter with a broad substrate specificity[34]. In contrast, the MjcD transporter exhibits high specificity for its cognate substrate, the lassopeptide Microcin J25[35]. In contrast to RiPPs, the specificity of GPA-associated ABC transporters has been less well characterized. Our study reveals that GPA ABC transporters exhibit specificity towards their cognate substrates based on backbone composition, whereas they show greater promiscuity with respect to modifications of it.

In a previous study[10], we showed that *tba* deletion in *A. balhimycina* leads to accumulation of balhimycin, which stands in contrast to the findings of the current study. Here, we demonstrated that deletion or functional inactivation of the Tba transporter significantly reduces the amount of produced balhimycin. The discrepancy may be attributed to differences in the quantification method used; Menges et al. relied on inhibition zone diameters for quantification[10], whereas we employed state-of-the-art HPLC-MS analysis. Importantly, impaired biosynthesis in the *A. balhimycina tba* mutants is not due to toxic effects of balhimycin, as resistance is facilitated constitutively through cell wall precursor remodeling[11,36–38]. Similar findings were reported previously, with deletion of transporter genes in particular BGCs leading to a reduction in the production levels of the respective compounds. Studies on the microcystin BGC in *Microcystis aeruginosa* showed that deletion of the transporter gene *mcyH* completely abolished the production of microcystin[39]. Comparable observations were made in *Streptomyces ghanaensis* ATCC14672, in which the deletion of two transporter genes *moeX5moeP5* and *moeD5moeJ5* led to a significant reduction in the production of moenomycin A[14]. Another example is provided by the NRPS derived compound cereulide and its cognate ABC transporter CesCD. The deletion or inactivation of CesCD resulted in a greatly reduced biosynthesis of cereulide by the producer strain *Bacillus cereus*[15]. These studies point to an important role of the cognate transporters in regulating the biosynthetic machinery of natural products and strongly highlights active transport as a crucial factor for continuous biosynthesis.

To elucidate the mechanism behind transport-dependent balhimycin production, we investigated different levels of regulation and showed that transcription is not affected by the

absence of Tba, as previously observed for microcystin biosynthesis in the absence of the cognate transporter[39]. Feedback regulatory mechanisms have been reported in balhimycin biosynthesis, involving enzymes and intermediates from the shikimate-pathway [40,41]. It was shown that the aromatic amino acids tyrosine and phenylalanine act as feedback enzymatic inhibitors of the shikimate pathway, thereby reducing the provision of balhimycin precursors. Our metabolomic analysis shows that none of the precursors accumulated in the *A. balhimycina Δtba* deletion mutant, ruling out an inhibitory effect on biosynthetic enzymes. Furthermore, none of the intermediates of the NRPS assembly line appeared to be present at significantly higher concentrations. We conclude that feedback inhibition is not the reason for reduced biosynthesis of balhimycin in the absence of a functional transporter.

The results of our proximity dependent biotinylation experiment provide an alternative explanation: we showed that many of the enzymes responsible for assembly and modification of the balhimycin backbone as well as supply of precursors are in close proximity to Tba, possibly forming a microcompartment specialized in GPA production. Based on the general ABC transporter mechanisms[42], we speculate that upon substrate binding, Tba undergoes a conformational change that is required for stimulation of balhimycin biosynthesis. However, it remains elusive whether the transporter directly interacts with the biosynthetic enzymes and how exactly this interaction impacts balhimycin biosynthesis. There is evidence that biosynthetic enzymes of secondary metabolites interact with the corresponding transporter and that compound biosynthesis might be membrane-associated, as exemplified by the cereulide and nisin biosynthesis. Both biosynthetic machineries were shown to build a complex and co-localize with the corresponding ABC transporter[15,43,44].
Among the biotinylated proteins with significant fold changes in the Tba^{TID} mutant, we identified a paraslipin protein, suggesting its close proximity to Tba and consequently to other proteins of the balhimycin biosynthetic machinery. Paraslipins belong to the stomatin, prohibitin, flotillin, and HflK/C (SPFH) superfamily and are known to be associated with the formation of functional membrane microdomains (FMM)[45]. Our finding suggests that Tba may be integrated into an FMM, serving as an anchor of biosynthetic enzymes. We speculate that the biosynthesis of balhimycin occurs within a specialized microcompartment and is likely regulated by the presence of catalytically active Tba, as suggested by our *in vivo trans* complementation assays. There is evidence that bacteria use microcompartments to optimize their metabolic processes. Carboxysomes are among the best studied and represent an example of how compartmentalization can increase the efficiency of metabolic reactions, in this specific case of carbon fixation by photoautotrophs and chemoautotrophs. Carboxysomes concentrate CO_2 owing to the activity of a carbonic anhydrase and the selective permeability of the protein shell, which in turn increases both the rates and specificity of the encapsulated carbon-fixing enzyme (ribulose 1,5-bisphosphate carboxylase/oxygenase) RuBisCO[46]. Another example is the propanediol-utilizing microcompartment (Pdu MCP) of enteric bacteria like *Salmonella*. The Pdu MCP allows the degradation of 1,2-propanediol and enteric pathogenesis[47]. However, further studies are required to confirm that in *A. balhimycina* GPA biosynthesis occurs in compartments associated with the membrane, and to fully elucidate the scope of the transporter's potential roles in this process.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Evi Stegmann (evi.stergmann@uni-tuebingen.de)

Materials availability

This study did not generate new unique reagents.

Data and code avilability

- All MD simulation trajectories, interaction data and MD quality control data, representative PDB files from the alphaFold models, phylogenetic trees are available under the DOIs: 10.5281/zenodo.7547342, 10.5281/zenodo.7732071 and 10.5281/zenodo.7547403 and are publicly available as of the date of publication.
  All RNA-Seq Illumina read files as well as the raw counts have been deposited in NCBI's Gene Expression Omnibus and are accessible under accession number GSE274067
  The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[102] partner repository with the dataset identifier PXD054387
- This study did not generate any unique code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Limitations of the study

Despite detailed investigation of the substrate specificity using *in-silico* models, *in vitro* studies are essential to fully characterize the binding pocket of the ABC transporter Tba for the GPA balhimycin. Structural elucidation of Tba-balhimycin complex would provide crucial insights into the binding mechanism.
Additionally, the hypothesis that Tba interacts with the biosynthetic machinery requires further validation. Investigating the regulatory function of Tba through direct protein-protein interaction studies would be key to understanding its precise impact on the balhimycin biosynthesis.

Acknowledgements

The work in the laboratory of S.W., E.S., W.W., N.Z., B.M., and K.N. related to this study was funded by the Deutsche Forschungsgemeinschaft (DFG) as part of the transregional collaborative research center TRR 261 'Cellular Mechanisms of Antibiotic Action and Production' (Projects B01/B02/B05/Z03). It was supported by the de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) and ELIXIR-DE (Forschungszentrum Jülich and W-de.NBI-001, W-de.NBI-004, W-de.NBI-008, W-de.NBI-010, W-de.NBI-013, W-de.NBI-014, W-de.NBI-016, W-de.NBI-022). NGS sequencing methods were performed with the support of the DFG-funded NGS Competence Center Tübingen (INST 37/1049-1). Data management and storage of raw data for this project were supported by the Quantitative Biology Center (QBiC), University of Tübingen, Germany. The authors would like to thank Libera Lo Presti for her support in writing, comments on the manuscript, and fruitful discussions. S.W., E.S., W.W., T.K., and N.Z. acknowledges the German Center for Infection Research (DZIF, TTU09.716). S.W., E.S., W.W., T.K., and N.Z. acknowledges funding by the Clusters of Excellence EXC2124 CMFI (project ID 390838134) and T.K. the TüCAD2 the Federal Ministry of Education and Research (BMBF) and the Baden-Württemberg Ministry of Science as part of the Excellence Strategy of the German Federal and State Governments – Germany, by the means of the program $TüCAD_2$, as well as the German Center for Infection Research (T.K.: DZIF, TTU06.716) (W.W. TTU09.826). The authors would like to thank the CSC - Finland for the generous provided resources, Ana Monica Daneliuc for contributing to refining HMM search for the creation of the BGC dataset and Anke Biedermann and Irina

Droste-Borel from the Proteome Center Tübingen for technical support. Graphical abstract was created in BioRender. Stegmann, E. (2025) https://BioRender.com/x66x647

Author contributions

N.G. performed *in silico* analysis of transporter sequences and phylogenetic analysis. N.G. and D.B. performed *in vivo* experiments and analyzed all data. T.K. performed MD simulations. S.W., E.S., and W.W. designed the study. N.G., D.B., T.K., S.W., E.S., and W.W. drafted the paper and wrote the original manuscript. A.K. performed HPLC-MS measurements. M.F.W. and B.M. performed proteomics. A.G. identified BGCs and putative transporters and did the correlation of BGCs and substrates. U.S. did statistical analysis of *in vivo* data. J.R. and H.L. performed metabolomics. T.H. and K.N performed RNA sequencing and analyzed the data. I.G. provided bioinformatic support.

Declaration of interest

The authors declare no competing interests.

STAR Methods
Key resource table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-FLAG® M2	Sigma-Aldrich	F3165
Rabbit polyclonal anti-BirA IgG	Thermo-Fisher	PA5-80251
goat anti-Mouse IgG DyLight™ 800	Thermo-Fisher	SA5-35521
goat anti-Rabbit IgG DyLight™ 800	Thermo-Fisher	SA5-35571
Streptavidin DyLight™ 800	Thermo-Fisher	21851
Bacterial and virus strains		
See Table S5 in the SI	This work	
Biological samples		
Chemicals, peptides, and recombinant proteins		
Q5® Hot Start High-Fidelity DNA Polymerase	NEB	M0491L
Phusion® High-Fidelity DNA Polymerase	Thermo-Fisher	F-530XL
Taq DNA Polymerase	Thermo-Fisher	EP0401
NdeI	Thermo-Fisher	ER0585
XbaI	Thermo-Fisher	ER0683
KpnI	Thermo-Fisher	ER0522
Apramycin	Genaxxon	M3450.0005
Ampicillin	Roth	K029.4
Erythromycin	Roth	4166.2
Critical commercial assays		
Deposited data		
MD simulation trajectories, interaction data and MD quality control data, representative PDB files from the AlphaFold models and phylogenetic trees	This work	10.5281/zenodo.7547342 10.5281/zenodo.7732071 10.5281/zenodo.7547403
RNA-Seq Illumina read files as well as the raw counts	This work	GSE274067
Mass spectrometry proteomics data	This work	PXD054387
Metabolomics raw data	This work	10.5281/zenodo.14918266
Experimental models: Cell lines		
Experimental models: Organisms/strains		
Oligonucleotides		
See Table S4 for cloning primers	This work	

Recombinant DNA		
See Table S6 for plasmids	This work	
Software and algorithms		
AlphaFold (v2.2.2)~~/Alphafold3~~	Jumper, J. *et al*. [26]	
Desmond	Bowers, K. J. *et al.* [74]	
OPLS4 force-field	Lu, C. *et al.* [75]	
Glide	Friesner, R. A. *et al.* [71]	
LigPrep	Shelley, J. C. *et al.* [69]	
Maestro (v2022.4)	www.schrodinger.com	
prokka (v1.14.16)	Seemann, T. [48]	
OrthoFinder platform (v2.3.11)	Emms, D. M. & Kelly, S. [49]	
blastP	https://blast.ncbi.nlm.nih.gov/	
scipy (v1.9.3)	Virtanen, P. *et al.* [50]	
ΔG prediction server (v1.0)	Hessa, T. *et al.* [19]	
MEGAX (v10.2.4)	Kumar, S. *et. al.* [51]	
MUSCLE algorithm	Edgar, R. C. [53]	
iTOL (v6.6)	Letunic, I. & Bork, P. [55]	
Other		
Serva, NativePAGE™ 3 to 12% (Bis-Tris, 1.0 mm, Mini Protein Gels)	Thermo-Fisher	BN1001BOX
Immun-Blot® polyvinylidene difluoride (PVDF) membrane (0.2 µm)	Bio-Rad	#1620177

Experimental model and study participant details
The bacterial strains and plasmids used in this study are described with the necessary information in Table S5 and Table S6.

Method details

1. Identification of transporter sequences of biosynthetic gene clusters

The biosynthetic gene clusters (BGC) of glycopeptide antibiotics (GPAs) were identified either by literature search[2] or by a hidden Markov model (HMM) search in the NCBI database using the X-domain sequence of non-ribosomal peptide synthetases (NRPS) of known GPAs[4]. If necessary, genomes and clusters were re-annotated with prokka (v1.14.16) using default parameters with specified genus[48]. Subsequently, genes encoding putative ATP binding cassette (ABC) transporters were identified by using the OrthoFinder platform (v2.3.11)[49]. This tool classifies all genes into groups whose members are orthologous with each other and therefore suitable for phylogenetic analyses. The correlation of BGC to GPA types was done manually. Every gene in the BGC was blasted and studied in order to classify the BGC in one of the five classical types of GPAs (I-V). The criteria used were the predicted amino acids at positions one and three of the backbone, the number of P450 monooxygenase genes and the presence of an acyltransferase-encoding gene.

2. Bioinformatic analysis of transporters

Amino acid sequence identities of the transporters were calculated by blastP pairwise comparison and displayed in a heatmap representation. The order of the transporters is based on a hierarchical clustering calculated with the python library scipy (v1.9.3) and using the function "scipy.cluster.hierarchy.linkage". The algorithm used was “Weighted Pair Group Method with Arithmetic Mean” (WPGMA), which builds a hierarchy of the clusters (of similar transporters) using an agglomerative approach[50]. The full protein scan option of the ΔG prediction server (v1.0)[19] was used to identify putative transmembrane (TM) helices (full protein scan; helix length 19-23; +length correction). Common sequence motifs of the NBDs of the ABC transporters[18] were manually identified after multiple sequence alignment (MSA) of all transporters using MEGAX (v10.2.4)[51].

3. Phylogenetic analysis of GPA associated ABC transporter

The evolutionary relationship between GPA associated ABC transporters was analyzed by phylogenetic tree construction. MSA, model testing, and calculation of the Maximum Likelihood phylogenetic tree were performed using the software MEGA-X (v10.2.4)[51]. Trimming of the MSA was done with the trimAl tool (v1.4.rev15 build[2013-12-17]) using default parameters[52], or manually. An unrelated ABC transporter (Abc30) of *A. balhimycina* was used as an outgroup. The phylogenetic tree was constructed as follows: all sequences were aligned using the MUSCLE algorithm[53] and then used to analyze the best amino acid substitution model. The model testing identified as the most suitable the model of Le & Gascuel (LG)[54] with frequencies (+F) and gamma distributed rates (number of discrete gamma categories = 5) (+G), and this was used by the ML algorithm with bootstrap test of 1000 repetitions to calculate phylogenetic trees. Visualization of phylogenetic trees was performed with iTOL (v6.6)[55].

4. Chemicals and reagents

Unless otherwise indicated, chemicals and enzymes were purchased from Sigma-Aldrich, Thermo-Fisher, and New England Biolabs. Primers listed in table S4 were synthesized by Eurofins Scientific or IDT. Antibodies were purchased from Sigma-Aldrich (Mouse monoclonal anti-FLAG® M2 (F3165)) and Thermo-Fisher (Rabbit polyclonal anti-BirA IgG (PA5-80251); goat anti-Mouse IgG DyLight™ 800 (SA5-35521); goat anti-Rabbit IgG DyLight™ 800 (SA5-35571); Streptavidin DyLight™ 800 (21851)). SERVAGel™ TG PRiME™ 8–16% precast gels were purchased from Serva, NativePAGE™ 3 to 12% (Bis-Tris, 1.0 mm, Mini Protein Gels) from Thermo-Fisher, and Immun-Blot® polyvinylidene difluoride (PVDF) membrane (0.2 µm) from Bio-Rad.

5. Bacterial strains, plasmids, and growth conditions

All bacterial strains and plasmids used in this study are listed in table S5 and table S6, respectively. *E. coli* was cultivated aerobically at 37°C in 10 ml tubes and shaken at 180 rpm in liquid LB medium (0.5% (w/v) yeast extract, 1% (w/v) tryptone, 85.56 mM NaCl) or on solid LB agar (1.5% (w/v)) medium. If necessary, the medium was supplemented with 100 µg/ml apramycin or 100 µg/ml ampicillin for plasmid selection. Unless otherwise indicated, *A. balhimycina* was cultivated aerobically at 29°C in baffled flasks with a steel coil, and shaken at 120 rpm in R5 medium[56] or on the respective solid agar (1.5% (w/v)) medium. Fifty µg/ml apramycin or 50 µg/ml erythromycin was used for plasmid selection. All *A. balhimycina* strains in this study are derivatives of DSM 44591.

6. Molecular cloning

Different DNA modification techniques were used in this study. Polymerase chain reaction (PCR) with “Q5® Hot Start High-Fidelity DNA Polymerase” and “Phusion® High-Fidelity DNA Polymerase” were used to amplify genes and DNA fragments, while “Taq DNA Polymerase” was used for colony PCR. pRM4[10] vector DNA was linearized using the restriction endonucleases *NdeI/XbaI*, pSP1[57] vector was linearized using the restriction endonuclease

KpnI. Assembly of DNA fragments was performed using standard gibson assembly method[58]. All primers used in this study are listed in Table S4.

7. Genetic manipulation of *A. balhimycina*

Chromosomal integration of genes into the genome of *A. balhimycina* was achieved by direct transformation (D-Trafo), as first described for *Amycolatopsis mediterranei*[59] and later adapted for *A. balhimycina*[57]. Demethylated plasmid DNA was isolated from *E. coli* JM110 or ET12567 and used for the transformation of *A. balhimycina* after 48 h of growth. Correct integration was confirmed by PCR.

8. Generation of *A. balhimycina* deletion mutants

In-frame *A. balhimycina* deletion mutants were generated using the non-replicative plasmid pSP1[57], which contains homologous flanking regions of ≈1.5 kB upstream and downstream of the target genes *tba* or *bgtfB*, resulting in the plasmids pSP1Δ*tba* and pSP1Δ*bgtfB*, respectively. *A. balhimycina* was transformed with the plasmids using D-Trafo. For single mutants, pSP1Δ*tba* and pSP1Δ*bgtfB* were introduced into *A. balhimycina* wildtype, whereas for the generation of the double mutant *A. balhimycina ΔbhaAΔbgtfB,* we introduced pSP1Δ*bgtfB* into the existing mutant *A. balhimycina ΔbhaA*[28]. Erythromycin resistant clones harboring either pSP1Δ*tba* or pSP1Δ*bgtfB* were confirmed by PCR using the primers P1-2 (Table S4), and subsequently used for the stress protocol to increase the frequency of double crossover events[28]. The resulting protoplasts were streaked onto R5 agar plates and R5 agar plates supplemented with 50 µg/ml erythromycin for selection. Erythromycin sensitive clones were screened for in-frame deletion by PCR using primers P9-10 and P15-16 (Table S4).

9. Crude membrane preparation

Crude membranes of *A. balhimycina* were prepared as reported previously by the group[60]. In brief, 300 mg of culture wet weight was used. The cells were washed with 1xPBS, resuspended in 750 µl buffer K (50 mM triethanolamine (TEA), pH 7.5, 250 mM sucrose, 1 mM EDTA, 1 mM $MgCl_2$, 10 µg/ml DNAse, 2 mg/mL lysozyme, 1:100 protease inhibitor cocktail (Sigma-Aldrich (P8849)) and incubated for 30 min at 4°C. Afterwards, the cells were mixed with glass beads (Ø=150-212 µm) and lysed through a bead mill (2 min, continuously). Cell debris was removed by centrifugation at 10,000 x *g* for 10 min. The crude membranes were precipitated by centrifugation at 55,000 x *g* at 4°C for 45 min and resuspended in 75 µl 1xPBS.

10. BN-PAGE, SDS-PAGE, and immunoblotting

Analysis of native transporter complexes was performed as described previously[61]. Crude membranes were solubilized for 1 h at 4°C using 1% Lauryl Maltose Neopentyl Glycol (LMNG). Non-solubilized materials were separated by centrifugation at 100,000 *x g* at 4°C for 30 min. Fifteen µg of solubilized membrane proteins were subjected to BN-PAGE (NativePAGE™ 3 to 12%) before transfer to a PVDF membrane. Similarly, for the analysis of denatured proteins, samples were subjected to SDS-PAGE (SERVAGel™ TG PRiME™ 8-16%). Membranes were probed with mouse anti-FLAG primary antibody (1:10,000), rabbit anti-BirA (1:5,000), anti-mouse secondary antibody (1:10,000), anti-rabbit secondary antibody (1:10,000), or Streptavidin DyLight™ 800 (1:10,000). Detection and analysis were performed using a Li-Cor Odyssey system and image Studio (Li-Cor) (v5.2).

11. Balhimycin production, export assay and bioactivity assay

All relevant strains were inoculated in 20 ml of R5 medium and incubated for 48 h as preculture. Five ml of the preculture was used for inoculation of 100 ml main culture in R5 medium. After 96 h of growth, 10 ml of every culture was sampled and separated by centrifugation into supernatant and mycelium. The supernatant was directly used for bioactivity assays as well as for High Performance Liquid Chromatography-Electrospray Ionisation-Mass Spectrometry (HPLC-ESI-MS) measurements. The mycelium was further processed in order to extract intracellular balhimycin. First, it was washed with 10 ml deionized water (diH_2O), 5 ml carbonate buffer (pH 9.7) and again with 10 ml diH_2O, to remove trace amounts of cell wall

bound balhimycin. Then, the mycelium was lyophilized and weighed for dry cell weight (DCW) determination and used for normalization. Five ml of methanol was used to extract balhimycin from the mycelium (12-16 h). The methanol fraction was separated from the mycelium by centrifugation and collected. The mycelium was washed with 5 ml diH_2O and the supernatant was also collected. The methanol and water wash fractions were pooled and the extract was evaporated using a Genevac EZ-2 (Genevac Ltd, Suffolk, UK). Finally, the evaporated extracts were dissolved in 0.3 ml of H_2O and used for bioactivity assays as well as HPLC-ESI-MS measurements. For bioactivity assays we used MM1 medium[62] indicator plates containing *B. subtilis* DSM10 spores ($5x10^6$ spores/mL) and applied 30-40 µl culture supernatant of various *A. balhimycina* strains. The plates were incubated at 37°C overnight.

12. HPLC-MS analysis and quantification of balhimycin

For the detection of balhimycin 2.5 µl of the culture supernatants and mycelium extracts were analyzed by means of HPLC–ESI-MS using a Nucleosil 100-C18 column (3 µm, 100 by 2 mm) (precolumn, 10 by 2 mm) (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) coupled to an ESI mass spectrometer. LC-MS measurements were obtained from the liquid chromatograph/mass selective detector (LC/MSD) Ultra Trap system XCT 6330 (Agilent Technologies, Waldbronn, Germany). Analysis was carried out at a flow rate of 400 µl/min with gradient elution. Solvent A was 0.1% formic acid, and solvent B 0.06% formic acid in acetonitrile. Gradient elution was performed as follows: t0=0% B, t15=t17=100% B, post time 5 min. 0% B. The flow rate was 400 µl/min, and the temperature was 40°C. For MS analysis an electrospray ionization (alternating positive and negative ionization) in Ultra Scan mode with a capillary voltage of 3.5 kV and a drying gas temperature of 350°C was used. The detection of m/z values was performed with Agilent DataAnalysis for 6300 series Ion Trap LC/MS 6.1 software (v3.4) (Bruker-Daltonik GmbH).

13. Molecular modeling and molecular dynamics simulation

13.1 Protein structure prediction. The structural model of relevant transporter members was retrieved from the AlphaFold (v2.2.2) Protein Structure Database using the multimer preset[26]. All structure models can be found in the Zenodo repository or upon reasonable request. Previous Sav1866 unbiased simulations[68], starting from the experimentally available outward open conformation (PDB 2HYD), revealed small conformational changes in the NBD. Alternatively, steered dynamics suggested that, despite the large opening-closing NBD movement, TMD conformational changes are indepent. Based on that, we hypothesized that our current simulation length would be insufficient to observe the inward-open transition. This prompted us to generate an inward-open Sav1866 model (UniProt ID: Q99T13, similar to PDB 5MKK), in order to allow a direct comparison against the Tva, Tri and Tba inward-open models. Sav1866 model structure was generated based on the heterodimeric ABC transporter TmrAB (PDB ID: 5MKK), selected based on sequence similarity, using the Homology Model package from Maestro (v2022.4).

13.2 Binding site prediction and molecular docking. System preparation and docking calculations were performed using the Schrödinger Drug Discovery suite for molecular modeling (v2022.4). Protein−ligand complex was prepared with the Protein Preparation Wizard to fix protonation states of amino acids, add hydrogens, and fix missing side-chain atoms. All ligands for docking were drawn using Maestro and prepared using LigPrep[69] to generate the 3D conformation, adjust the protonation state to physiological pH (7.4), and calculate the partial atomic charges with the OPLS4 force field. Docking studies with the prepared ligands were performed using Glide (v7.7)[70,71] with the flexible modality of induced-fit docking with extra precision (XP), followed by a side-chain minimization step using Prime. Ligands were docked within a grid around 12 Å from the centroid of the predicted binding site pocket, determined using SiteMap.

13.3 Molecular dynamics simulation. MD simulations for similar ABC transporters were previously validated by the group[72,73]. MD simulations were carried out using Desmond[74], with the OPLS4 force-field[75]. The simulated system encompassed the protein-ligand complexes, a predefined water model (TIP3P[76]) as a solvent and counterions. The system was treated in an orthorhombic box with periodic boundary conditions specifying the box's shape and size as 10 Å distance from the box edges to any atom of the protein. POPC membranes were assigned to the transmembrane helices using the System Setup, with standard options. In all simulations, we used a time step of 1 fs, the short-range coulombic interactions were treated using a cut-off value of 9.0 Å using the short-range method, while the Smooth Particle Mesh Ewald method (PME) handled long-range coulombic interactions[77]. Initially, the system's relaxation was performed using Steepest Descent and the limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithms in a hybrid manner, according to the established protocol available in the Desmond standard settings. During the equilibration step, the simulation was performed under the NPT ensemble for 5 ns implementing the Berendsen thermostat and barostat methods[78]. A constant temperature of 310 K was kept throughout the simulation using the Nose-Hoover thermostat algorithm[79] and Martyna-Tobias-Klein Barostat[80] algorithm to maintain 1 atm of pressure, respectively. After minimization and relaxation of the system, we continued with the production step of at least 800 ns, with frames being recorded/saved every 1,000 ps. Five independent replicas were produced for each substrate, resulting in a total of 4 µs simulation/ligand. Trajectories and interaction data are available on the Zenodo repository (see data availability session). The representative structures were selected by inspecting changes in the Root-mean-square deviation (RMSD), meaning for figures a representative frame was selected at random at points of the trajectory where the RMSD for were not fluctuating, after equilibration. Variation of the RMSD values along with the simulation, for both template crystal structures and simulations with docking pose are provided in the repository. Additionally, the changes in the Root-mean-square fluctuation (RMSF), normalized by residue for the protein backbone, are also provided in the Zenodo repository.

14. Transcriptomic analysis

For the transcriptomic analysis, *A. balhimycina* wildtype, Δ*tba* and Δ*bbr* were cultivated in balhimycin production conditions for 48 h followed by RNA extraction using the “Zymo Quick RNA Fungal/Bacterial Kit” (Zymo Research, CN#R2014) according to the manufacturer’s instructions. The RNA integrity was checked by agarose gel electrophoresis and residual DNA was removed by DNAseI treatment using the “DNA-free™ DNA Removal Kit” (ThermoFischer, CN#AM1906). Successful digestion was initially checked by PCR. The total RNA was quantified with a “Qubit RNA BR Assay Kit” (Thermo Fisher) and RNA integrity was checked by an Agilent 2100 BioAnalyzer with the “RNA 6000 Pico kit” (Agilent). Library preparation and ligation were performed using “Illumina Stranded Total RNA Prep” and “Ribo-Zero Plus Microbiome”, respectively, according to the manufacturer’s instructions. In brief, 100 ng of total RNA per sample were subjected to rRNA depletion, followed by cDNA library construction, adapter ligation, and 15 cycles of barcoding PCR. The obtained libraries were quantified with the “Qubit 1x DNA HS Assay Kit” (Thermo Fisher) and the fragment distribution was checked with an Agilent 2100 BioAnalyzer using the “High Sensitivity DNA Kit” (Agilent). Libraries were subsequently pooled and sequenced on an Illumina NovaSeq 6000 device using “NovaSeq 6000 SP Reagent Kit” (v1.5) (100 cycles) with a run mode 75,10,10,0. The average number of reads obtained was 13-21x106. The sequencing was demultiplexed with the latest version of the nextflow pipeline: nf-core/demultiplex. For demultiplexing, “bcl2fastq” was used and the quality was checked with “fastp”.

Sequencing statistics, including the quality per base and adapter content assessment, were conducted with FastQC (v0.11.8)[81]. All reads mappings were performed against the assembled strain *A. balhimycina* DSM 44591. The strain was annotated using BAKTA (v1.9.1)[82]. The annotation of the cluster genes (positions 6028489 to 6094273) was manually curated. The mappings of all samples were conducted with HISAT2 (v2.1.0)[83], using the following parameters: spliced alignment of reads was disabled and strand-specific information was set to reverse complemented (HISAT2 parameter --no-spliced-alignment and --rna-strandness

"R"). The resulting mapping files in SAM format were converted to BAM format using SAMtools (v1.9)[84]. Mapping statistics, including strand specificity estimation and percentage of mapped reads, were conducted with the RNA-Seq module of QualiMap2 (v2.2.2-a)[85]. Gene counts for all samples were computed with featureCounts (v1.6.4)[86], where the selected feature type was set to transcript records (featureCounts parameter -t transcript). A quality check for ribosomal rRNA was performed with a self-written script based on the absolute counts of annotated rRNAs. To assess variability across the replicates of each time series, a principal component analysis (PCA) was conducted with the DESeq2 package (v1.28.1)[87].
For the computation of genes differentially expressed between the mutants, DESeq2 was applied to the absolute gene counts as computed with featureCounts. For differences between the two mutants and the wildtype strain, genes with an adjusted p-value (FDR) < 0.05 and absolute log2 fold change (FC) > 1 were reported as differentially expressed.

15. Measurement of amino acid levels by metabolomic analysis
For metabolomic measurements, *A. balhimycina* wildtype and Δ*tba* were cultivated under balhimycin production conditions and harvested after 48 h. Two hundred µl of liquid cultures were filtered (Merck, Durapore® 0.45 µm PVDF) using a vacuum pump. The obtained biomass was extracted with 1 mL of an acetonitrile:methanol:water solution (40:40:20) for 1 h at -20°C. The mixture was transferred to a tube with glass beads (Ø=0.1-0.11 mm) and homogenized using a Precellys (2x 30 sec, 6.5 m/s). The mixture was centrifuged at -9°C and 13,000 rpm for 15 min. The supernatant was subsequently used for the LC-MS/MS measurement. Amino acids were measured via targeted LC-MS/MS as described previously[88]. The obtained ratios were further processed by normalizing the values to the dry cell weight (DCW). These values were then used to calculate the differences between the Δ*tba* mutant and the wildtype. The differences were expressed as log2 fold changes in relation to a 13C internal standard.

16. Proximity dependent biotinylation and proteomic analysis
For the proximity dependent biotinylation studies, *A. balhimycina* was cultivated under balhimycin producing conditions, as described in previous sections. Fifty ml of culture was harvested after 48 h at 5,000 *x g*, 4°C for 10 min. The cell pellet was washed with 20 ml of 1xPBS buffer and centrifuged at 4,600 *x g*, 4°C for 10 min. Biotin labeling was carried out in 20 ml 1xPBS supplemented with 500 µM biotin for 1 h at 37°C at 650 rpm. The reaction was stopped at 4°C and three subsequent wash steps with ice cold 1xPBS. The cells were resuspended in 15 ml lysis buffer (0.1 M Tris-HCl (pH 8.0), 0.15 M NaCl, 1 mM EDTA, 10 µg/ml DNAse, 2 mg/mL lysozyme, 1:100 protease inhibitor cocktail (Sigma-Aldrich (P8849), 1 mM $MgCl_2$) and sequentially disrupted by sonication (Branson sonifier 250) (output control "4", 35% duty cycle 2x 30 sec/10 sec pause), followed by homogenization by a Constant System CF-1 homogenizer (I&L Biosystems) (2x 40,000 psi). The cell lysate was cleared twice at 10,000 *x g*, 4°C for 3 and 10 min. The protein concentration was measured using Pierce™ BCA Protein Assay Kits (Thermo Fisher). Five mg of total protein (Input) was used for enrichment of biotinylated proteins using 150 µl of MagStrep® Strep-Tactin® beads (IBA). Binding of proteins to the streptavidin magnetic beads was performed by overhead rotation at 4°C overnight. Washing and elution were performed as described by the manufacturer. Input, flowthrough, and elution fractions were applied to an SDS-PAGE gel as a control and probed after western blotting with an anti-BirA antibody or streptavidin.
For subsequent mass spectrometry analysis, 20 µl of every eluate fraction was loaded on an SDS-PAGE gel and run until the loading front reached approx. 1 cm of the gel. The proteins were stained using a standard Coomassie staining solution. Until further processing, the gel was stored in 5% acetic acid. The bands were excised and cut further into small pieces. Proteins were then digested in the gel with trypsin, whereby all incubation steps were carried out under shaking conditions: for destaining, gel pieces were washed three times with 5 mM ammonium bicarbonate (ABC) in acetonitrile (ACN) (1:1, v/v) for 20 min. After a dehydration step with 100% ACN for 10 min, disulfide bonds were reduced by adding 10 mM dithiothreitol (DTT) in 20 mM ABC for 45 min at 56°C. Thiol groups were carbamidomethylated with 55 mM iodoacetamide (IAA) in 20 mM ABC for 45 min in the dark. Subsequently, gel pieces were

washed two times with 5 mM ABC in ACN (1:1, v/v) for 20 min and dehydrated with 100% ACN for 15 min. Liquid was evaporated by vacuum centrifugation for 10 min and sequencing grade trypsin (Promega) was added in a concentration of 12.5 ng/µl in 20 mM ABC, pH 8.0. Gel pieces were soaked for 10 min at room temperature (RT), then covered with 20 mM ABC, and proteins were digested at 37°C overnight. Peptide extraction was performed in three consecutive steps with different extraction buffers for 30 min: first 3% (v/v) trifluoroacetic acid (TFA) in 30% (v/v) ACN was added, followed by 0.5% (v/v) formic acid (FA) in 80% (v/v) ACN, and completed by 100% ACN. Supernatants were pooled and ACN was evaporated by vacuum centrifugation. Peptides were further purified with C18 StageTips[89] and analyzed on an EASY-nLC 1200 UHPLC coupled to a Q Exactive HF mass spectrometer (both Thermo Scientific) as described previously[90] with slight modification: peptides were eluted from the analytical column using a 46 min segmented gradient of 10-33-50% of HPLC solvent B (80% acetonitrile in 0.1% formic acid) at a flow rate of 200 nl/min. In the mass spectrometer, MS and MS/MS spectra were acquired at resolution 60k. Full MS target value and maximum IT were set to $3x10^6$ and 25 ms, respectively. In each scan cycle, the 7 most intense precursor ions were picked up, whereby MS/MS target value was set to 10^5 charges with a maximum IT of 110 ms.
MS data were processed using default parameters of the MaxQuant software (v2.4.12.0)[91]. Extracted peak lists were submitted to database search using the Andromeda search engine[92] to query a target-decoy[93] database of *A. balhimycina* (9,427 entries, downloaded on 30th of January 2024) and 286 commonly observed contaminants.
In database search, full tryptic specificity was required and up to two missed cleavages were allowed. Carbamidomethylation of cysteine was set as fixed modification, whereas protein N-terminal acetylation, and oxidation of methionine were set as variable modifications. For the main search, the peptide mass tolerance was set to 4.5 ppm, and 20 ppm at the fragment ion level. Peptide, protein, and modification site identifications were filtered at a false discovery rate (FDR) of 0.01. The iBAQ (Intensity Based Absolute Quantification) and LFQ (Label-Free Quantification) algorithms were enabled, as was the “match between runs” option[94,95].
Downstream statistical analysis was performed using LFQ values and according to Aguilan et al.[96] with an additional filtering step to remove hits that were only detected in one replicate and condition.

Quantification and statistical analysis

The extra- and intracellular amounts of balhimycin were quantified using the HPLC-MS data. For this purpose, the peak intensities of each data point in the MS chromatogram corresponding to the masses of balhimycin were summed to a total intensity value. The concentration of balhimycin in the analyzed samples was calculated by using a standard curve, in which the total intensities of pure balhimycin were used. The concentration values were used to determine the total amount of balhimycin in the supernatant or mycelium extracts, respectively. The export and accumulation levels of balhimycin were normalized to the dry cell weight (DCW). Normalized values were used for statistical analysis using R (v4.3.1)[63] including the additional packages ”dplyr” (v1.1.3)[64], “readxl” (v1.4.3)[65], “ggplot2” (v3.5.0)[66], and “reshape” (v.1.4.4)[67]. We used the “Wilcoxon signed-rank test” and the “Benjamini-Hochberg Procedure” to calculate significance levels (p-values) for all data analyzed. P-values < 0.05 were considered as statistically significant. The results of this analysis are found in Figure 2 and Figure 4 and the respective figure legend. Every datapoint represent one independent experiment and quantification (n). At least three independent experiments were performed per strain. Figure 2: *A. balhimycina* wildtype (n=11), Δ*tba* (n=12), Δ*tba* + *tba* (n=12), Δ*tba* + *tba*^{3xFLAG} (n=11), Δ*tba* + *tri*^{3xFLAG} (n=11), Δ*tba* + *tva*^{3xFLAG} (n=8), Δ*tba* + *tbaE545Q*^{3xFLAG} (n=6), Δ*tba* + *sav1866*^{3xFLAG} (n=8). Figure 4: *A. balhimycina* wildtype (n=3), Δ*tba* (n=3), Δ*tba* + *tba*^{3xFLAG} (n=3), Δ*tba* + *tbaQ309G*^{3xFLAG} (n=6), Δ*tba* + *tbaR310A*^{3xFLAG} (n=6), Δ*tba* + *tbaT316A*^{3xFLAG} (n=6).

Supplemental Excel table titles and legends

Table S1: List of used transporters with supporting information

Table S2: Statistical analysis of quantitative measurements

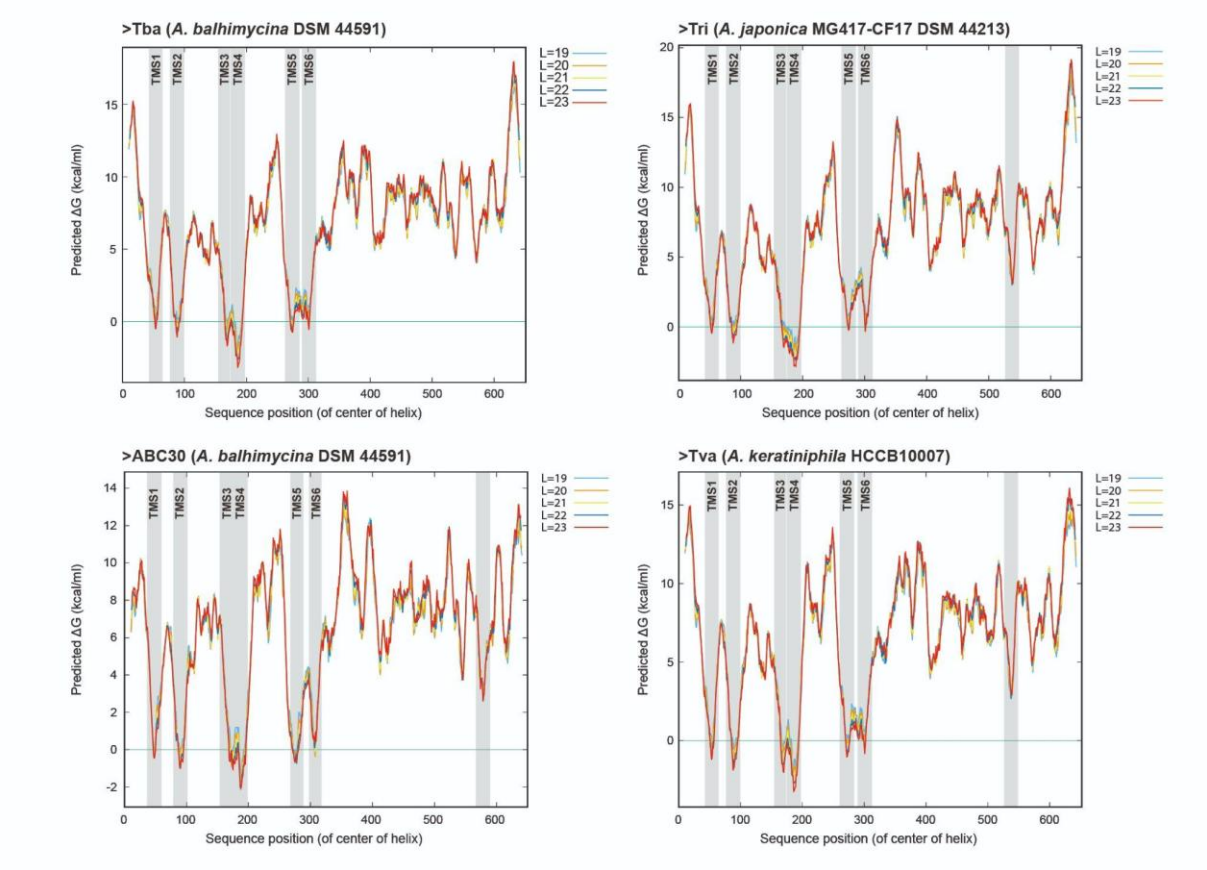


Figure S1. Prediction of transmembrane topology of different GPA ABC transporters. Prediction of transmembrane topology of ABC transporters encoded in the GPA BGC of *A. balhimycina* (Balhimycin producer), *A. japonica* (Ristomycin producer), *A. keratiniphila* (Vancomycin producer), and the non GPA transporter Abc30 of *A. balhimycina*. Topology was predicted based on the membrane integration propensity of segments (19-23 AA)[1]. Putative transmembrane segments (TMS) are marked in gray.

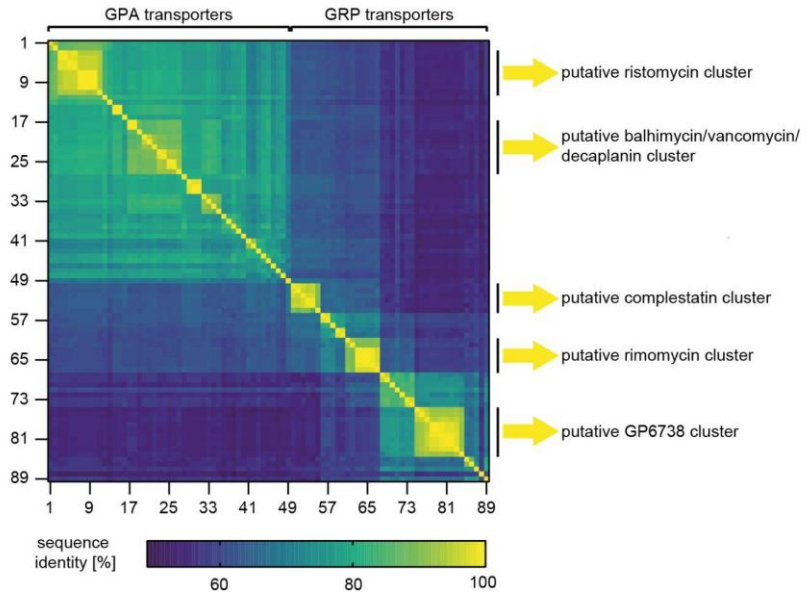


Figure S2. Similarity matrix of amino acid sequence identities of transporters. Heatmap representation of amino acid sequence identities of all 89 GPA and GRP associated ABC transporters. Identities were determined by a BlastP pairwise comparison analysis and are displayed in percent (%) of identity. Each number represents one transporter. The corresponding strains from which the sequences originate and numberings used in the heatmap are listed in supplementary table 1.

