The Crystal Structure of the Adenylation Enzyme VinN Reveals a Unique β-Amino Acid Recognition Mechanism*

Received for publication, August 1, 2014, and in revised form, September 17, 2014. Published, JBC Papers in Press, September 22, 2014, DOI 10.1074/jbc.M114.602326

Akimasa Miyanaga1, Jolanta Cieślak8, Yuji Shinohara5, Fumitaka Kudo9, and Tadashi Eguchi5
†*From the Departments of 1Chemistry and 5Chemistry and Materials Science, Tokyo Institute of Technology, O-okayama, Meguro-ku, Tokyo 152-8551, Japan

Background: The β-amino acid adenylation reaction is important for biosynthesis of natural products.

Results: We present the crystal structure and a mutational study of the adenylation enzyme VinN involved in vicenistatin biosynthesis.

Conclusion: VinN has a characteristic substrate-binding pocket that selectively accommodates β-amino acids.

Significance: This study could provide clues for β-amino acid specificity prediction and protein engineering of adenylation enzymes.

Adenylation enzymes play important roles in the biosynthesis and degradation of primary and secondary metabolites. Mechanistic insights into the recognition of α-amino acid substrates have been obtained for α-amino acid adenylation enzymes. The Asp residue is invariant and is essential for the stabilization of the α-amino group of the substrate. In contrast, the β-amino acid recognition mechanism of adenylation enzymes is still unclear despite the importance of β-amino acid activation for the biosynthesis of various natural products. Herein, we report the crystal structure of the stand-alone adenylation enzyme VinN, which specifically activates (2S,3S)-3-methylaspartate (3-MeAsp) in vicenistatin biosynthesis. VinN has an overall structure similar to that of other adenylation enzymes. The structure of the complex with 3-MeAsp revealed that a conserved Asp230 residue is used in the recognition of the β-amino group of 3-MeAsp similar to α-amino acid adenylation enzymes. A mutational analysis and structural comparison with α-amino acid adenylation enzymes showed that the substrate-binding pocket of VinN has a unique architecture to accommodate 3-MeAsp as a β-amino acid substrate. Thus, the VinN structure allows the first visualization of the interaction of an adenylation enzyme with a β-amino acid and provides new mechanistic insights into the selective recognition of β-amino acids in this family of enzymes.

Adenylation enzymes, which include the nonribosomal peptide synthetase (NRPS)2 adenylation domains, the acyl-CoA synthetases, and the luciferase enzymes, play important roles in the biosynthesis and degradation of primary and secondary metabolites (1). These enzymes first catalyze the adenylation of a carboxylate substrate using ATP to form an acyl-AMP intermediate. In most cases, the adenylation enzymes subsequently catalyze the formation of a thioester bond with the 4′-phosphopantetheine group of a carrier protein or CoA molecule. The carboxylate substrates of adenylation enzymes are structurally diverse and include acetate, amino acids, fatty acids, and aromatic acids. In general, each adenylation enzyme recognizes a specific carboxylate substrate.

Modular NRPS enzymes are involved in the biosynthesis of various secondary metabolites. NRPS adenylation domains, which mostly have α-amino acids as substrates, serve as gatekeepers of the NRPS assembly line via selective substrate recognition and activation. After the determination of the first crystal structure of an NRPS adenylation domain, PheA (see Fig. 1A) involved in gramicidin S biosynthesis (2), general rules for the assignment of substrate specificity for adenylation enzymes have been developed. These rules are based on classification of a specificity-conferring code derived from 10 amino acid residues that comprise the substrate-binding pocket (3, 4). The ~100-amino acid stretch of the N-terminal domain (corresponding to the positions Asp235–Cys331 in PheA) contains nine of the 10 residues in the substrate-binding pocket. For α-amino acid substrate recognition, the Asp residue at the 235 position is invariant and is essential for stabilization of the α-amino group of the α-amino acid. The remaining residue is the highly conserved C-terminal Lys residue, which is involved in two key interactions with both the carboxylate group of the α-amino acid and the ribos moiety of adenylate. These two residues (Asp and Lys) are important to fix the position and orientation of the α-amino acid for the reaction. The other eight specificity-conferring code residues are involved in the recognition of the side chain of the α-amino acid substrate. This specificity-conferring code rule can be applied to various NRPS-type adenylation enzymes because information on the relationship between the sequence and substrate specificity of these enzymes has been accumulated (3–7). This rule allows us to predict the substrate specificity of many biochemically
uncharacterized adenylation enzymes from their amino acid sequences and enables rational alteration of their substrate specificities (3, 8–12). However, this specificity-conferring code is not applicable for adenylation enzymes that activate unusual substrates such as β-amino acids.

Various natural products containing a β-amino acid unit have been isolated (13). In many cases, adenylation enzymes that activate β-amino acids are involved in their biosynthetic pathways (13–24). Similar to α-amino acid-activating enzymes, the interaction with the β-amino group and the carboxylate group should be important for β-amino acid-activating enzymes to fix the substrate orientation. However, the recognition model for α-amino acid-activating enzymes cannot be applied to that for β-amino acid-activating enzymes because the amino group position of the β-amino acid is different from that of the α-amino acid. Therefore, it can be assumed that β-amino acid adenylation enzymes have a structurally different active site for recognition of the β-amino acid. To date, no crystal structures of β-amino acid-activating enzymes are available, so the β-amino acid recognition mechanism of adenylation enzymes is poorly understood. What determines the substrate specificity between α-amino acid and β-amino acid adenylation enzymes also remains elusive.

