Activity, Binding, and Modeling Studies of a Reprogrammed Aryl Acid Adenylation Domain with an Enlarged Substrate Binding Pocket

Fumihiro Ishikawa,*a Hinano Kitayama,a Shinya Nakamura,b Katsuki Takashima,a Isao Nakanishi,b and Genzoh Tanabestar

*Pharmaceutical Organic Chemistry Lab, Faculty of Pharmacy, Kindai University; 3–4–1 Kowakae, Higashi-Osaka, Osaka 577–8502, Japan; and b Computational Drug Design and Discovery Lab, Faculty of Pharmacy, Kindai University; 3–4–1 Kowakae, Higashi-Osaka, Osaka 577–8502, Japan.

Received September 3, 2020; accepted November 26, 2020

The gatekeeping adenylation (A) domain of the non-ribosomal peptide synthetase (NRPS) selectively incorporates specific proteinogenic/non-proteinogenic amino acid into a growing peptide chain. The EntE of the enterobactin NRPS is a discrete aryl acid A-domain with 2,3-dihydroxybenzoic acid (DHB) substrate specificity. Reprogrammed EntE N235G variant possesses an enlarged substrate recognition site, and is capable of accepting non-native aryl acids. Biochemical characterization of this unique substrate recognition site should provide a better understanding of activi-site microenvironments. Here, we synthesized a non-hydrolysable adenylate analogue with 2-aminobenzoic acid (2-ABA), 3-aminobenzoic acid (3-ABA), and 4-aminobenzoic acid (4-ABA) and used them to calculate the apparent inhibition constants (Ki)app. Dose–response experiments using 3-ABA-sulfamoyladenosine (AMS) provided Kiapp values of $596 \text{nM}$ for wild-type EntE and $2.4 \text{nM}$ for the N235G variants. These results suggest that 3-amino group of benzoic acid plays an important role in substrate recognition by the N235G variant. These findings would help designing aryl acid substrates with substituents at the 2- and 3-positions.

Key words non-ribosomal peptide synthetase; adenylation domain; reprogramming; non-hydrolyzable acyl adenylate; aryl acid

Introduction

Modular non-ribosomal peptide synthetases (NRPSs) biosynthesize non-ribosomal peptides (NRPs). NRPs have a wide range of biologically important properties, including antimicrobial and antitumor activities.1) Each NRPS module is responsible for the incorporation of an amino acid building block into the growing peptides.2) Activities of peptidyl carrier protein (PCP) and adenylation (A) domains play a key role in the attachment of a wide variety of amino acid building blocks into the NRPS assembly lines.3) The A-domain functions as a gatekeeper, by selecting building blocks from a pool of 20 proteinogenic amino acids, non-proteinogenic amino acids, aryl acids, and α-hydroxyl acids.4) First, A-domain catalyzes the adenylate formation of an amino acid substrate at the expense of ATP and then transfers an aminoacyl adenylate to 4'-phosphopantetheine (Ppant) arm of the adjacent PCP domain, forming an aminoacyl-S-PCP.

Aryl acid building blocks contribute to the structural diversity and biological activities of NRPs. For example, 2,3-dihydroxybenzoic acid (DHB), 2-aminobenzoic acid (2-ABA), 3-aminobenzoic acid (3-ABA), and 4-aminobenzoic acid are observed in enterobactin,5) acetylaszonalenin,6) pactamycin,7) and albicidin,8) respectively.

EntB, EntE, and EntF are responsible for the biosynthesis of the iron chelator enterobactin in Escherichia coli9) (Fig. 1A). EntE is a discrete aryl acid A-domain that activates DHB and transfers it to aryl CP (ArCP) of EntB (Fig. 1A). We demonstrated that a single Asn-to-Gly (N-to-G) mutation in a DHB-activating A-domain makes EntE to generate an enzyme that accepts a variety of non-native aryl acids.9) The N235G variant was capable of adenylylating non-native aryl acids with nitro-, cyano-, chloro-, bromo-, and iodo-groups at 2- or 3-positions of benzoic acid (BA)9) (Fig. 1B). Crystal structures of the N235G variant complexed with sulfamoyloxyl-linked acyl-AMP analogues 3-cyanoBA-sulfamoyladenosine (AMS) and 2-nitroBA-AMS revealed the origins of an enlarged substrate recognition site for non-native aryl acids functionalized with diverse substituents.9) Recently, we facilitated the precise understanding of NRPS A-domain codes of aryl acids by gradually grafting the NRPS codes of salicylic acid-activating A-domains into the substrate binding pocket of EntE.10) More recently, we confirmed that the N235G variant displays enzymatic activity toward 2-ABA by a proof-of-concept study of an enzyme-linked immunosorbent assay system (ELISA) for the A-domains in NRPSs.11) The binding profiles only provide information on substrate candidates of wild-type EntE and the N235G variant, but unable to predict adenylation activity. However, substrate recognition is an important event for enzyme catalysis.

