Drosophila melanogaster nonribosomal peptide synthetase Ebony encodes an atypical condensation domain

Thierry Izoré[a,b,c,1], Julien Tailhades[a,b,c], Mathias Henning Hansen[a,b,c], Joe A. Kaczmarski[d], Colin J. Jackson[d], and Max J. Cryle[a,b,c,1]

[a]The Monash Biomedicine Discovery Institute, Monash University, Clayton, VIC 3800, Australia; [b]Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia; [c]EMBL Australia, Monash University, Clayton, VIC 3800, Australia; and [d]Research School of Chemistry, The Australian National University, Acton, ACT 2601, Australia

The protein Ebony from *Drosophila melanogaster* plays a central role in the regulation of histamine and dopamine in various tissues through condensation of these amines with β-alanine. Ebony is a rare example of a nonribosomal peptide synthetase (NRPS) from a higher eukaryote and contains a C-terminal sequence that does not correspond to any previously characterized NRPS domain. We have structurally characterized this C-terminal domain and have discovered that it adopts the aryl-alkylamine-*N*-acetyl transferase (AANAT) fold, which is unprecedented in NRPS biology. Through analysis of ligand-bound structures, activity assays, and binding measurements, we have determined how this atypical condensation domain is able to provide selectivity for both the carrier protein-bound amino acid and the amine substrates, a situation that remains unclear for standard condensation domains identified to date from NRPS assembly lines. These results demonstrate that the C terminus of Ebony encodes a eukaryotic example of an alternative type of NRPS condensation domain; they also illustrate how the catalytic components of such assembly lines are significantly more diverse than a minimal set of conserved functional domains.

nonribosomal peptide synthetase | NRPS | condensation reaction | C domain | aryl-alkylamine *N*-acetyl transferase

Nonribosomal peptides (NRPs) are secondary metabolites synthesized in many organisms, and to which they usually confer a significant fitness advantage. The diversity of NRPs stems from the structure of the nonribosomal peptide synthetase (NRPS) assembly lines that produce them (1, 2). NRPS systems are common in bacteria and fungi, where the products they synthesize include antibiotics, siderophore-sensing and bacterial quorum-sensing regulators, toxins, and even compounds used as anticancer agents and immunosuppressants (1). As these peptides are synthesized independently from the ribosome, this eliminates the requirement to utilize the "standard" proteinogenic pool of amino acids, and to date over 500 different monomers have been identified as being incorporated by NRPS machinery (1). However, further modification of the peptides is made possible via the mechanism of synthesis employed by typical NRPS assembly lines, which relies on a repeating modular architecture built from varying catalytic domains (1, 2). Each module is responsible for the incorporation of a single amino acid into the growing peptide; the core domains required for a minimal peptide extension module are an adenylation domain (A domain), a peptidyl carrier protein domain (PCP domain), and a condensation domain (C domain) (1). Catalytic activity begins with the A domain selecting the desired monomer, which is then activated using ATP before attachment to the phosphopantetheine moiety of the neighboring PCP domain. As a thioester, this residue is then delivered to the C domain, where a peptide bond is formed between the upstream PCP-bound

peptide and the PCP-bound amino acid. This then leads to the transfer of the peptide from the upstream PCP to the downstream PCP and extends the peptide by one residue (3). In addition to these essential domains, most NRPSs encode a terminal thioesterase domain (TE domain) in the last module of the assembly line to allow the release of the product from the synthesis machinery. The TE domain also serves as a point for further diversification of the peptide sequence through various pathways, which include macrocyclization or dimerization in addition to hydrolysis (1). Within modules, NRPSs may also contain additional domains acting in *cis* (such as epimerization domains) or external enzymes acting in *trans* that alter the structure of the NRP during synthesis (1, 2). This versatility provides a vast number of possible combinations within the NRPS assembly line and an even greater number of possible NRP products.

While bacterial and fungal NRPS machineries are common and largely conform to a standard architecture and domain arrangement, the presence of functional NRPS assembly lines in higher eukaryotes is rare (4). The NRPS-like proteins that have been identified in the latter group mostly consist of A and PCP domains

Significance

Nonribosomal peptide synthesis is responsible for the formation of many important peptide natural products in bacteria and fungi; it typically utilizes a modular architecture of repeating catalytic domains to produce these diverse peptide structures. The protein Ebony from *Drosophila melanogaster* is a rare example of such a nonribosomal peptide synthetase from a higher eukaryote, where it plays a central role in the regulation of amine neurotransmitters. Here, we reveal that the C-terminal portion of Ebony encodes an atypical peptide bond-forming nonribosomal peptide synthesis domain. Structural analysis shows that this domain adopts a fold not predicted by its primary sequence, and indicates how this domain maintains its high degree of substrate specificity.

