Biotechnological Access to Rare Non-proteinogenic Amino Acids via Metabolic Engineering

Dissertation

zur Erlangung des Grades

des Doktors der Naturwissenschaften

der Naturwissenschaftlich-Technischen Fakultät

Der Universität des Saarlandes

von

Ruslan Zhukovets

Saarbrücken

2026

Tag des Kolloquiums:	17.04.2026
Dekan:	Prof. Dr.-Ing. Dirk Bähre
Berichterstatter:	Prof. Dr. Andriy Luzhetskyy
	Prof. Dr. Christoph Wittmann
Vorsitz:	Prof. Dr. Uli Kazmaier
Akad. Mitglied:	Dr. Stefan Boettcher

Diese Arbeit entstand unter der Anleitung von Prof. Dr. Andriy Luzhetskyy in der Fachrichtung 8.2, Pharmazeutische Biotechnologie der Naturwissenschaftlich-Technischen Fakultät der Universität des Saarlandes von Marz 2021 bis Januar 2026.

Acknowledgements

I am deeply grateful to Prof. Dr. Andriy Luzhetskyy for the opportunity to conduct my PhD research under his supervision. His exceptional combination of academic expertise and entrepreneurial vision continuously inspired innovative ideas at the intersection of fundamental and applied science. One of these ideas became the foundation and inspiration for my project. The blend of scientific curiosity and practical potential was a driving force that helped me maintain motivation through many challenges. I sincerely thank him for his fresh perspectives, his guidance in recognizing what truly matters, and for granting me the freedom to make my own mistakes and learn from them.

I would like to express my sincere gratitude to Prof. Dr. Christoph Wittmann for being my second supervisor. His profound expertise in metabolic engineering and bioprocess development significantly strengthened this project. I am thankful for the professional insights and knowledge he shared, which helped bridge conceptual gaps and and deepen my scientific understanding.

I am very grateful to Dr. Maksym Myronovskyi for sharing his knowledge and for his continuous guidance throughout my PhD. His exceptional ability for critical evaluation not only helped me to design and conduct successful experiments but also served as an inspiring example of precise and logical scientific thinking. I sincerely thank him for the time and effort he devoted to explaining complex concepts in great detail, as well as for his patience and invaluable help in reviewing my manuscripts.

I would like to express my special gratitude to Anja Palusczak for her help and valuable advice in the development of many experimental methods, as well as for her constant support and patience throughout numerous failures and challenges. I also wish to thank all former and current members of AMEG for creating a friendly and supportive working environment, for their helpful discussions, and for the many enjoyable moments.

Exceptional thanks go to my love, Alyona, for her endless support, love, and joy that have accompanied me throughout this entire journey, and for all the wonderful moments we have shared together. I also want to thank my family for their unwavering support in my personal life and for continuously encouraging me to reach new frontiers in my professional development.

Publication

Zhukovets R., Stierhof M., Myronovskyi M., Luzhetskyy A. Biotechnological Access to Rare Cyclic Guanidine Amino Acids via Metabolic Engineering of Corynebacterium glutamicum. *Angewandte Chemie: under review.*

Patent application

Luzhetskyy A., **Zhukovets R**., Myronovskyi M., Stierhof-Wißgott M. (2025) Means and methods for producing arginine-derived non-proteinogenic amino acids (NPAAs) and uses thereof. EP25207928

Abstract

Modern science and industry demand the development of new bioactive molecules capable of addressing global challenges such as antimicrobial resistance, food insecurity, and environmental pollution. A promising strategy lies in expanding the repertoire of molecular building blocks beyond the limits of synthetic chemistry toward nature-derived scaffolds selected through evolution. Non-proteinogenic amino acids, discovered as residues of diverse bioactive natural products, represent potent examples. They possess unique structural and functional features essential for molecular stability and reactivity in biological systems. However, these metabolites are rarely accessible through conventional synthesis and often depend on fossil resources or toxic reagents.

This work establishes a sustainable platform for the microbial production of these valuable building blocks using metabolic engineering of bacterial cell factories. The main focus was on cyclic guanidine–containing amino acids as precursors for antibiotic development. Their biosynthetic pathways were optimized for efficient expression in amino acid–overproducing strains of *Corynebacterium glutamicum*. The engineered strains enabled gram-scale production of target metabolites during fermentation, followed by isolation and purification from the culture broth. The established platform was further extended to additional non-proteinogenic amino acids, demonstrating its versatility and potential for future expansion.

Zusammenfassung

Die moderne Wissenschaft und Industrie benötigen neue bioaktive Moleküle, die antimikrobielle Resistenz, Nahrungsmittelunsicherheit und Umweltverschmutzung adressieren. Eine tragfähige Strategie ist die Erweiterung des Baukastenrepertoires über die Grenzen der klassischen Synthese hinaus hin zu naturbasierten Gerüsten, die evolutionär selektiert wurden. Nicht-proteinogene Aminosäuren, identifiziert als Bausteine zahlreicher Naturstoffe, sind hierfür exemplarisch. Sie besitzen besondere strukturelle und funktionelle Eigenschaften, die Stabilität und Reaktivität in biologischen Systemen erhöhen. Gleichzeitig sind sie konventionell oft schwer zugänglich und an fossile Rohstoffe oder toxische Reagenzien gebunden.

Diese Arbeit etabliert eine nachhaltige Plattform zur mikrobiellen Produktion solcher Bausteine durch Stoffwechselingenieurwesen bakterieller Zellfabriken. Der Schwerpunkt liegt auf zyklisch guanidinhaltigen Aminosäuren als Vorstufen für die Entwicklung neuer Antibiotika. Entsprechende Biosynthesewege wurden für die effiziente Expression in aminosäureüberproduzierenden Stämmen von *Corynebacterium glutamicum* optimiert. Die konstruierten Stämme erreichten in der Fermentation Gramm-pro-Liter-Produktionen, und die Zielmetabolite wurden aus dem Kulturüberstand isoliert und gereinigt. Die Plattform wurde auf weitere nicht-proteinogene Aminosäuren ausgedehnt und zeigt damit Vielseitigkeit sowie Potenzial für eine künftige Erweiterung.

Content

Introduction

General introduction

In recent years, escalating challenges and global crises have placed increasing pressure on the stability and performance of essential sectors such as healthcare, agriculture, and industrial production [1]. Overcoming these challenges requires fundamentally new approaches that can deliver solutions beyond the reach of existing technologies [2].

In therapeutics, antimicrobial agents are foundational to modern medicine. Routine surgeries, intensive care, transplantation, and cancer therapies depend on reliable infection control [3]. Misuse and overuse of antibacterial compounds have accelerated the spread of drug-resistant and multidrug-resistant pathogens [4]. Antimicrobial resistance is now a rapidly developing global threat associated with millions of deaths each year (Fig.1) [5]. The World Health Organization has prioritized the most challenging drug-resistant pathogens and emphasized the urgent need for innovative antibacterial agents that can overcome current resistance mechanisms and preserve the effectiveness of infection control in the future [6].

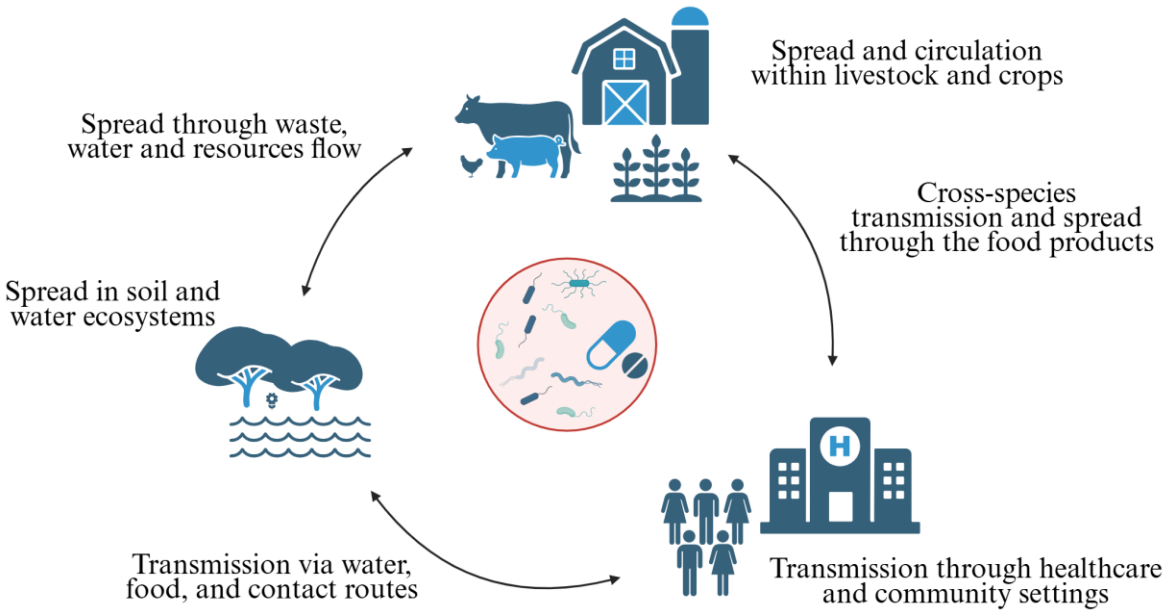


Figure 1. Spread of antimicrobial resistance across humans, agriculture, and the environment.

A parallel challenge exists in agriculture. Chemical pesticides remain essential for protecting crops from harmful organisms. By killing, repelling, or inhibiting insects, weeds, fungi, and other pests, they prevent losses, maintain quality, and support the high yields required to feed a growing population [7]. Yet intensive and widespread use of these chemicals has driven

resistance among target species, contributed to environmental contamination, and reduced biodiversity. Residues of these compounds in water sources and agricultural products have resulted in a negative impact on human health [8]. To sustain productivity while minimizing harm, there is a pressing need for next-generation compounds and bioproducts with higher specificity, reduced toxicity, and improved degradability [9].

Both examples illustrate the urgency for the search of new molecular classes with novel modes of action. Modern discovery pipelines rely on high-throughput and computational technologies (Fig. 2) [10-12].

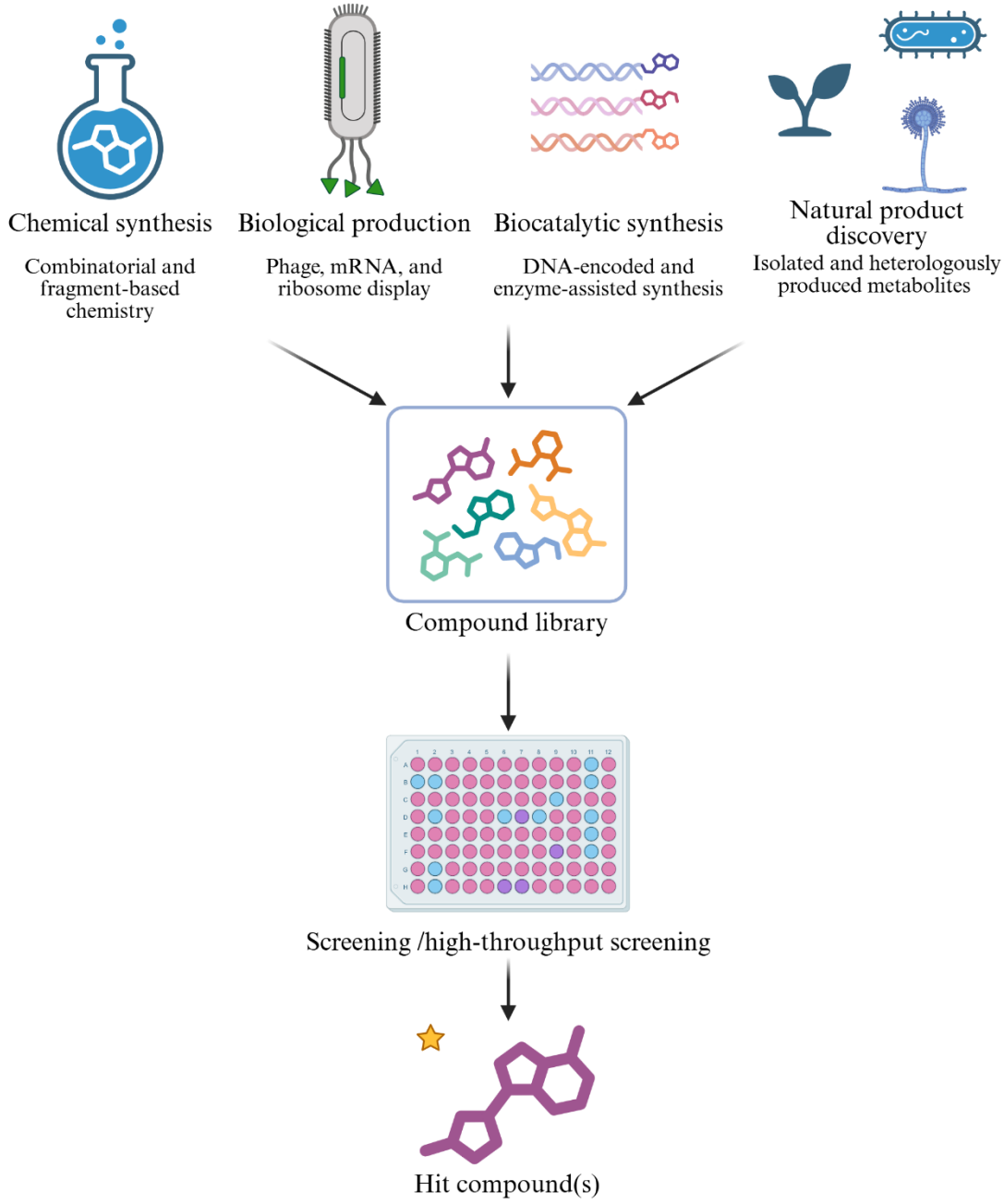


Figure 2. Early-stage drug discovery: generation and screening of compound libraries for hit identification.

Combinatorial chemistry, DNA-encoded libraries, AI-guided molecular design, and fragment-based screening are used to design and synthesize libraries containing millions of compounds [13-16]. High-throughput technologies allow these libraries to be rapidly screened against relevant biological targets, providing starting points for the development of new drug candidates [17]. Automated bioassays, microfluidics, and high-content imaging help to miniaturize the process, increase throughput, and improve assay specificity [18-20]. Machine learning, structural biology, and molecular modelling are employed to narrow down the initial chemical diversity of the libraries, optimize final structures based on in situ target interactions, and thereby reduce both time and cost per hit [21, 22].

The effectiveness of such platforms is constrained by their inputs. The diversity of the available building blocks and the scope of synthetic capabilities define the achievable outcomes and the novelty of what can be discovered [23, 24]. A similar dynamic is observed in materials science, where the performance of advanced materials is determined by the molecular structures of their primary building blocks. The introduction of new monomers has driven breakthroughs in polymer science, where even small changes in monomer structure can produce large differences in material properties [25, 26]. As in therapeutics and agrochemicals, meaningful progress in developing novel compounds with specific modes of action relies on broadening the repertoire of primary molecular components.

Introduction into non-proteinogenic amino acids

Over millions of years, nature has prioritized, tested, and validated a variety of privileged scaffolds, yielding molecules whose chemical architectures are finely tuned for specific biological roles. Many bioactive natural products contain unique residues of non-proteinogenic amino acids (npAAs). These residues frequently determine a compound`s binding specificity, conformational stability, or resistance to metabolic degradation, making them essential for biological activity. Inspired by these natural templates, npAAs have become widely recognized as versatile building blocks in molecular design. Their utilization in drug discovery pipeline expands the chemical diversity and availability of building blocks, enables rational incorporation into defined molecular positions, supports high-throughput screening, and facilitates subsequent optimization of hit compounds to improve pharmacokinetic properties (Fig. 3).

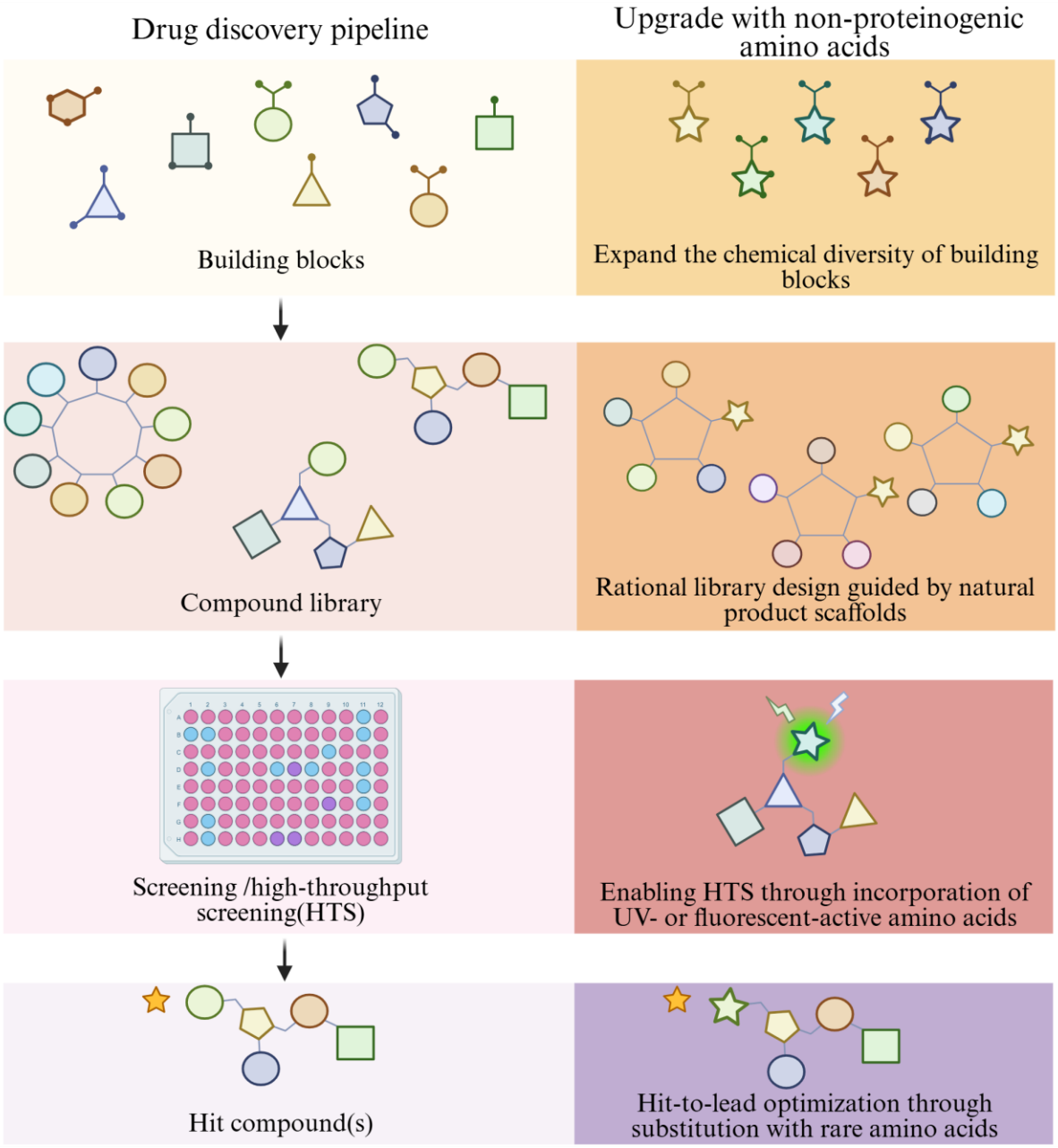


Figure 3. Potential utilization of non-proteinogenic amino acids in drug discovery

In peptide therapeutics, for example, the introduction of npAAs is used to enhance molecular interactions with the target, increase stability in physiological environments, and improve membrane permeability [27]. The effectiveness of this approach is evident in the fact that more than 100 FDA-approved drugs contain at least one npAA, underscoring their value in modern drug discovery [28].

Recent advances in genetic-code-expansion (GCE) technologies have significantly broadened the potential applications of npAAs in biotechnology. Through the use of engineered aminoacyl-tRNA synthetase (aaRS)/tRNA pairs, selected npAAs can be incorporated into peptides and proteins at site-specific positions during translation [29]. This approach utilizes

the most efficient and accurate system for peptide assembly in nature, while expanding its chemical repertoire beyond the 20 canonical amino acids. GCE has been successfully applied in enzyme engineering to create catalysts with enhanced activity, increased stability, or even entirely new chemical functions [30]. A notable example is the substitution of a single phenylalanine residue in the active site of (R)-amine transaminase with a npAA p-benzoyl-L-phenylalanine, which resulted in a 15-fold increase in catalytic activity [31].

In addition to their role as structural elements, certain npAAs serve as precursors in the biosynthesis of high-value bioactive compounds [32]. This role is particularly significant for molecules that cannot be efficiently synthesized by conventional chemical routes. For instance, 5-hydroxytryptophan is an intermediate in the biosynthesis of serotonin and melatonin, both of which are used clinically to modulate mood and regulate sleep patterns [33, 34]. Similarly, structurally related 4-hydroxytryptophan is a key precursor in the biosynthesis of psilocybin, a psychoactive compound currently being investigated for its therapeutic potential in treating depression, anxiety, and other psychiatric disorders [35].

Furthermore, npAAs have also made a significant impact in materials science. A notable example is L-3,4-dihydroxyphenylalanine (L-DOPA), a npAA best known as a biosynthetic precursor of the neurotransmitters dopamine, epinephrine, and norepinephrine. L-DOPA is also a key component of mussel adhesive proteins, where it is responsible for the remarkable ability to strongly bind to surfaces in wet environments. Incorporating L-DOPA into bioadhesive polymers has been shown to enhance interfacial binding in aqueous conditions, strongly advancing applications such as biomedical adhesives, antifouling coatings, and underwater repair materials [36-38].

Taken together, the above examples and many other studies highlight the strong potential of npAAs as molecular building blocks (Fig. 4). By expanding the design alphabet, npAAs introduce chemical structures that nature itself has already prioritized through evolution. Many of these residues were first identified in natural products whose properties have been explored for decades in academic, industrial, and medical research. This accumulated knowledge provides a solid foundation for reliable predictions of npAA interactions in biological contexts and supports their purposeful use in molecular design. In this way, npAAs stand out as evolution-tested scaffolds that can drive the next generation of innovation across medicine, agriculture, and materials science.

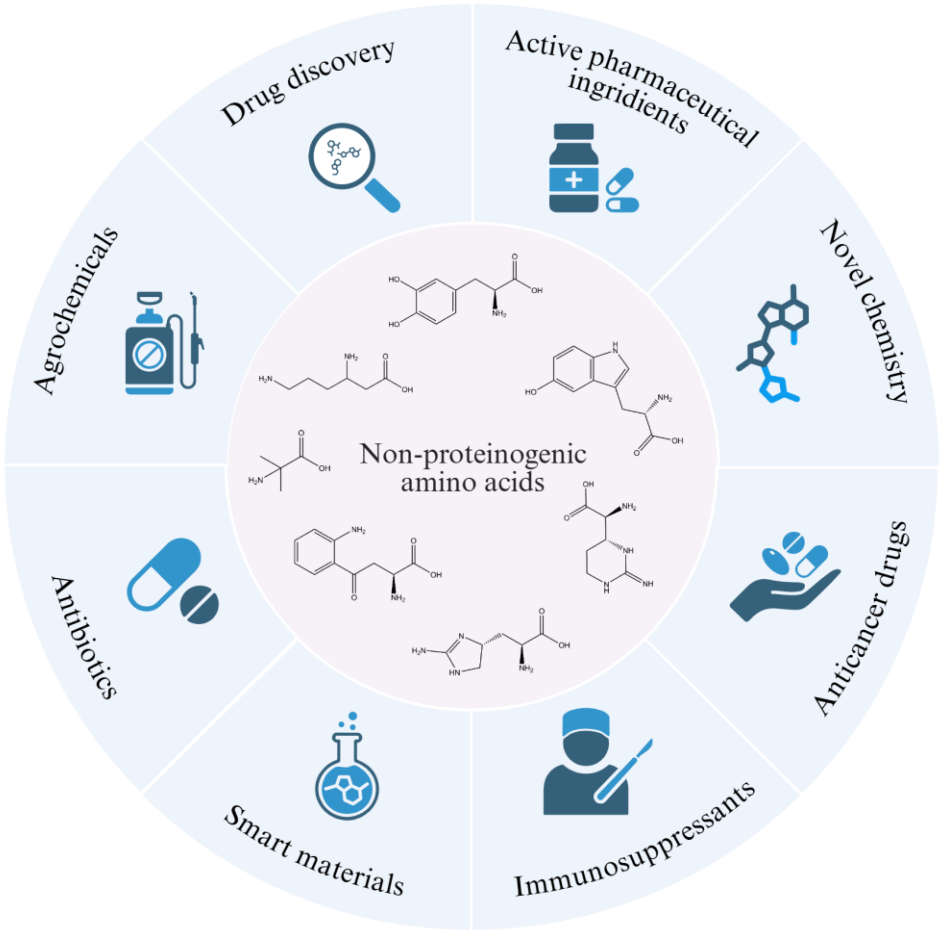


Figure. 4. Application fields of non-proteinogenic amino acids across science, medicine, and industry.

Such molecules will ultimately be evaluated not only by their performance in the laboratory or in practical applications, but also by the feasibility of their production. At present, the primary route to access npAAs remains chemical synthesis. However, this approach is often associated with significant challenges: multistep pathways, low overall yields, and a strong reliance on organic solvents and specialized catalysts. Many npAAs contain multiple stereocenters, making their production and purification even more demanding. As a result, chemical routes to numerous npAAs remain laborious, costly, or altogether impractical. Consequently, despite their great potential, many npAAs are either unavailable for research or commercially non-viable for industrial use.

To address these limitations, biotechnology provides a direct route to align innovation with sustainability. Engineered microbial systems are capable of producing complex molecules that are otherwise difficult or impossible to access through conventional synthesis. Since many npAAs originate from living organisms, nature has already evolved specialized sets of enzymes for their biosynthesis [39]. These pathways typically begin with canonical amino acids, which are transformed into the final npAA through cascades of in vivo enzymatic reactions.

Identifying these enzymes and introducing them into industrially suitable microbial hosts enables the use of microorganisms as efficient living factories (Fig. 5) [40].

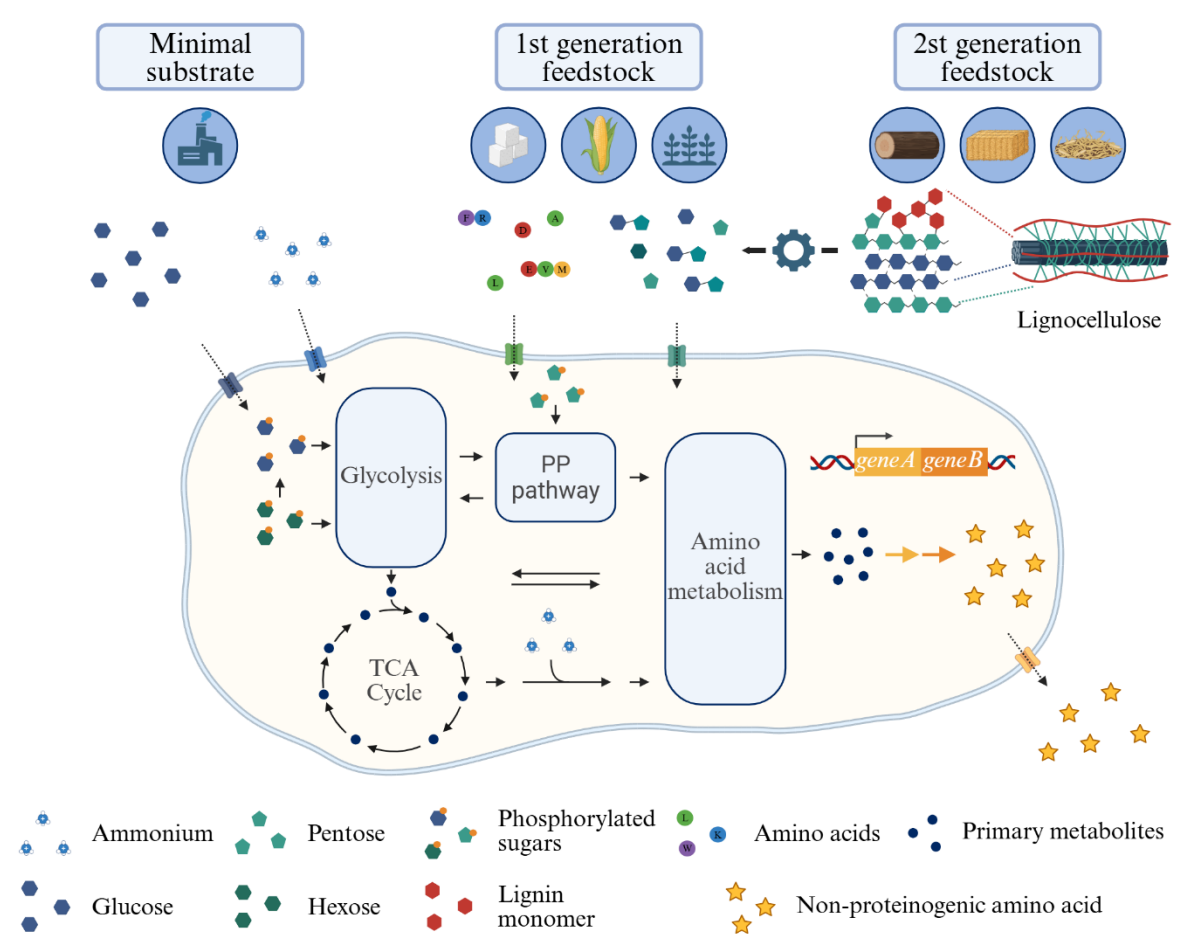


Figure. 5. Bacterial cell factories for sustainable production of high-value metabolites. Bacteria can function as efficient living factories, utilizing simple substrates such as glucose and ammonium to biosynthesize primary metabolites through central metabolic pathways. Genetic engineering enables redirection of carbon and nitrogen flux toward the overproduction of selected metabolites, which serve as precursors for high-value secondary products. Introduction of heterologous biosynthetic genes further expands metabolic capacity, allowing conversion of precursor molecules into non-proteinogenic amino acids. Industrial cultivation commonly employs first-generation feedstocks such as molasses and corn steep liquor as economical carbon and nitrogen sources, while advances in biotechnology now enable the utilization of second-generation lignocellulosic residues after pretreatment, supporting sustainable and circular bio-based production.

The main advantage of this approach lies in its compatibility with renewable feedstocks. Microorganisms can convert simple, sustainable substrates directly into high-value compounds during fermentation [41]. The process takes place in aqueous environments, significantly reducing chemical waste and the ecological footprint compared to traditional synthesis. Downstream purification using tangential flow filtration and ion-exchange chromatography eliminates the need for organic solvents, reinforcing the sustainability of the platform [42].

An additional advantage is the possibility of integrating microbial production with genetic-code-expansion (GCE) technologies. Such coupling enables the direct creation of novel libraries of compounds containing npAAs, while reducing the time and cost associated with producing and purifying each amino acid individually. In this way, biotechnological platforms deliver precisely what innovation pipelines demand: sustainable building blocks that expand chemical diversity while reducing environmental impact.

Outline of the work

Overall aim of this work was to establish a scalable, sustainable platform for the biotechnological production of high-value non-proteinogenic amino acids (npAAs) and their direct application in compound discovery.

Main pipeline:

1. **Candidate prioritization**: Identify potential npAAs as privileged building blocks.
2. **Pathway elucidation**: Define the biosynthetic pathways and key enzymes.
3. **Host selection**: Choose suitable microbial chassis for production.
4. **Pathway implementation**: Introduce and adapt the pathway to the host.
5. **Metabolic flux engineering**: Increase precursor supply and direct flux toward the target npAA.
6. **Bioprocess development**: Develop and optimize fermentation process.
7. **Downstream processing**: Develop sustainable purification and isolation method.

This work focused on arginine-derived cyclic npAAs—enduracididine, hydroxy-enduracididine, capreomycidine, and hydroxy-capreomycidine. These amino acids bear a distinctive cyclic guanidinium motif and occur as essential residues in several potent anti-infective peptides. The initial analysis evaluated their suitability and potential as building blocks and is presented in the theoretical section.

Subsequently, the corresponding biosynthetic pathways and associated enzymes were identified and described, including their origins, putative functions, and relevant homologs. Corynebacterium and Streptomyces were selected as potential bacterial hosts. Pathway introduction strategies were designed accordingly for each host, including cloning strategies, codon optimization, and tuning of transcription and translation levels.

Main efforts were directed at increasing titer and productivity through systematic metabolic engineering of the production host strain, focusing on enhancing precursor supply and developing an optimal fed-batch fermentation process. Finally, the produced compounds were isolated and purified using cation-exchange chromatography with water-based buffers, a method that achieved the required purity while maintaining a low environmental footprint.

Theoretical Basis and Design Strategy

Non-proteinogenic amino acids

Non-proteinogenic amino acids (npAAs) comprise all amino acids other than the 22 canonical residues that are incorporated into proteins by ribosomal translation under the standard genetic code. That is, npAAs exclude the 20 universal residues together with selenocysteine and pyrrolysine and instead cover the full breadth of amino acid chemistry. This term includes stereo chemical variants (L- and D- variants), backbone homologues defined by the position of the amino group relative to the carboxyl group (α, β, γ, δ and higher) and a vast spectrum of structurally modified derivatives.

In this work, the focus is on the biotechnological production of naturally occurring npAAs and their potential as building blocks in molecular engineering. A wide diversity of npAAs is found across many biological systems. These natural scaffolds have undergone evolutionary selection, their frequent presence in bioactive natural products provides strong evidence of their valuable biological functions. Knowledge of these functions forms a solid basis for their targeted and effective application in molecular design.

Reports in the literature suggest that more than 1,000 npAAs have been described, yet these figures are based on scattered evidence rather than a consolidated source[39, 43, 44]. The classification of npAAs is currently fragmented, often depending on how individual studies define or group them. While several attempts have been made to assemble databases of npAAs, most have been tailored to functional predictions and rely heavily on *in silico* assumptions[45-47]. Establishing such a database based on systematic and experimentally grounded resources would provide the critical foundation in this field. A structured and user-friendly platform which would include updated lists of known npAAs, their biological origins, biosynthetic pathways, and molecular models would organize existing knowledge and accelerate further research.

Classifying natural npAAs by biological origin offers early insight into their biosynthesis and functions and helps prioritize candidates as building blocks for molecular design. Biosynthesis of npAAs is tightly coupled to primary metabolism and to the networks surrounding the canonical amino acids. The most widespread class comprises primary metabolic intermediates that are common across taxa (Fig.6). Examples include ornithine and citrulline in the urea–arginine cycle, α-aminoadipate and meso-diaminopimelate in lysine biosynthetic pathways, homoserine and homocysteine in threonine and methionine biosynthesis, D-alanine and D-

glutamine in peptidoglycan assembly, β-alanine in pantothenate and coenzyme A biosynthesis. These npAAs are generally well studied, and where their properties are valuable, strain engineering for overproduction is established or under active development [48-50].

