様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　４ 年　 ６ 月　 ５ 日現在

機関番号：１２６０１
研究種目：若手研究
研究期間：2020～2021
課題番号：２０Ｋ１５４０７

研究課題名（和文）Characterization of novel biosynthetic gene clusters for combinatorial synthesis of thiopeptides

研究課題名（英文）Characterization of novel biosynthetic gene clusters for combinatorial synthesis of thiopeptides

研究代表者
ＶＩＮＯＧＲＡＤＯＶ　Ａｌｅｘａｎｄｅｒ（Vinogradov, Alexander）

東京大学・大学院理学系研究科（理学部）・特任助教

研究者番号：９０８４５８１９

交付決定額（研究期間全体）：（直接経費）　　3,200,000円

研究成果の概要（和文）：研究は、2つの方向性に焦点をあてて行われました。
1.新規生理活性を有する擬似天然物ラクタゾール様チオペプチドの探索のための mRNA-displayのプラットフォームの開発。我々は、提案したプラットフォームを確立し、設計された生物活性を持つ擬似天然物を発見することにより、その有用性を示した。この開発には、ラクタゾール生合成の再構築、ライブラリ構築プロトコルの開発、そして合成されたセレクションヒットの生化学的特性評価が含まれる。
2.ラクタゾール様生合成酵素の再構築。laz BGCに相同な 4 つの生合成遺伝子クラスターから 12種類の酵素の生産に成功した。発現した酵素はすべて活性を持っていた。

研究成果の学術的意義や社会的意義
Discovery of new pharmaceuticals is a major challenge in the modern biomedical research. This research addresses the need for new methods to discover compounds with good biological activities and favourable pharmacological profiles using bacterial natural products as an inspiration.

研究成果の概要（英文）：The work focused around two major directions. 1. Development of an mRNA-display based platform for de novo discovery of pseudo-natural lactazole-like thiopeptides with novel bioactivities. We have established the proposed platform and showed its utility by discovering dozens of pseudo-natural products with designed biological activities. The development entailed reengineering of in vitro lactazole biosynthesis, development of library construction protocols, establishment of a general synthetic strategy for modular access to the discovered structures, and finally, biochemical characterization of the synthesized selection hits. 2. In vitro reconstitution of lactazole-like biosynthetic enzymes. We have successfully produced 12 enzymes from 4 biosynthetic gene clusters (BGCs) homologous to laz BGC. All expressed enzymes were active, and catalyzed reactions analogous to Laz enzymes, which simplified the analysis, but crucially, had a divergent substrate specificity profiles.

研究分野：Biochemistry and chemical biology

キーワード：Biosynthesis　Enzymes　mRNA display　Bioengineering

１．研究開始当初の背景

Lactazole A is a natural product belonging to the thiopeptide group of ribosomally synthesized and post-translationally modified peptides (RiPPs). Thiopeptides are natural products with a lot of therapeutic potential, as they usually possess strong antibiotic activity against Gram-positive bacteria, including methicillin resistant *Staphyrococcus aureus* strains (MRSA). Lactazole A is biosynthesized from a minimal 9.8 kb biosynthetic gene cluster (**BGC**), which encodes five enzymes solely responsible for its biosynthesis (Chem. Biol. 2014, 21, 679–688). The biosynthesis is initiated with ribosomal production of LazA precursor peptide encoded inside *laz* BGC. The biosynthetic enzymes then utilize the N-terminal 38 residues of LazA as a recognition sequence (leader peptide; LP) and introduce post-translational modifications (PTMs) such as azole and dehydroalanine (Dha) to the C-terminal core peptide sequence (CP) of LazA. Eventually, a pyridine synthase LazC catalyzes formation of a pyridine ring and eliminates the LP, yielding the macrocyclic thiopeptide.

Lactazole is unique in many aspects. It has a 32-membered macrocycle, a low Cys/Ser/Thr content and it bears an unmodified amino acid in position 2. All of these features are unusual among thiopeptides. Recent bioinformatic studies indicate that the lactazole-like thiopeptides comprise close to half of all predicted thiopeptides (251 out of 508 annotated BGCs) and can be further subdivided into multiple subfamilies, and yet the prototypical laz BGC remains the only characterized member of this family to date. Overall, lactazole-like thiopeptides remain a rather enigmatic family of natural products, as close to nothing is known about their function, structural diversity, and biosynthesis.

Previously, we established a platform for in vitro biosynthesis of lactazole A, referred to as the FIT Laz system, via a combination of the flexible in vitro translation (FIT) system with recombinantly produced lactazole biosynthetic enzymes. The FIT-Laz system enabled a rapid and reliable access to lactazole A and related pseudo-natural thiopeptides, opening possibilities for thiopeptide bioengineering.

２．研究の目的

The work pursued three major goals: 1) establishment of an mRNA display selection platform utilizing pseudo-natural thiopeptides scaffolds based on the FIT-Laz system; 2) elucidation of substrate preferences for individual thiopeptide enzymes; 3) reconstitution and characterization of novel biosynthetic enzymes which are homologous to Laz enzymes.