followed by sequences usually not found in archetypical NRPS assembly lines (4). Examples of such NRPS enzymes are mostly those involved in mammalian lysine metabolism (5): AASDH, a 2-aminoadipic 6-semialdehyde dehydrogenase harboring an unusual A-PCP-PQQ arrangement (where PQQ represents a sequence containing seven binding motifs for pyrroloquinoline quinone); and Lys2 (6), an α-aminoadipate reductase with an A-PCP-NADPH–binding domain architecture. In addition to these well-characterized mammalian enzymes, Ebony from *D. melanogaster* contains an A-PCP di-domain followed by an uncharacterized sequence with no sequence homology to any known proteins based on standard search using domain prediction servers (Fig. 1). Ebony is an 879-residue protein (98.5 kDa) expressed in both glial and cuticular cells (7, 8). In glial cells, Ebony is involved in histamine regulation (the main neurotransmitter in the optical nerve system) and plays an essential role in neurotransmitter inactivation through conversion to carcinine [β-alanyl-histamine; Fig. 1 (9)]. Similarly, in cuticular cells Ebony catalyzes the condensation of β-alanine with dopamine to form β-alanyl-dopamine, a metabolite involved in the pigmentation and sclerotization of the insect cuticle. Mutants display strong phenotypes, with alteration of vision (10), circadian regulation of locomotor activity (11), and cuticle sclerotization in affected flies.

Ebony is an unusually fast NRPS enzyme (12), which can achieve a condensation reaction up to 60,000 times faster than the archetypical NRPS tyrocidine synthetase. While the A domain of Ebony is specific for β-alanine (13), the C-terminal domain appears to be versatile and can use a wide range of amines containing a planar ring. The C-terminal sequence of Ebony is thus of great interest among condensation-type domains given that this region seems to encode a type of NRPS condensation domain that is able to perform both the selection of an amine moiety (dopamine/histamine) and the condensation of these residues with β-alanine via an amide bond. Given that the C-terminal domain of Ebony appears to represent a previously unknown example of an NRPS condensation domain and displays intriguing catalytic properties, we sought to structurally characterize this eukaryotic NRPS domain and investigate its biochemical properties. To this end, we solved the crystal structure of the Ebony C-terminal domain both in its apo form and in complex with the amine substrates dopamine and histamine along with the resultant products β-alanyl-dopamine and carcinine (β-alanyl-histamine). Our results demonstrate that the Ebony C domain, unlike standard NRPS C domains [e.g., VibH (14); Fig. 1], unexpectedly adopts the aryl-alkylamine-*N*-acetyl transferase (AANAT) fold that was not directly apparent from standard sequence homology searches, and provides an understanding of the mechanism of selectivity of this condensation domain

Fig. 1. Structure and activity of a typical NRPS condensation domain and the protein Ebony. (*A*) Crystal structure of VibH, a typical NRPS condensation domain, together with the reaction it catalyzes. (*B*) Ebony NRPS (*Left*); note the unknown C-terminal domain and the reaction it catalyzes (*Right*).

for both the PCP-bound amino acid and aromatic amine substrates that fits the biological functions of Ebony.

Results

As the condensation function of Ebony appeared to rely on the unusual C-terminal portion of this enzyme, we concentrated on the characterization of this atypical condensation-like domain. Previous work has shown that Ebony is highly prone to degradation during *E. coli* expression, and we used this fact to our advantage to identify an optimal C-terminal construct based on proteolysis of the full-length protein during overexpression. This region, encompassing the residues from Leu666 to the C-terminal residue Lys879 (referred to here as C_N, for the amine-selecting C domain; Fig. 2), was well behaved and highly soluble when expressed with a C-terminal 6xHis tag. With the ability to access significant amounts of highly pure protein, we then turned to the structural characterization of this domain.

Structural Characterization and Substrate Binding of the Ebony Condensation-Like Domain. The optimized Ebony condensation-like domain (C_N) crystallized readily, forming numerous needle clusters in a wide range of conditions. To obtain diffraction-quality crystals, several rounds of condition optimization combined with crystal seeding were required, which yielded crystals in space group P2 (1) that diffracted to 2.0 Å. A molecular replacement model seeded from the *Paramecium bursaria* chlorella virus polyamine acetyltransferase (15) was successfully generated by the Robetta server (16), which makes use of both ab initio modeling and homology search routines. This proved necessary after crystallization of selenomethionine-labeled C_N protein as well as molecular replacement using potential homologs identified by Phyre2 (17) had failed. The density map thus obtained allowed us to build a model of C_N with high confidence and optimal geometry (*SI Appendix*, Table S1). The overall fold of the C_N domain is highly reminiscent of members of the AANAT family (18) despite a very low level of sequence identity (<20%, *SI Appendix*, Fig. S1); the closest structural homolog identified was the chlorella virus polyamine acetyltransferase (15), which aligned with an overall rmsd of 2.1 Å over 145 Cα atoms. The Ebony C_N domain is composed of seven α-helices arranged around a central β-sheet, with the 7-stranded β-sheet splayed between strands 4 and 5 (Fig. 2). While this structural feature in the AANAT superfamily serves to create a binding pocket for the pantetheine moiety of the acetyl-CoA (substrate) (18) (Fig. 2), in the case of Ebony it is instead required to accommodate the β-alanine–loaded phosphopantetheine arm from the upstream donor PCP domain. Although the structure of C_N is clearly highly divergent from the typical NRPS condensation domain—which is essentially a dimer of two chloramphenicol acetyltransferase (CAT) domains (Fig. 1)—the general feature of the binding site between two highly symmetrical half-domains is maintained in this C-type domain.