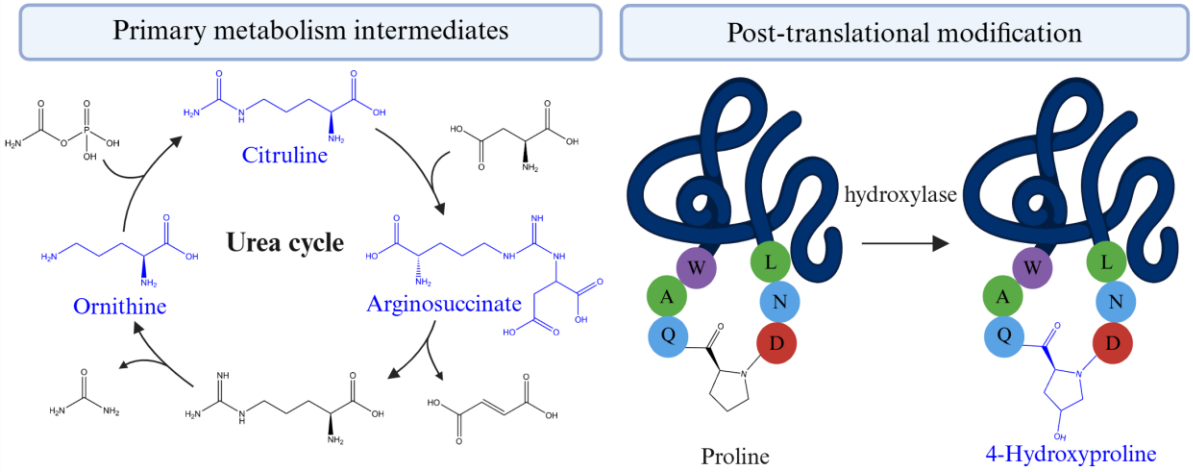


Figure 6. Non-proteinogenic amino acids originating from primary metabolism and post-translational modification. Primary metabolic npAAs (highlighted in blue) are the intermediates or side products of central biochemical pathways and are common across diverse taxa. Examples include ornithine, citrulline, and argininosuccinate, which are the part of the urea cycle but are not incorporated into proteins. A distinct group of npAAs originates from post-translational modifications (PTMs) of canonical residues, such as the hydroxylation of proline to form 4-hydroxyproline.

A second class consists of protein residues that arise through post-translational modifications (PTMs) (Fig. 5) [51]. Common PTMs include phosphorylation, acetylation, glycosylation, amidation, hydroxylation, S-nitrosylation, and methylation [52]. These modifications can stabilize protein folding, modulate activity, or act as regulatory marks. Since such residues are generated only after translation and are not synthesized as free amino acids, there is no direct biosynthetic pathway for their standalone production. For this reason, they are not suitable for the pipeline considered in this work.

A third class comprises free-standing amino acid metabolites and their dedicated biosynthetic intermediates (Fig. 7). These are often species-specific secondary metabolites with clear ecological or physiological roles. For example, N(ε)-acetyl-β-lysine is synthesized by methanogenic archaebacterial as a compatible solute under high salinity [53]. Canavanine, abundant in legume seeds, serves as nitrogen storage and as a defense compound against herbivores, it acts as an arginine antagonist and can be misincorporated into proteins [54]. Ibotenic acid found in certain *Amanita* mushrooms is the glutamate receptor agonist, which causes excitotoxicity in vertebrates but effect glutamate-gated chloride channels permeability in invertebrates[55, 56]. Because these amino acids accumulate in the free state, their production relies on the separated direct biosynthetic pathways. Additionally, these pathways

often include distinctive intermediates whose functions are not yet fully understood. This combination of unique chemical structure, defined function and tractable genetics makes them attractive building-block candidates for biotechnological production pipeline.

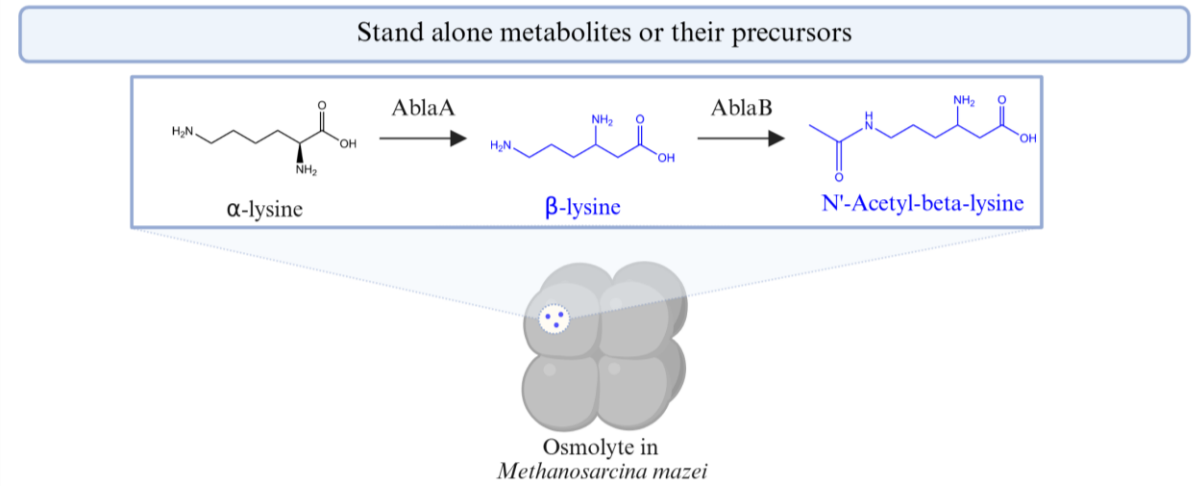


Figure 7. Stand-alone non-proteinogenic amino acids and their biosynthetic intermediates. A distinct class of npAAs (highlighted in blue) comprises free secondary metabolites synthesized independently of translation or peptide assembly systems. These amino acids are typically species-specific and perform ecological or physiological functions, often accumulating in the cytoplasm or extracellular environment. Examples include β-lysine and N-acetyl-β-lysine, which act as compatible solutes in *Methanosarcina mazei* under osmotic stress.

And the fourth class, most significant for the purposes of this work, comprises npAAs that occur as residues within larger secondary metabolites. These are typically incorporated by nonribosomal peptide synthetases (NRPS) or NRPS-hybrid assembly systems (Fig. 8). Such systems often generate cyclic or branched peptides enriched with npAAs, and the resulting products display a wide spectrum of biological activities and pharmacological properties [57]. In most cases, a nonribosomal peptide is synthesized by the set of enzymes encoded within a dedicated biosynthetic gene cluster (BGC). Such a cluster provides the complete enzymatic toolkit that uses common primary metabolites as precursors, converts them into specialized metabolites or npAAs, and incorporates them stepwise through the NRPS assembly line to yield the final active peptide [58]. There are three principal routes by which npAA residues are introduced. In the first case, the BGC contains enzymes that synthesize the npAA as a free metabolite, which is then activated and incorporated by the NRPS. In the second, the NRPS activates a canonical amino acid, and a tailoring domain within the assembly line modifies the residue. In the third, the NRPS incorporates a canonical amino acid residue and the modification occurs after the peptide has been released, through the action of post-NRPS tailoring enzymes.

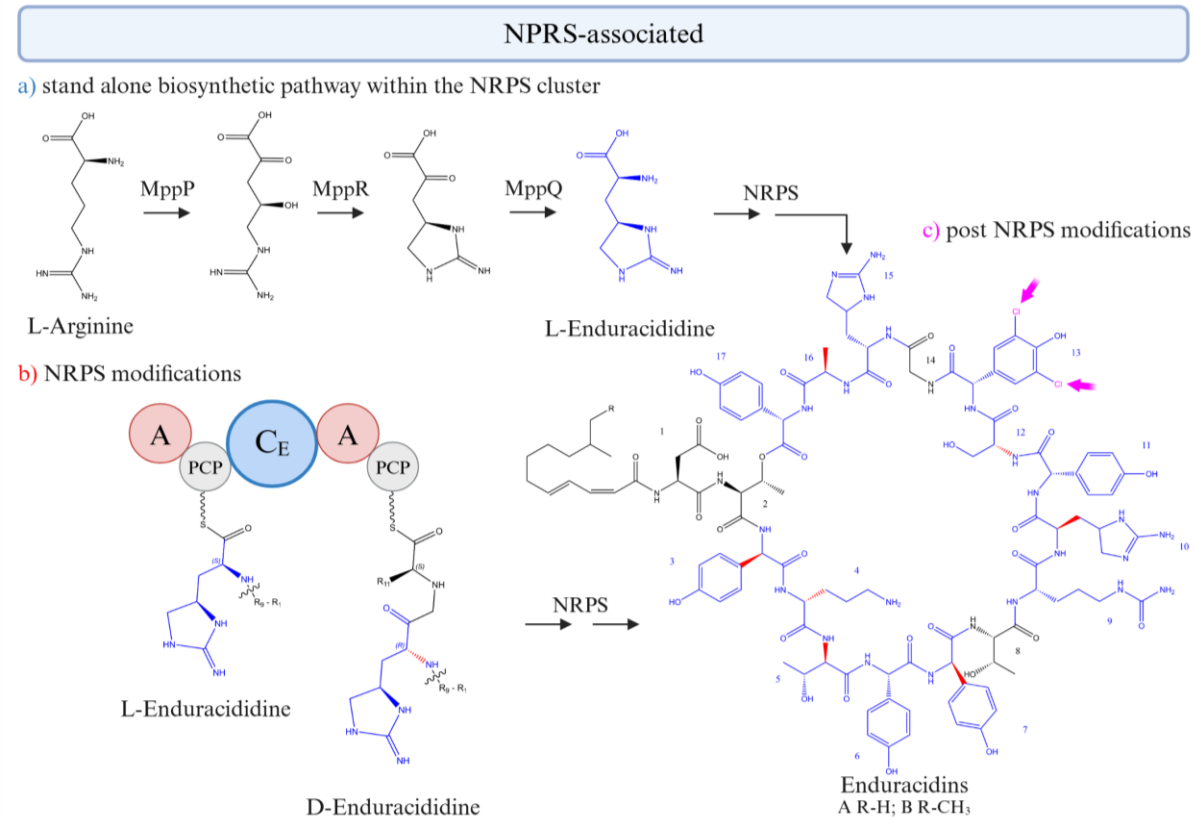


Figure 8. NRPS-associated non-proteinogenic amino acids. NRPS-associated npAAs (highlighted in blue) arise through enzymatic transformations linked to nonribosomal peptide assembly. In the first route, stand-alone enzymes within the NRPS cluster are responsible for the synthesis of the npAA as a free metabolite, as exemplified by the conversion of L-arginine to L-enduracididine via the MppPQR pathway. In the second, the NRPS machinery itself catalyzes on-line modifications, such as epimerization by condensation–epimerization (C/E) domains that generate D-residues (highlighted in red) from their L-counterparts. In the third, post-NRPS tailoring enzymes modify the released peptide, introducing chemical diversity through halogenation (highlighted in pink), hydroxylation, or glycosylation. The biosynthesis of enduracidins integrates all three mechanisms, representing a complete model of NRPS-associated nAA formation.stress.

The npAAs that are first made as free metabolites and later built into natural products represent the most promising candidates within the scope of this study. Secondary metabolites serve specific biological purposes that give the producing organism selective advantage. In many cases, npAA residues are essential structural elements that give the molecule its unique functions. This makes them valuable starting points because nature has already tested their utility. Another reason they stand out is that defined enzyme sets convert primary metabolites into free npAAs. This knowledge opens broad opportunities for metabolic engineering: choosing an optimal host, boosting enzyme performance, and developing efficient producer strains.

Applications of non-proteinogenic amino acids

The leading applications of npAAs are in peptide therapeutics because this field combines scientific innovation with a rapidly expanding market and strong economic return. The market of peptide therapeutics is growing fast, driven by new agents for metabolic, oncological, and neurological diseases [59]. In 2024, two of the top 10 world's best-selling drugs were peptide analogs used to treat type 2 diabetes and obesity: semaglutide (Ozempic, Novo Nordisk, 17.5 billion USD) and tirzepatide (Mounjaro, Eli Lilly, 11.5 billion USD) [60]. Both act on metabolic hormone pathways. Semaglutide is a glucagon-like peptide-1 (GLP-1) analog and GLP-1 receptor agonist. Tirzepatide is a gastric inhibitory polypeptide (GIP) analog with dual agonism at GIP and GLP-1 receptors. These receptors are activated after food intake, trigger insulin secretion, and tune metabolism. In type 2 diabetes this regulation is impaired, which makes GLP-1 and GIP attractive targets. The native hormones GLP-1 and GIP are rapidly degraded by dipeptidyl peptidase-4 (DPP-4), resulting in plasma half-lives of only about two minutes. Substitution of alanine at the DPP-4 cleavage site in GLP-1 with the npAA 2-aminoisobutyric acid prevents proteolytic degradation (Fig. 9). Together with an additional substitution of Lys34 by arginine and attachment of an 18-carbon diacid acyl chain that promotes albumin binding, the half-life of semaglutide was extended to approximately one week [61].

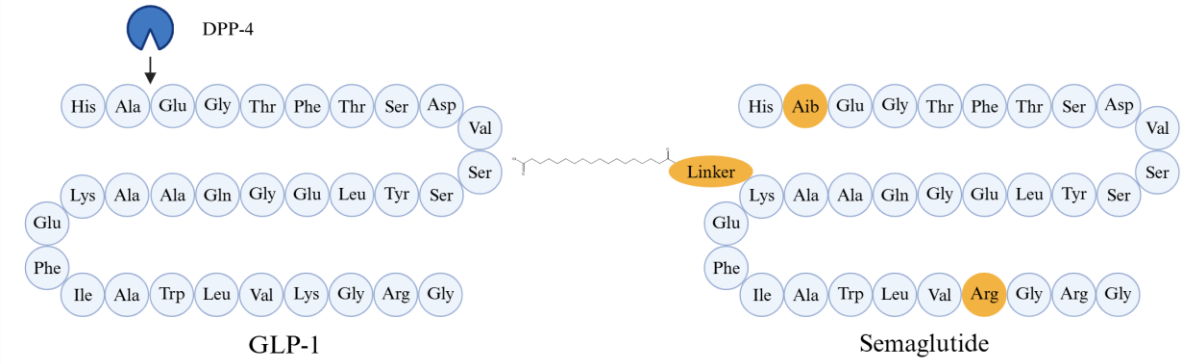


Figure 9. Structural comparison of GLP-1 and semaglutide. The natural hormone GLP-1 consists of 31 amino acids (blue circles) and is rapidly cleaved by DPP-4 at its recognition site. Semaglutide is a modified GLP-1 analogue in which substitutions and chemical modifications (yellow circles) protect the peptide from enzymatic degradation and extend its biological half-life.

These examples demonstrate both the main advantages and the central challenges of peptide therapeutics compared with the more widely established small molecules and biologics. In contrast to small molecules, which still represent the majority of marketed drugs, peptides can reach higher specificity and efficacy because their larger surfaces support strong and selective protein–protein interactions. This often reduces off-target effects and systemic toxicity [62].

Advances in solid-phase peptide synthesis (SPPS) and recombinant protein production have made peptide manufacturing cost-effective and commercially attractive. The major limitations of peptide drugs remain their poor membrane permeability and limited stability in vivo [63]. The example of GLP-1 analogs shows how targeting extracellular receptors can circumvent the permeability barrier. Short half-life is harder to avoid because proteases are abundant and conditions for peptide degradation are common in the physiological environment [64]. These challenges are addressed by rational molecular design, which includes the incorporation of npAAs, as well as strategies such as PEGylation, glycosylation, cyclization, and lipidation [65, 66].

Except enhancing proteolytic stability by masking proteases recognition site, incorporation of npAAs is used to diversify physicochemical properties of the peptides in order to enhance potency, selectivity and stability (Fig. 10) [67].

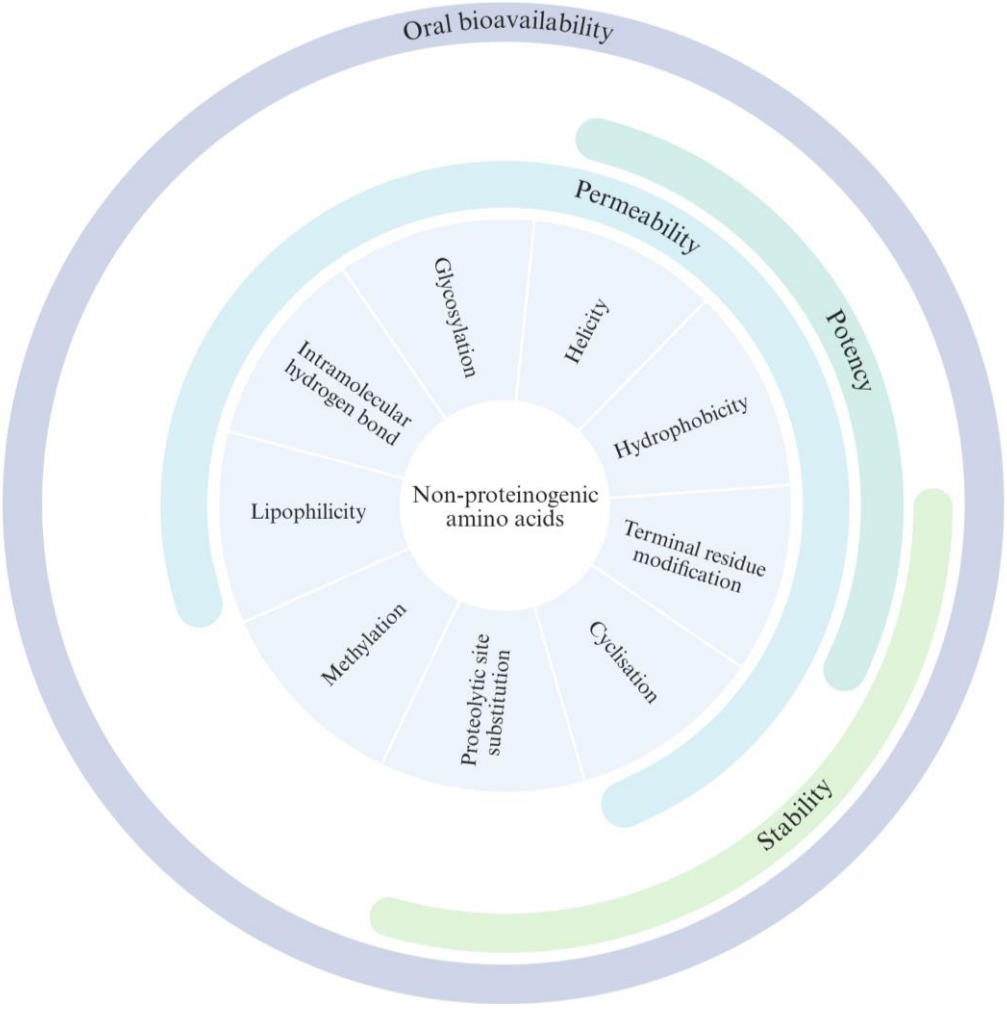


Figure 10. Application of non-proteinogenic amino acids in the development of peptide therapatics (adopted from [27]).

A clear example comes from study on the frog-derived antimicrobial peptide Esculentin-2P. Replacing all positively charged residues with homoarginine (hArg) increased broad-spectrum antibacterial activity and reduced cytotoxicity toward mammalian cells. The extended and more positively charged guanidinium side chains of hArg strengthened electrostatic interactions with negatively charged microbial membranes and improved peptide stability in serum [68].

Incorporation of methylated npAAs can increase local lipophilicity of peptides and enhance protein–protein interactions. A prominent example is provided by studies on compstatin, an inhibitor of complement component 3 (C3). C3 is a central protein of the human immune system and plays a critical role in the response to infections. Uncontrolled activation of C3, however, is associated with autoimmune diseases such as lupus and rheumatoid arthritis [69]. Compstatin itself was identified during phage-display screening as a lead peptide that required further optimization [70]. Substitution of the Val4 residue with 1-methyltryptophan eliminated a hydrogen bond and strengthened hydrophobic interactions with C3, resulting in a 264-fold increase in activity [71].

Peptides tend to adopt regular secondary structures such as α-helices and β-sheets. These conformations are typically stabilized by intramolecular hydrogen bonds, which reduce overall polarity of the peptide surface and thereby enhance membrane penetration [72, 73]. The α,β-unsaturated amino acid α,β-dihydrophenylalanine (ΔPhe), originally discovered in the natural product albonoursin, has been established as a useful tool to induce β-turns in short peptides [74]. This property was successfully exploited in the *de novo* design of short peptides, where ΔPhe residues were introduced in an i/i+3 arrangement to promote formation of a 3_{10}-helix. Additional lysine residues were incorporated to generate a polar face, what resulted in an amphipathic helical structure [75]. While incorporation of D-amino acids, most notably the D- and L-proline pair, is a classical strategy to stabilize β-hairpins connecting two β-strands [76].

All these examples illustrate how incorporation of npAAs can be used to fine-tune peptide chemistry, optimize intermolecular interactions, introduce new functionalities, stabilize secondary structure, and enhance stability or membrane penetration. Such properties have traditionally been exploited through rational design, site-specific substitution of residues based on structural and functional knowledge. These strategies are now increasingly complemented by computational approaches, which enable prediction of how a proposed sequence folds and interacts with the desired target prior to synthesis. This allows many candidate variants to be

evaluated *in silico* and a small, prioritized set to be selected for experimental validation, thereby accelerating the development of peptide therapeutics and peptide-based materials. [77].

A major advance in this direction was achieved with the development of deep learning system AlphaFold, which revolutionized protein structure prediction. Trained on known structures from the Protein Data Bank (PDB), it predicts 3D protein conformations from amino acid sequences with unprecedented accuracy[21]. Reliable structure prediction is essential for exploring structure–activity relationships (SAR) and subsequent design of novel compounds with specific functions [78]. However, AlphaFold does not currently support npAAs. In contrast, RoseTTAFold, a three-track neural network that simultaneously processes sequence, distance maps, and 3D coordinates, allows users to manually parameterize npAAs, predict peptide structures containing them, and model corresponding protein–protein interactions [79].

The predictive accuracy of these systems strongly depends on the amount and quality of available structural data. While they provide reliable models for canonical amino acid sequences, the limited representation of npAAs in training datasets makes their application to non-standard residues challenging [66]. Another important point is that these models build structures directly from sequence, which means that the incorporation of npAAs requires additional manual validation. At present, predicted structures and PPIs are mainly used as draft models, where the integration of npAAs is decided manually based on principles of peptide engineering. Although these tools already provide valuable insights into peptide structure and can substantially reduce the number of experiments required, careful validation, explicit parameterization of npAAs, and iterative cycles of experimental testing and model refinement are still necessary [77, 80].

A complementary generative approach is offered by diffusion-based models. RFdiffusion can create entirely new peptide backbones guided by predefined binding targets or structural motifs. Inspired by image generators such as DALL-E, it starts from random, “diffused” coordinates and gradually reconstructs ordered structures through iterative denoising steps (Fig. 11) [81]. The output is a 3D backbone model, representing the secondary structure where individual elements (α-helices, β-strands, loops) are arranged to fit the imposed geometric constraints and interaction requirements. This designed backbone can then be converted into a compatible amino acid sequence using ProteinMPNN, a message-passing neural network that assigns amino acid residues expected to fold into the generated structure [82].

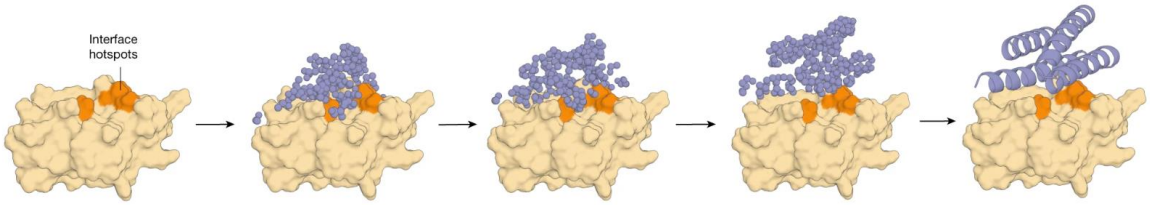


Figure 11. Principle workflow of the RFdiffusion model for generating protein binders. (adopted from [81]).

This generative strategy has already been successfully applied to the design of protein-binding short macrocyclic peptides. Cyclic peptides of 5–14 residues, inspired by many natural and clinically approved examples, are considered a promising therapeutic class because their cyclic structure provides greater stability and potency compared with linear peptides [66, 83]. The designed macrocycles displayed high-affinity binding to γ-aminobutyric acid type A receptor–associated protein (GABARAP), which mediates autophagy, and Rhombotarget A (RbtA), a cell-surface protein of *Acinetobacter baumannii*. Their experimental X-ray crystal structures closely matched the computational predictions [84]. The same study also aimed to extend the approach to npAAs, which is feasible through ProteinMPNN but would require custom residue parameterization and supporting experimental data.

The examples discussed above show that peptide structure optimization and chemical fine-tuning require a suitable lead compound as a starting point. In some cases, this lead may be a naturally derived peptide such as a nonribosomal product or a hormone that lacks stability or permeability to become a potent drug candidate. In other cases, de novo designed scaffolds provide the initial framework, which are further optimised to enhance activity, stability or specificity. Another major strategy for generating such leads is high-throughput screening of large peptide libraries [85].

Two main technologies are used to create the diverse labeled peptide libraries: phage display and mRNA display [86-88]. In phage display, DNA fragments encoding peptide variants are inserted into a phagemid vector that carries the genes required for phage assembly. The vector is introduced into a bacterial host, which produces phage particles displaying the peptide variants on their surface. These particles can then be screened in binding assays. Because each phage particle carries both the displayed peptide and its corresponding DNA sequence inside the capsid, hit peptides can be directly identified by sequencing [89].

In contrast, mRNA display libraries are generated entirely in vitro. First, a variant DNA library is transcribed into mRNA. Each mRNA molecule is then linked at its 3′ end to a DNA-puromycin linker. During in vitro translation, when the ribosome reaches the 3′ end, puromycin

enters the ribosomal A-site as a mimic of aminoacyl-tRNA and forms a covalent bond between the mRNA and the translated peptide. As in phage display, these libraries can be screened in binding assays, and hits can be identified by amplifying and sequencing the bound mRNA [90]. While phage display is limited by bacterial transformation efficiency, mRNA display can achieve library sizes of up to 10^{13} variants, about 1,000-fold greater than phage display [91].

Such libraries offer great potential for discovering new leads; however, their generation is restricted by the translational machinery and thus limited to the 20 canonical amino acids. A revolutionary solution emerged with the development of genetic code expansion (GCE) [92]. GCE was first implemented for *in vivo* applications using a naturally occurring tRNA/synthetase pair from *Methanosarcina* species [93]. In nature, pyrrolysyl-tRNA synthetase (PylRS) specifically charges $tRNA^{Pyl}$ with pyrrolysine. This tRNA carries a CUA anticodon, allowing it to recognize the amber stop codon (UAG) in mRNA. Instead of terminating translation, the UAG codon is suppressed and pyrrolysine is incorporated into the growing peptide chain [94]. Later, PylRS was modified and engineered to recognize substrates beyond pyrrolysine, which opened the way to incorporate a wide range of npAAs [95]. This system was subsequently adapted into nonsense-suppressing aaRS/tRNA pairs that remain orthogonal to the native translational machinery of *E. coli*, thereby enabling site-specific incorporation of npAAs into proteins[96]. To date, more than 400 different npAAs have been incorporated into proteins using natural or engineered aaRS/tRNA pairs [97-99]. Notably, one study demonstrated the use of a triply orthogonal aaRS/tRNA pairs, enabling the simultaneous incorporation of three distinct npAAs into a single peptide[100].

GCE is in an active stage of development for adaptation to display techniques. Initial efforts focused on improving the activity and stability of aminoacyl-tRNA synthetases to achieve efficient amber codon suppression and incorporation of npAAs at levels sufficient for generating large peptide libraries [101]. A major advancement was the development of orthogonal ribosomes, which were evolved to recognize only a specific ribosome binding site and exhibit enhanced efficiency of npAA incorporation [102]. The approach has since been extended to a broad spectrum of npAAs, including β-amino acids and α,α-disubstituted amino acids [103]. For example, incorporation of a bromoethyl-tyrosine derivative enabled intramolecular click reactions with cysteine residues, yielding cyclic peptides that, upon screening, produced high-affinity binders to the Keap1 Kelch domain and Sonic Hedgehog protein [104, 105]. Moreover, two distinct orthogonal aaRS/tRNA pairs have been used in parallel to incorporate propargyl-tyrosine and a cyclopropene-containing amino acid into

phage-displayed peptides, thereby enabling selective click reactions and further diversification of displayed libraries. [106].

Incorporation of npAAs into mRNA display systems is technically demanding because it requires reconstruction of all translational components *in vitro* and their adaptation for efficient activity [29]. The most widely used strategy relies on artificial ribozymes in combination with chemically activated amino acids [107]. In this method, tRNA-aminoacylating ribozymes (flexizymes) recognize the universal 3′-terminal CCA of tRNAs with reassigned anticodons and catalyze their acylation with pre-activated amino acid esters [108]. The resulting aminoacyl-tRNAs are then introduced into a cell-free translation system, where they direct the site-specific incorporation of npAAs into peptides that remain covalently linked to their encoding mRNA, enabling selection by the display method. The prominent example showed the construction of peptide library containing 12 different npAAs [109]. The same strategy was applied to incorporate N-chloroacetyl-D-tyrosine at the start codon, which cyclized with an internal cysteine to yield a library of thioether-linked macrocyclic peptides. Screening of this library produced ligands that specifically bind Lys48-linked ubiquitin chains, capable of modulating proteasomal degradation pathways that are directly implicated in cancer cell survival and proliferation [110]. Orthogonal aminoacyl-tRNA synthetases have been adapted for in vitro applications only recently. For example, a p-cyanophenylalanine–specific synthetase was used to incorporate p-cyanopyridylalanine, enabling pyridine–thiazoline cyclization in mRNA display[111]. Although *in vitro* display methods have the potential to overcome key limitations of cellular systems, such as toxicity from overexpressed orthogonal components, bottlenecks in transformation efficiency, and limited uptake or poor stability of npAAs, phage display remains more cost-effective and technically simpler, and therefore continues to be the most widely used platform.

The potential of npAAs is also being explored in the field of biocatalysts, where genetic code expansion provides a reliable method for *in vivo* incorporation. The technique has mainly been applied to enhance enzyme stability, introduce reactive handles for immobilization, and tune active site properties [112]. In one example, β-lactamase was engineered with npAAs bearing elongated thiol side chains to enable the formation of an additional, long-range disulfide bond. This artificial cross-link stabilized the protein fold and led to a significant improvement in thermostability [113]. In another study, 3,4-dihydroxyphenylalanine (DOPA) was incorporated at the terminal region of ω-transaminase. The catechol group of DOPA provided strong coordination with chitosan and polystyrene bead surfaces, enabling high-affinity enzyme

immobilization and improved operational stability under industrial reaction conditions[114]. An excellent example of functional optimization is the work on nitroreductase, a prodrug-activating enzyme in cancer therapy. Its therapeutic use has been limited by insufficient activity toward the prodrug CB1954. To address this, the active-site residue Phe124 was substituted with p-aminophenylalanine and p-nitrophenylalanine. These substitutions preserved the phenylalanine π-stacking with the electron-deficient aryl ring of the substrate, while the introduced amino or nitro groups enhanced ring polarization, which resulted in 20- and 30-fold increases in prodrug activation efficiency, respectively [115].

These examples demonstrate the considerable potential of npAAs as versatile building blocks across diverse areas of science and industry. Display technologies enable rapid generation of large libraries and high-throughput screening, while genetic code expansion provides access to their controlled incorporation *in vivo* [116]. Nevertheless, despite this wide potential, the practical use of npAAs remains constrained by the ability to synthesize sufficient quantities of the required building blocks and to develop scalable processes that meet industrial and market demands [27]. As outlined in this work, the presented study establishes a pipeline for the development of bacterial strains capable of producing unique npAAs. These strains can serve as platforms for constructing peptide libraries with directly incorporated npAAs, thereby eliminating the need for external supplementation and overcoming uptake limitations. In parallel, overproducer strains provide a sustainable and economically viable route to large-scale production of free npAAs through fermentation.

Guanidine containing amino acids

The guanidinium moiety is one of the most prominent functional groups in nature. It is widespread in biological systems as the side chain of arginine and participates in numerous molecular processes. Arginine residues are frequently located on protein surfaces and represent the second most abundant residues at protein–protein interfaces [117]. The positive charge of the guanidinium group is delocalized across three nitrogen atoms, resulting in a highly stable cation that remains protonated under all physiological conditions. This electronic structure enables the formation of multiple hydrogen bonds, which are widely involved in interactions with biological anions such as carboxylate and phosphate groups (Fig. 12). Notably, interactions between arginine and carboxylate groups account for approximately 40% of salt bridges observed in proteins [118].

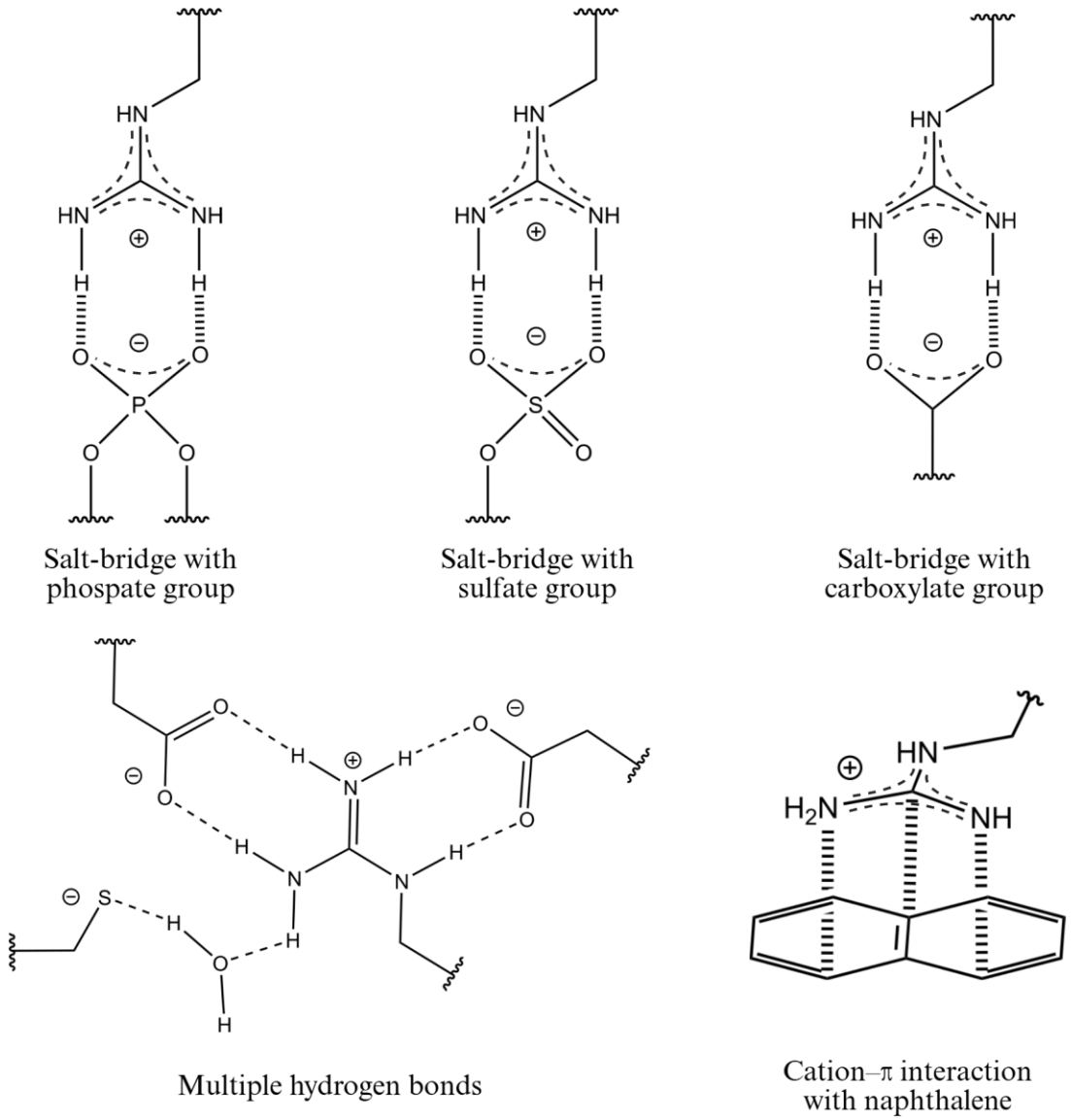


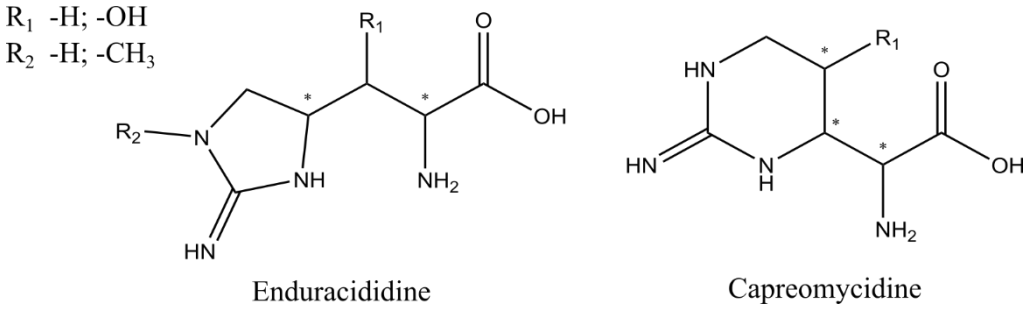
Figure 12. Diverse interactions of the guanidinium group in biological systems.

