FeeM, an N-Acyl Amino Acid Synthase from an Uncultured Soil Microbe: Structure, Mechanism, and Acyl Carrier Protein Binding

Ryan M. Van Wagoner1, 2 and Jon Clardy1,*
1 Department of Biological Chemistry and Molecular Pharmacology
Harvard Medical School
240 Longwood Avenue
Boston, Massachusetts 02115

Summary

Attempts to access antibiotics by capturing biosynthetic genes and pathways directly from environmental DNA, which is overwhelmingly derived from uncultured bacteria, have revealed a large and previously unknown family of N-acyl amino acid synthases (NASs). The structure of the NAS FeeM reveals structural similarity to the GCNS-related N-acyl transferases and acylhomoserine lactone synthases. The overall structure has a central β sheet with α helices on both sides. A bound product at a cleft in the β sheet identifies the active site and the structural basis for catalysis, and sequence conservation in this region indicates a bias for recognition over speed. FeeM interacts with an acyl carrier protein (FeeL), and the structure, mutagenesis, and enzymatic measurements reveal that a small hydrophobic pocket in α helix 5 dominates binding of FeeM to FeeL. The structural and mechanistic analyses suggest that the products of FeeM could be bacterial signaling agents.

Introduction

Bacteria, especially soil bacteria, have contributed many of our most useful small-molecule therapeutic agents (Clardy and Walsh, 2004). Soil bacteria produce well-known antibiotics such as erythromycin, vancomycin, and tetracycline, along with one of the most recently introduced antibiotics, daptomycin, a first-in-class antibiotic for skin infections (Miao et al., 2005). Similar examples could be given for many diseases, especially cancer and immunosuppression. These contributions took on a special significance with the discovery that the overwhelming majority of bacteria cannot be cultured by currently available techniques (Pace, 1997; Hugenholtz et al., 1998; Rappé and Giovannoni, 2003). Culture-free methods of assessing microbial diversity, typically PCR of ribosomal 16S sequences from environmental samples, indicate that less than 1%, perhaps less than 0.1%, of soil microbes can be cultured in the laboratory (Torsvik and Øvreås, 2002). Taken at face value, this analysis argues that the number of small-molecule therapeutic agents yet to be discovered from bacterial sources could outnumber those we already know by one or two orders of magnitude. However, accessing them will require placing either the growth of the producing microbes or the production of the active metabolites under investigator control.

Our laboratory has been pursuing an approach to the second strategy (Rendon et al., 2000). The initial approach involved cloning DNA extracted directly from soil (environmental DNA) to produce cosmid libraries in *Escherichia coli* and to screen for heterologous expression of antibiotic-producing pathways (Brady and Clardy, 2000; Brady et al., 2001). The approach is based on the observation that prokaryotic biosynthetic pathways are almost always on a continuous stretch of DNA that typically occupies 10–110 kb (Hopwood, 1999; Omura et al., 2001). In our approach, DNA fragments are limited to less than 50 kb, and expression has been investigated only in *E. coli*. In spite of these limitations, it has led to the discovery of many previously unreported antibiotic compounds (Brady et al., 2002; Brady and Clardy, 2000, 2004, 2005a, 2005b, 2005c).

In these studies, the most frequently encountered small molecules were N-acyl derivatives of amino acids (tyrosine, phenylalanine, tryptophan, or arginine) or polyamines in which the acyl group was a fatty acid. These small molecules were produced by a single enzyme, an N-acyl amino acid synthase (NAS), and every cosmid library has had multiple members with NAS genes that produced the corresponding small-molecule products. Many NASs have been cloned and sequenced, and sequence analyses revealed little sequence similarity to genes with an annotated function or with each other (Brady et al., 2004). The exception to this lack of relatedness was a small family of tyrosine NASs that were related to each other although they were unrelated by sequence to other tyrosine NASs. Some of the NASs appear as single genes, while others are part of more complex biosynthetic sequences. A particularly attractive NAS for further analysis was identified in the *fee* pathway (Brady et al., 2002), which encodes the production of enol esters from tyrosine and fatty acids by way of an N-acyl tyrosine intermediate (Figure 1). The NAS in this pathway, FeeM, is the major subject of this report. Its pathway neighbor, FeeL, has strong sequence similarity to an acyl carrier protein (ACP), and this tight genomic association suggested that FeeM used an ACP bound fatty acid as substrate. Given the lack of information available from sequence analysis, we elected to pursue the structural analysis of at least one NAS to see if structure could provide insights into relationships to proteins with known function, the structural basis of NAS catalysis, recognition of cognate ACPS, and most importantly their biological function.

Results and Discussion

Structure of FeeM

FeeM crystallized in the space group I4_22 with eight monomers arranged as four dimers in the asymmetric unit (see Table 1 for structure and data statistics). FeeM migrates as an apparent dimer during preparative
size-exclusion chromatography, but subsequent analytical chromatography showed that FeeM migrates as a monomer at concentrations <3 μM. Enzymatic activity was observed for FeeM in an assay at concentrations in the 50 nM range (data not shown). Although there are instances of conserved residues at the dimer interface, there are no patches of mutually conserved residues that interact with one another. These data indicate that dimerization is a crystallographic artifact unrelated to catalysis.

