Structure of the EntB Multidomain Nonribosomal Peptide Synthetase and Functional Analysis of Its Interaction with the EntE Adenylation Domain

Eric J. Drake,1,2 David A. Nicolai,1 and Andrew M. Gulick1,2,*
1Hautpman-Woodward Medical Research Institute
700 Ellicott Street
Buffalo, New York 14203
2Department of Structural Biology
The State University of New York at Buffalo
Buffalo, New York 14260

Summary

Nonribosomal peptide synthetases are modular proteins that operate in an assembly line fashion to bind, modify, and link amino acids. In the E. coli enterobactin NRPS system, the EntE adenylation domain catalyzes the transfer of a molecule of 2,3-dihydroxybenzoic acid to the pantetheine cofactor of EntB. We present here the crystal structure of the EntB protein that contains an N-terminal isochorismate lyase domain that functions in the synthesis of 2,3-dihydroxybenzoate and a C-terminal carrier protein domain. Functional analysis showed that the EntB-EntE interaction was surprisingly tolerant of a number of point mutations on the surface of EntB and EntE. Mutational studies on EntE support our previous hypothesis that members of the adenylate-forming family of enzymes adopt two distinct conformations to catalyze the two-step reactions.

Introduction

The nonribosomal peptide synthetases (NRPSs) are a family of modular proteins that catalyze the enzymatic synthesis of small peptides [1–3]. The secondary metabolites produced by the NRPS systems in bacteria and fungi include peptides with antibiotic, immunosuppressive, and anticancer activities, as well as some siderophores. The enzymatic synthesis of the NRPS products raises the possibility of engineering the NRPSs for the generation of catalysts that may synthesize peptides with novel activities [4].

Modular NRPSs contain multiple catalytic domains joined in a single protein that may be tens of thousands of residues in length. Usually, NRPSs contain one module for each amino acid that is incorporated into the final peptide. A minimal NRPS module contains two catalytic domains and a carrier protein domain, onto which the nascent peptide is attached during synthesis. The adenylation domain first activates the substrate through an adenylation reaction and then catalyzes the formation of a thioester between the acyl substrate and the phosphopantetheine (PPant) cofactor on the carrier protein domain. A condensation domain serves to catalyze peptide bond formation by transferring the amino acid from an upstream domain to an amino acid located on a downstream domain. Additional catalytic domains that exist in NRPSs include N-methylation and epimerization domains [1, 5]. The NRPS carrier protein domains contain a conserved serine that is posttranslationally modified by holo-ACP synthases with a PPant cofactor [6] and are similar to acyl carrier proteins involved in other cellular processes, including fatty acid biosynthesis and metabolism and polyketide biosynthesis.

E. coli contains a single NRPS system that is used in the synthesis of the siderophore enterobactin (Figure 1A). The enterobactin molecule contains three copies of a dipeptide of 2,3-dihydroxybenzoic acid (DHB) and serine [7]. The serine molecules form a trilactone ring between the carboxylate of one residue and the side chain hydroxyl of the next. Produced and secreted during iron-limiting conditions, the enterobactin molecule chelates an Fe3+ ion and is then taken up into the cell through the activity of the outer membrane transport protein FepA [7].

The enzymes involved in enterobactin synthesis are encoded by a six gene cluster (entA–F). The DHB molecules are produced from chorismate by the activities of EntA, EntB, and EntC (Figure 1B). EntC first catalyzes the conversion of chorismate to isochorismate [8]. The isochorismate lyase (ICL) domain of EntB hydrolyzes the pyruvate group of isochorismate to form 2,3-dihydro-2,3-dihydroxybenzoate [9], which is converted to DHB by the activity of EntA, a 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase [10].

EntE, EntB, and EntF are NRPS enzymes that generate the enterobactin molecule from three molecules of DHB and three molecules of serine, by using energy derived from the hydrolysis of six ATP molecules. The EntE protein contains an adenylation domain specific for DHB (Figure 1C); EntE transfers the DHB molecule to the PPant cofactor that is bound to the aryl carrier protein (ArCP) domain of the EntB protein [11]. EntB is thus a two-domain NRPS that contains both the ICL and ArCP domains [12]. EntF contains the complete serine module and a C-terminal domain that catalyzes the formation of the trilactone ring [13, 14]. The enterobactin NRPS system, therefore, contains only two modules that work iteratively for the complete synthesis of the enterobactin molecule. The remaining protein of the enterobactin cluster, EntD, is a PPant transferase that adds the cofactors to the carrier protein domains of EntB and EntF [6, 12].