The central carbon atom of the guanidinium group is sp^2-hybridized, which gives the moiety a planar geometry [119]. Together with π-charge delocalization this creates a form of aromaticity [120]. Electron-rich structures such as benzene generate regions of negative electron density above and below the aromatic ring, which can engage in cation–π interactions with the guanidinium group (Fig. 11) [121]. An important example is found at protein–RNA interfaces, where arginine guanidinium groups form hydrogen bonds with RNA phosphate groups, while additional stabilization arises from cation–π interactions between the guanidinium moiety and the aromatic nucleotide bases [117]. Due to this dual mode of interaction, arginine is involved in approximately 80% of protein–RNA complexes [122, 123].

Such characteristics and significant functional capabilities make the guanidinium containing amino acids an intriguing tool for molecular modeling. One of the most remarkable applications is in the field of cell-penetrating peptides. Arginine-rich peptides containing about 6–12 residues can rapidly cross cell membranes [124]. The guanidinium groups interact with negatively charged components on the cell surface and mediate entry by pore formation or through endocytosis [125]. Arginine residues can be combined with α,α-disubstituted amino acids and D-amino acids, which promote the formation of an α-helical peptide conformation. In this structure, residues interlock to form intramolecular bonds that create local hydrophobic regions, while the guanidinium groups of arginine are oriented outward. This arrangement produces an amphipathic peptide architecture. Because cell membranes consist of both polar head groups and hydrophobic lipid tails, such amphipathic peptides are able to directly translocate across the membrane [126].

Beyond the canonical arginine, the guanidinium group is also present in several npAAs, which are frequently incorporated as residues in biologically active natural products. In these compounds, the guanidinium functionality is often central to the biological effect. Guanidinium-containing npAAs act as key elements in molecular recognition and thereby dictate both the potency and the selectivity of the natural product.

The focus of this work is on naturally occurring cyclic guanidinium-containing non-proteinogenic amino acids, particularly enduracididine, capreomycidine, and their derivatives (Fig. 13). Enduracididine is an arginine analog in which the side chain forms a five-membered ring between the γ-carbon atom and the Nη2 nitrogen atom of the guanidinium group. The structure of enduracididine contains two chiral centers: one at the α-carbon, as in all amino acids, and one at the γ-carbon where cyclization occurs, giving rise to four possible stereoisomers. In addition to these, a β-hydroxylated derivative (β-hydroxyenduracididine) occurs in several glycopeptide antibiotics [127].



Enduracididine Capreomycidine

Figure 13. Structures of enduracididine and capreomycidine. Possible stereocenters are indicated with asterisks (*).

The first natural examples were uncovered in the depsipeptide antibiotic enduracidin (Fig. 14), where residues of L-enduracididine and D-allo-enduracididine were identified. Two congeners, enduracidin A and enduracidin B, were originally isolated. Both are 17-residue cyclic peptides, distinguished only by a single methyl group in their fatty acid side chain [128]. They display strong activity against Gram-positive pathogens, including *methicillin-resistant Staphylococcus aureus* (MRSA) and vancomycin-resistant enterococci (VRE) [129]. Enduracidins have even been manufactured on an industrial scale and used as feed additives for pigs and chickens, although their precise molecular mode of action remains unresolved. The well-known glycopeptide vancomycin acts by binding to the D-Ala–D-Ala termini of lipid II, thereby blocking cell wall biosynthesis and causing bacterial death [130]. Resistance arises when D-Ala–D-Ala is replaced by D-Ala–D-Lac, which abolishes vancomycin binding. enduracidins also bind lipid II, but their activity against resistant strains suggests a binding mode distinct from that of vancomycin [131].

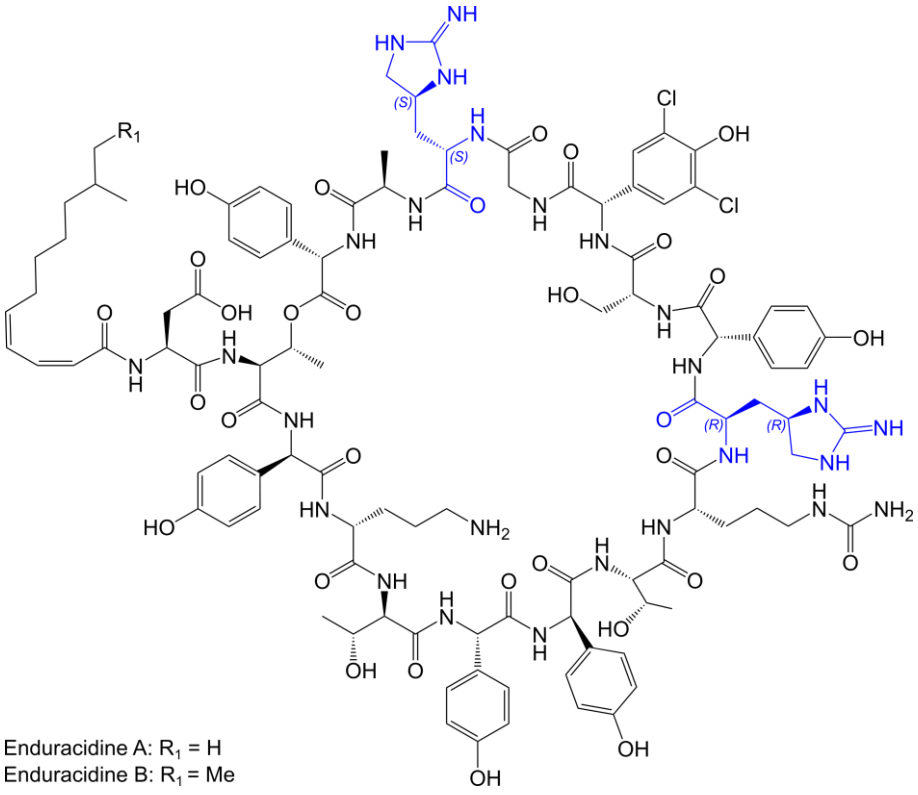


Figure 14. Chemical structure of enduracidin A and B. Enduracididine resides are highlighted in blue.

A further example is the mannopeptimycins (Fig. 15), glycopeptides that incorporate two residues of β-hydroxyenduracididine. These compounds are likewise active against MRSA and VRE [132]. Although their exact mechanism has not yet been fully defined, evidence indicates that they interact with the negatively charged pyrophosphate region of lipid II through their

positively charged residues [133]. Support for this model comes from inactivation of the enzyme *mppO*, which is responsible of the hydroxylation of enduracididine. Loss of this modification yields dideoxy mannopeptimycin analogs with impaired activity, underscoring the essential contribution of β-hydroxyenduracididine to lipid II recognition and antibacterial potency [134, 135].

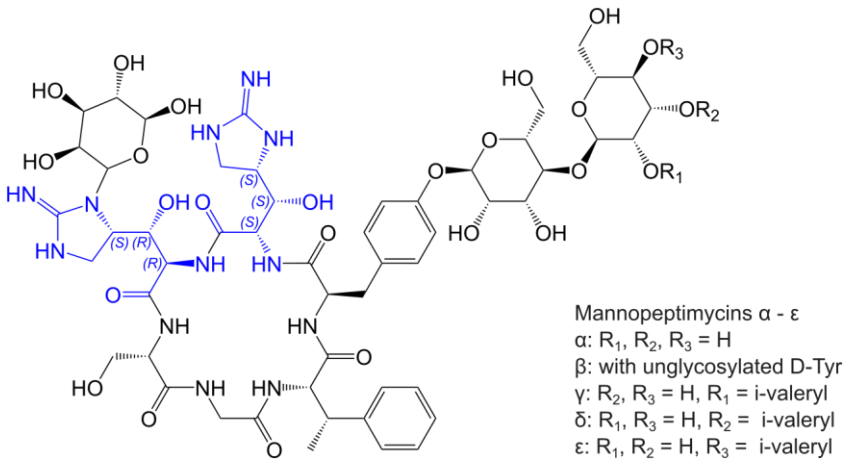


Figure 15. Chemical structure of mannopeptimycins. Enduracididine resides are highlighted in blue.

The most intensively studied enduracididine-containing antibiotic is teixobactin [136]. This 11-residue cyclic depsipeptide is closed by a lactone linkage that links the terminal isoleucine to D-threonine at position 8, creating a macrocycle that incorporates the last four residues, including L-*allo*-enduracididine (Fig. 16). The cyclic architecture places the enduracididine side chain on the outer surface of the molecule, leaving it exposed for target interaction. Teixobactin is exceptionally potent against Gram-positive bacteria, including MRSA, VRE, *Mycobacterium tuberculosis*, and *Clostridium difficile*, through binding to lipid II.

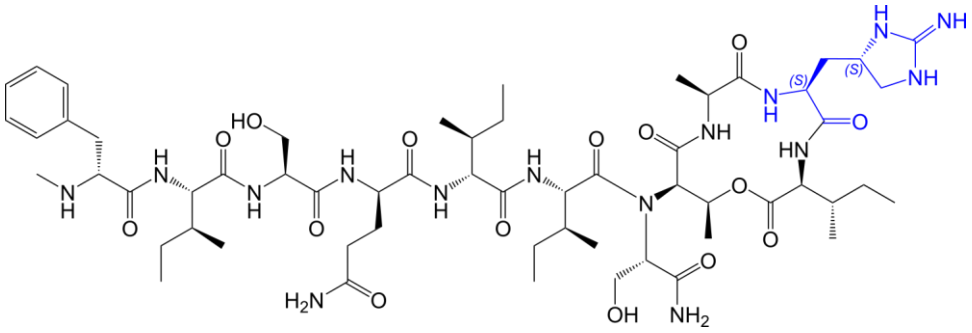


Figure 16. Chemical structure of teixobactin. Enduracididine reside is highlighted in blue.

Solid-state NMR studies have revealed that teixobactin specifically recognizes the highly conserved pyrophosphate–MurNAc moiety of lipid II (Fig. 17) [137]. Within this binding mode, enduracididine residue plays the central to this recognition: its side chain forms a hydrogen bond with the C6-hydroxyl group of MurNAc, while the positively charged

guanidinium group strengthens electrostatic contacts with the negatively charged pyrophosphate [138]. When enduracididine is replaced, these interactions become loose and poorly defined, and the affinity of the molecule for its target drops drastically.

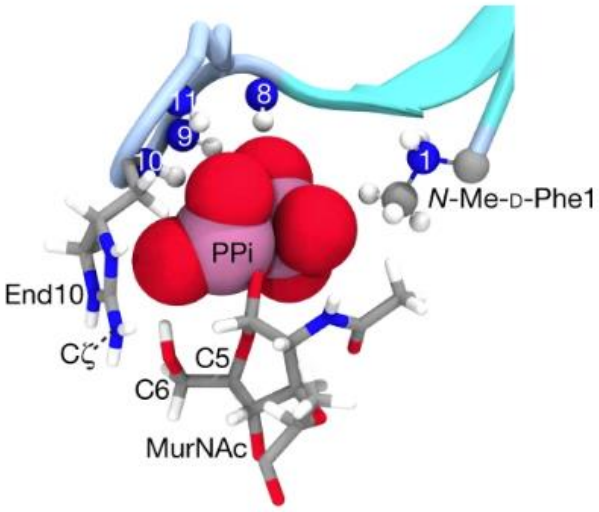


Figure 17. Interaction of the enduracididine residue in teixobactin with lipid II [138]. The enduracididine residue interacts with the pyrophosphate (PPi) group of lipid II and additionally stabilizes the complex through a hydrogen bond with the hydroxyl group of the MurNAc sugar.

A mechanistically distinct example is the aminoglycoside minosaminomycin, which contains L-allo-enduracididine as a core residue (Fig. 18). This compound inhibits the growth of mycobacteria, including *M. tuberculosis* [139, 140].

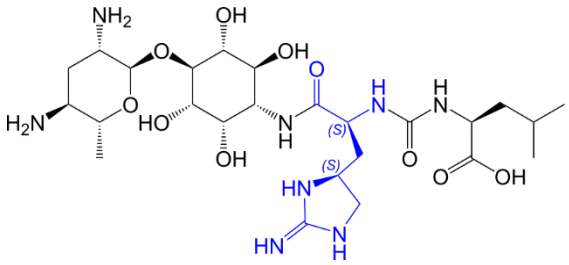


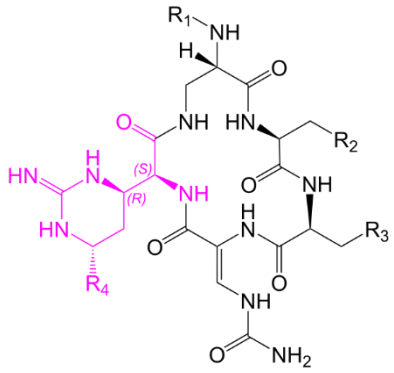
Figure18. Chemical structure of minosaminomycin. Enduracididine reside is highlighted in blue.

Its activity is limited to a narrow range of species, likely reflecting poor penetration through bacterial membranes combined with the presence of a specialized uptake system in mycobacteria. To separate uptake from intrinsic activity, minosaminomycin was tested in cell-free translation assays. In this environment it selectively blocked the initiation step of protein synthesis, leaving elongation and termination largely unaffected. Importantly, ribosomes resistant to the structurally related kasugamycin remained sensitive to minosaminomycin, which points to a different ribosomal binding region. The most plausible explanation is that

minosaminomycin interferes with initiation on the 30S subunit by engaging the negatively charged 16S rRNA, while the cationic enduracididine side chain enhances RNA recognition and binding affinity [141].

Capreomycidine is a cyclic guanidinium-containing non-proteinogenic amino acid structurally related to enduracididine. Unlike enduracididine, it forms a six-membered ring that closes at the β-carbon of arginine. The structure contains two stereogenic centers, which gives rise to four stereoisomers, and a δ-hydroxy derivative has also been identified [142].

Capreomycidine is best known as a residue of the tuberactinomycin family of antibiotics, most notably viomycin and capreomycin (Fig. 19). These small cyclic peptides were introduced as second-line agents against tuberculosis, particularly in the treatment of multidrug-resistant tuberculosis (MDR-TB). In structural terms, capreomycins contain L-capreomycidine, whereas viomycin incorporates δ-hydroxy-L-capreomycidine [142].



	R1	R2	R3	R4
Capreomycin IA:	H	OH	-NH-β-lysine	H
Capreomycin IB:	H	H	-NH-β-lysine	H
Capreomycin IIA:	H	OH	-NH2	H
Capreomycin IIB:	H	H	-NH2	H
Viomycin:	β-lysine	OH	OH	OH
Tuberactinomycin O:	β-lysine	OH	OH	H
Tuberactinomycin A:	γ-OH-β-lysine	OH	OH	OH
Tuberactinomycin N:	γ-OH-β-lysine	OH	OH	H
Tuberactinamin N:	H	OH	OH	H

Figure 19. Chemical structure of tuberactinomycins. Capreomycidine reside is highlighted in pink.

Unlike enduracididine-containing glycopeptides that act extracellularly on lipid II, tuberactinomycins must penetrate the mycobacterial cell envelope to reach their intracellular target. The envelope of *Mycobacterium tuberculosis* has a unique multilayered structure that combines features of both Gram-positive and Gram-negative bacteria. It has a thick peptidoglycan layer, which is covalently linked to arabinogalactan decorated with long-chain mycolic acids that form an outer membrane [143]. Such an architecture creates an exceptionally strong permeability barrier and contributes to the high level of intrinsic drug resistance in mycobacteria.

Despite this barrier, mycobacteria encode numerous outer membrane proteins that mediate the uptake of small metabolites. A plausible model is that tuberactinomycins, being polar cationic peptides, mimic natural substrates of peptide transport pathways and thereby gain access to the cytoplasm [144]. Once inside, they act on the ribosome, where they stabilize a pre-translocation

state and block protein synthesis. The capreomycidine moiety is a crucial component of this interaction, forming strong electrostatic contacts with the phosphate backbone of 16S rRNA, anchoring the antibiotic at the ribosomal interface [145]. Removal or alteration of this residue greatly reduces binding affinity and antibacterial potency.

Another representative example is the muraymycins, a family of nucleoside–peptide antibiotics that contain L-*epi*-capreomycidine (Fig. 20).

These compounds inhibit translocase I (MraY), an essential enzyme in peptidoglycan biosynthesis. MraY catalyzes the transfer of UDP-MurNAc-pentapeptide to the lipid carrier undecaprenyl phosphate to generate lipid I. Muraymycins mimic the natural substrate and thereby block this step in cell wall assembly [146]. The precise contribution of the capreomycidine residue to target binding is not fully understood, but experimental substitutions have shown that its replacement in muraymycin D2 leads to a substantial reduction or even complete loss of antibacterial activity [147].

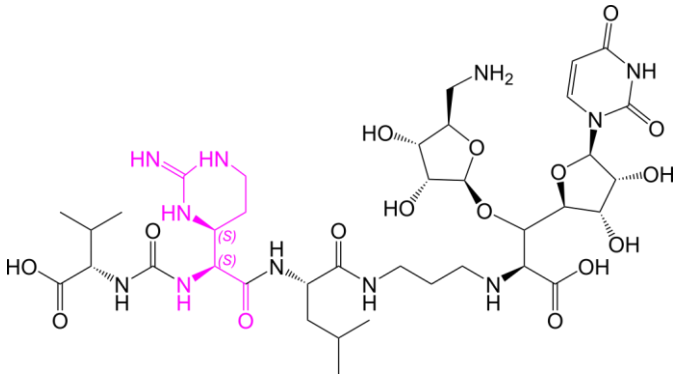


Figure 20. Chemical structure of muraymycin D2. Capreomycidine reside is highlighted in pink.

L-*epi*-capreomycidine is also found in chymostatin**s**, a family of natural peptides that block the activity of chymotrypsin-like serine proteases (Fig. 21) [148].

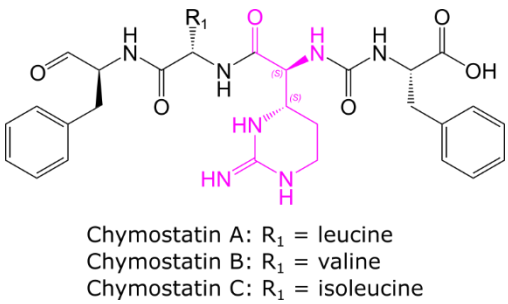


Figure 21. Chemical structure of chymostatin. Capreomycidine reside is highlighted in pink.

These enzymes have a deep hydrophobic pocket and normally cleave proteins after large hydrophobic or aromatic amino acids. In the chymostatin–protease complex, the phenylalanine residue of the inhibitor occupies this pocket, mimicking a natural substrate. The C-terminal aldehyde group reacts with the catalytic serine to form a covalent bond, which locks the enzyme in an inactive state. The capreomycidine residue does not sit inside the pocket but forms multiple hydrogen bonds that anchor the inhibitor and stabilize the inactive complex [149].

To conclude, the data presented here underline the exceptional importance of the guanidinium moiety in biological systems, particularly in mediating interactions at protein and nucleic acid interfaces. Natural antibacterial products provide striking examples: guanidinium-bearing residues guide these molecules to highly conserved sites and reinforce binding through dense networks of hydrogen bonds, electrostatic contacts, and cation–π interactions. These interactions are especially strong with organic phosphate groups, which are indispensable for cellular metabolism and therefore ideal targets for antibiotic action. The presence of guanidinium-containing npAAs further contributes to protease resistance, as their unusual cyclic structures are rarely encountered in standard protein substrates. Taken together, these features highlight guanidinium-containing npAAs as powerful and promising building blocks for the design of next-generation antibiotics and peptide therapeutics.

Tryptophan-derived non-proteinogenic amino acids

Tryptophan is a natural amino acid with an indole side chain. The indole is a bicyclic aromatic system that displays dual chemical character. The aromatic ring system is largely nonpolar, enabling hydrophobic interactions and π-stacking, whereas the N–H group of the pyrrole ring carries a partial positive charge and acts as a hydrogen-bond donor (Fig. 22).

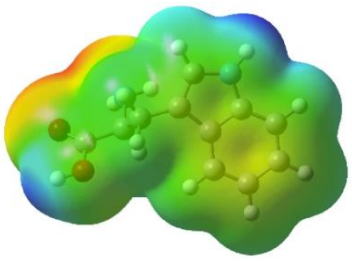


Figure 22. Tryptophan electrostatic potential [150].

As a residue in peptides, tryptophan is frequently enriched at membrane–water interfaces, where the bulky nonpolar moiety interacts with the lipid bilayer while the indole nitrogen forms hydrogen bonds with polar head groups at the aqueous boundary. Within protein structures,

tryptophan residues are often buried in hydrophobic cores, where they stabilize folding through aromatic stacking or cation–π interactions with positively charged side chains.

The indole ring of tryptophan shows a unique balance between stability and reactivity. Its side chain can undergo diverse modifications, including hydroxylation, prenylation, and halogenation, giving rise to stable bioactive metabolites [151]. This chemical versatility has led to the emergence of the tryptamine class of molecules, which includes major neurotransmitters such as serotonin and melatonin, as well as naturally occurring psychedelics like psilocin and N,N-dimethyltryptamine.

Tryptophan-derived npAAs are also often found in the structure of diverse natural products [152]. One of the most common modifications of tryptophan is halogenation of the indole ring, typically at the unsubstituted carbons of the benzene moiety, such as positions C4, C5, C6, and C7 [153-156]. A prominent example is the C7-specific halogenase RebH, originally identified in the rebeccamycin biosynthetic pathway, which generates 7-chlorotryptophan. Two such halogenated tryptophan units undergo intramolecular cyclization to form a planar indolocarbazole core structure. Subsequent glycosylation and methylation steps convert this scaffold into rebeccamycin, a promising antitumor agent that intercalates into DNA and inhibits topoisomerase I, an enzyme particularly active in rapidly dividing cancer cells [157]. This example also highlights another key feature of tryptophan metabolism, namely its ability to serve as a precursor for novel heterocyclic structures. In most cases, the indole side chain remains intact, while the α- and β-carbons of the amino acid backbone undergo enzymatic modifications that give rise to additional aromatic or cyclic scaffolds. These include pyrroloquinolines, pyrroloindoles, β-carbolines, and indolactams, among others [152].

The major catabolic route of tryptophan proceeds through the kynurenine pathway. This pathway represents a conserved part of primary metabolism, linking tryptophan degradation to the biosynthesis of nicotinamide adenine dinucleotide (NAD^+) via quinolinic acid as an intermediate [158]. Kynurenine is a central metabolite of this route. In humans, it is produced mainly in the liver, where it can be further metabolized or transported into the central nervous system, where it can cross the blood–brain barrier through the large neutral amino acid transporters. Beyond serving as a metabolic intermediate, kynurenine acts as an agonist of the aryl hydrocarbon receptor and thereby modulates immune signaling. Dysregulation of kynurenine metabolism has been linked to several pathological conditions, including Alzheimer's disease, Parkinson's disease, and schizophrenia [159]. Furthermore, kynurenine

and other intermediates of the kynurenine pathway, such as 3-hydroxykynurenine, absorb ultraviolet light because of their conjugated aromatic structures and are present in the human eye lens. These metabolites act as weak photosensitizers that capture photon energy and release it through non-radiative pathways, thereby dissipating it in a harmless manner. [160].

The bacterium *Streptomyces roseosporus* utilizes kynurenine for the biosynthesis of daptomycin, a calcium-dependent lipopeptide antibiotic active against Gram-positive pathogens including MRSA and VRE [161]. Daptomycin is typically used as a second- or last-line treatment because it must be administered intravenously and can cause dose-limiting side effects. Its mechanism of action requires calcium ions and phosphatidylglycerol and involves binding to lipid II, resulting in the formation of a tripartite Ca^{2+}–DAP–PG–lipid II complex. This interaction blocks peptidoglycan biosynthesis and induces membrane rearrangements and collapse [162]. Substitution of kynurenine with a structurally related residue reduces antibacterial activity by two- to fourfold, suggesting an important contribution of this residue. Although its precise functional role in the mechanism of action remains unclear, the kynurenine residue provides a unique experimental advantage. Owing to its distinct absorbance and fluorescence properties, kynurenine serves as an intrinsic fluorophore within the peptide. Changes in its emission intensity and wavelength upon binding to calcium and to phosphatidylglycerol-containing membranes have been exploited to monitor the conformational activation of daptomycin, the formation of the Ca^{2+}–DAP complex, and the subsequent insertion of the antibiotic into bacterial membranes [163].

In summary, the chemical plasticity of tryptophan-derived non-proteinogenic amino acids and metabolites establishes them as versatile building blocks for molecular design. The indole backbone enables strong hydrophobic interactions, while halogen substitutions introduce polar functionalities that enable interactions at the interface between nonpolar and aqueous phases. The intrinsic fluorescence of residues such as kynurenine provides a natural probe for experimental studies. Beyond their structural utility, many tryptophan-derived metabolites could function directly as active drugs, or neurotransmitter agonists.

Ergothioneine and hercynine

Ergothioneine is a histidine-derived metabolite with strong antioxidant properties, predominantly produced by fungi and many actinobacteria. Several biosynthetic pathways have been described, all converging at hercynine, which represents the Nα,Nα,Nα-trimethylated derivative of histidine (Fig. 23). Subsequent enzymatic conversions differ among

organisms but share a common principle: the incorporation of sulfur from cysteine into the imidazole ring of hercynine, catalyzed by distinct enzyme sets [164].

Ergothioneine exists in equilibrium between thiol and thione tautomers (Fig. 23). Under physiological pH, the thione form predominates, conferring exceptional stability and resistance to autoxidation. This tautomeric behavior underlies its potent antioxidant capacity and ability to maintain redox balance in biological systems [165].

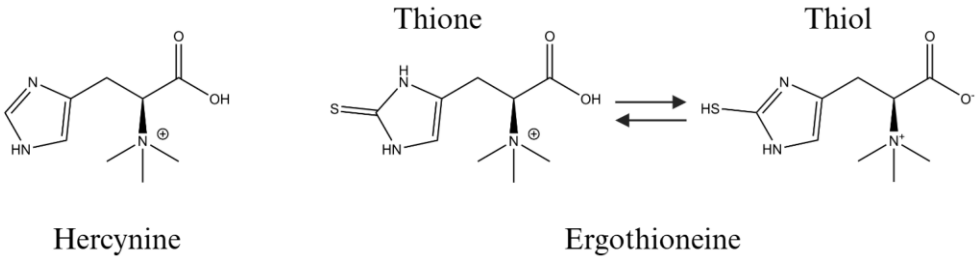


Figure 23. Structures of hercynine and tautomer forms of ergothioneine

Beyond its physiological roles, ergothioneine has attracted considerable industrial and biomedical interest. It is utilized as a bioactive additive in food, cosmetics, and nutraceutical formulations, and ongoing research explores its therapeutic potential due to its free-radical scavenging and metal-chelating properties [166]. Traditionally, ergothioneine is produced through fungal extraction, but recent advances in microbial fermentation have enabled efficient biosynthesis. Notably, engineered Escherichia coli strains have achieved titers exceeding 2 g L^{-1} in fed-batch cultivation, highlighting the feasibility of large-scale production [167].

Previously, the incorporation of ergothioneine into peptides was considered challenging due to its reactive thiol group, which was expected to interfere with coupling and deprotection reactions during solid-phase peptide synthesis (SPPS) [168]. However, recent studies demonstrated successful incorporation of ergothioneine without protective groups, revealing that the thione tautomeric form provides sufficient intrinsic stability to tolerate SPPS conditions. The resulting conjugates showed enhanced radical-scavenging capacity and strong copper-binding activity [169].

These findings highlight the versatility of ergothioneine and its biosynthetic precursor hercynine as functional building blocks for molecular engineering. Integrating their biosynthetic pathway into a strain compatible with genetic code expansion would enable site-specific incorporation of ergothioneine or hercynine analogs during ribosomal translation, providing a direct biological route to generate redox-active and metal-binding peptides and proteins that cannot be easily accessed by chemical synthesis alone.

Biotechnological production of non-proteinogenic amino acids

Biotechnological production of non-proteinogenic amino acids (npAAs) relies on the ability of microorganisms to convert organic substrates into valuable building blocks through fermentation. This process involves several critical choices that collectively determine its efficiency and economic feasibility. In this chapter, the interactions between the individual steps are discussed to develop a rationale approach for selecting an optimal strategy.

Substrate

The initial input into the process is an organic substrate that serves as both a nutrient source and a carbon backbone for the microorganisms. During fermentation, microorganisms grow and simultaneously redirect part of their metabolic flux toward the synthesis of the target compound. Fermentation usually takes place in a bioreactor, with aeration applied when aerobic species are employed. Because this process is based on living systems, substrate consumption must fulfill two purposes at once: biomass formation and product generation.

The composition of the fermentation medium strongly influences both parameters. Rich, undefined media contain a wide range of nutrients such as amino acids, sugars, lipids, vitamins, and mineral supplements. These media support rapid microbial growth and high biomass but are costly and may introduce impurities that complicate downstream purification. In contrast, minimal or defined media contain only the essential nutrients in controlled amounts. While this reduces growth rates due to the need for the cell to synthesize its own building blocks, it lowers production costs and simplifies purification.

For industrial processes, complex media based on agricultural byproducts are often preferred. Examples include molasses, which is rich in fermentable sugars, and corn steep liquor, which provides nitrogen and micronutrients. Such substrates are inexpensive, widely available, and already integrated into large-scale amino acid fermentations.

Current developments in sustainable biotechnology focus on broadening the spectrum of usable feedstocks. Efforts include engineering microorganisms capable of metabolizing lignocellulosic hydrolysates or even utilizing carbon dioxide as a carbon source. Such approaches aim to reduce dependency on food-related substrates and increase the environmental and economic sustainability of npAA production.

Microbial host selection

The selection of a microbial host is the most critical decision in the development of a biotechnological process for non-proteinogenic amino acid (npAA) production. The overall efficiency, robustness, and scalability of the process depend directly on the properties of the chosen strain.

Several factors determine host suitability. First, the organism must provide an environment in which the biosynthetic pathway of the npAA can be expressed, translated, and function optimally. Second, the host should offer sufficient precursor supply and metabolic robustness to channel flux toward the product. Additional considerations include the availability of a genetic engineering toolbox, knowledge of regulatory and transport mechanisms, tolerance to product accumulation, compatibility with large-scale fermentation, and ease of downstream processing.

One option is to employ the natural producer of a given npAA. This approach ensures the presence of a functional biosynthetic pathway without the need for extensive heterologous expression. However, it carries significant drawbacks. Natural producers must be optimized individually with strain-specific methods. Many of these organisms are poorly studied or even unculturable under laboratory conditions. Toolboxes for metabolic engineering are typically unavailable, and little is known about the regulation of their primary and secondary metabolism, making precursor optimization difficult. Furthermore, natural overproducers of npAAs are frequently actinomycetes such as *Streptomyces* species. While they possess remarkable potential for producing diverse natural products, each strain often has 10-30 biosynthetic gene clusters. In the absence of clear regulatory control, this leads to simultaneous synthesis of multiple secondary metabolites, which complicates purification of the desired compound.

To overcome these limitations, attention often shifts toward well-characterized microbial chassis. Such organisms offer a rich toolbox for metabolic engineering and a proven track record in the industrial production of amino acids and related compounds. In these hosts, the main challenge is to create optimal conditions for expression of the desired biosynthetic pathway. Present technologies enable multiple strategies: codon optimization ensures efficient transcription, promoter and ribosome binding site design improves translation, and protein engineering enhances enzyme stability and catalytic efficiency in the intracellular environment.

The availability of a well-developed engineering toolbox therefore represents one of the most decisive factors for host selection.

An alternative strategy involves the use of computational models to guide host choice. Genome-scale metabolic models can be applied to predict precursor availability, flux distribution, and theoretical yields. Coupled with machine learning, these approaches could identify hosts with the highest potential to sustain the precursors supply of a given pathway. Although such predictive methods are promising, their accuracy still depends heavily on the quality and completeness of experimental data, which currently limits their broad application.

Ultimately, these approaches inspire the concept of a platform strategy. Instead of relying on a single universal host, a portfolio of diversified chassis could be developed, each optimized for a defined set of precursors and environmental conditions. These hosts would be aligned with databases and genome-scale metabolic models to provide predictive insights into their capabilities. When a new biosynthetic pathway is discovered, it could first be screened against this portfolio to evaluate compatibility. The most promising host would then be selected and further tuned to achieve high titers of the desired npAA. Such a platform approach would significantly accelerate the biotechnological production of diverse npAAs and establish a scalable route toward their widespread use as building blocks.

In the scope of this work of the production of cyclic guanidiunm containing npAAs we focused on the two bacterial hosts: *Streptomyces albus* and *Corynebacterium glutamicum*.

Streptomyces albus

The biosynthetic pathways of cyclic guanidinium-containing npAAs originate from gene clusters of natural products found in various *Streptomyces* species. Utilization of *Streptomyces albus* therefore minimizes challenges of successful gene expression, since this host is naturally adapted to handle such biosynthetic systems. It also provides an intrinsic advantage, as it could be naturally less sensitive to the produced compounds and carries diverse transport systems that can move npAA across the membrane, which is a constant challenge in other hosts.

Streptomyces are established industrial producers of secondary metabolites, which demonstrates their potential for high-titer production and robustness during scale-up. In particular, the strain *Streptomyces albus* Δ14 represents a well-developed host with an advanced genetic engineering toolbox. This strain has been engineered by deletion of all of its own secondary metabolite clusters. That clears the background and simplifies purification of the desired compound and at the same time eliminate the competing metabolic pathways.