Macrolactam natural products are an important class of macrocyclic polyketides, and many of them contain various β-amino acids at the starter position for polyketide chain elongation. Vicenistatin is produced by *Streptomyces halstedii* HC34 and belongs to the family of macrolactam antibiotics. This enzyme possesses a unique β-amino acid unit, 3-amino-2-methylpropionate, at the starter position of the polyketide backbone (25). Recently, we elucidated that the stand-alone adenylation enzyme VinN is involved in vicenistatin biosynthesis (14) (see Fig. 1A). VinN recognizes (2S,3S)-3-methylaspartate (3-MeAsp) as a β-amino acid substrate and catalyzes adenylation of the carboxyl group at the C4 position (see Fig. 1A). Then VinN transfers the β-amino acid group onto the stand-alone acyl carrier protein VinL to give 3-MeAsp-VinL. After decarboxylation of the 3-MeAsp moiety by the pyridoxal 5'-phosphate-dependent enzyme VinO, the resulting 3-amino-2-methylpropionate unit is aminocacylated with l-alanine to give dipeptidyl-VinL by another stand-alone adenylation enzyme, VinM. The dipeptide moiety is then selectively transferred onto the loading acyl carrier protein domain of polyketide synthase for polyketide chain elongation. Thus, VinN appears to play a crucial role in the selection of the β-amino acid starter unit for polyketide chain elongation in vicenistatin biosynthesis. Previous biochemical studies have shown that VinN exhibits a strong preference for 3-MeAsp over other amino acids (14). VinN shows no activity against α-amino acids except for weak activity against l-aspartate. As VinN exhibits relatively low sequence identity with well studied adenylation enzymes such as the NRPS adenylation domain, the substrate specificity of VinN cannot be predicted from the specificity-conferring code. The crystal structure of SlgN1, which catalyzes the adenylation of 3-MeAsp in the biosynthesis of streptolydigin, has been reported recently (26). SlgN1 recognizes (2S,3S)-3-MeAsp as an α-amino acid substrate and adenylates the carboxyl group at the C1 position (Fig. 1A). The orientation of 3-MeAsp at the active site in SlgN1 is proposed to be opposite to that in VinN. In this study, we carried out kinetic, mutational, and structural studies on VinN to clarify how VinN selectively recognizes 3-MeAsp. The VinN crystal structure allows the first visualization of the interaction between an adenylation enzyme and a β-amino acid and provides new mechanistic insights into the selective recognition of β-amino acids in this family of enzymes.

**EXPERIMENTAL PROCEDURES**

**Preparation of Recombinant Wild-type and Mutant Enzymes—** The N-terminal domain region of the *vinN* gene (*vinNN*) was amplified from pCold I-*vinN* (14) with the primers 5'-acgctggccaggagaggg-3' and 5'-tttaagcttacggccgcaagtag-3' and then inserted between the Ndel and HindIII sites of the expression vector pCold I (Takara Biochemicals, Otsu, Japan), which is designed to attach a hexahistidine-tagged sequence to the N terminus of the target protein.

*Escherichia coli* BL21(DE3) cells (Takara Biochemicals) harboring a pCold I-*vinN* or pCold I-*vinNN* plasmid were grown at 37°C in Luria-Bertani broth containing ampicillin (50 μg/ml). After the optical density at 600 nm reached 0.6, protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside (0.2 mM), and the cells were then cultured for an additional 16 h at 15°C. The recombinant protein, which was collected from cell-free extracts prepared by sonication, was purified on a TALON affinity column (Clontech). The protein was then desalted and concentrated using a PD-10 column (GE Healthcare) and an Amicon Ultra centrifugal filter (Merck Millipore, Billerica, MA), respectively. For crystallization, the recombinant protein was further purified by ResourceQ (GE Healthcare) anion-exchange chromatography with a linear gradient from 0.15 to 0.45 M NaCl in 10 mM HEPES-Na buffer (pH 7.7) containing 10% (v/v) glycerol.

For construction of the F231A, F231L, Y234A, S299A, M323G, K330A, K330N, and R331A mutants, the pCold I-*vinN* plasmid was used for site-directed mutagenesis. Site-directed mutagenesis was performed with a QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA) using the following oligonucleotides: F231A, 5'-ggtctcctggacctcggtctctacaagg-3'; F231L, 5'-ggtctcgccaaggtgctgatctc-3'; M323G, 5'-ccagttcagcggcggctg-3'; K330A, 5'-ccagttcagcggcggctg-3'; K330N, 5'-ccagttcagcggcggctg-3'; K330A, 5'-ccagttcagcggcggctg-3'; and their complementary oligonucleotides. The mutations were confirmed by determining the nucleotide sequences. These plasmids were transformed into *E. coli* BL21(DE3) cells, and the mutated enzymes were prepared as described above. The CD experiment using a J-820 spectropolarimeter (Jasco, Tokyo, Japan) was performed to confirm that inactive mutant proteins such as M323G, K330A, and R331A adopt a fully folded state.

**Enzyme Assays and Determination of Kinetic Parameters—** A continuous spectroscopic assay that measures the release of inorganic pyrophosphate with the adenylation reaction was carried out according to the method of Webb (27). The assay solution (500-μl total volume) contained 50 mM Tris-HCl
buffer (pH 7.5), 10% glycerol, 1 mM MgCl₂, 1 mM ATP, 0.1 mM 2-amino-6-mercapto-7-methylpurine ribonucleoside, 1 unit/ml inorganic pyrophosphatase, 1 unit/ml purine nucleotide phosphorylase, and 1 mM amino acid (DL-threo-3-MeAsp (Sigma-Aldrich) or L-aspartate (Kanto Chemical, Tokyo, Japan)). For kinetic assays, the amino acid concentration was varied between 0.04 and 100 mM. The reaction was initiated by addition of VinN (1–10 μM) to the mixture and incubated at 28 °C. The increase in absorbance at 360 nm, attributable to the formation of 2-amino-6-mercapto-7-methylpurine, per second was monitored using a UV-2450 spectrophotometer (Shimadzu, Tokyo, Japan). The initial velocity was determined from the linear portion of the optical density profile (ε_{360nm} = 11,000 M⁻¹ cm⁻¹). Steady-state parameters were determined by the Michaelis-Menten equation.