Table 1. Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th></th>
<th>Inflection</th>
<th>Remote</th>
<th>Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>30–3.0 Å</td>
<td>20–3.05 Å</td>
<td>20–3.15 Å</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.081 (0.599)&lt;a&gt;</td>
<td>0.093 (0.777)</td>
<td>0.088 (0.622)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>7.9 (1.3)</td>
<td>7.5 (1.0)</td>
<td>7.6 (1.3)</td>
</tr>
<tr>
<td>R&lt;sub&gt;anom&lt;/sub&gt;</td>
<td>0.029 (0.158)</td>
<td>0.037 (0.215)</td>
<td>0.044 (0.189)</td>
</tr>
<tr>
<td>Completeness</td>
<td>100% (100%)</td>
<td>100% (100%)</td>
<td>99.9% (99.9%)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>16.4 (16.9)</td>
<td>15.3 (15.7)</td>
<td>13.7 (14.1)</td>
</tr>
<tr>
<td>Number of reflections</td>
<td>46,368</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution range</td>
<td>30–3.0 Å</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optical resolution</td>
<td>2.2 Å</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atoms</td>
<td>11,507</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond restraints</td>
<td>11,761</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angle restraints</td>
<td>15,961</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torsion restraints</td>
<td>8,025</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCS position restraints</td>
<td>10,616</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCS thermal restraints</td>
<td>10,616</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>0.25 (0.29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt;</td>
<td>0.29 (0.34)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOM</td>
<td>0.78 (0.77)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rmsd values</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonds</td>
<td>0.011 Å</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angles</td>
<td>0.953&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torsions</td>
<td>4.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most favored</td>
<td>93.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allowed</td>
<td>6.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generously allowed</td>
<td>0.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disallowed</td>
<td>0.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average B factor</td>
<td>101 Å&lt;sup&gt;2&lt;/sup&gt; (75 Å&lt;sup&gt;2&lt;/sup&gt;)&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPI&lt;sub&gt;free&lt;/sub&gt;</td>
<td>0.45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Values in parentheses are for the highest resolution shell unless stated otherwise.
<sup>b</sup>Values in parentheses are for dimer A-B.
Differences in structure among the eight monomers can be correlated with differences in crystal contacts and will not be considered further. The electron density was much better defined for two of the dimers (corresponding to chains A-B and E-F, with the dimer A-B being the best of all) than for the remaining two, especially the dimer G-H, which had several regions for which the density was too weak for model building. Even within the dimer A-B, monomer A lacked well-defined density for residues 112–115, whereas in monomer B, the corresponding residues were apparently stabilized by crystallographic contacts with the same region of a symmetry-related monomer. Electron density from the active site region of one of the monomers is shown in Figure 2. The remainder of this discussion will focus on chain B unless otherwise noted.

The structure of FeeM is shown in Figure 3. The center of FeeM is a β sheet with a cleft formed between two diverging β strands. One face of the β sheet, hereafter referred to as the concave face, wraps around the long amphipathic helix α4. The helix α5 is located on the same face and is directly adjacent to the entrance to the cleft, which forms an enclosed space in FeeM. The convex face is covered by a canopy formed from various secondary structures: the loop containing α1/α2, the loop containing the kinked helix α3, and the region containing the loop between β6 and β7. This canopy completely blocks solvent access to the active site cleft from the convex face. Helices α1 and α2 are antiparallel to one another and parallel to the sheet plane but roughly perpendicular to the β strands. The distorted helical structure in α3 appears to be partially stabilized by a three-way salt bridge formed by D67, D75, and R86. Electron density is clear for all three of these residues, and they are conserved among NASs related to FeeM (Figure 3C). Further possible stabilization of the distortion appears to come from packing of the side chain of the conserved Y78 against the curved C-terminal tip of β6 and a possible hydrogen bond from its side chain hydroxyl to the main chain carbonyl of A159 in β6. The turn connecting β6 and β7 at V166 forms a hydrogen bond with A33 in the turn connecting α1 and α2.

Product Binding and Substrate Recognition in FeeM

Based on structural precedent, the β cleft stood out even in low-resolution electron density maps as a likely active site, and further refinement of the phases revealed electron density within the cleft that resembled the product (Figure 2). This bound product was unexpected, as no product had been added to the crystallization solutions, and the product apparently copurified with FeeM and remained bound during crystallization. We modeled the density as N-lauroyl tyrosine (NLT) based on its apparent chain length. Unfortunately, we have not been able to independently verify the length of the fatty acid chain, as MALDI mass spectrometry of the protein solutions used for crystallization showed no evidence of a bound ligand (data not shown). The B factors for the bound product were similar to those of the surrounding atoms, which were among the lowest B factors in the structure.