Structurally, the most well studied domain of the NRPS systems is the adenylation domain. The adenylation domains are part of a larger family of adenylate-forming enzymes that include firefly luciferase and acyl-CoA synthetases [15, 16]. All members of this family catalyze two-step reactions in which the first reaction is an adenylation step resulting in the formation of an acyl-AMP. NRPS adenylation domains and the acyl-CoA synthetases catalyze the formation of a thioester with either PPant or with CoA in the second half-reaction. In its second half-reaction, luciferase catalyzes an oxidative decarboxylation that results in an activated molecule that decomposes to emit light [17].

*Correspondence: gulick@hwi.buffalo.edu
Crystal structures exist for firefly luciferase [17], NRPS adenylation domains PheA [18] and DhbE [19], and several acyl- and aryl-CoA synthetases [20–23]. The enzymes contain a large N-terminal domain of 400–520 residues and a smaller C-terminal domain of ~120 residues. On the basis of structural and biochemical work, we have previously proposed [20, 21] that the members of this family adopt one conformation to catalyze the initial adenylation reaction. The C-terminal domain then rotates by ~140° to adopt a second conformation that is used to catalyze the second half-reaction. This conformational change, which we term domain alternation, reconfigures a single active site to catalyze the two half-reactions.

We have determined the crystal structure of the EntB protein and present here further insight into the activity of the NRPS adenylation domains. By using the EntB structure as well as structures of EntE homologs in adenylate- and thioester-forming conformations, we identified residues near the potential recognition interfaces of EntB and EntE. We have created a series of point mutations in EntB and tested them through kinetic analyses with EntE. We have also created a series of mutations in the EntE adenylation domain. Our results demonstrate that the EntE adenylation domain is remarkably tolerant of the point mutations in the EntB protein, and that a variety of single-residue replacements have little effect on the EntE-EntB interaction. These results also support the domain rotation proposal for adenylate-forming enzymes.

Results

Structure Determination of EntB

Crystallization of EntB was achieved through initial screening with sparse matrix conditions by using a wide variety of PEG- and salt-based precipitants [24, 25]. The structure of the two-domain EntB protein was determined by molecular replacement with PHASER [26, 27]. The search model was a single domain of the 207 residue 2-amino-2-deoxyisochorismate lyase (PhzD) protein [28]. PhzD, involved in phenazine biosynthesis, is 45.5% identical to the ICL domain of EntB. A molecular replacement solution was found with a dimer of ICL domains in the asymmetric unit. Gel filtration experiments demonstrated that the EntB forms a dimer in solution as well. Electron density for portions of the main helices of the ArCP domain was initially visible in difference density maps, and the connecting loops became apparent as model building and refinement proceeded. Crystallographic and refinement data statistics are shown in Table 1.

The final model for EntB contains 558 residues; subunit A is missing the N-terminal 4 residues and the C-terminal 2 residues, while subunit B is missing 3 residues from each terminus. Several residues have disordered side chains and have been modeled as alanines. Linking the N-terminal ICL domain with the C-terminal ArCP domain is the 5 residue sequence Pro-Ala-Pro-Ile-Pro. This extended peptide positions the C-terminal carrier protein domain away from the ICL domain (Figure 2A). The N-terminal domains of the two subunits of EntB superimpose with a root-mean-square (rms) deviation of 0.23 Å. Greater variability is observed with the C-terminal domain. The rms deviation derived from overlapping the two C-terminal domains is 0.32 Å; however, the C-terminal domains of the two subunits adopt different conformations relative to the N-terminal domain. The difference in the orientation between the two C-terminal domains is ~23°.

The N-terminal domain of EntB contains a central six-stranded parallel β sheet (Figure 2). On one side of the sheet are three α helices that form the subunit interface of the EntB dimer. On the opposite side of the sheet is a long α helix at residues Pro51-Asn71. A short single-turn helix caps the C-terminal ends of the β sheet. A small pocket forms between the core of the ICL domain and an interrupted helix at Asp85-Met94. The distal side of this
helix makes hydrophobic interactions with the neighboring subunit, and Leu90 and Leu91 of one subunit interact with Trp24 and Tyr165 of the neighboring subunit of the dimer. A second set of interactions occurs between Met94 and Trp95 of one subunit and the methylene carboxylate of isochorismate. A second set of interactions occurs between subunit, and Leu90 and Leu91 of one subunit interact

The ArCP domain of EntB is a typical acyl carrier protein domain consisting of four helices. The Ser245 residue of EntB, which is the site of addition of the cofactor, is positioned at the start of the second helix. The Ser245 residue is located on the ArCP face that points away from the ICL domain and is directed into solvent (Figure 2A). This location allows access of this site to the cofactor and to the EntE and EntF proteins with which it will interact. In the crystal lattice, the Ser245 side chain is positioned close to a symmetry-related molecule, and no electron density is observed for the PPant cofactor. Prior to structure solution by molecular replacement, attempts were made, without success, to grow crystals of selenomethionine (SeMet)-labeled EntB. SeMet-labeled protein was produced by the metabolic inhibition method [29] in minimal media, which, lacking sufficient iron, induced the EntD PPant transferase [30]. As the Ser245 side chain is in close proximity to a neighboring molecule in the crystal lattice, the holoprotein containing the PPant cofactor could not crystallize under the current conditions in the crystal form observed with the apo-protein.