The two closest structural homologs of the Ebony C_N domain in *Drosophila* are the dopamine (19) (Fig. 2) and agmatine (20) AANATs (*SI Appendix*, Fig. S1), which display rmsd values of 2.4 and 2.5 Å despite the low sequence identities of these enzymes (16 and 12%, respectively). While the core fold of these enzymes is well conserved between C_N and the members of the AANAT family, there are substantial differences in the loop regions. The most striking difference is a mobile loop connecting β-strand 3 and α-helix 4 in Ebony, which is folded into a helix of different lengths in the structures of AANAT enzymes (*SI Appendix*, Fig. S2). This region is located on the opposite side of the protein to the donor PCP substrate-binding site and forms part of a highly acidic surface composed of residues Glu748, Glu750, Asp752, Glu762, Glu765, and Glu768 (Fig. 2). This acidic (and hence negatively charged) region is centered on an active site access channel anticipated to be involved in binding the positively
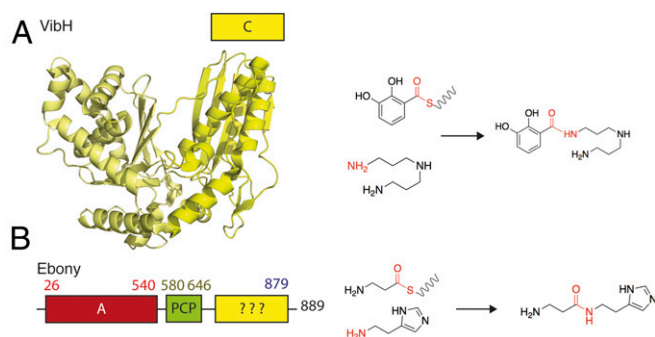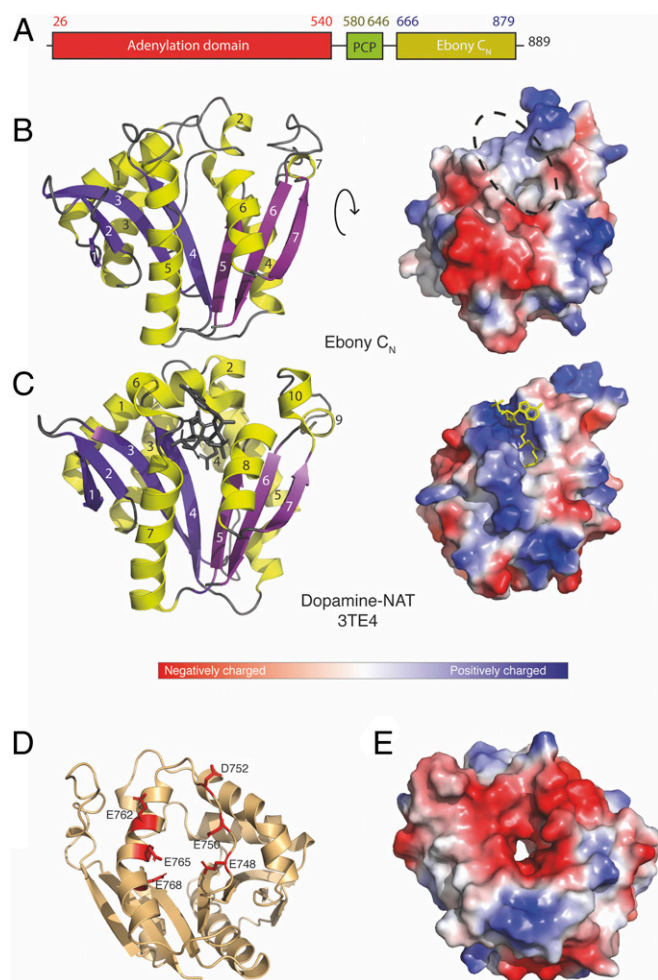
Fig. 2. Crystal structure of Ebony C_N and comparison with *D. melanogaster* dopamine NAT. (*A*) Primary architecture of Ebony. (*B*) Crystal structure of Ebony C_N shown as a cartoon (*Left*) and charge-colored surface (*Right*); note the splaying in the central β-sheet and the flat, hydrophobic surface in Ebony C_N compared with the CoA binding site in the dopamine-NAT homolog. (*C*) Crystal structure of dopamine-NAT shown as a cartoon (*Left*) and charge-colored surface (*Right*); β-strands are in white; helices, in black. (*D*) Localization of residues in the charged acidic region shown as a cartoon; and (*E*) overall negative surface charge as visualized in a charge-colored surface.

charged amine substrates utilized by Ebony (Fig. 2). This hypothesis is supported by the related AANAT enzymes displaying a similar charged region and utilizing amine substrates comparable to Ebony. The alteration of a secondary structure within this region to a flexible loop presents one possible route for accelerating the access of amine substrates to the C_N active site.