Crystallization, Data Collection, and Structure Determination—Crystals of VinNN were grown from a 1:1 (v/v) mixture of a VinNN protein solution (10 mg/ml in 10 mM HEPES-Na (pH 7.5) and 10% glycerol) and a reservoir solution (0.1 M Tris-HCl (pH 8.5), 0.2 M sodium acetate, and 30% polyethylene glycol 4000) using the sitting drop vapor diffusion method at 5 °C. The 3-MeAsp complex and L-aspartate complex were prepared by soaking apocrystals in 100 mM DL-threo-3-MeAsp for 4 h and 100 mM L-aspartate for 1 h, respectively. Prior to collection of the x-ray data, the crystals were flash frozen in a stream of liquid nitrogen. The x-ray diffraction data were collected on beamlines BL-5A and AR-NW-12A at the Photon Factory (Tsukuba, Japan) and subsequently indexed, integrated, and scaled using the iMosflm program (28). The initial phases were determined by the molecular replacement method using the Molrep pro-

FIGURE 1. Reaction of adenylation enzymes. A, reactions of α-amino acid adenylation enzymes PheA and SlgN1 and β-amino acid adenylation enzyme VinN. B, biosynthetic pathway of vicenistatin, including the VinN reaction. The 3-amino-2-methylpropionate unit is shown in red. ACP, acyl carrier protein; PKS, polyketide synthase; KS, ketosynthase; AT, acyltransferase; DH, dehydratase; KR, ketoreductase; TE, thioesterase.
TABLE 1

| Enzymes | Substrate | $K_m$ | $k_{cat}$ | $k_{cat}/K_m$ |
|---------|-----------|-------|-----------|---------------|
| Wild type | 3-MeAsp | 0.13 ± 0.02 | 0.39 ± 0.01 | 3000 |
| F231L | 3-MeAsp | 0.96 ± 0.24 | 0.095 ± 0.006 | 99 |
| S299A | 3-MeAsp | 0.77 ± 0.08 | 0.14 ± 0.03 | 180 |
| Wild type | l-Aspartate | 4.5 ± 0.5 | 0.79 ± 0.02 | 180 |
| F231L | l-Aspartate | 19 ± 5 | 0.27 ± 0.02 | 14 |
| S299A | l-Aspartate | 18 ± 5 | 0.12 ± 0.01 | 6.7 |

The Crystal Structure of the Adenylation Enzyme VinN

Kinetic analysis of VinN wild type and mutants

To determine the substrate recognition mechanism of VinN, we crystallized the VinNN protein containing only the N-terminal domain (Met$^1$–Arg$^{426}$) and then attempted to crystallize the VinNN protein. Finally, we succeeded in the crystalization of VinNN and determined its crystal structure at 2.15-Å resolution (Table 2).

Overall Structure of VinN—Two VinNN molecules are present in each crystallographic asymmetric unit. The VinNN structure has a five-layered $\alpha$-$\beta$-$\alpha$-$\beta$-$\alpha$ sandwich fold as observed in other crystal structures of adenylation enzymes (2, 26, 38–41) (Fig. 2A). The structure contains a disordered surface loop regions (Met$^1$–Lys$^{23}$, Gly$^{250}$–Asp$^{254}$, and Asp$^{399}$–Gly$^{403}$). A search for structurally related proteins using the DALI program (42) revealed that the closest functionally characterized protein is fatty acyl-CoA synthetase FadD13 (40) (Protein Data Bank code 3T5C, Z score = 49.8, root mean square deviation (r.m.s.d.) of 1.7 Å for 372 Ca atoms, and sequence identity of 23%). VinNN is also structurally similar to NRPS adenylation domains, including PheA (2) (Protein Data Bank code 1AMU, Z score = 45.7, r.m.s.d. of 2.0 Å for 365 Ca atoms, and sequence identity of 20%) and SlgN1 (26) (Protein Data Bank code 4GR5, Z score = 42.9, r.m.s.d. of 2.1 Å for 362 Ca atoms, and sequence identity of 23%).

A structural comparison of VinNN with other adenylation backbone adenylation enzymes, including PheA and SlgN1 (2, 26, 35, 41, 43), showed that VinNN has a similar cleft for adenylation binding on the surface of the N-terminal domain. VinNN has several conserved residues at that cleft (Fig. 3). For example, Thr$^{327}$ and Glu$^{328}$ that are supposed to interact with the $\alpha$-phosphate moiety of adenylate, Asp$^{411}$ that is supposed to interact with ribose hydroxyl groups, and Tyr$^{354}$ that is supposed to interact with the adenine moiety are located at almost the same positions in the structure (Fig. 4A). These structural observations suggest that the manner of ATP binding of VinNN is similar to that of other adenylation enzymes.