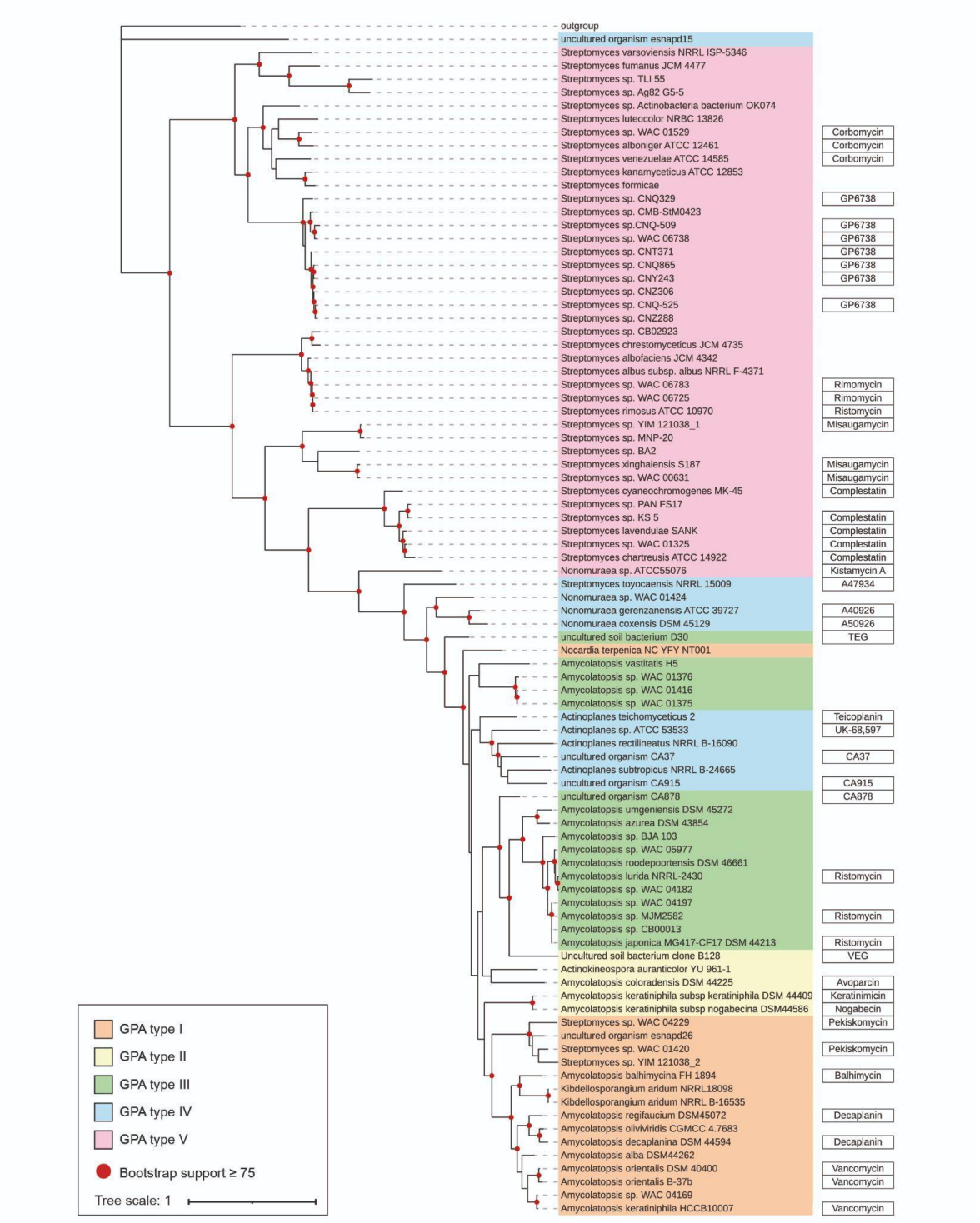


Figure S3. Evolutionary relationship within the group of GPA and GRP associated ABC exporters. Phylogenetic tree of ABC transporters encoded in BGC of GPAs and GRPs. The tree was calculated using the Maximum Likelihood (LG+F+G) algorithm with 1000 bootstrap repetitions. Nodes with bootstrap support ≥ 75 are marked by red dots. The tree was rooted using a putative ABC transporter encoded in an unknown BGC in *A. balhimycina* (Abc30). The strain from which the cluster and thus the associated ABC transporter originate is indicated in the tree. (Predicted) GPA types are labeled accordingly: type I (orange), type II (yellow), type III (green), type IV (blue), type V (pink). Known and predicted GPAs are mentioned behind the branches.

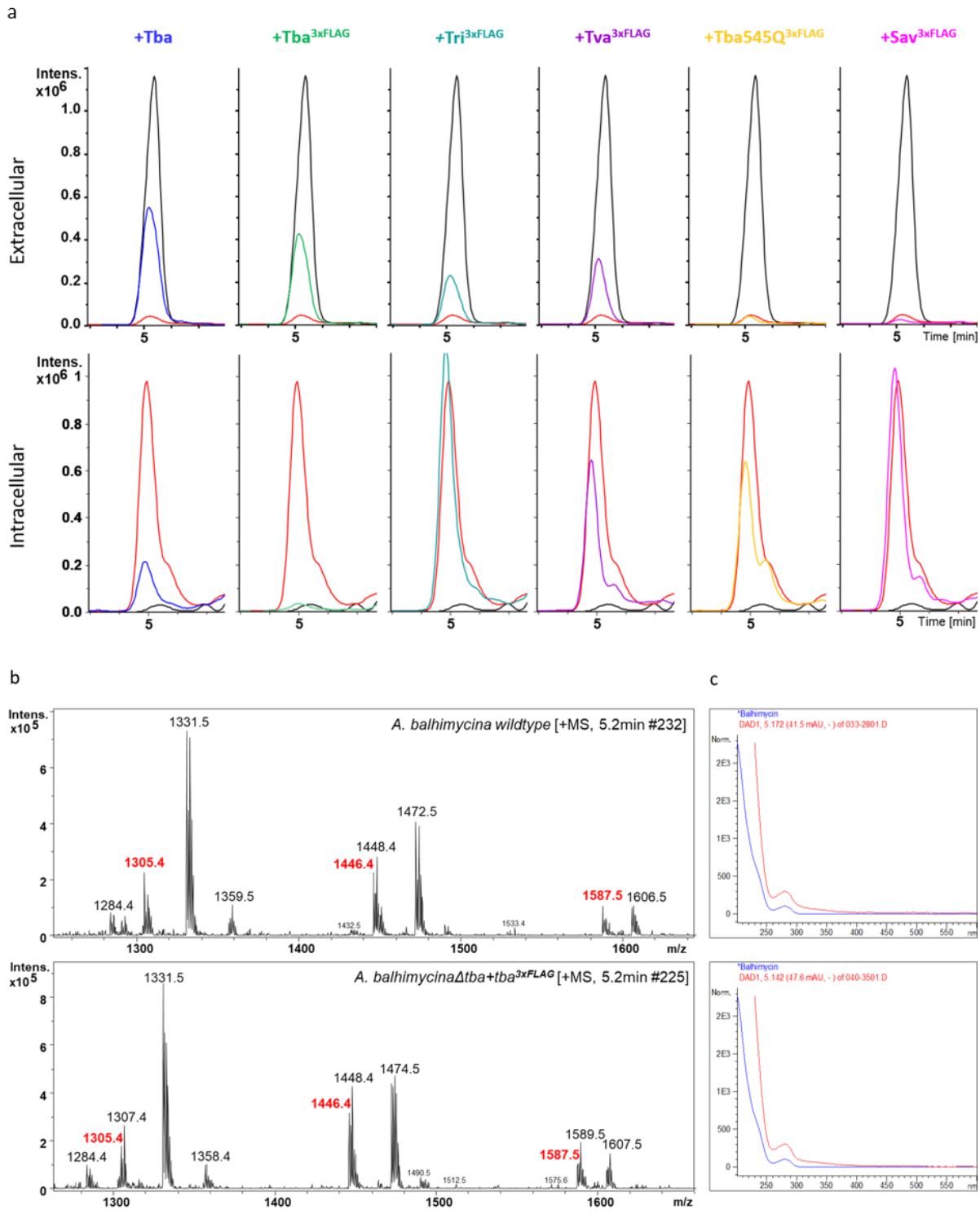


Figure S4. Export assay of balhimycin by diverse ABC transporters. a: HPLC-MS chromatograms of the culture supernatants (Extracellular) as well as mycelium extracts (Intracellular). The peaks (retention time 5-5.5 min) represent the extracted ion chromatogram (EIC) of the protonated balhimycin mass *m/z* 1446.41 $[M+H]^+$ (positive mode, smoothed 10.63-10.76 GA). In each chromatogram view, the EIC of *A. balhimycina* wildtype (black) and Δ*tba* (red) are overlaid with the respective EIC of the Δ*tba* mutant complemented with different transporters. **b:** Depicted mass spectra of *A. balhimycina* wildtype (top) and *A. balhimycina Δtba+tba*^{3xFLAG}(bottom*)* showing detected masses of the balhimycin molecules as $[M+H]^+$(red). **c:** UV-spectra of

balhimycin derived from HPLC-DAD analysis of supernatants of *A. balhimycina* wildtype (top) and *A. balhimycinaΔtba+tba*^{3xFLAG} (bottom) compared to in-house UV database of balhimycin standard.

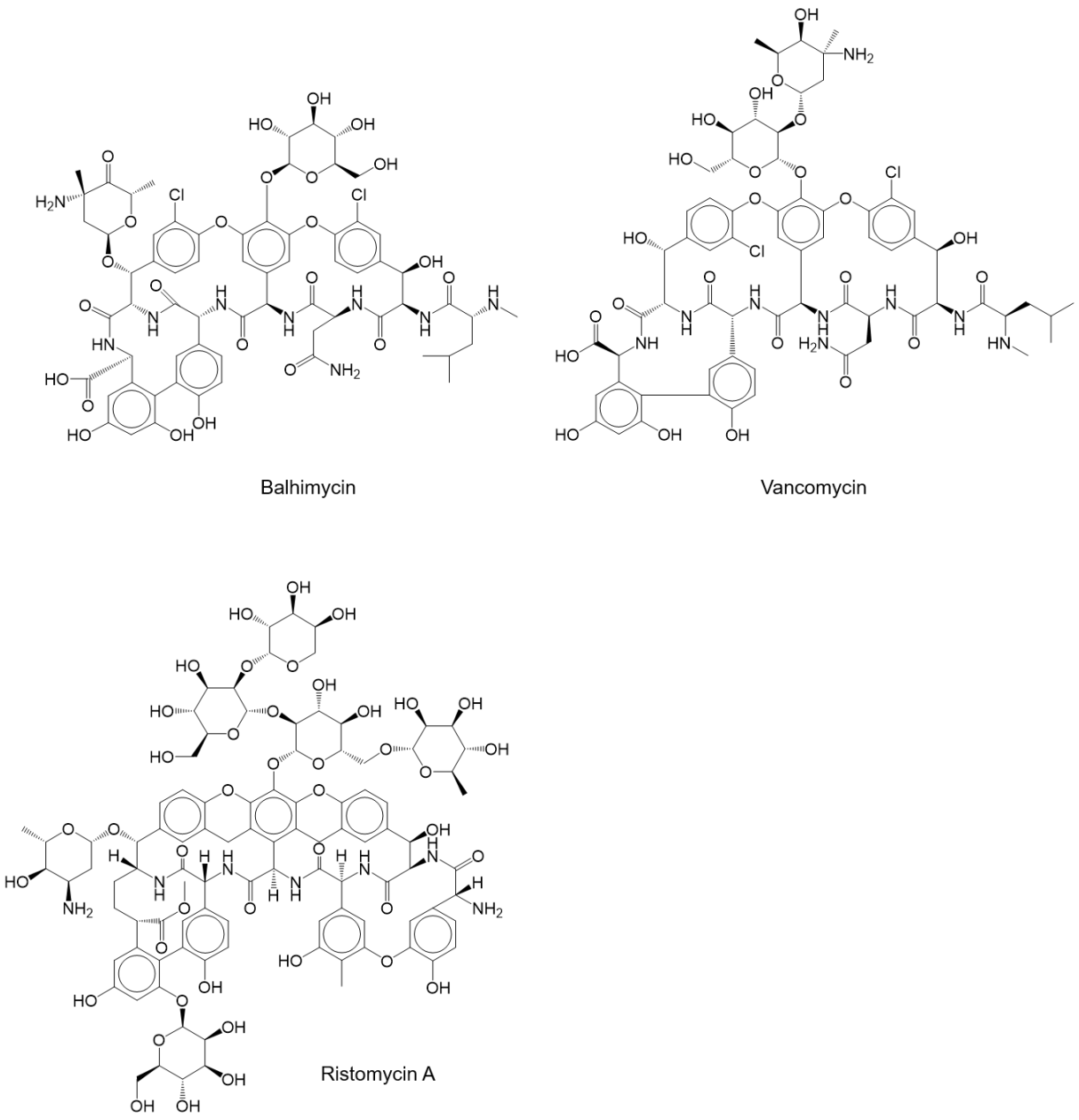


Figure S5. Chemical structure of the glycopeptides molecules.
Balhimycin (ChemSpider ID10249894), vancomycin (ChemSpider ID14253) and ristomycin A (ChemSpider ID16736169)

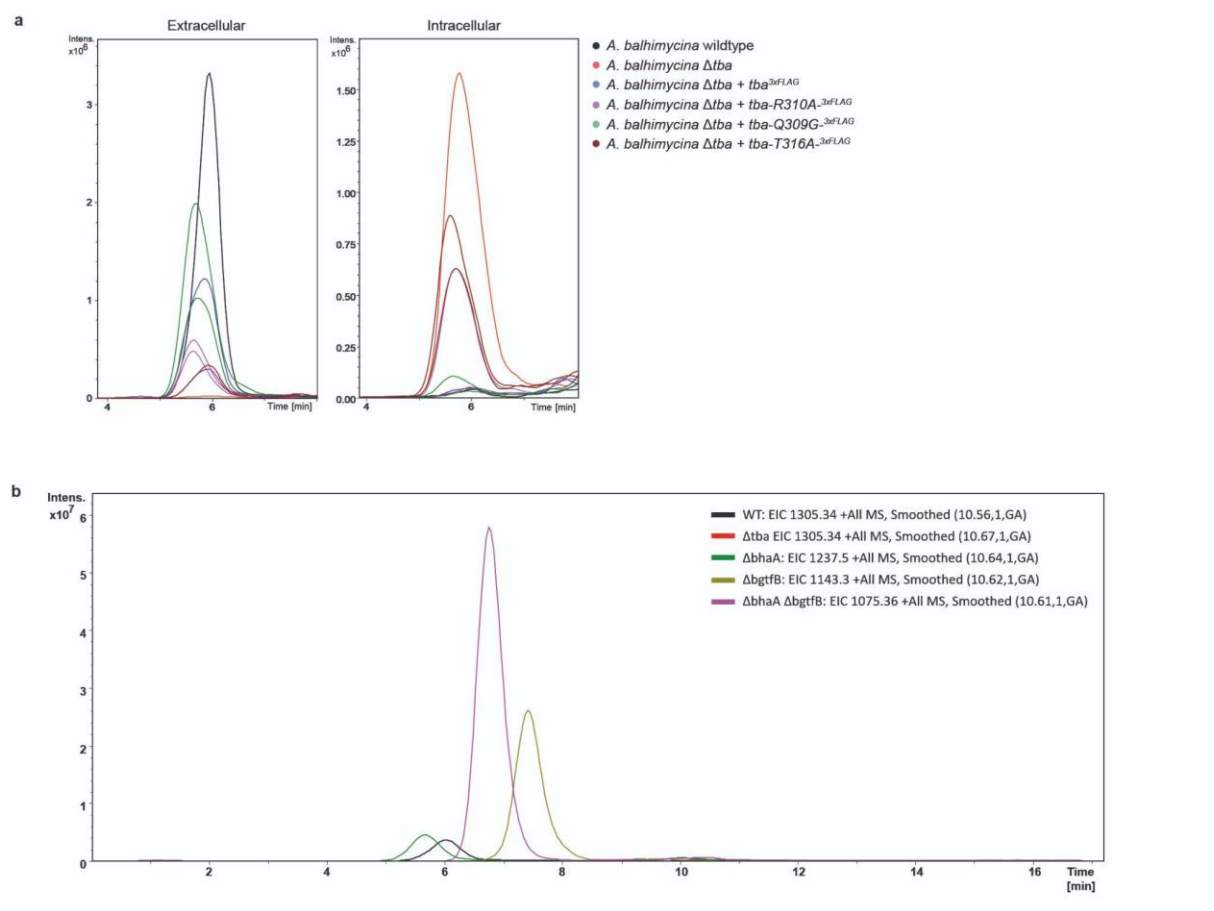


Figure S6. HPLC-MS chromatograms of mutants carrying point mutations in the TMD and mutants lacking the genes for chlorination and glycosylation. a: HPLC-MS chromatograms of the culture supernatants (Extracellular) as well as mycelium extracts (intracellular) of the mutants expressing *tba-Q309G-*^{3xFLAG}, *tba-R310A-*^{3xFLAG} and *tba-T316A-*^{3xFLAG}. The peaks (retention time 5-5,5 min) represent the extracted ion chromatogram (EIC) of the protonated balhimycin mass 1446.41 m/z [M+H] (positive mode, smoothed 10.63-10.76 GA). The EICs of each mutant are overlaid and compared to the wildtype (black) and Δ*tba* (orange) as well as *tba*^{3xFLAG}. **b:** HPLC-MS analyzed culture supernatants of mutants, missing the *bhaA* and *bgtfB* genes. Each chromatogram represents EICs of the corresponding derivative produced by the mutant. As controls, the EICs of wildtype (black) and Δ*tba* (red) are overlaid.

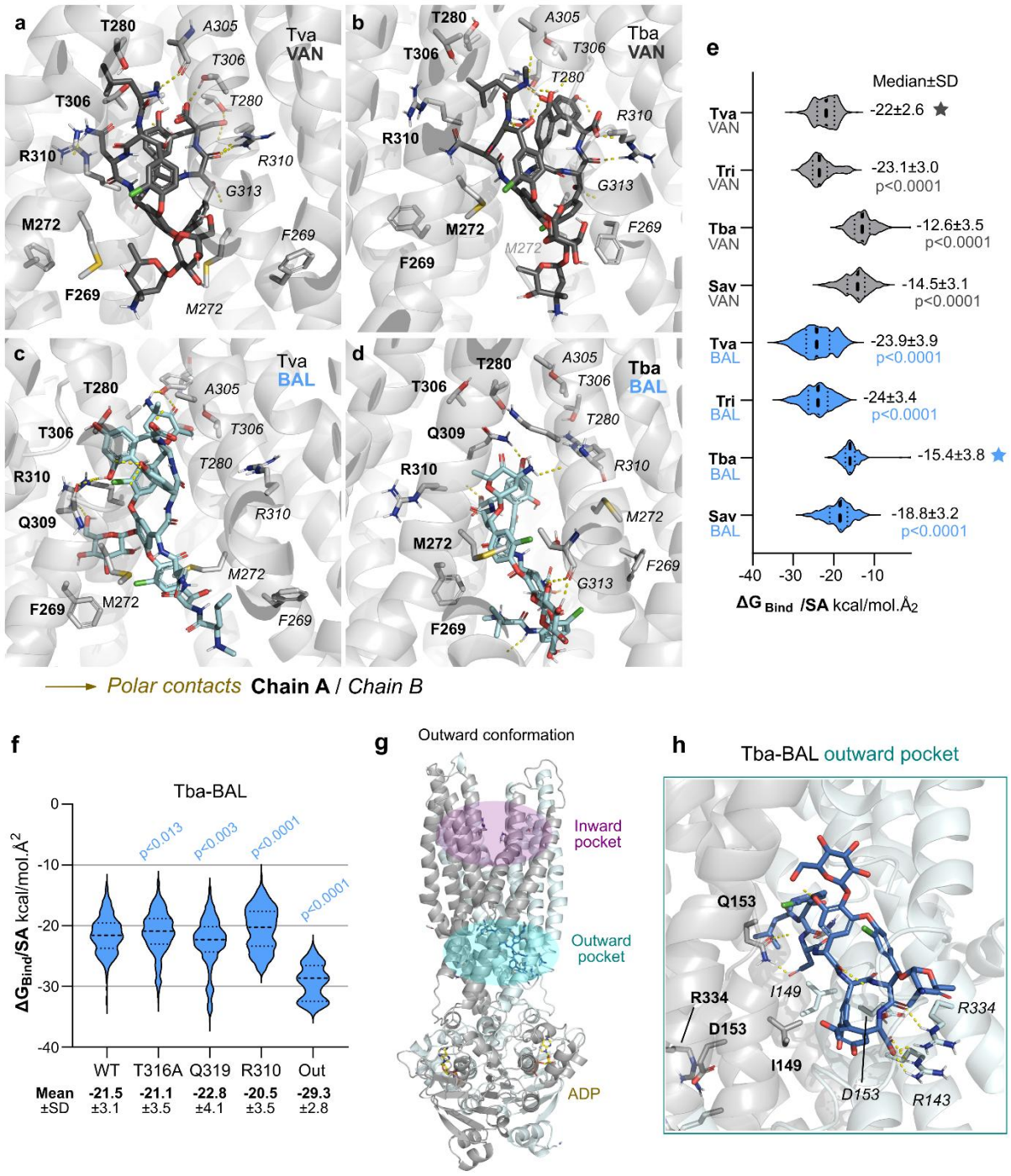


Figure S7. GPA potential binding mode and their respective predicted binding energies. Representative conformation from the MD simulation for the vancomycin (**VAN**) binding transporters Tva (**a**) and Tba (**b**), as well as their balhimycin (**BAL**) counterparts in (**c, d**). Conformations highlight that GPAs proposed binding mode points the R1-sugar moieties towards the cytoplasmic cavity beyond the substrate binding pocket (in the TMD). Predicted binding energy (**e, f**). Violin plot depicting the variation of free binding energy calculated along the trajectory normalized by the ligand's surface for: (**E**) WT-GPA transporters with VAN (gray) or BAL (blue) and (**f**) for different point mutations of Tba bound to BAL. Median energy values are depicted above each ligand. Mann-Whitney tests were performed to compare all groups against the control group's cumulative distribution, highlighted with a star from the respective group; exact p-values are depicted when available. Comparisons between

the same transporter with different binding GPA yielded no significant results and therefore are not displayed. **g**: 3D representation of a type IV GPA transporter Tba based on the SAV1866 (PDB 2HYD) followed by short MD simulations (5x200 ns). **h:** Proposed binding mode for relevant GPAs, generated from representative conformations from the MD simulation for the transporter Tba bound to balhimycin.

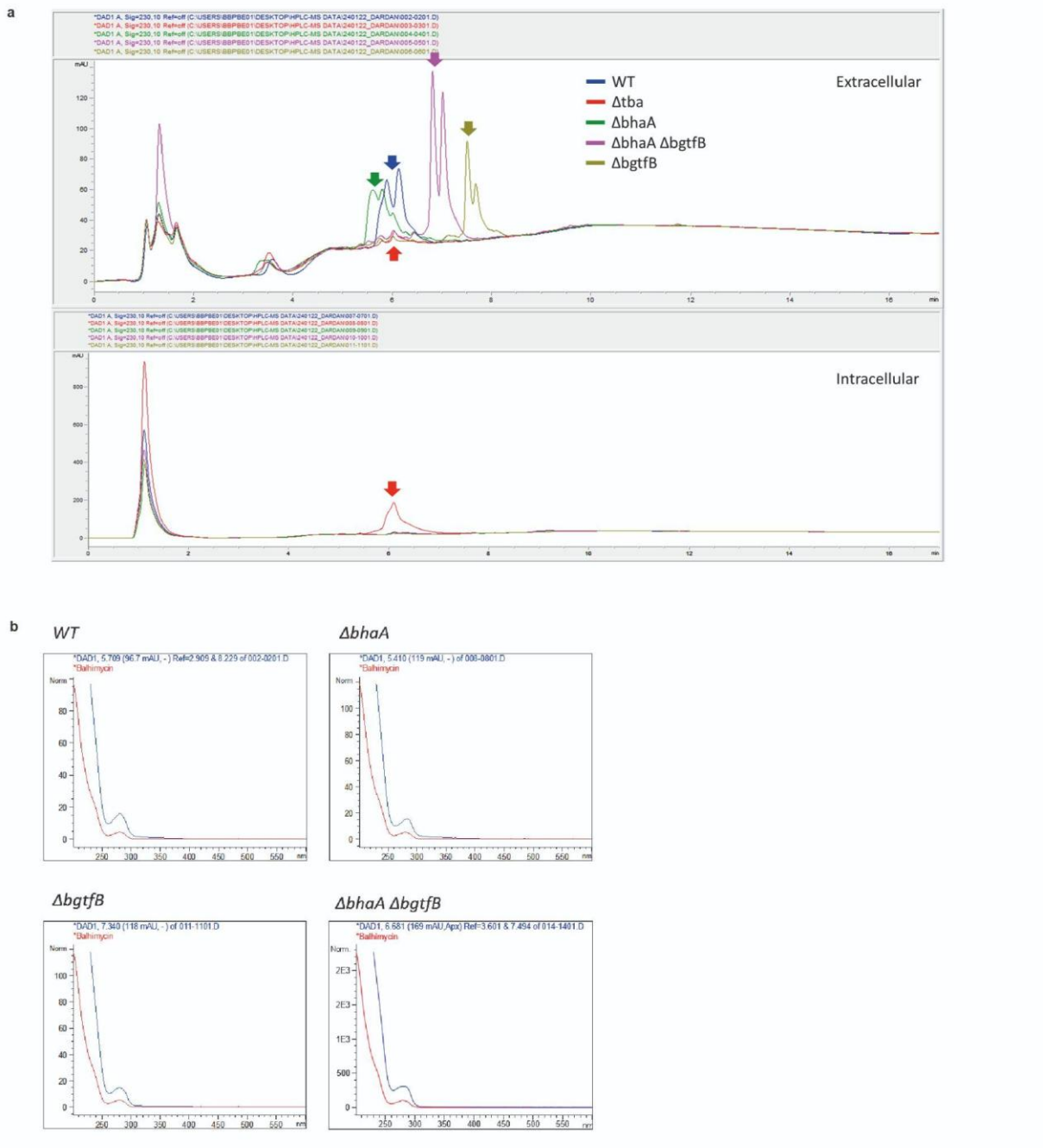


Figure S8. UV spectra of the analyzed deletion mutants Δ*bhaA*, Δ*bgtfB* and Δ*bhaA*Δ*bgtfB*. a: The UV spectrum at 230 nm after HPLC-DAD of each analyzed supernatant are overlaid and marked with arrows indicating the peak corresponding to the balhimycin derivatives produced by the respective mutant. **b:** UV-spectra of balhimycin and its derivatives derived from HPLC-DAD analysis of supernatants of

each *A. balhimycina* mutant (blue) compared to in-house UV database of balhimycin standard (red).

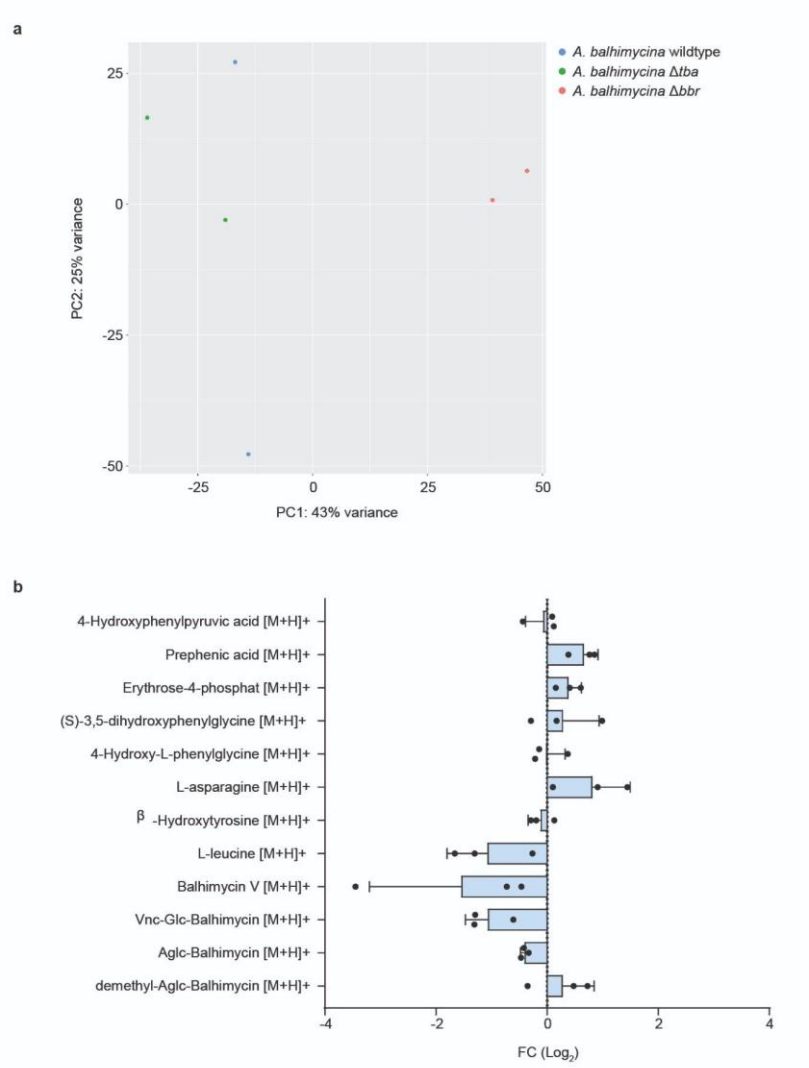


Figure S9. Transcriptome and metabolome analysis of *A. balhimycina Δtba*. a: Comparison of two individual replicates of *A. balhimycina* wildtype (blue), Δ*tba* (green), and Δ*bbr* (salmon) using principal component analysis (PCA). b: Levels of balhimycin intermediates after 48 h measured by flow-injection mass spectrometry. Values display *A. balhimycina* Δ*tba* in comparison to *A. balhimycina* wildtype.

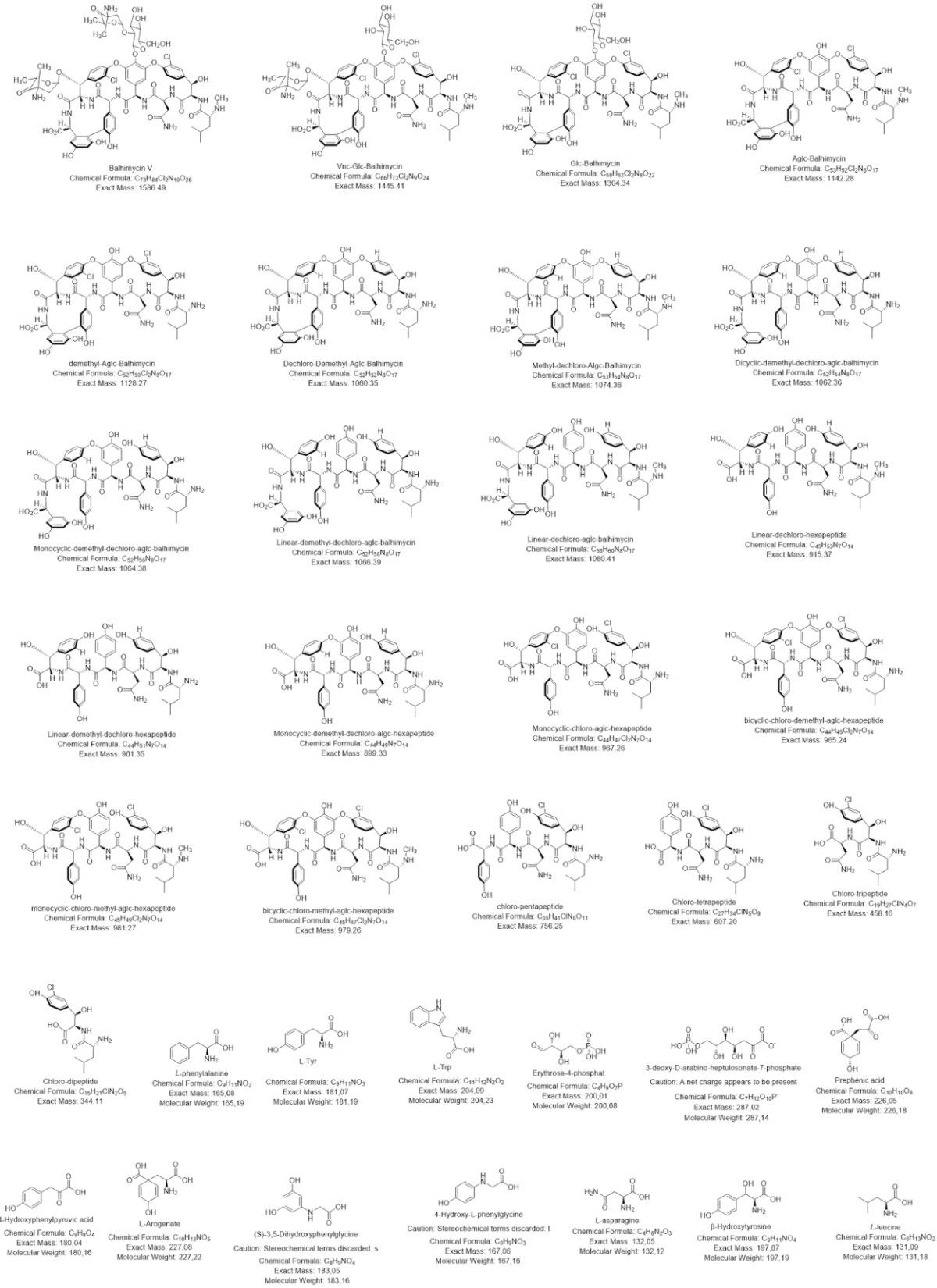


Figure S10. List of balhimycin intermediates. Structure, chemical formula, exact mass and molecular weight of balhimycin, putative intermediates, and precursors.

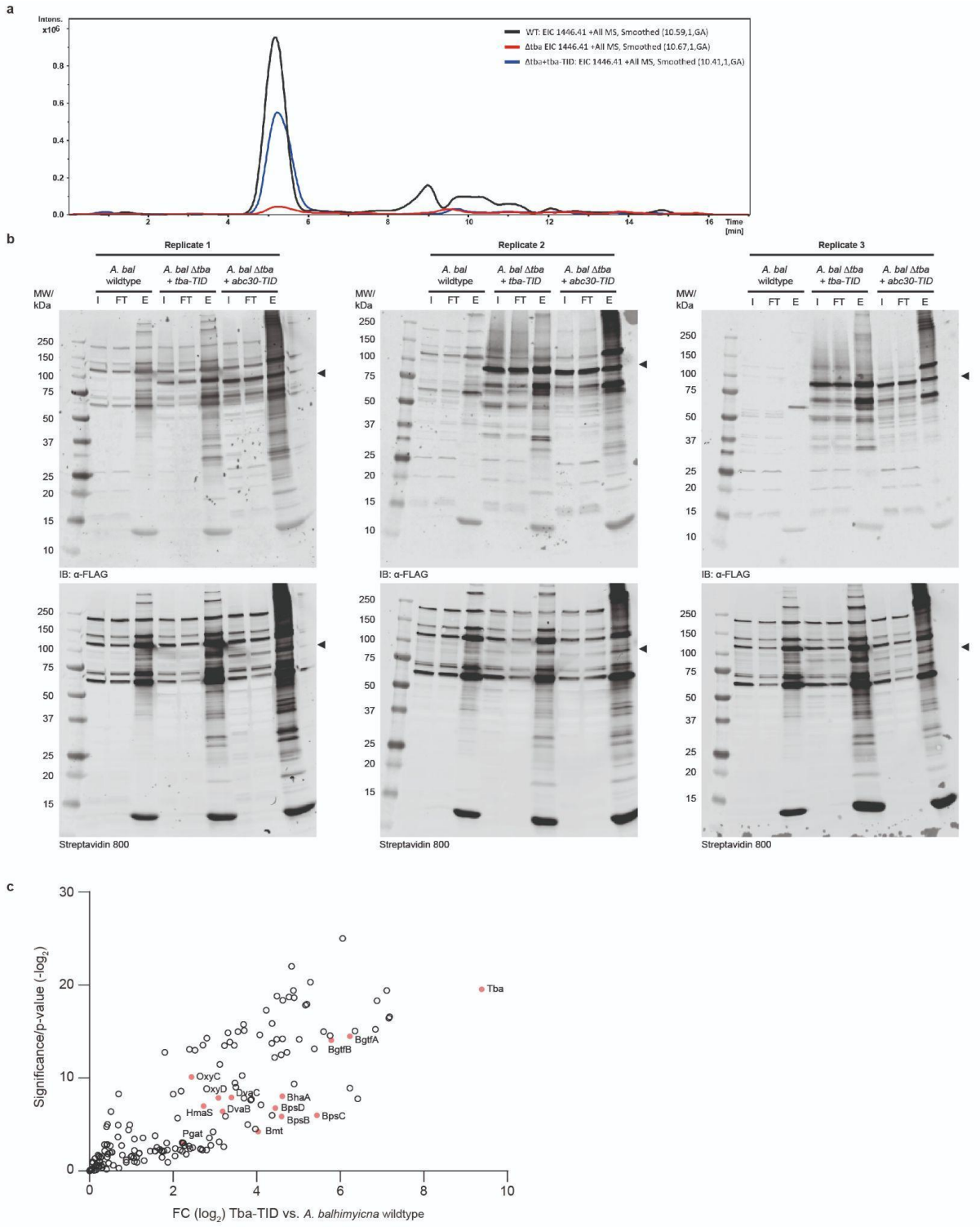


Figure S11. Proximity dependent biotinylation. a: HPLC-MS chromatograms of the culture supernatants of *A. balhimycina* Δ*tba* complemented with *tba-TID*. The peaks (retention time 5-5.5 min) represent the extracted ion chromatogram (EIC) of the protonated balhimycin mass *m/z* 1446.41 $[M+H]^+$ (positive mode, smoothed 10.41-10.67 GA). **b:** Immunoblotting (top) and fluorescence labeling with Streptavidin 800 (bottom) after SDS-PAGE of input (I) biotinylated proteins after enrichment (E) and flowthrough after enrichment (FT) from the recombinant strains expressing TID fusion proteins. **c:** Representation of all log2 fold change positive proteins detected

by LC-MS/MS after biotinylation. Hits of Tba-TID compared to the wildtype strain *A. balhimycin*a without TID.

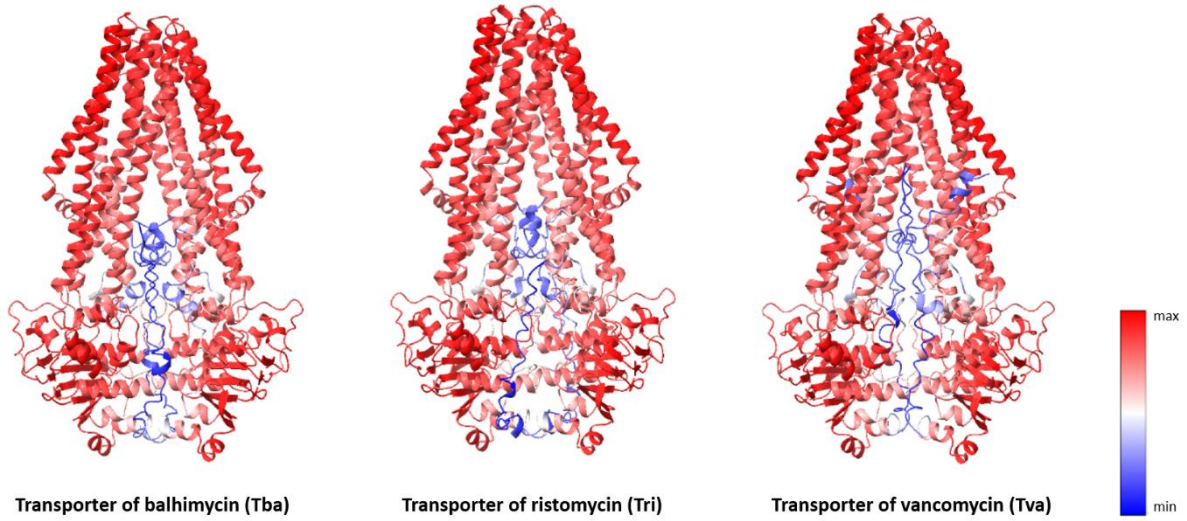


Figure S12. AlphaFold2 models of Tba, Tri and Tva and coloration based on B-factor coloured by model per-residue confidential score. The coloration is displayed in red as maximum and in blue as lowest value of the per-residue score. TMD regions were modelled with high confidence while the N-terminal could not, therefore it was removed from the calculations, to generate a truncated protein model.