NLT has three distinct regions: the aromatic ring of the tyrosine moiety, the portion of NLT near the amide bond, and the long chain fatty amide—each has extensive interactions with FeeM. The aromatic ring is located on the convex face where two of the three structural elements comprising the canopy of FeeM converge and forms both polar and nonpolar interactions with them. The location and interactions suggest that canopy formation and the binding of tyrosine are linked. Given the higher overall hydrophilicity of the convex face of the β sheet, it is possible that FeeM adopts an open conformation in which this face is exposed to solvent until binding of L-tyrosine promotes closure of the canopy. The aromatic hydroxyl group hydrogen bonds with the side chain amide nitrogen of N142 and the main chain carbonyl oxygen of P169—both of which are conserved. Nearly all of the other residues that contact the aromatic ring are absolutely conserved among related NASs as well (see Figure 3C), and the one exception, Y163, has minimal contacts in FeeM.

The amide bond portion of NLT, where catalysis occurs, interacts with FeeM via main chain hydrogen bonds. Hence, weak selective pressure for conservation is expected for this region of the NAS as the only apparent requirement is maintaining secondary structure. The sole exception is the absolutely conserved Y27 that hydrogen bonds with the product carboxylate through its side chain hydroxyl. The fatty amide penetrates the concave face of FeeM, normal to the plane of the β sheet. All interactions with FeeM are hydrophobic, and the residues involved are generally well conserved among NASs. One interesting exception is the residue pair V125/L126, which is located near the acyl chain terminus and is an A/Y pair in all other related NASs.
The overall conservation of residues forming contacts with the product exhibit what at first appears to be an unexpected pattern: the most highly conserved residues in the active site are involved in substrate recognition, while the residues where the actual enzymatic reaction takes place are more variable. If one takes into account the facility of the reaction catalyzed by FeeM—aminolysis of a thioester—this pattern makes sense. FeeM is built for recognition, not for speed. Bringing the correct substrates together in a productive configuration is a tougher challenge than increasing the rate of an already favorable reaction. Nonetheless, there

Figure 3. Overview of the Structure of FeeM

(A) View of FeeM looking onto the convex face of the β sheet of monomer B. The cartoon diagram is colored with a rainbow gradient from the N terminus (blue) to the C terminus (red). Residue numbers are shown at the Cα locations. The cocrystallized product is shown as ball and stick.

(B) View of the molecule after rotation by 90° as shown. The bottom face of the β sheet is the concave face, and the top face is the convex face as referred to in the text. The N and C termini are circled.

(C) The sequence and secondary structure of FeeM. The secondary sequence diagram is colored with the same gradient as A and B above. Also shown is a multiple sequence alignment of FeeM with various NASs derived from soil eDNA samples (CSL-3, CSL12.1, and CSL132), some putative NASs from the β-proteobacteria Nitrosomonas europaea (Chain et al., 2003) and Dechloromonas aromatica (unpublished sequence, NCBI accession number NZ_AADF0100001), and a structure-based alignment with AANAT (1CJW), Esal (IK4), and Lasi (IROS). Residues shaded yellow indicate similarity to a consensus sequence (not shown) with blue shading indicating identity with the consensus. Structure-based alignments were performed with INDONESIA (version April 15, 2004; http://alpha2.bmc.uu.se/dennis/). Shading with black indicates alignment of Cα positions to within 3.0 Å. Residues that contact the bound product are labeled according to the region of the product contacted: F, fatty amide; M, main chain; R, aromatic ring. The sequence alignments are truncated to regions similar to FeeM, although the sequences often extend at both the N and C termini.

The overall conservation of residues forming contacts with the product exhibit what at first appears to be an unexpected pattern: the most highly conserved residues in the active site are involved in substrate recognition, while the residues where the actual enzymatic reaction takes place are more variable. If one takes into account the facility of the reaction catalyzed by FeeM—aminolysis of a thioester—this pattern makes sense. FeeM is built for recognition, not for speed. Bringing the correct substrates together in a productive configuration is a tougher challenge than increasing the rate of an already favorable reaction. Nonetheless, there
Figure 4. Divergent Stereoview of the Superposition of the Active Sites of FeeM and AANAT

The backbone atoms and side chains believed to be important for catalysis in AANAT are shown in green. The amide portion (i.e., the bond formed during catalysis) of the product analog bound to AANAT is shown as a ball-and-stick model in green. The “proton wire” is shown in red. FeeM is shown in cyan with the amide portion of N-lauroyl tyrosine shown as a ball and stick in dark gray.

are aspects of FeeM’s structure that do facilitate the minimal catalysis required.