Structure of the 3-MeAsp and l-Aspartate Complexes—To understand the substrate recognition mechanism of VinN, we determined the structure of the complex with 3-MeAsp at 2.20-Å resolution. The position of the substrate-binding pocket of VinNN is similar to those of other adenylation enzymes. The entrance of the substrate-binding pocket is exposed to solvent because of the absence of the C-terminal domain. The structure shows clear electron density for (2S,3S)-3-MeAsp in the substrate-binding pocket but not for (2R,3R)-3-MeAsp (Fig. 2B). The binding of 3-MeAsp occurred with little overall structural perturbation to the VinNN polypeptide backbone (r.m.s.d. of 0.47 Å for 385 Ca atoms of chain A). Some side-chain movements are observed in the binding pocket. The side chains of Asp$^{236}$, Phe$^{231}$, and Ser$^{299}$ exhibit 19°, 15°, and 140° rotations, respectively, to accommodate 3-MeAsp (Fig. 2C). In the substrate-free structure, the backbone and side chains in the region of Ser$^{299}$ and Ala$^{300}$ seem to be flexible judging from their high B-factor values (65.5 Å$^2$). In contrast, this region of the structure in the complex exhibits significantly lower B-factor values (33.0 Å$^2$), suggesting that the binding of 3-MeAsp stabilizes it. Furthermore, we determined the structure of the complex with l-aspartate at 2.20-Å resolution. Similarly, the electron density of l-aspartate was clearly observed in the substrate-binding pocket (Fig. 2D). The overall structure of the complex with l-aspartate is almost identical to that with 3-MeAsp (r.m.s.d. of 0.47 Å for 385 Ca atoms of chain A).
TABLE 2
Data collection and refinement statistics

| Protein Data Bank code | Data set      | Apo 3WV4 | 3-MeAsp complex 3WV5 | l-Aspartate complex 3WVN |
|------------------------|---------------|----------|----------------------|--------------------------|
| Data collection statistics | Resolution (Å) | 50.00-2.15 | 37.41-2.20 | 42.27-2.20 |
| Beamline               | PF BL-5A     | PF AR-NW12A | PF AR-NW12A |
| Wavelength (Å)         | 0.9782       | 1.0000    | 1.0000    |
| Space group            | C222,        | C222,    | C222,    |
| Unit cell parameters   | α (Å)        | 81.79     | 80.66     | 80.53     |
| β (Å)                  | 91.96        | 109.41    | 109.30    |
| γ (Å)                  | 201.46       | 200.21    | 200.07    |
| Resolution (outer shell) (Å) | 50.00-2.15 (2.21-2.15) | 37.41-2.20 (2.28-2.20) | 42.27-2.20 (2.32-2.20) |
| Unique reflections     | 44,155 (3,311) | 44,288 (4,087) | 43,099 (5,734) |
| Redundancy             | 5.7 (4.7)    | 6.7 (6.1) | 6.1 (5.5) |
| Completeness (%)       | 89.5 (83.5)  | 97.8 (93.6) | 95.5 (88.3) |
| Rmerge (%)             | 7.3 (57.3)   | 6.3 (47.2) | 7.5 (61.0) |
| Mean I/σ(I)            | 12.2 (2.3)   | 17.0 (3.4) | 15.5 (2.6) |
| Wilson B-factor (Å²)   | 31.6         | 32.5      | 32.8      |
| Refinement statistics  | Resolution (Å) | 50.00-2.15 | 37.41-2.20 | 39.13-2.20 |
| Reflections used (F > 0σF) | 44,123     | 44,244    | 43,055    |
| Rwork (%)              | 20.2         | 20.5      | 20.7      |
| Rfree (%)              | 24.5         | 24.5      | 24.3      |
| No. of non-hydrogen atoms | 5,845      | 5,690     | 5,669     |
| Protein                | 191          | 203       | 176       |
| Water                  | 0            | 20        | 18        |
| Substrate              | 39.8         | 45.6      | 44.5      |
| Average B-factors (Å²) | 43.9         | 45.0      | 45.5      |
| Average B-factors (Å²) | 0.016        | 0.016     | 0.015     |
| Average B-factors (Å²) | 1.71         | 1.75      | 1.73      |
| r.m.s.d. from ideality | Bond lengths (Å) |          |          |
|                      | Bond angles (°) |          |          |
|                      | Ramachandran plot |          |          |
|                      | Favored region (%) |    |          |
|                      | Allowed region (%) |    |          |
|                      | Outer region (%) |    | 0.0      |

0.13 Å for 388 Cα atoms of chain A). There are no significant displacements in the substrate-binding pocket between the structures of these two complexes (Fig. 2C).

**Substrate Binding in the Complex Structures**—In both complex structures, the amino acid substrate is bound to the substrate-binding pocket of VinNN in the same position and orientation. The C4 carboxyl group of 3-MeAsp and l-aspartate is exposed to solvent and does not interact with any residues of the N-terminal domain.

In the structure of the complex with 3-MeAsp, two salt bridge interactions are formed between Arg331 and the C1 carboxyl group of 3-MeAsp (2.7 and 2.8 Å), and one salt bridge interaction is observed between Lys330 and the C1 carboxyl group of 3-MeAsp (3.0 Å) (Fig. 2B). These structural observations suggest the importance of Lys330 and Arg331 in 3-MeAsp recognition. In fact, the K330N mutant showed a significantly decreased activity (0.8% of the wild type). In addition, K330A and R331A mutants showed a complete loss of activity. Both Lys330 and Arg331 side chains are stacked between Tyr234 and Met323 and are located close together (3.5 Å) despite an electrostatic repulsion (Fig. 2E). Even in the substrate-free structure, these basic residues occupy the same position (Fig. 2C).