３．研究の方法

Each of the three identified goals was pursued independently using the appropriate techniques as elaborated below.

４．研究成果

1: **Establishment of an mRNA display selection platform utilizing pseudo-natural thiopeptides scaffolds**. [*J. Am. Chem. Soc*. **2021**, 143, 13358−13369; two other manuscripts in preparation].

Structural and functional reprogramming of ribosomally synthesized and post-translationally modified peptides (RiPPs) is a promising approach to utilizing natural product-like scaffolds in drug discovery. Mutagenesis of RiPP precursor peptides generates novel compounds if the underlying biosynthetic machinery can accommodate sequence altered substrates. For RiPPs assembled by promiscuous enzymes, combinatorial libraries can be generated and screened against a target of interest to discover new compounds with designer biological activities. Assessing whether a particular biosynthetic pathway is suitable for such purposes is usually a labor-intensive process because numerous mutants for each participating enzyme need to be analyzed. Contemporary mutagenesis approaches address this issue to some extent by enabling simultaneous profiling of dozens of compounds, but both in vivo and in vitro strategies mostly

provide qualitative output with no temporal resolution. For instance, analysis of enzyme preferences from in vivo data may be confounded by metabolic stability of the mutants as well as by export, self-immunity, and associated issues. We are particularly interested in functional reprogramming of lactazole, a thiopeptide from Streptomyces lactacystinaeus, assembled by five dedicated enzymes from a ribosomally produced thiopeptide precursor peptide (TPP). Our previous work established the FIT−Laz system, a combination of the flexible in vitro translation (FIT) system utilized to prepare TPPs with recombinantly produced Laz enzymes, as a platform for in vitro synthesis of lactazole and analogues. We demonstrated that every Laz enzyme can process extensively mutated TPPs and developed $LazA^{min}$, a TPP in which only seven residues (Ser1, Trp2, Ser10, Ser11, Ser12, Cys13, and Ala14) are required for complete maturation to a macrocyclic thiopeptide. Yet, more specific substrate preferences of the biosynthetic pathway, a prerequisite for successful library construction, remained unclear. In this work, we describe a statistical approach to investigate substrate−enzyme fitness landscapes for the lactazole biosynthesis pathway. We established the reactivity-profiling mRNA display assay to analyze a saturation mutagenesis library based on $LazA^{min}$ and developed a statistical framework to interpret display outcomes. We found that this workflow leads to an accurate, quantitative assessment of substrate fitness for 304 mutants in a single experiment. The platform is generalizable and should facilitate the assessment of RiPP biosynthetic pathways for bioengineering. For lactazole enzymes, our results comprehensively map substrate preferences and enable further engineering of $LazA^{min}$. Critically, this work constructed the first thiopeptide-mRNA display platform, which served as the foundation for the following work.
Next, we built upon this platform and showed its utility by discovering dozens of pseudo-natural products with designed biological activities. The development entailed reengineering of in vitro lactazole biosynthesis, development of library construction protocols, establishment of a general synthetic strategy for rapid and modular access to the discovered structures, and finally, thorough biochemical characterization of the synthesized selection hits. We constructed an mRNA display-based screening platform that enables the construction and screening of large ($>10^{13}$ unique compounds) combinatorial libraries of lactazole-like thiopeptides. We reengineered lactazole biosynthesis to improve the maturation efficiency and homogeneity of the resulting products and develop library construction protocols. We then designed, assembled, and screened a large combinatorial library of thiopeptides against Traf2- and NCK-interacting kinase (TNIK), a protein implicated in several forms of cancer, to identify a series of compounds acting as potent and selective kinase inhibitors (the best compound, TP15, had K_i of 3 nM). Two X-ray crystal structures of TNIK/thiopeptide complexes reveal how multiple post-translational modifications (PTMs) of the discovered thiopeptides promote the peptide folding and contribute to its interaction with the target protein. TP15 inhibited TNIK in HCT116 cell line assays, indicating that like many natural thiopeptides, de novo discovered compounds can access intracellular targets. Overall, we believe that Laz-RaPID opens new possibilities to utilize pseudo-natural product structures in early drug discovery.

2: **Elucidation of substrate preferences for individual thiopeptide enzymes** [ACS Central Science ASAP; https://doi.org/10.1021/acscentsci.2c00223; *J. Am. Chem. Soc*. **2021**, 143, 33, 13358–13369].