Comparison of FeeM to the GCN5-Related N-Acyl Transferases and the Acyl Homoserine Lactone Synthases: Implications for Catalysis and Function

Structure-based homology searches using the DALI server (Holm and Sander, 1993) indicated significant similarity between FeeM and members of the functionally related GCN5-related N-acyl transferases (GNATs). Members of this diverse family of proteins catalyze the acylation of nitrogen atoms with acetyl-coenzyme A as an acyl donor. Structural features that FeeM has in common with the GNATs include: the central β sheet; the active site cleft in parallel β strands; helix α4 along the bottom of the β sheet; and helix α1, which is present in most, but not all, GNATs.

The acyl homoserine lactone (AHL) synthases are a family of proteins that produce mediators of bacterial quorum sensing, and AHL synthases are also structurally similar to the GNATs. Unlike the GNATs but like the NASs, the AHL synthases utilize ACP rather than coenzyme A as the acyl donor for their amine substrates. To date, the structures of two AHL synthases have been determined by X-ray crystallography: Esal (Watson et al., 2002) and Las1 (Gould et al., 2004). Both enzymes are nearly identical to FeeM in their tertiary structure, with most differences being located on the convex face. In Esal the region corresponding to α1–α2 is not well defined. For Las1, α1 is similar to FeeM, while the remainder of the loop is different. All three enzymes contain β bulges at the active site cleft with two adjacent main chain amide protons pointing toward the cleft. There are no structures for AHL synthases with bound substrate, product, or structural analogs.

The structure of aryalkylamine-N-acetyltransferase (AANAT), a GNAT for which much is known about catalysis, in complex with a bisubstrate analog has been solved to 1.8 Å (PDB code 1CJW) (Hickman et al., 1999). Monomer B of FeeM superimposes with 123 residues in 1CJW with an rmsd of Cα positions of 1.56 Å. A superposition of the active site and side chains of catalytically important residues in AANAT that sheds light on the structural basis for catalysis is shown in Figure 4. Catalysis requires two features: deprotonation of the amine and stabilization of the tetrahedral oxyanion reaction intermediate. Deprotonation may be facilitated through E94, which is situated similarly to H120 in AANAT and therefore in contact with the “proton wire” — a set of crystallographically observed waters that provide a path for moving a proton from the substrate amine to solvent. Although our diffraction data were not sufficient to allow determination of water positions, the groups that should coordinate the two waters of the proton wire nearest the amine are situated in FeeM in a manner consistent with the existence of a proton wire. In GCN5, an analogously located glutamate is essential for catalysis (Tanner et al., 1999). Stabilization of the oxyanion intermediate in FeeM is likely promoted by both polar and steric interactions. Oxyanion stabilization by hydrogen bonding can be achieved through the amide hydrogens of the β bulge residues V96 and Q97, a common feature in GNATs. Optimal positioning of the thiester group for nucleophilic attack and oxyanion stabilization might be facilitated by Y149, which is oriented similarly to Y168 in AANAT, a residue required for catalysis, likely because of its steric and possibly polar interactions with the acyl donor sulfur atom (Hickman et al., 1999). For some GNATs, the significance of most conserved residues in the active site lies principally in orienting the substrates (Draker and Wright, 2004). For other GNATs, conserved residues participate directly in proton transfers (Tanner et al., 1999; Farazi et al., 2001; Scheibner et al., 2002). A hypothetical scheme for catalysis can be constructed based on similarities between FeeM and the GNATs and what is known about catalysis in the latter (see Figure 5).

Effects of Various Mutations on Interaction between FeeM and FeeL

FeeM joins tyrosine to an ACP-bound fatty acid. The tyrosine is recognized by FeeM, and biosynthetic fidelity for the fatty acid bound to FeeL must be mediated by protein-protein interactions between FeeM and FeeL.
To identify such potential interactions, the proximity of residues on the surface of FeeM to the tunnel leading to the active site was used in conjunction with the known sequence similarities between FeeM and similar NASs to design mutants of FeeM with compromised ability to bind FeeL. Binding was measured by a real-time assay that allows the determination of a pseudo-Michaelis-Menten kinetic binding constant of FeeM for FeeL in the presence of saturating amounts of L-tyrosine.

The surface adjacent to the active site tunnel entrance can be divided into three main regions on FeeL likely to be involved in interactions with FeeL (see Figure 6). The first region encompasses the loop region between β4 and β5, which normally serves as the diphosphate binding region in GNATs. In general, basic residues were considered potentially important based on the overall acidity of ACPs and earlier reports indicating that electrostatic interactions are important in interactions between ACPs and proteins that act on them (Gould et al., 2004; Parris et al., 2000). The mutants generated for this region are H102A and K112A. However, this region is not thought to be important since this loop is missing from many of the NASs similar to FeeM (Figure 3C). The second region is a spatial region that includes parts of α1 and the loop connecting β5 to α5 near the entrance to the product binding tunnel of FeeM. The mutants for this region are Q30R, K144A, and K144E. Q30 was selected because in all other related NASs the corresponding residue is a Lys or an Arg. The Q30R mutant could serve to increase the affinity of FeeM for FeeL. The third region is α5, which is covered by the mutants H145E, H145N, H145R, F148A, F148D, L151A, and L151E. The residues F148 and L151 were chosen because of their potential to participate in hydrophobic interactions with FeeL.