The main challenge lies in the regulation of precursor supply, which may limit flux toward cyclic guanidinium-containing npAAs and restrict maximum titers. Nevertheless, *S. albus* Δ14 remains highly valuable as a research platform. Even at lower intracellular production levels, it can be used for incorporation of heterologously produced npAAs into novel peptides via genetic code expansion or engineered nonribosomal peptide synthetase (NRPS) modules. This flexibility highlights its potential for the discovery and development of new peptide therapeutics.

Corynebacterium glutamicum

Corynebacterium glutamicum is the most widely used microbial host for amino acid production at the industrial scale and represents one of the most established workhorses in biotechnology. Belonging to the same class of Actinomycetes as *Streptomyces*, it shares evolutionary proximity that may facilitate compatibility with heterologous enzymes derived from secondary metabolite gene clusters. Its versatility has already been demonstrated by successful production of several npAAs, confirming its suitability as a chassis beyond canonical amino acids [170].

For cyclic guanidinium-containing npAAs, *C. glutamicum* provides a particular advantage because several strains have been engineered to overproduce arginine, the key precursor of this compound class. Detailed strategies for creating such strains are available, including removal of feedback regulation, optimization of precursor pathways, and amplification of key biosynthetic enzymes. The organism's metabolic network is thoroughly characterized, with robust fluxes through central carbon metabolism that can be redirected into the biosynthesis of primary metabolites.

Beyond metabolic robustness, *C. glutamicum* benefits from an advanced genetic toolbox and extensive industrial experience. Methods ranging from high-efficiency transformation to CRISPR-based genome editing and genome-scale metabolic modeling are well developed for this organism, enabling systematic optimization at multiple levels. Recent studies have also extended its substrate range toward sustainable feedstocks such as lignocellulosic hydrolysates, further increasing its relevance for large-scale applications.

Taken together, *C. glutamicum* combines three essential traits: an industrially proven background in amino acid production, a well-characterized and reliable metabolic framework, and advanced tools for genetic and metabolic engineering. These features establish it as a strong platform host for implementing biosynthetic pathways of cyclic guanidinium-containing npAAs and scaling them toward industrial relevance.

Material and methods

Bacterial strains and plasmids

All bacterial strains and plasmids used in this study are listed in Table 1.

Table 1. Bacterial strains and plasmids

Strain	Description	Reference
Escherichia coli		
DH10B	Strain used for DNA cloning and amplification	[171]
NM522	Strain used for DNA amplification and methylation	Promega
Streptomyces albus		
Del14	Strain used for enhanced production of secondary metabolites	[172]
Del14-End	Del14 derivative with integrated pRT801-End-s vector	This work
Corynebacterium glutamicum		
BCA	Strain used for integration of ϕC31 and ϕBT1 phage-based vectors	[173]
BCA-End	BCA derivative with integrated pRT801-End-c vector	This work
BCA-r-End	BCA derivative transformed with the replicative vector pClik-End	This work
BCA-r-tuf-End	BCA derivative transformed with the replicative vector pClik-tuf-End	This work
BCA-BCD-End	BCA derivative transformed with the replicative vector pClik-BCD-End	This work
BCA-BCD-Erg	BCA derivative transformed with the replicative vector pClik-BCD-Erg	This work
ATTC21831	Arginine overproducer strain	[174]
Arg0-BCD-End	ATTC21831 derivative transformed with the replicative vector pClik-BCD-End	This work
Arg0-BCD-Cap	ATTC21831 derivative transformed with the replicative vector pClik-BCD-Cap	This work
Arg0-BCD-hCap	ATTC21831 derivative transformed with the replicative vector pClik-BCD-Cap-OH	This work
Arg0-BCD-hEnd-	ATTC21831 derivative transformed with the replicative vector pClik-BCD-End-OH	This work
ATTC21850	Tryptophan overproducer strain	[175]
Trp-Kyn	ATTC21850 derivative transformed with the replicative vector pClik-BCD-Kyn	This work
ATTC21253	Lysine overproducer strain	[176]
Lys-Cad	ATTC21253 derivative transformed with the replicative vector pClik-BCD-Cad	

Plasmid	Description	Reference

pUC-GW-Kan	Cloning vector, kanR	Genewiz
pRT801	ϕBT1-based integrative plasmid, ampR	[177]
pRT801-End-s	pRT801 derivative carrying mppPQR genes under the control of A4 promotor	This work
pRT801-End-c	pRT801 derivative carrying codon optimized mppPQR genes under the control of P69, P68, and P67 promotors	This work
pTC	Cloning and methylation vector, tetR	[178]
pClik5a	Replicative vector for E. coli and C. glutamicum, kanR	[179]
pClik-End	pClik5a derivative carrying codon optimized mppPQR genes under the control of P69, P68, and P67 promotors, kanR	This work
pClik-tuf-End	pClik5a derivative carrying codon optimized *mppPQR* genes under the control of the promoter region of the tuf gene, kanR	This work
pClik-BCD-End	pClik5a derivative carrying codon optimized *mppPQR* genes under the control of BCD expression element, kanR	This work
pClik-BCD-Cap	pClik5a derivative carrying codon optimized *vioCD* genes under the control of BCD expression element, kanR	This work
pClik-BCD-hCap	pClik5a derivative carrying codon optimized *orfRP* genes under the control of BCD expression element, kanR	This work
pClik-BCD-hEnd	pClik5a derivative carrying codon optimized *mppPQRO* genes under the control of BCD expression element, kanR	This work
pClik-BCD-Kyn	pClik5a derivative carrying codon optimized genes tryptophan 2,3-dioxygenase and alpha/beta hydrolase (kynurenine synthase) under the control of BCD expression element, kanR	This work
pClik-BCD-Erg	pClik5a derivative carrying codon optimized *egtD12* genes under the control of BCD expression element, kanR	This work
pClik-BCD-Cad	pClik5a derivative carrying codon optimized *flvHG* genes under the control of BCD expression element, kanR	This work

DNA manipulation

Codon optimization and cloning strategies were developed using Geneious Prime (version 2024.0.2, Biomatters Ltd., Auckland, New Zealand). The designed DNA constructs were obtained through the Gene Synthesis service (Azenta Life Sciences, Chelmsford, MA, USA). Plasmid DNA was replicated in *Escherichia coli* DH5α and isolated using QIAGEN Plasmid Kit (QIAGEN, Hilden, Germany). Subsequent DNA manipulations, including PCR, restriction digestion, and ligation, were performed following standard protocols [180].

General media and growth conditions

The *E. coli* strains were routinely cultivated in the Lennox broth (LB). The *Streptomyces albus* strains were cultivated on Mannitol Soya Flour (MS) agar medium to produce spores [181]. *Corynebacterium glutamicum* strains were routinely cultivated in standard Brain Heart Infusion (BHI) medium. When necessary, the following antibiotics were added to the medium: kanamycin (50 μg ml^{-1} for *E. coli*, 25 μg ml^{-1} for *C. glutamicum*), apramycin (50 μg ml^{-1} for *E. coli* and *S. albus*, 25 μg ml^{-1} for *C. glutamicum*). Unless stated otherwise, *E. coli* strains were grown at 37 °C, *S. albus* strains were grown at 28 °C and *C. glutamicum* strains were grown at 30 °C.

Strain construction

For *S. albus* strains construction, the ϕBT1 site-specific recombinase-based vector pRT801 was introduced via conjugation [172]. *C. glutamicum* strains were transformed with the integrative vector pRT801 or the replicative vector pClick5a via electroporation. [173]. Successful integration was confirmed by PCR using the appropriate primers. All primers used in this study are listed in Table S1.

Batch cultivation in shake flasks

S. albus strains were grown on MS agar medium at 28°C for 4 days. Freshly obtained spores were then inoculated into the nutrient-rich liquid Tryptic Soy Broth (TSB) medium and cultivated at 28°C for 24 hours. This preculture served as the seed material for the main production cultivation in DNPM medium, which was carried out at 28°C for 96 hours [182]. The agar medium was supplemented with 50 μg ml^{-1} apramycin, while the liquid culture contained 25 μg ml^{-1}.

C. glutamicum strains were grown on BHI agar medium additionally supplemented with 40 g L^{-1} of glucose and 13.2 g L^{-1} of ammonium sulfate at 30°C for 2 days. Preculture was grown in the liquid BHI medium with glucose and ammonium sulfate for 24 hours. The main production cultivation was carried out for 48 hours in defined CGM medium. The CGM medium contained per litter: 80 g of glucose, 40 g of ammonium sulfate, 10 g of yeast extract, 4 g of dipotassium phosphate, 4 g of potassium phosphate, 2 g of citric acid, 500 mg of magnesium sulfate, 60 mg of protocatechuic acid, 40 mg of ferrous sulfate, 15 mg of zinc sulfate, 11 mg of Manganese(II) sulfate, 11 mg of calcium chloride, 5 mg of pyridoxal phosphate, 1 mg of thiamine and 1 mg of biotin. The agar medium was supplemented with 25

μg ml^{-1} of apramycin or kanamycin, while the liquid culture contained 12.5 μg ml^{-1} of apramycin or kanamycin.

Fed-batch cultivation

Fed-batch fermentation was carried out in a 3.7L bioreactor system Labfors 5 (Infors HT, Bottmingen, Switzerland). The aeration rate was set at 1.5 vvm by the integrated gas flow controller. Dissolved oxygen was determined by VisiFerm electrode (Hamilton Company, Reno, NV, USA)). pH was determined by EasyFerm Plus PHI electrode (Hamilton Company, Reno, NV, USA)) and maintained at the value of 6.9 by automated pumping of 28% ammonium hydroxide and 1M hydrochloric acid.

The first preculture was inoculated from a fresh BHI agar plate and cultivated for 12 hours in 20 mL of liquid BHI medium supplemented with 40 g L^{-1} of glucose, 13.2 g L^{-1} of ammonium sulfate, and 12.5 μg ml^{-1} of kanamycin at 30°C. This culture was used as the seed material for the second preculture. The volume of the second preculture matched the starting volume of the bioreactor. Two liters of CGM medium were inoculated with 20 mL of the first preculture. Kanamycin was then added to a final concentration of 12.5 μg ml^{-1}, and the medium was distributed into 500 mL flasks, each containing 50 mL of culture. This final preculture was cultivated at 30°C for 24 hours, after which the cell biomass was harvested by centrifugation (3 min, 8000 rcf), resuspended in 20 mL of CGM medium, and inoculated into the bioreactor.

The fed-batch cultivation consisted of two main stages. The initial batch phase, with a set stirring rate of 700 rpm and an aeration rate of 1.5 vvm, was carried out until complete glucose consumption. In the subsequent fed phase, dissolved oxygen was maintained at 30% saturation and controlled automatically by adjusting the stirrer speed, while the feeding strategy was manually regulated based on glucose concentration in the supernatant. Glucose concentration was measured using the D-Glucose Assay Kit (Megazyme, Bray, Ireland). The feed solution contained per liter: 500 g of glucose, 200 g of ammonium sulfate, 10 g of yeast extract, 2 g of dipotassium phosphate, 2 g of potassium phosphate, 2 g of citric acid, 500 mg of magnesium sulfate, 60 mg of protocatechuic acid, 40 mg of ferrous sulfate, 15 mg of zinc sulfate, 11 mg of Manganese(II) sulfate, 11 mg of calcium chloride, 5 mg of pyridoxal phosphate, 1 mg of thiamine and 1 mg of biotin.

Determination of Cell Dry Weight

The culture samples were diluted with Milli-Q water, and the optical density was measured photometrically at 660 nm. The correlation between cell dry weight (CDW) and optical density was determined according to a previously described method [183]. The CDW was calculated by multiplying the measured optical density by the experimentally determined coefficient (0.32) and expressed as g L^{-1}.

Marfey's derivatization and L-enduracididine verification

As a control, 1 mg of the enduracididine-containing peptide enramycin was hydrolyzed in 300 µL of 6N hydrochloric acid at 110°C for 45 minutes under a nitrogen atmosphere. After hydrolysis, the sample was dried using a nitrogen concentrator and dissolved in 100 µL of Milli-Q water. This sample and the *S. albus* Del14-End culture supernatant were subjected to Marfey's reagent derivatization as described earlier [184]. After the reaction, samples were analyzed using high-resolution LC/MS. The amino acid derivatives were separated on ACQUITY BEH C18 column (1.7 µm, 2.1 mm × 100 mm, Waters, Milford, MA, USA) at 40°C, with water + 0.1% formic acid and acetonitrile + 0.1% formic acid as the mobile phases. The mass of the compounds was determined using a maXis high-resolution LC-QTOF mass spectrometer (Bruker Daltonics, Bremen, Germany).

Preparative isolation of L-enduracididine and L-capreomycidine

The non-canonical amino acids were isolated from 1 L of culture broth after fed-batch cultivation. First, the cells were removed by centrifugation (10 min, 8000 rcf), and the remaining supernatant was evaporated to a final volume of 100 mL using a rotary evaporator at 40°C and 50 mbar. The concentrated sample was acidified to pH 2 with 2M hydrochloric acid and left overnight at 4°C to allow protein precipitation. The next day, the sample was filtered through a cellulose filter, and 20 mL of the filtrate was used for preparative isolation. The compounds were separated using an ÄKTA chromatography system (Cytiva, Uppsala, Sweden) and an OmniFit column (ID 25 mm × 400 mm, Diba Industries, Cambridge, UK) packed with Dowex 50WX8 ion-exchange resin Resin preparation and compound isolation were performed according to the method developed for the chromatography of basic amino acids [185].

NMR Spectroscopy for Structural Confirmation

The chemical structures of all the compounds were determined via multidimensional NMR analysis. ^{1}H-NMR, ^{13}C-NMR, and 2D spectra were recorded at 500 MHz (1H) and 125 MHz (13C), conducted in the Bruker Avance Neo 500 MHz, equipped with a Prodigy Cryo-probe. Samples were dissolved in D2O. Chemical shifts are reported in ppm relative to tetramethylsilane; the solvent was used as the internal standard. Coupling constants are reported in Hertz (Hz). Multiplicity is reported with the usual abbreviations (s: singlet, d: doublet, dd: doublet of doublets, t: triplet, dq: doublet of quartets, m: multiplet). Chiroptical measurements of all the compounds in H2O ([α]D20) were obtained on a model Jasco P-2000 Automatic Digital Polarimeter (JASCO, Easton, MD, USA) in a 3.5 x 50mm cell at 20°C.

Quantification of non-natural amino acids

The amino acid concentrating was quantified using the Accela UPLC system (Thermo Fisher Scientific, Waltham, MA, USA) coupled to the Amazon ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany). The separation of the underivatized amino acids was achieved using the ACQUITY BEH Amide column (1.7 µm, 2.1 mm × 100 mm, Waters Corporation, Milford, MA, USA) at 40°C, with water + 0.1% formic acid and acetonitrile + 0.1% formic acid as the mobile phases. Prior to the analyses supernatant samples were diluted in the 0.2M hydrochloric acid. The calibration curve for the L-enduracididine and L-capreomycidine was measured using the earlier isolated samples, the purity of the compounds was determined via NMR spectroscopy.

Results and discussion

Expression of the enduracididine biosynthetic pathway in *Streptomyces albus*

Enduracididine is a rare non-natural amino acid characterized by a unique 5-membered cyclic group containing guanidine moiety. A number of antibiotics contain enduracididine residue or its derivatives, which are essential for their biological activity [127]. For instance, L-enduracididine (2S,4R-enduracididine) and its stereoisomer D-*allo*-enduracididine are components of enramycin, L- and D-β-hydroxyenduracididine residues are present in mannopeptimycins, and L-*allo*-enduracididine is part of the antibiotic teixobactin. All these antibiotics target bacterial cell wall biosynthesis by binding to lipid II, a precursor of peptidoglycan, with enduracididine residues playing an essential role in this interaction [131, 138, 186]. These properties make enduracididine a promising building block for the development of novel antibiotics. However, its chemical synthesis is challenging, involving multiple steps and resulting in low yields.[127]. In this study, we describe a biotechnological approach to produce L-enduracididine through bacterial fermentation.

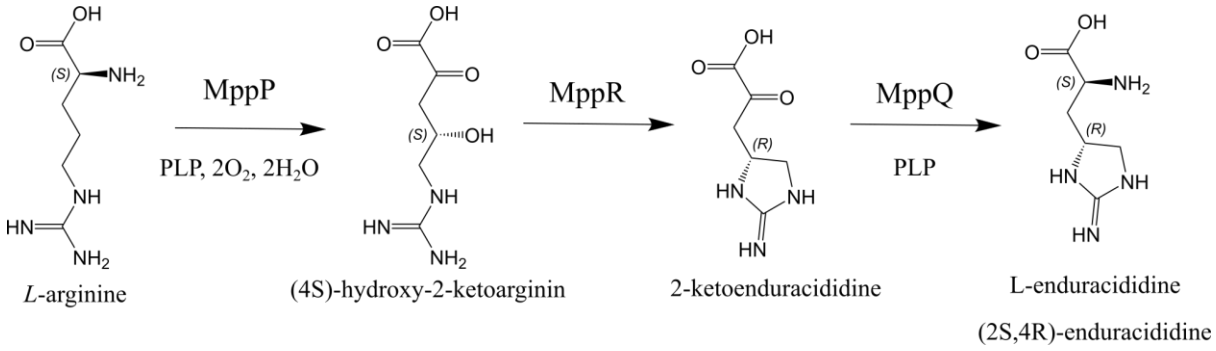


Figure 24. Biosynthetic pathway of L-enduracididine.

In vivo L-enduracididine is biosynthesized from L-arginine [187]. This conversion is catalyzed by three enzymes, MppP, MppQ, and MppR (Fig. 24), from the mannopeptimycin biosynthetic gene cluster [39], or their homologs EndP, EndQ, and EndR from the enramycin biosynthetic gene cluster [128]. Firstly, PLP-dependent oxidase MppP converts L-arginine into (4S)-hydroxy-2-ketoarginin [188]. In the next step, acetoacetate decarboxylase-like enzyme MppR catalyzes the dehydration of (4S)-hydroxy-2-ketoarginin at the C4-C5 position, which subsequently cyclizes to form 2-ketoenduracididine [135]. Finally, the PLP-dependent

aminotransferase MppQ converts 2-ketoenduracididine into L-enduracididine during a transamination reaction, with L-ornithine as the most efficient amino group donor [189].

To investigate the potential of fermentative production of L-enduracididine, we aimed to express biosynthetic genes *mppPQR* in *Streptomyces albus* Del14, a strain specifically engineered for enhanced production of secondary metabolites [172]. The genes were synthesized with native codon usage from *Streptomyces wadayamensis* and placed under the control of the strong synthetic A4 promotor [190]. This synthetic construct was cloned into the ϕBT1 integrase-based vector pRT801 and introduced into *S. albus* Del14 via conjugation. Successful chromosomal integration was confirmed by PCR.

The resulting strain, *S. albus* Del14-End, was cultivated in DNPM medium for 96 hours. L-enduracididine production was analyzed using Marfey's derivatization followed by LC-MS analysis [191]. The hydrolysate of the enduracididine-containing peptide enramycin was used as the standard. The mass and retention time of the Marfey's derivative of L-enduracididine from the *S. albus* Del14-End culture supernatant matched those of the standard, confirming the compound's identity (Fig. 25). No L-enduracididine was detected in the control strain *S. albus* Del14.

This experiment showed the successful expression of *mppPQR* genes and demonstrated their potential for fermentative production of the rare amino acid L-enduracididine. However, the production level of the L-enduracididine was extremely low. Efficient high-titer production of the amino acid requires optimal gene expression, high enzyme activity and sufficient precursor supply. Amino acid metabolism in *Streptomyces* remains largely unexplored, therefore, to avoid potential bottleneck in arginine supply, we decided to select a host with more favorable metabolic characteristics: *Corynebacterium glutamicum*.

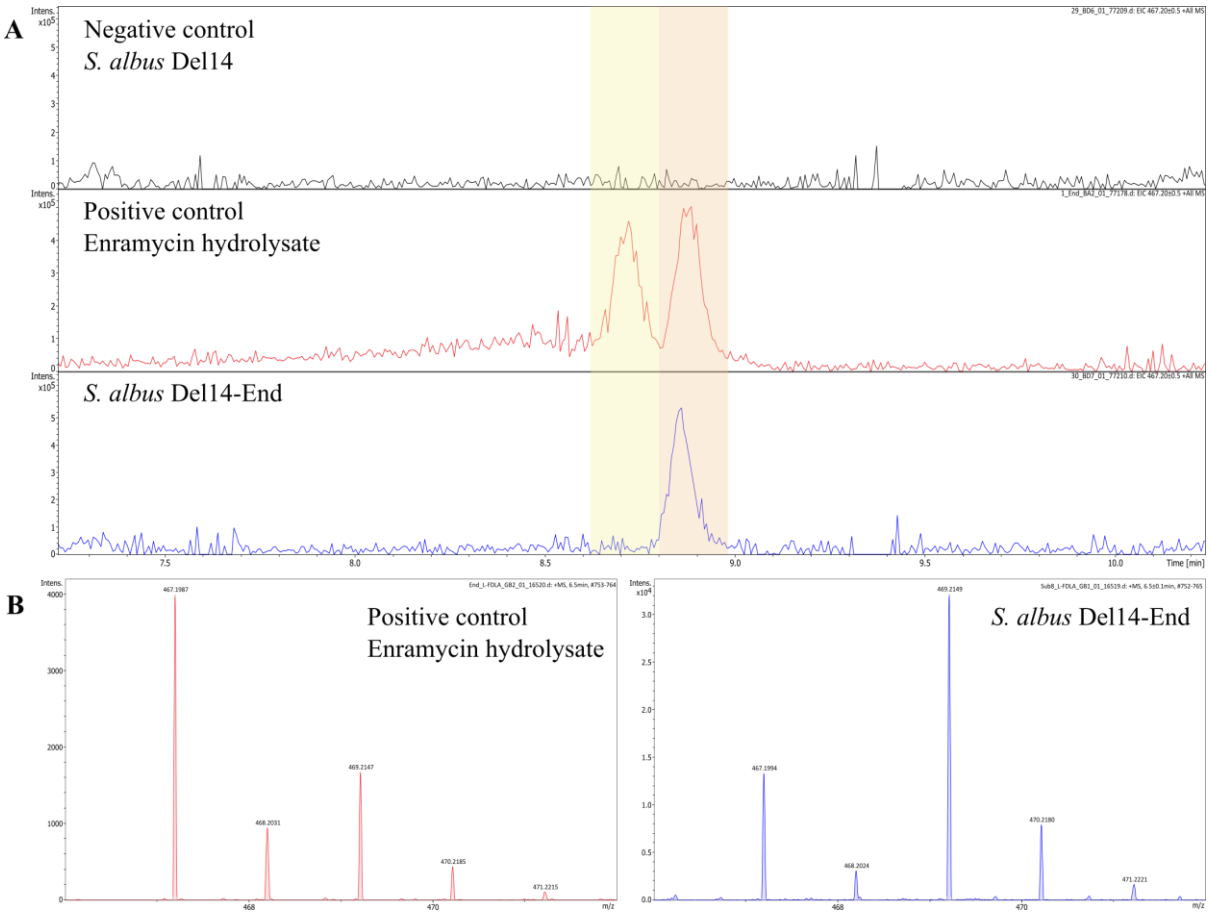


Figure 25. LC/MS chromatograms of *S. albus* Del14-End fermentation supernatant after Marfey's derivatization with corresponding controls for enduracididine production verification. (A) LC/MS chromatograms of Marfey-derivatized culture supernatants of *S. albus* Del14 (negative control, black) and *S. albus* Del14-End (blue), compared with the hydrolysate of the enduracididine-containing peptide enramycin (positive control, red). (B) Exact mass spectra of Marfey-derivatized enduracididine from the enramycin hydrolysate (left, red) and from the culture supernatant of *S. albus* Del14-End (right, blue). The signal at m/z 469.2149 corresponds to arginine, which is more abundant in the fermentative supernatant.

Expression of the enduracididine biosynthetic pathway in *Corynebacterium glutamicum*

Corynebacterium glutamicum is a well-known overproducer of natural amino acids and is widely used for their fermentative production [192]. Effective strategies for optimizing *C. glutamicum* strains have been extensively described, supported by well-established metabolic engineering and bioprocess methodologies [193]. Furthermore, *C. glutamicum* has been successfully used to produce several non-natural amino acids, including L-pipecolic acid, trans-4-hydroxy-L-proline and 3,4-dihydroxyphenyl-l-alanine [170, 194, 195].

For the expression of L-enduracididine biosynthetic genes, *C. glutamicum* strain BCA was selected as the host organism. This strain was developed as a derivative of the classical strain ATCC13032 and includes additional *attB* integration sites for ϕC31 and ϕBT1 phage-based

vectors [173]. The *mppPQR* genes were codon optimized for *C. glutamicum* and placed under the control of synthetic constitutive promoters P69, P68, and P67, respectively [196]. This synthetic construct was introduced into *C. glutamicum* BCA using integrative vector pRT801.

The resulting strain, *C. glutamicum BCA-End*, was cultivated in CGM medium for 48 hours. L-enduracididine was detected only in the biomass in trace amounts, while its precursor molecule, L-arginine, was detected at a concentration 0.25 mg g^{-1} cell dry weight (CDW) (Fig. 26). The presence of residual arginine suggests that the precursor supply is not the primary bottleneck for L-enduracididine production. To further increase the L-enduracididine production, we decided to focus on enhancing gene expression levels.

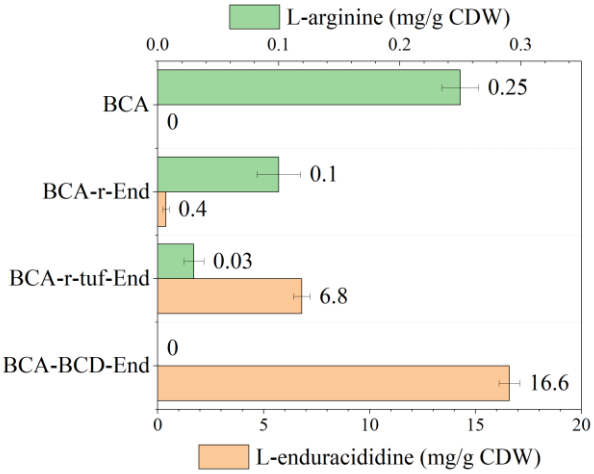


Figure 26. Effect of Gene Expression Systems on the Conversion Efficiency of L-Arginine to L-Enduracididine in *C. glutamicum* BCA strain.

To enhance the expression level of the *mppPQR* genes, we opted to increase their copy number within each cell. For this purpose, the multicopy replicative vector pClik5a was used instead of the integrative pRT801 vector [179]. The resulting strain, *C. glutamicum* BCA-r-End, demonstrated significantly higher production of intracellular L-enduracididine during 48-hour cultivation in the CGM medium. The higher amino acid concentration permits the use of a simplified detection method, using a HILIC column for direct amino acid separation, avoiding derivatization. *C. glutamicum* BCA-r-End with the replicative plasmid produced 0.4 mg g^{-1} CDW of L-enduracididine (Fig.26) while the residual concentration of L-arginine decreased 2.5-fold compared to the strain *C. glutamicum* BCA-End with the integrative plasmid. Increase of the copy number of the biosynthetic genes enhanced their expression level, resulting in higher conversion rates and greater L-enduracididine production. However, the residual L-arginine concentration of 0.1 mg g^{-1} CDW implies that further enhancement of enduracididine

gene expression might further increase the conversion of the precursor molecule into the non-natural amino acid.

To further enhance gene expression, the *mppPQR* genes were placed under the control of the native promoter region of the *tuf* gene [197]. This gene encodes elongation factor Tu, a highly conserved and essential protein required for translation, while its promoter is widely used by scientists to drive strong and constitutive gene expression [198]. Replacement of the synthetic promoter in front of *mppP* by the promoter region of the *tuf* gene resulted in a 17-fold increase in L-enduracididine production. The corresponding strain, *C. glutamicum* BCA-r-tuf-End, produced 6.8 mg g^{-1} CDW of L-enduracididine during 48-hour shake flask cultivation (Fig.26). The intracellular L-Arginine was detected at concentration of 0.03 mg g^{-1} CDW, highlighting the potential for further optimization. Encouraged by this result, we conducted an additional round of gene expression enhancement.

Besides the transcription, the high-level gene expression relies on efficient protein translation. The translation initiation as the most crucial factor depends on the absence of secondary mRNA structures affecting availability of the ribosome binding site (RBS). Interestingly, in prokaryotic polycistronic mRNA molecules, the translation efficiency of the first gene determines the translation efficiency of all downstream genes [199]. This feature has been applied for the development of the bicistronic design (BCD) elements for high-level protein expression [200]. Specifically, the BCD element includes a synthetic operon of two cistrons, in which the first short synthetic fore-cistron provides high-level translation for the second cistron, which is the gene of interest. To express the *mppPQR* genes, we created the fore-cistron, which consists of the promotor region of the *tuf* gene and its first 72 base pairs fused in frame with 72 terminal base pairs of the *rplP* gene (Fig. 27). The terminal sequence includes an internal Shine-Dalgarno (SD) sequence for initiation of translation of our genes of interest [201].

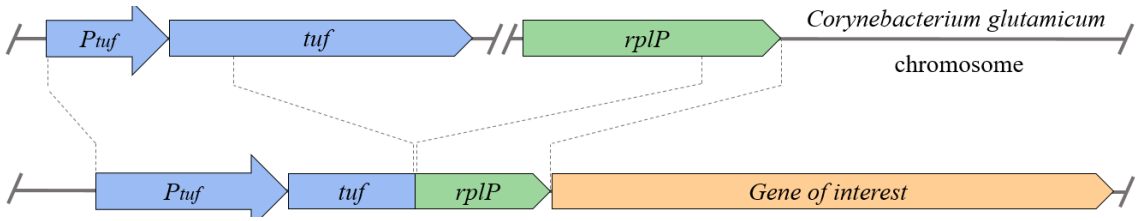


Figure 27. Architecture of the optimized bicistronic design element for efficient expression of biosynthetic genes in *C. glutamicum*.

Each of the genes *mppPQR* was synthesized with an upstream BCD expression fore-cistron and cloned into the pClik5a plasmid. This construct was introduced into the *C. glutamicum* BCA via transformation. The resulting strain, *C. glutamicum* BCA-BCD-End, produced 16.6 mg g^{-1} CDW of L-enduracididine in 48-hour shake flask cultivation (Fig. 26). This final strain demonstrated a 2.4-fold increase in the production compared to BCA r-tuf-End. No residual L-arginine could be detected in the biomass of the strain. The obtained data indicate that the overexpression of the *mppPQR* genes under control of the constructed BCD fore-cistron element led to the complete conversion of intracellular arginine in *C. glutamicum* BCA BCD-End. This finding also suggests that the precursor supply is now limiting the further increase of L-enduracididine production.

Expression of the enduracididine biosynthetic pathway in arginine overproducer *Corynebacterium glutamicum* ATTC21831

New host, *C. glutamicum* ATCC21831 (Arg0), was selected in order to overcome precursor supply limitation. This strain is a mutant of *C. glutamicum* ATCC13032, which was selected for its ability to produce high amounts of extracellular L-arginine [174]. Under our laboratory conditions the strain *C. glutamicum* ATCC21831 produced 0.79 g L^{-1} of extracellular L-arginine during 48-hour shake flask cultivation in CGM medium at 30 °C (Fig. 28).

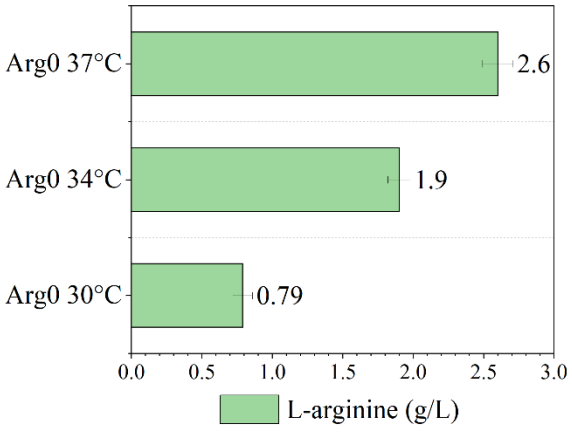


Figure 28. Impact of cultivation temperature on L-Arginine biosynthesis by *C. glutamicum* ATCC21831 strain in shake flasks.

It was previously reported that an increase of cultivation temperature can decrease the activity of the 2-oxoglutarate dehydrogenase complex, leading to the accumulation of 2-oxoglutarate and the redirection of its flux toward glutamate production [202, 203]. Since the L-arginine biosynthetic pathway begins with glutamate, a temperature shift may also enhance L-arginine

production. Increasing the cultivation temperature to 34 °C improved the productivity of the strain, yielding an arginine concentration of 1.9 g L^{-1}. The highest titer of 2.6 g L^{-1} was achieved at 37 °C, demonstrating the positive effect of temperature increase on the arginine production.

Arg0 was used for the expression of the *mppPQR* genes under the control of BCD element on the replicative pClik5a vector. The resulting strain, *C. glutamicum* Arg0-BCD-End, was cultivated in CGM medium in shaking flasks at 30 °C for 48 hours, producing 0.33 g L^{-1} of L-enduracididine, which was entirely secreted into the supernatant (Fig. 29). Interestingly, further increase of the cultivation temperature conditions led to the decrease of L-enduracididine production. This indicates that enduracididine biosynthetic enzymes have lower enzymatic activity at increased temperature.

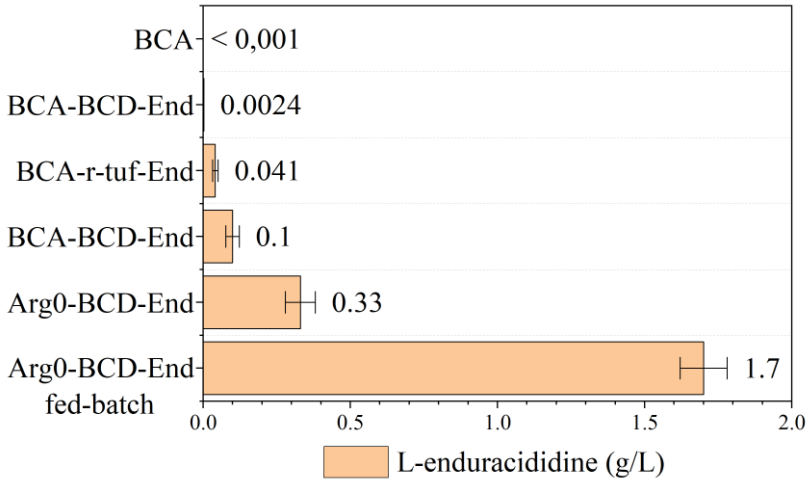


Figure 29. Shake flask and Fed-batch bioreactor production of L-Enduracididine by engineered *C. glutamicum* strains.

To determine the maximum production capacity of L-enduracididine, *C. glutamicum* Arg0-BCD-End was cultivated in a bioreactor via the fed-batch fermentation process. The biosynthesis of one molecule of L-arginine from the citric acid cycle intermediate 2-oxoglutarate requires three molecules of nicotinamide adenine dinucleotide phosphate (NADPH) [204]. NADPH is a crucial metabolite involved in maintaining intracellular redox balance and driving anabolic reactions, and it is primarily generated via the pentose phosphate pathway. During the exponential growth phase, *C. glutamicum* catabolizes glucose mainly through this pathway [205].