On the opposite side of the domain, the binding site for the aminoacyl-PCP domain is a relatively flat and hydrophobic surface, which stands in contrast to the highly positively charged "cradle" required to accommodate the phosphate groups of the CoA substrates in AANAT enzymes (Fig. 2 and *SI Appendix*, Fig. S1). The nature of this putative PCP interaction interface is in agreement with the vast majority of such PCP interaction interfaces identified within NRPS machineries to date [i.e., mainly hydrophobic (2)], likely because of the role played by the PCP-bound prosthetic linker in accessing the active sites of various NRPS domains. Replacement of a CoA substrate with a PCP-bound substrate in the case of Ebony also implies that binding of the thioester tethered β-alanyl substrate to C_N is controlled by

the activity of the upstream A domain through the hydrolysis of ATP during amino acid activation—a mechanism known as the A-domain alternation cycle (21, 22). The rapid rate of activity reported for Ebony appears to be governed by the activity of the upstream A domain, with the C_N domain also then able to catalyze peptide bond formation at a rate significantly faster than typical NRPS C domains (12). From isothermal calorimetry (ITC) measurements (*SI Appendix*, Table S2 and Fig. S7), Ebony C_N has a dissociation constant (K_d) for dopamine of ~30 μM (K_d of ~60 μM for the product β-alanyl-dopamine), which is consistent with processing under steady-state conditions given the concentration range of dopamine in cuticular cells (23). In contrast, Ebony C_N has a substantially lower affinity for histamine (K_d ~600 μM), which again is consistent with the millimolar concentrations of histamine released during neurotransmission in the brain and optic lobes [670 mM in synaptic vesicles (24)]. Interestingly, Ebony C_N has a slightly higher affinity for carcinine (product; K_d ~220 μM) than for the substrate, histamine. Given the physiological role of histamine as a neurotransmitter that is released in "bursts," this allows product concentration to regulate the activity of Ebony (24, 25), with such product inhibition also observed in other enzymes that modulate neurotransmitter levels [such as acetylcholinesterase (26)]. Thus, these affinity measurements are consistent with the physiological roles of the substrate molecules.

Substrate- and Product-Bound States of the Ebony C_N Domain. To gain insight into how the Ebony C_N domain functions to generate peptide bonds between PCP-bound β-alanine and dopamine/histamine, we determined four additional cocrystal structures of this domain in either substrate-bound (histamine and dopamine) or product-bound [carcinine (β-alanyl-histamine) and β-alanyl-dopamine] states; the relatively weak binding of histamine required higher soaking concentrations, consistent with the ITC measurements. All complexes produced clearly defined electron density for the additional ligands, which were easily identified in the C_N catalytic channel between β-strands 4 and 5 (see difference density maps and polder validation maps; *SI Appendix*, Fig. S3). The aromatic rings present in both the substrates and the β-alanine–conjugated products dock into a perfectly tailored hydrophobic cage consisting of residues Phe689, Val760, Phe761, and Leu764, which serve to trap the substrate (Fig. 3 *A* and *B*). To assess the importance of C_N residues in amine binding, we prepared F689A and F761A mutants within the hydrophobic cage as well as the E696L mutant of the central coordinating glutamate residue. In C_N condensation assays (assessing the formation of β-alanyl-dopamine; see below), mutant E696L showed no enzymatic activity, indicating the importance of this residue for coordination of the aromatic moiety of the amine acceptor. Of the two phenylalanine mutants, F689A showed only ~5% of the activity of the wild-type enzyme, while F761A retained almost 50% activity (Fig. 4). These results directly correlate with the relative distance of the phenylalanine rings from the amine substrate, with F689 closer (~3.8 Å) than F761 (~4.2 Å) (Fig. 3). Such residues are also conserved in members of the AANAT superfamily (*SI Appendix*, Fig. S1), which implies a general role for them in binding the amine substrates for these enzymes.