The results of the activity studies are shown in Table 2. None of the mutants tested involving the first region led to significant differences in $K_M$. Likewise, none of the mutants in the second region led to differences in $K_M$. Mutations in the third region produced much larger effects. Almost no effect was observed for mutations involving L151. The mutant F148A also had little to no effect. By contrast, the more radical mutant F148D showed a greatly decreased enzyme rate compared to native FeeM. Rate measurements at FeeL concentrations over ten times the apparent $K_M$ of native FeeM produced rates two orders of magnitude lower than native FeeM. The enzymatic rate dropped by more than 50% when the FeeL concentration was decreased by a factor of 2.5, indicating that the drop in observed rate was caused, at least in part, by a decreased affinity between FeeM and FeeL and cannot be wholly accounted for by misfolding of the mutant. Similarly, large decreases in rate were observed for the mutants H145E, H145N, and H145R, indicating that binding between FeeM and FeeL is quite sensitive to perturbations at this position.

Implications for Binding Interactions between FeeM and FeeL

Some insights into the basis for our mutation results can be gained by considering the observed mode of phosphopantetheine binding seen in most GNAT complex structures. In all cases so far, the phosphopantetheine cofactor binds as a β sheet mimic alongside the equivalent strand of β4 in FeeM. The amide groups of the cofactor form hydrogen bond interactions that a β strand would typically make. The conformation of FeeM is very similar to the GNATs in this region, indicating that this mode of binding is probably also in effect for the phosphopantetheine portion of FeeL. If this is the case, then the phosphopantetheine arm would be expected to fill the tunnel leading to the active site, and the surface surrounding the tunnel entrance would be the most likely place for FeeM and FeeL to interact.
Our results indicate that z5 is a crucial site of interaction between FeeM and FeeL (Figure 3B). All of the other tested mutants had negligible effects on the enzymatic parameters. Helix z5 is located next to the entry to the active site tunnel, though neither of the most sensitive residues contact the product directly. Of the two sensitive residues, H145 is by far the least tolerant of substitution. Sharp decreases in activity were seen upon replacement with acidic, basic, and small polar residues. The absolute conservation of H145 among related NASs and the ability of histidine to both hydrogen bond and provide hydrophobic interactions argue that histidine is uniquely suited to this site. The second residue, F148, is tolerant of substitution by Ala, but not by Asp, suggesting that this position may be buried via hydrophobic interactions in the FeeM/FeeL complex.

Table 2. Effects of Mutagenesis on Interaction between FeeM and FeeL

<table>
<thead>
<tr>
<th>Mutant</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_m$ (μM)</th>
<th>$ν$ (s$^{-1}$)$^a$</th>
<th>[FeeL] (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H102A</td>
<td>0.59 ± 0.02</td>
<td>10.1 ± 0.1</td>
<td>2.3 ± 0.1 × 10$^{-3}$</td>
<td>150</td>
</tr>
<tr>
<td>K112A</td>
<td>0.56 ± 0.01</td>
<td>13.0 ± 0.8</td>
<td>7 ± 1 × 10$^{-4}$</td>
<td>77.6</td>
</tr>
<tr>
<td>Region 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q30R</td>
<td>0.58 ± 0.02</td>
<td>8.3 ± 0.9</td>
<td>4.3 ± 0.2 × 10$^{-3}$</td>
<td>150</td>
</tr>
<tr>
<td>K144A</td>
<td>0.247 ± 0.006</td>
<td>8.2 ± 0.8</td>
<td>2.2 ± 0.2 × 10$^{-3}$</td>
<td>77.6</td>
</tr>
<tr>
<td>K144E</td>
<td>0.60 ± 0.01</td>
<td>5.8 ± 0.5</td>
<td>1.19 ± 0.09 × 10$^{-3}$</td>
<td>150</td>
</tr>
<tr>
<td>Region 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F148D</td>
<td></td>
<td></td>
<td>8 ± 1 × 10$^{-4}$</td>
<td>77.6</td>
</tr>
<tr>
<td>Region 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F148A</td>
<td>0.53 ± 0.01</td>
<td>10.3 ± 0.9</td>
<td>1.39 ± 0.4 × 10$^{-3}$</td>
<td>150</td>
</tr>
<tr>
<td>L151A</td>
<td>0.74 ± 0.02</td>
<td>11.5 ± 0.8</td>
<td>6.4 ± 0.4 × 10$^{-4}$</td>
<td>77.6</td>
</tr>
<tr>
<td>L151E</td>
<td>0.88 ± 0.03</td>
<td>12 ± 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All assays used octanoyl-FeeL as a substrate.

$^a$The rate measured at each of two concentrations is shown.
Although L151 is reasonably well conserved among related NASs and is directly adjacent to H145/F148 (see Figure 6), all L151 mutants tested had wild-type activity. These results, combined with the patterns of sequence conservation near $\alpha_5$, indicate that interactions between FeeM and FeeL in this region may be limited to H145/F148.