To evaluate the maximal production capacity of L-enduracididine, *C. glutamicum* Arg0-BCD-End and the parental Arg0 strain were cultivated in in a bioreactor via the fed-batch fermentation process (Fig. 30).

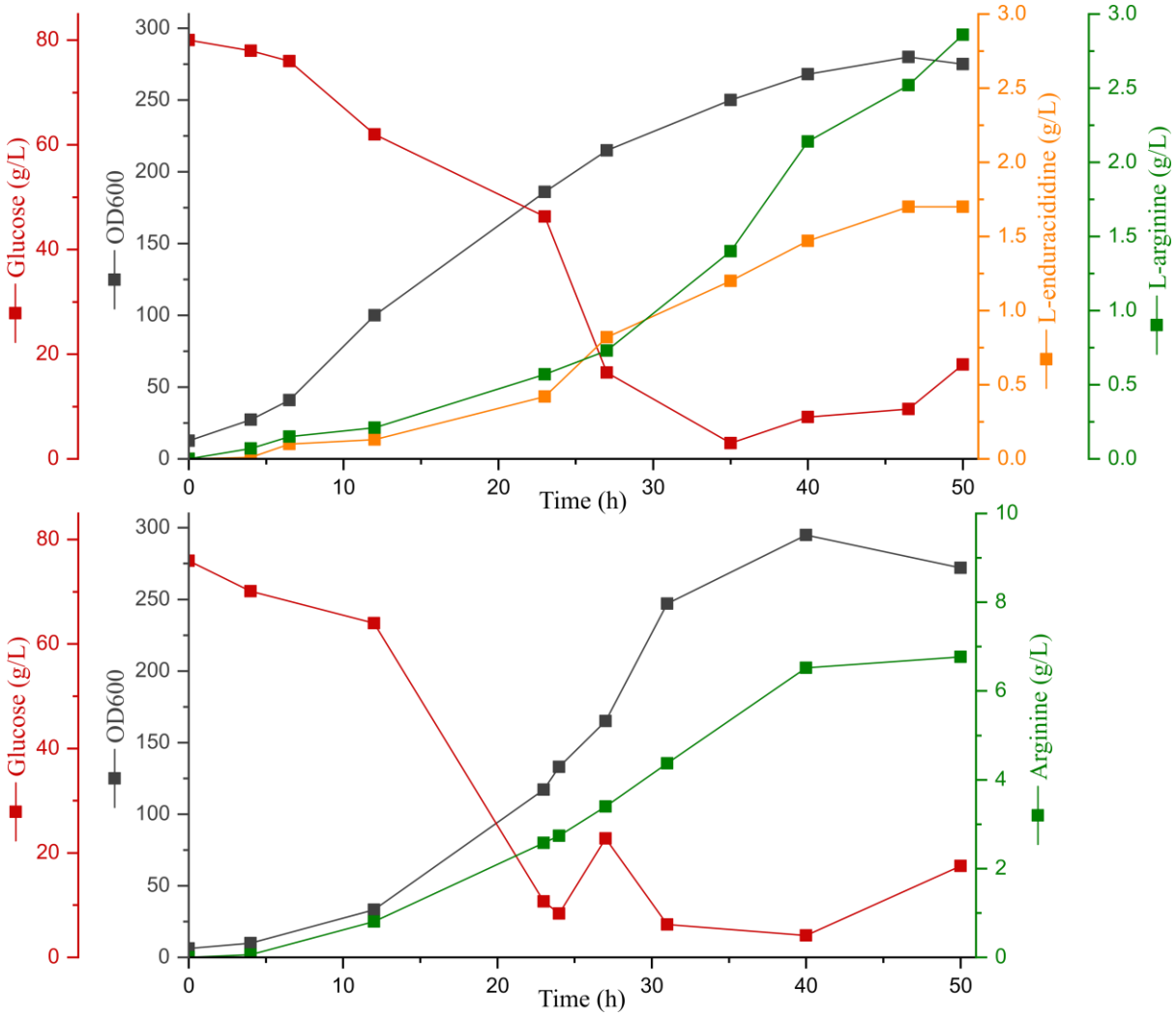


Figure 30. Fed-batch fermentation profile of *C. glutamicum* Arg0-BCD-End strain.

Both fermentations were run for 50 hours. The fermentation process was started with a batch phase, and feeding was initiated after the dissolved oxygen (DO) spike, which indicated complete glucose consumption. The control strain Arg0 produced 6.77 g L^{-1} of L-arginine, whereas the engineered strain Arg0-BCD-End produced 1.7 g L^{-1} of L-enduracididine together with 2.86 g L^{-1} of L-arginine. These data confirm that the engineered pathway remains active under fed-batch conditions and that substantial conversion of arginine into compound 1 occurs during high-density cultivation.

The overall yield for the production of L-enduracididine from glucose reached 14.8 mg g^{-1} glucose, corresponding to a volumetric productivity of 34 mg L^{-1} h^{-1}. During fermentation,

both Arg0 and Arg0-BCD-End reached OD_{600} values close to 300, reflecting the high specific growth rate and strong central carbon flux characteristic of *C. glutamicum*. However, rapid biomass formation directly competes with product biosynthesis and represents a major constraint on further increases in yield and productivity. In addition, high cell densities substantially elevate the oxygen demand of the culture, which can promote metabolic stress and redox imbalance and ultimately lead to cell lysis.

Although the engineered strain successfully produced gram-scale quantities of L-enduracididine, further strain development is required to redirect metabolic flux more efficiently toward arginine-derived products and to improve cellular energy and oxidative balance, thereby enhancing robustness. Notably, only 37% of the produced L-arginine was converted into compound 1, identifying precursor utilization as a second major limitation. This observation highlights the need for additional optimization of the expressed biosynthetic pathway to enable complete and efficient conversion of arginine into the target npAA.

The supernatant from the fed-batch fermentation of *C. glutamicum* Arg0-BCD-End was processed for the preparative isolation of L-enduracididine. It was concentrated via evaporation, acidified, and applied to a cation exchange column, as detailed in the Methods section. A single-step purification yielded 0.1 gram of the compound which was subsequently analyzed using NMR, confirming its identity to (2S,4R)-enduracididine (SI). This sample was then utilized to prepare a calibration curve and employed for quantitative measurements.

Expression of the hydroxyenduracididine biosynthetic pathway in arginine overproducer *Corynebacterium glutamicum* ATTC21831

The mannopeptimycin biosynthetic gene cluster also encodes an alpha-ketoglutarate-dependent hydroxylase MppO which selectively catalyzes the oxidation of the β-carbon of L-enduracididine leading to β-hydroxy-enduracididine (Fig. 31) [134].

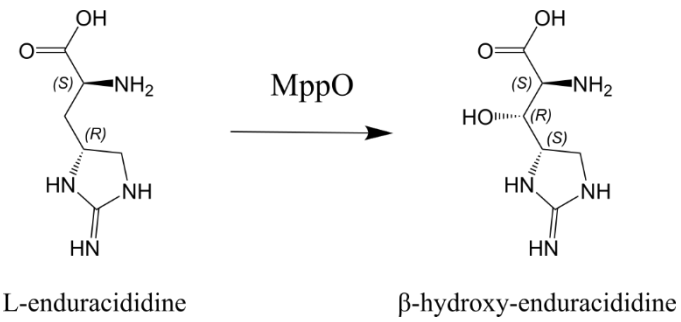


Figure 31. Conversion of L-enduracididine to β-hydroxy-enduracididine catalyzed by the MppO enzyme.

Although the specific contribution of this hydroxylated residue to the bioactivity of mannopeptimycins has not been fully elucidated, its substitution results in a marked decrease in biological activity, indicating that it plays an important functional role and requires further detailed investigation.

For the fermentative production of 3S-hydroxy-enduracididine the *mppO* gene was expressed alongside the L-enduracididine biosynthetic genes *mppPQR*. For this purpose, a codon-optimized version of *mppPQRO* genes under control of the BCD element was cloned into pClick5a vector and introduced in the arginine overproducer *C. glutamicum* Arg0. The resulting strain, *C. glutamicum* Arg0-BCD-hEnd, was cultivated in CGM medium under various temperature conditions. The concentration of produced amino acids are represented in the relative peak area, as the hydroxyenduracididine was not yet obtained in the pure form to perform the calibration curve to enable quantitative measurements (Fig. 32).

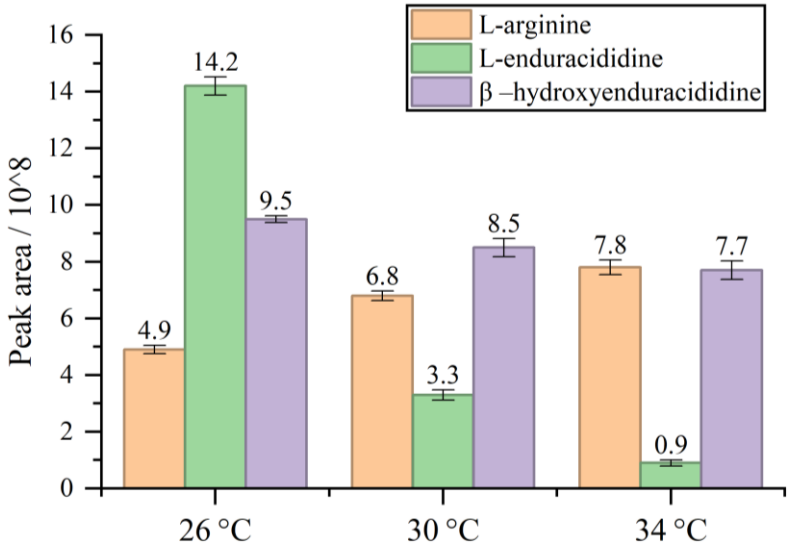


Figure 32. Production of L-arginine, L-enduracididine, and β-hydroxy-L-enduracididine during shake flask fermentation of *C. glutamicum* Arg0-BCD-hEnd at different temperatures.

The data indicate that lower fermentation temperatures result in the highest level of arginine conversion into enduracididine and hydroxyenduracididine. As the fermentation temperature increases, the extracellular concentration of arginine slightly rises, while the conversion to enduracididine decreases sharply. In contrast, the titer of the hydroxy form declines only slightly. These results suggest that the MppO enzyme is more stable and less temperature-sensitive within this range; however, its overall activity remains lower than that of the complete enduracididine biosynthetic pathway, as full conversion of enduracididine was not achieved.

Expression of the capreomycidine biosynthetic pathway in arginine overproducer *C. glutamicum* ATTC21831

Capreomycidine is another non-natural amino acid, structurally related to enduracididine. It differs from enduracididine by having a unique six-membered cyclic group containing a guanidine moiety, instead of the five-membered ring. Like enduracididine, the capreomycidine molecule contains two stereocenters what makes its chemical synthesis challenging. (2S,3R)-capreomycidine and its hydroxy derivative are present in the structures of the tuberactinomycin family of antibiotics, such as viomycin and capreomycin, which are effective against multidrug-resistant tuberculosis [206]. These antibiotics inhibit bacterial protein synthesis by binding simultaneously to the 16S rRNA of the small ribosome subunit and the 23S rRNA of the large subunit, thereby blocking the conformational changes required for translocation. The guanidine moiety of the amino acid capreomycidine is responsible for the formation of salt-bridge with the phosphate group of the adenosine residue A1493 of the 16S rRNA [207]. This data highlights that capreomycidine is another promising building block for the development of novel antibiotics. Since the reported chemical synthesis of capreomycidine incudes multiple steps and has extremely low yields, we aimed to apply previously developed biotechnological approach to produce this non-natural amino acid via fermentation [208].

Capreomycidine biosynthetic enzymes VioC and VioD are encoded in the viomycin biosynthetic gene cluster [209]. Similarly to enduracididine L-arginine serves as a main biosynthetic precursor of capreomycidine [210]. In the first step, an unusual alpha-ketoglutarate-dependent dioxygenase VioC catalyzes stereospecific hydroxylation of the L-arginine, yielding the (3S)-hydroxy-arginine (Fig. 33). In the next step, a PLP-dependent enzyme VioD catalyzes the intramolecular cyclization of (3S)-hydroxy-arginine, resulting in the (2S,3R)-capreomycidine formation.

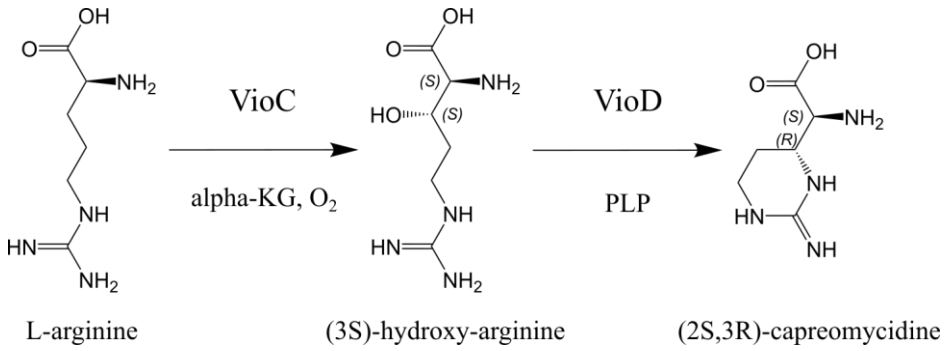


Figure 33. Biosynthetic pathway of (2S,3R)-capreomycidine.

The codon optimized genes *vioCD* from *Streptomyces vinaceus* ATCC11861 were synthesized and expressed under control of the BCD element on the pClik5a vector in the arginine overproducer *C. glutamicum* Arg0. The resulting strain, *C. glutamicum* Arg0-BCD-Cap, was cultivated in shake flasks for 48 hours under various fermentation temperatures. The highest production of (2S,3R)-capreomycidine was achieved at 34 °C, with a titer of 0.64g L^{-1} (Fig.34).

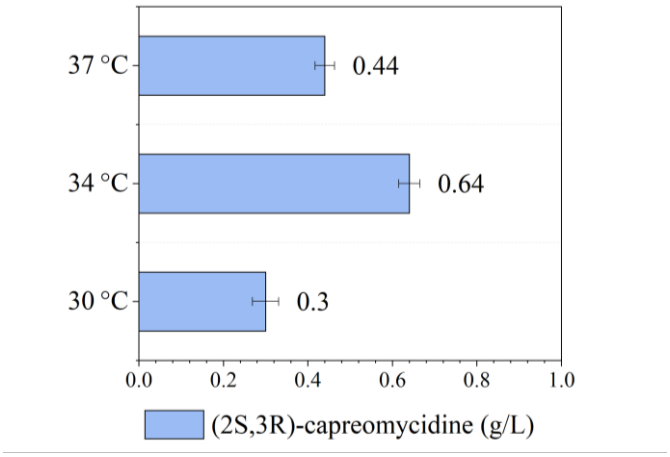


Figure 34. Impact of cultivation temperature on (2S,3R)-capreomycidine biosynthesis by *C. glutamicum* Arg0-BCD-Cap strain in shake flasks.

To determine the maximum production capacity of (2S,3R)-capreomycidine, *C. glutamicum* Arg0-BCD-Cap was cultivated in fed-batch mode at the optimal temperature of 34 °C (Fig. 35). The batch phase lasted 18 hours, after which feeding was initiated to maintain the glucose concentration at 10 g L^{-1}. After 50 hours of fermentation, a final titer of 9.1 g L^{-1} of (2S,3R)-capreomycidine and 1.39 g L^{-1} of L-arginine was achieved. The overall yield for the production of (2S,3R)-capreomycidine reached 70 mg g^{-1} glucose, corresponding to a volumetric productivity of 182 mg L^{-1} h^{-1}.

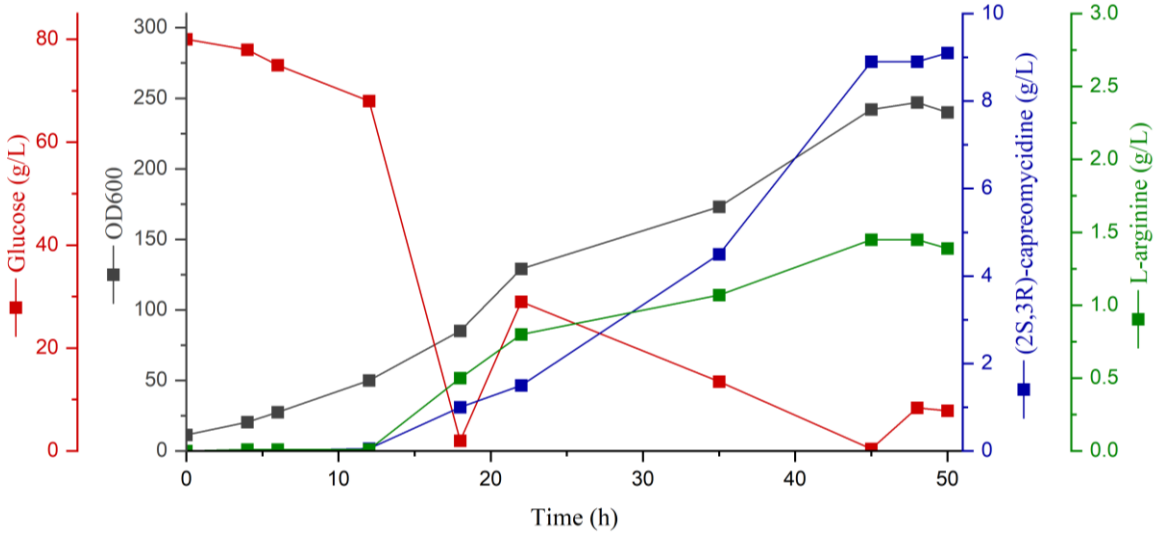


Figure 35. Fed-batch fermentation profile of *C. glutamicum* Arg0-BCD-Cap strain.

In contrast to biosynthetic pathway of L-enduracididine, the biosynthetic pathway of (2S,3R)-capreomycidine comprises only two enzymatic conversion steps. Together with the higher activity of the VioCD enzymes at elevated cultivation temperatures, this simpler pathway architecture results in higher overall productivity and improved precursor conversion. Notably, 86% of the produced L-arginine was converted into (2S,3R)-capreomycidine, demonstrating the high suitability of this pathway while also indicating that further optimization is required to achieve complete precursor utilization.

The supernatant from the fed-batch fermentation of C. glutamicum Arg0-BCD-Cap was processed for the isolation of (2S,3R)-capreomycidine following the same procedure outlined for L-enduracididine. After concentration by evaporation, acidification, and application to a cation exchange column, a single-step purification yielded 1 gram of the compound. NMR analysis confirmed its identity as (2S,3R)-capreomycidine (SI). The purified sample was subsequently used to prepare a calibration curve and employed for quantitative measurements.

Expression of the hydroxycapreomycidine biosynthetic pathway in arginine overproducer *C. glutamicum* ATTC21831

Hydroxycapreomycidine represents another promising non-proteinogenic building block. It is formed via post-NRPS hydroxylation of the capreomycidine residue in the viomycin biosynthetic intermediate and does not occur as a free amino acid within this pathway. However, the amino acid (4R)-hydroxycapreomycidine has been identified as a free intermediate in the streptolidine biosynthetic pathway. Two genes, orfP and orfR, encode enzymes responsible for its biosynthesis from L-arginine (Fig.36) [211].

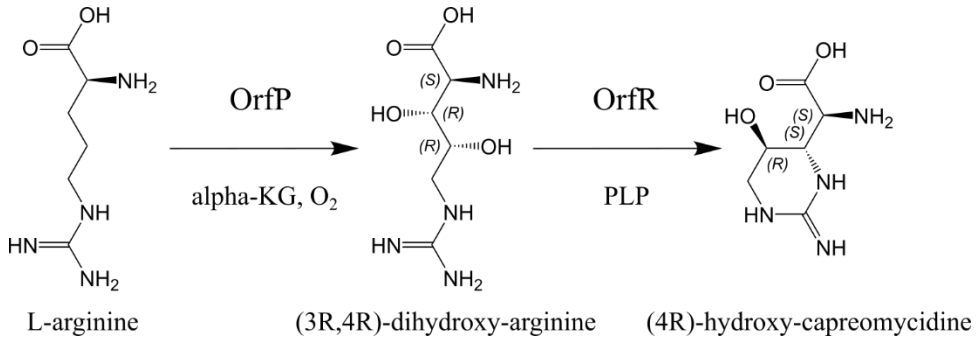


Figure 36. Biosynthetic pathway of (4R)-hydroxy-capreomycidine

The first enzyme, OrfP, is an α-ketoglutarate-dependent hydroxylase that converts L-arginine into (3R,4R)-dihydroxyarginine. In the subsequent step, the PLP-dependent enzyme OrfR

catalyzes an elimination/addition reaction, leading to cyclization of (3R,4R)-dihydroxyarginine and formation of (4R)-hydroxy-L-capreomycidine [211].

The *orfP* and *orfR* genes from *Streptomyces lavendulae* BRRC 12163 were codon-optimized and expressed under the control of a BCD element on the multicopy vector pClik5a in the arginine overproducer *C. glutamicum* Arg0. The resulting strain, *C. glutamicum* Arg0 BCD-hCap, was cultivated for 48 hours in shake flasks and subsequently analyzed (Fig.37).

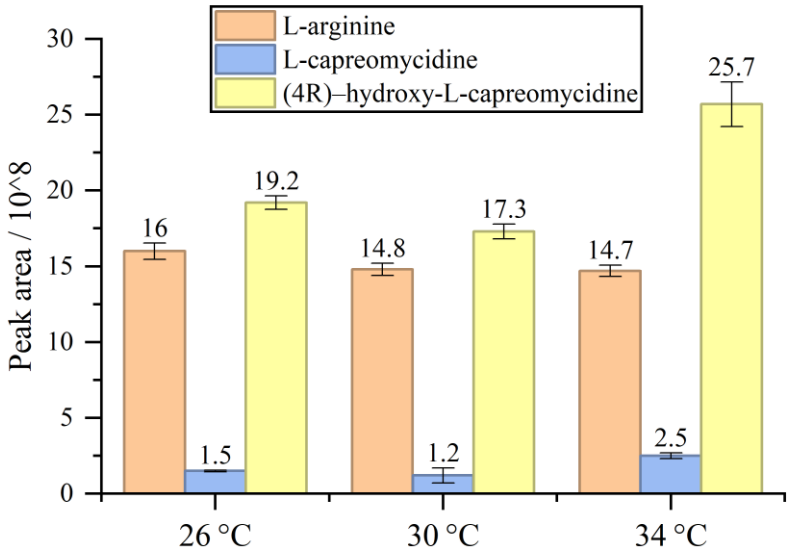


Figure 37. Production of L-arginine, L-capreomycidine, and (4R)-hydroxy-L-capreomycidine during shake flask fermentation of *C. glutamicum* Arg0-BCD-hCap at different temperatures.

The highest production of hydroxycapreomycidine was observed at 34 °C, although substantial production also occurred at lower temperatures, indicating that the enzymes OrfPR possess a broad temperature activity range. In addition to the main product, L-capreomycidine was detected in smaller quantities. This likely originates from partial enzyme promiscuity. Previous studies have shown that OrfP can generate singly hydroxylated arginine intermediates at either the C3 or C4 position. It is therefore plausible that 3-hydroxy-L-arginine is recognized as a substrate by OrfR, which subsequently catalyzes its cyclization to L-capreomycidine.

Expression of the kynurenine biosynthetic pathway in tryptophan overproducer *Corynebacterium glutamicum* ATTC21850

Kynurenine is a non-proteinogenic amino acid that serves as a conserved intermediate in the tryptophan catabolic pathway. It is also an integral component of the potent cyclic lipopeptide antibiotic daptomycin. During studies on the mode of action of daptomycin, the unique fluorescence properties of kynurenine, which depend on the polarity of its surrounding

environment, were utilized to monitor membrane targeting [212, 213]. This approach was later applied to calcium-dependent antibiotics (CDAs) produced by *Streptomyces coelicolor*, where tryptophan residues were replaced with kynurenine to use this spectroscopic property for membrane interaction studies [214]. Because of its intriguing physicochemical characteristics and occurrence in several peptide antibiotics, the catabolism of tryptophan in *S. coelicolor* was examined in greater detail. Two enzymes, tryptophan 2,3-dioxygenase (KynA) and kynurenine formamidase (KynB), were identified as key components of this pathway (Fig. 38) [215].

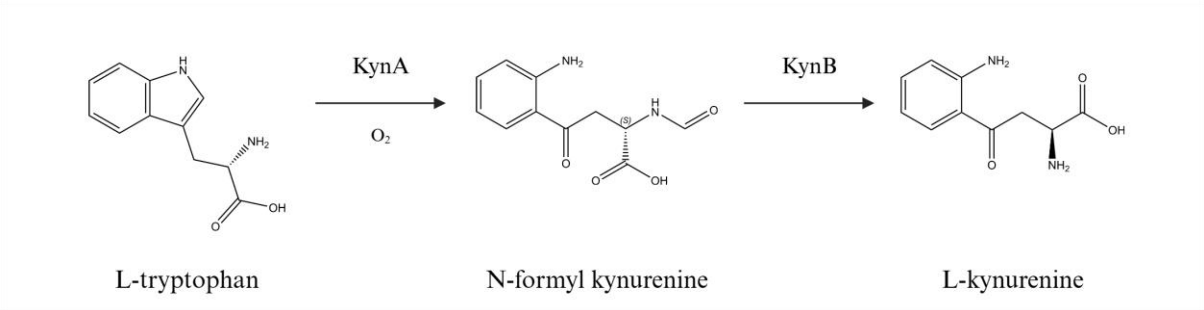


Figure 38. Biosynthetic pathway of L-kynurenine

To obtain this non-proteinogenic amino acid biotechnologically, the enzymes KynA and KynB were introduced into the tryptophan overproducer *C. glutamicum* ATCC 21850. The *kynA* and *kynB* genes were codon-optimized and expressed under the control of a BCD element on the multicopy vector pClik5a. The resulting strain, *C. glutamicum* Trp-Kyn, was cultivated in shake flasks for 48 hours at 30 °C and subsequently analyzed (Fig. 39).

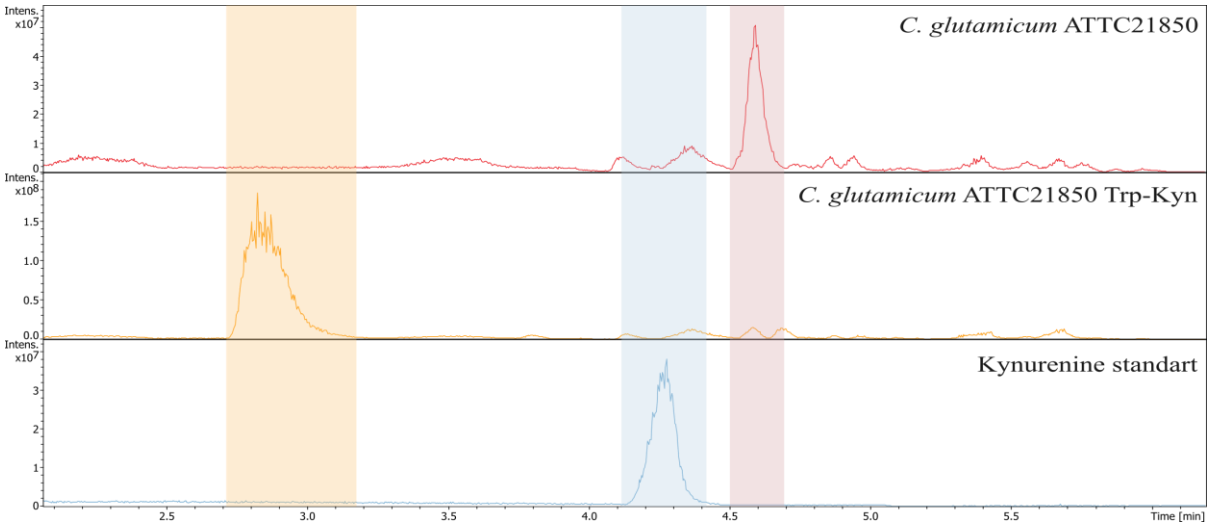


Figure 39. LC/MS chromatograms of *C. glutamicum* ATTC21850 and Trp-Kyn fermentation supernatant. The red bar indicates the retention time (RT) of L-tryptophan, which is abundant in the *C. glutamicum* ATCC 21850 control but nearly absent in the Trp-Kyn strain. The blue bar marks the RT of L-kynurenine, not detected in either strain. The yellow bar highlights a newly produced compound appearing exclusively in the Trp-Kyn fermentation supernatant, suggesting successful pathway redirection toward an alternative tryptophan-derived metabolite.

The data indicate that the newly engineered strain *C. glutamicum* Trp-Kyn produces only small amounts of extracellular tryptophan while simultaneously accumulating a high level of an unknown compound. This observation suggests that the compound is a tryptophan-derived metabolite, although not L-kynurenine, as its molecular mass and retention time do not correspond to those of the authentic standard. Kynurenine degradation is a conserved biosynthetic pathway, and *C. glutamicum* ATCC 21850 possesses homologs of the three main kynurenine degradation enzymes (Fig. 40).

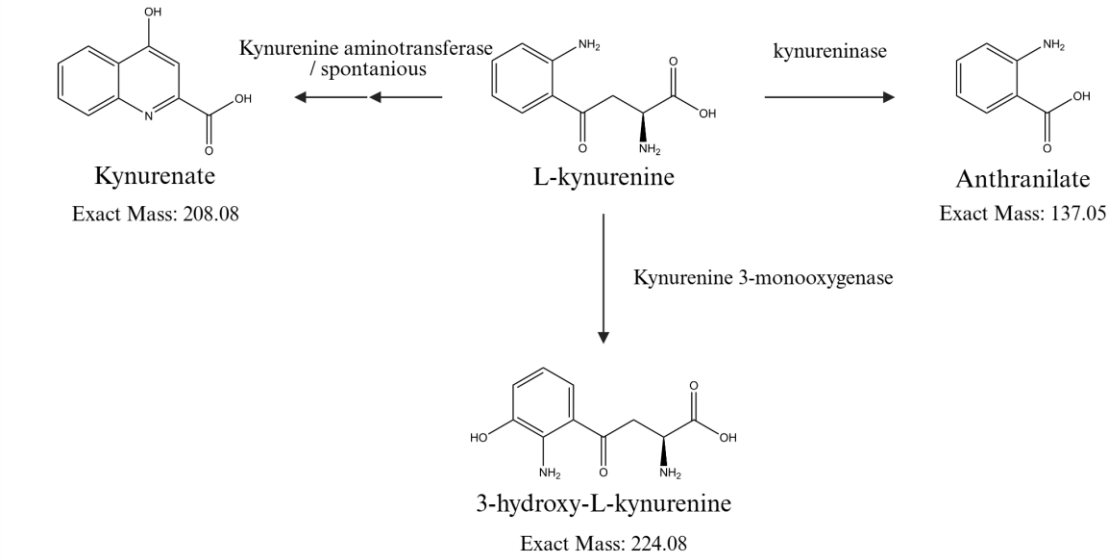


Figure 40. Kynurenine degradation pathways and initial metabolites with their exact masses.

The LC/MS data indicate that the unknown metabolite produced by *Corynebacterium glutamicum* Trp-Kyn matches the mass of kynurenate (kynurenic acid). This observation indicates the high activity of an endogenous kynurenine aminotransferase, which would convert intracellular kynurenine to kynurenate that is subsequently exported to the medium. To test this hypothesis, the next step is deletion of the gene encoding kynurenine aminotransferase and subsequent analysis of the obtained strain.

Expression of the ergothioneine biosynthetic pathway in *Corynebacterium glutamicum* BCA

Ergothioneine is widely recognized as a natural antioxidant and food preservative. However, ergothioneine and its precursor hercynine also hold potential as building blocks for the development of novel compounds. To evaluate this potential, bacterial production of these amino acids was first established. The biosynthesis begins with the conversion of histidine, during which EgtD and Egt1 catalyzes a SAM-dependent triple methylation of the α-amino nitrogen, forming hercynine. The enzyme Egt1 then conjugate hercynine with a cysteine utilizing molecular oxygen to generate hercynylcysteine sulfoxide. In the final step, the pyridoxal phosphate-dependent lyase Egt2 cleaves the cysteine residue at the sulfur atom, yielding ergothioneine. (Fig. 41) [216].

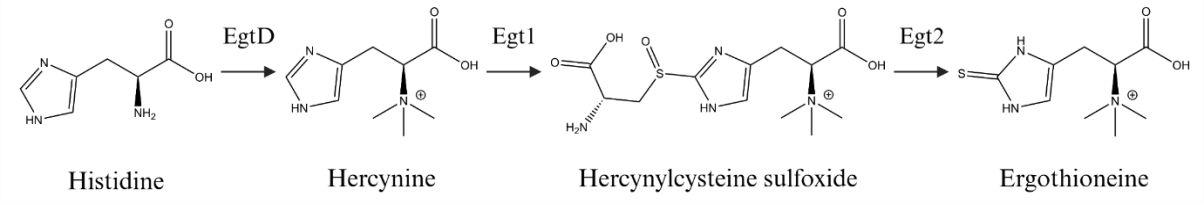


Figure 41. Biosynthetic pathway of ergothioneine

The genes *egtD* and *egt1-2* were codon-optimized and expressed under the control of a BCD element on the multicopy vector pClik5a in the *Corynebacterium glutamicum* BCA strain. The resulting strain, *C. glutamicum* BCA-BCD-Erg, was cultivated in shake flasks for 48 hours at 30 °C. Since this strain is not an overproducer of primary metabolites, the fermentation medium was supplemented with either histidine alone or a mixture of histidine and cysteine to a final concentration of 20 mM. The presence of target compounds was analyzed in both the biomass and the culture supernatant. For intracellular extraction, cells were collected by centrifugation, washed twice with Milli-Q water, and resuspended in 0.1 M HCl. The suspension was then heated at 100 °C for 20 minutes to release intracellular metabolites. After heating, the supernatant was adjusted to the original culture volume to maintain consistent concentration per liter and analyzed by LC/MS.

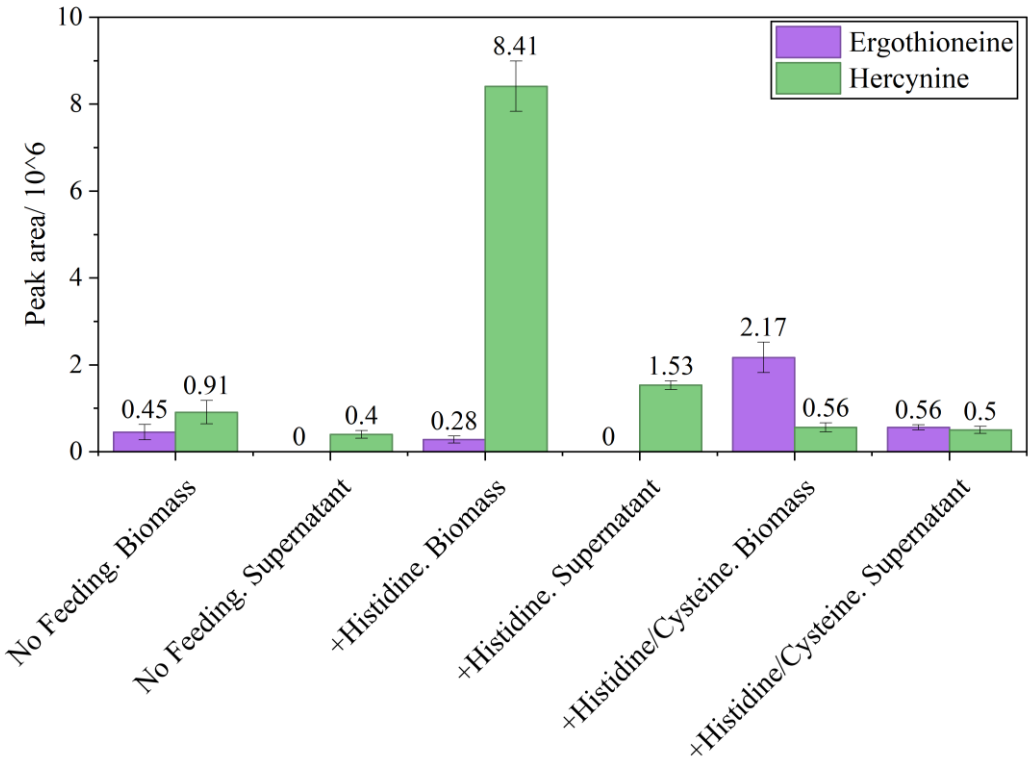


Figure 42. Production of hercynine, and ergothioneine during shake flask fermentation of *C. glutamicum* BCA-BCD-Erg with different feeding strategies.