Inside the hydrophobic cage, the catechol moiety of dopamine is coordinated by residue Glu696, which hydrogen-bonds to both hydroxyl groups (2.9 and 3.1 Å; Fig. 3). In the histamine-bound structure, the amine hydrogen on the histidine ring is coordinated via a single interaction with Glu696 (3.0 Å; Fig. 3), which, combined with the increased distance to F761 (5.5 vs. 4.2 Å), explains the lower affinity of C_N for histamine. Although typical NRPS condensation domains utilize a conserved active site histidine residue [albeit one whose role is somewhat unclear (3)], no such direct interaction is present in C_N. A superposition of the ligand-bound structures shows that the positioning of the
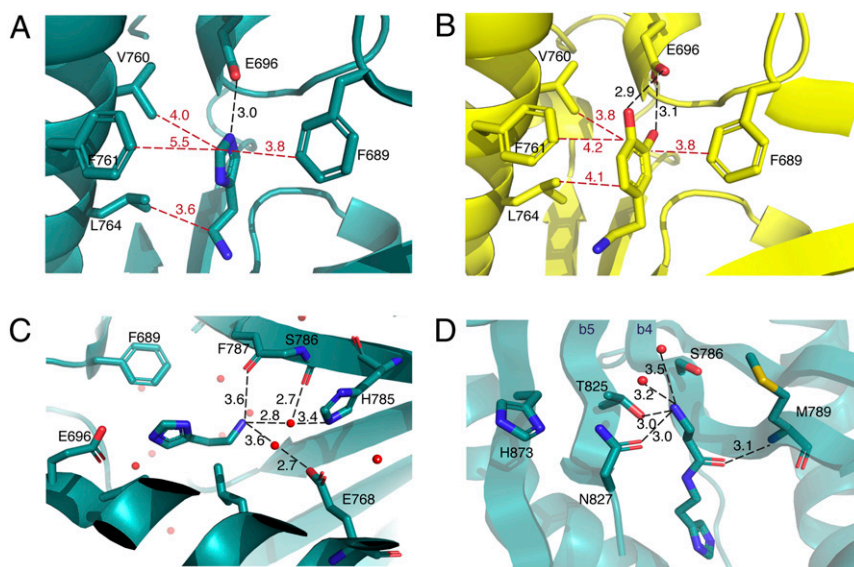
Fig. 3. Structures of substrate- and product-bound Ebony C_N and residues involved in orientation of the substrates. (*A* and *B*) Cartoon representation of histamine-bound (*A*) and dopamine-bound (*B*) Ebony C_N, with residues forming the "hydrophobic cage." Hydrophobic interactions are in red; hydrogen-bonding interactions, in black. (*C*) Cartoon representation of residues playing a role in positioning the amine substrate (histamine-bound C_N structure shown). These include the backbone carbonyl groups of F787 and S786 together with the side chains of residues H785 and E768 (T825 omitted for clarity). (*D*) Cartoon representation of the hydrophilic substrate-binding channel (carcinine-bound C_N structure shown), indicating how hydrogen-bonding interactions stabilize the amine in the β-alanine moiety of carcinine. The proximity of the N827 side chain relative to carcinine explains why compound **6** is not accepted as a substrate but its enantiomer (**7**) is (due to the cavity between carcinine and the M789 side chain accommodating the methyl branch of **7**).

aromatic ring of the amines is maintained—most likely due to the hydrophobic cage as discussed above—while there are much greater differences in the position of the aliphatic chain and the terminal amino group between the dopamine and the histamine molecules (Fig. 3). When ligand-bound C_N structures are compared with the dopamine AANAT in complex with acetyl-CoA [PDB ID 4TE3 (19)], it becomes clear that the histamine-bound structure represents the most likely substrate position for the condensation reaction. Indeed, the proximity of the reactive amine group and the angle of attack that this orientation provides are suitable for the reaction to take place (*SI Appendix*, Fig. S4). Also, the position of the β-alanyl moieties of the products found bound to C_N is consistent with attack of the amine (based on the histamine structure). Furthermore, the position of the peptide bond formed between the amine and β-alanine in both structures superimposes well on the carbonyl group of the acetyl-CoA substrate reported in the dopamine-AANAT complex (*SI Appendix*, Fig. S4). However, one important difference is that the β-alanyl moiety of the products bound to C_N projects into a hydrophilic cavity that is not present in the dopamine-AANAT structure.

Given the interactions of amino acid side chains with the amine group of β-alanine in this region (Fig. 3), it appears that this cavity may be an important determinant of the specificity of C_N for β-alanine (and a subset of related compounds) over α-amino acids (see below). Analysis of the structurally related dopamine-AANAT (3TE4) postulates an unusual Glu-Ser-Ser catalytic triad being involved in amide bond formation in AANAT catalysis. In C_N, the serine residues are replaced by threonine residues (Thr828 and Thr832) but the glutamate residue is not conserved. The postulated role of these threonine residues in AANAT catalysis is not directly supported in the activity of C_N due to the position of the amine groups of the substrate-bound C_N structures solved here. Indeed, the amine moiety in histamine, which adopts the most likely position for interception of the thioester moiety of the PCP-bound substrate, is instead coordinated by the oxygen atoms of backbone carbonyl groups (residues Phe787 and Thr826) as well as several water

molecules that themselves are further coordinated with the hydrophilic side chains of residues Glu768, His785, and Thr825 (Fig. 3). Mechanistic investigations of an enzyme different from, yet structurally related to, C_N (serotonin NAT) have implicated the equivalent residue to His785 in the mechanism of this NAT (27), although the histamine-bound C_N structure does not support direct interaction between this histidine residue and the amine group of histamines to allow deprotonation. Rather, the structures of C_N indicate that it is likely that this water-mediated interaction network is also able to orient the amine group of the bound substrate in such a way as to promote thioester attack. This does not require rearrangement of the amine substrate, as is required for the postulated mechanism for dopamine AANAT, which appears

Fig. 4. Mutational and substrate specificity studies of the Ebony C_N domain. (*Left*) Condensation activity observed for the wild-type C_N domain and mutants using β-alanyl-CoA **1** and dopamine (triplicate experiments). (*Right*) β-aminoacyl-CoA analogs tested with the C_N domain (**2–7**), indicating their viability as substrates for the condensation reaction with dopamine (glycyl-CoA and α-aminoacyl-CoAs are not accepted by C_N).