In this study, we determined the substrate profiles of wild-type EntE and the N235G variant toward naturally occurring aminobenzoic acid substrates (2-ABA, 3-ABA, and 4-ABA), synthesized a sulfamoyloxyl-linked acyl-AMP analogue with 2-ABA, 3-ABA, and 4-ABA (2-ABA-AMS, 3-ABA-AMS, and 4-ABA-AMS), and conducted binding and modeling studies of 2-ABA-AMS, 3-ABA-AMS, and 4-ABA-AMS toward wild-type EntE and the N235G variant (Fig. 1C). The study outcomes would help a better understanding of not only the unique substrate binding site of the N235G variant but also NRPS codes toward aryl acid building blocks, facilitating engineering of the aryl acid A-domain.
Results and Discussion

We measured enzymatic activity of wild-type EntE and the N235G variant toward 2-ABA, 3-ABA, and 4-ABA using a malachite green colorimetric assay (Fig. 2). These activity studies indicated that wild-type EntE could afford negligible level in enzymatic activity toward 3-ABA and 4-ABA. The substrate profile of the N235G variant demonstrated that it is capable of accepting 3-ABA and 4-ABA, and can catalyze their adenylation relative to wild-type EntE, although with relatively low \( K_{\text{app}} \) release.

We next synthesized 2-ABA-AMS, 3-ABA-AMS, and 4-ABA-AMS as described in the Supplementary materials (Charts S1, S2, and S3) to investigate the binding profiles of wild-type EntE and the N235G variant. We examined the inhibition of wild-type EntE and N235G variant by 2-ABA-AMS, 3-ABA-AMS, and 4-ABA-AMS. The two enzymes were expressed and purified as C-terminal His\(_6\)-tagged fusion proteins according to previously described procedure. The apparent inhibition constant \( K_{\text{app}} \) values were calculated at fixed concentrations of wild-type EntE (0.5 \( \mu \)M) or the N235G variant (1 \( \mu \)M), DHB substrate (100 \( \mu \)M), and ATP (200 \( \mu \)M). The resulting dose-response plot of 2-ABA-AMS was fit to the Morrison equation. A \( K_{\text{app}} \) value of 43 \( \pm \) 10 nM for N235G variant (Fig. S1 and Table 1) was observed, which is 11-fold lower than that of wild-type EntE \( (K_{\text{app}} = 457 \pm 51 \text{nM}) \) (Fig. S1 and Table 1). The competitors 3-ABA-AMS and 4-ABA-AMS gave \( K_{\text{app}} \) values of 2.4 \( \pm \) 1.4 nM and 2.0 \( \pm \) 0.2 \( \mu \)M, respectively, toward N235G variant (Fig. 3A and S1 and Table 1). On the basis of the N235G variant structure in complex with 2-nitroBA-AMS, we demonstrated that 2-nitro group

Fig. 1. Aryl Acid Adenylation (A) Domain EntE in (A) Non-ribosomal Peptide (NRP) Biosynthesis of the Siderophore Enterobactin; (B) Structural Analysis of the N235G Variant

The complex structures and active-site volumes (pink mesh) of the N235G variant were modeled using MOE. The ligands 3-cyanoBA-AMS and 2-nitroBA-AMS are shown as ball-and-stick models, with the following color code: nitrogen, blue; oxygen, red; sulfur, yellow [This figure was modified from ref. 9]. (C) Aryl acid substrates and the corresponding inhibitors described in this study. (Color figure can be accessed in the online version.)

Fig. 2. The Adenylation Activity of the Wild-Type EntE (A) and the N235G Variant (B) toward 2,3-Dihydroxybenzoic Acid (DHB), 2-Aminobenzoic Acid (2-ABA), 3-Aminobenzoic Acid (3-ABA), and 4-Aminobenzoic Acid (4-ABA)

The adenylation activity was measured in a malachite green phosphate assay. The activity was normalized to that of wild-type EntE toward DHB.

The activity was normalized to that of wild-type EntE toward DHB.
of the 2-nitroBA forms effective hydrogen-bonds with the backbone amide groups of Tyr236 and Ala335 in the N235G variant, and it orients toward the enlarged space created by the N235G substitution. The hydrogen-bond interactions enabled the remarkably selective binding of 2-nitroBA-AMS toward N235G variant (\(K_{i}^{app} = 50 \pm 3 \text{nM}\)) (Table 1). Significantly, the \(K_{i}^{app}\) value of 3-ABA-AMS toward N235G variant was approximately 21-fold lower than that of 2-nitroBA-AMS toward the N235G variant (Table 1). These results suggest that 3-ABA-AMS with \(K_{i}^{app}\) may be bound via highly selective protein interactions formed between the 3-amino group of 3-ABA and the enlarged substrate recognition site. To elucidate the recognition properties of 3-ABA-AMS by wild-type EntE and the N235G variant, we constructed models of wild-type EntE and N235G variant in complexed with 3-ABA-AMS. The modeled complexes of wild-type EntE and N235G variant with 3-ABA-AMS are shown in Figs. 3C–E. In wild-type EntE, the hydroxyl group of Ser240 enabled to engage hydrogen-bonding interaction with the amino group of 3-ABA-AMS, however, the amino group of Asn235 affected on the binding of the 3-ABA-AMS, probably because of steric hindrance between side chain of Asn235 and the 3-amino group of benzoic acid moiety (Fig. 3C). In contrast, N235G variant enlarges the substrate binding site relative to that of wild-type EntE, providing a cavity for 3-amino group of benzoic acid moiety, effectively forming hydrogen-bonding interaction with the hydroxyl group of Ser240 (Fig. 3D).