According to the obtained data, *C. glutamicum* BCA-BCD-Erg was able to produce both hercynine and ergothioneine during shake-flask fermentation (Fig. 42). The majority of both compounds accumulated intracellularly, although a small amount of hercynine was also detected in the culture supernatant, indicating that the strain possesses an intrinsic mechanism for exporting this metabolite. Supplementation with histidine significantly increased the intracellular production of hercynine and led to a corresponding rise in its extracellular concentration. However, this condition negatively affected intracellular ergothioneine formation, possibly due to metabolic imbalance caused by the high extracellular histidine concentration. When both histidine and cysteine were supplied simultaneously, production of both hercynine and ergothioneine increased, and both compounds were detected in the supernatant, suggesting that the strain is capable of exporting them. The final extracellular ergothioneine concentration reached approximately 1 mg L^{-1}. These results demonstrate the potential of *C. glutamicum* BCA as a host strain for ergothioneine and hercynine biosynthesis, supported by a functional pathway and metabolite export capacity. Furthermore, the feeding experiments highlight the need to enhance intracellular availability of histidine and cysteine to achieve higher product titers.

Expression of the dimethylcadaverine biosynthetic pathway in lysine overproducer *Corynebacterium glutamicum* ATTC21253

The recently discovered alkaloid terpenoid flavunoidine features a tetracyclic oxygenated sesquiterpene core that is conjugated to two amino-derived moieties, 5,5-dimethyl-L-pipecolate and dimethylcadaverine. The attachment of 5,5-dimethyl-L-pipecolate occurs through a hybrid NRPS-associated conjugation reaction, while dimethylcadaverine formation is catalyzed by the repurposed fungal terpene cyclase FlvF [217].

Dimethylcadaverine is generated through two consecutive enzymatic reactions that use L-lysine as the precursor. In the first step, the SAM-dependent N-methyltransferase FlvH catalyzes double methylation of the ε-amino group, producing Nε,Nε-dimethyl-L-lysine. In the second step, the PLP-dependent decarboxylase FlvG converts this intermediate into dimethylcadaverine (Fig. 43). Both intermediates, Nε,Nε-dimethyl-L-lysine and dimethylcadaverine, represent promising building blocks for novel compound development.

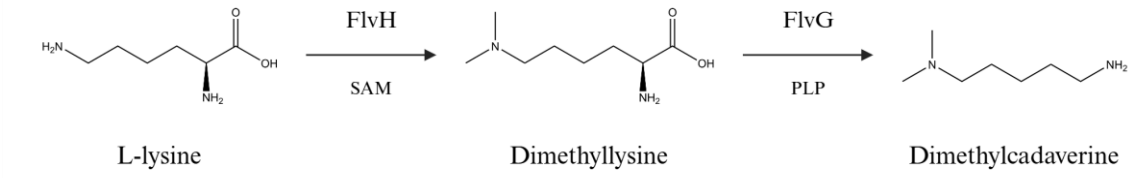


Figure 43. Biosynthetic pathway of Nε,Nε-dimethyl-L-lysine and dimethylcadaverine.

To establish biotechnological production of these metabolites, the *flvG* and *flvH* genes from *Aspergillus flavus* NRRL 3357 were codon-optimized and expressed under the control of a BCD element on the multicopy vector pClik5a in the lysine-overproducing strain *Corynebacterium glutamicum* ATCC 21253. The resulting strain, *C. glutamicum* Lys-Cad, was cultivated in shake flasks for 48 hours at 30 °C and subsequently analyzed by LC/MS (Fig. 44).

Despite the high-titer production of the precursor L-lysine by *C. glutamicum* Lys-Cad, no signal corresponding to dimethylcadaverine was detected, and only a minor peak matching the mass of protonated dimethyllysine was observed in the fermentation supernatant. Further strain development, enzyme optimization, and detailed compound purification and structural elucidation are required for the continued advancement of this project.

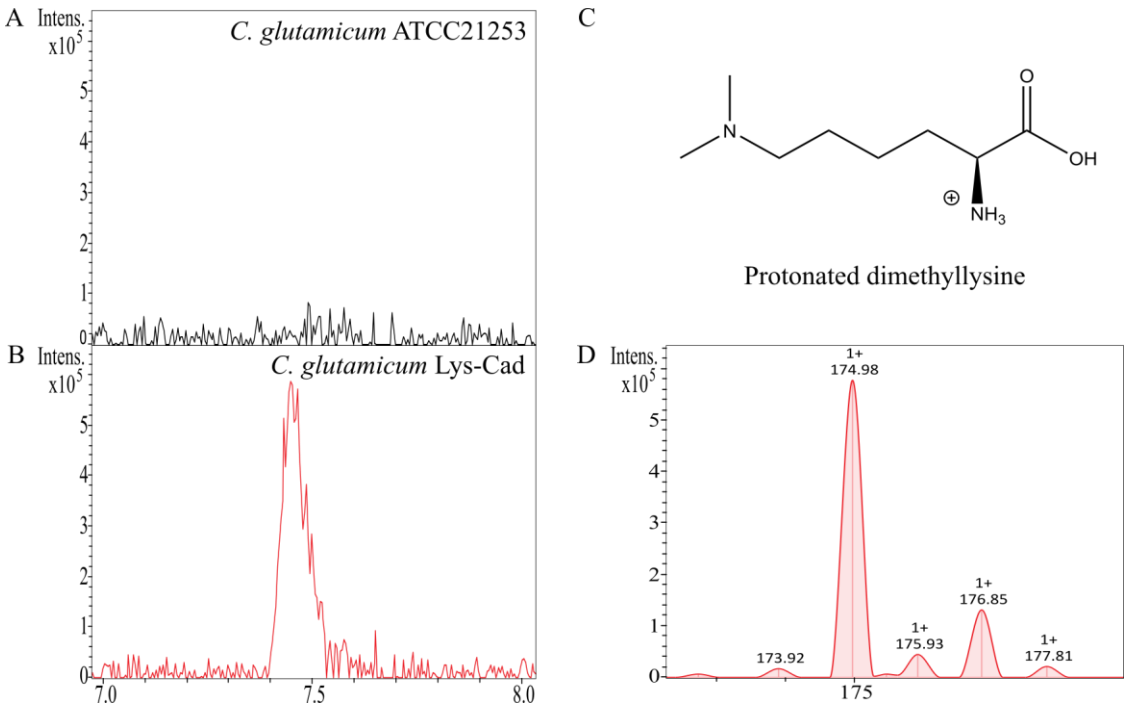


Figure 44. LC/MS chromatograms of *C. glutamicum* ATTC21253 and Lys-Cad fermentation supernatants. (A, B) Extracted ion chromatograms in the retention time range of 7.0–8.0 min from the fermentation supernatants of *C. glutamicum* ATCC 21253 (black) and Lys-Cad (red) strains, respectively. (C) Protonated structure of dimethyllysine. (D) Mass spectrum of the peak detected exclusively in the Lys-Cad strain. The observed m/z value corresponds to the theoretical mass of protonated dimethyllysine, confirming its formation during fermentation.

Conclusion and Outlook

The enduracididine biosynthetic pathway, composed of the three genes *mppPQR* from *Streptomyces wadayamensis*, was successfully expressed in *Streptomyces albus* Del14, enabling fermentative production of this amino acid. The obtained strain can be further optimized to increase production efficiency and may serve as a platform for in vivo incorporation of enduracididine into macromolecules through actinomycete-specific systems such as nonribosomal peptide synthetases (NRPS).

To achieve higher titers of this valuable metabolite, the *mppPQR* genes were subsequently optimized for expression in arginine-overproducing strain *Corynebacterium glutamicum* ATCC 21831. Implementation of a bicistronic expression element ensured efficient translation of all biosynthetic enzymes. The resulting strain, *C. glutamicum* Arg0-BCD-End, produced 1.7 g L^{-1} of L-enduracididine during a 50-hour fed-batch fermentation.

The same engineering pipeline was applied to express the capreomycidine biosynthetic pathway, consisting of the two genes *vioCD* from *Streptomyces vinaceus*. The resulting strain, *C. glutamicum* Arg0-BCD-Cap, produced 9.1 g L^{-1} of L-capreomycidine during a 50-hour fed-batch fermentation conducted at an elevated temperature of 34 °C. *C. glutamicum* ATCC 21831 accumulates higher levels of the precursor L-arginine under these temperature conditions. In contrast to the *mppPQR*-encoded enzymes, VioCD remained functional or even showed improved activity at 34 °C, resulting in significantly higher production titers.

Both L-enduracididine- and L-capreomycidine-containing supernatants obtained after fed-batch fermentations were collected and subjected to isolation and purification using strong cation exchange chromatography. The purification process employed exclusively water-based mobile phases, supporting the sustainable character of the production. Compound identity and purity were subsequently verified by NMR spectroscopy. The developed technology provides reliable access to these rare cyclic guanidinium building blocks and facilitates their future application in the synthesis of novel bioactive molecules. Furthermore, these compounds are already being utilized by collaborating partners for incorporation into NRPS- and RiPP-derived peptides.

The next development stages for the production of L-enduracididine- and L-capreomycidine should focus on metabolic engineering of *C. glutamicum* ATCC 21831 to redirect the metabolic flux more efficiently toward arginine biosynthesis, as precursor availability is clearly the main

bottleneck for achieving higher titers. In the case of L-enduracididine, the biosynthetic enzymes exhibit relatively low conversion efficiency and stability at higher temperature fermentation conditions. Therefore, protein optimization is required to improve enzyme stability at elevated temperatures and enhance catalytic activity within the intracellular environment of *Corynebacterium*.

Further expansion of the production platform enabled the bio-production of hydroxylated variants of L-enduracididine and L-capreomycidine through expression of additional pathway genes. The gene *mppO*, originating from the same mannopeptimycin biosynthetic gene cluster, was co-expressed with *mppPQR*. The resulting strain, *C. glutamicum* Arg0-BCD-hEnd, produced hydroxyenduracididine during shake-flask fermentation. Similarly, the genes orfPR from Streptomyces lavendulae, responsible for hydroxycapreomycidine biosynthesis, were introduced into *C. glutamicum* ATCC 21831. The obtained strain, *C. glutamicum* Arg0-BCD-hCap, produced hydroxycapreomycidine under shake-flask conditions. The hydroxy variants were analyzed by liquid chromatography and verified by mass spectrometry. However, this analysis was insufficient to determine the exact stereochemistry of the compounds. Further development of preparative purification methods is planned to enable NMR measurements, complete stereochemical characterization, and isolation of the purified compounds, which is essential for granting reliable access to these amino acids.

To assess the versatility of the established production system, several additional biosynthetic pathways from diverse origins were expressed in *C. glutamicum* strains. The genes *flvHG* from *Aspergillus flavus*, responsible for the biosynthesis of dimethylcadaverine and dimethyllysine, were cloned into the lysine-overproducing strain *C. glutamicum* ATCC 21253. The resulting strain produced only small amounts of dimethyllysine, which were detected by liquid chromatography and mass spectrometry, while no traces of dimethylcadaverine were observed. These results indicate that precursor supply is not the main limitation for the biosynthesis of these metabolites. To optimize and enhance production of the desired compounds, further experiments are required. The first step involves verification of enzyme translation levels to ensure adequate expression. Purified enzymes should then be used for *in vitro* assays to determine optimal catalytic parameters. As the corresponding genes originate from a distantly related genus, protein engineering may be necessary to adapt the enzymes to the intracellular environment of *Corynebacterium*. Additionally, alternative hosts such as *Escherichia coli* or *Saccharomyces cerevisiae* could be evaluated for improved enzyme activity and product formation.

Another example involved expression of the genes *egtD* from *Mycolicibacterium smegmatis* and *egt1* and *egt2* from *Neurospora crassa* in the *C. glutamicum* BCA strain. The resulting strain produced small amounts of the desired products, ergothioneine and hercynine, which were mainly retained within the biomass. Subsequent experiments involving precursor feeding with histidine and cysteine increased production severalfold and promoted extracellular transport of both compounds, leading to their accumulation in the supernatant. These findings indicate that precursor availability represents one of the main limitations for production of the desired metabolites. Further metabolic engineering to enhance sulfur assimilation and simultaneous biosynthesis of cysteine and histidine may substantially improve overall yields.

The final case involved expression of the kynurenine biosynthetic genes *kynAB* from *Streptomyces coelicolor* in the tryptophan-overproducing strain *C. glutamicum* ATCC 21850. The engineered strain converted nearly all produced tryptophan into a new metabolite. However, the mass and retention time of this compound did not match those of the kynurenine standard. Instead, the observed mass corresponded to one of the downstream products of kynurenine degradation, suggesting that native *C. glutamicum* enzymes involved in tryptophan catabolism further converted the desired amino acid. Genome analysis of the host revealed several candidate enzymes potentially responsible for this conversion. Future work will include targeted knockouts of these genes to prevent undesired degradation and stabilize kynurenine accumulation.

This work demonstrated the feasibility of microbial production of cyclic guanidinium-containing non-proteinogenic amino acids. Together, obtained results illustrate the potential of metabolic engineering of microbial cell factories to provide sustainable and scalable access to chemically diverse building blocks, paving the way for their integration into the development of novel bioactive compounds. The presented data also highlight that each biosynthetic pathway introduces distinct challenges that must be addressed to achieve efficient production of the target metabolites. Future work should focus on a modular design of the key development stages to enable effective adaptation of new biosynthetic pathways. Establishing a strain library encompassing phylogenetically diverse hosts could accelerate identification of the optimal cellular environments for different enzyme systems. Incorporation of routine analyses of protein expression levels and in vitro enzyme assays will provide valuable insights into pathway bottlenecks and guide further improvements in biosynthetic efficiency.

Supplementary information

Tablse S1. Primers used in this study.

Primer	Sequences	Description
pRT-int-F	TTCGGCCCCTTTTTTGGCC	Check and sequence of the insertions into the pRT801 vector
pRT-int-R	CCCCAGGCTTTACACTTTATGC	
End-S-F	GCCCACTACGACCACTACG	Check integration of the pRT801-End-s
End-S-R	GACCATCACGACCTCGCC	
End-S-1R	CGTAGTGGTCGTAGTGGGC	Sequence of *mppPQR* genes
End-S-2R	CATCACCTGCTCCGGGCC	
End-S-3R	CGATCCTCATCCTCGACGC	
End-S-4R	GGCGAGGTCGTGATGGTC	
End-S-1F	GAAGGTCAGGGTGTGGGC	
pC-1F	AACTGGATGGCTTTCTTGCC	Sequence of pClik5a vector
pC-2F	ACCAAGCGAAACATCGCATC	
pC-3F	CACTCAAAGGCGGTAATACGG	
pC-4F	TGATCCGGCAAACAAACCAC	
pC-5F	AATCAGTTGCGCCTACTGC	
pC-6F	AATTTCCAGGGCAGCAAGAC	
pC-7F	GCAGGAGTTTTCTAGCGGAC	
End-C-1F	CCGAGTTGAACATCGCGGA	Sequence of the codon-optimized *mppPQR* genes for *C. glutamicum*
End-C-2F	TATTCTGCTCGGCGTGTCC	
End-C-3F	ACCTACGGCGAAAACCAGG	
End-C-4F	CGGCGGTGCATTGAATCAC	
End-C-5F	CTCTCCCCCGGTCCAGATA	
End-C-1R	TAGCAGGACAACACGCGAC	
Ptuf-F	AAAGTCGACTGGCCGTTACCCTGCGAATG	Amplification of the promotor region of *tuf* gene
Ptuf-R	CTTTCAGTTGTGGCTGGGTAGTCATTGTATGTC CTCCTGGACTTCGTGGT	
Insert-F	ACCACGAAGTCCAGGAGGACATACAATGACT ACCCAGCCACAACTGAAAG	Amplification of the codon-optimized *mppPQR* genes under the control of the promotor region of *tuf* gene
Insert-R	AAATCTAGATTAATGATGGTCGGTAGTCAAGA CG	
End-B-1F	CCGAGTTGAACATCGCGGA	Sequence of the codon-optimized *mppPQR* genes for *C. glutamicum* under control of the BCD element
End-B-2F	TATTCTGCTCGGCGTGTCC	
End-B-3F	ACCTACGGCGAAAACCAGG	
End-B-4F	CGGCGGTGCATTGAATCAC	
End-B-5F	CTCTCCCCCGGTCCAGATA	
End-B-1R	TTCCCTTGTCCAGATAGCCC	
End-B-2R	TAGCAGGACAACACGCGAC	
Cap-B-1F	CACCCTGGACGTGCTCTTT	Sequence of the codon-optimized *vioCD* genes for *C. glutamicum* under control of the BCD element
Cap-B-2F	TCCGCGAGCGTTACTTTCA	
Cap-B-3F	CCCTGTCCAAGGCATACGG	
Cap-B-4F	TGGCCGACAATCAATGAAGC	
Cap-B-1R	GGGAGCTCATGCGCGATAA	
MppO-1F	GTGCCGTCTTGACTACCGAC	Sequence of the codon-optimized *mppO* gene
MppO-2F	TCTATGCAGACGAGCTCCCT	
MppO-3F	GATGACGAGACCCGCGAAA	
hCap-1F	ACATACAGTGGCAAAGGCGA	Sequence of the codon-optimized *orfPR* genes
hCap-2F	GGTATGGGCTCCAAGCAACT	
hCap-3F	TACCCTGCGAATGTCCACAG	
hCap-4F	CTGATCGTGTGCTGGTCACT	

hCap-5F	TAATCTCGCCCACTTGACCG	Sequence of the codon-optimized *orfPR* genes
flvGH-1F	ACATACAGTGGCAAAGGCGA	Sequence of the codon-optimized *flvGH* genes
flvGH-2F	CTCCTGGCCGGTAAGCAAAT	
flvGH-3F	TGAATGCGATGAACTCCGGT	
flvGH-4F	ATTATCGGCCGTCGCGATG	
Kyn-1F	TAAACATCGGCACCATCGGT	Sequence of the codon-optimized *kynAB* genes
Kyn-2F	GCGGTTACGACATCCCAGAT	
Kyn-3F	GCCAGAAGCTTCCATGCAAG	
Kyn-4F	CTCCCTGGCTGACCGACC	
Erg-1F	ATAGGCAGATAACGGTTCCCC	Sequence of the codon-optimized *egtD* and *egt1-2* genes
Erg-2F	GCCACCACGAAGTCCAGG	
Erg-3F	ATTTCGAGGAACACCTCGGC	
Erg-4F	TTGAAAATCAACGCCGTTGCCC	
Erg-5F	GTAACTTGCGCAAAGTTAATCTCC	
Erg-6F	AAGTACTCCAAAGAGGAATCCG	
Erg-7F	TCATCCCGACTCACATCCC	
Erg-8F	GCCGTTAATGCTCATCTGACC	
Erg-9F	TGTGCTCAGTCTTCCAGGC	
Erg-10F	GGACGGCAAAGATGAAATTCTG	
Erg-11F	AAGGACGCCATCAAATGGCG	
Erg-1R	GTTTACTTTGCAGGGCTTCCC	

Quantification of L-enduracididine and L-capreomycidine

Calibration curves were constructed based on the compound purity determined via NMR and are presented in Fig. S1 and S2. The measured samples were quantitatively diluted to ensure that the peak areas fell within the linear range of the calibration curves.

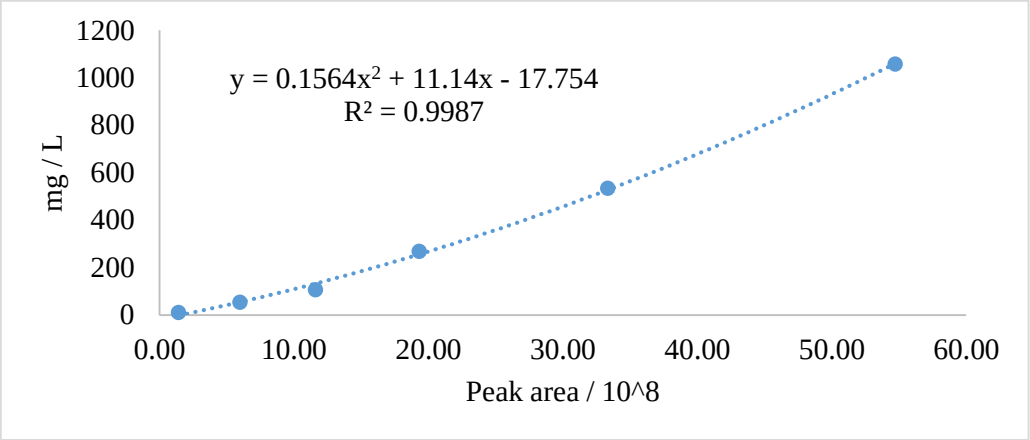


Figure S1. Enduracididine calibration curve.

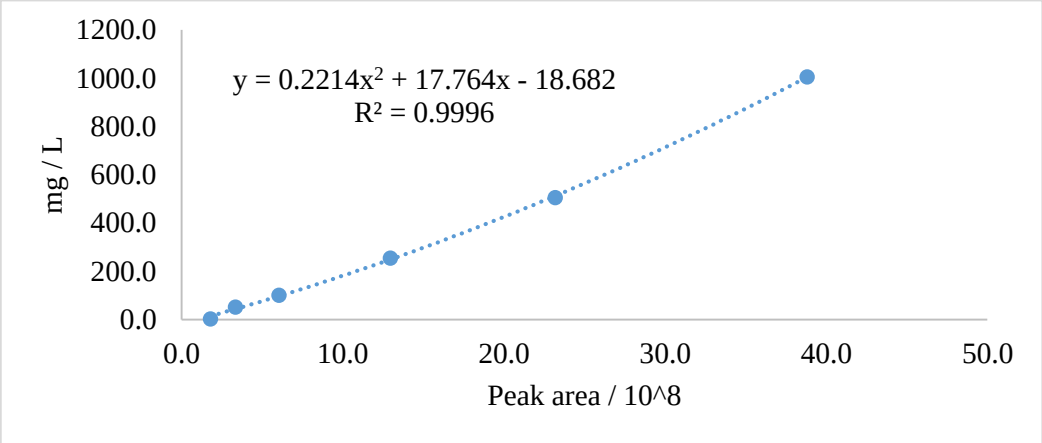


Figure S2. Capreomycidine calibration curve.

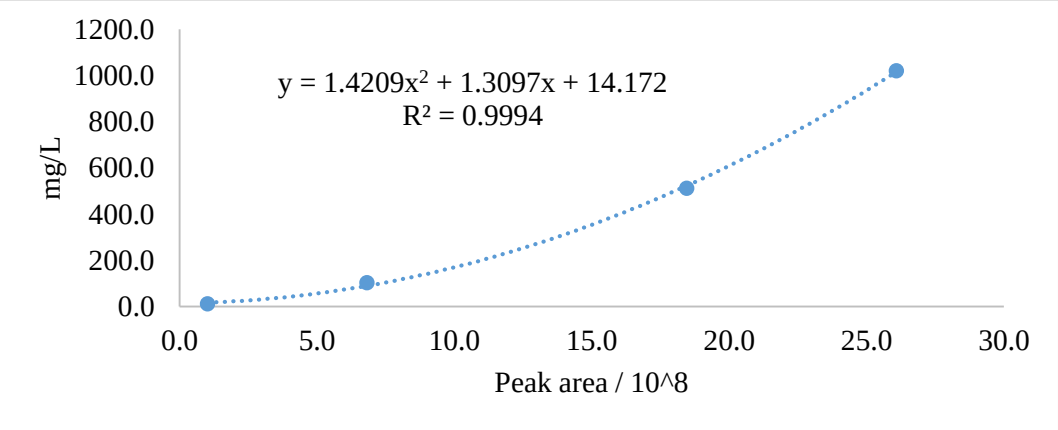


Figure S3. Ergothioneine calibration curve.

Structure confirmation for the L-enduracididine and L-capreomycidine

The structure of L-Enduracididine was confirmed by NMR and optical rotation measurements which were consistent with literature [218-220].

The structure of L-Capreomycidine was confirmed by NMR and Marfey's method.

Physical data

L-enduracididine:

White powder; [α]D20 +68° (C = 0.49 in H2O); 1H NMR (500 MHz, D2O) δ 4.09 (dq, J = 9.2, 6.8 Hz, 1H), 3.76 (t, J = 9.5 Hz, 1H), 3.33 (dd, J = 9.6, 6,9 Hz, 1H), 3.22 (dd, J = 7.5, 6.5 Hz, 1H), 1.89 – 1.79 (m, 2H); 13C NMR (125 MHz, D2O) δ 182.5, 159.9, 53.5, 53.3, 48.6, 39.9; 15N-HMBC (D2O) δ 93.8, 81.3, 31.0; HRMS (ESI-TOF) m/z: [M+H]+ Calcd for C6H13N4O2+ 173.1033 Da; Found 173.1032 Da.

L-capreomycidine:

White powder; 1H NMR (500 MHz, D2O) δ 3.43 (ddd, J = 8.8, 5.6, 4.6 Hz, 1H), 3.22 (dt, J = 12.8, 5.1 Hz, 1H), 3.14 (m, 1H), 3.10 (d, J = 5.8, 1H), 1.80 (dq, J = 13.7, 4.7 Hz, 1H), 1.64 (m, 1H); 13C NMR (125 MHz, D2O) δ 179.5, 154.3, 59.0, 52.0, 36.8, 22.5; 15N-HMBC (D2O) δ 81.1, 77.6, 22.3; HRMS (ESI-TOF) m/z: [M+H]+ Calcd for C6H13N4O2+ 173.1033 Da; Found 173.1032 Da.

NMR Spectrometer and Polarimeter

The chemical structures of all the compounds were determined via multidimensional NMR analysis. 1H-NMR, 13C-NMR, and 2D spectra were recorded at 500 MHz (1H) and 125 MHz (13C), conducted in the Bruker Avance Neo 500 MHz, equipped with a Prodigy Cryo-probe. Samples were dissolved in D2O. Chemical shifts are reported in ppm relative to tetramethylsilane; the solvent was used as the internal standard. Coupling constants are reported in Hertz (Hz). Multiplicity is reported with the usual abbreviations (s: singlet, d: doublet, dd: doublet of doublets, t: triplet, dq: doublet of quartets, m: multiplet). Chiroptical measurements of all the compounds in H2O ([α]D20) were obtained on a model Jasco P-2000 Automatic Digital Polarimeter (JASCO, Easton, MD, USA) in a 3.5 x 50mm cell at 20 °C.

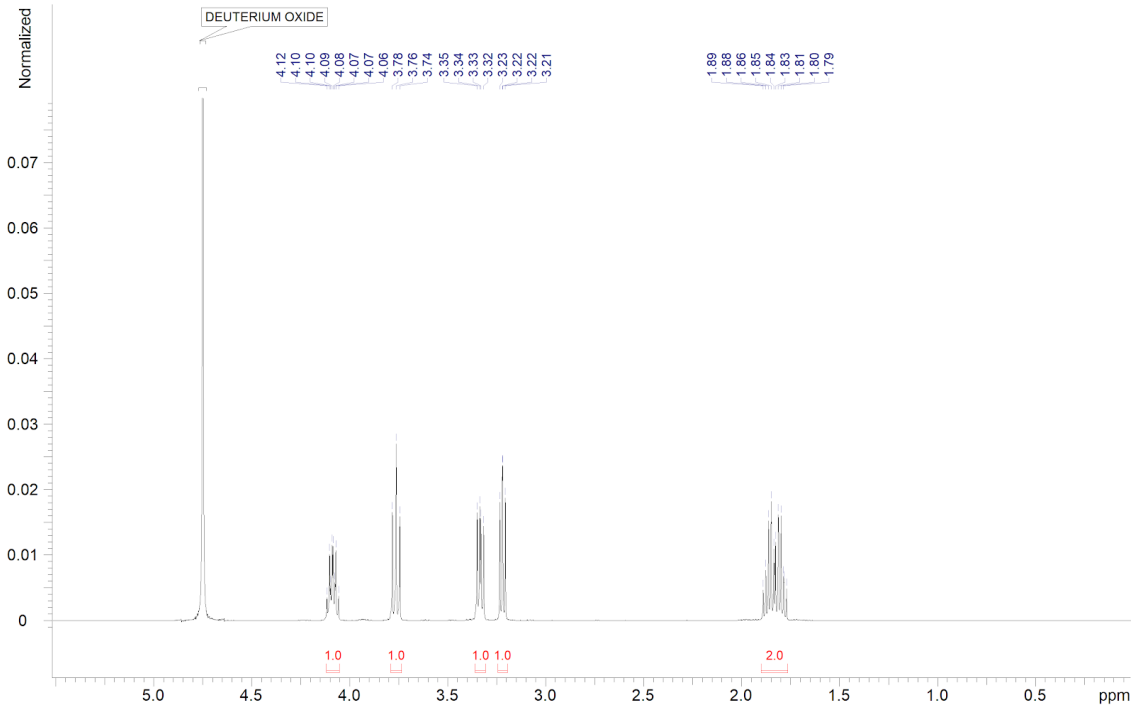


Figure S3: 1H NMR spectrum (500 MHz, D2O) of L-enduracididine.

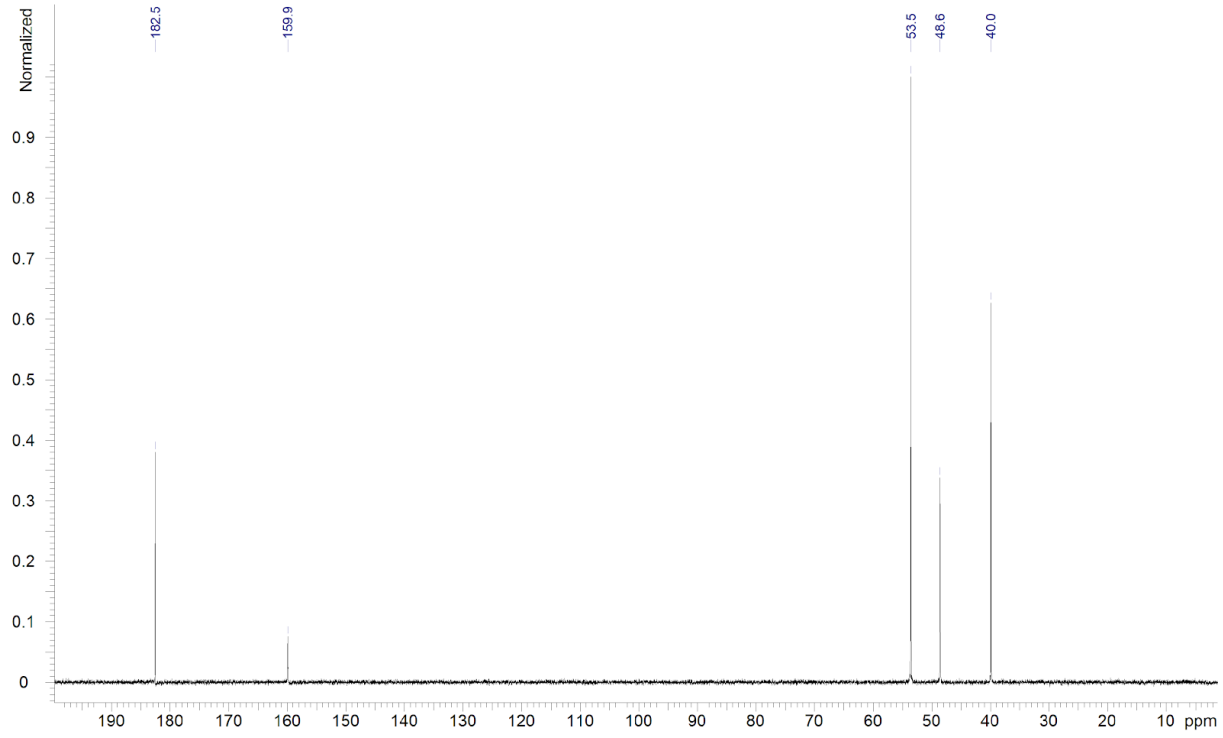


Figure S4: 13C NMR spectrum (125 MHz, D2O) of L-enduracididine.

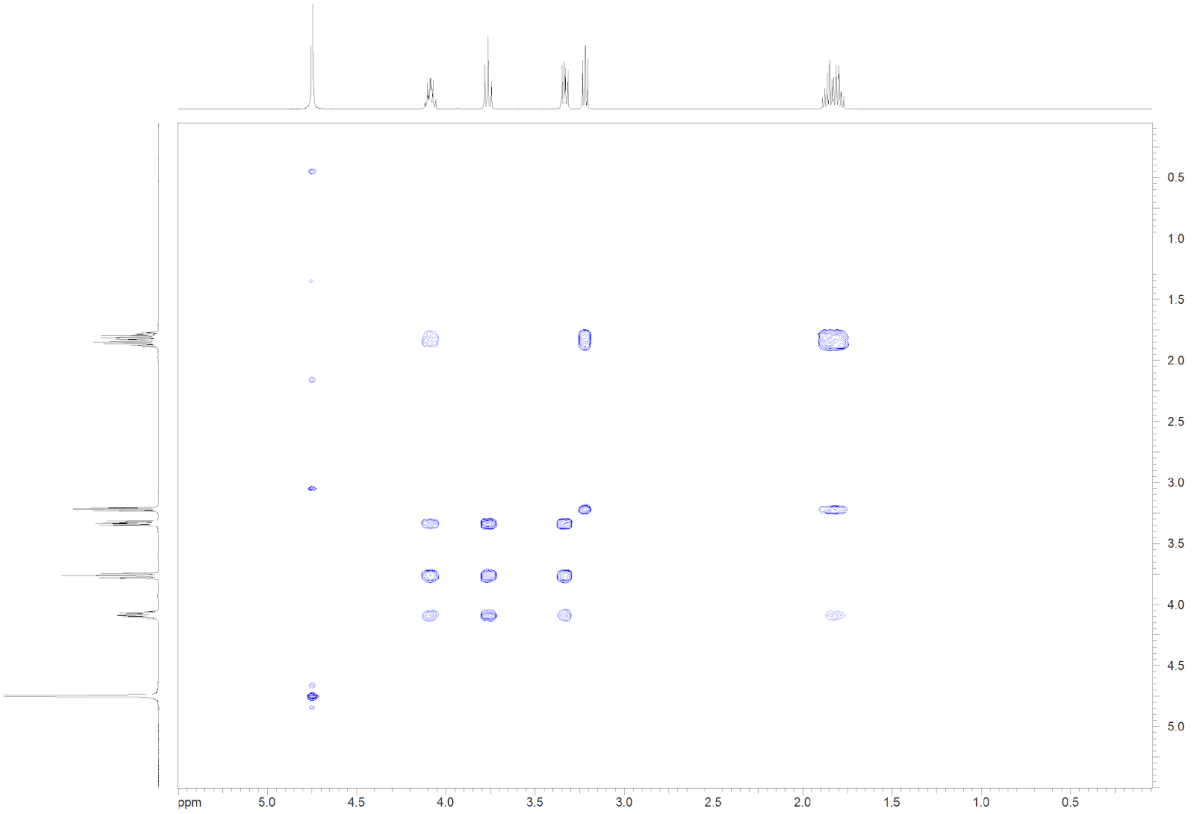


Figure S5. H-H COSY spectrum (D2O) of L-enduracididine.