C_N Construct	Yield	Error
WT	100%	6.0%
E696L	ND	-
E768Q	76.3%	7.2%
F689A	4.7%	3.9%
F761A	47.2%	5.6%

Izoré et al.

unlikely given the effectiveness of the hydrophobic cage in binding the aromatic side chain of the amine substrates in this case.

To further investigate the role of these residues in the activity of the C_N domain, H785F and E768Q mutants were prepared. The H785F mutant could not be expressed in soluble form, which highlights a likely structural role of His785 in domain folding and stability (including interactions with Glu748 and Glu768). Mutation of Glu768 to Gln did not significantly reduce the activity of C_N (~75% residual activity) (Fig. 4; numbers in bold throughout reference the structures in Fig. 4), supporting the indirect interaction of this residue with the amine group of β-alanine through a network of ordered water molecules as determined in the product-bound structures of the C_N domain. It has been suggested that the differences observed in catalytic mechanisms of AANAT-like enzymes result from the ease of the thioester aminolysis reaction, which does not require a specific catalytic site (27). This argument also appears to hold for traditional NRPS C domains, where the role of the central active site histidine residue is debated across different systems (3). One clear difference, however, is that in standard NRPS C domains the central histidine residue is believed to directly coordinate the substrates during peptide bond formation. In the case of C_N, this is not the case, and the impact of the protein backbone appears to be mostly indirect, serving instead to assist in orienting the amine substrate via a coordinated water network.

Amino Acid Specificity of the Ebony C_N Domain. Within NRPS-catalyzed biosynthesis, amino acid selectivity is largely determined by the activity of specific adenylation domains within the NRPS machinery (1). C domains are believed to play a reduced role in the selection of peptide structure, but rather play a role as stereochemical gatekeepers together with the related epimerization (E) domains (1). Given the structure of C_N and the atypical structure of the β-alanine substrate, we were curious how selective this domain is for different amino acid acceptor substrates. We first confirmed the acceptance of β-alanyl-CoA as a substitute for the β-alanine–loaded PCP domain (28) by Ebony C_N in reactions along with dopamine as the amine acceptor; we confirmed the formation of β-alanyl-dopamine through comparison with an authentic standard (*SI Appendix*, Fig. S5). Next, we utilized the same assay with the C_N domain and a set of 6 β-alanyl-CoA analogs, 7 α-amino acid CoAs, and glycine CoA (Fig. 4 and *SI Appendix*, Fig. S6). These experiments showed that, of the 14 substrates tested, only 3 out of the 6 β-alanine analogs were accepted, while α-amino acids/glycine were not accepted. The lack of acceptance of α-amino acids by C_N appears to be caused by steric hindrance around the α-position of these amino acids through the active site channel formed by parts of β-strands 4 and 5.

The product-bound C_N structures reveal that the side chains of Ser786, Thr825, and Asn827 all coordinate the amine group of β-alanine (Fig. 3) and aid the correct orientation of the PCP-bound (or CoA-bound) amino acid in the C_N active site. The side chain of Asn827, in particular, appears very important to this process, as it is held via hydrogen-bonding interactions from His873 [2.9 Å; His873 also interacts with Thr843 (3.4 Å)] in such a way that it fits into a β-sheet–type orientation that is then also adopted by the backbone of the bound β-alanine molecule and the β-strand 4 Met789 (Fig. 3). The lack of C_N activity with glycine can be rationalized by the loss of interaction between the terminal amino acid and these residues. This is also supported by the lack of acceptance of **2** by C_N, as the terminal amine is missing in this compound. Furthermore, the addition of a methyl substitution to the amine moiety of β-alanine in **4** prevents this compound from being a viable substrate for the C_N domain. This is likely reconciled by the additional steric bulk that this substitution introduces around the crucial amine group. The