Tyr234 and Met323 do not directly interact with the 3-MeAsp substrate but are likely to modulate the conformation of the Lys330 and Arg331 side chains by steric effects. To evaluate this steric effect of Tyr234 and Met323, we constructed Y234A and M323G mutants. The Y234A mutant showed very weak activity (3.7% of the wild type), and the M323G mutant completely lost activity. In these mutants, the smaller side chain of the mutated residue cannot have any steric interactions with the Lys330 or Arg331 side chains. This probably causes displacement of these basic residues to reduce the electrostatic repulsion between them, resulting in the observed significant reduction in activity.

Ser299 forms a hydrogen bond with the C1 carboxyl group of 3-MeAsp. The S299A mutant exhibited a 6-fold higher Km value compared with the wild type, suggesting that the side-chain hydroxyl group of Ser299 participates in the recognition of the C1 carboxyl group of 3-MeAsp. The side-chain carboxyl group of Asp230 interacts with the β-amino group of the 3-MeAsp substrate with a salt bridge interaction (2.6 Å). The main-chain carbonyl oxygen of Lys330 also interacts with the β-amino group through a hydrogen bond (2.8 Å). The methyl group of 3-MeAsp seems to be recognized by the side-chain aromatic group of Phe231 (3.8–4.3 Å) through CH–π interactions (44) and van der Waals interactions (Fig. 2, C and E). The presence of these interactions appears to give 3-MeAsp selectivity over l-aspartate. In the structure of the complex with l-aspartate, the C3 carbon of l-aspartate is relatively far from the side chain of Phe231 (5.0 Å). F231A and F231L mutations resulted in higher Km values for both 3-MeAsp and l-aspartate. The F231L mutant showed a 7-fold higher Km value for 3-MeAsp compared with the wild type, although the Leu residue should retain the van der Waals interactions with the methyl group of 3-MeAsp. This mutant also showed a 4-fold higher Km value for
L-aspartate (Table 1). Furthermore, the F231A mutant exhibited very weak activity toward 3-MeAsp (1.8% of the wild type) and no detectable activity with L-aspartate. These results suggest that the mutation of Phe231 affects substrate specificity both for 3-MeAsp and L-aspartate. The side chain of Phe231 might be important for the recognition of not only the methyl group but also other parts of 3-MeAsp.

**Structural Comparison with the Crystal Structures of Other Adenylation Enzymes**—We compared the structure of VinNN with those of other adenylation enzymes, including PheA, which recognizes L-phenylalanine as a substrate (2). The carboxyl group at the C4 position of 3-MeAsp that is adenylated in VinN occupies a position similar to that of the α-carboxyl group of L-phenylalanine in PheA (Fig. 4B), suggesting that the same interaction exists with the conserved Lys residue of the C-terminal domain (Lys310 in VinN) as observed in other adenylation enzymes (Fig. 5). The position of the amino nitrogen atom of 3-MeAsp is similar to that of L-phenylalanine in PheA. In both structures, the conserved Asp residue forms the same interaction with the amino group of the substrate. Thus, the distance between the amino and carboxylate groups of the amino acids at the active site of both enzymes is similar, although the β-amino acid has one carbon atom inserted between these termini compared with the α-amino acid. Consequently, in the VinNN structure, the C2–C3 bond of 3-MeAsp appears to be pushed into a relatively large space of the active site. The positions of the three carbon atoms (C1, C2, and C3) of 3-MeAsp individually show 1.1–1.5-Å displacement from the corresponding atoms of L-phenylalanine in the PheA structure.

Further structural comparisons revealed that two factors likely give rise to the difference in the overall architecture of the substrate-binding pocket. First, the β13β14 loop containing the two important residues Lys330 and Arg331 is one residue shorter in VinN than in most other adenylation enzymes, including PheA (Figs. 3 and 4C). This one-residue-shorter loop has also been reported in malonyl-CoA synthetase ScMatB (45). In the VinNN structure, this shorter loop causes a 0.8–2.0-Å movement of the polypeptide backbone between Lys330 and Ile332 (Fig. 4, B and D), providing space for fixing the location and
The Crystal Structure of the Adenylation Enzyme VinN

FIGURE 3. Amino acid sequence comparison of VinN with PheA. Specificity-conferring code residues are shown in red. The residues involved in AMP binding in PheA are shown in blue. The secondary structural elements of VinNN and PheA are indicated by bars above or below the sequence. The VinN β13β14 loop containing Lys330 and Arg331 is indicated by a broken line above the sequence. The C-terminal domain sequence of VinN is shown in gray. The conserved positions are marked with an asterisk.

orientation of the C1–C2 bond of 3-MeAsp. The C1 carboxyl group of 3-MeAsp might be pulled in this space by interactions with Ser299, Lys330, and Arg331, resulting in its slight movement toward the β13β14 loop. Second, VinN has bulky residues in the substrate-binding pocket. Phe231 and Ser299, which are equivalent to Ala236 and Ala301, respectively, in PheA, are located on the opposite side of the substrate-binding pocket from the β13β14 loop (Fig. 4B). The position of the C1, C2, and C3 carbon atoms of the 3-MeAsp substrate could be affected by the presence of these two larger residues. They might be pushed toward the β13β14 loop by the steric constraints of these residues. The importance of Phe231 was confirmed by the mutational studies where mutation of Phe231 significantly reduced the activity against both 3-MeAsp and L-aspartate (Table 1).

Thus, these two factors are thought to be important for the activity against both 3-MeAsp and L-aspartate (Table 1). The importance of Phe231 was confirmed by the mutational studies where mutation of Phe231 significantly reduced the activity against both 3-MeAsp and L-aspartate (Table 1).