Table S3. Mass list of balhimycin intermediates and precursors

Nr	Name	Molecular formula	Exact Mass
1	Chloro-dipeptide	C15H21ClN2O5	344.113901
2	Chloro-tripeptide	C19H27ClN4O7	458.156829
3	Chloro-tetrapeptide	C27H34ClN5O9	607.204508
4	chloro-pentapeptide	C35H41ClN6O11	756.252187
5	bicyclic-chloro-methyl-aglc-hexapeptide	C45H47Cl2N7O14	979.255809
6	monocyclic-chloro-methyl-aglc-hexapeptide	C45H49Cl2N7O14	981.271459
7	bicyclic-chloro-demethyl-aglc-hexapeptide	C44H45Cl2N7O14	965.240159
8	Monocyclic-chloro-aglc-hexapeptide	C44H47Cl2N7O14	967.255809
9	Monocyclic-demethyl-dechloro-algc-hexapeptide	C44H49N7O14	899.333753
10	Linear-demethyl-dechloro-hexapeptide	C44H51N7O14	901.349402
11	Linear-dechloro-hexapeptide	C45H53N7O14	915.365053
12	Linear-dechloro-aglc-balhimycin	C53H60N8O17	1080.407647
13	Linear-demethyl-dechloro-aglc-balhimycin	C52H58N8O17	1066.391997
14	Monocyclic-demethyl-dechloro-aglc-balhimycin	C52H56N8O17	1064.376347
15	Dicyclic-demethyl-dechloro-aglc-balhimycin	C52H54N8O17	1062.360697

16	Methyl-dechloro-Algc-Balhimycin	C53H54N8O17	1074.360697
17	Dechloro-Demethyl-Aglc-Balhimycin	C52H52N8O17	1060.345047
18	demethyl-Aglc-Balhimycin	C52H50Cl2N8O17	1128.267103
19	Aglc-Balhimycin	C53H52Cl2N8O17	1142.282753
20	Glc-Balhimycin	C59H62Cl2N8O22	1304.335578
21	Vnc-Glc-Balhimycin	C66H73Cl2N9O24	1445.414557
22	Balhimycin V	C73H84Cl2N10O26	1586.493536
Precursors			
NR	Name	Molecular formula	Exact mass
1	L-leucine	C6H13NO2	131.094629
2	β-Hydroxytyrosine	C9H11NO4	197.068809
3	L-asparagine	C4H8N2O3	132.053493
4	4-Hydroxy-L-phenylglycine	C8H9NO3	167.058244
5	(S)-3,5-dihydroxyphenylglycine	C8H9NO4	183.053159
6	Erythrose-4-phosphat	C4H9O7P	200.008593
7	3-deoxy-D-arabino-heptulosonate-7-phosphate	C7H12O10P	287.016812
8	Prephenic acid	C10H10O6	226.04774
9	4-Hydroxyphenylpyruvic acid	C9H8O4	180.04226

Table S4 | Primers used in this study

ID	Primer	Sequence (5'□3')
P1	seq_ermE_prom_f	TTG TGG GCA CAA TCG TGC CGG
P2	seq_lacZa_r	GCA CTG GCC GTC GTT TTA CAA CGT
P3	tba_UP_f	TAA AGG GAG AGA CGA ATT CGA GCT CGG TAC GCC AAG CCG CAC CCG C
P4	tba_UP_r	GCT GCG GAA TTC ATC CAT CCG GTC ACC TCT CT
P5	tba_Down_f	AGA GGT GAC CGG ATG GAT GAA TTC CGC AGC GCG GAC CA
P6	tba_Down_r	GCA GGT CGA CTC TAG AGG ATC CCC GGG TAC GCT CGG TGC CCG CCC
P7	tba_pRM4_f	CTG CAG GAA TTC GAT ATC AAG CTT TCA TCC TCC GTA GCC CAT GTG
P8	tba_pRM4_r	CCA GGG GAG GAC CCA TAT GAT GGA CAT GGT GTT GCG TTT
P9	KO_tba_f	ACG CCC CGC ACG TGG TGG CGT
P10	KO_tba_r	TCG GAC GCG CCC TCC TCG GCG GCG
P11	bgtfB_down_f	GCG ATA TCC GCG AAA CTG CTG CTC G
P12	bgtfB_down_r	GGA TCG TCT AGA TGG CCA CCT TCG C

P13	bgtfB_UP_f	GTG GAA TTC GGC AGC TCG TCC GGA C
P14	bgtfB_UP_r	TTG ATA TCC GGT TGT TCC GCT CCC
P15	KO_bgtfB_f	AAC CTA CGG AAC AGA GGG TGC
P16	KO_bgtfB_r	GTC TCC TCT TTG CTT TGT CTT CGA AGA
P11	gib_uni_pRM4_f	TCT AGA GTC GAC CTG CAG CCC GAG
P12	gib_uni_tba_r	TCC TCC GTA GCC CAT GTG TTG GAT C
P13	gib_pRM4_3xFLAG_f	GAT CCA ACA CAT GGG CTA CGG AGG ATC TAG AGA CTA CAA AGA CCA TGA C
P14	gib_pRM4_3xFLAG_r	CTC GGG CTG CAG GTC GAC TCT AGA TCA TTT GTC ATC GTC ATC CTT GTA ATC
P15	gib_pRM4_Tri2_f	ATA AGC TAG CCA GGG GAG GAC CCA TAT GGA AGT AAT GTT GCG CTT CGG
P16	gib_pRM4_Tri_r	GCT CGG GCT GCA GGT CGA CTC TAG ATC ATC CTC CGT AGA CCA CGG T
P17	gib_tri_3xFLAG_f	GGA CAC CGT GGT CTA CGG AGG ATC TAG AGA CTA CAA AGA CCA TGA C
P18	gib_pRM4_Sav1866_f	ATA AGC TAG CCA GGG GAG GAC CCA TAT GAT TAA ACG ATA TTT GCA ATT TGT TAA GCC
P19	gib_pRM4_Sav1866_r	CTC GGG CTG CAG GTC GAC TCT AGA TTA TAA GTT TTG AAT GCT ATA TAA ATG CTC GTA AG
P20	gib_Sav1866_FLAG_f	AGG TGC TTA CGA GCA TTT ATA TAG CAT TCA AAA CTT ATC TAG AGA CTA CAA AGA CCA TGA C
P21	gib_FLAG_Sav1866_r	CAC CGT CAT GGT CTT TGT AGT CTC TAG ATA AGT TTT GAA TGC TAT ATA AAT GCT CGT AAG
P22	gib_pRM4_Tva_f	ATA AGC TAG CCA GGG GAG GAC CCA TAT GGA AGT GGT GTT GCG CTT C
P23	gib_pRM4_Tva_r	CTC GGG CTG CAG GTC GAC TCT AGA TCA TCC TCC GAA GCC AAT GG
P24	gib_FLAG_Tva_r	CAC CGT CAT GGT CTT TGT AGT CTC TAG ATC CTC CGA AGC CAA TGG GTT G
P25	gib_Tva_FLAG_f	GAT CCA ACC CAT TGG CTT CGG AGG ATC TAG AGA CTA CAA AGA CCA TGA C
P26	gib2_pRM4_Tba_f	ATA AGC TAG CCA GGG GAG GAC CCA TAT GGA CAT GGT GTT GCG TTT CGA G
P27	gib_TbaE545Q_r	GGG CGG TGG CTT GGT CGA GGA CGA CGA TC
P28	gib_Tba545Q_f	GAT CGT CGT CCT CGA CCA AGC CAC CGC CC
P29	gib_pRM4_3XFLAG_r	CTC GGG CTG CAG GTC GAC TCT AGA TCA TTT GTC ATC GTC ATC CTT GTA ATC
P30	gib_TbaR310A_r	CGG CCC GAA CAG CGC CTG GAG CAG GGT G
P31	gib_TbaR310A_f	CAC CCT GCT CCA GGC GCT GTT CGG GCC G
P32	gib_TbaQ309G_r	CGA ACA GCC GCC CGA GCA GGG TGG CGA TG
P33	gib_TbaQ309G_f	CAT CGC CAC CCT GCT CGG GCG GCT GTT CG
P34	gib_TbaT316A_r	CCC GGA CAG CTG GGC GAT CGG CCC GAA CA
P35	gib_TbaT316A_f	TGT TCG GGC CGA TCG CCC AGC TGT CCG GG

P36	gib_GSTurboID_Tba_r	GAG CGA TCA GCT TCA GAG GCA CAG TAT TGT CTT TGC TCC CTC CTC CGT AGC CCA TGT GTT GG
P37	gib_Tba_GSTurboID_f	CGG TGA TCC AAC ACA TGG GCT ACG GAG GAG GGA GCA AAG ACA ATA CTG TGC CTC TGA AGC
P38	gib_pRM4_His6_r	CAA GCT CGG GCT GCA GGT CGA CTC TAG ATC AGT GGT GGT GGT GGT GGT GCT C
P39	gib_pRM4_ABC30_f	ATA AGC TAG CCA GGG GAG GAC CCA TAT GAC CAT TGG AGA CGA CCC GAG C
P40	gib_GSTurboID_ABC30_r	GAG CGA TCA GCT TCA GAG GCA CAG TAT TGT CTT TGC TCC CGC CCA AAG CAC GGT TCC CGC
P41	gib_pRM4_TurboID_f	CCA AGA GCG GGA ACC GTG CTT TGG GCG GGA GCA AAG ACA ATA CTG TGC CTC TGA AGC

Table S5 | Strains used in this study

Strain	Relevant features	Reference
E. coli NEB5α	Cloning host for plasmid generation	New England Biolabs
E. coli NovaBlue	Cloning host for plasmid generation	Novagen®
E. coli JM110	*lacY dam dcm* [F' *lacI*^{q} Z Δ *M15*] (Methylation deficient strain)	[97]
E. coli ET12567	F^{-} *dam*-13::Tn9 *dcm*-6 *hsdM hsdR* (Methylation deficient)	[98]
A. balhimycina DSM 44591	Balhimycin producing wild type	[99,100]
A. balhimycina Δ*tba*	*tba* deletion mutant	This study
A. balhimycina Δ*tba* [+*tba*]	Apr^{R}, *tba* deletion mutant, ΦC31attB(pRM4-*tba*)	This study
A. balhimycina Δ*tba* [+*tba-3xFLAG*]	Apr^{R}, *tba* deletion mutant, ΦC31attB(pRM4-*tba-3xFLAG*)	This study
A. balhimycina Δ*tba* [+*tri-3xFLAG*]	Apr^{R}, *tba* deletion mutant, ΦC31attB(pRM4-*tri-3xFLAG*)	This study
A. balhimycina Δ*tba* [+*tva*-3x*FLAG*]	Apr^{R}, *tba* deletion mutant, ΦC31attB(pRM4-*amoh-3xFLAG*)	This study
A. balhimycina Δ*tba* [+*sav1866*-3x*FLAG*]	Apr^{R}, *tba* deletion mutant, ΦC31attB(pRM4-*sav1866-3xFLAG*)	This study
A. balhimycina Δ*tba* [+*tbaE545Q*-3x*FLAG*]	Apr^{R}, *tba* deletion mutant, ΦC31attB(pRM4-*tbaE545Q-3xFLAG*)	This study
A. balhimycina Δ*tba* [+*tbaR310A*-3x*FLAG*]	Apr^{R}, *tba* deletion mutant, ΦC31attB(pRM4-*tbaR310A-3xFLAG*)	This study
A. balhimycina Δ*tba* [+*tbaQ309G*-3x*FLAG*]	Apr^{R}, *tba* deletion mutant, ΦC31attB(pRM4-*tbaQ309G-3xFLAG*)	This study

A. balhimycina Δ*tba* [+*tbaT316A*-3*xFLAG*]	Apr^R, *tba* deletion mutant, ΦC31attB(pRM4-*tbaT316A*-*3xFLAG*)	This study
A. balhimycina Δ*bhaA*	*bhaA* deletion mutant	[28]
A. balhimycina Δ*bgtfB*	*bgtfB* deletion mutant	This study
A. balhimycina Δ*bgtfB*Δ*bhaA*	*bhaA and bgtfB* double-deletion mutant	This study
A. balhimycina Δ*tba* [+*tba*-*TID*]	Apr^R, *tba* deletion mutant, ΦC31attB(pRM4-tba-*TID*)	This study
A. balhimycina [+*TID*]	ΦC31attB(pRM4-*TID*)	This study
B. subtilis DSM10	Indicator strain for bioactivity assay	[101]

*Apr^R (Apramycin resistance)

Table S6 | Plasmids used in this study

Plasmid	Annotation	Resistance	Reference
pRM4	pSET152*ermEp** derived Φ31 integration vector with artificial ribosomal binding site	Apr^R	[10]
pSP1	pT7/T3-α19 derived gene disruption vector, with *ermE* gene in *Sap*I site	Ery^R/ Amp^R	[57]
pSP1Δ*tba*	pSP1 derived deletion vector for *tba* gene	Ery^R/ Amp^R	This study
pSP1Δ*bgtfB*	pSP1 derived deletion vector for *bgtfB* gene	Ery^R/ Amp^R	This study
pRM4-tba	pRM4 derived integrative *tba* expression vector	Apr^R	This study
pRM4-tba-3xFLAG	pRM4 derived integrative *tba-3xFLAG* expression vector	Apr^R	This study
pRM4-tri-3xFLAG	pRM4 derived integrative *tri-3xFLAG* expression vector	Apr^R	This study
pRM4-tva-3xFLAG	pRM4 derived integrative *tva-3xFLAG* expression vector	Apr^R	This study
pRM4-sav1866-3xFLAG	pRM4 derived integrative *sav1866-3xFLAG* (multidrug exporter of *S. aureus*) expression vector	Apr^R	This study
pRM4-tbaE545Q-3xFLAG	pRM4 derived integrative *tbaE545Q-3xFLAG* expression vector	Apr^R	This study
pRM4-tbaR310A-3xFLAG	pRM4 derived integrative *tbaR310A-3xFLAG* expression vector	Apr^R	This study
pRM4-tbaQ309G-3xFLAG	pRM4 derived integrative *tbaQ309G-3xFLAG* expression vector	Apr^R	This study
pRM4-tbaT316A-3xFLAG	pRM4 derived integrative *tbaT316A-3xFLAG* expression vector	Apr^R	This study
pET21a-TurboID-His6 (Addgene #107177)	TurboID expression vector; under control of P_{T7}	Amp^R	[31]
pRM4-tba-TID	pRM4 derived integrative *tba-TID* expression vector	Apr^R	This study
pRM4-TID	pRM4 derived integrative *TID* expression vector	Apr^R	This study
pRM4-abc30-TID	pRM4 derived integrative *abc30-TID* expression vector	Apr^R	This study

*Apr^R (Apramycin resistance), *Ery^R (Erythromycin resistance),*Amp^R (Ampicillin resistance)

References supplementary information

1. Hessa, T. *et al.* Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* **450**, 1026–1030 (2007).

References

1. Zeng, D. *et al.* Approved glycopeptide antibacterial drugs: Mechanism of action and resistance. *Cold Spring Harb. Perspect. Med.* **6**, a026989 (2016).
2. Culp, E. J. *et al.* Evolution-guided discovery of antibiotics that inhibit peptidoglycan remodelling. *Nature* **578**, 582-587 (2020).
3. Nicolaou, K. C., Boddy, C. N., Bräse, S. & Winssinger, N. Chemistry, biology, and medicine of the glycopeptide antibiotics. *Angew. Chem. Int. Ed Engl.* **38**, 2096-2152 (1999).
4. Gavriilidou, A. *et al.* Phylogenetic distance and structural diversity directing a reclassification of glycopeptide antibiotics. Preprint at https://doi.org/10.1101/2023.02.10.526856 (2023).
5. Recktenwald, J. *et al.* Nonribosomal biosynthesis of vancomycin-type antibiotics: a heptapeptide backbone and eight peptide synthetase modules. *Microbiol. Read. Engl.* **148**, 1105-1118 (2002).
6. Bischoff, D. *et al.* The Biosynthesis of vancomycin-type glycopeptide antibiotics –New insights into the cyclization steps. *Angew. Chem. Int. Ed.* **40**, 1693-1696 (2001).
7. Stegmann, E. *et al.* Genetic analysis of the balhimycin (vancomycin-type) oxygenase genes. *J. Biotechnol.* **124**, 640-653 (2006).
8. Haslinger, K., Peschke, M., Brieke, C., Maximowitsch, E. & Cryle, M. J. X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. *Nature* **521**, 105-109 (2015).
9. Yim, G., Thaker, M. N., Koteva, K. & Wright, G. Glycopeptide antibiotic biosynthesis. *J. Antibiot. (Tokyo)* **67**, 31-41 (2014).
10. Menges, R., Muth, G., Wohlleben, W. & Stegmann, E. The ABC transporter Tba of *Amycolatopsis balhimycina* is required for efficient export of the glycopeptide antibiotic balhimycin. *Appl. Microbiol. Biotechnol.* **77**, 125-134 (2007).
11. Schäberle, T. F. *et al.* Self-resistance and cell wall composition in the glycopeptide producer *Amycolatopsis balhimycina*. *Antimicrob. Agents Chemother.* **55**, 4283-4289 (2011).
12. Biemans-Oldehinkel, E., Doeven, M. K. & Poolman, B. ABC transporter architecture and regulatory roles of accessory domains. *FEBS Lett.* **580**, 1023-1035 (2006).
13. Biemans-Oldehinkel, E., Mahmood, N. A. B. N. & Poolman, B. A sensor for intracellular ionic strength. *Proc. Natl. Acad. Sci.* **103**, 10624-10629 (2006).
14. Ostash, B., Doud, E. & Walker, S. ABC transporter genes from *Streptomyces ghanaensis* moenomycin biosynthetic gene cluster: roles in antibiotic production and export. *Arch. Microbiol.* **194**, 915-922 (2012).
15. Gacek-Matthews, A. *et al.* Beyond toxin transport: Novel role of ABC Transporter for enzymatic machinery of cereulide NRPS assembly line. *mBio* **11**, e01577-20 (2020).
16. Neumüller, A. M., Konz, D. & Marahiel, M. A. The two-component regulatory system BacRS is associated with bacitracin 'self-resistance' of *Bacillus licheniformis* ATCC 10716. *Eur. J. Biochem.* **268**, 3180-3189 (2001).
17. Ohki, R. *et al.* The BceRS two-component regulatory system induces expression of the bacitracin transporter, BceAB, in *Bacillus subtilis*. *Mol. Microbiol.* **49**, 1135-1144 (2003).
18. Thomas, C. & Tampé, R. Structural and mechanistic principles of ABC transporters. *Annu. Rev. Biochem.* **89**, 605-636 (2020).
19. Hessa, T. *et al.* Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* **450**, 1026-1030 (2007).
20. Brigham, R. B. & Pittenger, R. C. *Streptomyces orientalis*, n. sp., the source of vancomycin. *Antibiot. Chemother. Northfield Ill* **6**, 642-647 (1956).
21. Xu, L. *et al.* Complete genome sequence and comparative genomic analyses of the vancomycin-producing *Amycolatopsis orientalis*. *BMC Genomics* **15**, 363 (2014).
22. Tan, G. Y. A. *et al. Amycolatopsis regifaucium* sp. nov., a novel actinomycete that produces kigamicins. *Int. J. Syst. Evol. Microbiol.* **57**, 2562-2567 (2007).
23. Hansen, M. H. *et al.* Resurrecting ancestral antibiotics: unveiling the origins of modern lipid II targeting glycopeptides. *Nat. Commun.* **14**, 7842 (2023).
24. Bibb, M. J., Janssen, G. R. & Ward, J. M. Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*. *Gene* **38**, 215-226 (1985).
25. Dawson, R. J. P. & Locher, K. P. Structure of a bacterial multidrug ABC transporter. *Nature* **443**, 180-185 (2006).
26. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
27. Thomas, C. *et al.* Structural and functional diversity calls for a new classification of ABC transporters. *FEBS Lett.* **594**, 3767-3775 (2020).

28. Puk, O. *et al.* Glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908: Function of a halogenase and a haloperoxidase/perhydrolase. *Chem. Biol.* **9**, 225-235 (2002).
29. Stegmann, E., Frasch, H.-J. & Wohlleben, W. Glycopeptide biosynthesis in the context of basic cellular functions. *Curr. Opin. Microbiol.* **13**, 595-602 (2010).
30. Shawky, R. M. *et al.* The border sequence of the balhimycin biosynthesis gene cluster from *Amycolatopsis balhimycina* contains *bbr*, encoding a StrR-like pathway-specific regulator. *Microb. Physiol.* **13**, 76-88 (2007).
31. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880-887 (2018).
32. Kim, D. I. *et al.* Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc. Natl. Acad. Sci.* **111**, E2453–E2461 (2014).
33. Waglechner, N., McArthur, A. G. & Wright, G. D. Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. *Nat. Microbiol.* **4**, 1862-1871 (2019).
34. Kuipers, A. *et al.* NisT, the transporter of the lantibiotic nisin, can transport fully modified, dehydrated, and unmodified prenisin and fusions of the leader peptide with non-lantibiotic peptides*. *J. Biol. Chem.* **279**, 22176-22182 (2004).
35. Romano, M. *et al.* Structural basis for natural product selection and export by bacterial ABC transporters. *ACS Chem. Biol.* **13**, 1598-1609 (2018).
36. Frasch, H.-J. *et al.* Alternative pathway to a glycopeptide-resistant cell wall in the balhimycin producer *Amycolatopsis balhimycina*. *ACS Infect. Dis.* **1**, 243-252 (2015).
37. Kilian, R., Frasch, H.-J., Kulik, A., Wohlleben, W. & Stegmann, E. The VanRS homologous two-component system VnlRSAb of the glycopeptide producer *Amycolatopsis balhimycina* activates transcription of the *vanHAX*$_{Sc}$ genes in *Streptomyces coelicolor*, but not in *A. balhimycina*. *Microb. Drug Resist.* **22**, 499-509 (2016).
38. Unsleber, S., Wohlleben, W. & Stegmann, E. Diversity of peptidoglycan structure – Modifications and their physiological role in resistance in antibiotic producers. *Int. J. Med. Microbiol.* **309**, 151332 (2019).
39. Pearson, L. A., Hisbergues, M., Börner, T., Dittmann, E. & Neilan, B. A. Inactivation of an ABC transporter gene, *mcyH*, results in loss of microcystin production in the *Cyanobacterium microcystis aeruginosa* PCC 7806. *Appl. Environ. Microbiol.* **70**, 6370-6378 (2004).
40. Thykaer, J. *et al.* Increased glycopeptide production after overexpression of shikimate pathway genes being part of the balhimycin biosynthetic gene cluster. *Metab. Eng.* **12**, 455-461 (2010).
41. Goldfinger, V. *et al.* Metabolic engineering of the shikimate pathway in *Amycolatopsis* strains for optimized glycopeptide antibiotic production. *Metab. Eng.* **78**, 84-92 (2023).
42. Thomas, C. & Tampé, R. Structural and mechanistic principles of ABC transporters. *Annu. Rev. Biochem.* **89**, 605-636 (2020).
43. Chen, J., van Heel, A. J. & Kuipers, O. P. Subcellular localization and assembly process of the nisin biosynthesis machinery in *Lactococcus lactis*. *mBio* **11**, e02825-20 (2020).
44. Chen, J., van Heel, A. J. & Kuipers, O. P. Visualization and analysis of the dynamic assembly of a heterologous lantibiotic biosynbhesis complex in *Bacillus subtilis*. *mBio* **12**, e0121921 (2021).
45. Browman, D. T., Hoegg, M. B. & Robbins, S. M. The SPFH domain-containing proteins: more than lipid raft markers. *Trends Cell Biol.* **17**, 394-402 (2007).
46. Kerfeld, C. A., Aussignargues, C., Zarzycki, J., Cai, F. & Sutter, M. Bacterial microcompartments. *Nat. Rev. Microbiol.* **16**, 277–290 (2018).
47. Yang, M. *et al.* Decoding the stoichiometric composition and organisation of bacterial metabolosomes. *Nat. Commun.* **11**, 1976 (2020).
48. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinforma. Oxf. Engl.* **30**, 2068-2069 (2014).
49. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).
50. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261-272 (2020).
51. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547-1549 (2018).
52. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973 (2009).
53. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
54. Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307-1320 (2008).

55. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293-W296 (2021).
56. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. *Practical Streptomyces Genetics*. (John Innes Foundation, Norwich Research Park, Colney, Norwich NR4 7UH, UK, 2000).
57. Pelzer, S., Reichert, W., Huppert, M., Heckmann, D. & Wohlleben, W. Cloning and analysis of a peptide synthetase gene of the balhimycin producer *Amycolatopsis mediterranei* DSM5908 and development of a gene disruption/replacement system. *J. Biotechnol.* **56**, 115-128 (1997).
58. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343-345 (2009).
59. Madoń, J. & Hütter, R. Transformation system for *Amycolatopsis* (*Nocardia*) *mediterranei*: direct transformation of mycelium with plasmid DNA. *J. Bacteriol.* **173**, 6325-6331 (1991).
60. Dietsche, T. *et al.* Structural and functional characterization of the bacterial type III secretion export apparatus. *PLOS Pathog.* **12**, e1006071 (2016).
61. Zilkenat, S. *et al.* Blue native PAGE analysis of bacterial secretion complexes in bacterial secretion systems: *Methods and Protocols* (eds. Journet, L. & Cascales, E.) 331-362 (Springer US, New York, NY, 2024). doi:10.1007/978-1-0716-3445-5_22.
62. Davis, B. D. & Mingioli, E. S. Mutants of *Escherichia coli* requiring methionine or vitamin B12. *J. Bacteriol.* **60**, 17-28 (1950).
63. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2023).
64. Wickham, H. *et al.* dplyr: A Grammar of Data Manipulation. (2023).
65. Wickham, H. *et al.* readxl: Read Excel Files. (2023).
66. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verl. N. Y.* (2016).
67. Wickham, H. Reshaping Data with the reshape Package. *J. Stat. Softw.* **21**, 1-20 (2007).
68. St-Pierre, J.-F., Bunker, A., Róg, T., Karttunen, M. & Mousseau, N. Molecular Dynamics Simulations of the bacterial ABC transporter SAV1866 in the closed form. *J. Phys. Chem. B* **116**, 2934-2942 (2012).
69. Shelley, J. C. *et al.* Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *J. Comput. Aided Mol. Des.* **21**, 681-691 (2007).
69 Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47**, 1739-1749 (2004).
71. Friesner, R. A. *et al.* Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **49**, 6177-6196 (2006).
72. Vesga, L. C. *et al.* Tetrahydroquinoline/4,5-dihydroisoxazole molecular hybrids as inhibitors of breast cancer resistance protein (BCRP/ABCG2). *ChemMedChem* **16**, 2686-2694 (2021).
73. Zattoni, I. F. *et al.* A new porphyrin as selective substrate-based inhibitor of breast cancer resistance protein (BCRP/ABCG2). *Chem. Biol. Interact.* **351**, 109718 (2022).
74. Bowers, K. J. *et al.* Scalable algorithms for molecular dynamics simulations on commodity clusters. in *Proceedings of the 2006 ACM/IEEE conference on Supercomputing* 84-es (Association for Computing Machinery, New York, NY, USA, 2006). doi:10.1145/1188455.1188544.
75. Lu, C. *et al.* OPLS4: Improving force field accuracy on challenging regimes of chemical space. *J. Chem. Theory Comput.* **17**, 4291-4300 (2021).
76. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935 (1983).
77. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092 (1993).
78. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684-3690 (1984).
79. Martyna, G. J., Klein, M. L. & Tuckerman, M. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **97**, 2635-2643 (1992).
80. Martyna, G. J., Tuckerman, M. E., Tobias, D. J. & Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* **87**, 1117-1157 (1996).
81. Andrews, S. FastQC: a quality control tool for high throughput sequence data (2018).
82. Schwengers, O. *et al.* Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb. Genomics* **7**, 000685 (2021).
83. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357-360 (2015).
84. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079

(2009).
85. Okonechnikov, K., Conesa, A. & García-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292-294 (2016).
86. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
87. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
88. Guder, J. C., Schramm, T., Sander, T. & Link, H. Time-optimized isotope ratio LC-MS/MS for high-throughput quantification of primary metabolites. *Anal. Chem.* **89**, 1624-1631 (2017).
89. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896-1906 (2007).
90. Semanjski, M. *et al.* The kinases HipA and HipA7 phosphorylate different substrate pools in Escherichia coli to promote multidrug tolerance. *Sci. Signal.* **11**, eaat5750 (2018).
91. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367-1372 (2008).
92. Cox, J. *et al.* Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794-1805 (2011).
93. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4**, 207-214 (2007).
94. Luber, C. A. *et al.* Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. *Immunity* **32**, 279-289 (2010).
95. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337-342 (2011).
96. Aguilan, J. T., Kulej, K. & Sidoli, S. Guide for protein fold change and *p* -value calculation for non-experts in proteomics. *Mol. Omics* **16**, 573-582 (2020).
97. Yanisch-Perron, C., Vieira, J. & Messing, J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene* **33**, 103-119 (1985).
98. MacNeil, D. J. *et al.* Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**, 61-68 (1992).
99. Nadkarni, S. R. *et al.* Balhimycin, a new glycopeptide antibiotic produced by *Amycolatopsis* sp. Y-86, 21022 Taxonomy, production, isolation, and biological activity. *J. Antibiot. (Tokyo)* **47**, 334-341 (1994).
100. Wink, J. M. *et al.* Three new antibiotic producing species of the genus *Amycolatopsis*, *Amycolatopsis balhimycina* sp. nov., *A. tolypomycina* sp. nov., *A. vancoresmycina* sp. nov., and description of *Amycolatopsis keratiniphila* subsp. *keratiniphila* subsp. nov. and *A. keratiniphila* subsp. *nogabecina* subsp. nov. *Syst. Appl. Microbiol.* **26**, 38-46 (2003).
101. Ansorg, R., H, Z. & R, T. Bedeutung des Nachweises antibakterieller Stoffe im Urin für die bakteriologische Diagnostik und die Kontrolle der Chemotherapie von Harnwegsinfektionen. *Bedeut. Nachweises Antibakterieller Stoffe Im Urin Für Bakteriol. Diagn. Kontrolle Chemother. Von Harnwegsinfekt.* (1975).
102. Perez-Riverol, Y. *et al.* The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543-D552 (2022).

4 Manuscript 1: Molecular interplay of the glycopepeptide antibiotics biosynthetic machinery

Dardan Beqaj*, Nina Pfahler, Samuel Wagner, Thilo Stehle, Mirita Franz-Wachtel, Boris Macek, Andreas Kulik, Wolfgang Wohlleben, Evi Stegmann#

**First author*

#Corresponding author

4.1 Contributions

All experiments were planned and designed by Dardan Beqaj and Evi Stegmann.
The mutants of *A. balhimycina* lacking the docking domains were generated by Dardan Beqaj, who performed all the *in vivo* experiments and analyzed and interpreted the HPLC-MS derived data, resulting in Figure 4 and Figure S2. The HPLC-MS measurements included in Figure 4 and Figure S2 were done by Andreas Kulik.
The MST measurements were designed and performed by Nina Pfahler, resulting in Figure 3.
All plasmids and experiments for the interaction study using the bacterial two-hybrid assay were done by Dardan Beqaj which resulted in Figure 5A-C, Figure 6C-H and Figure 7B.
The Blue Native PAGE experiments were done by Dardan Beqaj, with initial supervision from Samuel Wagner. The proteomic HPLC-MS/MS measurements after BN-PAGE were conducted by Mirita Franz-Wachtel and Boris Macek. The provided data with protein list and intensities were analyzed and interpreted by Dardan Beqaj resulting in Figure 7A, Figure S7-S12.
The plasmid *pRM4-bpsC-turboID* for proximity dependent biotinylation was done by Dardan Beqaj as well as the complementation and all downstream experiments. The proteomic HPLC-MS/MS mesaurement was done by Mirita Franz-Wachtel and Boris Macek and the obtained data were further analyzed and interpreted by Dardan Beqaj, resulting in Figure 7C and Figure S13. The MSAs, analysis of AlphaFold models, Alphabridge analysis and PISA analysis were done by Dardan Beqaj and resulted in Figure 5D, Figures 6A-B, and Figure S3-S6.

The initial draft of the manuscript was written by Dardan Beqaj with comments and proof-reads of Evi Stegmann and Wolfgang Wohlleben. Nina Pfahler wrote the original draft for the MST measurements of docking domains.

4.2 Summary of results

This study explores the molecular interactions within the biosynthetic machinery of GPAs, specifically concentrating on balhimycin biosynthesis. While the core enzymes involved in balhimycin production have been well-defined, the intricate interplay between NRPSs and accessory modification enzymes needs further investigation.
To bridge this knowledge gap, this study aims to elucidate the molecular interplay within the balhimycin biosynthetic complex, uncovering so far unknown roles of such interactions, and provide a broader understanding of GPA production.

To systematically analyze these interactions, we adopted a stepwise approach; we initially investigated the NRPS enzymes responsible for assembling the heptapeptide core structure and subsequently focused on the enzymes that introduce key modifications to this scaffold. Since no single method was suitable for capturing all protein-protein interactions within the biosynthetic machinery, we employed a range of complementary techniques tailored to the specific nature of each interaction.
It is well established that non-ribosomal peptide synthetases (NRPSs) interact through specialized docking domains (DDs), which mediate precise intermodular communication and ensure the directional transfer of the growing peptide chain. These domains play a critical role in maintaining the integrity and efficiency of the biosynthetic assembly line [71].
In the balhimycin biosynthetic machinery, we identified such docking domains (DD) in each of the three NRPSs BpsA, BpsB, and BpsC. Specifically, two interacting pairs of docking domains (DD) were found: C-terminal docking domains (cDDs) on the upstream modules and N-terminal docking domains (nDDs) on the downstream partners, forming DD1 (BpsA-BpsB) and DD2 (BpsB-BpsC) (Manuscript 1, Figure 1B and Figure 2).
We investigated the function of these docking domains both *in vitro* and *in vivo*. For the *in vitro* experiments, we did not express the DDs in isolation but instead co-expressed them together with their adjacent catalytic domains (PCP-cDD1^{BpsA} and nDD1-C^{BpsB} for DD1 and PCP-cDD1^{BpsA} and nDD1-C^{BpsB} for DD2) in *E. coli*. This strategy was chosen to preserve the native structural context of the protein-protein interfaces, as the conformation and stability of DD interactions are often influenced by neighboring domains.

Microscale thermophoresis (MST) analysis using these constructs determined dissociation constants (K_D) of ~250 µM for DD1 and ~320 µM for DD2, indicating weak but specific interactions (Manuscript 1, Figure 3A, B). Complementary *in vivo* analyses using deletion mutants lacking the respective DDs (ΔcDD1, ΔcDD2, and ΔcDD1-2) exhibited a significant decrease in balhimycin production, confirming that these domains play a critical role in

biosynthesis (Manuscript 1, Figure 4). Interestingly, low levels of balhimycin production could still be detected in the deletion mutants—both in bioassays against *Bacillus subtilis* and in analytical HPLC-MS measurements—suggesting that the NRPS modules were still able to assemble and interact, albeit less efficiently. This observation implies that neighboring catalytic domains may partially compensate for the absence of DD-mediated interactions. Supporting this notion, MST measurements with DD-deficient constructs still revealed low but detectable binding affinities, specifically, between PCP^{BpsA}/nDD1-C^{BpsB} (Manuscript 1, Figure 3C), and PCP-cDD^{BpsB}/C^{BpsC} (Manuscript 1, Figure 3D). Overall, the experiments clearly showed that the DDs play a mediating role in the NRPS communications but can be supported by structural features of adjacent catalytic domains. These results highlight that although DDs are key mediators of NRPS interactions, adjacent catalytic domains contribute to maintaining connectivity within the biosynthetic assembly line.

The study further investigates protein-protein interactions between NRPSs and accessory proteins, first focusing on MbtH-like proteins (MLPs), small proteins that act as essential cofactors for certain NRPSs [84,93]. They stabilize and enhance the activity of specific adenylation domains within NRPSs and are thought to be involved in the activation of amino acids, thereby playing a critical role in the biosynthesis of various secondary metabolites, including antibiotics and siderophores.

In the balhimycin biosynthetic gene cluster (BGC), a gene encoding an MLP is located directly downstream of *bpsC.* Bacterial two-hybrid (B2H) assays confirmed that MLP_{bal} interacts with all modules of the NRPSs BpsA, BpsB and BpsC (Manuscript 1, Figure 5A). Based on MLP-NRPS interaction from other NP pathways [82,93], it is suggested that A-domains, within modules M1-7, are the target interface of MLP_{bal}. To validate this, we performed predictions of protein complexes using AlphaFold3. The analysis confirmed that MLP_{bal} interacts with A-domains in each module (Manuscript 1, Figure S5).

Despite its essential interaction profile, deletion of MLP_{bal} had no observable effect on balhimycin production, suggesting functional complementation by homologous proteins (Ref). Indeed, homologous MLPs encoded by other BGCs of *A. balhimycina* exhibited interactions with NRPS modules M1-7 of the NRPSs BpsA-C (Manuscript 1, Figure 1B, Table 1), which we tested by B2H assays. Although five homologs have been identified sharing high sequence homology (Manuscript 1, Table 1, Figure S2), only two were capable of interacting with the NRPS modules (Manuscript 1, Figure 5B-C), suggesting potential functional redundancy and a crosstalk between different biosynthetic pathways.

To further explore the specificity of these interactions, we used AlphaFold3 to model complexes between MLP_{bal} and the A-domains of each NRPS module (Manuscript 1, Figure S6). Complementary analyses using multiple sequence alignments (MSAs), Alphabridge, and

PISA identified a putative binding interface (Manuscript 1, Figure 6A,B, Figure S2). These insights served as a basis for a targeted mutagenesis approach aimed to systematically assess the sequence determinants responsible for the lack of interaction in MLP_{8030}. The mutational analysis revealed that multiple substitutions were necessary to establish a functional binding interface. Specifically, replacing the N-terminal residues MNGEPRHVV- with MSNPFDNEDGSFFVL-, the C-terminal residues -RSVRQ with -KSLIREASA, and the RDTG motif with TRVH enabled binding to A-domain-containing modules. Additionally, the substitution of the HRQN motif with FAEV further enhanced binding, suggesting that these residues, located in close spatial proximity to the A-domain interface in the structural models (Figure 6B), are critical for high-affinity binding (Manuscript 1, Figure 6C-H).

In addition to the specific interactions between individual NRPSs and the interactions of the NRPS modules and MLPs, which ensure correct assembly of the heptapeptide backbone, further enzymatic modifications are required to produce the final, biologically active structure of balhimycin. Among these are key oxidative cross-linking reactions that define the three-dimensional conformation essential for target binding and antimicrobial activity. These cross-links between aromatic amino acid residues are catalyzed by cytochrome P450 monooxygenases OxyA, OxyB, and OxyC, which act in concert with the NRPS machinery (Manuscript 1, Figure 1B). Structural analyses [62] and inactivation studies [63,65] have demonstrated that these reactions critically depend on the presence of the X-domain, a unique, non-catalytic domain within GPA NRPSs. It is responsible for the recruitment of the P450 monooxygenases to the peptidyl carrier protein (PCP)-bound intermediate, enabling site-specific cross-linking [62].

Furthermore, another important modification is the halogenation of incorporated β-hydroxytyrosine residues by the halogenase BhaA. Biochemical assays [67] and gene inactivation experiments [44] have shown that this reaction, like the oxidative cross-linking steps, occurs exclusively on PCP-tethered peptides. These findings strongly indicate that physical interactions between the NRPS core and the modifying enzymes are necessary for efficient processing during balhimycin biosynthesis.

To explore these interactions and uncover potential additional protein-protein assemblies, we employed blue native polyacrylamide gel electrophoresis (BN-PAGE) coupled to HPLC-MS/MS. For better resolution, cytosolic and membrane protein fractions were separated and further analyzed independently. Gel lanes were sliced according to molecular weight, enabling fractionation of protein complexes based on their native size. Each slice was subjected to HPLC-MS to identify the present proteins. Overlapping migration patterns and consistent co-elution profiles were used to infer the presence of potential protein complexes. Proteins

displaying similar migration behavior and closely correlated intensity profiles strongly suggest the formation of stable protein assemblies.
In *A. balhimycina* wildtype, we observed that BpsA, BpsB, and BpsC co-migrated in the cytosolic fraction with peak intensities detected in the range of 720 kDa to 1.2 MDa in the cytosolic fraction (Manuscript 1, Figure 7A, Figure S7). These results confirmed their interaction, likely mediated by the DDs. Additionally, the halogenase BhaA was found to be highly abundant within this molecular range, supporting its association with NRPSs, particularly with BpsA and BpsB, which incorporate β-hydroxytyrosines.

In contrast, P450 monooxygenases were found at lower molecular weights in the samples of *A. balhimycina* wildtype strain (Manuscript 1, Figure 7A, Figure S7), suggesting a monomeric state indicating weak or transient interactions, possibly interrupted due to extraction conditions. To further explore the interaction between the P450 monooxygenases OxyA-C and the X-domain of BpsC, we conducted B2H assays. We could show that OxyA and OxyC interact with the X-domain of BpsC, while no interaction was observed for OxyB, suggesting a distinct binding site.
We used AlphaFold3 to predict protein complexes between the X-domain and P450 monooxygenases OxyA-C (Figure S4), revealing that all three enzymes interact with the X-domain through the same interface. These interactions appear transient; however, B2H data do not support the models for OxyB (Manuscript 1, Figure 7C). Notably, in BN-PAGE experiments, OxyB was found in the cytosolic fractions of *A. balhimycina ΔbpsC* at a molecular weight similar to NRPSs BpsA and BpsB, despite the absence of BpsC and consequently the X-domain within it.

To investigate whether OxyB functions earlier in balhimycin peptide formation, we tested its interaction with various domains of BpsB's module 6 using AlphaFold3 but found no significant interactions due to low model accuracy.
In wildtype membrane fractions, the NRPS core complex (BpsA, BpsB, BpsC) co-migrated with BhaA and the ABC transporter Tba, indicating Tba's interaction with the balhimycin biosynthetic machinery (Manuscript 1, Figure 7B, Figure S7B). However, the intensities of NRPSs and BhaA in membrane fractions were lower than in cytosolic fractions. In *A. balhimycina ΔbpsC*, BpsA and BpsB also co-migrated with Tba and other modification enzymes (Figure S7D). Intensity profile comparisons between wildtype and *A. balhimycina ΔbpsC* strains showed that BpsC's absence destabilizes the NRPS complex, evidenced by a broader enzyme distribution across gel slices. These findings suggest that modification enzymes interact with the NRPS core complex in both strains, but the lack of BpsC affects the

stability and formation of these protein complexes, potentially impacting the overall biosynthetic process.

Having observed these interactions, we sought to further explore the protein complexes in a different layer by using proximity-dependent biotinylation using TurboID (TID) as earlier described in this thesis (Paper 1). We generated a recombinant strain by *introducing bpsC-TID* at the ΦC31 attachment site in the genome of *A. balhimycinaΔbpsC* under the control of a constitutive promoter (Bibb et al., 1985). To compare the biotinylated proteins in proximity to BpsC, we used a control strain producing cytosolic TID, generated by integrating the TID-encoding gene into the genome of *A. balhimycina* wildtype as described above. Following enrichment of biotinylated proteins using streptavidin beads, HPLC-MS/MS analysis was performed. Comparative analysis revealed approximately 470 proteins, 10 linked to balhimycin biosynthesis.

Notably, the halogenase BhaA was one of the most prominent proteins enriched (~ 1 $\log_2$ fold change) in the BpsC-TiD strain, confirming its proximity to BpsC and its putative interaction with the NRPSs responsible for incorporation of β-hydroxytyrosines chlorinated by BhaA (Manuscript 1, Figure 7D). This finding complemented the BN-PAGE data, where BhaA co-migrated with the NRPSs (Figure 7A-B, Figure S7A-D). Additionally, two glycosyltransferases, BgtfA and BgtfB, were also enriched (~2 $\log_2$ fold change), indicating their proximity to the biosynthetic machinery (Manuscript 1, Figure 7D). Although less prominent, the P450 monooxygenase OxyC and the NRPSs BpsA and BpsB were also enriched (0.5-0.8 $\log_2$ fold changes), suggesting potential co-localization and interaction.

However, no other biosynthesis enzymes were detected in proximity to BpsC. We attribute this to the protein size of BpsC and the spatial orientation of the TiD tag, which may exceed the ~10 nm distance required for effective biotinylation, potentially limiting the detection of nearby proteins that are not within this range.

Overall, this study provides detailed insights into the protein-protein interactions that underpin GPA biosynthesis. Docking domains facilitate NRPS communication, while accessory proteins, including MLPs, further support proper function of A-domains and even are complemented by distinct homologs in the absence of the BGC encode. Additionally, P450 monooxygenases OxyA-C, the halogenase BhaA and glycosyltransferases BgtfB-C, contribute the chemical complexity by interacting with the NRPS core machinery. BN-PAGE and TurboID experiments support the formation of a dynamic, multi-enzyme complex essential for efficient balhimycin production. These findings deepen our understanding of GPA biosynthesis and build a foundation for further investigation to engineer biosynthetic enzymes from the perspective of protein interaction to putatively generate novel derivatives or enhance production rates.

4.3 Manuscript 1

Title:

Molecular interplay of the glycopepeptide antibiotics biosynthetic machinery

Author list

Dardan Beqaj*[1], Nina Pfahler[5], Thilo Stehle[5], Samuel Wagner[3], Mirita Franz-Wachtel[4], Boris Macek[4], Wolfgang Wohlleben[2], Evi Stegmann[1#]

Affiliations:

1. Interfaculty Institute of Microbiology and Infection Medicine Tübingen, Microbial Bioactive Compounds, University of Tübingen, Auf der Morgenstelle 28, 72076, Tübingen, Germany.
2. Interfaculty Institute of Microbiology and Infection Medicine Tübingen, Microbiology/Biotechnology, University of Tübingen, Auf der Morgenstelle 28, 72076, Tübingen, Germany
3. Section of Cellular and Molecular Microbiology, Interfaculty Institute of Microbiology and Infection Medicine (IMIT), University of Tübingen, Tübingen, Germany.
4. Interfaculty Institute of Cell Biology, University of Tübingen, Tübingen, Germany.

*First author

#Corresponding author: evi.stegmann@uni-tuebingen.de

Abstract

Glycopeptide antibiotics (GPAs) are an important class of natural products, including key clinical drugs such as vancomycin and teicoplanin. The biosynthetic pathway of GPAs involves the coordinated action of non-ribosomal peptide synthetases (NRPS) responsible for biosynthesis of the peptide backbone. Further, the peptide backbone undergoes modifications facilitated by various enzymes, increasing the chemical complexity of GPAs. While many reactions occur during peptide assembly, the interactions between the NRPSs and modification enzymes remain poorly understood. We employed diverse approaches to investigate the enzymatic interplay in the biosynthesis of the vancomycin type GPA, balhimycin. In vitro and In vivo experiments showed that, NRPSs is facilitated by docking domains and their adjacent domains. Bacterial two-hybrid assays and AlphaFold3 predictions revealed interactions between NRPSs and other proteins and further elucidated the interaction interface. Furthermore, BN-PAGE and proximity labelling uncovered interaction between modification enzymes and NRPSs.

Introduction

Natural products are a rich source of clinically relevant compounds, including antibiotics, which have significantly advanced human health. Among these, glycopeptide antibiotics (GPAs) like vancomycin and teicoplanin are vital in combating infections with Gram-positive bacteria, particularly strains resistant to other antibiotics (Butler et al., 2014). GPAs are synthesized by non-ribosomal peptide synthetases (NRPSs) - mega-enzymes, consisting of modules, each responsible for incorporating a specific amino acid into a growing peptide chain (Pelzer et al., 1997; Sosio et al., 2004, 2003; van Wageningen et al., 1998). A typical module contains catalytic domains such as adenylation (A), condensation (C), and thiolation (T) domains, and can also include various domains, such as an epimerization (E) or an S-adenosylmethionine (SAM)-dependent methyltransferase domain, increasing peptide diversity (Marahiel et al., 1997).

GPAs are classified into five main types (I–V) based on their peptide backbone composition (Nicolaou et al., 1999). The biosynthesis of GPAs has been extensively studied in *Amycolatopsis balhimycina*, the producer of balhimycin. Balhimycin is a type I GPA that differs from vancomycin only in its glycosylation pattern while sharing the same mechanism of action. The balhimycin biosynthetic assembly line consists of three NRPSs (BpsA–C), which synthesize the heptapeptide backbone by incorporating the proteinogenic amino acids *L*-leucine and asparagine, along with the non-proteinogenic amino acids β-hydroxytyrosine, hydroxyphenylglycine, and dihydroxyphenylglycine, in a precise order. Additional chemical modifications are introduced by various enzymes, including three P450 monooxygenases (OxyA–C) for cross-linking, a halogenase (BhaA), three glycosyltransferases (BgtfA–C), and a methyltransferase (Bmt). The final compound is exported via the ABC transporter Tba (Gericke et al., 2025; Menges et al., 2007). Nearly all genes encoding these essential enzymes are clustered within a biosynthetic gene cluster (BGC)(Gavriilidou et al., 2024; Stegmann et al., 2010).

The biosynthesis of GPAs requires a highly coordinated series of steps: precursor amino acid biosynthesis, heptapeptide assembly, peptide modification, and final export. Among these, peptide modifications such as chlorination by BhaA (Puk et al., 2002) and cross-linking by OxyA-C occur on the NRPS complex during assembly(Bischoff et al., 2001; Gavriilidou et al., 2024, 2024; Evi Stegmann et al., 2006; Woithe et al., 2007), suggesting that protein-protein interactions are necessary to bring modification enzymes into proximity with the NRPS-bound peptide intermediates. Evidence supporting this multi-enzyme complex hypothesis includes *in vitro* studies showing that specific NRPS domains, such as the X-domain, serve as platforms for recruiting P450 monooxygenases (Haslinger et al., 2015). In addition, the role of the ABC transporter Tba was shown to go beyond export, indicating its proximity to and potential interaction with the biosynthetic machinery, coupling balhimycin biosynthesis to the presence of a functional ABC transporter (Gericke et al., 2025). However, interactions involving other biosynthetic enzymes remain largely unexplored.