The residues H145 and F148 apparently form a small hydrophobic pocket that may bind to FeeL. Although H145 is not strictly speaking a part of $\alpha_5$, it can be considered part of a slightly expanded turn that extends $\alpha_5$. In both the A and B protomers of FeeM, there appears to be something bound to the pocket, indicating that this pocket is well suited for binding ligands. In chain A, P115 binds to the pocket, while in chain B, there is some residual unmodeled density that resembles the product molecule in shape. There is evidence that a similar hydrophobic pocket is present in many GNATs that can bind the adenine base of coenzyme A (Clements et al., 1999), the gem-dimethyl group of the phosphopantetheine arm (Penneff et al., 2001), or general buffer components (Rojas et al., 1999). In principle, a similar hydrophobic pocket is possible in the AHL synthases Esal and Lasl but would require expansion of the top turn of the corresponding $\alpha$ helix. One attractive candidate residue on FeeL for binding in this pocket is I41, which is directly adjacent to the phosphopantetheinylated residue S40. Residues corresponding to D93, S40, and I41 in FeeL have all been shown to be important in complexes between ACPs and modifying enzymes (Parris et al., 2000).

The most surprising outcome of the mutation studies is the relative lack of importance of basic residues to the activity of FeeM. As was mentioned previously, ACPs are typically very acidic proteins. The X-ray structure of a complex between holo-ACP synthase from Bacillus subtilis in complex with its ACP indicated that substantial electrostatic interactions took place between the acidic helix $\alpha_3$ of ACP and the basic helix $\alpha_1$ of the synthase (Parris et al., 2000). Also, mutational studies with the AHL synthase Lasl, a protein very similar in structure to FeeM, was sensitive to mutations in a basic residue located on $\alpha_5$. It appears that FeeM differs from these other functionally and structurally related enzymes due to its close genetic association with FeeL, which has many fewer acidic residues along $\alpha_3$ than other ACPs (net charge of $-1$ in FeeL versus $-8$ in AcpP from E. coli). Perhaps mutual adaptation between FeeM and FeeL to enhance biosynthetic fidelity has eliminated the selective pressure for basicity to allow complex formation between FeeM and FeeL and replaced it with a more demanding shape complementarity. It is interesting to note that the NASs related to FeeM do have conserved basic residues in $\alpha_5$, although FeeM does not, and in these others no tightly linked ACP was in the gene clusters. The other family members may rely on the ACPs of fatty acid synthesis as substrates—as is presumed to be the case for the AHL synthases (Val and Cronan, 1998). Since the ACPs of fatty acid synthesis are essential proteins that need to serve as substrates for many different proteins, there is less opportunity for loss of acidic residues by mutation. This may make maintenance of basicity more crucial for the other NASs.

Perspectives on NAS Function

There have been frequent allusions in the preceding discussion to members of the AHL synthase family of proteins. These proteins synthesize small-molecule signals important for bacterial quorum sensing. Pathways containing AHL synthases have been shown to be associated with bioluminescence (Eberhard et al., 1981), production of virulence factors (Passador et al., 1993), biofilm formation (Davies et al., 1998), and the production of antibiotics (Stead et al., 1996) among other phenotypes. In addition to strong similarities in structure between FeeM and the AHL synthases Esal (Watson et al., 2002) and Lasl (Gould et al., 2004), there are other intriguing connections between the NASs and AHL synthases. Both NASs and the AHL synthases join an amino acid to a fatty acid by forming an amide bond. Both the NASs and the AHL synthases utilize ACPs as acyl donors, whereas most GNAT-like proteins use coenzyme A derivatives as acyl donors. These similarities in structure and substrate could indicate a functional link between the products of the respective enzymatic families, and the long chain N-acyl amino acids and their downstream metabolites may function in bacterial signaling pathways in the organisms that produce them.

Although the immediate products of FeeM have antibiotic activity against B. subtilis, the products that arise by subsequent reaction along the fee pathway do not (S.F. Brady, personal communication). A recent study by Davies et al. found that subinhibitory concentrations of antibiotic compounds led to patterns of gene induction in bacteria, and the patterns observed could not be adequately described as SOS responses to environmental stresses (Goh et al., 2002). These observations support the hypothesis that many antibiotics may serve as signaling molecules when present at low concentrations (Goh et al., 2002). In our study on FeeM, we see perhaps another link between signaling molecules and antibiotics. The fact that a small molecule of significant antibiotic activity can be an early intermediate in a pathway for which antibiotic activity does not appear to be the ultimate goal provides support for some sort of role in signaling. What would be useful in confirming such a role for the NAS products are either direct studies with cultivable hosts known to possess NAS genes or extensive in silico studies that examine the NAS genes in their genomic context to identify potential specific functions for the NASs and the pathways that contain them.