A number of acyl, fatty acyl, and peptidyl carrier protein structures have been determined by crystallography and NMR [31–42]. There is considerable structural variation; however, nearly all structures contain the four main helices. A long loop joins helices 1 and 2, and the cofactor binding serine residue is located at the start of helix 2. Helices 1, 2, and 4 are similar in length and orientation, while helix 3 is shorter and nearly perpendicular to the other helices. A simple overlay of the Cz position of the different structures shows differences in the orientations of the helices and results in relatively high rms deviations of ~2.5 Å over the complete domains; this largely results from differences in the orientations of the main Cz helices. To best compare the structures of the different carrier protein domains, we used the program PROMOTIF [43] to determine the angles of interaction between different helices. For EntB, the angles between the helices were very similar to the values observed for the peptidyl carrier protein domain of the TycC3 NRPS protein [31]. The orientations of the helices show some consistency between three different classes of carrier protein domains: NRPS PCP domains, fatty acyl carrier protein domains, and ACP domains used in polyketide synthesis (Table 2). The angles observed between helices of the two NRPS proteins are more similar to those of the fatty acyl carrier protein domains than to the interhelical angles of the polyketide domains. Representative carrier proteins are shown in Figure 3C. We also compared the angles of the structures determined by crystallography with those determined by NMR and showed that the values did not correlate with experimental method (Table 2). Although the number of structures limits a full statistical correlation, the NRPS carrier proteins are more similar with respect to helical orientation to the acyl carrier proteins from primary metabolism than to the carrier protein domains involved in polyketide secondary metabolism.

### Functional Analysis of EntB and EntE Mutants

By using our structure of EntB and prior structures of adenylate-forming enzymes [19–21] as a guide, we constructed point mutations in residues at the surfaces of
EntB (Figures 4A and 4B) and EntE (Figure 4C) and tested these mutants for the intermolecular transfer of DHB from EntE to the EntB ArCP domain. Three loops (Leu225–Ser229, Ile239–Arg256, Asp261–Asn270) were identified on the surface of the ArCP domain near the PPant binding site of Ser245 (Figure 4B). Examination of the 32 residues that formed the interaction loops of EntB demonstrated that 10 of these residues were directed into the core of the protein and were unlikely to be involved in specific side chain interactions with EntE. Of the remaining 22 residues, we chose 9 residues for mutagenesis that might contribute to interactions with EntE; these residues were more centrally located near Ser245 and were also positioned where they might introduce charge or hydrophobic contacts toward the interface (Figure 4A). To determine if the residues formed specific interactions across the EntB/EntE interface, residues were replaced with amino acids that exhibited significantly different biochemical properties.

It was first necessary to produce the holo-EntB mutants and confirm that the mutations did not affect the ability of EntD to attach the PPant cofactor on the EntB mutants. We relied on the in vivo pantetheinylation of EntB by producing protein in minimal media. We used MALDI-TOF mass spectrometry to analyze cofactor content. Initially, two samples of wild-type protein were submitted, one purified from rich media and one purified from minimal media (Table 3). The difference in molecular weight between the two samples was 341.2 Da, representing the presence of the PPant cofactor on the protein produced in minimal media. To gauge the ability to use mass spectrometry as an analytical tool, a 50:50 mixture of the two samples was analyzed. The [M+2H]^2+ peak clearly distinguished the two samples and showed equal peak heights. Although mass spectrometry is not quantitative, we felt that this could be used to provide a rough estimate of the degree of cofactor addition. The remaining mutants were grown in M9 and subjected to mass spectrometric analysis. The cofactor was properly added by EntD to all mutants except one (Table 3). The remaining mutant, D244R, showed a mixture of apo- and holo-protein, with the peak of the holo-protein about 40% the height of the apo-protein peak. Three independent preparations of the D244R mutant were performed, and all generated the same result. This suggests that mutation of the Asp244 residue reduces the recognition of EntB by the EntD PPant transferase; however, a small fraction (≈35% of the total EntB D244R protein produced) was still modified. In E. coli ACP, a homologous D35C mutant is also unable to be properly converted to the holo-ACP [44].