In living systems, multi-enzyme complexes contribute to efficient functioning of biosynthetic pathways by facilitating substrate channelling, enhancing reaction rates, and minimizing undesirable side reactions (Kerfeld and Erbilgin, 2015). The organization of enzymes into complexes is especially critical in secondary metabolism, where high substrate specificity and coordination are required to produce complex molecules (Chen et al., 2021, 2020; Gacek-Matthews et al., 2020). Thus, understanding the precise interactions that drive the assembly

of multi-enzyme complexes is essential for deciphering the biosynthesis of natural products and engineering enhanced or novel bioactive compounds.

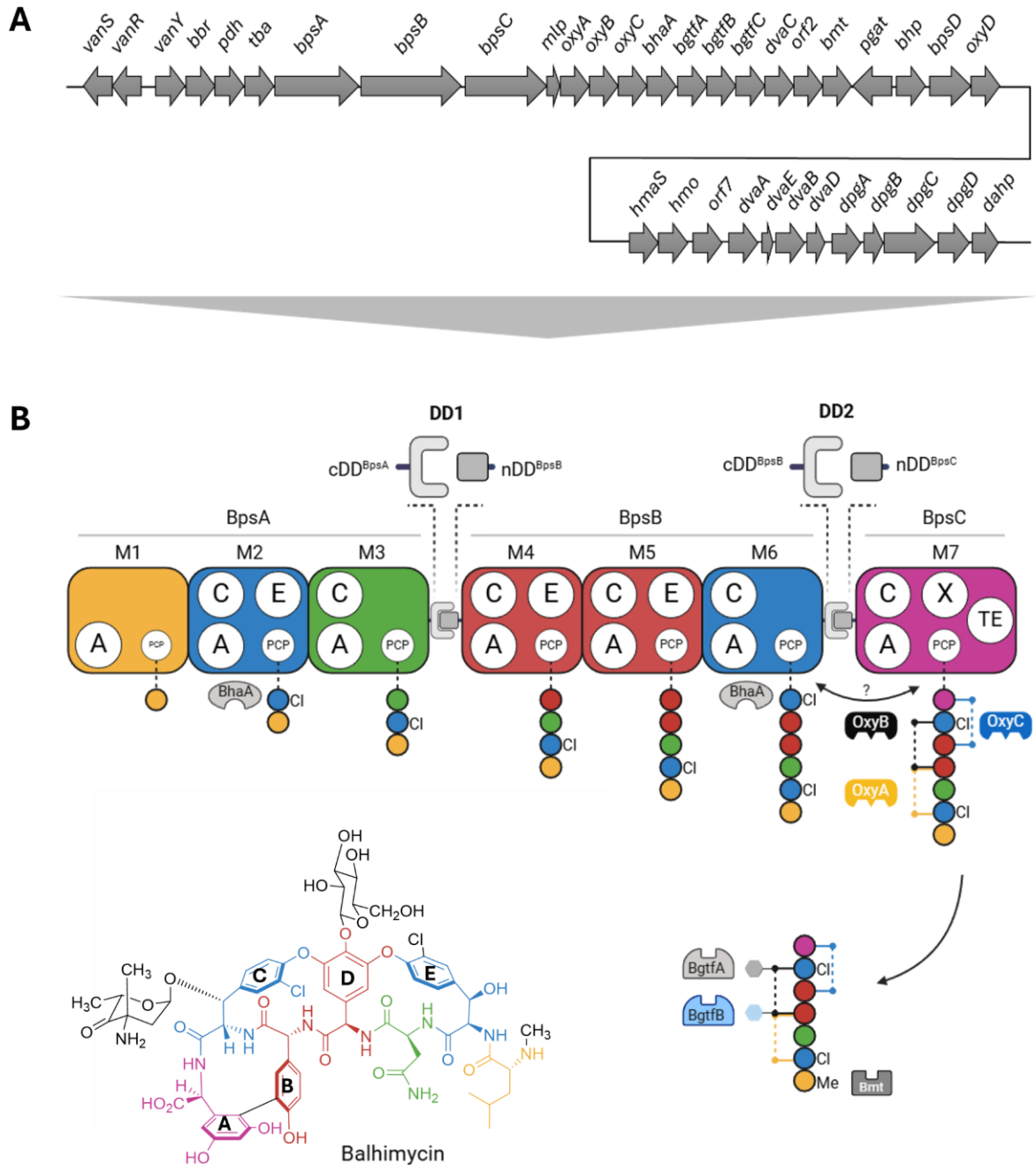


Figure 1: Schematic overview of the balhimycin BGC and biosynthesis. (A) The BGC contains nearly all biosynthetic genes encoding for a two component system (*vanRS*), a carboxypeptidase (*vanY*), an StrR-like transcriptional regulator (*bbr*), a prephenate dehydrogenase (*pdh*), an ABC transporter (*tba*), the non-ribosomal peptide synthetases (*bspA-C*), an MbtH-like protein (*mlp*), the P450 monooxygenases (*oxyA-C*), a halogenase (*bhaA*), the glycosyltransferases (*bgtfABC*), a putative deacetylase (*orf2*), the N-methyltransferase (*bmt*), a phenylglycine aminotransferase (*pgat*), the genes involved in the biosynthesis of β-hydroxytyrosine (*bhp, bpsD, oxyD*), the genes involved in the biosynthesis of hydroxyphenylgylcine (*hmaS, hmo*), a putative Na+/H+ antiporter (*orf7*), the genes involved in the biosynthesis of the sugar moieties (*dvaA-D*), the genes involved in the biosynthesis of

dihydroxyphenylglycine (*dpgA-D*) and the DAHP synthase (*dahp*). **(B)** The biosynthesis is catalyzed by the NRPSs BpsA-C. Colors indicate modules (M1-M7) and the corresponding amino acids incorporated into a growing peptide chain. The halogenase (BhaA) chlorinates amino acids two and six. The P450 monooxygenases (OxyA-C) catalyze crosslinking reactions, and glycosyltransferases (BgtfA and BgtfB) attach sugar moieties. The N-methyltransferase (Bmt) attaches a methyl group to the amino acid in position one.

This study aimed to deepen our understanding of the protein-protein interactions between individual NRPSs and the associated enzymes, using balhimycin biosynthesis as a model (Figure 1). A combined approach of biochemical, genetic and chemical analyses identified specific interactions between NRPS domains and confirmed that modification enzymes and small accessory proteins are integral components of the biosynthetic machinery. Computational modelling via AlphaFold3 provided insights into potential interaction interfaces and key residues involved in these interactions. Together, these findings enhance our understanding of the structural organization and functionality of multi-enzyme biosynthetic complexes.

Results

Specific docking domain-mediated communication between GPA NRPSs enables efficient GPA production

The assembly of the balhimycin heptapeptide backbone is catalyzed by three NRPSs (BpsA, BpsB and BpsC), requiring precise coordination between these enzymes to ensure the efficient biosynthesis of a peptide with a unique amino acid sequence (Figure 1). While Stachelhaus previously investigated communication-mediating domains (COMs), also referred to as docking domains (DDs) in NRPSs (ref.), their role in GPA NRPSs remains largely unexplored. We initiated our analysis by identifying the non-catalytic DD in various GPA NRPSs. *In silico* sequence analysis based on multiple sequence alignments (MSA) and gene annotation revealed the presence of putative C-terminal DDs (cDDs), consisting of ~37 amino acids, and N-terminal DDs (nDDs), short segments of 5–12 amino acids. These cDDs and nDDs form two distinct DD pairs, DD1 and DD2, at the interface between adjacent NRPSs. To investigate their predicted role, we analyzed the DD pairs of the balhimycin NRPS (Figure 2).

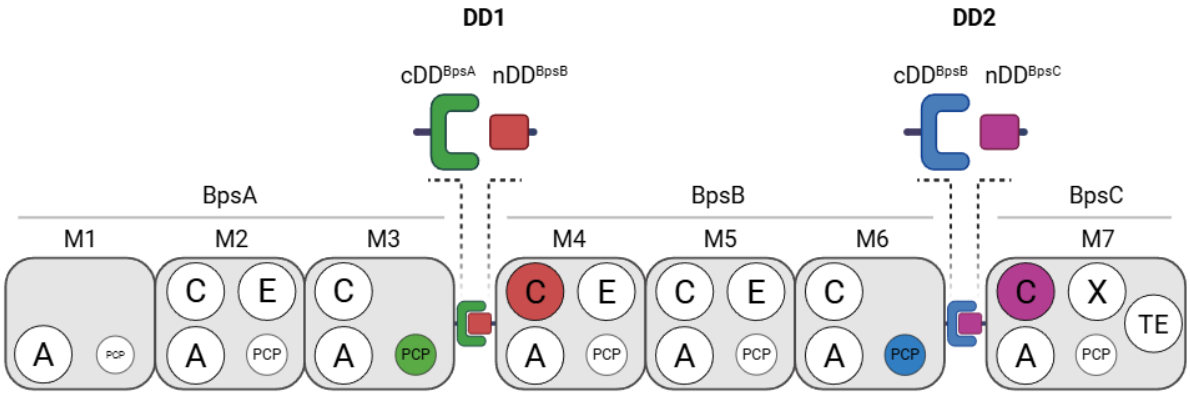


Figure 2: **Schematic representation of the balhimycin assembly line highlighting the location of DDs.** Shown are the NRPSs BpsA-C containing modules M1-7. In green is shown the region of PCP-cDD1^{BpsA} and in red the region of nDD1-C^{BpsB} forming the pair DD1. In blue is shown PCP-cDD2^{BpsB} and in dark pink nDD2-C^{BpsC} forming the pair DD2. The colored regions were used in our MST measurements and deleted in our knock-out experiments.

We first set out to determine the affinities between individual DDs of each pair by performing microscale thermophoresis (MST). We expressed the DD encoding region with its up- or downstream adjacent catalytic domain, which resulted in the proteins PCP^{BpsA}-cDD1, nDD1-C^{BpsB}, PCP^{BpsB}-cDD2, and nDD2-C^{BpsC} (Figure 2). For the BpsA/BpsB interface, mimicked by PCP^{BpsA}-cDD1/nDD1-C^{BpsB} , a K_D of ~250 µM with a 95% confidence level could be measured (Figure 3A). Similar results were obtained for the BpsB/BpsC DD interface, where the interaction between PCP^{BpsB}-cDD2 and nDD2-C^{BpsC} showed a K_D of ~320 µM with a 95% confidence interval (Figure 3B). Both K_D values indicate weak protein-protein interactions, meaning that only apparent affinities can be reported. This limitation arises because weak interactions require high ligand concentrations to reach saturation, which is not feasible in MST measurements. However, relative comparisons between affinities remain valid. Based on our data, the BpsA/BpsB interface exhibits a higher binding affinity compared to the interface of BpsB/BpsC.

To better understand the functional role of the DDs in GPA biosynthesis, we employed a gene knock-out approach to generate *A. balhimycina* mutants lacking these domains. Specifically, we constructed *A. balhimycinaΔcDD1*, lacking the C-terminal DD of BpsA; *A. balhimycinaΔcDD2*, lacking the C-terminal DD of BpsB; and *A. balhimycinaΔcDD1-2*, a double mutant lacking both DDs. These mutants are hereafter referred to as ΔcDD1, ΔcDD2, and ΔcDD1-2, respectively.

Culture filtrates from each mutant exhibited significantly reduced inhibitory activity against *Bacillus subtilis* (Figure 4A), with inhibition observed only after 120 hours of cultivation. To confirm that the observed inhibition was due to balhimycin production, we performed high-performance liquid chromatography coupled with electrospray ionization mass spectrometry (HPLC-ESI-MS) to detect masses corresponding to balhimycin (Figure 4B). In all three mutants (ΔcDD1, ΔcDD2, and ΔcDD1-2), peaks corresponding to balhimycin appeared at retention times of 5–5.5 minutes, consistent with the wildtype control (Figure S2). The analysis identified peaks with m/z 1305.34 $[M+H]^+$ for mono-glycosylated balhimycin and m/z 1446.46 $[M+H]^+$ for bi-glycosylated balhimycin. Notably, in the mutants, the mono-glycosylated form (m/z 1305.34 $[M+H]^+$) was predominant, except in the ΔcDD1 mutant (Figure S2). For quantitative comparison, HPLC-ESI-MS data were used to measure production levels, revealing a correlation between the inhibitory activity and chromatographic profiles (Figure 4C). The production level of ΔcDD1 was slightly higher than ΔcDD2 and ΔcDD1-2. These findings suggest that DDs play a crucial role in balhimycin biosynthesis, potentially influencing glycosylation efficiency and overall GPA production.

However, the fact that biosynthesis was not entirely abolished in the DD deletion mutants indicated that additional domains located up- or downstream the DDs, may also contribute to NRPS interaction. To investigate this possibility at the protein level, we conducted biochemical affinity measurements using the adjacent domains lacking the DD-encoding regions. We generated the constructs expressing PCP^{BpsA}, C^{BpsB}, PCP^{BpsB}, and C^{BpsC} and assessed their interactions. Interestingly, when measuring the affinity between PCP^{BpsA} and nDD1-C^{BpsB}, which mimics the *in vivo* situation in the *A. balhimycinaΔcDD1* mutant, the K_D value exceeded 10mM (Figure 3C), indicating a weak interaction between the tested proteins. Comparable values were obtained for the interaction between PCP^{BpsB}-cDD2 and C^{BpsC} (Figure 3D). These

results suggested that domains adjacent to the DDs may play a crucial role in facilitating NRPS interactions, even when one DD is absent. This aligns with our *in vivo* experiments, where balhimycin production was detected after extended cultivation. While MST measurements indicated a slightly higher affinity between PCP^{BpsA}-cDD1 and nDD1-C^{BpsB}, consistent with the *in vivo* findings that the ΔcDD1 mutant displays higher inhibitory activity and balhimycin production than ΔcDD2, statistical analysis (t-test) revealed no significant differences in production titers among the mutants (Figure 4C).

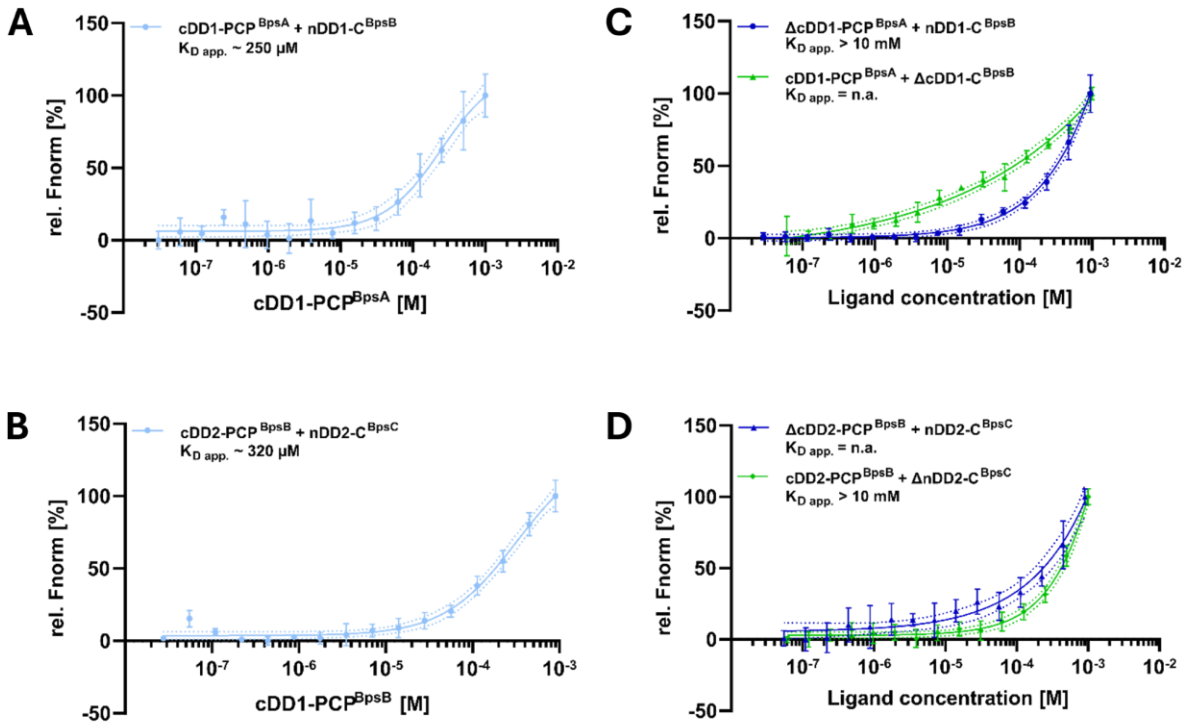


Figure 3 Affinity between adjacent DD pairs as measured by MST. MST measurements of various constructs expressing the DDs and adjacent domains. (**A**) Affinity between PCP^{BpsA}-cDD1 and nDD1-C^{BpsB}. (**B**) Affinity between PCP^{BpsB}-cDD2 and nDD2-C^{BpsC}. In (**C**) is shown the affinity between PCP^{BpsA} and nDD-C^{BpsB} and between PCP-$cDD1^{BpsA}$ and C^{BpsB}. **(D)** Shows the affinity between PCP^{BpsB} and nDD2-C^{BpsC} and between cDD2-PCP^{BpsB} and C^{BpsC}. The measurements show an approximate binding affinity K_D in the presence of DDs which ranges from 250-320µM (**A** and **B**). The affinity decreases in the absence of DDs to a K_D of 10mM (**C** and **D**). All data is plotted in relative, normalized fluorescence [%] against the ligand concentration [M] in a $\log_{10}$ scale. MST measurements: MST power = medium, excitation power = 20-100% (auto-detect). Error bars indicate standard deviations.

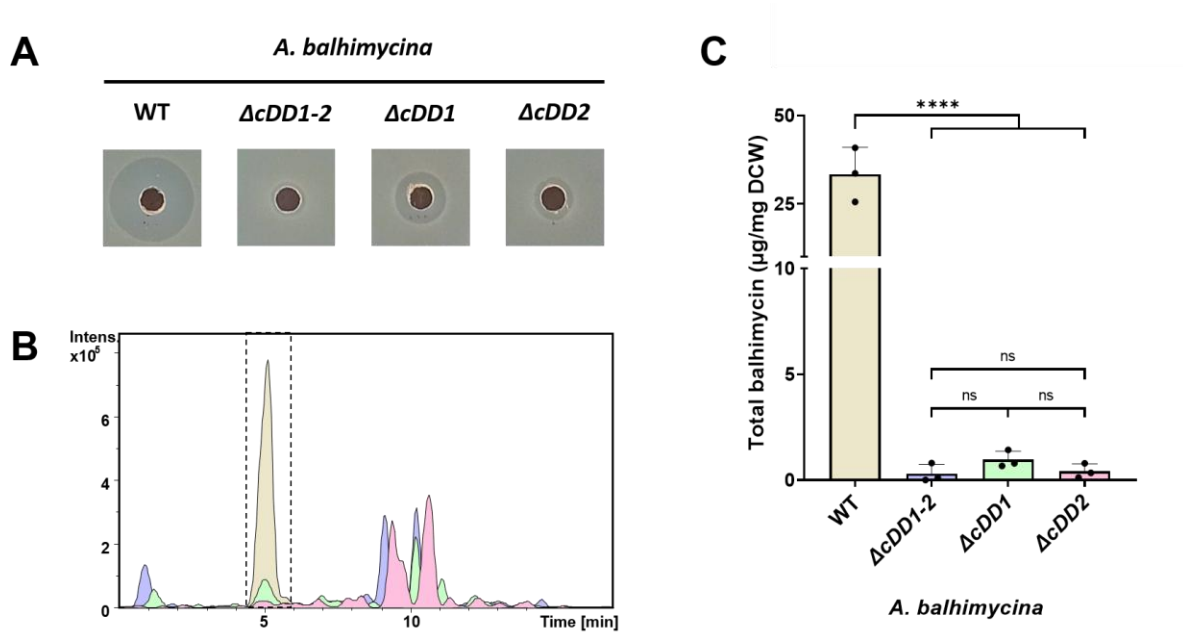


Figure 4: Bioactivity test and HPLC-MS analysis of balhimycin production in the DD mutants. (A) Bioactivity assay against *B. subtilis* exposed to the culture filtrate of the wild-type (WT) and respective mutants. **(B)** Overlayed HPLC-MS chromatograms of the culture filtrates. Dashed lines indicate the peaks corresponding to balhimycin (see Figure S2). **(C)** Quantified amounts of balhimycin produced by each mutant based on the HPLC-MS analysis. Error bars indicate standard deviations of three biological replicates, (****) p<0.05; (ns) not significant.

Multiple MLPs can interact with balhimycin NRPSs

The NRPS interactions via the DDs is necessary for the formation of the NRPS core complex, which assembles the heptapeptide backbone. Moreover, the assembly of the GPA backbone requires prior recognition and activation of the amino acids, a process catalyzed by the adenylation domains of the NRPSs (Gulick, 2009). The proper function of these domains has been reported to depend on an accessory protein known as MbtH-like protein (MLP), co-encoded in NRPS BGCs (Herbst et al., 2013). In the balhimycin BGC, the MLP-encoding gene *mlp*$_{bal}$, is located downstream of the NRPS genes *bpsA-C* (Figure 1A), similar to all other GPA BGCs (Table S6). Previous deletions of *mlp*$_{bal}$ from the balhimycin BGC had no effect on production (Stegmann et al., 2006)

Using the bacterial two-hybrid (B2H) assay, we performed pairwise experiments to identify potential interactions between MLP$_{bal}$ and NRPSs. For this, we used pKT25 to separately clone the NRPS modules (M1-7, Figure 1B) of *bpsA-C,* each containing the respective A-domain (Figure 1B and Table S3), and cloned the *mlp*$_{bal}$ gene into pUT18c. Our results showed that MLP$_{bal}$ interacts with all A-domain containing modules (Figure 5A), highlighting their role in NRPS functionality. This supports the hypothesis that MLPs encoded by homologous genes elsewhere in the genome might compensate for the MLP function in the absence of MLP$_{bal}$ (Stegmann et al., 2006b).

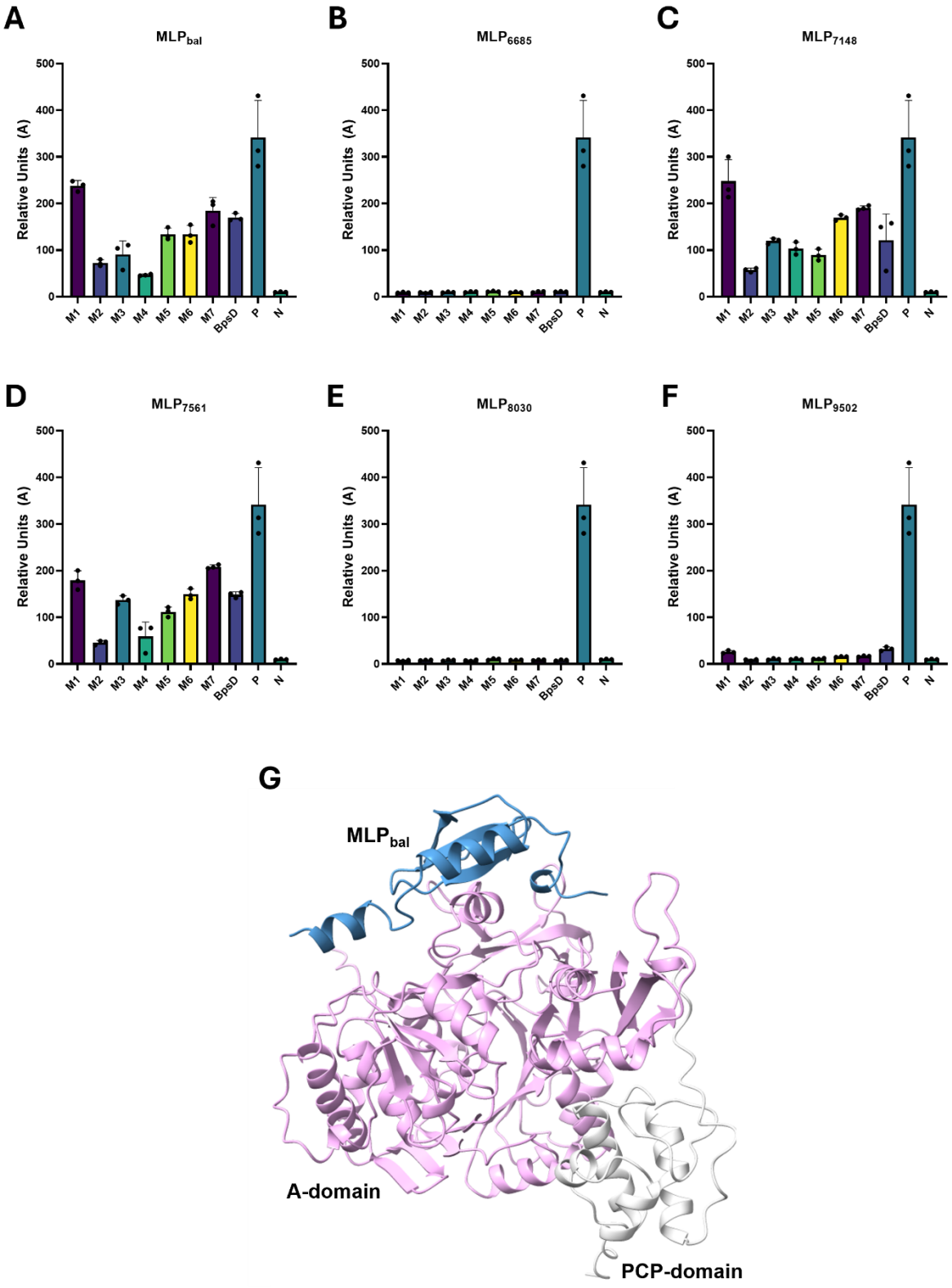


Figure 5: Bacterial two-hybrid interaction assays of modules 1-7 (M1-M7) of the NRPSs BpsA-C with MLPs and AlphaFold3 generated protein complex. (A-F) BTH101 was co-transformed with plasmids encoding the respective T25 and T18 cya fusions. Bars indicate units of activity of at least three independent co-transformants, represented as averages. pKT25-zip/pUT18c-zip represent the positive control (P) and a negative control (N), BTH101 carrying the empty vector plasmids pKT25 and pUT18c was used. Dots represent replicates; error bars indicate standard deviations.**(G)** The complex displays the interaction between M1(consisting of A-domain in light pink and PCP-domain in light grey)

and MLP_{bal} (blue), showing its binding on the A-domain. The model scored an ipTM: 0.96 and pTM: 0.84 suggesting high accuracy of prediction. Further models are shown in Figure S5.

To test this hypthesis, we mined the genome of *A. bahlimycina* and identified five additional MLP-encoding genes (*mlp_{7148}, mlp_{7561}, mlp_{9502}, mlp_{6685}* and *mlp_{8030}*), each associated with predicted NRPS BGCs (Table S5). Sequence analysis revealed that MLP_{7148} and MLP_{7561} share the highest similarity with MLP_{bal}, while MLP_{9502}, MLP_{6685}, and MLP_{8030} display lower sequence similarities (Table 1).

To assess whether these genes can functionally replace MLP_{bal}, we tested them (Table 1) in B2H assays to evaluate their interactions with the NRPS modules M1-7 (Figure 1B). Only MLP_{7148} and MLP_{7561} exhibited interactions across all NRPS modules (Figure 5C and D, while MLP_{9502}, MLP_{6685}, and MLP_{8030} showed no interaction (Figure 5B, E and F). This suggested that only MLP_{7148} and MLP_{7561} could potentially compensate for *mlp_{bal}* deletion and restore NRPS biosynthetic function in its absence. To further confirm that A-domains in the modules M1-7 mediate the interaction with MLPs, we predicted complexes between each module (M1-7) and MLP_{bal} using AlphaFold3 (Figure 5D, Figure S5). The models showed that in each complex the MLP_{bal} interacting site is located in the A-domain.

Table 1: Information on the MLPs encoded in the A. balhimycina genome.

Nr	Name	Accession number	% Similarity*	Length**
1	MLP_{bal}	RSM50666.1	100 %	69
2	MLP_{7148}	RSM36180.1	70.31%	74
3	MLP_{7561}	RSM36806.1	65.08%	69
4	MLP_{9502}	RSM46668.1	58.06%	66
5	MLP_{6685}	RSM48734.1	55.77%	64
6	MLP_{8030}	RSM43807.1	45.61%	60

*Similarity is compared to MLP_{bal}. **Number of amino acids.

Specific structural features are required for stable interactions between MLPs and A-domains

Despite their high sequence similarity, our B2H experiments shows that specific structural features are essential for interaction between MLPs and A-domains. To explore this, we conducted MSAs of MLPs encoded in GPA BGCs, including MLP_{bal} and all *A. balhimycina* MLPs homologs (Table 1), to identify conserved residues and motifs (Figure S3). The non-interacting MLPs of *A. balhimycina* (Table1, Nr 4-5) differ from the GPA BGC-encoded MLPs and the interacting *A. balhimycina* MLP homologs (Table 1, Nr 2 and 3) in their N-terminal sequences as well as in distinct C-terminal variations. In addition, the MSAs revealed differences in two motifs, FAEV and TRVH, which are absent in non-interacting MLPs, leading to the hypothesis that these residues may play a crucial role in stabilizing the interface needed for A-domain docking.

To further validate this hypothesis, we used AlphaFold3 to predict protein complexes between MLP_{bal} and the A-domains (Figure S6) in addition to the previously generated models using M1-7 (Figure 5D, Figure S5). We set out to evaluate the prediction using the ipTM and pTM scores provided by AlphaFold3. The pTM score, a measure of the overall accuracy, ranged between 0.84 and 0.93, suggesting high confidence models. The ipTM score, a measure of the accuracy of the relative position of the subunits, ranged from 0.95 to 0.96, showing high confidence of the models. In addition to this, Alphabridge and PISA analysis allowed us to

identify residues in close contact forming the putative binding interface (Figure 6A,B). Our models revealed that MLP_{bal} consist of an N-terminal loop followed by a three-stranded β-sheet, and two α-helices. Notably, the β-sheet and the first α-helix form the primary interface exposed to an α-helix of the A-domain. Based on these structural insights, we identified key interacting parts, which we then targeted in a follow-up mutagenesis experiment to prove their role in binding.

To systematically assess the impact of changes in sequence, we introduced targeted substitutions in the non-interacting MLP_{8030}, beginning with the N-terminus, followed by the C-terminal region, and finally the TRVH and FAEV motifs (Table 2). We then evaluated their effects on A-domain binding using B2H assays. This analysis revealed that multiple substitutions were necessary to establish an effective binding interface. Specifically, replacing the N-terminal residues MNGEPRHVV- with MSNPFDNEDGSFFVL-, the C-terminal residues -RSVRQ with -KSLIREASA, and the RDTG motif with TRVH enabled interaction with A-domain-containing modules. Additionally, replacing the HRQN motif in MLP_{8030} with FAEV further enhanced binding, suggesting that these residues, which are positioned towards the A-domain residues in our models (Figure 6B), are essential for high-affinity binding.

Table 2: Mutational variants of the MLP_{8030} sequence.

Variant	Exhanged sequence*
MLP_{v1}	MSNPFDNEDGSFFVLVNHEEQYSIWPAHRQNPPGWRDTGHEGTRAECLDHVERVWTDLRPRSVRQ*
MLP_{v2}	MSNPFDNEDGSFFVLVNHEEQYSIWPAHRQNPPGWRDTGHEGTRAECLDHVERVWTDLRPKSLIREASA
MLP_{v3}	MSNPFDNEDGSFFVLVNHEEQYSIWPAHRQNPPGWTRVHHEGTRAECLDHVERVWTDLRPKSLIREASA
MLP_{v4}	MSNPFDNEDGSFFVLVNHEEQYSIWPAFAEVPPGWTRVHHEGTRAECLDHVERVWTDLRPKSLIREASA
MLP_{v5}	MSNPFDNEDGSFFVLVNHEEQYSIWPAFAEVPPGWRDTGHEGTRAECLDHVERVWTDLRPRSVRQ
MLP_{v6}	MNGEPRHRVVVNHEEQYSIWPAFAEVPPGWRDTGHEGTRAECLDHVERVWTDLRPRSVRQ

Underlined are the exchanged residues deriving from MLP_{bal}

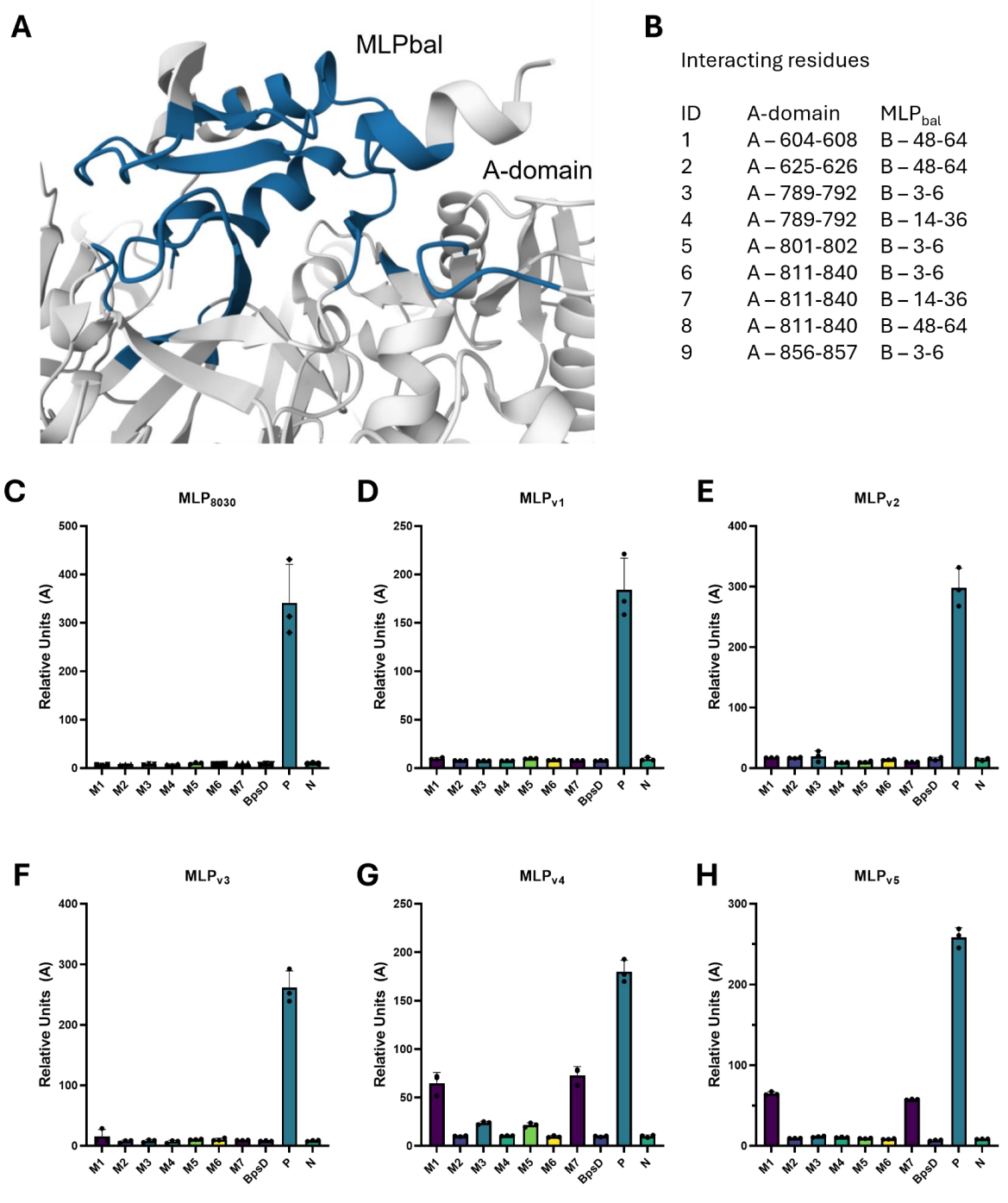


Figure 6. Analyzed interaction interface between MLPbal and A-domains of BpsA-C and the validation of important predicted residues via mutagenesis of MLP8030 and evaluation with the B2H assay. (A) Interaction interface between MLPbal and A-domain of M1. Highlighted are the residues predicted from A-domain and MLPbal to be involved in the interaction and additionally listed in **(B)**. **(C-H)** BTH101 was co-transformed with plasmids encoding the respective T25 and T18 cya fusions of M1-7 and MLP_{8030} variants listed in Table 2. Bars indicate units of activity of at least three independent co-transformants, represented as averages. pKT25-zip/pUT18c-zip represent the positive control (P) and a negative control (N), BTH101 carrying the empty vector plasmids pKT25 and pUT18c was used. Dots represent replicates; error bars indicate standard deviations

Modification enzymes interact with the NRPSs

After confirming the interaction between NRPSs and MLPs, we focused on modification enzymes. Previous studies on GPA biosynthesis have shown that halogenases and oxygenases catalyze their reactions on the NRPS-bound peptide, suggesting direct interactions between these enzymes and NRPS (ref). Specifically, it was demonstrated that the X-domain of the final NRPS module (BpsC in the case of bahlimycin biosynthesis) is essential for recruiting oxygenases. However, the full range of GPA biosynthetic enzymes that may physically interact with NRPSs remains unknown. To investigate this, we used native complex separation by coupling blue native polyacrylamide gel electrophoresis (BN-PAGE) with mass spectrometry (MS) to identify co-migrating proteins. This approach provided a snapshot of balhimycin biosynthetic enzymes involved in potential protein-protein interactions and helped uncover additional enzymes that may be part of the multi-enzyme complex.

To obtain a better resolution, we separated membrane and cytosolic protein fractions, generating lane intensity profiles by slicing the gel and analyzing each segment by MS. We then compared the profiles of the wildtype *A. balhimycina* with those of the control strain, *A. balhimycina ΔbpsC*, which lacks BpsC, the NRPS responsible for incorporating the seventh amino acid, (S)-3,5-dihydroxyphenylglycine, into the balhimycin backbone. MS analysis identified approximately 3000 proteins, including 21 directly involved in balhimycin biosynthesis, spanning precursor supply, backbone synthesis, or modification. Lane intensity profiles showed that proteins co-migrating within specific gel slices likely form complexes, particularly when their molecular weights and intensity distributions are aligned .

In *A. balhimycina* wildtype, BpsA, BpsB, and BpsC exhibited peak intensities within the 720 kDa–1.2 MDa range in the cytosolic fraction (Figure 7A), indicating their interaction presumably due to the DDs. Notably, the halogenase BhaA was also highly abundant in this range, suggesting its association with NRPSs, specifically BpsA and BpsB, which incorporate the β-hydroxytyrosines that are subsequently chlorinated by Bha. Conversely, P450 monooxygenases were detected at lower molecular weights in the wildtype strain, indicating a monomeric state. We speculate that the BN-PAGE extraction conditions may have disrupted weak or transient interactions. To further investigate the interaction of the P450 monooxygenases OxyA-C with the X-domain of BpsC, we performed B2H assays. The X-domain was expressed using pKT25, while OxyA-C were expressed from pUT18c.

Interestingly, only OxyA and OxyC interacted with the X-domain, while no interaction was observed for OxyB (Figure 5). To exclude the possibility that C-terminal fusion of the cya domain T18 to OxyB may have interfered with binding, we generated constructs with an N-terminal T18 fusion to OxyB. However, also N-terminally tagged OxyB did not show interaction with the X-domain of BpsC (data not shown). . To determine putative interactions by means of an orthogonal approach, we predicted protein complexes between the X-domain and the P450 monooxygenases OxyA-C using AlphaFold3. The models revealed that all three enzymes interact with the X-domain (Figure S4) via the same interaction interface. These results suggested that the interactions are transient . However, the B2H data do not support these models for OxyB. Intriguingly, in the BN-PAGE experiments OxyB was detected in the *A. balhimycina ΔbpsC* cytosolic fractions at the same molecular weight range as the NRPSs BpsA and BpsB (Figure S7), despite the absence of BpsC. To explore the possibility that OxyB may act at an earlier stage of balhimycin peptide formation, as previously reported (Bischoff et al.,

2005; Evi Stegmann et al., 2006), we tested its interaction with various domains of module 6 of BpsB but found no significant interactions. Additionally, attempts to predict possible binding interfaces using AlphaFold3 were unsuccessful due to the low accuracy of the models.

Interestingly, in membrane fractions of the wildtype, the NRPS core complex (BpsA, BpsB, and BpsC) co-migrated alongside BhaA and the ABC transporter Tba, suggesting that Tba interacts with the balhimycin biosynthetic machinery , as previously shown(Gericke et al., 2025). However, the intensities of the NRPSs and BhaA in the membrane fractions were lower than in the cytosolic fractions. Similarly, in *A. balhimycinaΔbpsC* BpsA and BpsB showed co-migration with Tba and BhaA and other modification enzymes (Figure S7C,D).

Comparison of intensity profiles between the wildtype and *A. balhimycinaΔbpsC* strains indicated that the absence of BpsC destabilizes the NRPS complex, as evidenced by a broader distribution of the enzymes across gel slices (Figure S7A,C). These findings suggest that modification enzymes interact with the NRPS core complex in both the wildtype and *A. balhimycina ΔbpsC* strains, although the absence of BpsC alters the formation and stability of these protein complexes. Specifically, the absence of BpsC appears to destabilize the NRPS complex, resulting in a broader distribution of associated proteins, which may impact the overall biosynthetic process.

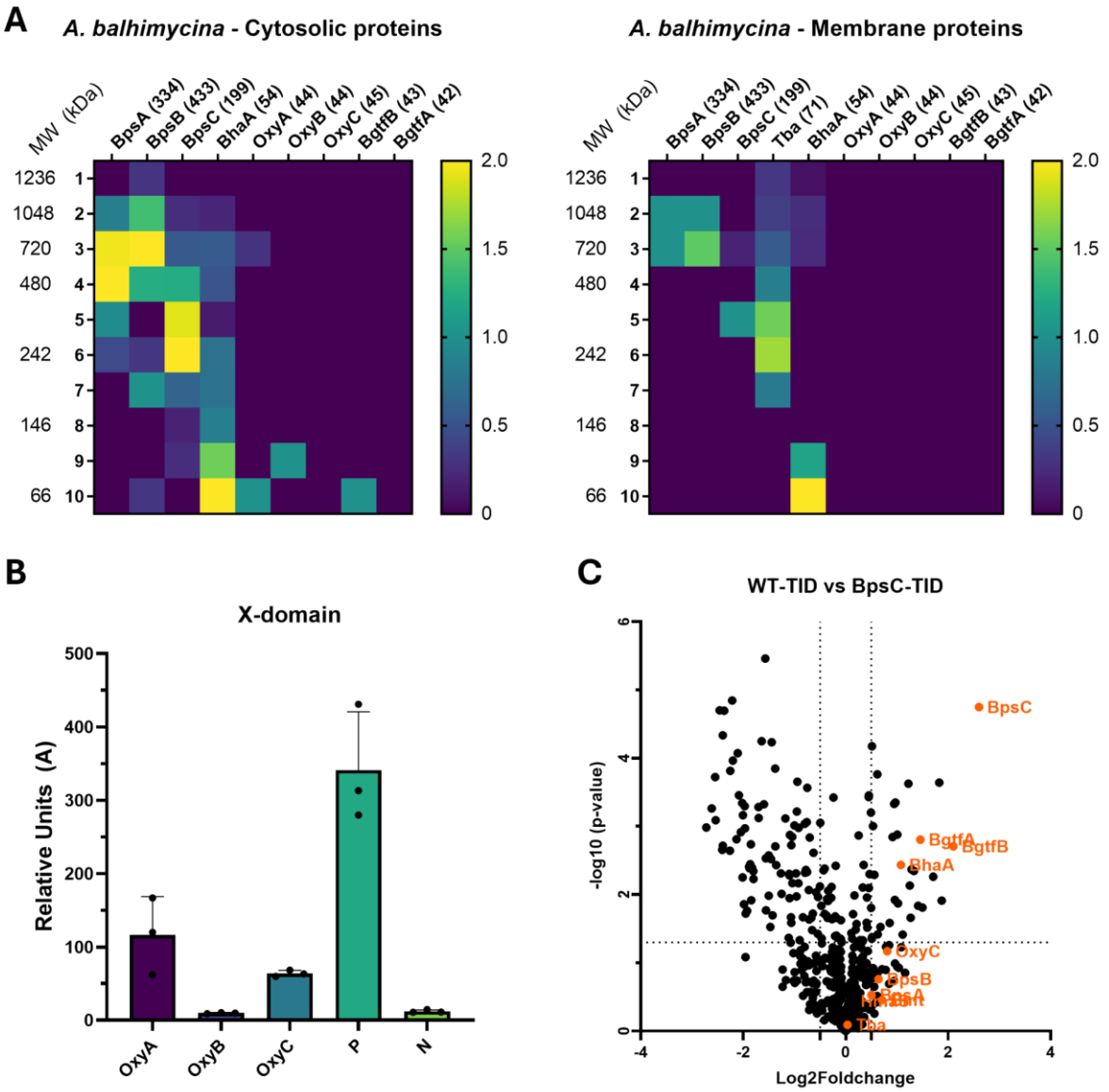


Figure 7. Identification of interacting proteins using BN-PAGE, B2H assays and proximity dependent labelling. (A) Heatmap after analysis of the LC-MS/MS data from BN-PAGE, showing high intensity profiles of proteins putatively present in a complexed form. (B) BTH101 was co-transformed with plasmids encoding the respective T25 and T18 cya fusions (Table S3). Bars indicate units of activity of at least three independent co-transformants, represented as averages. pKT25-zip/pUT18c-zip represent the positive control (P) and a negative control (N), BTH101 carrying the empty vector plasmids pKT25 and pUT18c was used. Dots represent replicates; error bars indicate standard deviations (C) Volcano plot after proximity dependent biotinylation experiments in which BpsC was tagged using TurboID. On the right are the proteins with a positive log2 fold change, where balhimycin BGC encoded enzymes are colored in orange.