PTM enzymes often display non-obvious substrate preferences by acting on diverse yet well-defined sets of peptides and/or proteins. Understanding of substrate fitness landscapes for PTM enzymes is important in many areas of contemporary science, including natural product biosynthesis, molecular biology and biotechnology. Here, we report an integrated platform for accurate profiling of substrate preferences for PTM enzymes. The platform features i) a combination of mRNA display with next generation sequencing as an ultrahigh throughput technique for data acquisition and ii) deep learning for data analysis. The high accuracy (>0.99 in each of two studies) of the resulting deep learning models enables comprehensive analysis of enzymatic substrate preferences. The models can quantify fitness across sequence space, map modification sites, and identify important amino acids in the substrate. To benchmark the platform, we perform profiling of a Ser dehydratase (LazBF) and a Cys/Ser cyclodehydratase (LazDEF), two enzymes from the lactazole biosynthesis pathway. In both studies, our results point to complex enzymatic preferences, which, particularly for LazBF, cannot be reduced to a set of simple rules. The ability of the constructed models to dissect such complexity suggests that the developed platform can facilitate the wider study of PTM enzymes. This platform also enables rapid, streamlined analysis of substrate preferences for Laz enzyme homologs (see **3**).
In the second part of these investigation, we focused on the substrate preferences of glutamate elimination domains (LazF and homologs) in greater detail. Formation of dehydroalanine and dehydrobutyrine residues via tRNA-dependent dehydration of serine and threonine is a key post-

translational modification in the biosynthesis of lanthipeptide and thiopeptide RiPPs. The dehydration process involves two reactions, wherein the O–glutamyl Ser/Thr intermediate, accessed by a dedicated enzyme utilizing Glu-tRNA^{Glu} as the acyl donor, is recognized by the second enzyme, referred to as the glutamate elimination domain (ED), which catalyzes the eponymous reaction yielding a dehydroamino acid. Many details of ED catalysis remain unexplored because the scope of available substrates for testing is limited to those that the upstream enzymes can furnish. Here, we report two complementary strategies for direct, nonenzymatic access to diverse ED substrates. We establish that a thiol-thioester exchange reaction between a Cys-containing peptide and an $^{\alpha}$thioester of glutamic acid leads an S–glutamylated intermediate which can act as a substrate for EDs. Furthermore, we show that the native O–glutamylated substrates can be accessible from S–glutamylated peptides upon a site-specific S-to-O acyl transfer reaction. Combined with flexible in vitro translation utilized for rapid peptide production, these chemistries enabled us to dissect the substrate recognition requirements of three known EDs. Our results establish that EDs are uniquely promiscuous enzymes capable of acting on substrates with arbitrary amino acid sequences, and performing retro-Michael reaction beyond the canonical glutamate elimination. To facilitate substrate recruitment, EDs apparently engage in nonspecific hydrophobic interactions with their substrates. Altogether, our results establish the substrate scope of EDs and provide clues to their catalysis.

3. **Reconstitution and characterization of novel biosynthetic enzymes which are homologous to Laz enzymes**.

So far, we have successfully produced 12 enzymes from 4 BGCs homologous to the prototypical laz BGC. To do this, we established an *E. coli* expression system that greatly facilitated heterologous protein expression from these BGCs. All expressed enzymes were active, and catalyzed reactions analogous to Laz enzymes, which simplified the analysis, but crucially, had divergent substrate specificity profiles. The ongoing work is directed toward establishing hybrid biosynthetic pathways by combining the most active enzymes from their respective BGCs with Laz enzymes. To this end, we are pursuing three lines of inquiry: i) investigation of substrate preferences for the individual enzymes using the techniques developed in **2**; ii) engineering of hybrid LPs which activate near-cognate enzymes from homologous BGCs; iii) the combination of new enzymes with the mRNA display platform described in **1**.

５．主な発表論文等

〔雑誌論文〕　計3件（うち査読付論文　3件／うち国際共著　0件／うちオープンアクセス　2件）

１．著者名 Alexander A. Vinogradov, Masanobu Nagano, Yuki Goto, and Hiroaki Suga	４．巻 143
２．論文標題 Site-Specific Nonenzymatic Peptide S/O-Glutamylation Reveals the Extent of Substrate Promiscuity in Glutamate Elimination Domains	５．発行年 2021年
３．雑誌名 Journal of the American Chemical Society	６．最初と最後の頁 13358～13369
掲載論文のDOI（デジタルオブジェクト識別子） 10.1021/jacs.1c06470	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

１．著者名 Vinogradov Alexander, Nagai Emiko, Chang Jun Shi, Narumi Kakeru, Onaka Hiroyasu, Goto Yuki, Suga Hiroaki	４．巻 142
２．論文標題 Accurate Broadcasting of Substrate Fitness for Lactazole Biosynthetic Pathway from Reactivity-Profiling mRNA Display	５．発行年 2020年
３．雑誌名 Journal of the American Chemical Society	６．最初と最後の頁 20329～20334
掲載論文のDOI（デジタルオブジェクト識別子） 10.1021/jacs.0c10374	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 Alexander A. Vinogradov, Jun Shi Chang, Hiroyasu Onaka, Yuki Goto, and Hiroaki Suga	４．巻 -
２．論文標題 Accurate Models of Substrate Preferences of Post-Translational Modification Enzymes from a Combination of mRNA Display and Deep Learning	５．発行年 2022年
３．雑誌名 ACS Central Science	６．最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1021/acscentsci.2c00223	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

〔学会発表〕　計0件

〔図書〕　計0件

〔産業財産権〕

〔その他〕

－

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関