**DISCUSSION**

Many crystal structures of adenylation enzymes such as NRPS adenylation domains have been reported (2, 26, 35, 41, 43, 46–48). However, no crystal structure of a β-amino acid-activating enzyme has been available so far. In this study, we determined the structure of the N-terminal domain of VinN, which is involved in biosynthesis of the macrolactam antibiotic vicenstatin. The structures of complexes with both 3-MeAsp and L-aspartate clearly demonstrate that VinN recognizes 3-MeAsp as a β-amino acid. The VinNN structure can provide important mechanistic insights into β-amino acid recognition as described below.

Sequence alignment and structural comparison with other adenylation enzymes revealed that the β-amino acid substrate specificity of VinN is also governed by the specificity-conferring code (Figs. 3 and 5). The VinNN structure and the mutational study clearly show that five of the 10 residues (Asp230, Phe231, Ser299, Lys330, and Arg331) are directly involved in substrate binding (Fig. 2, B and E). Two other residues (Tyr234 and Met322) derived from the specificity-conferring code seem to be indirectly involved in substrate binding by conformational modulation of two essential basic residues, Lys330 and Arg331. In addition, the conserved Lys330 of the C-terminal domain is thought to interact with the substrate carboxyl group that is adenylated as reported for other adenylation enzymes. Importantly, the Asp residue important for recognition of the α-amino group of the α-amino acid is also conserved in VinN. The VinNN structure shows that Asp301 is used for the recognition of the β-amino group of 3-MeAsp. However, the overall archi-
The Crystal Structure of the Adenylation Enzyme VinN

The structure of the substrate-binding pocket of VinN is significantly different compared with those of other adenylation enzymes. The β-amino acid specificity of VinN seems to be dictated by the surrounding regions, including the shorter β13β14 loop and the specificity-conferring code residues such as Phe231. These factors seem to largely contribute to control of the conformation of the β-amino acid substrate so that the β-amino group of the substrate is placed adjacent to the Asp230 side-chain carboxyl group.

A comparison of the specificity-conferring code of VinN with other adenylation enzymes revealed that VinN has a special code for the β-amino acid substrate. The unique feature involves the presence of one polar residue, Ser299, and two basic residues, Lys330 and Arg331 (corresponding to Ala301, Ile330, and Cys331, respectively, in PheA). The crystal structure shows that all of these residues are involved in the recognition of the carboxyl group at the C1 position that is not adenylated (Fig. 2B). Other α-amino acid-activating enzymes generally have small aliphatic residues at the positions corresponding to Ser299 and Lys330 in VinN (3–7). In addition, they contain neither Arg nor Lys at the position equivalent to Arg231 in VinN, although they have various types of residues, including Asp and His, at this position. Thus, VinN has distinct amino acids at these positions. These polar and basic residues might be specific for the recognition of dicarboxylic β-amino acids. Even adenylation enzymes using dicarboxylic α-amino acid substrates, including l-aspartate and l-glutamate, have no basic residues at corresponding positions (7). These enzymes have a basic or polar residue at a different position such as the 234 or 278 position in PheA numbering for recognition of a carboxyl group that is not adenylated. Another feature is that VinN has a bulky aromatic residue, Phe231, adjacent to the invariant Asp230. The crystal structure shows that this Phe residue likely provides a steric constraint to control the substrate conformation. Normally, α-amino acid-activating enzymes have a small aliphatic residue such as Ala or Val at this position. Thus, the substrate-conferring code of VinN is quite different from those of α-amino acid adenylation enzymes. In contrast to VinN, SlgN1 shows the above mentioned α-amino acid adenylation enzyme-like features. SlgN1 contains Gln305 at the position equivalent to Thr279 in PheA (Fig. 5). The SlgN1 structure suggests that this polar Gln305 residue might be involved in carboxyl group recognition, although the structure of the complex with 3-MeAsp was not reported (26) (Fig. 4D). SlgN1 also has a small Ala266 residue adjacent to the invariant Asp265 residue. In addition, the length of the β13β14 loop in SlgN1 is the same as those in

**FIGURE 4. Structural comparison of VinNN with other adenylation enzymes.** A, the structure of the adenylate-binding site. The superimposed structures of VinNN (green) and PheA (purple) are shown. B, the structure of the substrate-binding pocket. The superimposed structures of VinNN (green) and PheA (purple) are shown. C, the alignment of the β13β14 loop region of VinN with other adenylation enzymes. The sequences of the following 10 adenylation enzymes were used for the alignment analysis: α-amino acid-activating enzymes, including Bacillus brevis PheA (2), Streptomyces lydicus SlgN1 (26), Bacillus subtilis SrfA-C (45), and Bacillus cereus DltA (35); β-amino acid-activating enzymes, including VinN, Streptomyces sp. ML694-90F3 IdnL1 (18), Salinispora tropica Stro_2775 (21), Streptomyces sp. MJ635-86F5 CmiS6 (19), and Streptomyces sp. DSM 21069 BecJ (20); and malonyl-CoA synthetase Streptomyces coelicolor ScMatB (44). The specificity-conferring code residues are shown in red. The VinN β13β14 loop is indicated by a broken line above the sequence. The conserved positions are marked with an asterisk. D, the superimposed substrate-binding pocket structures of VinNN (green) and SlgN1 (gray).
The Crystal Structure of the Adenylation Enzyme VinN

The sequences of the following 15 adenylation enzymes were used for the phylogenetic analysis: VinN, B. subtilis DhbE (41), Streptomyces globisporus SgcC1 (16), Streptomyces rochei Orf5 and Orf19 (24), Streptomyces sp. ML694-90F3 IdnL1 (18), S. tropica Stro_2775 (21), Streptomyces sp. MJ835-86F5 CmiS6 (19), Streptomyces sp. DSM 21069 BecJ (20), Chondromyces crocatus CmdD-A7 (17), B. brevis SrfA-C (45), Pantoaea agglomerans AdmI (23), and B. cereus DltA (35). The specificity-conferring codes are also shown. VinN and other β-amino acid-activating enzymes are underlined.