Having observed potential interactions between the NRPSs and various biosynthetic enzymes through BN-PAGE, we sought to further investigate these interactions at a more granular level. To achieve this, we employed proximity-dependent labeling, a technique previously established for *A. balhimycina* in our group (Gericke et al., 2025). This approach allows for the

identification of proteins in close proximity to a target protein in the native producer, providing deeper insights into the molecular interactions within the balhimycin biosynthetic machinery.

The TurboID-tag (TiD), an optimized biotin ligase variant fused to the protein of interest, enables the identification of proteins within a range of ~10 nm. We generated a recombinant strain by introducing *bpsC-TiD* at the ΦC31 attachment site in the genome of *A. balhimycinaΔbpsC* under the control of a constitutive promoter (Bibb et al., 1985). Prior to protein extraction and analysis, we confirmed the restored production of fully active balhimycin by HPLC-MS and bioactivity assays, ensuring that the BpsC-TiD fusion protein was functional (Figure S13A). To compare the biotinylated proteins in proximity to BpsC, we used a control strain producing cytosolic TiD, generated by integrating the TiD-encoding gene into the genome of *A. balhimycina* wildtype as described above. Following enrichment of biotinylated proteins using streptavidin beads, HPLC-MS/MS analysis was performed. Comparative analysis revealed approximately 470 proteins, 10 linked to balhimycin biosynthesis. For robust statistical validation, we focused on proteins detected in at least two replicates to avoid biased conclusions.

Notably, the halogenase BhaA was one of the most prominent proteins enriched in the BpsC-TiD strain (~ 1 $\log_2$ fold change), confirming its proximity to BpsC and its putative interaction with the NRPSs responsible for incorporation of β-hydroxytyrosines chlorinated by BhaA after its incorporation to the growing peptide chain. This finding complemented the BN-PAGE data, where BhaA co-migrated with the NRPSs. Additionally, two glycosyltransferases, BgtfA and BgtfB, were also enriched (~2 $\log_2$ fold change), indicating their proximity to the biosynthetic machinery. Although less prominent, the P450 monooxygenase OxyC and the NRPSs BpsA and BpsB were also enriched, (0.5-0.8 $\log_2$ fold changes), suggesting potential co-localization and interaction.However, no other biosynthesis enzymes were detected in proximity to BpsC. We attribute this to the protein size of BpsC and the spatial orientation of the TiD tag, which may exceed the ~10 nm distance required for effective biotinylation, potentially limiting the detection of nearby proteins that are not within this range.

Discussion

Our study provides valuable insights into the enzymatic interplay within the GPA biosynthetic machinery, focusing on the biosynthesis of the type I GPA balhimycin. A central theme of our research was understanding the interaction between NRPSs and the pivotal role of docking domains (DDs) in stabilizing these interactions. Our *in silico* predictions based on sequence homology suggested the presence of DDs in glycopeptide NRPSs, these findings provide the first experimental evidence, both *in vitro* and *in vivo,* of their functional role in GPA biosynthesis.

They were initially discovered in polyketide synthases (PKS), and since then undergone investigation in several megasynthases including NRPSs (Staunton & Weissman 2001; Smith et al., 2018). Previous studies on elements mediating interactions in hybrid assembly lines, such as proteins from the rapamycin, yersiniabactin, and epothilone pathways, highlighted the requirement for domains adjacent to DDs for functional interactions. These studies demonstrated that hybrid systems exhibited lower production rates compared to natural assembly lines (O'Connor et al., 2003). Likewise, studies on PaxA and PaxB, key proteins in the biosynthesis of "peptide-antimicrobial-Xenorhabdus" (PAX) in *Xenorhabdus bovienii SS-2004*, demonstrated that the interaction between these proteins requires cooperation between DDs and adjacent domains, such as the T-domain of PaxA binding to the nDD of PaxB (Watzel et al., 2021). Our results align with these observations, suggesting that DDs in balhimycin production facilitate NRPS assembly by stabilizing interfaces between BpsA-BpsB and BpsB-BpsC, enabling efficient substrate transfer between donor and acceptor modules. While DD deletion weakened the binding affinity, interactions persisted, likely due to residues in adjacent domains. This mechanism mirrors the PaxA/PaxB domain architecture, where interactions are facilitated by a PCP domain upstream of the cDD and a C-domain downstream of the nDD. Further structural studies are needed to fully elucidate these interactions in the context of balhimycin biosynthesis.

In addition to NRPS-NRPS interactions, our study highlights the importance of interactions between NRPSs and modification enzymes such as the P450 monooxygenases OxyA, OxyB and OxyC, and the halogenase BhaA. These interactions are crucial for balhimycin halogenation and cross-linking. Our experiments confirmed the interactions between OxyA and OxyC with the X-domain of BpsC, which is in line with previous studies that identified the X-domain as a unique recruiting interface for P450s on GPA NRPSs (Haslinger et al., 2015). However, OxyB did not interact with the X-domain in our heterologous *in vivo* system, suggesting that OxyB targets a different interface on NRPSs. Previous work aligns with this observation, showing that OxyB catalyzes the first C-O-D cross-link on hexapeptides prior to the incorporation of the seventh amino acid, (S)-3,5-dihydroxyphenylglycine. Examples, are shown in studies investigating the function and cascade of P450 monooxygenases (Bischoff et al., 2005, 2001; Evi Stegmann et al., 2006). The absence of OxyB affinity in our assays may have resulted from experimental limitations, such as the exclusion of PCP-bound hexapeptides, its likely natural substrate. Similar limitations might have contributed to the low-confidence models predicted by AlphaFold due to the inability to model the substrate bound to the respective domains.

Our investigation also highlighted the potential interaction of the halogenase BhaA with the NRPS machinery, essential for chlorination of PCP-bound substrates (Kittilä et al., 2017). BN-PAGE protein profiling revealed that BhaA co-migrated with NRPSs BpsA-C, and proximity-dependent biotinylation further confirmed its proximity to BpsC.

Additionally, our analysis of MLP_{bal} suggests its role in supporting A-domain functionality during amino acid activation and selection. We also demonstrated that genome encoded homologs could compensate for *mlp*$_{bal}$ deletion, which did not significantly reduce balhimycin production (Stegmann et al., 2006b). In support of this hypothesis, we identified interactions between balhimycin NRPSs and other MLPs encoded by distinct predicted BGCs, such as MLP_{7561} from a siderophore BGC and MLP_{7148} in an unknown NRP BGC. These results suggest potential crosstalk between otherwise unrelated BGCs. Interestingly, transcriptomic analysis revealed that under balhimycin production conditions only the genes encoding the interacting MLPs are transcribed, however MLP_{7148} in a lower rate compared to MLP_{bal} and $MLP_{7561,}$ but higher than MLP_{6685}, MLP_{9502} and MLP_{8030} (Figure S14). Similar cross-complementation has been observed in *Streptomyces scabiei* during thaxtomin A production, where the deletion of the MLP-encoding gene *txtH* reduced production, but a further deletion of non-cognate MLPs completely abolished production (Li et al., 2020). Likewise, MLPs from different BGCs in *Streptomyces coelicolor,* such as *cdaX* from the CDA BGC and *cchK* from the coelichelin BGC, have been shown to complement each other's NRPSs (Lautru et al., 2007). Despite high sequence similarity among MLPs, our *in silico* and mutagenesis analysis indicate that specific motifs dictate the interaction of MLP_{bal} and functional homologs with NRPSs. Attempts to universally exchange MLPs in NRPS systems have often failed to yield consistent results, highlighting the necessity for a detailed understanding of MLP-NRPS compatibility(Esquilín-Lebrón et al., 2018; Mori et al., 2018; Schomer and Thomas, 2017).

Overall, our study significantly advances the understanding of the balhimycin biosynthetic machinery by revealing various protein-protein interactions. This work underscores the critical importance of maintaining these interactions for efficient GPA biosynthesis. The insights gained here can guide future strategies to optimize enzyme activity in biotechnological applications, to generate novel GPAs or to enhance the efficiency of microbial pathways used in industrial drug production.

Material and Methods

Bacterial strains, cultivation conditions and molecular cloning

Cultivation and DNA manipulation of *Escherichia coli* NovaBlue were performed as previously described (Sambrook and Russell, 2001). For plasmid selection, the medium was supplemented with apramycin 100 µg/ml, ampicillin 150 µg/ml or kanamycin 50 µg/ml. *Amycolatopsis balhimycina* and all derived mutants were cultivated at 29°C, in baffled flasks with a steel coil, and shaken at 120 rpm in R5 medium (Kieser et al., 2000) or on solid agar (1.5% (w/v)) medium. Apramycin 50 µg/ml or erythromycin 50 µg/ml was used for plasmid selection.

Oligonucleotides used for mutant constructions and bacterial two-hybrid (B2H) analyses are listed in Tables S1. All used strains and plasmids are listed in Table S2 and S3, respectively.

For all polymerase chain reactions (PCRs), the "Q5 High-Fidelity DNA Polymerase (NEB©)" was used to amplify the genes with the overhangs for the respective target vector. To construct the plasmids for the B2H studies, pKT25/pNKT25 and pUT18c/pUT18 vectors were linearized with *BamH*I/*Kpn*I restriction endonucleases. All enzymes were from Thermofischer. For deletions in *A. balhimycina* the pSP1 vector was linearized using the restriction endonuclease *Kpn*I. DNA fragments were assembled by Gibson assembly (Gibson et al., 2009) using the NEBuilder HiFi DNA assembly Mix (NEB©). DNA gel extraction and plasmid isolation were

performed using the purification kits from Macherey-Nagel (#REF 740609.50) and VWR (#13-6945-00).

Genetic manipulation of *A. balhimycina*

Genetic manipulation of *A. balhimycina* in this study was achieved using a well-established direct transformation protocol (Pelzer et al. 1997). For this method, demethylated DNA was isolated from the *E. coli* strains JM110 and ET12567. For the analysis of the docking domains, in-frame deletion mutants of *A. balhimycina* were generated, in which the C-terminal docking domain encoding region was deleted using a homologous recombination approach. Deletion was performed using pSP1 (Pelzer et al. 1997) derivatives, containing flanking homologous regions. The flanking regions of approximately 1,5 kB up- and downstream of the DD region were cloned using primers P229-P232 (Table S1.) and the plasmid was integrated by direct transformation. Erythromycin-resistant clones harboring pSP1ΔcDD1, or pSP1ΔcDD2 on the chromosome, were confirmed by PCR and subsequently used in the stress protocol to induce a crossover event according to Puk et al. (2002). The generated protoplasts were streaked on R5-agar plates as well as R5-agar plates supplemented with 50 µg/ml erythromycin. Erythromycin sensitive clones were screened for in-frame deletion by PCR using primers P235-P236 and P248-P249 (Table S1).

Balhimycin production and HPLC-MS analysis

For the production assay, the generated mutants and the wildtype strain were inoculated in 20 ml R5-Medium and incubated for 48h as preculture as described above. After growth, 5 ml were taken to inoculate 100 ml of R5-Medium as the main culture.
After 120h of cultivation, 10 ml of cultures were harvested and separated by centrifugation into supernatant and mycelium. The supernatant was directly used for bioactivity assays as well as for HPLC-MS measurements. The mycelium was lyophilized and weighed for dry cell weight determination and subsequent balhimycin quantification.

Balhimycin detection was performed by analyzing 2.5 µl of the different supernatants by means of high-performance liquid chromatography-electrospray ionization-mass spectrometry (HPLC–ESI-MS) using a Nucleosil 100-C18 column (3 µm, 100 by 2 mm) (precolumn, 10 by 2 mm) (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) coupled to an ESI mass spectrometer. LC-MS measurements were obtained from a LC/MSD Ultra Trap system XCT 6330 (Agilent Technologies, Waldbronn, Germany). Analysis was carried out at a flow rate of 400 µl/min with gradient elution. Solvent A was 0.1% formic acid, and solvent B 0.06% formic acid in acetonitrile. Gradient elution was performed as follows: t0=0% B, t15=t17=100% B, post time 5 min. 0% B. The flow rate was 400 µl/min, and the temperature was 40°C. For MS analysis an electrospray ionization (alternating positive and negative ionization) in Ultra Scan mode with a capillary voltage of 3.5 kV and a drying gas temperature of 350°C was used. Detection of m/z values was conducted with Agilent DataAnalysis for 6300 series Ion Trap LC/MS 6.1 version 3.4 software (Bruker-Daltonik GmbH).

To quantify the balhimycin amounts produced by the mutants we used the obtained HPLC-MS data. For this purpose, the peak intensities of each data point in the MS chromatogram corresponding to the masses of balhimycin was summed to a total intensity, which was used to determine the concentration of balhimycin by comparing it to a calibration curve generated by measuring pure balhimycin. To compare the produced amounts, all production levels of balhimycin were normalized to the cell dry weight.

Bacterial two-hybrid interaction assays

Qualitative assay

For the bacterial two hybrid experiments, genes encoding for several balhimycin biosynthesis enzymes were cloned into the respective vectors as described above to generate translational fusions with the catalytic domains of the *B. pertussis* adenylate cyclase. All steps were based on the protocols described by (Karimova et al., 2017). To detect protein interactions, *E. coli* BTH101 was co-transformed with pKT25/pNKT25 and pUT18c/pUT18 derivatives and spotted on McConkey agar or LB-X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside 40 µg/ml) plates supplemented with maltose as sole carbon source, IPTG 0,5mM, and appropriate antibiotics. Plates were incubated at 30°C for 3-4 days. Each interaction was repeated three times. The ability of co-transformants to catabolize maltose resulting in red-colored colonies on MacConkey agar or blue-colored colonies on LB-X-gal plates, is based on functional reconstitution of the adenylate cyclase due to the interaction of the fusion proteins.

Quantitative assay

All interactions were quantified using a β-galactosidase assay, as described by (Karimova et al., 2017) and performed each time in triplicates. In brief, BTH101 co-transformants were inoculated in 5 ml LB- broth supplemented with 0.5 mM IPTG and the appropriate antibiotics and incubated overnight at 30°C. A 1:5 dilution with M63-medium was prepared and the optical density at 600 nm was measured on a microplate reader (Tecan Infinite M200 Pro). Two hundred microlitres of the diluted cells were permeabilized with 7µl of 0.1 SDS and 10µl of chloroform and incubated at room temperature for 20-30 minutes. Twenty microliters of the permeabilized cells were mixed with PM2 reaction buffer containing 100mM β-mercaptoethanol and 0,1% o-nitrophenyl-β-D-galactopyranosid (ONPG) as substrate for the β-Galactosidase and incubated at 30°C for 30-40 minutes. The reaction was stopped with 50µl of 1M Na2CO3 and the optical density was measured at 420 nm using the microplate reader. The values of each well were blanked, and the enzymatic activity was calculated using following formula Dilution*[1000*((OD420/OD600)/(Time)].

MicroScale Thermophoresis (MST)

To perform MST measurements, docking domain (DD) constructs with the respective predicted sequence and one adjacent domain were included in the construct design. For BpsA, the C-terminal PCP-domain was attached to the subsequent cDD1 region, resulting in a PCP^{BpsA}-cDD1 construct. The same was done for PCP^{BpsB}-cDD2, consisting of the C-terminal PCP-domain of BpsB and the respective cDD2. For the N-terminal DD constructs, the nDD sequences were expressed with the adjacent C-domains in BpsB and BpsC, resulting in nDD1-C^{BpsB} and nDD2-C^{BpsC} constructs. For control experiments, the N-terminal and C-terminal DD's were excluded, giving a construct containing the adjacent domain which is either a PCP- or C-domain. For all MST measurements, the RED – MALEIMIDE 2nd Generation Kit (NanoTemper) was used. The coupling of the fluorophore to the target protein was done as described in the manual appended to the Kit. All measurements were performed at room temperature using the Monolith NT.115 NanoTemper MST device with the instrument settings: MST power = medium and excitation power = 100% (auto-detect). MST buffer (25 mM HEPES/NaOH, 50 mM NaCl, pH 7.5) was used to dilute the target protein prior to the dilution series. Maleimide labelling buffer (NanoTemper) was used to perform the dilution series prior to the actual measurement. The unlabelled ligand, in this case the binding protein, was concentrated to at least 2 mM for all experiments. The concentration of the labelled target

protein was adjusted according to the labelling efficiency (200 – 400 nM). All samples were centrifuged at max. speed for 10 min prior use to sediment possible aggregates. For the measurement, Monolith NT.115 Premium Capillaries were used (NanoTemper) to avoid adsorption of the molecules to the inner surface of the capillaries. All measurements were done in triplicate. Data points were initially normalized and then fitted using a sigmoidal, 4PL fit (GraphPad).

Blue native PAGE separation and mass spectrometry

For the extraction of proteins from *A. balhimycina*, a protocol suitable for our strain was established. Twenty millilitres of R5 medium were used to inoculate pre-cultures with *A. balhiymcina* WT and *A. balhimycinaΔbpsC,* which were incubated for 48h. Five millilitres of the pre-culture were used to inoculate 100 ml R5 medium as a main culture and subsequently incubated for 35h. For further steps, the pellet was washed 1x with buffer K (50 mM triethanolamine (TEA), pH 7.5, 250 mM sucrose, 1 mM EDTA), resuspended in 20 ml buffer K supplemented with DNAseI 10 µg/ml, lysozym 500µg/ml, protease inhibitor cocktail (Sigma-Aldrich (P8849)) 1:100, and MgCl2 1 mM, and incubated at 37°C for 30 min. Afterwards the cells were homogenised using a French press and subsequently the cell debris were separated by centrifugation for 20 min at 20.000 rpm.

The cell lysate was further centrifuged for 45 min at 55,000 x *g* and 4°C, to precipitate crude membranes from the cytosolic fractions. Using the dounce homogenizer the crude membranes were resuspended in 1-2 ml 1x PBS buffer. The total protein concentration of the cytosolic and membrane fractions was measured using a Pierce™ BCA Protein Assay Kit (ThermoFischer) and a concentration of 4 mg/ml was used for further steps. The crude membranes were further solubilized using Lauryl Maltose Neopentyl Glycol (LMNG) by incubation for 1 h at 4°C under shaking at 500 rpm. Then, the samples were centrifuged for 30 min at 40,000 rpm, 4°C and the supernatant was taken for further analysis. For the subsequent separation of protein complexes from the cytosolic and membrane solubilized fractions a blue native polyacrylamide gel electrophoresis (BN-PAGE) was applied using a 4-16% Bis-Tris precast gel (Invitrogen). The gel was stained using standard Coomassie staining procedure and scanned using a Li-Cor Odyssey system. The gel was further cut in 4mm pieces along the separated lane using a razor blade, starting from the well on the top (Figure S7E). The gel slices were stored in 5% acetic and subsequently analyzed using HPLC-MS/MS.

The proteins derived from the gel slices after the blue native PAGE separation were analyzed by LC-MS/MS analysis on an Easy-nLC 1200 (Thermo Scientific) coupled to a QExactive HF mass spectrometer (Thermo Scientific); method: 60 min, Top07, HCD. The processing of data was performed using MaxQuant software (vs 1.6.7.0). The spectra were searched against an *Amycolatopsis balhimycina* database obtained from Uniprot (downloaded 11th of December 2019; 9,427 entries), and 286 commonly observed contaminants. The data were processed with a setting of 1% for the false discovery rate (FDR).

Data analysis

The acquired data from the mass spectrometry analysis were further processed using Perseus v. 2.0.9.0 (Tyanova et al., 2016). For each replicate, a matrix was generated and all the detected proteins were annotated. Subsequently, the data was transformed to Log2 values. To normalize the values, the scale to interval feature of Perseus was used. In this, the

intensities of the detected proteins are normalized to the total intensity of all proteins in the respective slice and the calculated value is represented on a scale, which in our analysis was set from 0-2.
For all analyzed slices a histogram was generated (Figure S8) to display data distribution. In addition, for each set of replicates a Pearson correlation was calculated to evaluate the reproducibility of the obtained data between the replicates.
The final values were plotted as a heatmap which shows the intensity profiles of the balhimycin biosynthetic enzymes in the analyzed gel slices along the lanes in the BN-PAGE separation. To correlate the interaction of the proteins, we compared the maxima of the intensities in the gel slices, meaning that proteins showing a high intensity in the same fraction, were considered interacting proteins.

Proximity labelling using TurboID
For the proximity labelling studies, *A. balhimycina* was cultivated in production conditions, as described in previous sections. Fifty millilitres of cell culture was harvested after 48h at 5,000 x g, 4°C for 10 min. The cell pellet was washed with 20 ml of 1xPBS buffer and centrifuged at 4,600xg, 4°C for 10 min. To start the labelling, 20 ml of 1xPBS supplemented with 500µM Biotin was added and the cells were incubated for 60 minutes at 37°C under shaking at 650 rpm. The reaction was stopped by placing the cells on ice and washing 3x with an ice-cold 1xPBS buffer. The cells were subsequently lysed by adding 15 mL lysis buffer W, containing protease inhibitor, 10 µg/ml DNase, lysozyme (200 µg/ml) and 1 mM MgCl2. After initial cell lysis, cells were further disrupted by sonication (Branson sonifier 250) and homogenization using a CF1 Homogenisator (I&L Biosystems). The obtained cell lysate was centrifuged at 10,000xg, 4°C for 3 min to remove cell debris. The supernatant was centrifuged again at 10,000xg, 4°C for 10 min. The protein concentration was measured using a BCA assay kit (ThermoFischer) and a total protein amount of 5 mg was used for the enrichment steps.

For the streptavidin-based affinity enrichment, 150 µl of MagStrep® Strep-Tactin® beads (IBA) were used and mixed with the previously mentioned protein amount. To bind the biotinylated proteins on the beads, the mixture was incubated overnight at 4°C using an overhead rotator. The beads were washed and eluted as described by the manufacturer. For subsequent mass spectrometry analysis, 20µl of the eluate were loaded on a 4-16% (Invitrogen) SDS-PAGE gel and run until the loaded proteins reached ca. 1cm of the gel. The gel was stored in 5% acetic acid until further processing.

To check for protein enrichment, immunodetection was performed using samples from the cell lysate, flowthrough and eluate. For this, samples were separated by SDS-PAGE and transferred to a PVDF membrane (Immun-Blot® PVDF Membrane 0.2 µm (Bio-Rad)). The membrane was probed with α-TurboID (rabbit, 1:10,000), Streptavidin DyLigth 800 (1:10,000) and α-rabbit 800 (1:10,000) antibodies. Scanning of membranes and analysis was performed using a Li-Cor Odyssey system and image Studio (Li-Cor).

Proteins were then digested in the gel with trypsin under shaking conditions: for destaining, gel pieces were washed three times with 5 mM ammonium bicarbonate (ABC) in acetonitrile (ACN) (1:1, v/v) for 20 min. After a dehydration step with 100% ACN for 10 min, disulfide bonds were reduced by adding 10 mM dithiothreitol (DTT) in 20 mM ABC for 45 min at 56°C. Thiol groups were carbamidomethylated with 55 mM iodoacetamide (IAA) in 20 mM ABC for 45 min in the dark. Subsequently, gel pieces were washed two times with 5 mM ABC in ACN (1:1, v/v)

for 20 min and dehydrated with 100% ACN for 15 min. Liquid was evaporated by vacuum centrifugation for 10 min and sequencing grade trypsin (Promega) was added in a concentration of 12.5 ng/µl in 20 mM ABC, pH 8.0. Gel pieces were soaked for 10 min at room temperature (RT), then covered with 20 mM ABC, and proteins were digested at 37°C overnight. Peptide extraction was performed in three consecutive steps with different extraction buffers for 30 min: first 3% (v/v) trifluoroacetic acid (TFA) in 30% (v/v) ACN was added, followed by 0.5% (v/v) formic acid (FA) in 80% (v/v) ACN, and completed by 100% ACN. Supernatants were pooled and ACN was evaporated by vacuum centrifugation. Peptides were further purified with C18 StageTips and analyzed on an EASY-nLC 1200 UHPLC coupled to a Q Exactive HF mass spectrometer (both Thermo Scientific) as described previously90 with slight modification: peptides were eluted from the analytical column using a 46 min segmented gradient of 10-33-50% of HPLC solvent B (80% acetonitrile in 0.1% formic acid) at a flow rate of 200 nl/min. In the mass spectrometer, MS and MS/MS spectra were acquired at resolution 60k. Full MS target value and maximum IT were set to 3x106 and 25 ms, respectively. In each scan cycle, the 7 most intense precursor ions were picked up, whereby MS/MS target value was set to 105 charges with a maximum IT of 110 ms.

MS data were processed using default parameters of the MaxQuant software (v2.4.12.0). Extracted peak lists were submitted to database search using the Andromeda search engine to query a target-decoy database of *A. balhimycina* (9,427 entries, downloaded on 30th of January 2024) and 286 commonly observed contaminants.

In database search, full tryptic specificity was required and up to two missed cleavages were allowed. Carbamidomethylation of cysteine was set as fixed modification, whereas protein N-terminal acetylation, and oxidation of methionine were set as variable modifications. For the main search, the peptide mass tolerance was set to 4.5 ppm, and 20 ppm at the fragment ion level. Peptide, protein, and modification site identifications were filtered at a false discovery rate (FDR) of 0.01. The iBAQ (Intensity Based Absolute Quantification) and LFQ (Label-Free Quantification) algorithms were enabled, as was the “match between runs” option.

Downstream statistical analysis was performed using iBAQ values and processed with Perseus. The intensities were transformed as Log2 and filtered using two valid numbers per replicate. The data were then normalized using the Z-score feature of Persus. Missing values of replicates were imputed from the normal distribution. Multiple t-test were performed for each protein and data were plotted as volcano plot using an false discovery rate (FDR) of calculation deriving in p-values (p) and foldchange (FC). Values with a $p<0.05$ and $FC >0.5$ were considered significant.

Prediction of protein complexes using AlphaFold3

For prediction of protein complexes the latest version of AlphaFold using the AlphaFold3 online server was used. The models were further analyzed using Alphabridge(Álvarez-Salmoral et al., 2024) and PISA (Krissinel and Henrick, 2007) to determine the putative interaction interface. For all visualization of the models Chimera X (Meng et al., 2023) was used.

Supplementary information

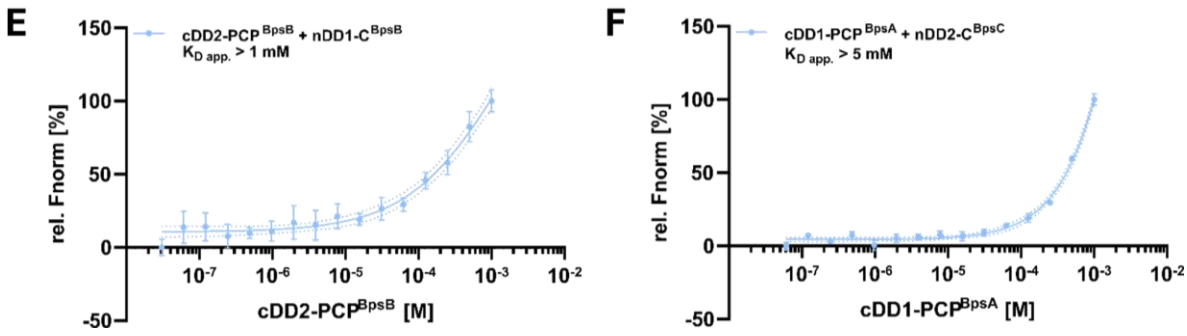


Figure S1: MST measurements of cross-interaction between non-pair DDs.

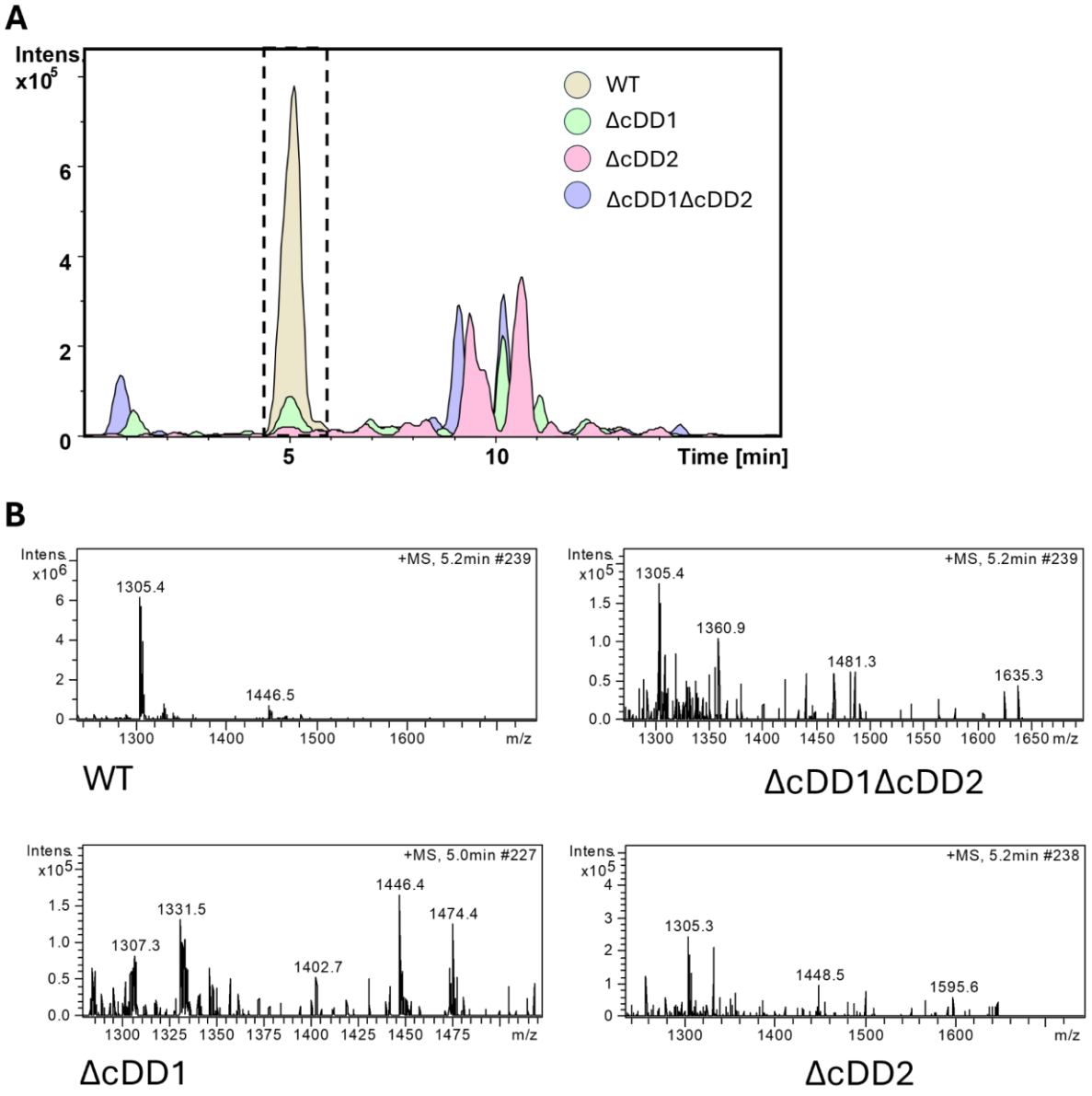


Figure S2: HPLC-MS analysis of the *A. balhimycina* DD mutants. (A) Chromatograms showing the peak for balhimycin detection (RT) and (B) the corresponding MS spectra of each chromatogram displaying the masses of balhimycin.

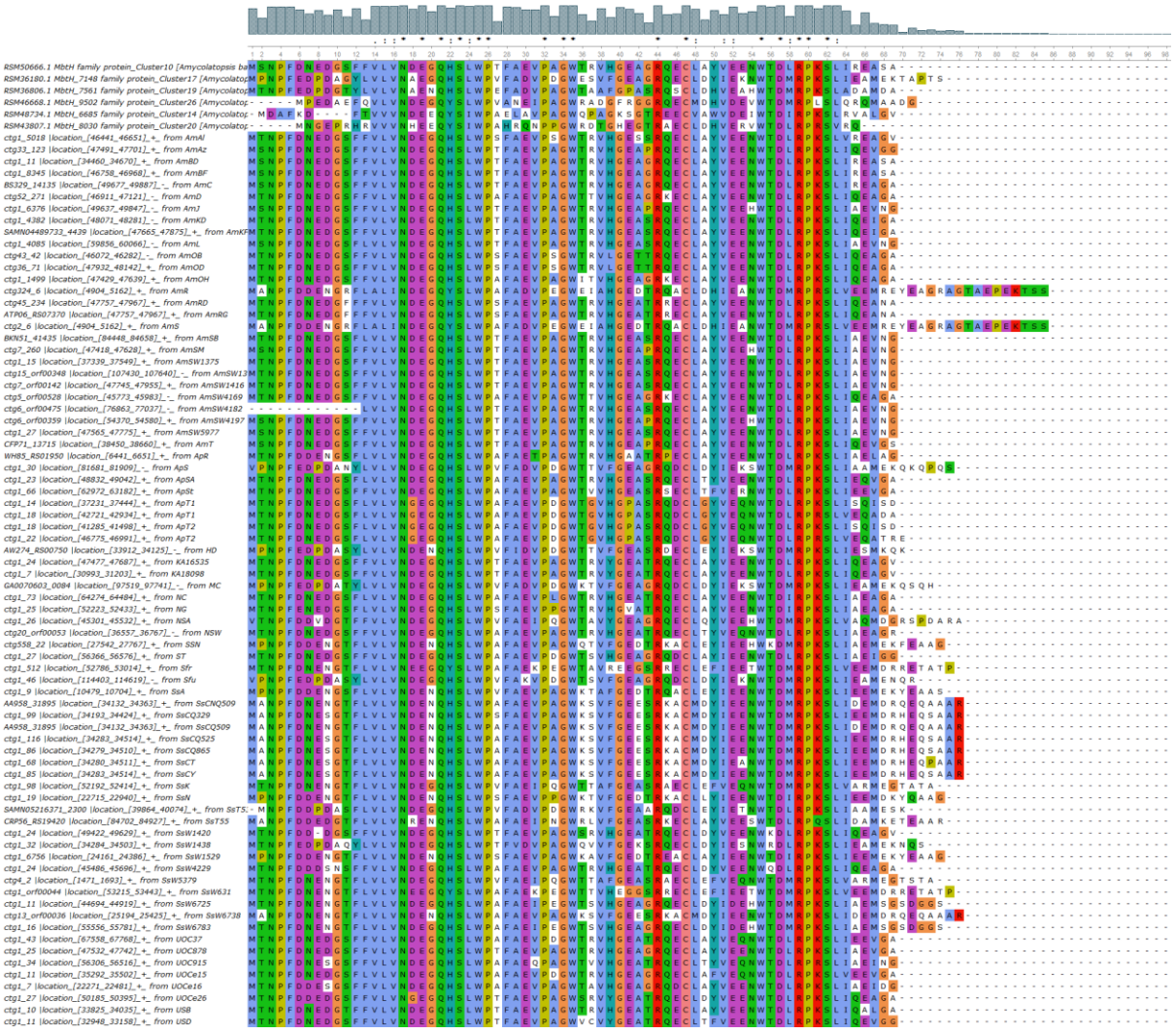


Figure S3: Multiple sequence alignment (MSA) of all MLPs encoded by *A. balhimycina* as well as all GPA BGC encoded MLPs acquired from publicly available sequences. The MSA was conducted using the Muscle algorithm and visualized using uGene.

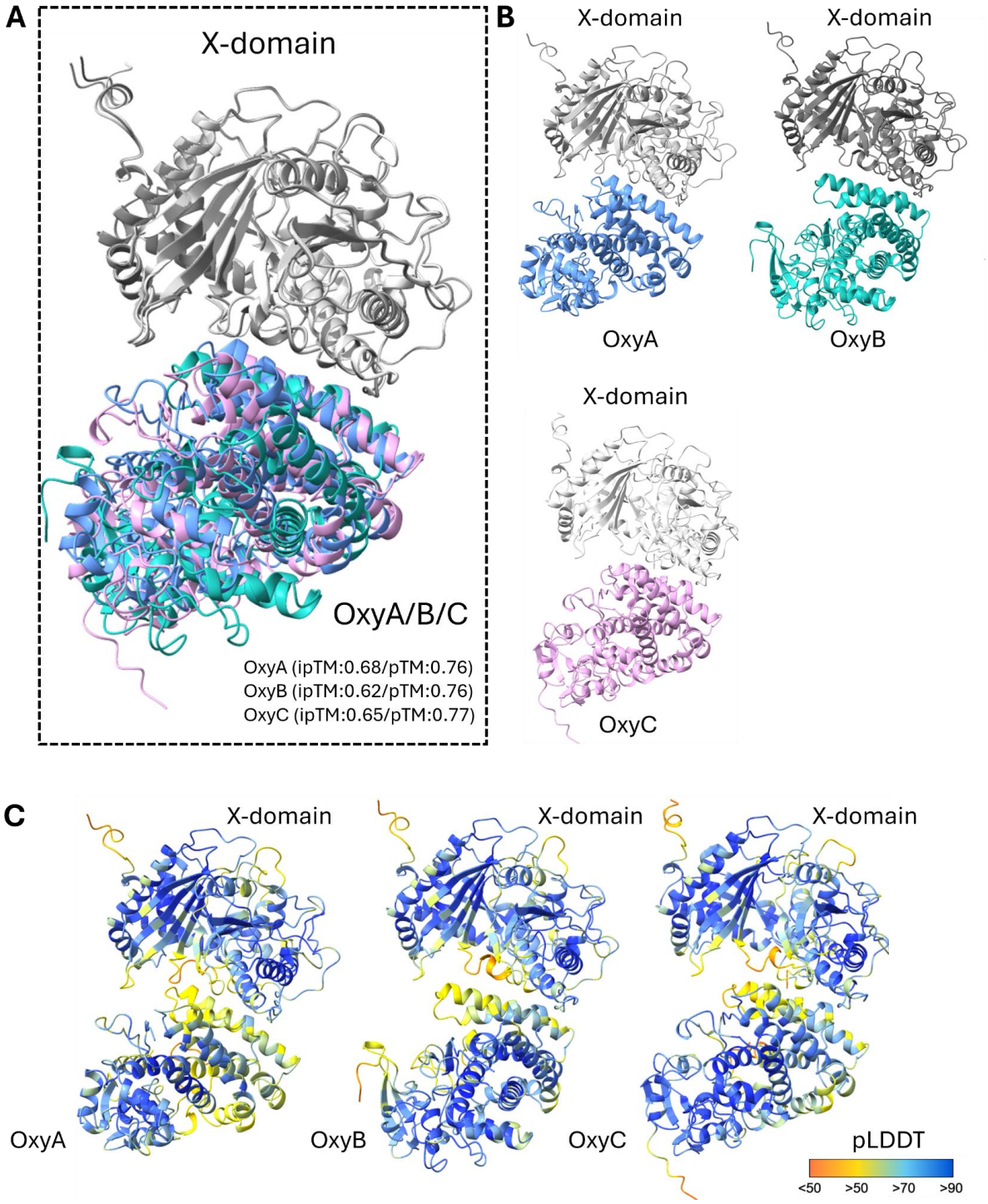


Figure S4: Predicted AlphaFold3 models of X-domain and OxyA-C interactions. (A) Superimposed models of all three comxplexes. The X-domain is shown in grey, OxyA in blue, OxyB in cyan and OxyC in light pink. ipTM and pTM scores are shown as well.. (B) All three models separately, in the same orientation. (C) Respective models with colored chains based on pLDDT sore.

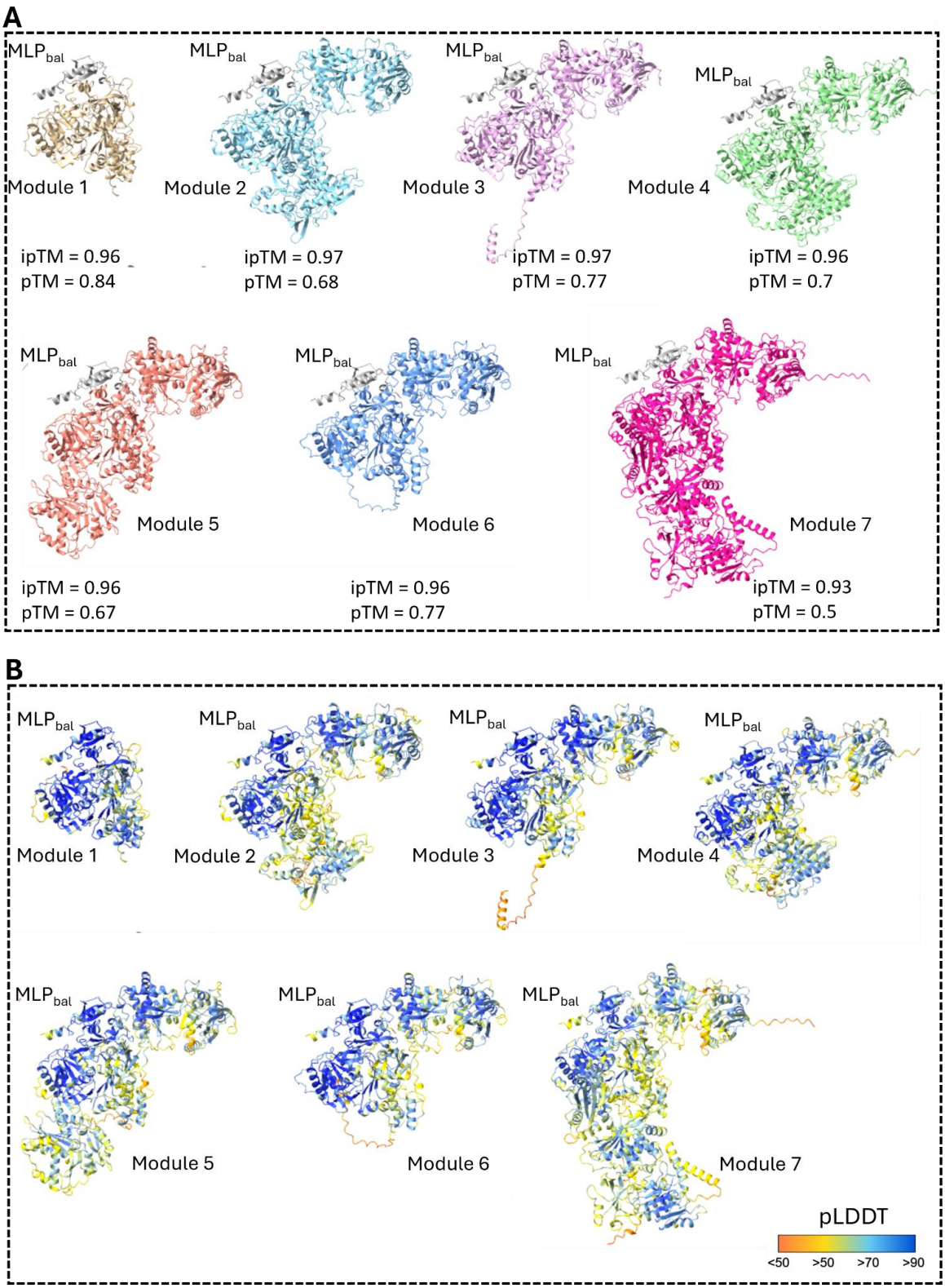


Figure S5: Predicted AlphaFold3 models of Modules 1-7 and MLP_{bal} interactions and confidence scores. (A) Models of all comxplexes between M1-7 and MLPbal and the respective pTM and ipTM scores shown below. **(B)** Respective models in the same orientation with colored chains based on pLDDT sore.