Experimental Procedures

Production of FeeM

The gene feeM was cloned from the original cosmid (CSLC-2) (Brady et al., 2002) into a pET42a-based vector (Novagen) modified by Joanne Widom (Cornell University Department of Chemistry and Chemical Biology) to produce a fusion protein containing an N-terminal maltose binding protein from the MalC-C2x vector (New England Biolabs) followed by a hexahistidine tag and thrombin cleavage site. The vector was transformed into BL21(DE3) cells (Novagen) that were grown in MM media (Pryor and Leiting, 1997) supplemented with 5 g/l casamino acids and 30 g/l kanamycin sulfate and subsequently in MM media containing 50 mg/ml L-selenomethionine (Guerrero et al., 2001), 400 $\mu$M isopropyl-$\beta$-D-thiogalactopyranoside (IPTG) for growth overnight at 20°C. The harvested cells were resuspended in ~50 ml of buffer A (50 mM 4-[(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 300 mM NaCl, 10 mM $\beta$-mercaptoethanol, 10% glycerol [pH 8.0]).
The cell suspension was sonicated on ice, and the clarified supernatant was subjected to affinity purification using Ni-NTA resin (Qiagen). The recovered protein was subjected to dialysis overnight against 1 liter of buffer A while simultaneously undergoing cleavage with 150 NIH units of bovine phosphoenolpyruvate-thrombin (Calbiochem) per liter culture at room temperature. FeeM was separated from the clipped fusion partner by treating the thrombin lysate with Ni-NTA resin and subsequently purifying FeeM with a Sephacryl S-100 column (Amersham Biosciences). The final buffer system contained 10 mM HEPES, 50 mM NaCl, 1 mM DL-dithiothreitol (DTT), 1 mM ethylenediaminetetraacetic acid (EDTA), 0.01% sodium azide (pH 8.0). The solution was concentrated to 20 mg/ml for crystallization with UltraFree 15 concentrators (Millipore).

Crystallization and X-Ray Diffraction Data Collection

The conditions that produced the crystal used for the structure determination were 4.06 M NaCl, 50 mM glycine (pH 10.0) by sitting drop vapor diffusion at room temperature. Crystals formed and grew over a 4 week period. The crystal was collected on a mounted cryoloop (Hampton Research) and immersed in 200 μl of cryoprotectant solution for 2–5 min. The cryoprotectant solution was prepared by dissolving 0.350 g of dextrose in 1.0 ml of 1 M NaCl and adding 1 μl of 1.0 M DTT. The crystal was then immersed in a liquid nitrogen bath and stored under liquid nitrogen until data collection.

Data were collected on the Northeastern Collaborative Access Team (NE-CAT) beamline 8BM at the Advanced Photon Source at Argonne National Laboratories. Data were collected at 100 K at the energies 12,660 (reflection), 13,160 (remote), and 12,662 (peak) keV. A φ rotation of 1° and a detector distance of 400 mm were used for data collection.

Reflection Data Processing

The CCP4 4.2.2 suite of programs (CCP4, 1994) was used for data processing. The diffraction spots were integrated, scaled, and truncated with the programs MOSFLM 6.2.2 (Leslie, 1994; Campbell, 1995), ODS (Stetter et al., 1997), SCALA 3.1.20 (Evans, 1997), and TRUNCATE 4.2 (French and Wilson, 1978). Only fully recorded reflections were used for scaling and merging. Reflection peaks were selected for Rfree validation (Brünger, 1992) (5% of total) in thin data shells with the program DATAMAN 6.2.8 (Kleywegt and Jones, 1996b) from the RAVE suite of programs.

Experimental Phasing and Model Refinement

Selenium sites were located and refined with the programs SnB 2.2 (Weeks and Miller, 1999), DREAM (Blessing and Smith, 1999), and SHARP 2.0.1 (de La Fortelle and Bricogne, 1997). A density modified map revealed clear elements of secondary structure and boundaries between protein and solvent. The noncrystallographic symmetry (NCS) operators were identified for the eight monomers in the asymmetric unit with the programs MAMM 6.0.4 (Kleywegt and Jones, 1994, 1999), MAPMAN 7.6 (Kleywegt and Jones, 1996b), NCSD, and IMP (Jones, 1992), iteratively with DM 2.1 (Cowtan, 1994). The final operators were used as input to the program DM, which was used to perform solvent flattening, histogram matching, phase extension, and NCS averaging on the data to improve the phases. Phase extension to 3.0 Å was done over 30 steps starting at 4.5 Å. The solvent content was set to 0.6.

The atomic model for monomer “A” was built with the baton, build facilities of the program O 3.0.11 (Jones et al., 1991). The additional models in the asymmetric unit were built with the NCS operators in conjunction with the program XPAND 1.4.4 from the XUTL suite of programs (Kleywegt et al., 2001). Model refinement was carried out using the program REFMAC 5.1.24 (Murshudov et al., 1997).