We then tested the ability of EntE to recognize the holo-EntB mutants and transfer the DHB molecule to
the PPant cofactor of EntB by using a coupled AMP-formation assay. The S245A mutant and the wild-type EntB purified from cells grown in iron-rich media showed very limited activity as substrates for EntE. We then measured specific activity by using a single concentration of the mutant EntB proteins. The mutant enzymes all showed detectable activity with EntE; several served as better acceptor substrates than the wild-type EntB (Table 3).

Surprisingly, none of the dramatic mutations that we constructed hindered the specific activity; only two residues (Asp240 and Asp263) showed reductions in activity. While the D244R mutant showed only 13% activity, this assay was performed with the mixture of apo- and holo-protein and is thus difficult to interpret. We performed additional kinetic analyses with a subset of the mutant EntB enzymes that included the two mutant proteins that had the lowest specific activity and one of the mutants that showed elevated activity, F264E. Apparent kinetic constants were determined at constant concentrations of DHB and ATP and varying concentrations of the holo-EntB. As would be expected for residues located at the surface that may impact the recognition of EntB and EntE, there is only a minimal (less than 2-fold) effect on $k_{cat}$. In contrast, the two mutations with a mild impact on relative specific activity show increases in $K_m$.
Mutational Analysis of EntE

We have proposed that members of the adenylate-forming family adopt two conformations to catalyze the two-step reaction [20, 21]. Starting with the crystal structure of DhbE [19], we created a model in which the C terminus of DhbE is in the thioester-forming conformation observed in the structure of acetyl-CoA synthetase bound to adenosine-5'-pentaphosphate and CoA [21]. From this model, several residues were identified that project toward the surface of the DhbE protein in close proximity to the PPant binding tunnel. In the crystal structure of DhbE, these residues are more than 20 Å away from the active site in the adenylate bound conformation (Figure 4C). These residues were mutated to test for the ability to affect the EntB-EntE interaction.

EntE mutations made were R437D, K473D, R494D, and a double mutant containing R437D and K473D. Neither the wild-type nor any of the EntE mutants showed spontaneous hydrolysis of ATP to produce AMP in the absence of an EntB thiol acceptor (data not shown). Importantly, all four mutants were able to catalyze the exchange of radiolabeled pyrophosphate into ATP; thus, all four mutants were competent to catalyze the adenylate reaction (Table 3). In assays of the complete thioester-forming reaction, the K473D and R494D mutants had near wild-type activity, while R437D and the double mutant R437D/K473D were reduced to 10% and 3% specific activity, respectively, with the wild-type EntB substrate. These activity levels are comparable to the activities with wild-type EntE and the Ser245 mutant of EntB. Arg437 immediately precedes Gly438, a 100% conserved residue that is located on a loop that projects into the active site in the thioester-forming conformation observed in Acs [21]. These results therefore support domain alternation, as the ability of this mutation to affect the catalytic activity of the thioester-forming reaction specifically implies that this loop indeed rotates into the active site for the catalysis of the second half-reaction.

Discussion

We present here the structure of the two-domain NRPS protein EntB containing an N-terminal ICL domain and a C-terminal ArCP domain. The structure provides a view of the ICL catalytic active site. The EntB ICL domain is a member of a family of isochorismatase enzymes that may also serve as targets for inhibitor design [45]. There has been recent interest in the development of inhibitors of the enzymes involved in chorismate metabolism that may serve as leads for antibiotic development [46–48]. Also, the structure of a multidomain NRPS suggests how the domains are organized, and our structural and functional studies can provide additional clues to the interaction of NRPS domains.

The comparison of the ICL domain with the active site of PhzD [28] demonstrates strong conservation of important residues. The catalytic aspartic acid (Asp37 in EntB) is retained, as are the residues that form the mostly hydrophobic binding pocket (Phe42, Trp95, and Tyr126). Residues that interact with the aryl carboxylate (Gln79 and Arg88) are conserved, as is Lys123, which is positioned to interact with the pyruvate carboxylate. Indeed, the only inner shell residue of the active site that differs is the replacement of Val54 of PhzD with Ile155 in EntB (Figure 3B). Interestingly, this residue is located near the C4 position of the isochorismate ligand of the PhzD structure. The catalytic efficiency of PhzD is reduced by a factor of 50 with chorismate compared to isochorismate [28]. In contrast, the EntB protein displays a 103-fold decrease in catalytic efficiency for chorismate compared to isochorismate [9]. Our structural results suggest that the specificity reflects the inability of EntB to accommodate the hydroxyl at the C4 position of chorismate because of the replacement of Val154 in PhzD with Ile155 in EntB. Additional functional and structural studies of the ICL domain will further clarify the roles of individual active site residues.