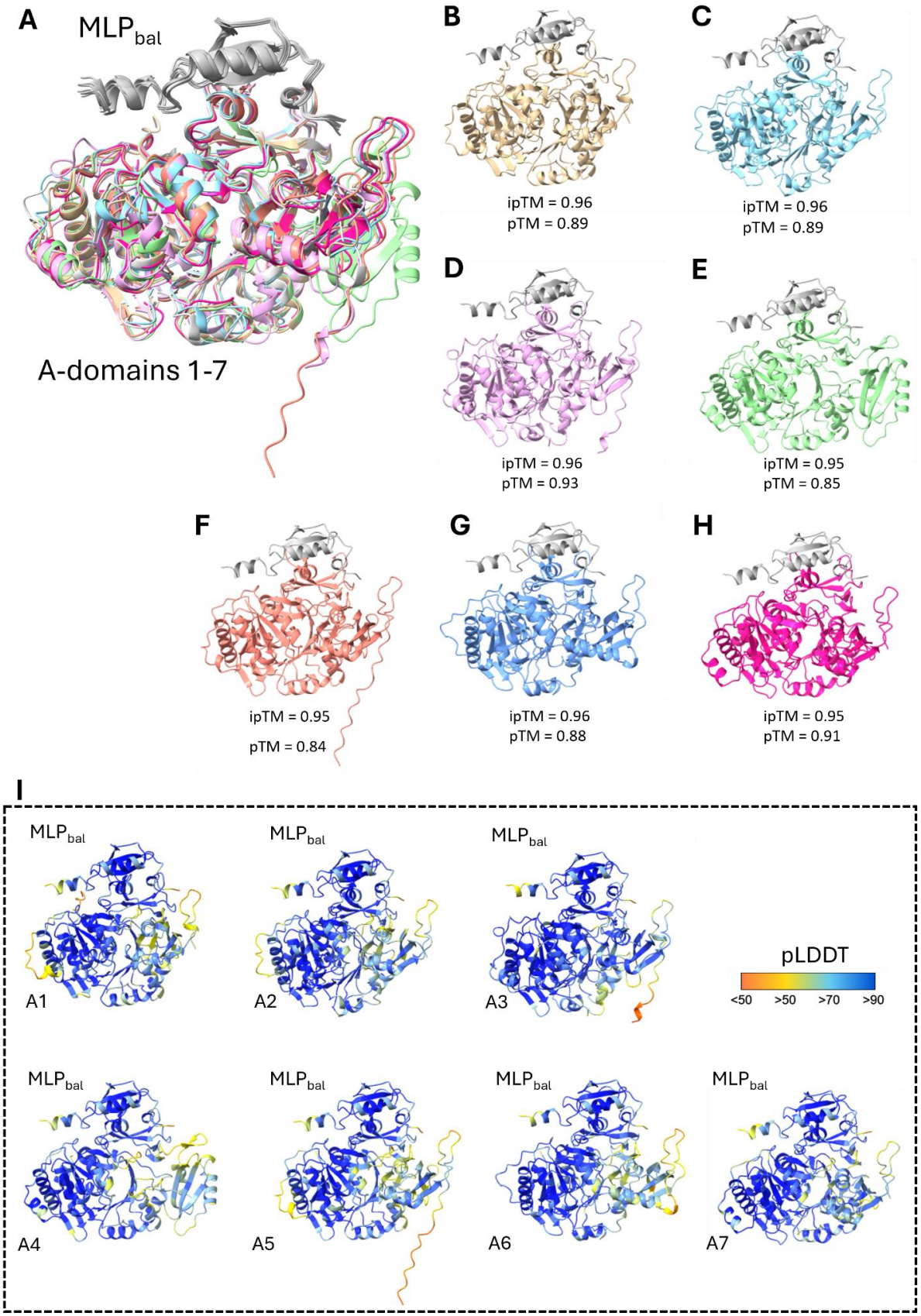


Figure S6: Predicetd AlphaFold3 models of A-domains (from M1-7) and MLP_{bal} interactions. (A) Superimposed models of all predicted complexes of A-domains and MLPbal. **(B-H)** All seven models separately with the respective ipTM and pTM score. **(I)** Respective models in the same orientation with colored chains based on pLDDT sore.

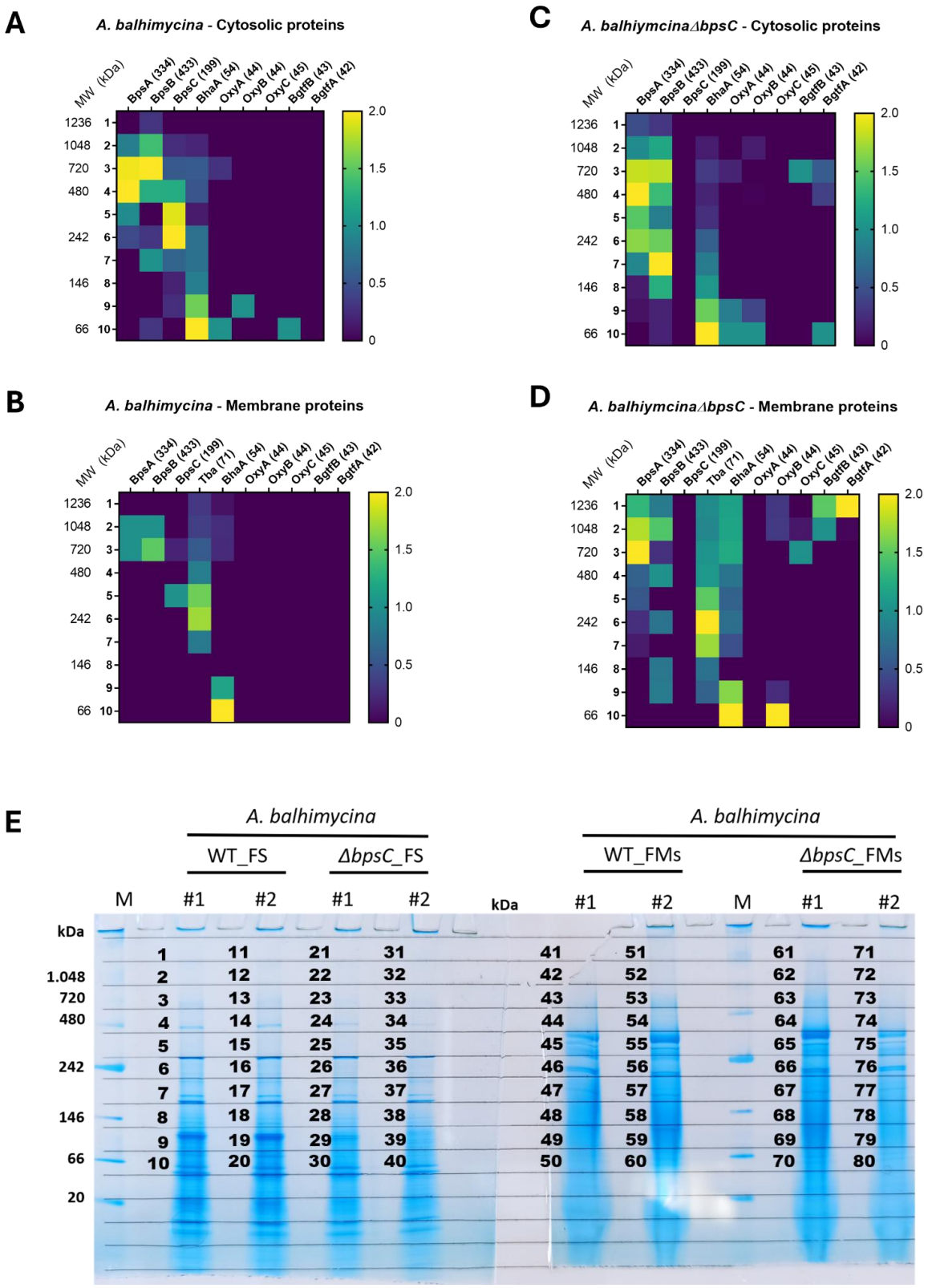


Figure S7: Heatmaps of identified proteins after BN-PAGE and HPLC-MS/MS analysis. The heatmaps shows the normalized signals (0-2) of each protein over the lane. (A) The proteins identified in the cytosolic fractions of *A. balhimycina* wildtype and (B) in the membrane fractions of *A. balhimycina* wildtype. (C) Shows the cytosolic proteins from *A. balhiymcina ΔbpsC* and (D) shows the membrane proteins of *A. balhimycina ΔbpsC*. **(E)** Shows the BN-PAGE gels after commassie staining with the respective cut raster used for sampling. FS is the cytosolic protein fraction, FMs is the membrane solubilized fraction.

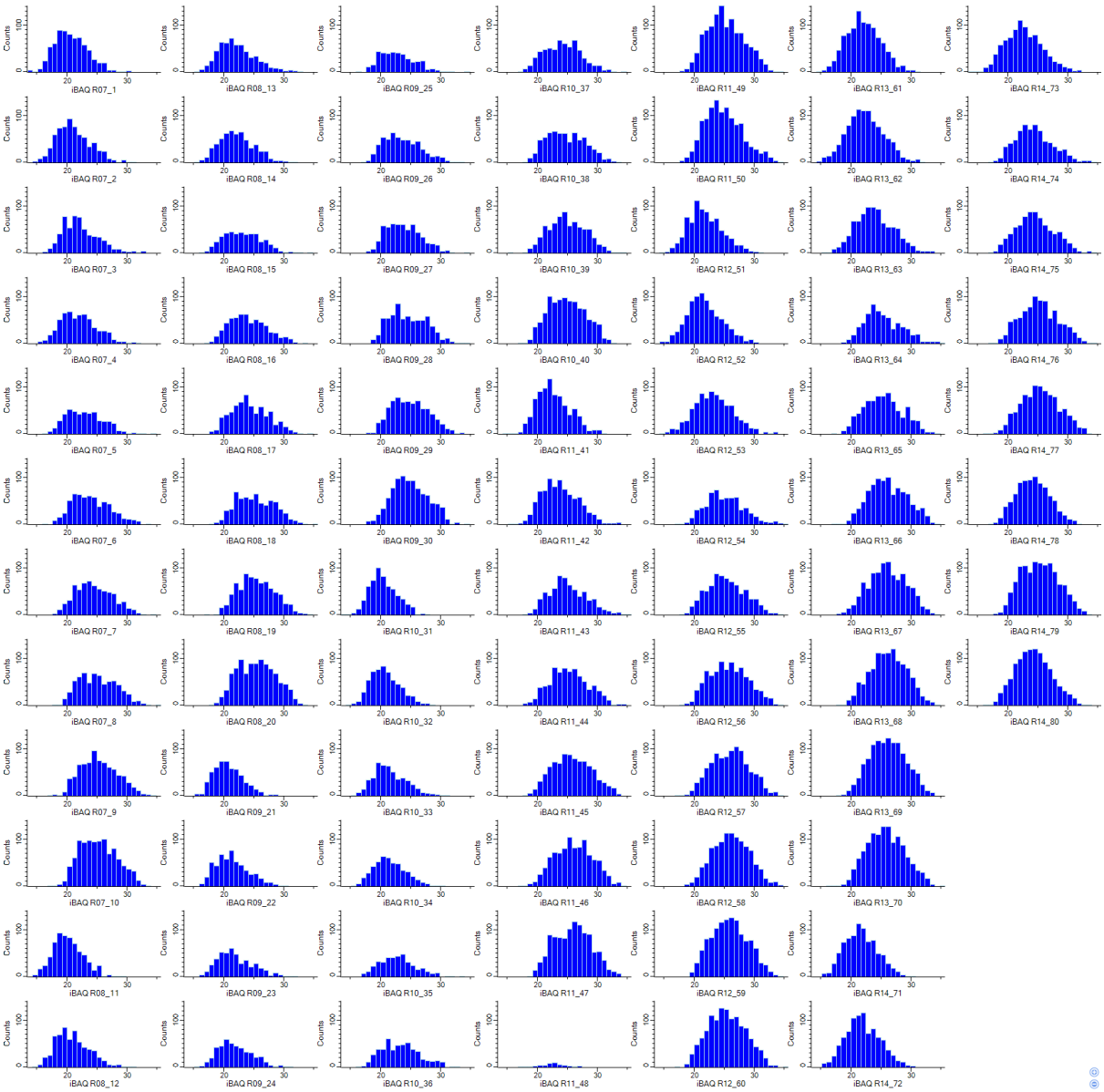


Figure S8. Normal distribution of protein intensities after HPLC-MS/MS detection. The plots show the normal distribution of each sample (gel slice) from cytosolic and membrane fraction of *A. balhimycina* wildtype and *A. balhimycina ΔbpsC, respectively*. Plotted are transformed and normalized data. The intensities represent the iBAQ values.

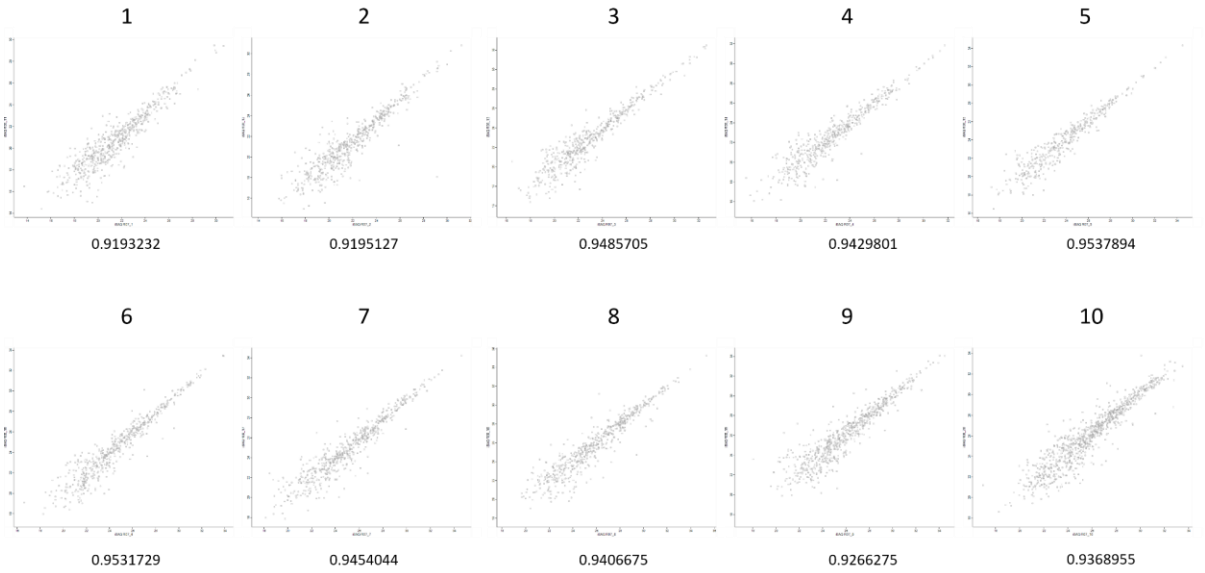


Figure S9: Pearson correlation of each replicate from *A. balhimycina* wildtype of the cytosolic fractions. The correlation coefficient (0-1) is displayed below each scatter plott. The data on the scatter plots represents values of each detected protein in the replicates. The higher the coefficient, the higher the confidence of reproducibility.

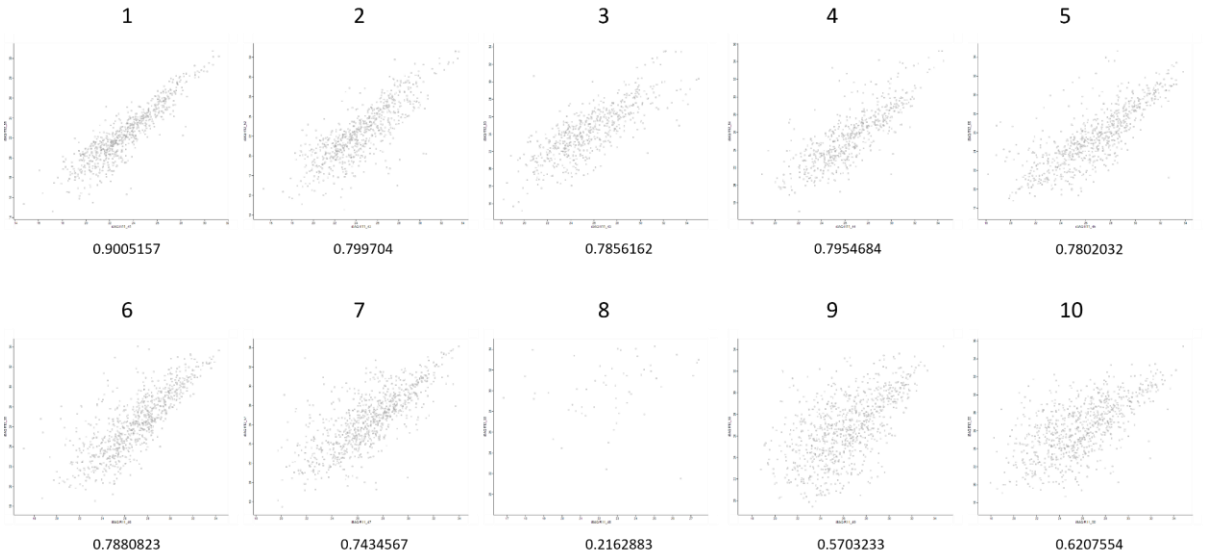


Figure S10: Pearson correlation of each replicate from *A. balhimycina* wildtype of the membrane fractions. The correlation coefficient (0-1) is displayed below each scatter plott. The data on the scatter plots represents values of each detected protein in the replicates.The higher the coefficient, the higher the confidence of reproducibility.

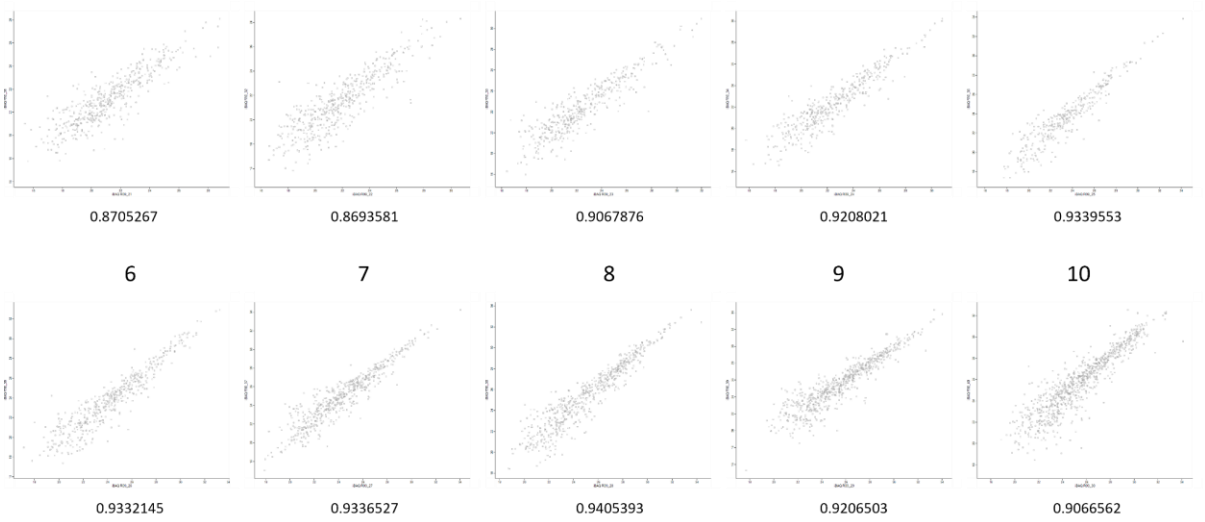


Figure S11: Pearson correlation of each replicate from *A. balhimycina ΔbpsC* of the cytosolic fractions. The correlation coefficient (0-1) is displayed below each scatter plott. The data on the scatter plots represents values of each detected protein in the replicates.The higher the coefficient, the higher the confidence of reproducibility.

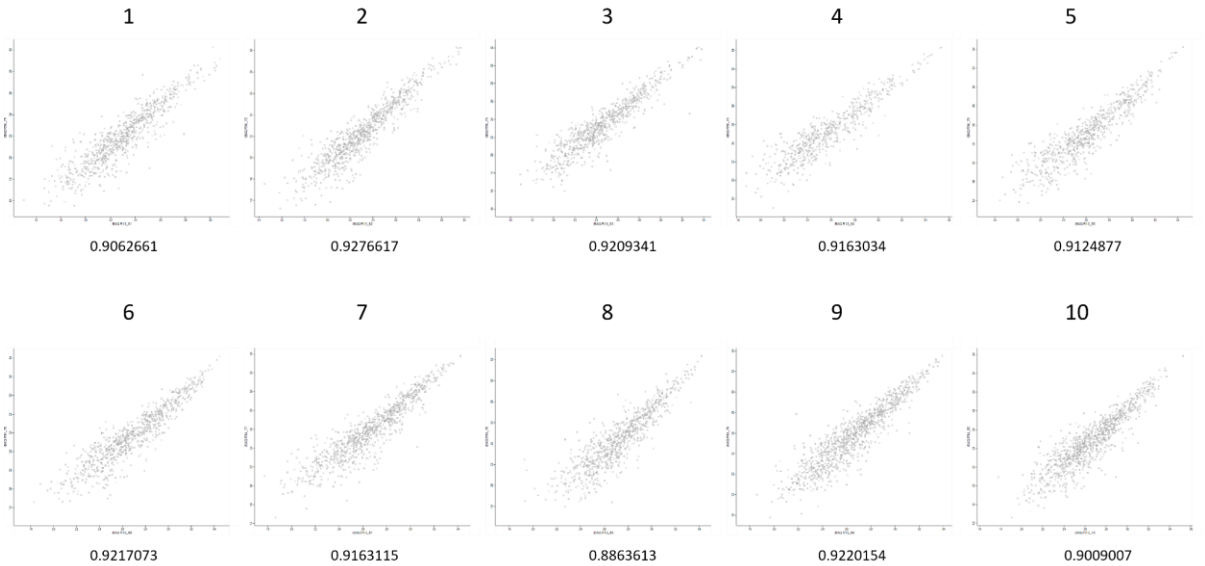


Figure S12: Pearson correlation of each replicate from *A. balhimycina ΔbpsC* of the membrane fractions. The correlation coefficient (0-1) is displayed below each scatter plott. The data on the scatter plots represents values of each detected protein in the replicates.The higher the coefficient, the higher the confidence of reproducibility.

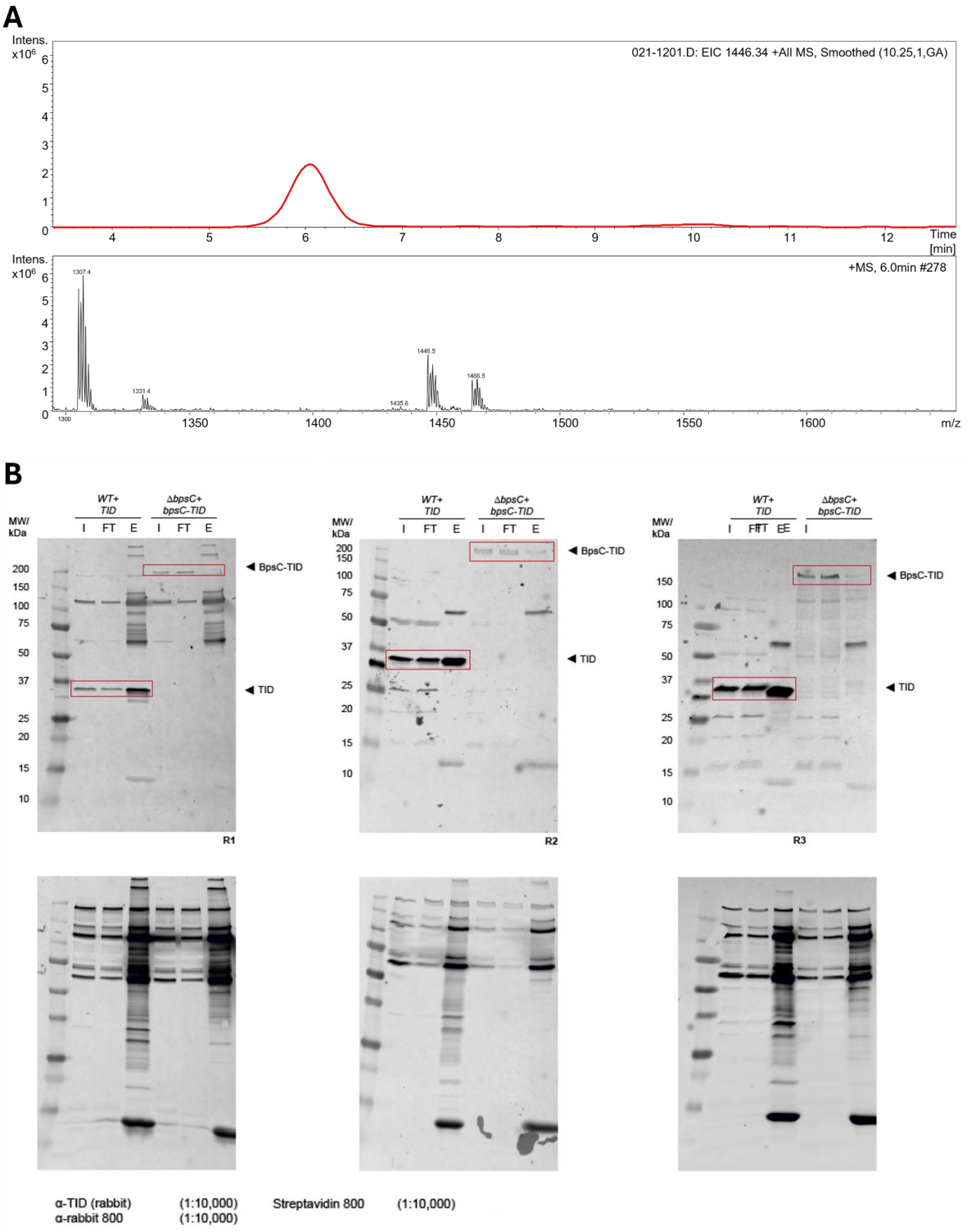


Figure S13: Proximity dependent biotinylation. (A) HPLC-MS chromatograms of the culture supernatants of *A. balhimycina ΔbpsC* complemented *with bpsC-TID*. The peaks (retention time 6 min) represent the extracted ion chromatogram (EIC) of the protonated balhimycin mass m/z 1446.41 $[M+H]^+$ (positive mode, smoothed 10.25 GA). **(B)** Immunoblotting (top) and fluorescence labeling with Streptavidin 800 (bottom) after SDS-PAGE of input (I) biotinylated proteins after enrichment (E) and flowthrough after enrichment (FT) from the recombinant strains expressing TID fusion proteins.

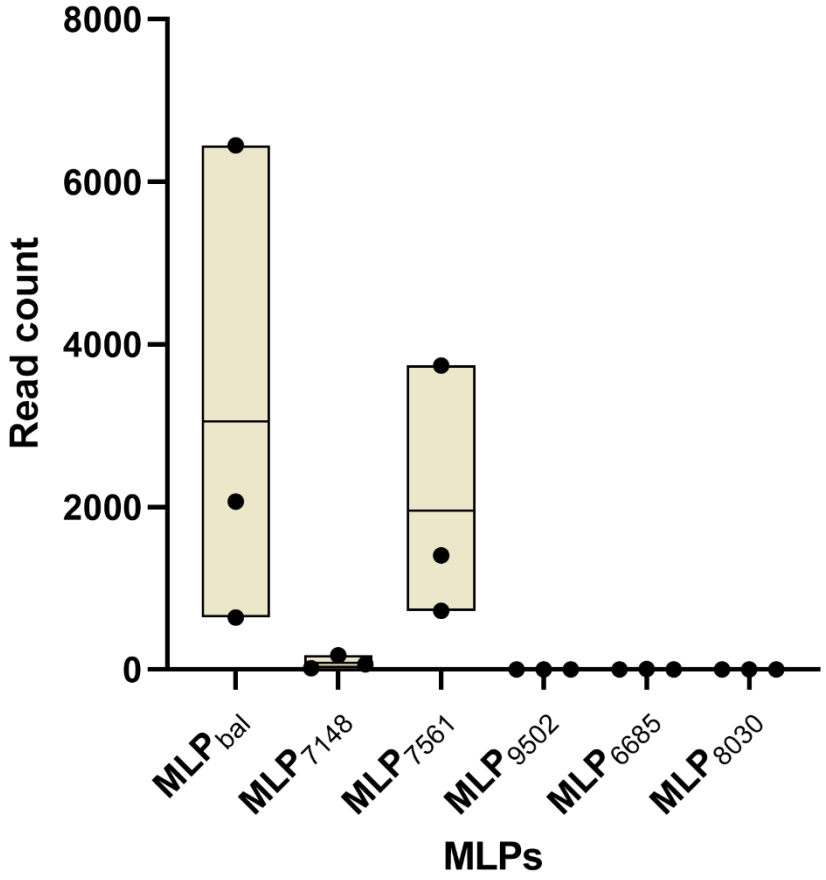


Figure S14: Transcriptomic analysis of *A. balhimycina* wildtype. Shown are the read counts of each mlp-encoding gene after 73h of cultivation in balhimycin production conditions. Data are represented in box plots (line indicating median), dots represent values of each replicate.

Table S1: Primers used in this study

ID	Primer name	Sequence
P1	T25_M1-bpsA_f	CGCGCACGCGGCGGGCTGCAGGGTCGACTCTAGAGatgaattccgcagcgcg
P2	T25_M1-bpsA_r	TAAAACGACGGCCGAATTCTTAGTTACTTAGGTACcgccacggcggcca
P3	T25_M2-bpsA_f	CGCGCACGCGGCGGGCTGCAGGGTCGACTCTAGAGgaggagatctggccgctg
P4	T25_M2-bpsA_r	TAAAACGACGGCCGAATTCTTAGTTACTTAGGTACggccacgtccagccgt
P5	T25_M3-bpsA_f	CGCGCACGCGGCGGGCTGCAGGGTCGACTCTAGAGctggccgaggtgctgc
P6	T25_M3-bpsA_r	TAAAACGACGGCCGAATTCTTAGTTACTTAGGTACtcaccggccgctttcgaa
P7	T25_M4-bpsB_f	CGCGCACGCGGCGGGCTGCAGGGTCGACTCTAGAGatcgaggaaatctggccgc

P8	T25_M4-bpsB_r	TAAAACGACGGCCGAATTCTTAGTTACTTAGGTACcagttcgagacggaccgg
P9	T25_M5-bpsB_f	CGCGCACGCGGCGGGCTGCAGGGTCGACTCTAGAGctgacggacatctggccg
P10	T25_M5-bpsB_r	TAAAACGACGGCCGAATTCTTAGTTACTTAGGTACgacgtccagccggagcg
P11	T25_M6-bpsB_f	CCTGGCGCGCACGCGGCGGGCTGCAGGGTCGACTCTAGAGctcgacatctggccgct
P12	T25_M6-bpsB_r	CGTTGTAAAACGACGGCCGAATTCTTAGTTACTTAGGTACttcctggacgatcaccgcc
P13	T25_M7-bpsC_f	CGCGCACGCGGCGGGCTGCAGGGTCGACTCTAGAGgtgaccgttgacgacactcg
P14	T25_M7-bpsC_r	TAAAACGACGGCCGAATTCTTAGTTACTTAGGTACtcatcgttgtgtcctttcaccct
P15	T18_bha_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGatgtcggtcgaagacttcgac
P16	T18_bha_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACtcacgcagggtggtgagg
P17	T18_bgtfA_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGatgcgcgtgttgatctcgg
P18	T18_bgtfA_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACtcatgcgggaacggcg
P19	T18_bgtfB_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGgtgaagcgtgtgctgttgt
P20	T18_bgtfB_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACttacgcgggaacagccgg
P21	T18_bgtfC_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGatgcgtgtgttgctgtcgac
P22	T18_bgtfC_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACtcacgcgggaacggtcag
P23	T18_bmt_f	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACtcaccgcgcggctg
P24	T18_bmt_r	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGgtgagcggtcaactcgagc
P25	T18_orf1_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGatgagcaatccgttcgacaacg
P26	T18_orf1_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACtcaggcgctcgcttccc
P27	T18_orf7_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGatgctgcacacctttgcc
P28	T18_orf7_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACctagagactttgctcctgcgg
P29	T18_oxyA_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGgtgttcgaggagagcaacgc
P30	T18_pxyA_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACtcaccactccagcagcagg

P31	T18_oxyB_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGttgaatgatgacgacccgcggcc
P32	T18_oxyB_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACtcaccaggcgaccatcagc
P33	T18_oxyC_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGatggggcacgatatcggt
P34	T18_oxyC_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACctaccaggtgaccggaaccc
P35	T18c_M4-bpsB_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGatcgaggaaatctggccgc
P36	T18c_M4-bpsB_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACcagttcgagacggaccgg
P37	T18_tba_f	GCGCAGTGGAACGCCACTGCAGGTCGACTCTAGAGatggacatggtgttgcgttt
P38	T18_tba_r	TTACTTAGTTATATCGATGAATTCGAGCTCGGTACtcatcctccgtagcccatgtg
P39	T25_bgtfA_f	CAGGGTCGACTCTAGAGatgcgcgtgttgatctcgg
P40	T25_bgtfA_r	CGAATTCTTAGTTACTTAGGTACtcatgcgggaacggcg
P41	T25_bgtfB_f	CAGGGTCGACTCTAGAGgtgaagcgtgtgctgttgt
P42	T25_bgtfB_r	CGAATTCTTAGTTACTTAGGTACttacgcgggaacagccgg
P43	T25_bgtfC_f	CAGGGTCGACTCTAGAGatgcgtgtgttgctgtcgac
P44	T25_bgtfC_r	CGAATTCTTAGTTACTTAGGTACtcacgcgggaacggtcag
P45	T18n_M4_f	GCAGGTCGACTCTAGAGatgagccagtcgcgga
P46	T18n_M4_r	GAATTCGAGCTCGGTACcagttcgagacggaccgg
P47	T18c_M7_f	GCAGGTCGACTCTAGAGgtgaccgttgacgacactcg
P48	T18c_M7_r	GAATTCGAGCTCGGTACtcatcgttgtgtcctttcaccct
P49	T25n_nDDbpsC_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGatgaccgtggacgataccc
P50	T25n_nDDbpsC_r	CTGCATGGTCATTGAATTCGAGCTCGGTACgtccagacgaacacgca
P51	T25n_nDDbpsB_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGatgagccagagccgtatcg
P52	T25n_nDDbpsB_r	CTGCATGGTCATTGAATTCGAGCTCGGTACcagacgcacacgaacc
P53	T18c_cDDbpsA_f	GTGGAACGCCACTGCAGGTCGACTCTAGAGgagcgtctgctgtgc
P54	T18c_cDDbpsA_r	TAGTTATATCGATGAATTCGAGCTCGGTACttaacggccgctttcaaactcgct
P55	T18c_cDDbpsB_f	GTGGAACGCCACTGCAGGTCGACTCTAGAGgaagcgcgtctgtgc
P56	T18c_cDDbpsB_r	TAGTTATATCGATGAATTCGAGCTCGGTACttattcacccgccagacccgcc
P57	T+Xdom_f	GCGGGCTGCAGGGTCGACTCTAGAGgagaaggtgctgtgcgcgct
P58	T+Xdom_r	ATTCTTAGTTACTTAGGTACCCGGGcgagacgtcggtgcgc

P59	T25_C-M6_f	GCGGGCTGCAGGGTCGACTCTAGAGctcgacatctggccgctt
P60	T25_C-M6_r	GCCGAATTCTTAGTTACTTAGGTACgaggcgcgcccgga
P61	T25_A-M6_f	GCGGGCTGCAGGGTCGACTCTAGAGttccgcaggcaggcggagcggtca
P62	T25_A-M6_r	GCCGAATTCTTAGTTACTTAGGTACccgcaccttgacctgcg
P63	T25_T+DD-M6_f	GCGGGCTGCAGGGTCGACTCTAGAGgaggcgaggctgtgcgc
P64	T25_T+DD-M6_r	GCCGAATTCTTAGTTACTTAGGTACttcgccggccagccc
P65	T25-DD+C_M7_f	GCGGGCTGCAGGGTCGACTCTAGAGgtgaccgttgacgacactcg
P66	T25-DD+C_M7_r	GCCGAATTCTTAGTTACTTAGGTACgtcgagccggaccc
P67	T25-A-M7_f	GCGGGCTGCAGGGTCGACTCTAGAGttcgcccggcgggtggccgc
P68	T25-A-M7_r	GCCGAATTCTTAGTTACTTAGGTACgcggatcttcacctggt
P69	T25n_A_M2_bpsA_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGcggatggcggcgc
P70	T25n_A_M2_bpsA_r	CTGCATGGTCATTGAATTCGAGCTCGGTACgcggatcttgactggt
P71	T25n_C_M2_bpsA_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGgaggagatctggccgctgg
P72	T25n_C_M2_bpsA_r	CTGCATGGTCATTGAATTCGAGCTCGGTACgtccagcggcacccggac
P73	T25n_E_M2_bpsA_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGgtcggtgacgtcgcgtgga
P74	T25n_E_M2_bpsA_r	CTGCATGGTCATTGAATTCGAGCTCGGTACggccacgtccagccgtac
P75	T25n_T_M2_bpsA_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGgagcgggtgctgtgcgagt
P76	T25n_T_M2_bpsA_r	CTGCATGGTCATTGAATTCGAGCTCGGTACcgactccgccgccagctgc
P77	T25n_A_M6_bpsB_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGttccgcaggcaggcgg
P78	T25n_A_M6_bpsB_r	CTGCATGGTCATTGAATTCGAGCTCGGTACccgcaccttgacctgcg
P79	T25n_C_M6_bpsB_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGctcgacatctggccgctt
P80	T25n_C_M6_bpsB_r	CTGCATGGTCATTGAATTCGAGCTCGGTACgaggcgcgcccgga
P81	T25n_T_M6_bpsB_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGgaggcgaggctgtgcg
P82	T25n_T_M6_bpsB_r	CTGCATGGTCATTGAATTCGAGCTCGGTACttcctggacgatcaccgccagccgtt
P83	T25n_A_M7_bpsC_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGttcgcccggcgggt
P84	T25n_A_M7_bpsC_r	CTGCATGGTCATTGAATTCGAGCTCGGTACgcggatcttcacctggtcgtcg

P85	T25n_C_M7_bpsC_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGgtcgaggacgtctggcctctttcg
P86	T25n_C_M7_bpsC_r	CTGCATGGTCATTGAATTCGAGCTCGGTACgtcgagccggaccc
P87	T25n_T+X_M7_bpsC_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGgagaaggtgctgtgcgcgct
P88	T25n_T+X_M7_bpsC_r	CTGCATGGTCATTGAATTCGAGCTCGGTACcgagacgtcggtgcgca
P89	T25n_TE_M7_bpsC_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGccgcccctgttctgcgtcc
P90	T25n_TE_M7_bpsC_r	CTGCATGGTCATTGAATTCGAGCTCGGTACcagcgccgcggattg
P91	T18n_bha_f	GCATGCCTGCAGGTCGACTCTAGAGggagatgtcggtcgaagact
P92	T18n_bha_r	TGGCGGCTGAATTCGAGCTCGGTACtcacgcagggtggtgaggca
P93	T18n_tba_f	GCATGCCTGCAGGTCGACTCTAGAGgatggacatggtgttgcgtt
P94	T18n_tba_r	TGGCGGCTGAATTCGAGCTCGGTACtcatcctccgtagcccatgtgttggatcaccg
P95	T18n_bgtfB_f	GCATGCCTGCAGGTCGACTCTAGAGgtgccgggaaccggcc
P96	T18n_bgtfB_r	TGGCGGCTGAATTCGAGCTCGGTACttacgcgggaacagccggctt
P97	T18n_oxyB_f	GCATGCCTGCAGGTCGACTCTAGAGgttgaatgatgacgacccgcggc
P98	T18n_oxyB_r	TGGCGGCTGAATTCGAGCTCGGTACtcaccaggcgaccatcagct
P99	T18c_mbtH_7148_f	ACGCCACTGCAGGTCGACTCTAGAGatgccgaatcctttcgaggac
P100	T18c_mbtH_7148_r	ATATCGATGAATTCGAGCTCGGTACtcatgacgtcggcgctg
P101	T18c_mbtH_7561_f	ACGCCACTGCAGGTCGACTCTAGAGatgaccaacccgttcgaagac
P102	T18c_mbtH_7561_r	ATATCGATGAATTCGAGCTCGGTACtcaggcgtccatggcgt
P103	T18c_mbtH_6685_f	ACGCCACTGCAGGTCGACTCTAGAGgtggacgcgttcaaggact
P104	T18c_mbtH_6685_r	ATATCGATGAATTCGAGCTCGGTACtcagaccccgagcgcg
P105	T18c_mbtH_8030_f	ACGCCACTGCAGGTCGACTCTAGAGatgaacggcgagccacg
P106	T18c_mbtH_8030_r	ATATCGATGAATTCGAGCTCGGTACctactggcggacgctgc
P107	T18c_mbtH_9502_f	ACGCCACTGCAGGTCGACTCTAGAGatgccggaagacgccg
P108	T18c_mbtH_9502_r	ATATCGATGAATTCGAGCTCGGTACtcagccgtcggccgc
P109	T25_A_M1_f	ACGCGGCGGGCTGCAGGGTCGACTCTAGAGatgaattccgcagcgcgga
P110	T25_A_M1_r	CGACGGCCGAATTCTTAGTTACTTAGGTACgcggatcttcaccctggt
P111	T25_T_M1_f	ACGCGGCGGGCTGCAGGGTCGACTCTAGAGgaacgtgtcctctgtggactct

P112	T25_T_M1_r	CGACGGCCGAATTCTTAGTTACTTAGGTACcgccacggcggcca
P113	T25_C_M2_f	ACGCGGCGGGCTGCAGGGTCGACTCTAGAGgaggagatctggccgctggcccc
P114	T25_C_M2_r	CGACGGCCGAATTCTTAGTTACTTAGGTACgtccagcggcaccc
P115	T25_A_M2_f	ACGCGGCGGGCTGCAGGGTCGACTCTAGAGcggatggcggcgcggc
P116	T25_A_M2_r	CGACGGCCGAATTCTTAGTTACTTAGGTACgcggatcttgacctggt
P117	T25_T_M2_f	ACGCGGCGGGCTGCAGGGTCGACTCTAGAGgagcgggtgctgtgcgagttgtt
P118	T25_T_M2_r	CGACGGCCGAATTCTTAGTTACTTAGGTACcgactccgccgccagct
P119	T25_E_M2_f	ACGCGGCGGGCTGCAGGGTCGACTCTAGAGgtcggtgacgtcgcgtggacgccggtgatg
P120	T25_E_M2_r	CGACGGCCGAATTCTTAGTTACTTAGGTACggccacgtccagcc
P121	T25_M3_A-T-cDD_f	ACGCGGCGGGCTGCAGGGTCGACTCTAGAGttcgccgcgcaggcc
P122	T25_M3_A-T-cDD_r	GCCGAATTCTTAGTTACTTAGGTACCCGGGtcaccggccgcttttcgaat
P123	M4_nDD-C-A_T18n_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGtgaatgagccagtcgcgga
P124	M4_nDD-C-A_T18n_f	TGGCGGCTGAATTCGAGCTCGGTACCCGGGgcggatcttcacctggtcg
P125	T25_M6_A-T-cDD_f	ACGCGGCGGGCTGCAGGGTCGACTCTAGAGttccgcaggcaggcgg
P126	T25_M6_A-T-cDD_r	GCCGAATTCTTAGTTACTTAGGTACCCGGGtcattcgccggccagcc
P127	M7_nDD-C-A_T18n_f	AGCTTGCATGCCTGCAGGTCGACTCTAGAGtgagtgaccgttgacgacactcg
P128	M7_nDD-C-A_T18n_f	CTCGCTGGCGGCTGAATTCGAGCTCGGTACgcggatcttcacctggtcg
P129	T25_Xdom_full_f	ACGCGGCGGGCTGCAGGGTCGACTCTAGAGgcgaagtcacgccccg
P130	T25_Xdom_full_r	GCCGAATTCTTAGTTACTTAGGTACCCGGGtcacggacggccgcg
P131	pUT18n-nDDbpsB-F1_f	atgagccagtcgcggCGGGTACCGAGCTCGAA
P132	pUT18n-nDDbpsB-F1_r	AAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCT
P133	pUT18n-nDDbpsB-F2_f	GATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGAT
P134	pUT18n-nDDbpsB-F2_r	ccgcgactggctcatAGCTGTTTCCTGTGTGAAATTGTTATCC