The optimal procedure for utilizing NCS restraints was found to be the use of medium restraints for residues 17–54, 60–102, and 121–186 of all monomers. Medium NCS restraints were applied groupwise for residues 1–16, 55–59, 103–120, and 187–193 of monomers A, D, E, and H and for monomers B, C, F, and G. TLS groups (Schomaker and Trueblood, 1968) were defined to encompass each of the monomers, and TLS parameters were refined. The use of TLS was critical for obtaining a good fit between the model and scattering factors. Rebuilding was accomplished with the program OOPS2 (Kleywegt and Jones, 1996a) in conjunction with O. After several rounds of rebuilding, a model was added for NLT into residual density that had been apparent from the initial experimental map. Additional tools utilized in improving the stereochemical quality of the model include the Penumeter Rotoram Library (Lovell et al., 2000); Probe, Reduce, and Flipkin (Word et al., 1999a, 1999b); and Mage (Richardson and Richardson, 1992). The program PyMOL 0.88 (Dyalo, 2002) was used for figure preparation.

Production and Purification of FeeM Mutants

The desired mutations in the feeM gene were made directly in the construct described above with the QuickChange II (Stratagene) Site-Directed Mutagenesis Kit. The primers used for mutagenesis are described in the Supplemental Data. The mutant proteins were produced by overexpression in BL21 (DE3) cells (Novagen) as described above for FeeM, except that expression was performed directly in MM media supplemented with casamino acids. Likewise, metal immobilization affinity chromatography, thrombin cleavage, and size exclusion chromatography were carried out for each mutant protein by a procedure essentially identical to that described for FeeM above. None of the purification or storage solutions for the mutants contained added reducing agents. The purified mutants were concentrated to 20 mg/ml as described for FeeM and were flash frozen in liquid nitrogen and stored at −80 °C.

Determination of Binding of FeeL and FeeL Mutants

The preparation of octanoyl-holo-FeeL was achieved with apo-FeeL, octanoyl-coenzyme A, and the phosphopantetheinyl transferase Sfp (Quad et al., 1998). The procedure used will be described in a separate manuscript. Enzymatic rate was measured by a spectrophotometric assay that detects the thiol released upon thioester aminolysis via its reaction with 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB). The assay is very similar to one previously used for AANAT (De Angelis et al., 1998) except that our assay used real-time measurements of the absorbance at 412 nm, whereas the previous study utilized a discontinuous assay. Each 60 μl reaction contained 100 mM HEPES, 50 mM NaCl, 10 mM EDTA, 167 μM DTNB, and 533 mM L-tyrosine. Previous assays indicated FeeM has a pseudo-Km of ~80 μM for L-tyrosine (data not shown). Each reaction contained varying amounts of FeeL. Each reaction was initiated by addition of 3.00 μl of mutant protein solution, the concentration of which depended on the level of enzymatic activity present. The reactions were then immediately placed in a spectrophotometer at ambient temperature (~75 F–85 F) in which the absorbance at 412 nm was measured every 10 s for a total of 60 s.

The rate for each reaction was determined by least squares linear regression of the observed absorbance versus time data. Rate measurements were made at five or six different concentrations of FeeL for the mutant proteins that have reported kinetic parameters in Table 2. The concentrations spanned at least the range 0.1-4 μM to 5.0 × Kd. Five replicate measurements were made at each FeeL concentration. The rate measurements were converted to the rate of production of FeeL with an extinction coefficient of 14,000 for DTNB that had been determined by using a similar buffer with L-cysteine as a standard. The rates were also normalized for mutant protein concentration. The complete data for each mutant protein were fit to the Michaelis-Menten kinetic model by nonlinear least squares optimization. For several mutants, the apparent Kd would have required very high concentrations of FeeL for full determination. In these cases, we measured the rates at FeeL concentrations of 150 μM and 62.6 μM with five replicates each.

Supplemental Data

Supplemental Data include graphs of the nonlinear regressions of enzymatic rate data fit to a Michaelis-Menten curve, a list of primer sequences used for cloning and mutagenesis work, and a paragraph discussing dimer contacts and are available at http://www. structure.org/cgi/content/full/14/9/1425/DC1/.

Acknowledgments

We are grateful for the help of Joanne Widom during the expression phase of this project for her advice and constructs. We also extend appreciation to Craig Ogata for help during the collection of diffraction data at APS. We acknowledge the support of a postdoctoral
fellowship provided by the National Institutes of Health to R.M.V. (NRSA, GM065017) and research support from National Institutes of Health CA59021 (J.C.). This work is based upon research conducted at the Northeastern Collaborative Access Team (NE-CAT) beamlines of the Advanced Photon Source, supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under contract number W-31-109-ENG-38.

Received: March 15, 2006
Revised: July 9, 2006
Accepted: July 16, 2006
Published: September 12, 2006

References


Accession Numbers

The coordinates and structure factors for the structure of FeeM have been deposited into the Protein Data Bank under the ID Code 2G0B.