Prior structural studies on NRPS proteins have been limited to single domains, either self-standing proteins like the adenylation domain DhbE [19] and the condensation domain VibH [49], or genetically truncated domains such as the phenylalanine-activating domain of GrsA [18], the PCP domain of TyS [31], and the cyclization/thioesterase domain of SrfA-C [50]. Although the two domains of EntB do not interact functionally, EntB serves as a suitable small model to demonstrate that some multidomain NRPS proteins retain sufficient structural rigidity for crystallization.

EntB uses a proline-rich linker to join the two domains. This peptide limits interactions between the two domains and maintains the active site of the N-terminal ICL domain and the PPant cofactor binding site of the ArCP domain at ~45 Å from each other. Given the apparent functional and structural independence of these ICL and ArCP regions of the protein, one could ask

---

Table 2. Interhelical Angles Observed in Different Carrier Protein Domains

<table>
<thead>
<tr>
<th></th>
<th>H1–H2 Angle</th>
<th>H1–H4 Angle</th>
<th>H2–H3 Angle</th>
<th>H2–H4 Angle</th>
<th>H3–H4 Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>PKS (n = 3)</td>
<td>134.9 ± 4.7</td>
<td>118.2 ± 6.5</td>
<td>141.3 ± 1.6</td>
<td>−20.8 ± 3.6</td>
<td>148.4 ± 7.5</td>
</tr>
<tr>
<td>NRPS (n = 2)</td>
<td>141.8 ± 0.8</td>
<td>132.9 ± 5.0</td>
<td>113.4 ± 2.3</td>
<td>−24.2 ± 11.5</td>
<td>128.5 ± 1.6</td>
</tr>
<tr>
<td>ACP (n = 8)</td>
<td>141.8 ± 15.1</td>
<td>134.9 ± 15.4</td>
<td>126.2 ± 13.1</td>
<td>−23.5 ± 9.9</td>
<td>137.2 ± 13.7</td>
</tr>
<tr>
<td>NMR (n = 8)</td>
<td>138.1 ± 15.7</td>
<td>132.7 ± 17.3</td>
<td>131.0 ± 16.0</td>
<td>−21.9 ± 9.5</td>
<td>133.0 ± 13.7</td>
</tr>
<tr>
<td>X-ray (n = 5)</td>
<td>138.1 ± 11.1</td>
<td>127.5 ± 7.1</td>
<td>122.4 ± 10.8</td>
<td>−25.0 ± 5.9</td>
<td>143.3 ± 9.9</td>
</tr>
</tbody>
</table>

The average interhelical angles are listed for carrier proteins grouped either by experimental method or by biological function (PKS, polyketide synthase; NRPS, NRPS acyl or peptidyl carrier protein domains; ACP, acyl or fatty acyl carrier proteins). A total of 13 proteins were used in the analysis. These include three PKS ACP domains (1NQ4, oxytetracycline ACP from S. rimosus [39]; 1F80, frenolicin ACP from S. roseofulvus [40]; 2AF6, actinorhodin ACP from S. coelicolor [35]), two NRPS carrier protein domains (1DNY, tyrocidin PCP domain from TyS from B. brevis [31]; ENTB, E. coli protein presented here), and eight acyl carrier protein domains (1L0H, 1T8K, and 1ACP, three independently determined structures of B. subtilis acyl carrier protein [36, 37]; 1DV5, 1VD8, 1KL7, 1FU, 1OR5, frenolicin ACP from S. rimosus [39]; 1F80 and 1HY8, two independently determined structures of B. subtilis acyl carrier protein [36, 37]; 1DV5, D-alanyl carrier protein from L. casei [41]; 1KL7, acyl carrier protein from M. tuberculosis [42]; 1VKU, acyl carrier protein from T. maritima).
why the two domains are expressed as a single protein. Perhaps this represents an extreme level of coregulation of these two elements in enterobactin synthesis. The synthesis of DHB from chorismate involves three steps catalyzed sequentially by EntC, EntB, and EntA. The reaction catalyzed by EntC is reversible; the $K_{eq}$ for the reaction is 0.56 in favor of the chorismate-to-isochorismate direction [8]. The ICL step catalyzed by EntB is thus the committed step toward DHB production. By combining the ICL and ArCP domains, nature has
evolved to ensure the availability of the EntB ArCP domain when DHB is produced.

A long-standing goal, with obvious technical challenges, is to determine the structure of even larger multidomain NRPSs. Until such studies are successful, structures of individual domains will continue to provide details that can be used in the generation of models for the interaction of neighboring domains. The creation of such models will benefit from the limitations imposed by available structural and biochemical data. For example, the generation of a model for the interaction of the EntE adenylation domain and the EntB ArCP domain is constrained by the lengths of linker peptides that occur between fused adenylation and carrier protein domains. To identify the linker region between the adenylation and PCP domains, the C-terminal Ala536 of EntE was taken as the N-terminal residue of the ICL domain in a way that enables it to interact with only moderate affinity with any one partner, but one corresponding to the apo-protein and one corresponding to the holo-protein. The relative kinetic activity represents the activity observed with the mixture of apo- and holo-proteins.