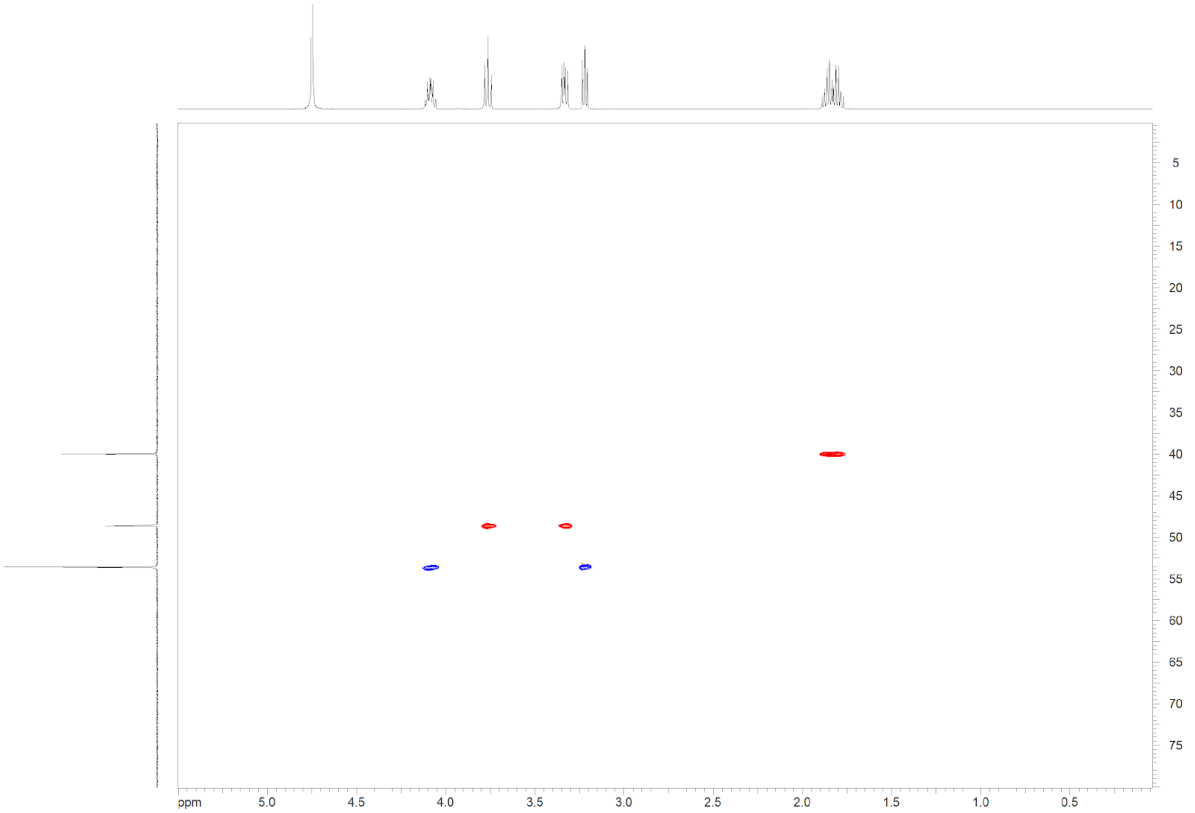


Figure S6. Edited-HSQC spectrum (D2O) of L-enduracididine.

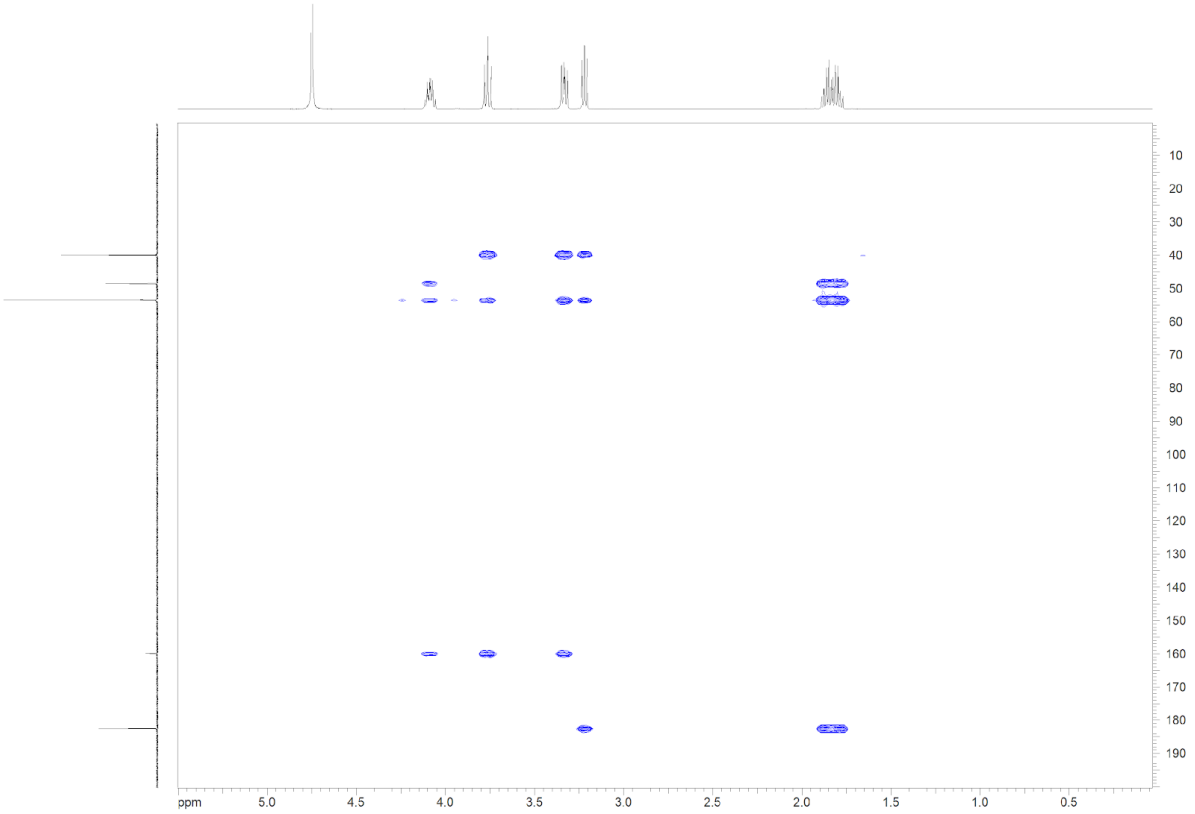


Figure S7. HMBC spectrum (D2O) of L-enduracididine.

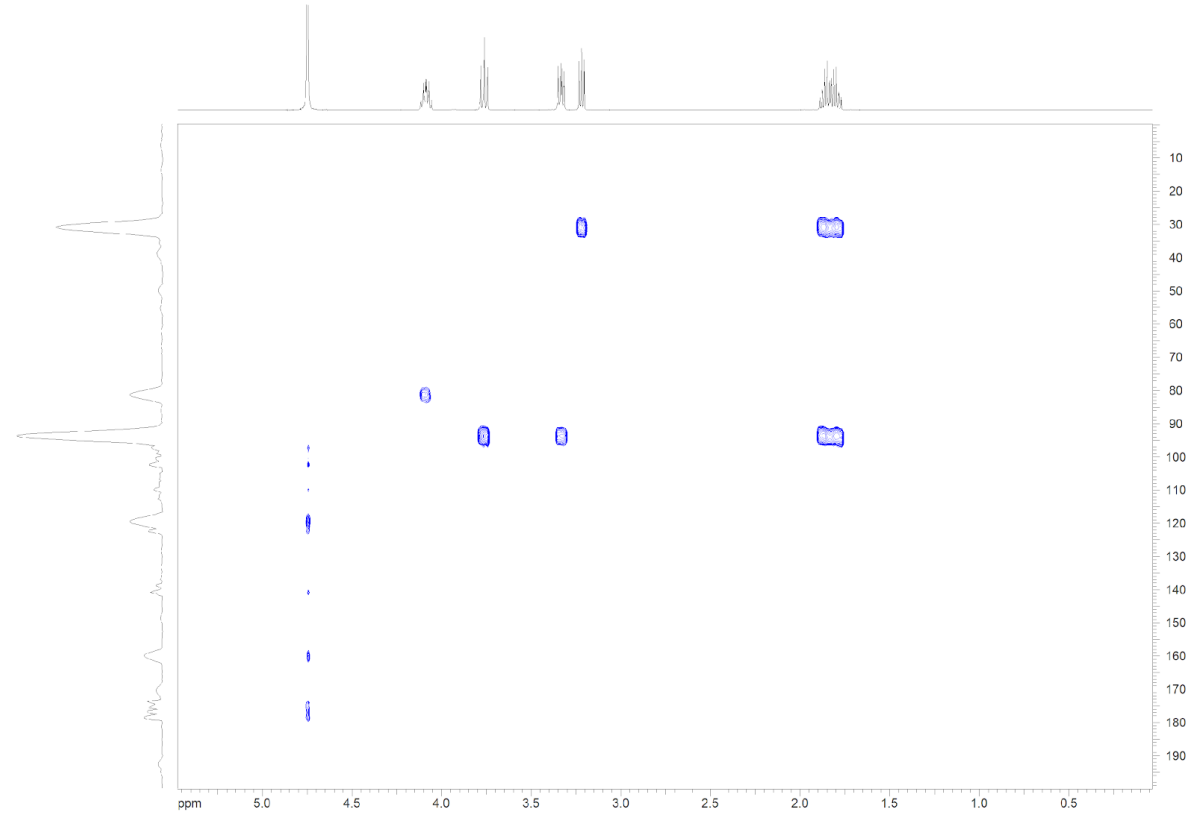


Figure S8. 15N-HMBC spectrum (D2O) of L-enduracididine.

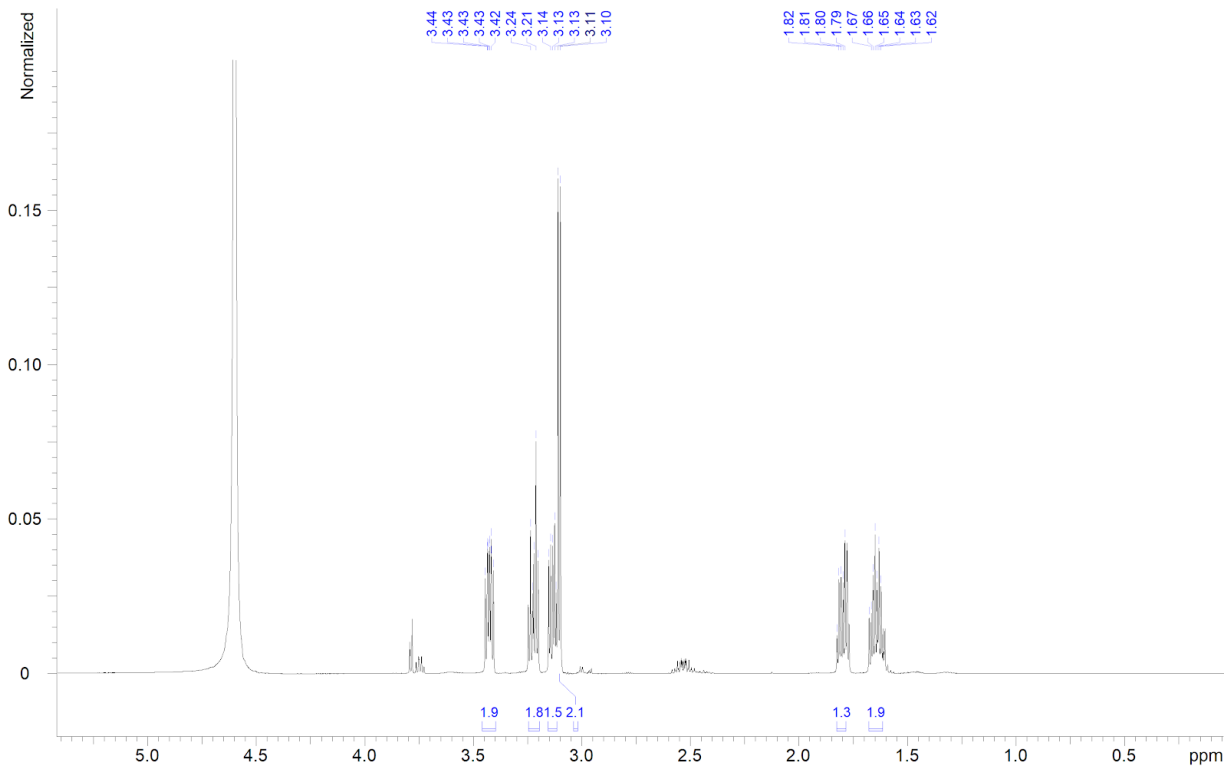


Figure S9. 1H NMR spectrum (500 MHz, D2O) of L-capreomycidine.

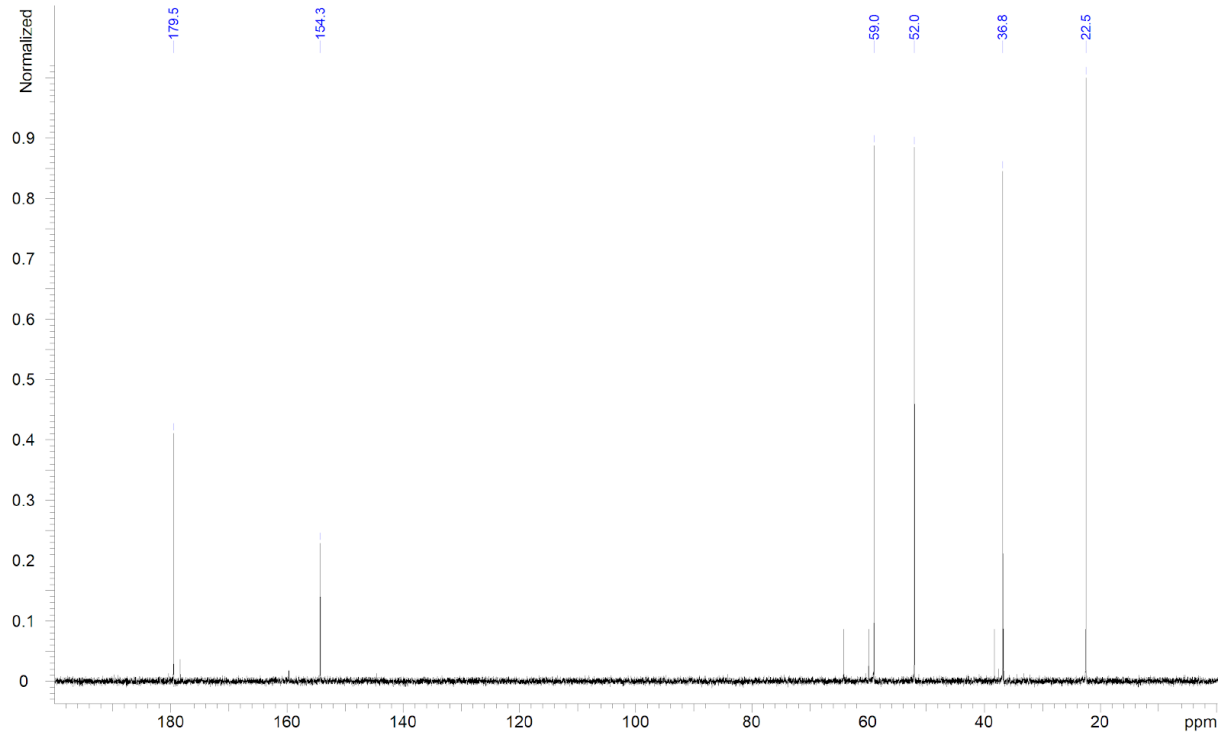


Figure S10. 13C NMR spectrum (125 MHz, D2O) of L-capreomycidine.

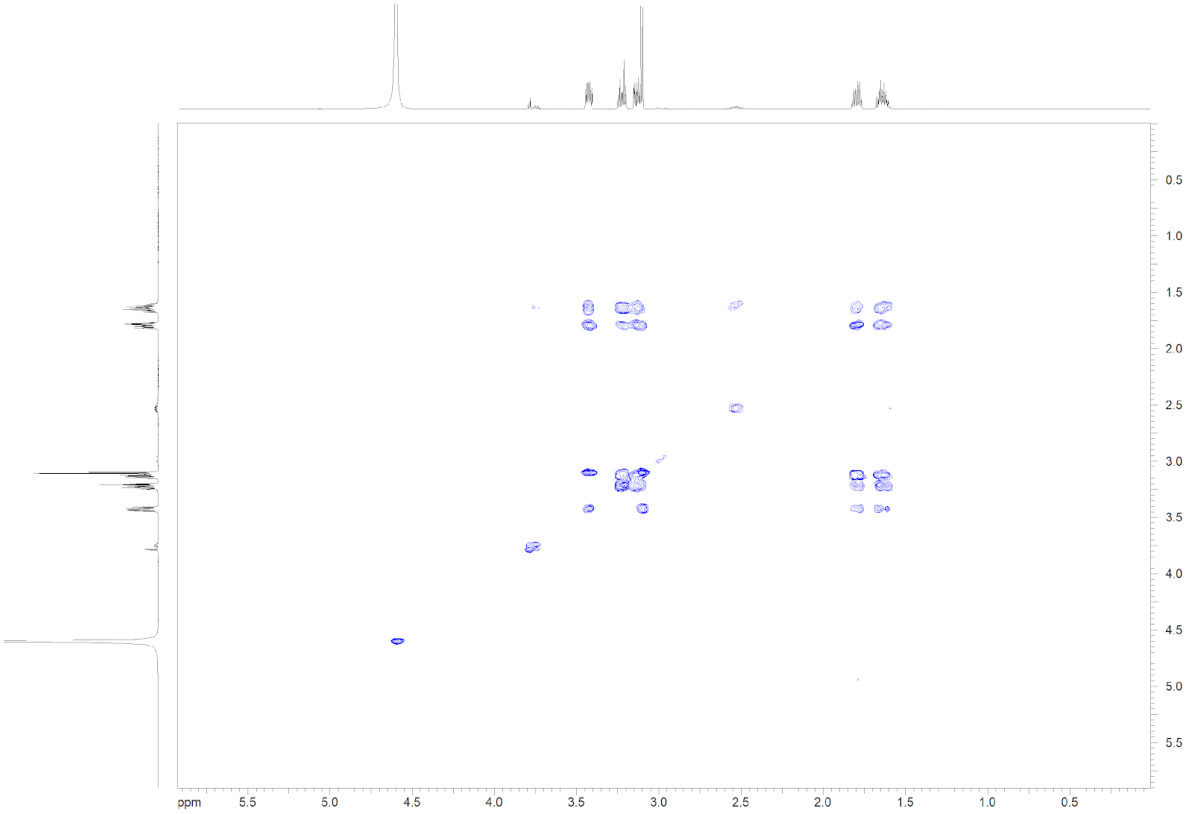


Figure S11. H-H COSY spectrum (D2O) of L-capreomycidine.

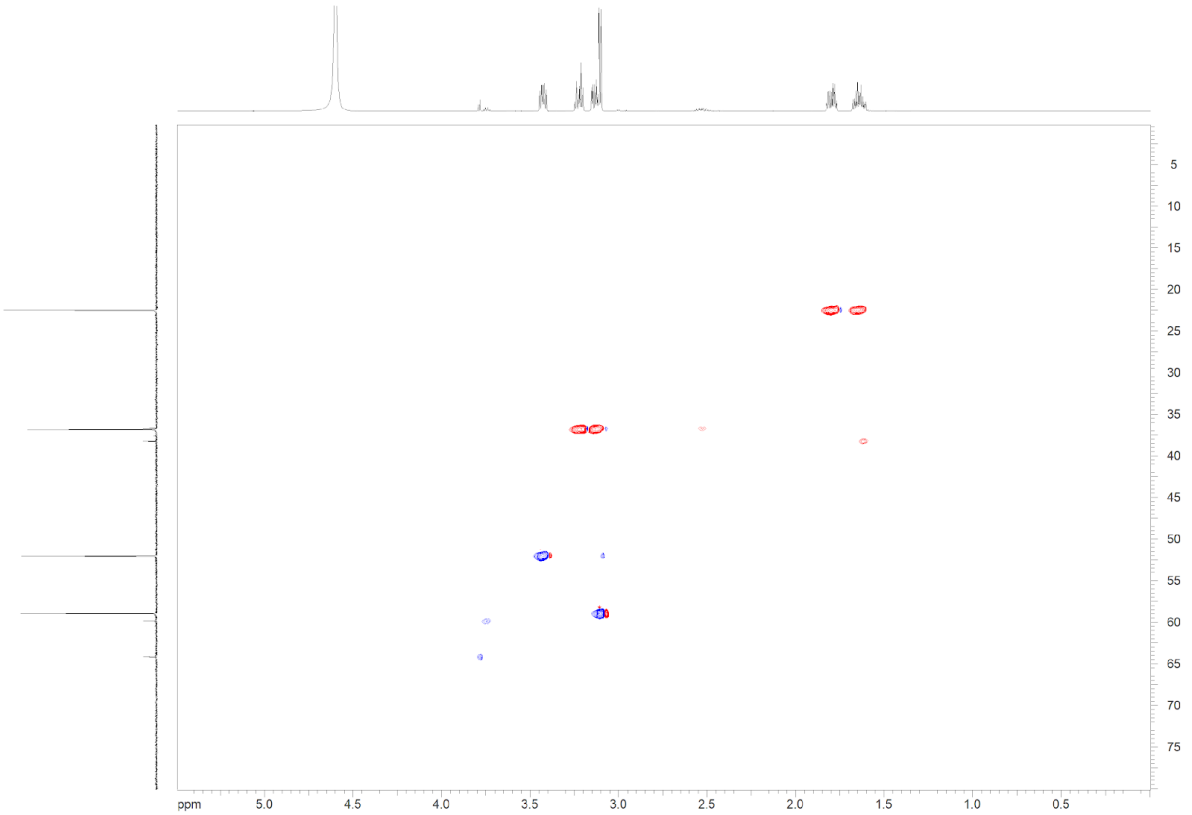


Figure S12. Edited-HSQC spectrum (D2O) of L-capreomycidine.

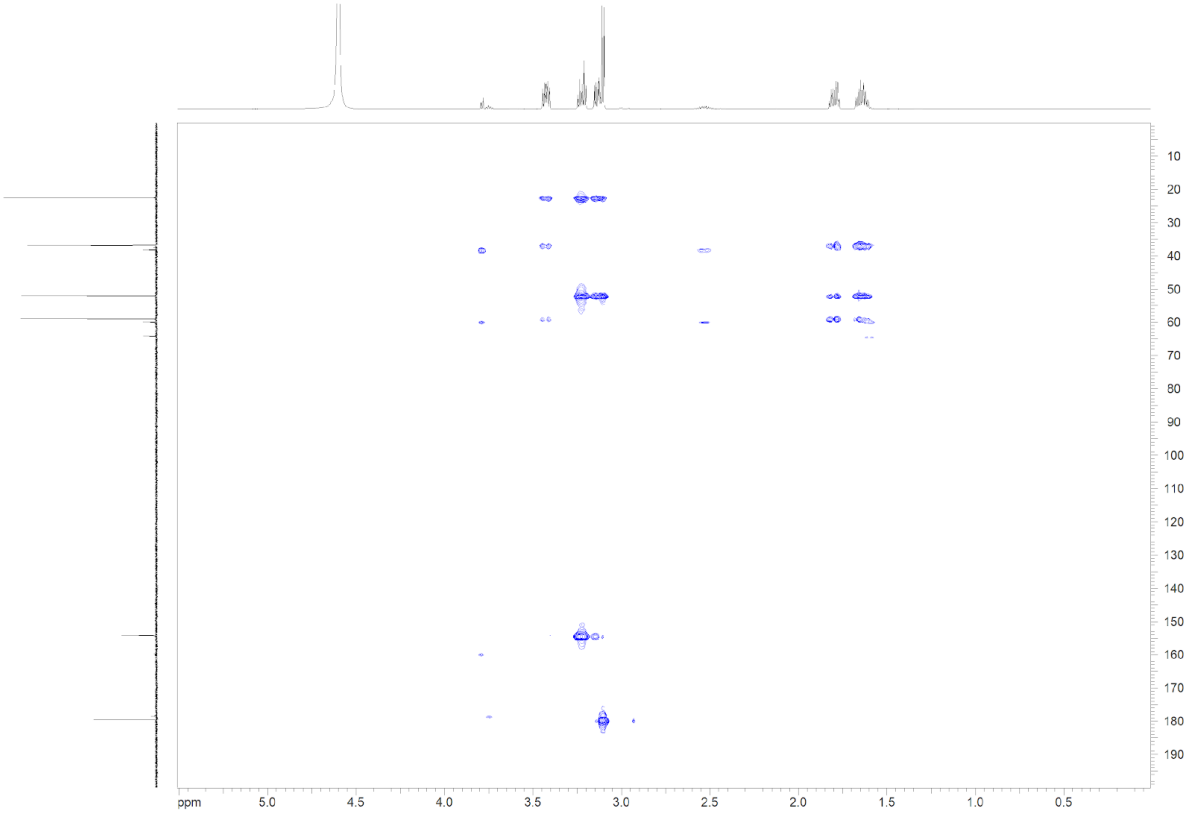


Figure S13. HMBC spectrum (D2O) of L-capreomycidine.

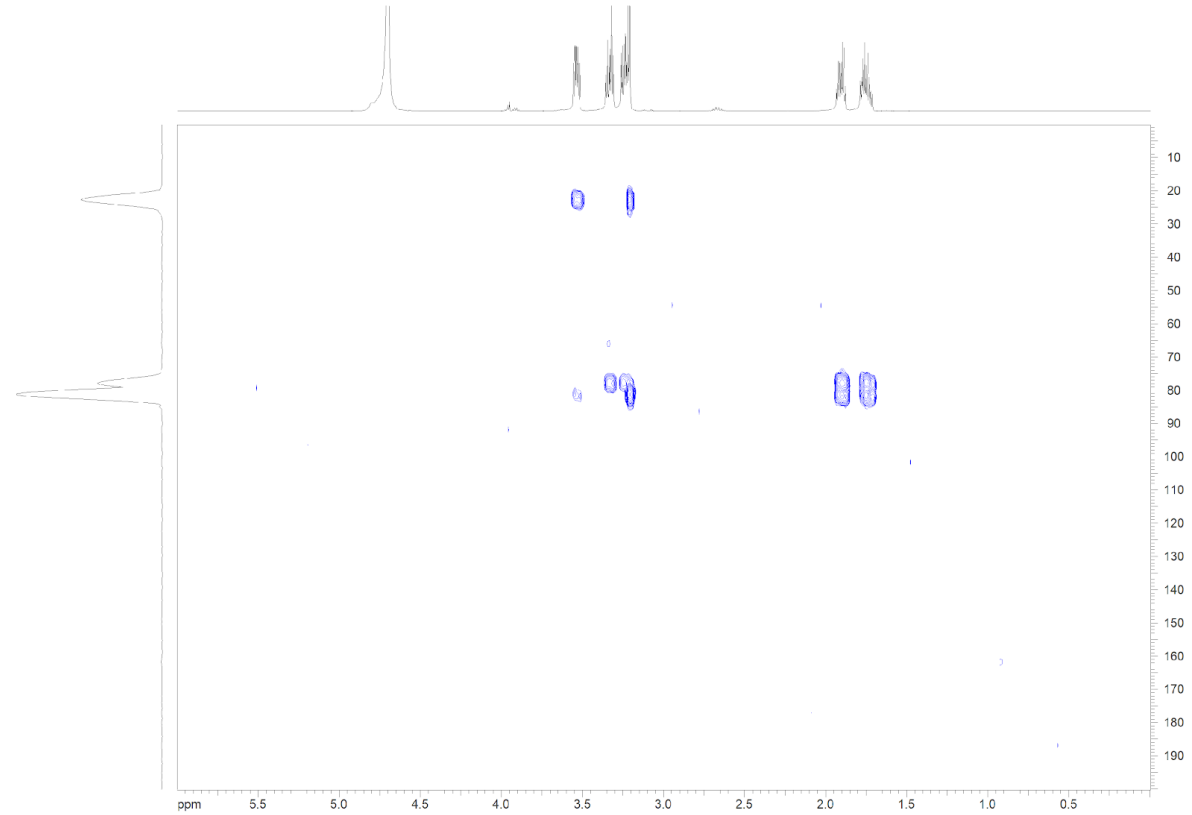


Figure S14. 15N-HMBC spectrum (D2O) of L-capreomycidine.

References

1. McNeely, J.A., *Nature and COVID-19: The pandemic, the environment, and the way ahead.* Ambio, 2021. **50**(4): p. 767-781.
2. Adenle, A.A., et al., *Global UN 2030 agenda: how can science, technology and innovation accelerate the achievement of sustainable development goals for all?* PLOS Sustainability and Transformation, 2023. **2**(10): p. e0000085.
3. Organization, W.H., *Global antimicrobial resistance and use surveillance system (GLASS) report 2022*. 2022: World Health Organization.
4. Salam, M.A., et al. *Antimicrobial resistance: a growing serious threat for global public health*. in *Healthcare*. 2023. MDPI.
5. Murray, C.J., et al., *Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis.* The lancet, 2022. **399**(10325): p. 629-655.
6. Organization, W.H., *WHO bacterial priority pathogens list, 2024: bacterial pathogens of public health importance, to guide research, development, and strategies to prevent and control antimicrobial resistance*. 2024: World Health Organization.
7. Casida, J.E., *Pest toxicology: the primary mechanisms of pesticide action.* Chemical research in toxicology, 2009. **22**(4): p. 609-619.
8. Deguine, J.-P., et al., *Agroecological crop protection for sustainable agriculture.* Advances in agronomy, 2023. **178**: p. 1-59.
9. Umetsu, N. and Y. Shirai, *Development of novel pesticides in the 21st century.* Journal of Pesticide Science, 2020. **45**(2): p. 54-74.
10. Szymański, P., M. Markowicz, and E. Mikiciuk-Olasik, *Adaptation of high-throughput screening in drug discovery—toxicological screening tests.* International journal of molecular sciences, 2011. **13**(1): p. 427-452.
11. Zhang, K., et al., *Artificial intelligence in drug development.* Nature medicine, 2025. **31**(1): p. 45-59.
12. Ishibashi, M., et al., *High-throughput simultaneous analysis of pesticides by supercritical fluid chromatography coupled with high-resolution mass spectrometry.* Journal of agricultural and food chemistry, 2015. **63**(18): p. 4457-4463.
13. Liu, R., X. Li, and K.S. Lam, *Combinatorial chemistry in drug discovery.* Current opinion in chemical biology, 2017. **38**: p. 117-126.
14. Satz, A.L., et al., *DNA-encoded chemical libraries.* Nature Reviews Methods Primers, 2022. **2**(1): p. 3.
15. Durant, G., et al., *The future of machine learning for small-molecule drug discovery will be driven by data.* Nature computational science, 2024. **4**(10): p. 735-743.
16. Erlanson, D.A., et al., *Twenty years on: the impact of fragments on drug discovery.* Nature reviews Drug discovery, 2016. **15**(9): p. 605-619.
17. Miethke, M., et al., *Towards the sustainable discovery and development of new antibiotics.* Nature Reviews Chemistry, 2021. **5**(10): p. 726-749.
18. Way, G.P., et al., *Evolution and impact of high content imaging.* Slas Discovery, 2023. **28**(7): p. 292-305.
19. Morato, N.M., et al., *Automated high-throughput system combining small-scale synthesis with bioassays and reaction screening.* Slas technology: translating life sciences innovation, 2021. **26**(6): p. 555-571.
20. De Stefano, P., E. Bianchi, and G. Dubini, *The impact of microfluidics in high-throughput drug-screening applications.* Biomicrofluidics, 2022. **16**(3).
21. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* nature, 2021. **596**(7873): p. 583-589.
22. Adelusi, T.I., et al., *Molecular modeling in drug discovery.* Informatics in Medicine Unlocked, 2022. **29**: p. 100880.

23. Galloway, W.R., A. Isidro-Llobet, and D.R. Spring, *Diversity-oriented synthesis as a tool for the discovery of novel biologically active small molecules.* Nature communications, 2010. **1**(1): p. 80.
24. Fitzgerald, P.R., et al., *Building block-centric approach to DNA-encoded library design.* Journal of Chemical Information and Modeling, 2024. **64**(12): p. 4661-4672.
25. Righetti, G.I.C., F. Faedi, and A. Famulari, *Embracing sustainability: The World of bio-based polymers in a mini review.* Polymers, 2024. **16**(7): p. 950.
26. Mohanty, A.K., et al., *Sustainable polymers.* Nature Reviews Methods Primers, 2022. **2**(1): p. 46.
27. Ding, Y., et al., *Impact of non-proteinogenic amino acids in the discovery and development of peptide therapeutics.* Amino Acids, 2020. **52**(9): p. 1207-1226.
28. Sharma, K.K., et al., *Unnatural Amino Acids: Strategies, Designs, and Applications in Medicinal Chemistry and Drug Discovery.* Journal of Medicinal Chemistry, 2024.
29. Huang, Y., et al., *Genetic Code Expansion: Recent Developments and Emerging Applications.* Chemical Reviews, 2024. **125**(2): p. 523-598.
30. Li, Y. and P.A. Dalby, *Engineering of enzymes using non-natural amino acids.* Bioscience Reports, 2022. **42**(8): p. BSR20220168.
31. Pagar, A.D., et al., *Non-canonical amino acid-based engineering of (R)-amine transaminase.* Frontiers in chemistry, 2022. **10**: p. 839636.
32. Chen, L., et al., *Advances in biosynthesis of non-canonical amino acids (ncAAs) and the methods of ncAAs incorporation into proteins.* Molecules, 2023. **28**(18): p. 6745.
33. Zhao, D., et al., *Melatonin synthesis and function: evolutionary history in animals and plants.* Frontiers in endocrinology, 2019. **10**: p. 441357.
34. Sutanto, C.N., et al., *The impact of 5-hydroxytryptophan supplementation on sleep quality and gut microbiota composition in older adults: A randomized controlled trial.* Clinical Nutrition, 2024. **43**(3): p. 593-602.
35. Meng, C., et al., *Structural basis for psilocybin biosynthesis.* Nature Communications, 2025. **16**(1): p. 2827.
36. Giuri, D., P. Ravarino, and C. Tomasini, *L-Dopa in small peptides: an amazing functionality to form supramolecular materials.* Organic & Biomolecular Chemistry, 2021. **19**(21): p. 4622-4636.
37. Taghizadeh, A., et al., *Mussel-inspired biomaterials: from chemistry to clinic.* Bioengineering & Translational Medicine, 2022. **7**(3): p. e10385.
38. Zhang, W., et al., *Catechol-functionalized hydrogels: biomimetic design, adhesion mechanism, and biomedical applications.* Chemical Society Reviews, 2020. **49**(2): p. 433-464.
39. Hedges, J.B. and K.S. Ryan, *Biosynthetic pathways to nonproteinogenic α-amino acids.* Chemical Reviews, 2019. **120**(6): p. 3161-3209.
40. Eisenstein, M., *Living factories of the future.* Nature, 2016. **531**(7594): p. 401-403.
41. Qumsani, A.T., *The contribution of microorganisms to sustainable development: towards a green future through synthetic biology and systems biology.* Journal of Umm Al-Qura University for Applied Sciences, 2024: p. 1-17.
42. Xu, M., et al., *Optimization of l-arginine purification from Corynebacterium crenatum fermentation broth.* Journal of Separation Science, 2020. **43**(14): p. 2936-2948.
43. Lu, Y. and S. Freeland, *On the evolution of the standard amino-acid alphabet.* Genome biology, 2006. **7**(1): p. 102.
44. Walsh, C.T., R.V. O'Brien, and C. Khosla, *Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds.* Angewandte Chemie International Edition, 2013. **52**(28): p. 7098-7124.
45. Revilla-López, G., et al., *NCAD, a database integrating the intrinsic conformational preferences of non-coded amino acids.* The Journal of Physical Chemistry B, 2010. **114**(21): p. 7413-7422.