α-amino acid-activating enzymes (Fig. 4C). These differences could explain why SlgN1 shows the opposite orientation of 3-MeAsp binding to VinN. A few alterations in the substrate-binding pocket affect the orientation of substrate in this case.

This study might be important from an evolutionary point of view. Adenylation enzymes might gain β-amino acid specificity from α-amino acid specificity with a few critical alterations of the substrate-binding pocket although they maintain the overall fold and catalytic mechanism. The replacement of some substrate-binding pocket residues is likely a key evolutionary event to generate various β-amino acid specificities. Sequence alignment and phylogenetic analysis show that other β-amino acid-activating enzymes can be divided into several subfamilies (Fig. 5). These enzymes possess an Asp residue at the position corresponding to Asp230 of VinN and are thought to use an Asp residue for the recognition of the β-amino group of the β-amino acid substrate in a manner similar to that of VinN. Of these, IdnL1 (18), CmiS6 (19), BecJ (20), and Stro_2775 (21), which use β-amino fatty acids in macrolactam antibiotic biosynthesis, might have a similar β-amino acid recognition mechanism because they share some residues in common with VinN.

In particular, Phe231 and Ser299 of VinN are conserved in these enzymes (Fig. 5), which might use both the Phe and Ser residues to control substrate conformation by steric constraints similar to VinN. In addition, these enzymes might have a β138β14 loop that is one residue shorter than that in α-amino acid-activating enzymes, although this is difficult to predict because of the low sequence identity (Fig. 4C). The identification of these conserved features could help the prediction of β-amino acid substrate specificity of biochemically uncharacterized adenylation enzymes on the basis of their amino acid sequences. Conversely, other types of β-amino acid-activating enzymes seem to use a different amino acid residue for recognition. For example, SgcC1, which activates β-tyrosine in C-1027 biosynthesis, contains a Pro residue adjacent to the invariant Asp (Fig. 5). This Pro residue is thought to be involved in the catalysis and/or specificity for β-amino acids as evidenced from mutational studies (16). Other enzymes have different amino acids in this position. CmdD-A7, which activates β-phenylalanine in chondramide biosynthesis, has a smaller Gly residue (17), and Orf5, which activates β-lysine in streptothricin biosynthesis, contains a Thr residue (24). These enzymes might adopt a different strategy for β-amino acid recognition. The elucidation of the structures of these enzymes is necessary for the further understanding of β-amino acid recognition by adenylation enzymes.

In conclusion, we determined the first crystal structure of a β-amino acid adenylation enzyme. This structure sheds on the molecular basis for selective activation of β-amino acids by visualizing the orientation and conformation of the β-amino acid substrate. This study represents a significant contribution to the expansion of knowledge of the adenylation enzyme family. In addition, the VinN structure could provide clues useful for the protein engineering of an adenylation enzyme to enable alteration of the substrate specificity to introduce β-amino acids instead of α-amino acids.

Acknowledgments—We thank Dr. Takatoshi Arakawa at The University of Tokyo for assistance with the CD experiment. This work was performed with the approval of the Photon Factory Program Advisory Committee (Proposal 2012G0508).

REFERENCES

1. Gulick, A. M. (2009) Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. ACS Chem. Biol. 4, 811–827
2. Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997) Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. EMBO J. 16, 4174–4183
3. Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem. Biol. 6, 493–505
4. Challis, G. L., Ravel, J., and Townsend, C. A. (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem. Biol. 7, 211–224
5. Rausch, C., Weber, T., Kohlbacher, O., Wohlsleben, W., and Huson, D. H. (2005) Specificity prediction of adenylation domains in nonribosomal peptide synthetase (NRPS) using transductive support vector machines (TSVMs). Nucleic Acids Res. 33, 5799–5808
6. Wackler, B., Lackner, G., Chooi, Y. H., and Hoffmeister, D. (2012) Characterization of the Stuillas grevillei quinone synthetase GreA supports a nonribosomal code for aromatic α-keto acids. ChemBioChem 13, 1798–1804
7. Khayatt, B. I., Overmars, L., Siezen, R. J., and Francke, C. (2013) Classification of the adenylation and acyl-transferase activity of NRPS and PKS systems using ensembles of substrate specific hidden Markov models. PLoS One 8, e62136
8. Eppelmann, K., Stachelhaus, T., and Marahiel, M. A. (2002) Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. Biochemistry 41, 9718–9726
9. Uguru, G. C., Milne, C., Borg, M., Flett, F., Smith, C. P., and Micklefield, J. (2004) Active-site modifications of adenylation domains lead to hydrolysis of upstream nonribosomal peptidyl thioester intermediates. J. Am. Chem. Soc.
10. Stevens, B. W., Lilien, R. H., Georgiev, I., Donald, B. R., and Anderson, A. C. (2006) Redesigning the PhoA domain of gramicidin synthetase leads to a new understanding of the enzyme’s mechanism and selectivity. *Biochemistry* **45**, 15495–15504

11. Chen, C. Y., Georgiev, I., Anderson, A. C., and Donald, B. R. (2009) Computational structure-based redesign of enzyme activity. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3764–3769

12. Thirlway, J., Lewis, R., Nunns, L., Al Nakeeb, M., Styles, M., Struck, A. W., Smith, C. P., and Micklefield, J. (2012) Introduction of a non-natural amino acid into a nonribosomal peptide antibiotic by modification of adenylation domain specificity. *Angew. Chem. Int. Ed. Engl.* **51**, 7181–7184