Table 3. Kinetic Analysis of the EntB-EntE Interaction

<table>
<thead>
<tr>
<th>EntB Protein</th>
<th>Expected MW</th>
<th>Observed MW*</th>
<th>Relative Specific Activityb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type</td>
<td>33,091.7</td>
<td>33,100</td>
<td>1.00 ± 0.05</td>
</tr>
<tr>
<td>WT-rich media</td>
<td>32,750.3</td>
<td>32,759</td>
<td>0.05 ± 0.01</td>
</tr>
<tr>
<td>S245A</td>
<td>32,734.3</td>
<td>32,726</td>
<td>0.02 ± 0.01</td>
</tr>
<tr>
<td>D227R</td>
<td>33,132.6</td>
<td>33,119</td>
<td>3.64 ± 0.35</td>
</tr>
<tr>
<td>I239D</td>
<td>33,093.4</td>
<td>33,093</td>
<td>2.68 ± 0.13</td>
</tr>
<tr>
<td>D240R</td>
<td>33,132.6</td>
<td>33,121</td>
<td>0.60 ± 0.11</td>
</tr>
<tr>
<td>D244R apo³</td>
<td>32,791.4</td>
<td>32,781</td>
<td>0.13 ± 0.02</td>
</tr>
<tr>
<td>D244R holo³</td>
<td>33,132.6</td>
<td>33,123</td>
<td>—</td>
</tr>
<tr>
<td>R247E</td>
<td>33,064.4</td>
<td>33,057</td>
<td>2.15 ± 0.02</td>
</tr>
<tr>
<td>M249E</td>
<td>33,089.4</td>
<td>33,092</td>
<td>1.04 ± 0.07</td>
</tr>
<tr>
<td>D263R</td>
<td>33,132.6</td>
<td>33,117</td>
<td>0.66 ± 0.01</td>
</tr>
<tr>
<td>F264E</td>
<td>33,073.4</td>
<td>33,070</td>
<td>3.15 ± 0.26</td>
</tr>
<tr>
<td>K269E</td>
<td>33,092.4</td>
<td>33,090</td>
<td>2.42 ± 0.37</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EntE Mutant</th>
<th>Relative EntE Specific Activity (Complete Reaction)b</th>
<th>Relative PPI-Exchange Activityb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type</td>
<td>1.00 ± 0.05</td>
<td>1.00 ± 0.11</td>
</tr>
<tr>
<td>R437D</td>
<td>0.10 ± 0.01</td>
<td>0.94 ± 0.15</td>
</tr>
<tr>
<td>K473D</td>
<td>0.87 ± 0.08</td>
<td>0.98 ± 0.17</td>
</tr>
<tr>
<td>R494D</td>
<td>0.92 ± 0.11</td>
<td>0.90 ± 0.22</td>
</tr>
<tr>
<td>R437D/K473D</td>
<td>0.03 ± 0.01</td>
<td>0.84 ± 0.12</td>
</tr>
</tbody>
</table>

* Molecular weight was observed by MALDI-TOF Mass spectrometry as described in the Experimental Procedures.

b Activity was calculated for the EntE-catalyzed reactions by using the complete reaction (measuring AMP production with the coupled assay) or the adenylation reaction by using PPi-exchange. Results are expressed relative to wild-type.

³ Two peaks were observed in the MS analysis of the D244R mutant, one corresponding to the apo-proteins.

NRPSs are involved in the synthesis of small peptides that can have antibiotic, immunosuppressant, or other activities. The modular nature of the NRPSs makes them suitable targets for engineering through the modification of substrate binding pockets or through domain or module rearrangements. The ability to reposition domains, however, requires an understanding of the nature of the interactions between specific domains to avoid the juxtaposition of incompatible domains. The crystal structure of the EntB protein provides insight into the nature of an NRPS two-domain protein and demonstrates that the ArCP domain protrudes from the ICL domain in a way that enables it to interact with EntE and the condensation domain of EntF. Our functional analysis suggests that the nature of the interaction of EntB with EntE is rather
promiscuous, and that multiple single-residue point mutations have little effect on the ability of EntE to recognize EntB. Our structural results also demonstrate that the NRPS carrier protein domains may share more overall structural similarities with ACP domains involved in fatty acid synthesis than with the ACP domains of polyketide biosynthesis. The mutagenesis studies presented here support the domain alternation hypothesis, which suggests that adenylation domains will undergo a 140° domain rotation to catalyze the two sequential chemical steps. This feature, together with consideration of the lengths of interdomain linkers, will assist in the modeling of the interaction of adenylation domains with carrier protein domains in both inter- and intramolecular transfers of acyl substrates.