46. Milchevskiy, Y.V., G.I. Kravatskaya, and Y.V. Kravatsky, *AAindexNC: estimating the physicochemical properties of non-canonical amino acids, including those derived from the PDB and PDBeChem Databank.* International Journal of Molecular Sciences, 2024. **25**(23): p. 12555.
47. Gfeller, D., O. Michielin, and V. Zoete, *SwissSidechain: a molecular and structural database of non-natural sidechains.* Nucleic acids research, 2012. **41**(D1): p. D327-D332.
48. Wu, X.-Y., et al., *Recent advances of L-ornithine biosynthesis in metabolically engineered Corynebacterium glutamicum.* Frontiers in Bioengineering and Biotechnology, 2020. **7**: p. 440.
49. Schramm, T., et al., *High-throughput enrichment of temperature-sensitive argininosuccinate synthetase for two-stage citrulline production in E. coli.* Metabolic Engineering, 2020. **60**: p. 14-24.
50. Jin, X., et al., *Metabolic engineering strategies for L-Homoserine production in Escherichia coli.* Microbial Cell Factories, 2024. **23**(1): p. 338.
51. Uy, R. and F. Wold, *Posttranslational Covalent Modification of Proteins: Only 20 amino acids are used in. protein synthesis, yet some 140" amino acids" are found in various proteins.* Science, 1977. **198**(4320): p. 890-896.
52. Khoury, G.A., R.C. Baliban, and C.A. Floudas, *Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database.* Scientific reports, 2011. **1**(1): p. 90.
53. Sowers, K.R., et al., *N epsilon-acetyl-beta-lysine: an osmolyte synthesized by methanogenic archaebacteria.* Proceedings of the National Academy of Sciences, 1990. **87**(23): p. 9083-9087.
54. Staszek, P., et al., *L-Canavanine: How does a simple non-protein amino acid inhibit cellular function in a diverse living system?* Phytochemistry Reviews, 2017. **16**(6): p. 1269-1282.
55. Zinkand, W.C., et al., *Ibotenic acid mediates neurotoxicity and phosphoinositide hydrolysis by independent receptor mechanisms.* Molecular and chemical neuropathology, 1992. **16**(1): p. 1-10.
56. Obermaier, S. and M. Müller, *Ibotenic acid biosynthesis in the fly agaric is initiated by glutamate hydroxylation.* Angewandte Chemie, 2020. **132**(30): p. 12532-12535.
57. Finking, R. and M.A. Marahiel, *Biosynthesis of nonribosomal peptides.* Annu. Rev. Microbiol., 2004. **58**(1): p. 453-488.
58. Süssmuth, R.D. and A. Mainz, *Nonribosomal peptide synthesis—principles and prospects.* Angewandte Chemie International Edition, 2017. **56**(14): p. 3770-3821.
59. Gare, C.L., A.M. White, and L.R. Malins, *From lead to market: chemical approaches to transform peptides into therapeutics.* Trends in Biochemical Sciences, 2025.
60. Verdin, P., *Top companies and drugs by sales in 2024.* Nature reviews. Drug discovery, 2025.
61. Hall, S., D. Isaacs, and J.N. Clements, *Pharmacokinetics and clinical implications of semaglutide: a new glucagon-like peptide (GLP)-1 receptor agonist.* Clinical pharmacokinetics, 2018. **57**(12): p. 1529-1538.
62. Muttenthaler, M., et al., *Trends in peptide drug discovery.* Nature reviews Drug discovery, 2021. **20**(4): p. 309-325.
63. Xiao, W., et al., *Advance in peptide-based drug development: delivery platforms, therapeutics and vaccines.* Signal Transduction and Targeted Therapy, 2025. **10**(1): p. 74.
64. Wang, L., et al., *Therapeutic peptides: current applications and future directions.* Signal transduction and targeted therapy, 2022. **7**(1): p. 48.
65. Li, Q., W. Chao, and L. Qiu, *Therapeutic peptides: chemical strategies fortify peptides for enhanced disease treatment efficacy.* Amino Acids, 2025. **57**(1): p. 25.
66. Hickey, J.L., et al., *Beyond 20 in the 21st century: prospects and challenges of non-canonical amino acids in peptide drug discovery.* ACS Medicinal Chemistry Letters, 2023. **14**(5): p. 557-565.
67. Baumann, T., et al., *Prospects of in vivo incorporation of non-canonical amino acids for the chemical diversification of antimicrobial peptides.* Frontiers in microbiology, 2017. **8**: p. 124.

68. Yao, K., et al., *Enhancing the selectivity and conditional sensitivity of an antimicrobial peptide through cleavage simulations and homoarginine incorporation to combat drug-resistant bacteria.* Scientific Reports, 2025. **15**(1): p. 21798.
69. Soulika, A.M., et al., *Compstatin inhibits complement activation by binding to the β-chain of complement factor 3.* Molecular immunology, 2006. **43**(12): p. 2023-2029.
70. Sahu, A., B.K. Kay, and J.D. Lambris, *Inhibition of human complement by a C3-binding peptide isolated from a phage-displayed random peptide library.* The Journal of Immunology, 1996. **157**(2): p. 884-891.
71. Katragadda, M., et al., *Hydrophobic effect and hydrogen bonds account for the improved activity of a complement inhibitor, compstatin.* Journal of medicinal chemistry, 2006. **49**(15): p. 4616-4622.
72. Pauling, L., R.B. Corey, and H.R. Branson, *The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain.* Proceedings of the National Academy of Sciences, 1951. **37**(4): p. 205-211.
73. Eisenberg, D., *The discovery of the α-helix and β-sheet, the principal structural features of proteins.* Proceedings of the National Academy of Sciences, 2003. **100**(20): p. 11207-11210.
74. Mathur, P., S. Ramakumar, and V. Chauhan, *Peptide design using α, β-dehydro amino acids: From β-turns to helical hairpins.* Peptide Science: Original Research on Biomolecules, 2004. **76**(2): p. 150-161.
75. Pathak, S. and V.S. Chauhan, *Rationale-based, de novo design of dehydrophenylalanine-containing antibiotic peptides and systematic modification in sequence for enhanced potency.* Antimicrobial agents and chemotherapy, 2011. **55**(5): p. 2178-2188.
76. Zhao, B., et al., *A Thioether-Stabilized d-Proline–l-Proline-Induced β-Hairpin Peptide of Defensin Segment Increases Its Anti-Candida albicans Ability.* ChemBioChem, 2016. **17**(15): p. 1416-1420.
77. Bennett, N.R., et al., *Improving de novo protein binder design with deep learning.* Nature Communications, 2023. **14**(1): p. 2625.
78. Guha, R., *On exploring structure–activity relationships.* In silico models for drug discovery, 2013: p. 81-94.
79. Baek, M., et al., *Accurate prediction of protein structures and interactions using a three-track neural network.* Science, 2021. **373**(6557): p. 871-876.
80. Vázquez Torres, S., et al., *De novo design of high-affinity binders of bioactive helical peptides.* Nature, 2024. **626**(7998): p. 435-442.
81. Watson, J.L., et al., *De novo design of protein structure and function with RFdiffusion.* Nature, 2023. **620**(7976): p. 1089-1100.
82. Dauparas, J., et al., *Robust deep learning–based protein sequence design using ProteinMPNN.* Science, 2022. **378**(6615): p. 49-56.
83. Zhang, H. and S. Chen, *Cyclic peptide drugs approved in the last two decades (2001–2021).* RSC Chemical Biology, 2022. **3**(1): p. 18-31.
84. Rettie, S.A., et al., *Accurate de novo design of high-affinity protein-binding macrocycles using deep learning.* Nature Chemical Biology, 2025: p. 1-9.
85. Quartararo, A.J., et al., *Ultra-large chemical libraries for the discovery of high-affinity peptide binders.* Nature communications, 2020. **11**(1): p. 3183.
86. Driggers, E.M., et al., *The exploration of macrocycles for drug discovery—an underexploited structural class.* Nature Reviews Drug Discovery, 2008. **7**(7): p. 608-624.
87. Smith, G.P., *Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface.* Science, 1985. **228**(4705): p. 1315-1317.
88. Nemoto, N., et al., *In vitro virus: bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro.* FEBS letters, 1997. **414**(2): p. 405-408.

89. Song, B.P.C., A.C.W. Ch'ng, and T.S. Lim, *Review of phage display: A jack-of-all-trades and master of most biomolecule display.* International journal of biological macromolecules, 2024. **256**: p. 128455.
90. Barendt, P.A., et al., *Streamlined protocol for mRNA display.* ACS combinatorial science, 2013. **15**(2): p. 77-81.
91. Jaroszewicz, W., et al., *Phage display and other peptide display technologies.* FEMS microbiology reviews, 2022. **46**(2): p. fuab052.
92. Burke, S.A., S.L. Lo, and J.A. Krzycki, *Clustered genes encoding the methyltransferases of methanogenesis from monomethylamine.* Journal of bacteriology, 1998. **180**(13): p. 3432-3440.
93. Hao, B., et al., *A new UAG-encoded residue in the structure of a methanogen methyltransferase.* Science, 2002. **296**(5572): p. 1462-1466.
94. Noren, C.J., et al., *A general method for site-specific incorporation of unnatural amino acids into proteins.* Science, 1989. **244**(4901): p. 182-188.
95. Mursinna, R.S. and S.A. Martinis, *Rational design to block amino acid editing of a tRNA synthetase.* Journal of the American Chemical Society, 2002. **124**(25): p. 7286-7287.
96. Liu, D.R. and P.G. Schultz, *Progress toward the evolution of an organism with an expanded genetic code.* Proceedings of the National Academy of Sciences, 1999. **96**(9): p. 4780-4785.
97. Koch, N.G. and N. Budisa, *Evolution of pyrrolysyl-tRNA synthetase: From methanogenesis to genetic code expansion.* Chemical Reviews, 2024. **124**(16): p. 9580-9608.
98. Icking, L.-S., et al., *iNClusive: a database collecting useful information on non-canonical amino acids and their incorporation into proteins for easier genetic code expansion implementation.* Nucleic Acids Research, 2024. **52**(D1): p. D476-D482.
99. Kimoto, M. and I. Hirao, *Genetic alphabet expansion technology by creating unnatural base pairs.* Chemical Society Reviews, 2020. **49**(21): p. 7602-7626.
100. Dunkelmann, D.L., et al., *Engineered triply orthogonal pyrrolysyl–tRNA synthetase/tRNA pairs enable the genetic encoding of three distinct non-canonical amino acids.* Nature chemistry, 2020. **12**(6): p. 535-544.
101. Bryson, D.I., et al., *Continuous directed evolution of aminoacyl-tRNA synthetases.* Nature chemical biology, 2017. **13**(12): p. 1253-1260.
102. Wang, K., et al., *Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion.* Nature biotechnology, 2007. **25**(7): p. 770-777.
103. Hampton, J.T. and W.R. Liu, *Diversification of phage-displayed peptide libraries with noncanonical amino acid mutagenesis and chemical modification.* Chemical reviews, 2024. **124**(9): p. 6051-6077.
104. Bionda, N., A.L. Cryan, and R. Fasan, *Bioinspired strategy for the ribosomal synthesis of thioether-bridged macrocyclic peptides in bacteria.* ACS chemical biology, 2014. **9**(9): p. 2008-2013.
105. Owens, A.E., et al., *MOrPH-PhD: an integrated phage display platform for the discovery of functional genetically encoded peptide macrocycles.* ACS central science, 2020. **6**(3): p. 368-381.
106. Oller-Salvia, B. and J.W. Chin, *Efficient phage display with multiple distinct non-canonical amino acids using orthogonal ribosome-mediated genetic code expansion.* Angewandte Chemie, 2019. **131**(32): p. 10960-10964.
107. Goto, Y. and H. Suga, *The RaPID platform for the discovery of pseudo-natural macrocyclic peptides.* Accounts of Chemical Research, 2021. **54**(18): p. 3604-3617.
108. Goto, Y., T. Katoh, and H. Suga, *Flexizymes for genetic code reprogramming.* Nature protocols, 2011. **6**(6): p. 779-790.
109. Guillen Schlippe, Y.V., et al., *In vitro selection of highly modified cyclic peptides that act as tight binding inhibitors.* Journal of the American Chemical Society, 2012. **134**(25): p. 10469-10477.

110. Nawatha, M., et al., *De novo macrocyclic peptides that specifically modulate Lys48-linked ubiquitin chains.* Nature Chemistry, 2019. **11**(7): p. 644-652.
111. Iskandar, S.E., et al., *Enabling genetic code expansion and peptide macrocyclization in mRNA display via a promiscuous orthogonal aminoacyl-tRNA synthetase.* Journal of the American Chemical Society, 2023. **145**(3): p. 1512-1517.
112. Won, Y., et al., *Recent advances in enzyme engineering through incorporation of unnatural amino acids.* Biotechnology and Bioprocess Engineering, 2019. **24**(4): p. 592-604.
113. Liu, T., et al., *Enhancing protein stability with extended disulfide bonds.* Proceedings of the National Academy of Sciences, 2016. **113**(21): p. 5910-5915.
114. Deepankumar, K., et al., *Engineering Transaminase for Stability Enhancement and Site-Specific Immobilization through Multiple Noncanonical Amino Acids Incorporation.* ChemCatChem, 2015. **7**(3): p. 417-421.
115. Jackson, J.C., et al., *Improving nature's enzyme active site with genetically encoded unnatural amino acids.* Journal of the American Chemical Society, 2006. **128**(34): p. 11124-11127.
116. Liu, T., *Recent advances in genetic code expansion: from cell engineering to protein design*. 2022. p. 167565.
117. Gupta, M.N. and V.N. Uversky, *Biological importance of arginine: A comprehensive review of the roles in structure, disorder, and functionality of peptides and proteins.* International Journal of Biological Macromolecules, 2024. **257**: p. 128646.
118. Neves, M.A., M. Yeager, and R. Abagyan, *Unusual arginine formations in protein function and assembly: rings, strings, and stacks.* The journal of physical chemistry B, 2012. **116**(23): p. 7006-7013.
119. Mason, P., et al., *The hydration structure of guanidinium and thiocyanate ions: implications for protein stability in aqueous solution.* Proceedings of the National Academy of Sciences, 2003. **100**(8): p. 4557-4561.
120. Williams, M.L. and J.E. Gready, *Guanidinium-Type resonance stabilization and its biological implications. I. the guanidine and extended-guanidine series.* Journal of Computational Chemistry, 1989. **10**(1): p. 35-54.
121. Marshall, M.S., et al., *Potential energy curves for cation− π interactions: off-axis configurations are also attractive.* The Journal of Physical Chemistry A, 2009. **113**(48): p. 13628-13632.
122. Bayer, T.S., et al., *Arginine-rich motifs present multiple interfaces for specific binding by RNA.* Rna, 2005. **11**(12): p. 1848-1857.
123. Barik, A., et al., *Molecular architecture of protein-RNA recognition sites.* Journal of Biomolecular Structure and Dynamics, 2015. **33**(12): p. 2738-2751.
124. Stanzl, E.G., et al., *Fifteen years of cell-penetrating, guanidinium-rich molecular transporters: basic science, research tools, and clinical applications.* Accounts of chemical research, 2013. **46**(12): p. 2944-2954.
125. Allolio, C., et al., *Arginine-rich cell-penetrating peptides induce membrane multilamellarity and subsequently enter via formation of a fusion pore.* Proceedings of the National Academy of Sciences, 2018. **115**(47): p. 11923-11928.
126. Oba, M., et al., *Secondary structures and cell-penetrating abilities of arginine-rich peptide foldamers.* Scientific Reports, 2019. **9**(1): p. 1349.
127. Atkinson, D.J., et al., *Enduracididine, a rare amino acid component of peptide antibiotics: Natural products and synthesis.* Beilstein journal of organic chemistry, 2016. **12**(1): p. 2325-2342.
128. Yin, X. and T.M. Zabriskie, *The enduracidin biosynthetic gene cluster from Streptomyces fungicidicus.* Microbiology, 2006. **152**(10): p. 2969-2983.
129. Kawakami, M., et al., *Anti-microbial activities of enduracidin (enramycin) in vitro and in vivo.* The journal of antibiotics, 1971. **24**(9): p. 583-586.

130. de Kruijff, B., V. van Dam, and E. Breukink, *Lipid II: a central component in bacterial cell wall synthesis and a target for antibiotics.* Prostaglandins, Leukotrienes and Essential Fatty Acids, 2008. **79**(3-5): p. 117-121.
131. Fang, X., et al., *The mechanism of action of ramoplanin and enduracidin.* Molecular BioSystems, 2006. **2**(1): p. 69-76.
132. Wang, B., et al., *Total Synthesis of Mannopeptimycins α and β.* Journal of the American Chemical Society, 2016. **138**(11): p. 3926-3932.
133. Ruzin, A., et al., *Mechanism of action of the mannopeptimycins, a novel class of glycopeptide antibiotics active against vancomycin-resistant gram-positive bacteria.* Antimicrobial agents and chemotherapy, 2004. **48**(3): p. 728-738.
134. Haltli, B., et al., *Investigating β-hydroxyenduracididine formation in the biosynthesis of the mannopeptimycins.* Chemistry & biology, 2005. **12**(11): p. 1163-1168.
135. Burroughs, A.M., et al., *Structural and functional characterization of MppR, an enduracididine biosynthetic enzyme from streptomyces hygroscopicus: functional diversity in the acetoacetate decarboxylase-like superfamily.* Biochemistry, 2013. **52**(26): p. 4492-4506.
136. Ling, L.L., et al., *A new antibiotic kills pathogens without detectable resistance.* Nature, 2015. **517**(7535): p. 455-459.
137. Shukla, R., et al., *Mode of action of teixobactins in cellular membranes.* Nature communications, 2020. **11**(1): p. 2848.
138. Shukla, R., et al., *Teixobactin kills bacteria by a two-pronged attack on the cell envelope.* Nature, 2022. **608**(7922): p. 390-396.
139. Hamada, M., et al., *Minosaminomycin, a new antibiotic containing myo-inosamine.* The Journal of Antibiotics, 1974. **27**(1): p. 81-83.
140. Olesker, A., et al., *Conformation of the structure of myo-inosamine components of minosaminomycin synthesis of derivatives of D-myo-inosamine-1.* The Journal of Antibiotics, 1975. **28**(6): p. 490-491.
141. Suzukake, K., et al., *Biochemical study of minosaminomycin in relation to the kasugamycin group antibiotics.* The Journal of Antibiotics, 1977. **30**(2): p. 132-140.
142. Tryon, J.H., et al., *Genome mining and metabolomics uncover a rare d-capreomycidine containing natural product and its biosynthetic gene cluster.* ACS chemical biology, 2020. **15**(11): p. 3013-3020.
143. Chen, H., et al., *The mycobacterial membrane: a novel target space for anti-tubercular drugs.* Frontiers in microbiology, 2018. **9**: p. 1627.
144. Palande, A., et al., *Outer membrane proteins inMycobacterium tuberculosisand its potential role in small molecule permeation.* 2023.
145. Laughlin, Z.T. and G.L. Conn, *Tuberactinomycin antibiotics: Biosynthesis, anti-mycobacterial action, and mechanisms of resistance.* Frontiers in Microbiology, 2022. **13**: p. 961921.
146. Wiegmann, D., et al., *Muraymycin nucleoside-peptide antibiotics: Uridine-derived natural products as lead structures for the development of novel antibacterial agents.* Beilstein journal of organic chemistry, 2016. **12**(1): p. 769-795.
147. Löckener, I., et al., *The MraY Inhibitor Muraymycin D2 and Its Derivatives Induce Enlarged Cells in Obligate Intracellular Chlamydia and Wolbachia and Break the Persistence Phenotype in Chlamydia.* Antibiotics, 2024. **13**(5): p. 421.
148. Tatsuta, K., et al., *The structure of chymostatin, a chymotrypsin inhibitor.* The Journal of Antibiotics, 1973. **26**(11): p. 625-646.
149. Delbaere, L.T. and G.D. Brayer, *The 1.8 Å structure of the complex between chymostatin and Streptomyces griseus protease A: A model for serine protease catalytic tetrahedral intermediates.* Journal of molecular biology, 1985. **183**(1): p. 89-103.
150. Marincean, S., M. Al-Modhafir, and D.B. Lawson, *π-π stacking interactions in tryptophan-lumiflavin-tyrosine: a structural model for riboflavin insertion into riboflavin-binding protein.* Journal of Molecular Modeling, 2025. **31**(2): p. 38.

151. Roose, B.W. and D.W. Christianson, *Structural basis of tryptophan reverse N-prenylation catalyzed by CymD.* Biochemistry, 2019. **58**(30): p. 3232-3242.
152. Alkhalaf, L.M. and K.S. Ryan, *Biosynthetic manipulation of tryptophan in bacteria: pathways and mechanisms.* Chemistry & biology, 2015. **22**(3): p. 317-328.
153. Liu, Z., et al., *Functional characterization of the halogenase SpmH and discovery of new deschloro-tryptophan dimers.* Organic & Biomolecular Chemistry, 2019. **17**(5): p. 1053-1057.
154. Luhavaya, H., et al., *Biosynthesis of l-4-chlorokynurenine, an antidepressant prodrug and a non-proteinogenic amino acid found in lipopeptide antibiotics.* Angewandte Chemie International Edition, 2019. **58**(25): p. 8394-8399.
155. Frese, M., et al., *Regioselective enzymatic halogenation of substituted tryptophan derivatives using the FAD-dependent halogenase RebH.* ChemCatChem, 2014. **6**(5): p. 1270-1276.
156. Zuo, R. and Y. Ding, *Direct aromatic nitration system for synthesis of nitrotryptophans in Escherichia coli.* ACS Synthetic Biology, 2019. **8**(4): p. 857-865.
157. Sánchez, C., et al., *The biosynthetic gene cluster for the antitumor rebeccamycin: characterization and generation of indolocarbazole derivatives.* Chemistry & biology, 2002. **9**(4): p. 519-531.
158. Davis, I. and A. Liu, *What is the tryptophan kynurenine pathway and why is it important to neurotherapeutics?* 2015, Taylor & Francis. p. 719-721.
159. Fujikawa, M., M. Ueda, and K. Maruyama, *Role of kynurenine and its derivatives in the neuroimmune system.* International Journal of Molecular Sciences, 2024. **25**(13): p. 7144.
160. Sherin, P., et al., *Photoactivity of kynurenine-derived UV filters.* Journal of Photochemistry and Photobiology B: Biology, 2008. **93**(3): p. 127-132.
161. Persha, H., et al., *Real-World Clinical Characteristics and Outcomes with Daptomycin Use in Pediatric Patients: A Retrospective Case Series.* Antibiotics, 2024. **13**(9): p. 833.
162. Grein, F., et al., *Ca2+-Daptomycin targets cell wall biosynthesis by forming a tripartite complex with undecaprenyl-coupled intermediates and membrane lipids.* Nature communications, 2020. **11**(1): p. 1455.
163. Taylor, S.D. and M. Palmer, *The action mechanism of daptomycin.* Bioorganic & medicinal chemistry, 2016. **24**(24): p. 6253-6268.
164. Stampfli, A.R., W. Blankenfeldt, and F.P. Seebeck, *Structural basis of ergothioneine biosynthesis.* Current opinion in structural biology, 2020. **65**: p. 1-8.
165. Sato, S., et al., *Biosynthesis of ergothioneine: current state, achievements, and perspectives.* Applied Microbiology and Biotechnology, 2025. **109**(1): p. 1-14.
166. Kang, X., et al., *Advances on the biosynthesis of ergothioneine using microbial chassis.* Biotechnology Advances, 2025: p. 108701.
167. Kang, X., et al., *Enhancing Ergothioneine Production by Combined Protein and Metabolic Engineering Strategies.* Journal of Agricultural and Food Chemistry, 2025. **73**(15): p. 9234-9245.
168. Maggiora, L.L., C.W. Smith, and A. Hsi, *L-2-thiol-histidine: Introduction of conformational constraints into peptides via thioether linkage.* Tetrahedron letters, 1990. **31**(20): p. 2837-2840.
169. Jenny, K.A., et al., *Ergothioneine in a peptide: Substitution of histidine with 2-thiohistidine in bioactive peptides.* Journal of Peptide Science, 2021. **27**(10): p. e3339.
170. Pauli, S., et al., *Systems metabolic engineering upgrades Corynebacterium glutamicum for selective high-level production of the chiral drug precursor and cell-protective extremolyte L-pipecolic acid.* Metabolic Engineering, 2023. **77**: p. 100-117.
171. Grant, S.G., et al., *Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants.* Proceedings of the National Academy of Sciences, 1990. **87**(12): p. 4645-4649.

172. Myronovskyi, M., et al., *Generation of a cluster-free Streptomyces albus chassis strains for improved heterologous expression of secondary metabolite clusters.* Metabolic engineering, 2018. **49**: p. 316-324.
173. Marques, F., A. Luzhetskyy, and M.V. Mendes, *Engineering Corynebacterium glutamicum with a comprehensive genomic library and phage-based vectors.* Metabolic Engineering, 2020. **62**: p. 221-234.
174. K Nakayama, H.Y., *Process for producing l-arginine by fermentation*. 1973: United States.
175. Nakayama, K. and H. Hagino, *Process for producing l-tryptophan*. 1974, Google Patents.
176. Nakayama, K. and K. Araki, *Process for producing L-lysine*. 1973, Google Patents.
177. Gregory, M.A., R. Till, and M.C. Smith, *Integration site for Streptomyces phage φBT1 and development of site-specific integrating vectors.* Journal of bacteriology, 2003. **185**(17): p. 5320-5323.
178. Becker, J., et al., *From zero to hero—design-based systems metabolic engineering of Corynebacterium glutamicum for l-lysine production.* Metabolic engineering, 2011. **13**(2): p. 159-168.
179. Buschke, N., H. Schröder, and C. Wittmann, *Metabolic engineering of Corynebacterium glutamicum for production of 1, 5-diaminopentane from hemicellulose.* Biotechnology Journal, 2011. **6**(3): p. 306-317.
180. Green, M.R. and J. Sambrook, *Molecular cloning.* A Laboratory Manual 4th, 2012. **448**.
181. Hobbs, G., et al., *Dispersed growth of Streptomyces in liquid culture.* Applied microbiology and biotechnology, 1989. **31**: p. 272-277.
182. Bilyk, O., et al., *Cloning and heterologous expression of the grecocycline biosynthetic gene cluster.* PloS one, 2016. **11**(7): p. e0158682.
183. Rohles, C.M., et al., *Systems metabolic engineering of Corynebacterium glutamicum for the production of the carbon-5 platform chemicals 5-aminovalerate and glutarate.* Microbial cell factories, 2016. **15**: p. 1-13.
184. Horbal, L., et al., *Cyclofaulknamycin with the rare amino acid D-capreomycidine isolated from a well-characterized Streptomyces albus strain.* Microorganisms, 2021. **9**(8): p. 1609.
185. Moore, S. and W.H. Stein, *Chromatography of amino acids on sulfonated polystyrene resins.* 1951.
186. Wang, J., et al., *Total synthesis of mannopeptimycin β via β-hydroxyenduracididine ligation.* Journal of the American Chemical Society, 2021. **143**(32): p. 12784-12790.
187. Hatano, K., et al., *Biosynthesis of enduracidin: origin of enduracididine and other amino acids.* Agricultural and biological chemistry, 1984. **48**(6): p. 1503-1508.
188. Han, L., et al., *Streptomyces wadayamensis MppP is a PLP-dependent oxidase, not an oxygenase.* Biochemistry, 2018. **57**(23): p. 3252-3264.
189. Vuksanovic, N., et al., *Structural and Biochemical Characterization of MppQ, an L-Enduracididine Biosynthetic Enzyme from Streptomyces hygroscopicus.* Biochemistry, 2023. **62**(21): p. 3105-3115.
190. Siegl, T., et al., *Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes.* Metabolic engineering, 2013. **19**: p. 98-106.
191. Harada, K.-i., et al., *Application of D, L-FDLA derivatization to determination of absolute configuration of constituent amino acids in peptide by advanced Marfey's method.* Tetrahedron letters, 1996. **37**(17): p. 3001-3004.
192. Udaka, S., *The discovery of Corynebacterium glutamicum and birth of amino acid fermentation industry in Japan.* Corynebacteria: genomics and molecular biology, 2008: p. 1-6.
193. Tatsumi, N. and M. Inui, *Corynebacterium glutamicum: biology and biotechnology*. Vol. 23. 2012: Springer Science & Business Media.
194. Falcioni, F., B. Bühler, and A. Schmid, *Efficient hydroxyproline production from glucose in minimal media by Corynebacterium glutamicum.* Biotechnology and Bioengineering, 2015. **112**(2): p. 322-330.

195. Kurpejović, E., V.F. Wendisch, and B. Sariyar Akbulut, *Tyrosinase-based production of L-DOPA by Corynebacterium glutamicum.* Applied Microbiology and Biotechnology, 2021. **105**(24): p. 9103-9111.
196. Zhang, S., et al., *Model-based reconstruction of synthetic promoter library in Corynebacterium glutamicum.* Biotechnology letters, 2018. **40**: p. 819-827.
197. Becker, J., et al., *Amplified expression of fructose 1, 6-bisphosphatase in Corynebacterium glutamicum increases in vivo flux through the pentose phosphate pathway and lysine production on different carbon sources.* Applied and environmental microbiology, 2005. **71**(12): p. 8587-8596.
198. Wei, H., et al., *Identification and application of a novel strong constitutive promoter in Corynebacterium glutamicum.* Annals of Microbiology, 2018. **68**: p. 375-382.
199. Kacar, B., et al., *Functional constraints on replacing an essential gene with its ancient and modern homologs.* MBio, 2017. **8**(4): p. 10.1128/mbio. 01276-17.
200. Claassens, N.J., et al., *Bicistronic design-based continuous and high-level membrane protein production in Escherichia coli.* ACS synthetic biology, 2019. **8**(7): p. 1685-1690.
201. Mutalik, V.K., et al., *Precise and reliable gene expression via standard transcription and translation initiation elements.* Nature methods, 2013. **10**(4): p. 354-360.
202. Uy, D., et al., *Instability of glutamate production by Corynebacterium glutamicum 2262 in continuous culture using the temperature-triggered process.* Journal of biotechnology, 2003. **104**(1-3): p. 173-184.
203. Stansen, C., et al., *Characterization of a Corynebacterium glutamicum lactate utilization operon induced during temperature-triggered glutamate production.* Applied and environmental microbiology, 2005. **71**(10): p. 5920-5928.
204. Dominguez, H., et al., *Modified carbon flux during oxygen limited growth of Corynebacterium glutamicum and the consequences for amino acid overproduction.* Biotechnology letters, 1993. **15**: p. 449-454.
205. Eggeling, L. and M. Bott, *Handbook of Corynebacterium glutamicum*. 2005: CRC press.
206. Thomas, M.G., Y.A. Chan, and S.G. Ozanick, *Deciphering tuberactinomycin biosynthesis: isolation, sequencing, and annotation of the viomycin biosynthetic gene cluster.* Antimicrobial agents and chemotherapy, 2003. **47**(9): p. 2823-2830.
207. Stanley, R.E., et al., *The structures of the anti-tuberculosis antibiotics viomycin and capreomycin bound to the 70S ribosome.* Nature structural & molecular biology, 2010. **17**(3): p. 289-293.
208. DeMong, D.E. and R.M. Williams, *Asymmetric synthesis of (2 S, 3 R)-capreomycidine and the total synthesis of capreomycin IB.* Journal of the American Chemical Society, 2003. **125**(28): p. 8561-8565.
209. Carter, J.H., et al., *Biosynthesis of viomycin. II. Origin of β-lysine and viomycidine.* Biochemistry, 1974. **13**(6): p. 1227-1233.
210. Ju, J., et al., *Conversion of (2S)-arginine to (2S, 3R)-capreomycidine by VioC and VioD from the viomycin biosynthetic pathway of Streptomyces sp. strain ATCC11861.* ChemBioChem, 2004. **5**(9): p. 1281-1285.
211. Chang, C.Y., et al., *Biosynthesis of streptolidine involved two unexpected intermediates produced by a dihydroxylase and a cyclase through unusual mechanisms.* Angewandte Chemie, 2014. **126**(7): p. 1974-1979.
212. Lakey, J.H. and M. Ptak, *Fluorescence indicates a calcium-dependent interaction between the lipopeptide antibiotic LY 146032 and phospholipid membranes.* Biochemistry, 1988. **27**(13): p. 4639-4645.
213. Moreira, R. and S.D. Taylor, *The chiral target of daptomycin is the 2R, 2′ S stereoisomer of phosphatidylglycerol.* Angewandte Chemie, 2022. **134**(4): p. e202114858.
214. Goodyear, J., et al., *The Calcium-Dependent Antibiotics: Structure–Activity Relationships and Determination of Their Lipid Target.* ACS Infectious Diseases, 2024. **11**(1): p. 226-237.

215. Zummo, F., et al., *Tryptophan catabolism via kynurenine production in Streptomyces coelicolor: identification of three genes coding for the enzymes of tryptophan to anthranilate pathway.* Applied microbiology and biotechnology, 2012. **94**(3): p. 719-728.
216. Zhang, H., et al., *A Review of Novel Antioxidant Ergothioneine: Biosynthesis Pathways, Production, Function and Food Applications.* Foods, 2025. **14**(9): p. 1588.
217. Tararina, M.A., et al., *Structure of the Repurposed Fungal Terpene Cyclase FlvF Implicated in the C–N Bond-Forming Reaction of Flavunoidine Biosynthesis.* Biochemistry, 2022. **61**(18): p. 2014-2024.
218. SATOSHI, H. and K. YUKIHIKO, *Enduracidin, a new antibiotic-5-Structures of new basic amino acids, enduracididine and alloenduracididine (Communications to the editors).* The Journal of Antibiotics, 1968. **21**(11): p. 665-667.
219. Craig, W., et al., *A highly stereoselective and scalable synthesis of L-allo-enduracididine.* Organic Letters, 2015. **17**(18): p. 4620-4623.
220. Shen, K. and Q. Wang, *Copper-catalyzed alkene aminoazidation as a rapid entry to 1, 2-diamines and installation of an azide reporter onto azahetereocycles.* Journal of the American Chemical Society, 2017. **139**(37): p. 13110-13116.