importance of coordination with the amine moiety in β-alanine is further supported, albeit indirectly, by the acceptance of **3**; this indicates that the hydrogen-bonding interaction with the protein—although required—can accommodate alternate groups (such as a carboxylate) through alteration of the interacting water network. The methylene extended compound **5** is also accepted by C_N, which can also be rationalized by a rearrangement of the water network that maintains the essential interactions between the terminal amine and the C_N domain. The influence of the C_N binding site on the acceptance of branched substrates is clearly seen in the acceptance of (*R*)-β-homoalanine **7** as a substrate while the enantiomer **6** is not accepted. Inspection of the product-bound C_N structures shows that the *S*-methyl group sterically clashes with the carbonyl group of Asn827 but that the *R*-methyl group is easily accommodated in the cavity toward the more distant Met789 side chain (Fig. 3). These results indicate that the C_N domain requires a substrate with a terminal moiety able to hydrogen-bond effectively within the substrate-binding channel. The ability of the C_N domain to tolerate longer substrates as well as some branching within substrates depends entirely upon the ability of the narrow channel of the amino acyl-PCP binding site to accept them. However, these results indicate the potential for the C_N domain to accept substrates other than β-alanine, which is enforced in full-length Ebony by the selectivity of the A domain.

Discussion

The modular architecture of NRPS assembly lines leads to tremendous diversity within the products assembled by them, with the specificity of the assembly lines largely ascribed to the functions of adenylation domains (1). It has been recognized that C domains—which are essential catalytic NRPS domains—can play much wider roles in generating product diversity within NRPS pathways (3). Examples of atypical catalytic functions for traditional NRPS C domains include the generation of a β-lactam ring in nocardicin biosynthesis (29), multiple-step heterocyclization reactions, and elimination and rearrangement to produce methoxyvinyl-containing amino acids (30); the noncatalytic roles for these domains include the recruitment of multiple cytochrome P450 enzymes to perform oxidative cross-linking of aromatic side chains within glycopeptide antibiotic biosynthesis (31). It is clear that, within classical NRPS-type architectures, there is significant potential to identify additional functions within "standard" catalytic domains.

Our results demonstrate that Ebony C_N is a condensation domain in NRPS machinery, having specific selectivity requirements for the PCP-bound amino acid. The C_N domain also adopts a totally different fold, belonging to the AANAT superfamily of enzymes despite very low levels of sequence identity. The relative chemical ease of the reaction performed by both Ebony C_N and standard NRPS C domains is reflected in the active sites found in both folds, although in the case of Ebony C_N it appears that the enzyme controls the orientation of the amine nucleophile through a water network rather than a central histidine residue. Use of a highly charged active site channel for the amine substrate in Ebony C_N somewhat diverges from the structure of traditional C domains with related amine acceptor substrates such as VibH. This likely stems from the rapid rate of catalysis required by Ebony compared with NRPS systems from secondary metabolism pathways. The structure of Ebony C_N also appears to be highly rigid, with little if any rearrangement of the protein upon substrate binding. Such rigidity also stands in contrast to traditional NRPS C domains, in which the relative orientation of the two CAT-like subdomains generates differences in the accessibility of the acceptor site. The ability of traditional C domains to adopt open and closed conformations (2, 3) with regard to the acceptor site is one way that an NRPS assembly line can generate directionality during synthesis across

multiple modules and hence multiple peptide bond formation steps. In the case of Ebony, the ability to prealign substrates in a rigid active site leads to a significant increase in the maximal rate of this reaction, which is crucial for its function in vivo. Although binding affinities for the amine substrates differ considerably [K_d: ~30 μM (dopamine) vs. ~600 μM (histamine)] in accordance with their different physiological roles (*SI Appendix*, Table S2), the use of the AANAT fold to perform this condensation reaction appears to be an effective way for this eukaryotic NRPS to control the rate of reaction by dispensing with a traditional, significantly slower C domain.

In utilizing a different fold to perform the role traditionally held by a C domain, Ebony C_N is reminiscent of other recent examples of reactions found in NRPS catalysis within different enzyme folds [e.g., NRPS offloading via a penicillin-binding protein (32)]. Our results obtained with Ebony C_N serve once again to demonstrate that significant diversity exists in the enzymatic machinery behind NRPS pathways and that these alternate enzymatic systems provide clear advantages for the specific function of these assembly lines. Given the importance of the products of NRPS pathways for human health, it is crucial that we gain an understanding of the different strategies adopted by naturally occurring assembly lines for peptide synthesis if we are to undertake the reengineering of NRPS assembly lines to produce new molecules with targeted structures and function.

Methods

Protein Expression and Purification. The Ebony C_N construct was cloned via PCR into the pHIS-17 plasmid, expressed in *E. coli* and purified using NiNTA affinity and gel filtration. Mutants were generated using standard PCR procedures, expressed, and purified as wild type (for details see *SI Appendix*).

Protein Crystallization and Structure Determination. Datasets were collected at the Australian Synchrotron (Victoria, Australia) on either beamline MX1 or MX2 equipped with an Eiger detector (Dectris) (*SI Appendix*, Table S1) (33–37). For details see *SI Appendix*.

Compound Synthesis. Compounds were synthesized using standard peptide synthesis procedures (for details see *SI Appendix*).

Activity Assays. Peptide bond formation using Ebony C_N was assessed with different substrates (for details see *SI Appendix*). ITC experiments are described in the *SI Appendix*.