13. Kudo, F., Miyamura, A., and Eguchi, T. (2014) Biosynthesis of natural products containing β-amino acids. *Nat. Prod. Rep.* **31**, 1056–1073

14. Shinohara, Y., Kudo, F., and Eguchi, T. (2011) A natural protecting group strategy to carry an amino acid starter unit in the biosynthesis of macrolactam polyketide antibiotics. *J. Am. Chem. Soc.* **133**, 18134–18137

15. Mootz, H. D., and Marahiel, M. A. (1997) The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* **179**, 6843–6850

16. Van Lanen, S. G., Lin, S., Dorrestein, P. C., Kelleher, N. L., and Shen, B. (2006) Substrate specificity of the adenylation enzyme SggC1 involved in the biosynthesis of the enediyne antitumor antibiotic C-1027. *J. Biol. Chem.* **281**, 29633–29640

17. Racid, S., Krug, D., Weissman, K. J., and Müller, R. (2007) Biosynthesis of (R)-β-tyrosine and its incorporation into the highly cytotoxic chondramides produced by *Chromobacterium crocutas*. *J. Biol. Chem.* **282**, 21810–21817

18. Takaiishi, M., Kudo, F., and Eguchi, T. (2013) Identification of inedine biosynthetic gene cluster: characterization of novel β-glutamate-β-decarboxylase IdhL3. *J. Antibiot.* **66**, 691–699

19. Amagai, K., Takaku, R., Kudo, F., and Eguchi, T. (2013) A unique amino transfer mechanism for constructing the β-amino fatty acid starter unit in the biosynthesis of the macrolactam antibiotic cremicyn. *ChemBioChem* **14**, 1998–2006

20. Jorgensen, H., Degnes, K. F., Sletta, H., Fjaervik, E., Svendsen, G., Bruheim, P., Bredholt, H., Nygård, G., Døskeland, S. O., Elingsen, T. E., and Zocher, S. B. (2009) Biosynthesis of macrolactam BE14106 involves two distinct PKS systems and amino acid processing enzymes for generation of the aminoaoyl starter unit. *Chem. Biol.* **16**, 1109–1121

21. Udvary, D. W., Zeiger, L., Asolkar, R. N., Singan, V., Lapidus, A., Fenical, W., Jensen, P. R., and Moore, B. S. (2007) Genome sequencing reveals the basis of the interaction of MbtH-like proteins, putative regulators of non-ribosomal peptides in *Mycobacterium tuberculosis* biosynthesis. *J. Antibiot.* **60**, 450–450

22. Rouhiainen, L., Vakkilainen, T., Siemer, B. L., Buikema, W., Haselkorn, R., Smith, C. P., and Micklefield, J. (2012) Introduction of a non-natural amino acid processing enzyme to a new understanding of the enzyme’s mechanism and selectivity. *Acta Crystallogr. Sect. F Struct. Biol. Crystallogr.* **68**, 5032–5033

23. Lovell, S. C., Davis, I. W., Arendall, W. B., Satchell, P. C., and Micklefield, J. (2012) Redesigning the PheA domain of gramicidin synthetase leads to a new understanding of the enzyme’s mechanism and selectivity. *Acta Crystallogr. Sect. F Struct. Biol. Crystallogr.* **68**, 5032–5033

24. Villiers, B. R., and Hollfelder, F. (2009) Mapping the limits of substrate specificity of the adenylation domain of TyCA. *ChemBioChem* **10**, 671–682

25. Reiger, A. S., Wu, R., Dunaway-Mariano, D., and Gulick, A. M. (2008) Structural characterization of a 140° domain movement in the two-step reaction catalyzed by 4-chlorobenzoate:CoA ligase. *Biochemistry* **47**, 8016–8025

26. Goyal, A., Yousuf, M., Rajakumar, E., Arora, P., Gokhale, R. S., and Sivonen, K. (2004) Genes coding for hepatotoxic heptapeptides in the cyanobacterium *Microcystis* sp. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1076–1081

27. Arora, P., Goyal, A., Natarajan, V. T., Rajakumar, E., Verma, P., Gupta, R., Yousuf, M., Trivedi, O. A., Mohanty, D., Yagi, A., Sankaranarayanan, R., and Gokhale, R. S. (2009) Mechanistic and functional insights into fatty acid activation in *Mycobacterium tuberculosis*. *Nat. Chem. Biol.* **5**, 166–173

28. Goyal, A., Verma, P., Anandhakrishnan, M., Gokhale, R. S., and Sankaranarayanan, R. (2012) Molecular basis of the functional divergence of fatty acyl-AMP ligase biosynthetic enzymes of *Mycobacterium tuberculosis*. *J. Biol. Chem.* **287**, 221–228

29. Hayashi, S., Ariga, M., Nakanishi, K., and Suyama, M. (2008) Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12120–12125

30. Holm, L., and Sander, C. (1995) DALI: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478–480

31. Yonshus, H., Neumann, P., Zimmermann, S., May, J. J., Marahiel, M. A., and Stubbs, M. T. (2008) Crystal structure of DltA: Implications for the reaction mechanism of non-ribosomal peptide synthetase adenylation domains. *J. Biol. Chem.* **283**, 32484–32491

32. Drakes, E. J., Duckworth, B. P., Neres, J., Aldrich, C. C., and Gulick, A. M. (2010) Biochemical and structural characterization of bisubstrate inhibitors of BaE, the self-standing nonribosomal peptide synthetase adenylation enzyme of acinetobactin synthetase. *Biochemistry* **49**, 9292–9305