P135	pUT18n-nDDbpsC-F1_f	atgaccgttgacgacactcgcgcgaagcgccgctccagCCGGGTACCGAGCTCGAA
P136	pUT18n-nDDbpsC-F2_r	gctggagcggcgcttcgcgcgagtgtcgtcaacggtcatAGCTGTTTCCTGTGTGAAATTGTTATCC
P137	T25_cDDbpsA_f	GCGGGCTGCAGGGTCGACTCTAGAGgttgcgggcgatcgt
P138	T25_cDDbpsA_r	ATTCTTAGTTACTTAGGTACCCGGGtcaccggccgctttcgaattcgga
P139	T25_cDDbpsB_f	GCGGGCTGCAGGGTCGACTCTAGAGctcgcggccgaggg
P140	T25_cDDbpsB_r	ATTCTTAGTTACTTAGGTACCCGGGtcattcgccggccag
P141	T25_C_long_M6_f	GCGGGCTGCAGGGTCGACTCTAGAGctcgacatctggccgctt
P142	T25_C_long_M6_r	ATTCTTAGTTACTTAGGTACCCGGGtcacgccgccagcgg
P143	T25_A_long_M6_f	GCGGGCTGCAGGGTCGACTCTAGAGgagacgctgccggtgttgttccgca
P144	T25_A_long_M6_r	ATTCTTAGTTACTTAGGTACCCGGGtcagcgggggcccgc
P145	T25_PCP_long_M6_f	GCGGGCTGCAGGGTCGACTCTAGAGatggaggcgaggctgtgcgc
P146	T25_PCP_long_M6_r	ATTCTTAGTTACTTAGGTACCCGGGtcattcgccggccag
P147	T25_TM1-M2_f	GCGGGCTGCAGGGTCGACTCTAGAGagcaccgatgcggaacgtgtcctct
P148	T25_TM1-M2_r	GCCGAATTCTTAGTTACTTAGGTACttacgccggaccgccg
P149	T25_EM2-M3_f	GCGGGCTGCAGGGTCGACTCTAGAGgtgccggagcccggtcgtt
P150	T25_EM2-M3_r	GCCGAATTCTTAGTTACTTAGGTACtcaccggccgctttcgaat
P151	T25_bpsD_f	GCGGGCTGCAGGGTCGACTCTAGAGaccggcgcgatcgt
P152	T25_bpsD_r	GCCGAATTCTTAGTTACTTAGGTACtcacgtcagggcggcgt
P153	T18c-oxyD_f	ACGCCACTGCAGGTCGACTCTAGAGcagacgacgaacgccgt
P154	T18c-oxyD_r	ATATCGATGAATTCGAGCTCGGTACtcagcgcccggtgaaccgca
P155	T18c-bhp_f	ACGCCACTGCAGGTCGACTCTAGAGctgatgacgactgagcacg
P156	T18c-bhp_r	ATATCGATGAATTCGAGCTCGGTACtcacagttcttcgggggtccgggg
P157	pUT18c_58_I51V_f	CCTCGACTACgtgGAGAAGAACTGGACCG
P158	pUT18c_58_I51V_r	CATTCCTGCCTGCCGGCC
P159	T25_oxyA_f	GCGGGCTGCAGGGTCGACTCTAGAGttcgaggagagcaacgc
P160	T25_oxyA_r	GCCGAATTCTTAGTTACTTAGGTACtcaccactccagcagcaggg
P161	T25_oxyB_f	GCGGGCTGCAGGGTCGACTCTAGAGttgaatgatgacgacccgcggc
P162	T25_oxyB_r	GCCGAATTCTTAGTTACTTAGGTACtcaccaggcgaccatcagct
P163	T25_oxyC_f	GCGGGCTGCAGGGTCGACTCTAGAGgggcacgatatcggtcag
P164	T25_oxyC_r	GCCGAATTCTTAGTTACTTAGGTACctaccaggtgaccggaaccc

P165	T25_TM1-M2_2_f	GCGGGCTGCAGGGTCGACTCTAGAGagttcgcggcgagtgccgccg
P166	T25_TM1-M2_2_r	GCCGAATTCTTAGTTACTTAGGTACcgccggaccgccg
P167	T25_TM1_M2_3_f	GCGGGCTGCAGGGTCGACTCTAGAGcccgagttcgcggcgagtgccgccg
P168	T25_TM1_M2_3_r	GCCGAATTCTTAGTTACTTAGGTACcggaccgccggcg
P169	T25_EM4_M5_f	GCGGGCTGCAGGGTCGACTCTAGAGccacgcgccgaggtcgtcacggg
P170	T25_EM4_M5_r	GCCGAATTCTTAGTTACTTAGGTACcgcgtcgccgccgg
P171	T25_EM5_M6_f	GCGGGCTGCAGGGTCGACTCTAGAGagccgcccggccgg
P172	T25_EM5_M6_r	GCCGAATTCTTAGTTACTTAGGTACtcattcgccggccagcccg
P173	T25_EM4_M5_2_f	GCGGGCTGCAGGGTCGACTCTAGAGgtgggcgagatcccgtggacgccggtgatgcgtgc
P174	T25_EM4_M5_2_r	GCCGAATTCTTAGTTACTTAGGTACcgaccccggtgacgtccagccggag
P175	T25_EM5_M6_2_f	GCGGGCTGCAGGGTCGACTCTAGAGgacggcgtcggcgaggttccgtgga
P176	T25_EM5_M6_2_r	GCCGAATTCTTAGTTACTTAGGTACtcattcgccggccagcccg
P177	T18c_NBD_tba_f	ACGCCACTGCAGGTCGACTCTAGAGgtggtgccggacg
P178	T18c_NBD_tba_r	ATATCGATGAATTCGAGCTCGGTACTTAtcctccgtagcccatgtgttgg
P179	T18n_NBD_tba_f	GCATGCCTGCAGGTCGACTCTAGAGgtggtgccggacg
P180	T18n_NBD_tba_r	TGGCGGCTGAATTCGAGCTCGGTACtcctccgtagcccatgtgttgg
P181	T18c-15_hMLP_f	gtccgccagtagCTAAGTAATATGGTGCACTCTCAGTACAATCTG
P182	T18c-15_hMLP_r	cttcgtggttcaccagcacgaaaaaggagccg
P183	T18c-60_hMLP_f	tttttcgtgctggtgaaccacgaagagcagtactcg
P184	T18c-60_hMLP_r	CCATATTACTTAGctactggcggacgctgcgcgg
P185	T25_M2_-C_f	GCGGGCTGCAGGGTCGACTCTAGAGctggtggggcggatggcggcgc
P186	T25_M2_-C_r	GCCGAATTCTTAGTTACTTAGGTACcgccggaccgcc
P187	T18c_hMLPv2_vec_f	gtctggaccgatctccggccgaagagcc
P188	T18c_hMLPv2_vec_r	acggattgctcatCTCTAGAGTCGA
P189	T18c_hMLPv2_ins_f	TCGACTCTAGAGatgagcaatccgt
P190	T18c_hMLPv2_ins_r	tcttcggccggagatcggtccagac
P191	T25n_bpsD_f	GCATGCCTGCAGGTCGACTCTAGAGatgaccggcgcgatc
P192	T25n_bpsD_r	TGGTCATTGAATTCGAGCTCGGTACcgtcagggcggcgtcgag
P193	T18n_bhp_f	GCATGCCTGCAGGTCGACTCTAGAGatgctgatgacgact
P194	T18n_bhp_r	TGGCGGCTGAATTCGAGCTCGGTACcagttcttcggggg
P195	T18n_oxyD_f	GCATGCCTGCAGGTCGACTCTAGAGatgcagacgacgaac
P196	T18n_oxyD_r	TGGCGGCTGAATTCGAGCTCGGTACgcgcccggtgaa
P197	hMLP_v3_vec_f	acccgcgtgcacCACGAGGGCACCCGC
P198	hMLP_v3_ins_r	gtgcacgcgggtCCAGCCCGGCGGGTT

P199	hMLP_9502_vec_f	gccgacggctgaGTACCGAGCTCGAA
P200	hMLP_9502_vec_r	cgttgaccagtacgaaaaaggagcc
P201	hMLP_9502_ins_f	ggctccttttcgtactggtcaacgacgaaggc
P202	hMLP_9502_ins_r	TCGAGCTCGGTACtcagccgtcggc
P203	hMLP_6685_vec_f	ctcggggtctgaGTACCGAGCTCGAATTCAT
P204	hMLP_6685_vec_r	gctcctcgtcgttcagcacgaaaaaggagcc
P205	hMLP_6685_ins_f	tttttcgtgctgaacgacgaggagcagtattcg
P206	hMLP_6685_ins_r	TCGAGCTCGGTACtcagaccccgag
P207	hMLP4_vec_f	gcgagcgcctgaGTACCGAGCTCG
P208	hMLP4_vec_r	tCCAGCCCGGCGGcacctccgcgaacgcgggccagat
P209	hMLP4_ins_f	atctggcccgcgttcgcggaggtgCCGCCGGGCTGG
P210	hMLP4_ins_r	TCGAGCTCGGTACtcaggcgctcgc
P211	hMLPv7_vec_r	cgcgggccagatcgagtactgctcttcGTCgttcaccagcac
P212	hMLPv7_ins_f	gtgctggtgaacGACgaagagcagtactcgatctggcccgcg
P213	hMLPv8_vec_r	gaacgcgggccagatcgagtactgTCCttcgtggttcac
P214	hMLPv8_ins_f	gtgaaccacgaaGGAcagtactcgatctggcccgcgttc
P215	hMLPv9_vec_r	cacctccgcgaacgcgggccagatcgaGTGctgctcttcgtggtt
P216	hMLPv9_ins_f	aaccacgaagagcagCACtcgatctggcccgcgttcgcggaggtg
P217	hMLPv10_vec_r	cctccagccTGCcggcacctccgcgaacgcgggccagat
P218	hMLPv10_ins_f	atctggcccgcgttcgcggaggtgccgGCAggctggagg
P219	hMLPv11_vec_r	ggtgccctcTCCgccggtgtccctccagcccggcggcac
P220	hMLPv11_ins_f	gtgccgccgggctggagggacaccggcGGAgagggcacc
P221	hMLPv12_vec_f	cgttgtcgaacggattgctcatCTCTAGAGTCGACCTGCA
P222	hMLPv12_vec_r	gagggcacccgcCAAgagtgcctggaccacgtcgagcgg
P223	hMLPv12_ins_f	TGCAGGTCGACTCTAGAGatgagcaatccgttcgacaacg
P224	hMLPv12_ins_r	ccgctcgacgtggtccaggcactcTTGgcgggtgccctc
P225	M2_A+T_f	gcgggctgcagggtcgactctagagcccgccggcctggtgggg
P226	M2_A+T_r	agtgaattcttacttacttaggtaccgactccgccgccagct
P227	M5_A+T_f	gcgggctgcagggtcgactctagagccttcgctgccggtggg
P228	M5_A+T_r	agtgaattcttacttacttaggtacgccacgctcggcga
P229	cDDbpsB_UP_f	GGAGAGACGAATTCGAGCTCGGTACatggccgtgtcgttcgc
P230	cDDbpsB_UP_r	cttatggcggtcattcctggacgatcaccgc
P231	cDDbpsB_DO_f	atcgtccaggaatgaccgccataaggagcaga
P232	cDDbpsB_DO_r	ACTCACTATAGGGAAAGCTTGCATGggtcaggttcgcgccg
P233	pSP1_lacZ_f	ATGTGCTGCAAGGCGATTAAGT
P234	pSP1_lacZ_r	GAAACAGCTATGACCATGATTACGCC
P235	KO_cDDbpsB_f	atgcagctgtcggcccg
P236	KO_cDDbpsB_r	gcattccctcctgcagcg
P237	KO_cDDbpsB_r_2	agacgaacgagatccgctcggc
P238	cDDbpsA_up_f	GGAGAGACGAATTCGAGCTCGGTACtccgggctatccggcc
P239	cDDbpsA_up_r	ccagccacgttcaccgttccacgatcgccg
P240	cDDbpsA_down_f	atcgtggaacggtgaacgtggctggacatgttgt

P241	cDDbpsA_down_r	ACTCACTATAGGGAAAGCTTGCATGggcaccccattccccg
P242	pSP1ΔcDDbpsA_seq1	atcgccttcgatctgacggtcacc
P243	pSP1ΔcDDbpsA_seq2	agctggtgttcgccggccg
P244	pSP1ΔcDDbpsA_seq3	agcagcgcacgcccgcg
P245	pSP1ΔcDDbpsA_seq4	accgcctggtggtgacctgcc
P246	pSP1ΔcDDbpsA_seq5	acacgctgctcgtcttcgagaag
P247	KO_cDDbpsA_f	agcggctgctgtgtgc
P248	KO_cDDbpsA_r	gccagatttcctcgatccgcga

Table S2: Strains used in this study

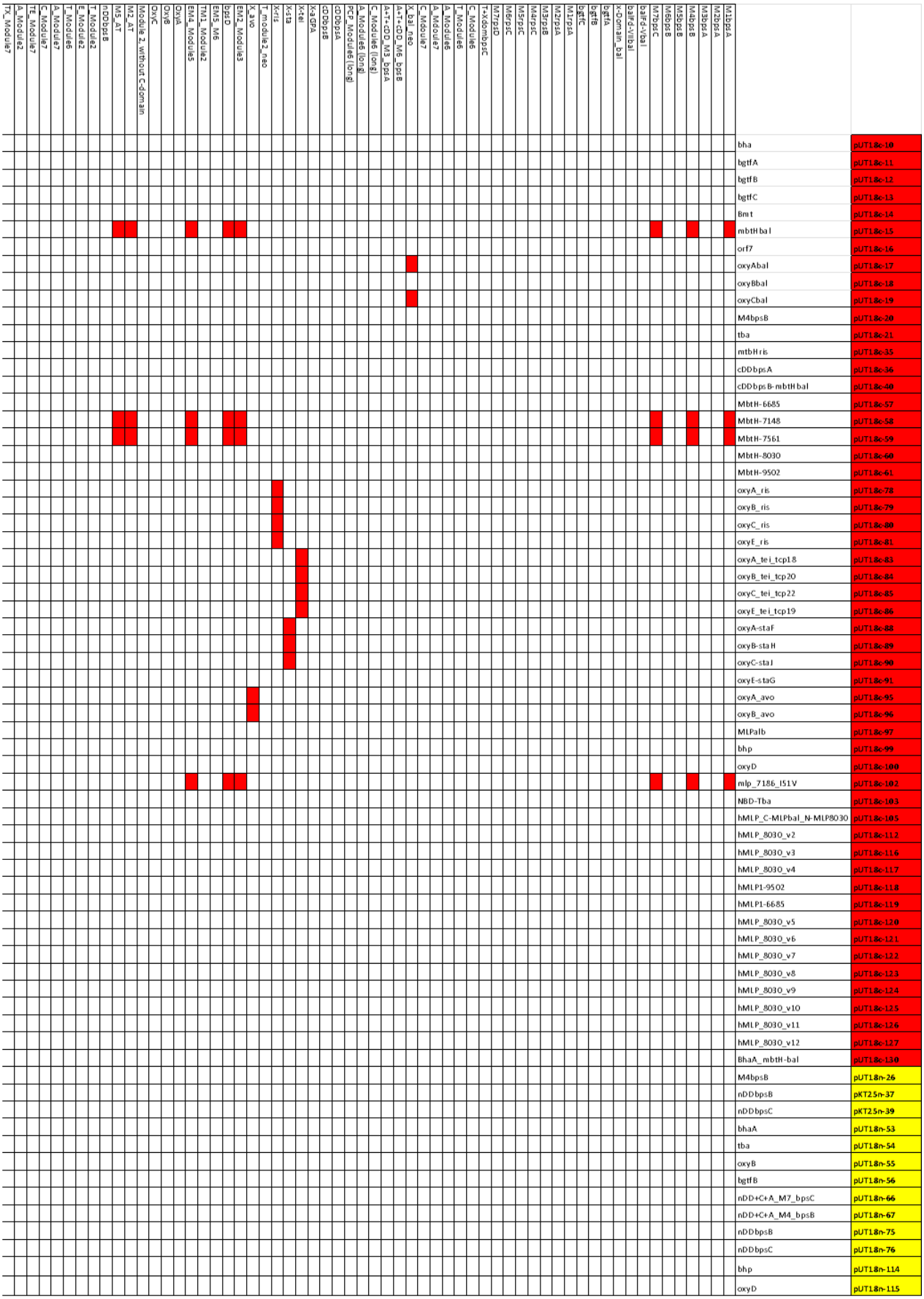
Strain	Description	Source
A. balhimycina DSM 44591	wildtype strain	(Nadkarni et al., 1994; Wink et al., 2003)
A. balhimycina ΔbpsC	Deletion of the third NRPS gene bpsC of the balhimycin biosynthetic gene cluster	This study
A. balhimycinaΔcDD bpsB (ΔcDD2)	Deletion of C-terminal docking domain of the NRPS gene *bpsB* from the balhimycin biosynthetic gene cluster	This study
A. balhimycinaΔcDD 1	Deletion of C-terminal docking domain of the NRPS gene *bpsA* from the balhimycin biosynthetic gene cluster	This study
A. balhimycinaΔcDD 1ΔcDD2	Deletion of C-terminal docking domain of the NRPS gene *bpsA* and *bpsB* from the balhimycin biosynthetic gene cluster	This study
A. balhimycinaΔbps C::pRM4-bpsC-TurboID-His	bpsC deletion mutant, ΦC31attB(pRM4-bpsC-TurboID),AprR	This study
E. coli JM110	lacY dam dcm [F' lacIq Z Δ M15] (Methylation deficient strain)	(Yanisch-Perron et al., 1985)
E. coli ET12567	F− dam-13::Tn9 dcm-6 hsdM hsdR (Methylation deficient strain)	(MacNeil et al., 1992)
E. coli NovaBlue	Cloning host for plasmid generation: endA1 hsdR17 (rK12– mK12+) supE44 thi-1 recA1 gyrA96 relA1 lac F′[proA+B+ lacIqZΔM15::Tn10] (TetR)	Novagen®
E. coli BTH101	F-, cya-99, araD139, galE15, galK16, rpsL1 (StrR), hsdR2, mcrA1, mcrB1	Karimova et. al.
*AprR (Apramycin resistance) *EryR (Erythromycin resistance) *TetR (Tetrycyclin resistance) *StrR (Streptomycin resistance)		

Table S3: Plasmids used in this study

Plasmid	Relevant characteristics	Reference
pSP1	pT7/T3-α19 derived gene disruption vector, with ermE gene in SapI site EryR/AmpR	(Pelzer et al., 1997)
pRM4	pSET152ermEp* derived Φ31 integration vector with artificial ribosomal binding site, AprR	(Menges et al., 2007)
pKT25	aph, cya-T25	Karimova et al.
pUT18c	bla, cya-T18	Karimova et al.
pNKT25	aph, cya-T25	Karimova et al.
pUT18	bla, cya-T18	Karimova et al.
pKT25-zip	pKT25-derivative contanining the leucine zipper of GCN4 fused to T25	Karimova et al.
pUT18c-zip	pUT18c-derivative contanining the leucine zipper of GCN5 fused to T18	Karimova et al.
pKT25-1	M1bpsA	This study
pKT25-2	M2bpsA	This study
pKT25-3	M3bpsA	This study
pKT25-4	M4bpsB	This study
pKT25-5	M5bpsB	This study
pKT25-6	M6bpsB	This study
pKT25-7	M7bpsC	This study
pKT25-8	balFd-Vbal	This study
pKT25-9	balFd-VIIbal	This study
pUT18c-10	bha	This study
pUT18c-11	bgtfA	This study
pUT18c-12	bgtfB	This study
pUT18c-13	bgtfC	This study
pUT18c-14	Bmt	This study
pUT18c-15	mbtHbal	This study
pUT18c-16	orf7	This study
pUT18c-17	oxyAbal	This study
pUT18c-18	oxyBbal	This study
pUT18c-19	oxyCbal	This study
pUT18c-20	M4bpsB	This study
pUT18c-21	tba	This study
pKT25-22	x-Domainbal	This study
pKT25-23	bgtfA	This study
pKT25-24	bgtfB	This study
pKT25-25	bgtfC	This study
pUT18n-26	M4bpsB	This study
pUT18c-27	M7bpsC	This study
pUT18c-36	cDDbpsA	This study
pKT25n-37	nDDbpsB	This study

pUT18c-38	cDDbpsB	This study
pKT25n-39	nDDbpsC	This study
pUT18c-40	cDDbpsB-mbtHbal	This study
pKT25-41	T+XdombpsC	This study
pKT25-42	C_Module6	This study
pKT25-43	T_Module6	This study
pKT25-44	A_Module6	This study
pKT25-45	A_Mdoule7	This study
pKT25-46	C_Mdoule7	This study
pNKT25-47	T_Module2	This study
pNKT25-48	E_Module2	This study
pNKT25-49	T_Module6	This study
pNKT25-50	A_Module7	This study
pNTK25-51	C_Module7	This study
pNTK25-52	TE_Module7	This study
pUT18n-53	bhaA	This study
pUT18n-54	tba	This study
pUT18n-55	oxyB	This study
pUT18n-56	bgtfB	This study
pUT18c-57	MbtH-6685	This study
pUT18c-58	MbtH-7148	This study
pUT18c-59	MbtH-7561	This study
pUT18c-60	MbtH-8030	This study
pUT18c-61	MbtH-9502	This study
pNKT25-62	A_Module2	This study
pNKT25-63	TX_Module7	This study
pKT25-65	X_bal_neo	This study
pUT18n-66	nDD+C+A_M7_bpsC	This study
pUT18n-67	nDD+C+A_M4_bpsB	This study
pKT25-68	A+T+cDD_M6_bpsB	This study
pKT25-69	A+T+cDD_M3_bpsA	This study
pKT25-70	C_Module6 (long)	This study
pKT25-71	A_Module6 (long)	This study
pKT25-72	PCP_Module6 (long)	This study
pTK25-93	T_module2_neo	This study
pKT25-94	X_avo	This study
pUT18c-95	oxyA_avo	This study
pUT18c-96	oxyB_avo	This study
pUT18c-97	MLPalb	This study
pKT25-98	EM2_M3	This study
pUT18c-99	Bhp	This study
pUT18c-100	oxyD	This study

pKT25-101	bpsD	This study
pUT18c-102	mlp_7184_l51V	This study
pUT18c-103	NBD-Tba	This study
pKT25-104	EM5_M6	This study
pUT18c-105	hMLP_hMLP_8030_v1	This study
pKT25-106	TM1_Module2	This study
pKT25-107	EM4_Module5	This study
pKT25-108	OxyA	This study
pKT25-109	OxyB	This study
pKT25-110	OxyC	This study
pKT25-111	Module 2, without C-domain	This study
pUT18c-112	hMLP_8030_v2	This study
pUT18n-114	bhp	This study
pUT18n-115	oxyD	This study
pUT18c-116	hMLP_8030_v3	This study
pUT18c-117	hMLP_8030_v4	This study
pUT18c-118	hMLP1-9502	This study
pUT18c-119	hMLP1-6685	This study
pUT18c-120	hMLP_8030_v5	This study
pUT18c-121	hMLP_8030_v6	This study
pUT18c-122	hMLP_8030_v7	This study
pUT18c-123	hMLP_8030_v8	This study
pUT18c-124	hMLP_8030_v9	This study
pUT18c-125	hMLP_8030_v10	This study
pUT18c-126	hMLP_8030_v11	This study
pUT18c-127	hMLP_8030_v12	This study
pKT25-128	M2_AT	This study
pKT25-129	M5_AT	This study



bha	pUT18c-10
bgtfA	pUT18c-11
bgtfB	pUT18c-12
bgtfC	pUT18c-13
Bmt	pUT18c-14
mbtH bal	pUT18c-15
orf7	pUT18c-16
oxyAbal	pUT18c-17
oxyBbal	pUT18c-18
oxyCbal	pUT18c-19
M4bpsB	pUT18c-20
tba	pUT18c-21
mtbH ris	pUT18c-35
cDDbpsA	pUT18c-36
cDDbpsB-mbtH bal	pUT18c-40
MbtH-6685	pUT18c-57
MbtH-7148	pUT18c-58
MbtH-7561	pUT18c-59
MbtH-8030	pUT18c-60
MbtH-9502	pUT18c-61
oxyA_ris	pUT18c-78
oxyB_ris	pUT18c-79
oxyC_ris	pUT18c-80
oxyE_ris	pUT18c-81
oxyA_tei_tcp18	pUT18c-83
oxyB_tei_tcp20	pUT18c-84
oxyC_tei_tcp22	pUT18c-85
oxyE_tei_tcp19	pUT18c-86
oxyA-staF	pUT18c-88
oxyB-staH	pUT18c-89
oxyC-staJ	pUT18c-90
oxyE-staG	pUT18c-91
oxyA_avo	pUT18c-95
oxyB_avo	pUT18c-96
MLPalb	pUT18c-97
bhp	pUT18c-99
oxyD	pUT18c-100
mlp_7186_I51V	pUT18c-102
NBD-Tba	pUT18c-103
hMLP_C-MLPbal_N-MLP8030	pUT18c-105
hMLP_8030_v2	pUT18c-112
hMLP_8030_v3	pUT18c-116
hMLP_8030_v4	pUT18c-117
hMLP1-9502	pUT18c-118
hMLP1-6685	pUT18c-119
hMLP_8030_v5	pUT18c-120
hMLP_8030_v6	pUT18c-121
hMLP_8030_v7	pUT18c-122
hMLP_8030_v8	pUT18c-123
hMLP_8030_v9	pUT18c-124
hMLP_8030_v10	pUT18c-125
hMLP_8030_v11	pUT18c-126
hMLP_8030_v12	pUT18c-127
BhaA_mbtH-bal	pUT18c-130
M4bpsB	pUT18n-26
nDDbpsB	pKT25n-37
nDDbpsC	pKT25n-39
bhaA	pUT18n-53
tba	pUT18n-54
oxyB	pUT18n-55
bgtfB	pUT18n-56
nDD+C+A_M7_bpsC	pUT18n-66
nDD+C+A_M4_bpsB	pUT18n-67
nDDbpsB	pUT18n-75
nDDbpsC	pUT18n-76
bhp	pUT18n-114
oxyD	pUT18n-115

Table S4: Pairwise interactions performed in the Bacterial Two-Hybrid system.

Region	Type	From	To	Most similar known cluster	Most similar known cluster_1	MLP
Region 1	bacteriocin	1,281	11,485			
Region 2	terpene	373,382	392,979	isorenieratene	Terpene	
Region 3	terpene	1,439,560	1,461,064	geosmin	Terpene	
Region 4	lanthipeptide	2,361,844	2,381,483	Ery-9 / Ery-6 / Ery-8 / Ery-7 / Ery-5 / Ery-4 / Ery-3	RiPP:Lanthipeptide	
Region 5	ectoine	3,609,448	3,619,150	ectoine	Other	
Region 6	NRPS-like	4,824,012	4,863,200	thiotetroamide	Polyketide	
Region 7	NRPS,bacteriocin	5,001,399	5,042,645	atratumycin	NRP	
Region 8	thiopeptide,T1PKS	5,245,653	5,330,527	amycolamycin A / amycolamycin B	Polyketide	
Region 9	indole	5,752,184	5,773,134	ulleungmycin	NRP	
Region 10	NRPS,T3PKS,T1PKS,NRPS-like	6,016,717	6,120,327	balhimycin	NRP	MLP_{bal}
Region 11	NRPS,bacteriocin	6,184,085	6,231,149			
Region 12	ectoine	6,289,389	6,297,415	kosinostatin	NRP + Polyketide	
Region 13	NRPS,T1PKS	6,720,182	6,788,285	stambomycin A / stambomycin B / stambomycin C / stambomycin D	Polyketide:Modular type I + Saccharide:Hybrid/tailoring	
Region 14	NRPS,bacteriocin,lanthipeptide	7,087,574	7,175,113	RP-1776	Polyketide + NRP:Cyclic depsipeptide	MLP_{6685}
Region 15	NRPS,terpene	7,213,946	7,265,822	isorenieratene	Terpene	
Region 16	hglE-KS,T1PKS	7,599,051	7,648,097	melilingmycin	Polyketide	
Region 17	NRPS,NRPS-like,T1PKS	7,664,020	7,984,262	nystatin A1	Polyketide:Modular type I + Saccharide:Hybrid/tailoring	MLP_{7148}
Region 18	NRPS,T1PKS,bacteriocin	8,091,004	8,160,164	tyrobetaine	NRP	
Region 19	NRPS	8,315,353	8,369,024	scabichelin	NRP	MLP_{7561}
Region 20	PKS-like,NRPS,T1PKS	8,852,630	8,910,596	tubulysin A	NRP + Polyketide	MLP_{8030}
Region 21	NRPS-like	9,002,144	9,043,722			
Region 22	T2PKS,NRPS,nucleoside	9,072,806	9,162,869	azicemicin B	Polyketide	
Region 23	NRPS,T1PKS	9,502,312	9,555,599	tyrobetaine	NRP	
Region 24	T2PKS,NRPS,oligosaccharide,T1PKS	9,629,964	9,721,626	arixanthomycin A / arixanthomycin B / arixanthomycin C	Polyketide:Type II + Saccharide:Hybrid/tailoring	
Region 25	NRPS-like	10,385,905	10,428,423	herboxidiene	Polyketide	
Region 26	NRPS,T1PKS,ectoine	10,516,163	10,622,461	A83543A	Polyketide	MLP_{9502}
Region 27	siderophore	10,820,402	10,832,132	macrotetrolide	Polyketide	

Table S5: AntiSmash analysis of the *A. balhimycina* genome, used to predict the BGCs. Depicted in blue are the predicted BGCs in which the MLP-encoding genes were identified.

References

1. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M.R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., Jumper, J.M., 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630, 493–500. https://doi.org/10.1038/s41586-024-07487-w
2. Álvarez-Salmoral, D., Borza, R., Xie, R., Joosten, R.P., Hekkelman, M.L., Perrakis, A., 2024. AlphaBridge: tools for the analysis of predicted macromolecular complexes. https://doi.org/10.1101/2024.10.23.619601
3. Bibb, M.J., Janssen, G.R., Ward, J.M., 1985. Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*. Gene 38, 215–226. https://doi.org/10.1016/0378-1119(85)90220-3
4. Bischoff, D., Bister, B., Bertazzo, M., Pfeifer, V., Stegmann, E., Nicholson, G.J., Keller, S., Pelzer, S., Wohlleben, W., Süssmuth, R.D., 2005. The Biosynthesis of Vancomycin-Type Glycopeptide Antibiotics—A Model for Oxidative Side-Chain Cross-Linking by Oxygenases Coupled to the Action of Peptide Synthetases. ChemBioChem 6, 267–272. https://doi.org/10.1002/cbic.200400328
5. Bischoff, D., Pelzer, S., Bister, B., Nicholson, G.J., Stockert, S., Schirle, M., Wohlleben, W., Jung, G., Süssmuth, R.D., 2001. The Biosynthesis of Vancomycin-Type Glycopeptide Antibiotics-The Order of the Cyclization Steps. Angew. Chem. Int. Ed. 40, 4688–4691. https://doi.org/10.1002/1521-3773(20011217)40:24<4688::AID-ANIE4688>3.0.CO;2-M
6. Butler, M.S., Hansford, K.A., Blaskovich, M.A.T., Halai, R., Cooper, M.A., 2014. Glycopeptide antibiotics: Back to the future. J Antibiot 67, 631–644. https://doi.org/10.1038/ja.2014.111
7. Chen, J., van Heel, A.J., Kuipers, O.P., 2021. Visualization and Analysis of the Dynamic Assembly of a Heterologous Lantibiotic Biosynthesis Complex in Bacillus subtilis. mBio 12, 10.1128/mbio.01219-21. https://doi.org/10.1128/mbio.01219-21
8. Chen, J., van Heel, A.J., Kuipers, O.P., 2020. Subcellular Localization and Assembly Process of the Nisin Biosynthesis Machinery in Lactococcus lactis. mBio 11, 10.1128/mbio.02825-20. https://doi.org/10.1128/mbio.02825-20
9. Esquilín-Lebrón, K.J., Boynton, T.O., Shimkets, L.J., Thomas, M.G., 2018. An Orphan MbtH-Like Protein Interacts with Multiple Nonribosomal Peptide Synthetases in Myxococcus xanthus DK1622. Journal of Bacteriology 200. https://doi.org/10.1128/jb.00346-18
10. Gacek-Matthews, A., Chromiková, Z., Sulyok, M., Lücking, G., Barák, I., Ehling-Schulz, M., 2020. Beyond Toxin Transport: Novel Role of ABC Transporter for Enzymatic Machinery of Cereulide NRPS Assembly Line. mBio 11, e01577-20. https://doi.org/10.1128/mBio.01577-20
11. Gavriilidou, A., Adamek, M., Rodler, J.-P., Kubach, N., Voigtländer, A., Kokkoliadis, L., Hughes, C.C., Cryle, M.J., Stegmann, E., Ziemert, N., 2024. Animating insights into the biosynthesis of glycopeptide antibiotics. Current Opinion in Microbiology 82, 102561. https://doi.org/10.1016/j.mib.2024.102561

12. Gericke, N., Beqaj, D., Kronenberger, T., Kulik, A., Gavriilidou, A., Franz-Wachtel, M., Schoppmeier, U., Harbig, T., Rapp, J., Grin, I., Ziemert, N., Link, H., Nieselt, K., Macek, B., Wohlleben, W., Stegmann, E., Wagner, S., 2025. Unveiling the substrate specificity of the ABC transporter Tba and its role in glycopeptide biosynthesis. iScience 112135. https://doi.org/10.1016/j.isci.2025.112135
13. Gulick, A.M., 2009. Conformational Dynamics in the Acyl-CoA Synthetases, Adenylation Domains of Non-ribosomal Peptide Synthetases, and Firefly Luciferase. ACS Chem. Biol. 4, 811–827. https://doi.org/10.1021/cb900156h
14. Haslinger, K., Peschke, M., Brieke, C., Maximowitsch, E., Cryle, M.J., 2015. X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. Nature 521, 105–109. https://doi.org/10.1038/nature14141
15. Herbst, D.A., Herbst, Dominik A., Bjoern Boll, Boll, B., Zocher, G., Stehle, T., Heide, L., 2013. Structural Basis of the Interaction of MbtH-like Proteins, Putative Regulators of Nonribosomal Peptide Biosynthesis, with Adenylating Enzymes. Journal of Biological Chemistry 288, 1991–2003. https://doi.org/10.1074/jbc.m112.420182
16. Kerfeld, C.A., Erbilgin, O., 2015. Bacterial microcompartments and the modular construction of microbial metabolism. Trends in Microbiology 23, 22–34. https://doi.org/10.1016/j.tim.2014.10.003
17. Kittilä, T., Kittel, C., Tailhades, J., Butz, D., Schoppet, M., Büttner, A., Goode, R.J.A., Schittenhelm, R.B., Van Pee, K.-H., Süssmuth, R.D., Wohlleben, W., Cryle, M.J., Stegmann, E., 2017. Halogenation of glycopeptide antibiotics occurs at the amino acid level during non-ribosomal peptide synthesis. Chem. Sci. 8, 5992–6004. https://doi.org/10.1039/C7SC00460E
18. Krissinel, E., Henrick, K., 2007. Inference of Macromolecular Assemblies from Crystalline State. Journal of Molecular Biology 372, 774–797. https://doi.org/10.1016/j.jmb.2007.05.022
19. Lautru, S., Oves-Costales, D., Pernodet, J.-L., Challis, G.L., 2007. MbtH-like protein-mediated cross-talk between non-ribosomal peptide antibiotic and siderophore biosynthetic pathways in Streptomyces coelicolor M145. Microbiology 153, 1405–1412. https://doi.org/10.1099/mic.0.2006/003145-0
20. Li, Y., Tahlan, K., Bignell, D.R.D., 2020. Functional Cross-Talk of MbtH-Like Proteins During Thaxtomin Biosynthesis in the Potato Common Scab Pathogen Streptomyces scabiei. Front. Microbiol. 11, 585456. https://doi.org/10.3389/fmicb.2020.585456
21. Marahiel, M.A., Stachelhaus, T., Mootz, H.D., 1997. Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. Chem. Rev. 97, 2651–2674. https://doi.org/10.1021/cr960029e
22. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., Ferrin, T.E., 2023. UCSF ChimeraX: Tools for structure building and analysis. Protein Science 32, e4792. https://doi.org/10.1002/pro.4792
23. Menges, R., Muth, G., Wohlleben, W., Stegmann, E., 2007. The ABC transporter Tba of *Amycolatopsis balhimycina* is required for efficient export of the glycopeptide antibiotic balhimycin. Appl Microbiol Biotechnol 77, 125–134. https://doi.org/10.1007/s00253-007-1139-x
24. Mori, S., Green, K.D., Choi, R., Buchko, G.W., Fried, M., Garneau-Tsodikova, S., 2018. Using MbtH-Like Proteins to Alter the Substrate Profile of a Nonribosomal Peptide Adenylation Enzyme. ChemBioChem 19, 2186–2194. https://doi.org/10.1002/cbic.201800240

25. Nicolaou, K.C., Boddy, C.N.C., Bräse, S., Winssinger, N., 1999. Chemistry, Biology, and Medicine of the Glycopeptide Antibiotics. Angew. Chem. Int. Ed. 38, 2096–2152. https://doi.org/10.1002/(SICI)1521-3773(19990802)38:15<2096::AID-ANIE2096>3.0.CO;2-F
26. O'Connor, S.E., Walsh, C.T., Liu, F., 2003. Biosynthesis of Epothilone Intermediates with Alternate Starter Units: Engineering Polyketide–Nonribosomal Interfaces. Angewandte Chemie International Edition 42, 3917–3921. https://doi.org/10.1002/anie.200352077
27. Pelzer, S., Reichert, W., Huppert, M., Heckmann, D., Wohlleben, W., 1997. Cloning and analysis of a peptide synthetase gene of the balhimycin producer Amycolatopsis mediterranei DSM5908 and development of a gene disruption/replacement system. Journal of Biotechnology 56, 115–128. https://doi.org/10.1016/S0168-1656(97)00082-5
28. Puk, O., Huber, P., Bischoff, D., Recktenwald, J., Jung, G., Süssmuth, R.D., van Pée, K.H., Wohlleben, W., Pelzer, S., 2002. Glycopeptide biosynthesis in Amycolatopsis mediterranei DSM5908: function of a halogenase and a haloperoxidase/perhydrolase. Chem Biol 9, 225–235. https://doi.org/10.1016/s1074-5521(02)00101-1
29. Schomer, R.A., Thomas, M.G., 2017. Characterization of the Functional Variance in MbtH-like Protein Interactions with a Nonribosomal Peptide Synthetase. Biochemistry 56, 5380–5390. https://doi.org/10.1021/acs.biochem.7b00517
30. Sosio, M., Kloosterman, H., Bianchi, A., de Vreugd, P., Dijkhuizen, L., Donadio, S., 2004. Organization of the teicoplanin gene cluster in Actinoplanes teichomyceticus. Microbiology (Reading) 150, 95–102. https://doi.org/10.1099/mic.0.26507-0
31. Sosio, M., Stinchi, S., Beltrametti, F., Lazzarini, A., Donadio, S., 2003. The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by nonomuraea species. Chem Biol 10, 541–549. https://doi.org/10.1016/s1074-5521(03)00120-0
32. Stegmann, E., Frasch, H.-J., Wohlleben, W., 2010. Glycopeptide biosynthesis in the context of basic cellular functions. Current Opinion in Microbiology 13, 595–602. https://doi.org/10.1016/j.mib.2010.08.011
33. Stegmann, Evi, Pelzer, S., Bischoff, D., Puk, O., Stockert, S., Butz, D., Zerbe, K., Robinson, J., Süssmuth, R.D., Wohlleben, W., 2006. Genetic analysis of the balhimycin (vancomycin-type) oxygenase genes. Journal of Biotechnology 124, 640–653. https://doi.org/10.1016/j.jbiotec.2006.04.009
34. Stegmann, Efthimia, Rausch, C., Stockert, S., Burkert, D., Wohlleben, W., 2006a. The small MbtH-like protein encoded by an internal gene of the balhimycin biosynthetic gene cluster is not required for glycopeptide production. FEMS Microbiology Letters 262, 85–92. https://doi.org/10.1111/j.1574-6968.2006.00368.x
35. Stegmann, Efthimia, Rausch, C., Stockert, S., Burkert, D., Wohlleben, W., 2006b. The small MbtH-like protein encoded by an internal gene of the balhimycin biosynthetic gene cluster is not required for glycopeptide production. FEMS Microbiology Letters 262, 85–92. https://doi.org/10.1111/j.1574-6968.2006.00368.x
36. van Wageningen, AndréM.A., Kirkpatrick, P.N., Williams, D.H., Harris, B.R., Kershaw, J.K., Lennard, N.J., Jones, M., Jones, S.J.M., Solenberg, P.J., 1998. Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. Chemistry & Biology 5, 155–162. https://doi.org/10.1016/S1074-5521(98)90060-6
37. Watzel, J., Duchardt-Ferner, E., Sarawi, S., Bode, H.B., Wöhnert, J., 2021. Cooperation between a T Domain and a Minimal C-Terminal Docking Domain to Enable Specific Assembly in a Multiprotein NRPS. Angewandte Chemie International Edition 60, 14171–14178. https://doi.org/10.1002/anie.202103498

38. Woithe, K., Geib, N., Zerbe, K., Li, D.B., Heck, M., Fournier-Rousset, S., Meyer, O., Vitali, F., Matoba, N., Abou-Hadeed, K., Robinson, J.A., 2007. Oxidative Phenol Coupling Reactions Catalyzed by OxyB: A Cytochrome P450 from the Vancomycin Producing Organism. Implications for Vancomycin Biosynthesis. J. Am. Chem. Soc. 129, 6887–6895. https://doi.org/10.1021/ja071038f

Aknowledgements
The work in the laboratory of E.S.,S.W., W.W.,T.S., N.Z. and B.M., related to this study was funded by the Deutsche Forschungsgemeinschaft (DFG) as part of the transregional collaborative research center TRR 261 ‘Cellular Mechanisms of Antibiotic Action and Production’ (Projects B01/B02/B05/Z03). Data management and storage of raw data for this project were supported by the Quantitative Biology Center (QBiC), University of Tübingen, Germany. The authors would like to thank Libera Lo Presti for her support in writing, comments on the manuscript, and fruitful discussions.

Author contributions
D.B. performed the *in vivo* experiments of docking domains, bacterial two-hybrid assays, BN-PAGE experiments, AlphaFold3 predictions, proximity labelling and analyzed all derived data. N.P. performed the *in vitro* analysis of docking domains and analzyed the data and wrote the original draft related to it. A.K. performed all HPLC-MS measurements in this study. M.F.W. and B.M. performed proteomics. D.B. and E.S., drafted the paper and wrote the original manuscript.

Conflict of interest
The authors declare no competing interests.

5 Discussion

Backbone-guided specificity of Tba reframes GPA-transporter compatibility

Our analysis of the evolutionary relationships of ABC transporters involved in GPA biosynthesis suggested a co-evolution between transporters and their respective substrates. This is supported by the fact that GPA- associated ABC transporters cluster phylogenetically according to GPA type and are encoded within their corresponding BGC. These findings highlight a structural dependency between GPA and its cognate ABC transporters, likely resulting from evolutionary adaptation.

A central determinant of this specificity appears to be the peptide backbone composition of the GPAs. While post-NRPS modifications, such as glycosylation or halogenation, can differ even within the same GPA type, our MD simulations and in vivo export assays indicate that these modifications play a lesser role in substrate recognition. Instead, our data show that backbone structure is the primary factor guiding transporter-substrate compatibility.
For instance, ABC transporters associated with type I GPAs, such as Tba and Tva, export comparable levels of balhimycin, whereas Tri, the transporter linked to a type III GPA, exhibits significantly reduced export efficiency. Supporting this, mutants producing non-chlorinated and non-glycosylated balhimycin derivatives exhibit higher export rates compared to the wildtype, reinforcing the minor contribution of decorations to transporter recognition.

Although a detailed mechanistic model of substrate binding is currently hindered by the lack of structural data on Tba–balhimycin interactions, our MD simulations provide first hypotheses.
In its predicted inward-open conformation, Tba appears to interact with the balhimycin backbone, particularly residues 4-7 via residues F269, Q309, and G313 in the predicted binding pocket. Sugar moieties, in contrast, point toward the cytoplasm and likely play only a marginal role in binding. Interestingly, while the major backbone differences among GPA types occur at residues 1-3, these do not appear central to substrate interaction in our models.

In addition, cross-linking differences between type I and type III/IV GPAs, particularly involving amino acid residues 1 and 3, may influence the molecule's structural rigidity or charge distribution, factors likely to affect recognition and binding by their cognate ABC transporters.
Structural studies examining GPA interactions with their cellular target, the D-Ala-D-Ala terminus of lipid II, suggested that structural differences between the type I GPAs (vancomycin and balhimycin) and type III GPA (ristocetin A) can influence the dimerization and supramolecular organization [94,95]. Type I GPAs such as vancomycin and balhimycin, likely act as dimers during target binding [96,97]. These findings underscore the role of specific structural

elements, particularly sugar moieties, in modulating molecular interactions, which we could also confirm experimentally.

The timing and location of balhimycin dimerization remain elusive. It is conceivable that dimerization occurs intracellularly and may represent a crucial prerequisite for effective transport, a hypothesis not addressed in this study but warranting further investigation.
Notably, substrate specificity among natural product-associated ABC transporters varies considerably. While Tba displays narrow specificity, the nisin transporter NisT has been reported to exhibit a broader substrate range [98]. Conversely, MjcD, the transporter of microcin J25, shows high selectivity for its native substrate [99], underscoring the diversity of recognition mechanisms among different transporter–substrate systems.
In this context, the backbone-guided specificity observed in GPAs is of particular interest, especially for the rational design of pathway engineering strategies aimed at generating novel derivatives. Our findings demonstrate that even subtle alterations in specific residues or chemical modifications can markedly influence export efficiency and, by extension, the overall yield. However, modulating export without compromising bioactivity poses a significant challenge and highlights the need for a nuanced understanding of structure–function relationships in GPA transport.