Experimental Procedures

Cloning, Expression, and Purification of EntE

The entE gene was cloned from E. coli JM109 DNA by PCR. Template DNA was prepared by boiling 10 μl of an overnight culture in 90 μl water for 5 min. The primers used for PCR were 5'-CCAGTGCG GATGCGCGTGAC-3' and 5'-CCACTCGAGTCAGGCT CGATGGCCTAGGCT-3'. The 1.6 kb gene product was cloned initially into pgEM7z and then subcloned into a modified pET15b plasmid that contained a TEV protease site [51]. All sequences were confirmed by DNA sequencing. This plasmid was used for expression in BL21(DE3) cells. Cultures of cells were grown to an OD 600 of ~0.6 at 37°C and were induced with 0.5 mM IPTG for 3 hr. Cells were lysed by sonication in a buffer containing 40 mM HEPES (pH 7.5 at 4°C), 150 mM KCl, 20 mM imidazole, and 0.2 mM tris(carboxyethyl)phosphine (TCEP). Protein was purified by metal chelate affinity chromatography. The column was washed with 50 mM imidazole, followed by elution of tagged EntE with the lysis buffer containing 300 mM imidazole. The purified protein was dialyzed into cleavage buffer (50 mM N-2-hydroxyethylpiperazine-N'-3-propanesulfonic acid [HEPPS] [pH 8.0 at 4°C], 150 mM NaCl, 0.5 mM EDTA, and 0.2 mM TCEP). TEV protease (~0.3 mg) was added to the tagged protein and allowed to react overnight at 4°C in the dialysis bag. The cleaved protein was passed over a metal chelate column, and the untagged protein was collected from the flowthrough. The final protein was dialyzed into 10 mM HEPES (pH 7.5 at 4°C), 0.2 mM TCEP, and 1 mM Na2S. From 1 L of cells, ~45 mg protein was obtained. This protein was frozen by pipetting directly into liquid nitrogen and was stored at ~80°C.

Cloning and Expression of EntB

The primers used to clone entB were 5'-CCAGTGCGCATATGGC TATGGACATTCCATTCACCCG-3' and 5'-CCACTCGAGTATCGCGCT GATGGCCTAGGCT-3'. The amplified gene was cloned into pgEM7z and then pET15b/TEV. Protein expression in BL21(DE3) cells was induced with 0.5 mM IPTG overnight at 16°C. Purification was again obtained by two metal chelate chromatography steps separated by the dialysis bag. The cleaved protein was passed over a metal chelate column, and the untagged protein was collected from the flowthrough. The final protein was dialyzed into 10 mM HEPES (pH 7.5 at 4°C), 0.2 mM TCEP, and 1 mM Na2S. From 1 L of cells, 45 mg protein was obtained. This protein was frozen by pipetting directly into liquid nitrogen and was stored at ~80°C.

Expression of Holo-EntB

Attempts to grow crystals of SeMet-labeled protein, produced by metabolic inhibition in M9 minimal media, were never successful. Once the final structure was obtained by using crystals grown of na- native protein, it was discovered that Ser245 was clearly unmodified with the PPant cofactor. We reasoned that production of SeMet pro- tein by growth in minimal media induced the host enter operon, resulting in the production of the EntD PPant transferase. The presence of the cofactor on Ser245 sterically hinders lattice packing and provides an explanation for why we were unable to crystallize the SeMet protein. holo-EntB, used for functional assays, was produced by growing cells in M9 minimal media supplemented with 0.18% glucose, 148 μM thiamine, and 50 μg/ml ampicillin. Reproducing the protocol that we had used to make SeMet-labeled protein, an amino acid cocktail was added to a final concentration of 27 μM lysine, 42 μM threonine, 30 μM phenylalanine, 19 μM leucine, 19 μM isoleucine, 21 μM valine, and 17 μM methionine 30 min prior to induction. The purification was the same as for wild-type protein. EntB samples were submitted to the Roswell Park Cancer Institute Proteomics Fa- cility for mass spectrometry analysis with a Waters Micromass MALDI Micro MX mass spectrometer with 10 mg/ml sinapinic acid as the matrix. The molecular weight was calculated from the [M+2H]+ ions because of improved resolution.