In summary, we synthesized 2-ABA-AMS, 3-ABA-AMS, and 4-ABA-AMS and tested its inhibitory properties against the wild-type EntE and the N235G variant. Inhibition studies provided a \(K_{i}^{app}\) value 2.4 \(\pm\) 1.4 nM for the N235G variant, demonstrating highly selective and tight binding properties of 3-ABA-AMS. We envision that these findings could help designing aryl acid substrates with substituents at the 2- and 3-positions, expanding the possibility of an adenylating catalyst with an enlarged substrate binding pocket.

### Experimental

The experimental section is provided as a supplementary material through an Internet.

---

**Table 1.** Apparent \(K_{i}^{app}\) Values of Sulfamoyloxy-Linked Acyl-AMP Analogues toward Wild-Type EntE and the N235G Variant

| Enzyme | Ligand-AMS | \(K_{i}^{app}\) (nM) |
|--------|------------|------------------|
| EntE   | DHB        | 180 \(\pm\) 45 |
|        | 2-NitroBA   | 28 \(\pm\) 2  |
|        | 2-ABA      | 457 \(\pm\) 94 |
|        | 3-ABA      | 85 \(\pm\) 28 |
| N235G  | DHB        | 3.4 \(\pm\) 1.6 |
|        | 2-NitroBA   | 50 \(\pm\) 3   |
|        | 2-ABA      | 43 \(\pm\) 10  |
|        | 3-ABA      | 2.4 \(\pm\) 1.4|
|        | 4-ABA      | 2.0 \(\pm\) 0.2|

*a* Each value represents the mean \(\pm\) standard error of the mean (S.E.M.) of duplicate independent measurements. *b* Values taken from ref 9.
Acknowledgments This work was supported by Grants-in-Aid for Scientific Research on Innovative Areas (17H05438 and 19H04664 to F.I.) from the Japan Society of the Promotion of Science (JSPS) and by a grant from the Takeda Science Foundation (F.I.). This work was partly supported by a Grant-in-Aid for Research (C) (19K05722 to F.I. and 17K08377 to G.T.) from JSPS and by grants from the Noda Institute for Scientific Research (F.I.) and the Hoansha Foundation (G.T.). We are also thankful for the financial support extended by the Antiaging Project for Private Universities.

Conflict of Interest The authors declare no conflict of interest.

Supplementary Materials The online version of this article contains supplementary materials.

References
1) Süssmuth R. D., Mainz A., Angew. Chem. Int. Ed., 56, 3770–3821 (2017).
2) Hur G. H., Vickery C. R., Burkart M. D., Nat. Prod. Rep., 29, 1074–1098 (2012).
3) Kudo F., Miyanaga A., Eguchi T., J. Ind. Microbiol. Biotechnol., 46, 515–536 (2019).
4) Gehring A. M., Mori I., Walsh C. T., Biochemistry, 37, 2648–2659 (1998).
5) Ames B. D., Walsh C. T., Biochemistry, 49, 3351–3365 (2010).
6) Kudo F., Zhang J., Sato S., Eguchi T., ChemBioChem, 20, 2458–2462 (2019).
7) Cociancich S., Pesic A., Petras D., Uhlmann S., Kretz J., Schubert V., Vieweg L., Duplan S., Marguerettaz M., Noell J., Pieretti I., Haggendall M., Kemper S., Mainz A., Rott P., Royer M., Süssmuth R. D., Nat. Chem. Biol., 11, 195–197 (2015).
8) Walsh C. T., Liu J., Rusnak F., Sakaitani M., Chem. Rev., 90, 1105–1129 (1990).
9) Ishikawa F., Miyanaga A., Kitayama H., Nakamura S., Nakanishi I., Kudo F., Eguchi T., Tanabe G., Angew. Chem. Int. Ed., 58, 6906–6910 (2019).
10) Ishikawa F., Nohara M., Nakamura S., Nakanishi I., Tanabe G., Biochemistry, 59, 351–363 (2020).
11) Ishikawa F., Nohara M., Takashima K., Tanabe G., ChemBioChem, 21, 3056–3061 (2020).
12) Molecular Operating Environment, version 2019.01 (2019) Chemical Computing Group, Montreal.
13) McQuade T. J., Shallop A. D., Sheoran A., DelProposto J. E., Toodikov O. V., Garneau-Tsodikova S., Anal. Biochem., 386, 244–250 (2009).
14) Ishikawa F., Miyamoto K., Konno S., Kasai S., Kakeya H., ACS Chem. Biol., 10, 2816–2826 (2015).