Dual role of Tba: exporter and biosynthesis switch

While our findings elucidate the molecular basis of substrate recognition by Tba, they also unexpectedly uncovered a secondary, equally crucial function for this transporter. Beyond its established role in export, Tba appears to actively participate in regulating balhimycin biosynthesis itself, expanding our understanding of transporter functionality in GPA-producing bacteria.
The biosynthesis of balhimycin represents a complex and tightly regulated process characteristic of GPA production. It involves a coordinated interplay of numerous enzymes that assemble the amino acid building blocks into a peptide backbone, which is subsequently modified through various tailoring reactions, leading to the final bioactive compound [46,100]. The ABC transporter Tba is responsible for exporting the active balhimycin into the extracellular environment[40] where it targets the D-alanyl-D-alanine (D-Ala-D-Ala) termini of the peptidoglycan pentapeptide involved in bacterial cell wall biosynthesis, presumably specifically that of competing bacteria in the surrounding environment [41].
ABC transporters belong to the protein superfamily of ABC systems and are known for mediating substrate translocation across cellular membranes, facilitating exchange between

intra- and extracellular environments. Depending on their structural and functional characteristics, they're functionally categorized as either importers or exporters [101–103]. These functions are attributed to the classical role of ABC transporters, suggesting that their primary role is substrate binding and translocation in a defined direction. However, recent studies have revealed that certain members of the ABC system family can perform noncanonical functions, such as involvement in mRNA translation. For example, the ABC protein ABCE1 regulates mRNA translation by physically remodeling the ribosome during no-go decay. It binds the stalled ribosome along with a class I release factor (RF1). The ribosome is physically split by the ATP-driven conformational shift of ABCE1. This activity helps control whether translation should restart, or the faulty mRNA and incomplete polypeptide should be targeted for degradation [104].

ABC transporters in antibiotic-producing bacteria are generally known to function as exporters and are often associated with such self-resistance mechanisms by actively extruding the compound from the cytoplasm [105]. However, the ABC transporter Tba does not contribute to resistance in *A. balhimycina* [40]. Self-resistance against the glycopeptide antibiotic balhimycin is primarily mediated by the *vanHAX* genes [41] but also involves additional genomic determinants that contribute to protecting the producer from its own antibiotic [106,107].

Our investigation of Tba uncovered a previously unrecognized functional role in GPA biosynthesis. Specifically, deletion of *tba* resulted in a drastic reduction of balhimycin production, which could not be restored by complementation with transporter genes from biosynthetic gene clusters (BGCs) encoding different GPA types. Only the native *tba* gene or its counterpart from the vancomycin BGC , belonging to the same GPA structural class, was able to recover production. These findings suggest two key insights: first, that balhimycin export is highly specific and requires an ABC transporter from a BGC encoding the same GPA type; and second, that active transport is intrinsically linked to the biosynthetic process itself.

The transporter-dependent regulation of biosynthesis is not unique to balhimycin. Similar phenomena have been observed in other natural product biosynthetic pathways. For instance, deletion of the transporter gene *mcyH* in the cynobacterium *Microcystis aeruginosa,* completely abolished microcystin production [108]. In *Steptomyces ghanaensis* ATCC14672, the deletion of two transporter gene pairs, *moeX5/moeP5* and *moeD5/moeJ5*, significantly reduced moenomycin A production [109]. Furthermore, production of the NRPS derived toxin cereulide in *Bacillus cereus* was decreased when the ABC transporter gene *cesCD* was deteled [89].

In the case of Tba, we could show experimentally that the observed impact on balhimycin biosynthesis is not due to transcriptional or general metabolic changes. This parallels observations from other systems, where transcriptional regulation could not explain the loss or

reduction of compound production, as in microcystin and moenomycin A biosynthesis [108,109]. These patterns suggest an additional, more direct layer of regulation. To explore this, we employed proximity-dependent labeling using a Tba^{TID} fusion protein to identify proteins in the immediate vicinity of Tba. This approach provided the first evidence that Tba spatially co-localizes with the balhimycin biosynthetic machinery. Most notably, several core biosynthetic enzymes were found in close proximity to Tba, suggesting a possible physical or functional interaction between the transporter and the NRPS assembly line.

While direct interaction between Tba and the NRPS machinery remains to be definitively proven, our data strongly suggest that Tba plays a structural or regulatory role in facilitating balhimycin biosynthesis. This hypothesis is supported by the fact that a catalytically inactive variant of Tba (Tba^{E545Q}) failed to restore balhimycin production, despite being present. This indicates that ATP-driven conformational changes of Tba are likely essential, not only for substrate export but also for triggering or maintaining productive interactions with the biosynthetic machinery. One possible mechanism involves transient, ATP-dependent conformational states of Tba that enable such interactions. The presence of conserved short C-terminal regions in Tba might contribute to mediating these interactions, although experimental evidence for this remains to be established.

Direct interactions between ABC transporters and biosynthetic enzymes have been demonstrated in other systems. In cereulide biosynthesis, Gacek et al. could show by bacterial two-hybrid assays and microscopy that the NRPSs CesAB physically interact with the ABC transporter CesCD [89]. In the case of the ribosomally synthesized and post-translationally modified peptide (RiPP) nisin, the transporter NisT was shown to directly interact with the modification enzymes cyclase and dehydratase NisBC [90,91].

Even though direct evidence for interactions of Tba with the NRPS machinery is still lacking, our findings support the existence of a similar scenario in the balhimycin biosynthesis and possibly in other GPA-producing systems.

Altogether, these insights revealed an unsuspected dual role of Tba, not only as an exporter of balhimycin but also as a regulatory component essential for its biosynthesis. This opens new perspectives on the traditional role of transporters in GPA biosynthesis and points to the potential for modulating the biosynthesis, which could serve as a basis for further biotechnological improvements in production yields of GPA.

Structure driven interplay of the balhimycin biosynthesis enzymes directs the cellular assembly of the biosynthetic machinery

Given Tba's apparent influence on both export and biosynthetic processes, we next turned our attention to the broader organization of the balhimycin biosynthetic machinery. Specifically, we sought to dissect how protein–protein interactions within the NRPS assembly line and its associated tailoring enzymes contribute to the spatial and functional organization of GPA biosynthesis. The ability to generate novel GPA variants depends on a detailed understanding of these complex biosynthetic machineries. While genetic and biochemical studies of balhimycin biosynthesis have provided valuable insights, the dynamic nature of GPA NRPS systems presents ongoing challenges.

In our study, we demonstrate through proximity-dependent biotinylation that key biosynthetic enzymes are closely associated with Tba, suggesting compartmentalized GPA biosynthesis. We investigated molecular interactions within the balhimycin biosynthetic complex, aiming to uncover previously unrecognized roles and advance strategies for GPA production and engineering.

A key focus of our research was the investigation of the interactions between the NRPSs BpsA-C and the functional importance of docking domains (DDs) in maintaining these interactions. While *in silico* analyses based on sequence homology previously predicted the occurrence of DDs in GPA NRPSs, our work provides the first experimental evidence of their functional role in GPA biosynthesis, both *in vitro* and *in vivo.* This was done by measuring the protein binding affinity of DD pairs, which revealed binding mediated by the respective DD and putative involvement of adjacent domains. In addition, deletion mutants were generated, lacking the DD regions, which showed lower production levels, however not entirely abolished. To date, two types of DD have been characterized exclusively in NRPS systems: the communication (COM) domains [70] and peptide-antimicrobial-Xenorhabdus (PAX) domains [75]. An additional class has been described in PKS-NRPS hybrids are also described [77,110].

Our findings indicated that DDs in balhimycin biosynthesis enhance NRPS assembly by stabilizing the interfaces between BpsA-BpsB and BpsB-BpsC, thereby facilitating effective substrate transfer between donor and acceptor modules. Although DD deletion reduced binding affinity, residual interactions persisted, likely due to contributions from residues in adjacent catalytic domains. This finding parallels the studies on PaxA/PaxB system [75], where interactions are promoted by a PCP domain located upstream of the cDD and a C-domain downstream of the nDD. Whether similar structural principles underlie the DDs in GPA biosynthesis remains to be elucidated.

Beyond NRPS-NRPS interactions, this study emphasizes the significance of interactions between NRPSs and modification enzymes, notably the P450 monooxygenases OxyA, OxyB, and OxyC, along with the halogenase BhaA. These interactions are vital for halogenation and cross-linking steps in balhimycin biosynthesis. Our investigations validated the interactions between OxyA and OxyC with the X-domain of BpsC, which supports previous studies that identified the X-domain as a specific recruiting interface for P450s on NRPSs, unique in GPAs [62]. Nonetheless, OxyB did not interact with the X-domain in our *in vivo* system using the B2H assay, indicating that OxyB interacts through a distinct interface on NRPSs. This aligns with earlier findings showing that OxyB catalyzes the first C–O–D cross-link on PCP-bound hexapeptides before incorporation of the seventh amino acid, (S)-3,5-dihydroxyphenylglycine. Examples are presented in studies examining the function and cascade of P450 monooxygenases in GPA biosynthesis [63,65,111,112]. However, P450 monooxygenases from other GPA pathways like teicoplanin, A47934 and chloroerymomycin show substrate affinities only when the X-domain is present, pointing out differences between P450 monooxygenases in GPA pathways [62].

P450–PCP interactions have also been reported in other NRPS pathways. For instance, structural studies have demonstrated the formation of PCP–P450 complexes in skyllamycin biosynthesis, essential for generating β-hydroxylated precursors [85]. The absence of OxyB binding in our assays may be due to experimental constraints, such as the lack of PCP-bound hexapeptides, the presumed natural substrate. Similar limitations may also explain the low-confidence structural predictions from AlphaFold3, which could not account for substrate-loaded PCP domains.

We also present the first evidence of interaction between the halogenase BhaA and the NRPS machinery. This was demonstrated through co-migration with BpsA–C and confirmed by proximity-dependent biotinylation, revealing BhaA's spatial proximity to BpsC. Although no structural data currently exist on NRPS-halogenase interactions in GPA biosynthesis, genetic and biochemical evidence suggests that chlorination of balhimycin intermediates occurs exclusively on PCP-bound peptides [67].

Taken together, these findings underscore the complexity of protein-protein interactions involved in GPA biosynthesis, which go beyond the NRPS core assembly and include modification enzymes such as P450 monooxygenases and halogenases. These interactions are critical for the stepwise modification of the growing peptide chain and highlight the importance of specific recruitment domains like the X-domain. However, the fidelity and efficiency of NRPS functions are not solely determined by modification enzyme interactions. Another key layer of regulation lies in the activation and selection of amino acid substrates by A-domain processes that often rely on accessory proteins for proper function.

In this context, we next examined the role of the MbtH-like protein MLP_{bal}. MLPs are known to interact with A-domains and contribute to their catalytic function [43].

We provide first evidence that genome-encoded homologs can compensate for the deletion of *mlp_{bal}*, as its deletion did not significantly reduce balhimycin synthesis [113]. This hypothesis was supported by the identification of physical interactions between balhimycin NRPSs and two MLP homologs, MLP_{7561} encoded by a siderophore BGC, and MLP_{7148} encoded by an unidentified NRP BGC. Interestingly, transcriptomic data confirmed expression of these *mlp* genes under balhimycin production conditions. This suggests potential co-expression of multiple BGCs, which may point to simultaneous biosynthesis of their respective products, though this study does not provide direct experimental evidence for such co-activity.

Comparable cross-complementation phenomena have been reported in *Streptomyces scabiei* during thaxtomin A biosynthesis. While deletion of the BGC-encoded *mlp* gene *txtH* led to reduced yields, complete loss of production only occurred when additional non-cognate *mlp* genes were also deleted [114]. Similarly, MLPs encoded from several BGCs in *Streptomyces coelicolor*, including *cdaX* from the CDA BGC and *cchK* from the coelichelin BGC, were shown to complement the function of each other's NRPS systems [115]. Despite the high sequence similarity among MLPs, the *in silico* and mutagenesis analyses suggested that specific motifs are required for the functional interaction between MLP_{bal} or its homologs and balhimycin NRPSs. Previous attempts to universally exchange MLPs in different NRPS systems have often yielded inconsistent results, underlining the need for a deeper understanding of MLP-NRPS compatibility [116,117].

This study thus highlights the important significance of preserving these interactions for effective GPA biosynthesis. The gained insights may inform future approaches to engineer biosynthetic pathways, either to improve enzyme functionality, develop novel GPA analogues, or optimize microbial production systems for biotechnological applications.

The GPA biosynthesis is spatially organized at the membrane

The observed network of protein interactions hints at a level of spatial organization within the cell, raising the question of whether balhimycin biosynthesis might be compartmentalized in a manner analogous to other secondary metabolite systems. Although bacteria lack membrane-bound organelles characteristic of eukaryotic cells, they nonetheless exhibit a degree of functional compartmentalization, often mediated through specialized membrane structures [118]. The export of secondary metabolites in bacteria involves complex membrane-embedded transport systems. These systems mediate the translocation of bioactive compounds across the membrane [105].

Recent studies suggest that biosynthetic enzymes of secondary metabolites interact with their respective transporters, indicating that biosynthesis may occur at or near the membrane, as demonstrated by cereulide and nisin biosynthesis. Both biosynthetic machineries were demonstrated to co-localize and form complexes with the cognate ABC transporters [89–91].

In this study, multiple lines of evidence suggest that Tba may be integrated into a functional membrane microdomain (FMM), where it serves as an anchor for the balhimycin biosynthetic machinery, including the associated modification enzymes.

Our *in vivo* trans-complementation experiments using transporters from various GPA BGCs, together with proximity-dependent biotinylation assays and the protein-protein interaction analysis by BN-PAGE and B2H assays, support the hypothesis that balhimycin biosynthesis is spatially organized within a membrane-associated microcompartment (Figure 4), and that this organization may be dependent on catalytically active Tba.

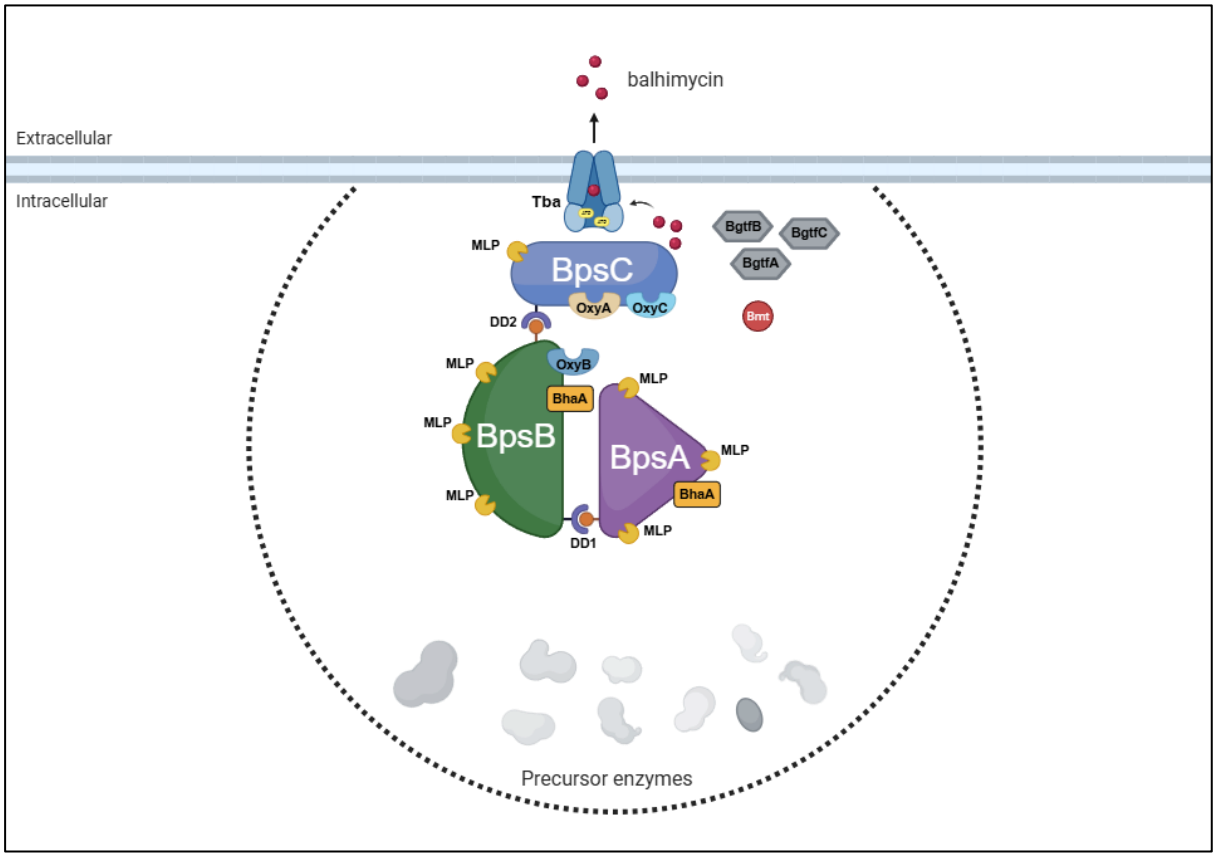


Figure 4: Proposed model of the organization and functional formation of the balhimycin biosynthetic complex. Shown are the enzymes BpsA-C, MLPs, BhaA and OxyA-C forming a core complex. BgtfA-C and Bmt are shown as independent enzymes catalyzing their reactions after release of the balhimycin heptapeptide from the NRPS. The core machinery is in close proximity to Tba. The dashed lines represent the proposed microcompartmentalization of the balhimycin biosynthesis, including the precursor enzymes shown in grey colors.

The identification of a paraslipin protein among the biotinylated interactors further supports this model. Given that paraslipins are known markers of functional FMMs, their proximity to Tba implies a close spatial association with the balhimycin biosynthetic complex and highlights that *A. balhimycina* has the ability to produce proteins associated with the formation of microcompartments. BN-PAGE analyses additionally confirmed the production of both paraslipin and ferritins in *A. balhimycina*. While paraslipins are associated with the formation of membrane microdomains [119], ferritins are known to act as iron-storage shells and may contribute to local metal homeostasis within biosynthetic niches [120].

Evidence indicates that bacteria employ microcompartments to enhance their metabolic activities. Carboxysomes, for example, represent protein-based microcompartments that enhance the efficiency of metabolic reactions, particularly in carbon fixation by photoautotrophs and chemoautotrophs. They achieve this by concentrating CO_2 due to the action of carbonic anhydrase and the selective permeability of the protein shell, thereby enhancing both the rates and specificity of the encapsulated carbon-fixing enzyme, ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) [121]. An extra example is the propanediol-utilizing microcompartment (Pdu MCP) seen in intestinal bacteria such as *Salmonella* [122]. The Pdu MCP facilitates the degradation of 1,2-propanediol and enteric pathogenesis.

Taken together, these findings point toward a potential membrane-associated organization of GPA biosynthesis in *A. balhimycina*, with the transporter Tba acting as a scaffold for the assembly of biosynthetic and modification enzymes. Nevertheless, further studies are required to definitively establish the nature of this microcompartmentalization and to clarify the precise functional role of Tba within this spatially organized biosynthetic system. This integrated biosynthetic model advances our mechanistic understanding of GPA production and provides a valuable framework for rational engineering of antibiotic biosynthesis.

6 Literature

1. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).
2. Staunton, J. & Weissman, K. J. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **18**, 380–416 (2001).
3. Klaenhammer, T. R. Genetics of bacteriocins produced by lactic acid bacteria*. *FEMS Microbiology Reviews* **12**, 39–85 (1993).
4. Kahne, D., Leimkuhler, C., Lu, W. & Walsh, C. Glycopeptide and Lipoglycopeptide Antibiotics. *Chem. Rev.* **105**, 425–448 (2005).
5. Paritala, S. T., Gandhi, G., Agrawal, K., Sengupta, P. & Sharma, N. Glycopeptides: Insights Towards Resistance, Clinical Pharmacokinetics and Pharmacodynamics. *Indian J Microbiol* (2024) doi:10.1007/s12088-024-01273-y.
6. Gavriilidou, A. *et al.* Compendium of specialized metabolite biosynthetic diversity encoded in bacterial genomes. *Nat Microbiol* **7**, 726–735 (2022).
7. Roller, C., Ludwig, W. & Schleifer, K. H. Gram-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes. *Microbiology* **138**, 1167–1175 (1992).
8. Mccormick, M. H., Mcguire, J. M., Pittenger, G. E., Pittenger, R. C. & Stark, W. M. Vancomycin, a new antibiotic. I. Chemical and biologic properties. *Antibiot Annu* **3**, 606–611 (1955).
9. Butler, M. S., Hansford, K. A., Blaskovich, M. A. T., Halai, R. & Cooper, M. A. Glycopeptide antibiotics: Back to the future. *J Antibiot* **67**, 631–644 (2014).
10. Nadkarni, S. R. *et al.* BALHIMYCIN, A NEW GLYCOPEPTIDE ANTIBIOTIC PRODUCED BY Amycolatopsis sp. Y-86, 21022 TAXONOMY, PRODUCTION, ISOLATION AND BIOLOGICAL ACTIVITY. *J. Antibiot.* **47**, 334–341 (1994).

11. Spohn, M. *et al.* Overproduction of Ristomycin A by Activation of a Silent Gene Cluster in Amycolatopsis japonicum MG417-CF17. *Antimicrob Agents Chemother* **58**, 6185–6196 (2014).
12. Kunstmann, M. P., Mitscher, L. A., Porter, J. N., Shay, A. J. & Darken, M. A. LL-AV290, a new antibiotic. I. Fermentation, isolation, and characterization. *Antimicrob Agents Chemother (Bethesda)* **8**, 242–245 (1968).
13. Xu, F. *et al.* A genetics-free method for high-throughput discovery of cryptic microbial metabolites. *Nat Chem Biol* **15**, 161–168 (2019).
14. Nishikiori, T. *et al.* Production by actinomycetes of (S,S)-N,N'-ethylenediamine-disuccinic acid, an inhibitor of phospholipase C. *J Antibiot (Tokyo)* **37**, 426–427 (1984).
15. Spohn, M., Wohlleben, W. & Stegmann, E. Elucidation of the zinc-dependent regulation in Amycolatopsis japonicum enabled the identification of the ethylenediamine-disuccinate ([S,S]-EDDS) genes. *Environmental Microbiology* **18**, 1249–1263 (2016).
16. Kodani, S., Komaki, H., Suzuki, M., Hemmi, H. & Ohnishi-Kameyama, M. Isolation and structure determination of new siderophore albachelin from Amycolatopsis alba. *Biometals* **28**, 381–389 (2015).
17. Seyedsayamdost, M. R., Traxler, M. F., Zheng, S.-L., Kolter, R. & Clardy, J. Structure and Biosynthesis of Amychelin, an Unusual Mixed-Ligand Siderophore from Amycolatopsis sp. AA4. *J. Am. Chem. Soc.* **133**, 11434–11437 (2011).
18. Banskota, A. H. *et al.* Genomic Analyses Lead to Novel Secondary Metabolites. *J Antibiot* **59**, 533–542 (2006).
19. Sensi, P., Margalith, P. & Timbal, M. T. Rifomycin, a new antibiotic; preliminary report. *Farmaco Sci* **14**, 146–147 (1959).
20. Stegmann, E., Frasch, H.-J. & Wohlleben, W. Glycopeptide biosynthesis in the context of basic cellular functions. *Current Opinion in Microbiology* **13**, 595–602 (2010).
21. Yim, G., Thaker, M. N., Koteva, K. & Wright, G. Glycopeptide antibiotic biosynthesis. *J Antibiot* **67**, 31–41 (2014).

22. Nicolaou, K. C., Boddy, C. N. C., Bräse, S. & Winssinger, N. Chemistry, Biology, and Medicine of the Glycopeptide Antibiotics. *Angew. Chem. Int. Ed.* **38**, 2096–2152 (1999).
23. Gavriilidou, A. *et al.* Phylogenetic distance and structural diversity directing a reclassification of glycopeptide antibiotics. Preprint at https://doi.org/10.1101/2023.02.10.526856 (2023).
24. Xu, M. *et al.* Phylogeny-Informed Synthetic Biology Reveals Unprecedented Structural Novelty in Type V Glycopeptide Antibiotics. *ACS Cent. Sci.* **8**, 615–626 (2022).
25. Waglechner, N., McArthur, A. G. & Wright, G. D. Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. *Nat Microbiol* **4**, 1862–1871 (2019).
26. van Wageningen, AndréM. A. *et al.* Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chemistry & Biology* **5**, 155–162 (1998).
27. Pelzer, S. *et al.* Identification and Analysis of the Balhimycin Biosynthetic Gene Cluster and Its Use for Manipulating Glycopeptide Biosynthesis in Amycolatopsis mediterranei DSM5908. *Antimicrobial Agents and Chemotherapy* **43**, 1565–1573 (1999).
28. Xu, L. *et al.* Complete genome sequence and comparative genomic analyses of the vancomycin-producing Amycolatopsis orientalis. *BMC Genomics* **15**, 363 (2014).
29. Sosio, M. *et al.* Organization of the teicoplanin gene cluster in Actinoplanes teichomyceticus. *Microbiology (Reading)* **150**, 95–102 (2004).
30. Li, T.-L. *et al.* Biosynthetic Gene Cluster of the Glycopeptide Antibiotic Teicoplanin: Characterization of Two Glycosyltransferases and the Key Acyltransferase. *Chemistry & Biology* **11**, 107–119 (2004).
31. Thaker, M. N. *et al.* Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat Biotechnol* **31**, 922–927 (2013).
32. Sosio, M., Stinchi, S., Beltrametti, F., Lazzarini, A. & Donadio, S. The Gene Cluster for the Biosynthesis of the Glycopeptide Antibiotic A40926 by *Nonomuraea* Species. *Chemistry & Biology* **10**, 541–549 (2003).

33. Pootoolal, J. *et al.* Assembling the glycopeptide antibiotic scaffold: The biosynthesis of from Streptomyces toyocaensis NRRL15009. *Proceedings of the National Academy of Sciences* **99**, 8962–8967 (2002).
34. Pfeifer, V. *et al.* A polyketide synthase in glycopeptide biosynthesis: the biosynthesis of the non-proteinogenic amino acid (S)-3,5-dihydroxyphenylglycine. *J Biol Chem* **276**, 38370–38377 (2001).
35. Chen, H., Tseng, C. C., Hubbard, B. K. & Walsh, C. T. Glycopeptide antibiotic biosynthesis: Enzymatic assembly of the dedicated amino acid monomer (*S*)-3,5-dihydroxyphenylglycine. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14901–14906 (2001).
36. Hubbard, B. K., Thomas, M. G. & Walsh, C. T. Biosynthesis of L-p-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics. *Chemistry & Biology* **7**, 931–942 (2000).
37. Puk, O. *et al.* Biosynthesis of Chloro-ᴺʟ-Hydroxytyrosine, a Nonproteinogenic Amino Acid of the Peptidic Backbone of Glycopeptide Antibiotics. *J. BACTERIOL.* **186**, (2004).
38. Thykaer, J. *et al.* Increased glycopeptide production after overexpression of shikimate pathway genes being part of the balhimycin biosynthetic gene cluster. *Metab Eng* **12**, 455–461 (2010).
39. Goldfinger, V. *et al.* Metabolic engineering of the shikimate pathway in *Amycolatopsis* strains for optimized glycopeptide antibiotic production. *Metabolic Engineering* **78**, 84–92 (2023).
40. Menges, R., Muth, G., Wohlleben, W. & Stegmann, E. The ABC transporter Tba of *Amycolatopsis balhimycina* is required for efficient export of the glycopeptide antibiotic balhimycin. *Appl Microbiol Biotechnol* **77**, 125–134 (2007).
41. Schäberle, T. F. *et al.* Self-resistance and cell wall composition in the glycopeptide producer *Amycolatopsis balhimycina*. *Antimicrob Agents Chemother* **55**, 4283–4289 (2011).
42. Kilian, R., Frasch, H.-J., Kulik, A., Wohlleben, W. & Stegmann, E. The VanRS Homologous Two-Component System VnlRSAb of the Glycopeptide Producer *Amycolatopsis*

balhimycina Activates Transcription of the vanHAXSc Genes in Streptomyces coelicolor, but not in *A. balhimycina*. *Microb Drug Resist* **22**, 499–509 (2016).

43. Boll, B., Taubitz, T. & Heide, L. Role of MbtH-like Proteins in the Adenylation of Tyrosine during Aminocoumarin and Vancomycin Biosynthesis. *Journal of Biological Chemistry* **286**, 36281–36290 (2011).
44. Puk, O. *et al.* Glycopeptide biosynthesis in Amycolatopsis mediterranei DSM5908: function of a halogenase and a haloperoxidase/perhydrolase. *Chem Biol* **9**, 225–235 (2002).
45. O'Brien, D. P. *et al.* Expression and assay of an N-methyltransferase involved in the biosynthesis of a vancomycin group antibiotic. *Chem. Commun.* 103–104 (2000) doi:10.1039/A907953J.
46. Stegmann, E., Frasch, H.-J. & Wohlleben, W. Glycopeptide biosynthesis in the context of basic cellular functions. *Current Opinion in Microbiology* **13**, 595–602 (2010).
47. Lu, W. *et al.* Characterization of a regiospecific epivancosaminyl transferase GtfA and enzymatic reconstitution of the antibiotic chloroeremomycin. *Proceedings of the National Academy of Sciences* **101**, 4390–4395 (2004).
48. Howard-Jones, A. R. *et al.* Kinetic Analysis of Teicoplanin Glycosyltransferases and Acyltransferase Reveal Ordered Tailoring of Aglycone Scaffold to Reconstitute Mature Teicoplanin. *J. Am. Chem. Soc.* **129**, 10082–10083 (2007).
49. Donadio, S., Sosio, M., Stegmann, E., Weber, T. & Wohlleben, W. Comparative analysis and insights into the evolution of gene clusters for glycopeptide antibiotic biosynthesis. *Mol Genet Genomics* **274**, 40–50 (2005).
50. Sandercock, A. M. *et al.* Biosynthesis of the di-meta-hydroxyphenylglycine constituent of the vancomycin-group antibiotic chloroeremomycin. *Chem. Commun.* 1252–1253 (2001) doi:10.1039/B103668H.
51. Marahiel, M. A., Stachelhaus, T. & Mootz, H. D. Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chem. Rev.* **97**, 2651–2674 (1997).
52. Stachelhaus, T., Schneider, A. & Marahiel, M. A. Rational Design of Peptide Antibiotics by Targeted Replacement of Bacterial and Fungal Domains. *Science* **269**, 69–72 (1995).

53. Süssmuth, R. D. & Mainz, A. Nonribosomal Peptide Synthesis—Principles and Prospects. *Angew Chem Int Ed* **56**, 3770–3821 (2017).
54. Wenski, S. L., Thiengmag, S. & Helfrich, E. J. N. Complex peptide natural products: Biosynthetic principles, challenges and opportunities for pathway engineering. *Synthetic and Systems Biotechnology* **7**, 631–647 (2022).
55. Gulick, A. M. Conformational Dynamics in the Acyl-CoA Synthetases, Adenylation Domains of Non-ribosomal Peptide Synthetases, and Firefly Luciferase. *ACS Chem. Biol.* **4**, 811–827 (2009).
56. Izoré, T. & Cryle, M. J. The many faces and important roles of protein–protein interactions during non-ribosomal peptide synthesis. *Nat. Prod. Rep.* **35**, 1120–1139 (2018).
57. Crosby, J. & Crump, M. P. The structural role of the carrier protein – active controller or passive carrier. *Nat. Prod. Rep.* **29**, 1111–1137 (2012).
58. Stachelhaus, T., Mootz, H. D., Bergendahl, V. & Marahiel, M. A. Peptide Bond Formation in Nonribosomal Peptide Biosynthesis: CATALYTIC ROLE OF THE CONDENSATION DOMAIN*. *Journal of Biological Chemistry* **273**, 22773–22781 (1998).
59. Keating, T. A., Marshall, C. G., Walsh, C. T. & Keating, A. E. The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. *Nat Struct Mol Biol* **9**, 522–526 (2002).
60. Schneider, A. & Marahiel, M. A. Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in Bacillus subtilis. *Arch Microbiol* **169**, 404–410 (1998).
61. Du, L. & Lou, L. PKS and NRPS release mechanisms. *Nat. Prod. Rep.* **27**, 255–278 (2010).
62. Haslinger, K., Peschke, M., Brieke, C., Maximowitsch, E. & Cryle, M. J. X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. *Nature* **521**, 105–109 (2015).
63. Bischoff, D. *et al.* The Biosynthesis of Vancomycin-Type Glycopeptide Antibiotics—A Model for Oxidative Side-Chain Cross-Linking by Oxygenases Coupled to the Action of Peptide Synthetases. *ChemBioChem* **6**, 267–272 (2005).

64. Hadatsch, B. *et al.* The Biosynthesis of Teicoplanin-Type Glycopeptide Antibiotics: Assignment of P450 Mono-Oxygenases to Side Chain Cyclizations of Glycopeptide A47934. *Chemistry & Biology* **14**, 1078–1089 (2007).
65. Stegmann, E. *et al.* Genetic analysis of the balhimycin (vancomycin-type) oxygenase genes. *Journal of Biotechnology* **124**, 640–653 (2006).
66. Gerhard, U., Mackay, J. P., Maplestone, R. A. & Williams, D. H. The role of the sugar and chlorine substituents in the dimerization of vancomycin antibiotics. *J. Am. Chem. Soc.* **115**, 232–237 (1993).
67. Kittilä, T. *et al.* Halogenation of glycopeptide antibiotics occurs at the amino acid level during non-ribosomal peptide synthesis. *Chem. Sci.* **8**, 5992–6004 (2017).
68. Schmartz, P. C. *et al.* Substituent Effects on the Phenol Coupling Reaction Catalyzed by the Vancomycin Biosynthetic P450 Enzyme OxyB. *Angewandte Chemie International Edition* **51**, 11468–11472 (2012).
69. Zhao, Y., Ho, Y. T. C., Tailhades, J. & Cryle, M. Understanding the Glycopeptide Antibiotic Crosslinking Cascade: *In Vitro* Approaches Reveal the Details of a Complex Biosynthesis Pathway. *ChemBioChem* **22**, 43–51 (2021).
70. Hahn, M. & Stachelhaus, T. Selective interaction between nonribosomal peptide synthetases is facilitated by short communication-mediating domains. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15585–15590 (2004).
71. Hahn, M. & Stachelhaus, T. Harnessing the potential of communication-mediating domains for the biocombinatorial synthesis of nonribosomal peptides. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 275–280 (2006).
72. Jaremko, M. J., Davis, T. D., Corpuz, J. C. & Burkart, M. D. Type II non-ribosomal peptide synthetase proteins: structure, mechanism, and protein–protein interactions. *Nat. Prod. Rep.* **37**, 355–379 (2020).
73. Hacker, C. *et al.* Structure-based redesign of docking domain interactions modulates the product spectrum of a rhabdopeptide-synthesizing NRPS. *Nat Commun* **9**, 4366 (2018).

74. Kegler, C. & Bode, H. B. Artificial Splitting of a Non-Ribosomal Peptide Synthetase by Inserting Natural Docking Domains. *Angew Chem Int Ed* **59**, 13463–13467 (2020).
75. Watzel, J., Hacker, C., Duchardt-Ferner, E., Bode, H. B. & Wöhnert, J. A New Docking Domain Type in the Peptide-Antimicrobial-Xenorhabdus Peptide Producing Nonribosomal Peptide Synthetase from *Xenorhabdus bovienii*. *ACS Chem. Biol.* **15**, 982–989 (2020).
76. Watzel, J., Duchardt-Ferner, E., Sarawi, S., Bode, H. B. & Wöhnert, J. Cooperation between a T Domain and a Minimal C-Terminal Docking Domain to Enable Specific Assembly in a Multiprotein NRPS. *Angewandte Chemie International Edition* **60**, 14171–14178 (2021).
77. Dowling, D. P. *et al.* Structural elements of an NRPS cyclization domain and its intermodule docking domain. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12432–12437 (2016).
78. Sundlov, J. A., Shi, C., Wilson, D. J., Aldrich, C. C. & Gulick, A. M. Structural and Functional Investigation of the Intermolecular Interaction between NRPS Adenylation and Carrier Protein Domains. *Chemistry & Biology* **19**, 188–198 (2012).
79. Sundlov, J. A. & Gulick, A. M. Structure determination of the functional domain interaction of a chimeric nonribosomal peptide synthetase from a challenging crystal with noncrystallographic translational symmetry. *Acta Cryst D* **69**, 1482–1492 (2013).
80. Lai, J. R., Fischbach, M. A., Liu, D. R. & Walsh, C. T. A protein interaction surface in nonribosomal peptide synthesis mapped by combinatorial mutagenesis and selection. *Proceedings of the National Academy of Sciences* **103**, 5314–5319 (2006).
81. Frueh, D. P. *et al.* Dynamic thiolation–thioesterase structure of a non-ribosomal peptide synthetase. *Nature* **454**, 903–906 (2008).
82. Miller, B. R., Drake, E. J., Shi, C., Aldrich, C. C. & Gulick, A. M. Structures of a Nonribosomal Peptide Synthetase Module Bound to MbtH-like Proteins Support a Highly Dynamic Domain Architecture. *Journal of Biological Chemistry* **291**, 22559–22571 (2016).
83. Schomer, R. A. & Thomas, M. G. Characterization of the Functional Variance in MbtH-like Protein Interactions with a Nonribosomal Peptide Synthetase. *Biochemistry* **56**, 5380–5390 (2017).

84. Baltz, R. H. Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery. *J Ind Microbiol Biotechnol* **38**, 1747–1760 (2011).

85. Haslinger, K. *et al.* The Structure of a Transient Complex of a Nonribosomal Peptide Synthetase and a Cytochrome P450 Monooxygenase. *Angewandte Chemie International Edition* **53**, 8518–8522 (2014).

86. Greening, C. & Lithgow, T. Formation and function of bacterial organelles. *Nat Rev Microbiol* **18**, 677–689 (2020).

87. Drews, G. & Niklowitz, W. Beiträge zur Cytologie der Blaualgen. *Archiv. Mikrobiol.* **24**, 147–162 (1956).

88. Kerfeld, C. A. *et al.* Protein Structures Forming the Shell of Primitive Bacterial Organelles. *Science* **309**, 936–938 (2005).

89. Gacek-Matthews, A. *et al.* Beyond Toxin Transport: Novel Role of ABC Transporter for Enzymatic Machinery of Cereulide NRPS Assembly Line. *mBio* **11**, e01577-20 (2020).

90. Chen, J., van Heel, A. J. & Kuipers, O. P. Subcellular Localization and Assembly Process of the Nisin Biosynthesis Machinery in Lactococcus lactis. *mBio* **11**, 10.1128/mbio.02825-20 (2020).

91. Chen, J., van Heel, A. J. & Kuipers, O. P. Visualization and Analysis of the Dynamic Assembly of a Heterologous Lantibiotic Biosynthesis Complex in Bacillus subtilis. *mBio* **12**, 10.1128/mbio.01219-21 (2021).

92. Shawky, R. M. *et al.* The border sequence of the balhimycin biosynthesis gene cluster from *Amycolatopsis balhimycina* contains bbr, encoding a StrR-like pathway-specific regulator. *J Mol Microbiol Biotechnol* **13**, 76–88 (2007).

93. Herbst, D. A. *et al.* Structural Basis of the Interaction of MbtH-like Proteins, Putative Regulators of Nonribosomal Peptide Biosynthesis, with Adenylating Enzymes. *Journal of Biological Chemistry* **288**, 1991–2003 (2013).

94. Nahoum, V., Spector, S. & Loll, P. J. Structure of ristocetin A in complex with a bacterial cell-wall mimetic. *Acta Crystallographica Section D* **65**, 832–838 (2009).

95. Loll, P. J. *et al.* Vancomycin forms ligand-mediated supramolecular complexes. *J Mol Biol* **385**, 200–211 (2009).

96. Schäfer, M., Schneider, T. R. & Sheldrick, G. M. Crystal structure of vancomycin. *Structure* **4**, 1509–1515 (1996).

97. Lehmann, C., Bunkóczi, G., Vértesy, L. & Sheldrick, G. M. Structures of Glycopeptide Antibiotics with Peptides that Model Bacterial Cell-Wall Precursors. *Journal of Molecular Biology* **318**, 723–732 (2002).

98. Kuipers, A. *et al.* NisT, the Transporter of the Lantibiotic Nisin, Can Transport Fully Modified, Dehydrated, and Unmodified Prenisin and Fusions of the Leader Peptide with Non-lantibiotic Peptides *. *Journal of Biological Chemistry* **279**, 22176–22182 (2004).

99. Romano, M. *et al.* Structural Basis for Natural Product Selection and Export by Bacterial ABC Transporters. *ACS Chem. Biol.* **13**, 1598–1609 (2018).

100. Gavriilidou, A. *et al.* Animating insights into the biosynthesis of glycopeptide antibiotics. *Current Opinion in Microbiology* **82**, 102561 (2024).

101. Dassa, E. & Bouige, P. The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms. *Res Microbiol* **152**, 211–229 (2001).

102. Bouige, P., Laurent, D., Piloyan, L. & Dassa, E. Phylogenetic and functional classification of ATP-binding cassette (ABC) systems. *Curr Protein Pept Sci* **3**, 541–559 (2002).

103. Saurin, W., Hofnung, M. & Dassa, E. Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *J Mol Evol* **48**, 22–41 (1999).

104. Gerovac, M. & Tampé, R. Control of mRNA Translation by Versatile ATP-Driven Machines. *Trends Biochem Sci* **44**, 167–180 (2019).

105. Méndez, C. & Salas, J. A. The role of ABC transporters in antibiotic-producing organisms: drug secretion and resistance mechanisms. *Research in Microbiology* **152**, 341–350 (2001).

106. Unsleber, S., Wohlleben, W. & Stegmann, E. Diversity of peptidoglycan structure-Modifications and their physiological role in resistance in antibiotic producers. *Int J Med Microbiol* **309**, 151332 (2019).

107. Stegmann, E., Frasch, H.-J., Kilian, R. & Pozzi, R. Self-resistance mechanisms of actinomycetes producing lipid II-targeting antibiotics. *Int J Med Microbiol* **305**, 190–195 (2015).

108. Pearson, L. A., Hisbergues, M., Börner, T., Dittmann, E. & Neilan, B. A. Inactivation of an ABC transporter gene, mcyH, results in loss of microcystin production in the cyanobacterium Microcystis aeruginosa PCC 7806. *Appl Environ Microbiol* **70**, 6370–6378 (2004).

109. Ostash, B., Doud, E. & Walker, S. ABC transporter genes from Streptomyces ghanaensis moenomycin biosynthetic gene cluster: roles in antibiotic production and export. *Arch Microbiol* **194**, 915–922 (2012).

110. Richter, C. D., Nietlispach, D., Broadhurst, R. W. & Weissman, K. J. Multienzyme docking in hybrid megasynthetases. *Nat Chem Biol* **4**, 75–81 (2008).

111. Bischoff, D. *et al.* The Biosynthesis of Vancomycin-Type Glycopeptide Antibiotics-The Order of the Cyclization Steps. *Angew. Chem. Int. Ed.* **40**, 4688–4691 (2001).

112. Schoppet, M. *et al.* The biosynthetic implications of late-stage condensation domain selectivity during glycopeptide antibiotic biosynthesis. *Chem. Sci.* **10**, 118–133 (2019).

113. Stegmann, E., Rausch, C., Stockert, S., Burkert, D. & Wohlleben, W. The small MbtH-like protein encoded by an internal gene of the balhimycin biosynthetic gene cluster is not required for glycopeptide production. *FEMS Microbiology Letters* **262**, 85–92 (2006).

114. Li, Y., Tahlan, K. & Bignell, D. R. D. Functional Cross-Talk of MbtH-Like Proteins During Thaxtomin Biosynthesis in the Potato Common Scab Pathogen Streptomyces scabiei. *Front. Microbiol.* **11**, 585456 (2020).

115. Lautru, S., Oves-Costales, D., Pernodet, J.-L. & Challis, G. L. MbtH-like protein-mediated cross-talk between non-ribosomal peptide antibiotic and siderophore

biosynthetic pathways in Streptomyces coelicolor M145. *Microbiology* **153**, 1405–1412 (2007).

116. Esquilín-Lebrón, K. J., Boynton, T. O., Shimkets, L. J. & Thomas, M. G. An Orphan MbtH-Like Protein Interacts with Multiple Nonribosomal Peptide Synthetases in Myxococcus xanthus DK1622. *Journal of Bacteriology* **200**, (2018).

117. Mori, S. *et al.* Using MbtH-Like Proteins to Alter the Substrate Profile of a Nonribosomal Peptide Adenylation Enzyme. *ChemBioChem* **19**, 2186–2194 (2018).

118. Roze, L. V., Chanda, A. & Linz, J. E. Compartmentalization and molecular traffic in secondary metabolism: a new understanding of established cellular processes. *Fungal Genet Biol* **48**, 35–48 (2011).

119. Browman, D. T., Hoegg, M. B. & Robbins, S. M. The SPFH domain-containing proteins: more than lipid raft markers. *Trends in Cell Biology* **17**, 394–402 (2007).

120. Arosio, P., Ingrassia, R. & Cavadini, P. Ferritins: A family of molecules for iron storage, antioxidation and more. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1790**, 589–599 (2009).

121. Kerfeld, C. A., Aussignargues, C., Zarzycki, J., Cai, F. & Sutter, M. Bacterial microcompartments. *Nat Rev Microbiol* **16**, 277–290 (2018).

122. Yang, M. *et al.* Decoding the stoichiometric composition and organisation of bacterial metabolosomes. *Nat Commun* **11**, 1976 (2020).