Crystallization of EntB and Structure Determination

Crystals of EntB were grown at 14°C by hanging drop vapor diffusion with a precipitant containing 10%−15% methyl ether PEG 3000, 0.7–0.9 M MgCl2, 10% ethylene glycol, and 50 mM HEPPS (pH 8.0). The crystals tended to be highly anisotropic, and multiple additives were tested to improve the diffraction. The best crystals were obtained by transferring the crystals from mother liquor to solutions containing CaCl2 instead of MgCl2. The crystals were simultaneously cryoprotected by increasing the concentration of ethylene glycol to 22%. Crystals were soaked for 9 min in 15% MePEG5000, 0.75 M CaCl2, 11% ethylene glycol, 50 mM HEPPS (pH 8.0), and they were soaked for 8 min in the same solution with 22% ethylene glycol. Diffraction data were collected at ~170°C by using a Rigaku RU-H3RB rotating anode, Osmic Max-Flux confocal focusing mirrors, and an R-axis IV image plate. Diffraction images were processed with MOSFILM [52], merged with SCALA, and converted to structure factors with TRUNCATE of the CCP4 software suite [53].

The EntE structure was determined by molecular replacement by using a single subunit of the PhzD protein [28] (PDB code: 1NF9) as a search model. The program PHASER [27] was used to identify the location of the two molecules in the asymmetric unit. Refinement of the initial solution with REFMAC resulted in a model with a crystallo- graphic Rfactor of 38.3% (Rfree of 42.5%). Statistics for data collection and refinement are presented in Table 1. Representative electron density is shown in Figure 2B.

Mutagenesis of EntB and EntE

Site-directed mutagenesis was carried out by using the QuikChange Mutagenesis Kit (Stratagene) and by following the manufacturer’s protocol. Oligonucleotides were designed to contain the desired point mutation plus additional silent mutations to validate the reaction by using a restriction digest. PCR was performed under the following conditions: denaturation at 95°C for 30 s, followed by 16 cy- cles of denaturation at 95°C for 30 s, annealing at 55°C for 1 min, and extension at 68°C for 10 min. Plasmid DNA was screened for the inserted silent restriction change, and the entire coding region was sequenced to confirm the introduction of the desired mutation only.

Assay for Adenylation Activity of EntE with Wild-Type or Mutant EntB Proteins

Specific activity was determined by using an NADH+ consumption assay monitored at OD530. Standard assays contained 50 mM HEPES buffer (pH 7.5), 5 mM MgCl2, 1 mM TCEP, 1 mM ATP, 2 mM DHB, 3 mM phosphoenolpyruvate, 4 U myokinase, 4.5 U pyruvate kinase, 6.5 U lactate dehydrogenase, and 250 μM NADH+. EntE and EntB proteins were introduced into the reaction at 2 μM and 60 μM, respectively. Specific activity was calculated from the linear disappearance of NADH+, monitored for 180 s or until the absorbance due to NADH+ reached a plateau (due to covalent modifica- tion of all holo-EntB). Wild-type activity was 0.18 μmol/min/mg EntE. To determine apparent kinetic constants, the EntB concentra- tion was varied from 10 to 90 μM. The program DYNAFIT [54] was used to determine kinetic constants from the initial velocities.

Pyrophosphate Exchange Assay

The pyrophosphate exchange assay was performed by using standard techniques [11, 15]. In this assay, EntE is allowed to catalyze the initial adenylation half-reaction in the absence of a thiol acceptor molecule; during the reverse reaction, radiolabeled PPI is incorpo- rated into ATP. Specifically, 1.0 μM EntE was added to 2 mM ATP, 0.2 mM NaPPI, 50 mM HEPES (pH 8.0), 100 mM NaCl, 10 mM
MgCl₂, 0.15 μCi 32P-PPi, and 5 mM DHB. Reactions (100 μl) were incubated for 10 min at 37°C, then quenched with 0.5 ml 1% charcoal, 0.1 M unlabelled PPI, and 0.35 M perchloric acid. The charcoal was pelleted, washed twice with 1 ml H₂O, and resuspended in 0.5 ml H₂O for scintillation counting. Five replicate assays were performed for each protein; results were recorded as μm mol radiolabeled ATP incorporated per min per mg of enzyme. Wild-type activity was 210 μm mol radiolabeled ATP produced per min per mg of EntE.

Acknowledgments
We would like to thank Kristen M. Homick for assistance with cloning entE and entB and Jeff Mocny for assistance in crystallizing the EntB protein. We would also like to thank Drs. Young-Mee Park, Rama Dey-Rao, and Liguo Song of the Roswell Park Proteomics facility for the mass spectrometry analysis. This research was supported by National Institutes of Health Grant GM-068440 (to A.M.G.) and start-up funds from the John R. Oishei Foundation.

Received: July 13, 2005
Revised: January 30, 2006
Accepted: February 10, 2006
Published: April 21, 2006

References