1. Süssmuth RD, Mainz A (2017) Nonribosomal peptide synthesis-principles and prospects. *Angew Chem Int Ed Engl* 56:3770–3821.
2. Izoré T, Cryle MJ (2018) The many faces and important roles of protein-protein interactions during non-ribosomal peptide synthesis. *Nat Prod Rep* 35:1120–1139.
3. Bloudoff K, Schmeing TM (2017) Structural and functional aspects of the non-ribosomal peptide synthetase condensation domain superfamily: Discovery, dissection and diversity. *Biochim Biophys Acta Proteins Proteomics* 1865:1587–1604.
4. Di Vincenzo L, Grgurina I, Pascarella S (2005) In silico analysis of the adenylation domains of the freestanding enzymes belonging to the eucaryotic nonribosomal peptide synthetase-like family. *FEBS J* 272:929–941.
5. Kasahara T, Kato T (2003) Nutritional biochemistry: A new redox-cofactor vitamin for mammals. *Nature* 422:832.
6. Ehmann DE, Gehring AM, Walsh CT (1999) Lysine biosynthesis in Saccharomyces cerevisiae: Mechanism of alpha-aminoadipate reductase (Lys2) involves posttranslational phosphopantetheinylation by Lys5. *Biochemistry* 38:6171–6177.
7. Hovemann BT, et al. (1998) The *Drosophila* ebony gene is closely related to microbial peptide synthetases and shows specific cuticle and nervous system expression. *Gene* 221:1–9.
8. Richardt A, Rybak J, Störtkuhl KF, Meinertzhagen IA, Hovemann BT (2002) Ebony protein in the *Drosophila* nervous system: Optic neuropile expression in glial cells. *J Comp Neurol* 452:93–102.
9. Borycz J, Borycz JA, Loubani M, Meinertzhagen IA (2002) Tan and ebony genes regulate a novel pathway for transmitter metabolism at fly photoreceptor terminals. *J Neurosci* 22:10549–10557.
10. Hotta Y, Benzer S (1969) Abnormal electroretinograms in visual mutants of *Drosophila*. *Nature* 222:354–356.
11. Suh J, Jackson FR (2007) *Drosophila* ebony activity is required in glia for the circadian regulation of locomotor activity. *Neuron* 55:435–447.
12. Hartwig S, Dovengerds C, Herrmann C, Hovemann BT (2014) *Drosophila* Ebony: A novel type of nonribosomal peptide synthetase related enzyme with unusually fast peptide bond formation kinetics. *FEBS J* 281:5147–5158.
13. Richardt A, et al. (2003) Ebony, a novel nonribosomal peptide synthetase for beta-alanine conjugation with biogenic amines in *Drosophila*. *J Biol Chem* 278: 41160–41166.
14. Keating TA, Marshall CG, Walsh CT, Keating AE (2002) The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. *Nat Struct Biol* 9:522–526.
15. Charlop-Powers Z, Jakoncic J, Gurnon JR, Van Etten JL, Zhou MM (2012) *Paramecium bursaria* chlorella virus 1 encodes a polyamine acetyltransferase. *J Biol Chem* 287: 9547–9551.
16. Kim DE, Chivian D, Baker D (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* 32:W526–W531.
17. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10:845–858.
18. Vetting MW, et al. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. *Arch Biochem Biophys* 433:212–226.
19. Cheng KC, Liao JN, Lyu PC (2012) Crystal structure of the dopamine N-acetyltransferase-acetyl-CoA complex provides insights into the catalytic mechanism. *Biochem J* 446:395–404.
20. Dempsey DR, et al. (2017) Structural and mechanistic analysis of *Drosophila melanogaster* agmatine N-acetyltransferase, an enzyme that catalyzes the formation of N-acetylagmatine. *Sci Rep* 7:13432.
21. Kittilä T, Mollo A, Charkoudian LK, Cryle MJ (2016) New structural data reveal the motion of carrier proteins in nonribosomal peptide synthesis. *Angew Chem Int Ed Engl* 55:9834–9840.
22. Wu R, Reger AS, Lu X, Gulick AM, Dunaway-Mariano D (2009) The mechanism of domain alternation in the acyl-adenylate forming ligase superfamily member 4-chlorobenzoate: Coenzyme A ligase. *Biochemistry* 48:4115–4125.
23. Denno ME, Privman E, Borman RP, Wolin DC, Venton BJ (2016) Quantification of histamine and carcinine in *Drosophila melanogaster* tissues. *ACS Chem Neurosci* 7: 407–414.
24. Borycz JA, Borycz J, Kubów A, Kostyleva R, Meinertzhagen IA (2005) Histamine compartments of the *Drosophila* brain with an estimate of the quantum content at the photoreceptor synapse. *J Neurophysiol* 93:1611–1619.
25. de Ruyter van Steveninck RR, Laughlin SB (1996) The rate of information transfer at graded-potential synapses. *Nature* 379:642–645.
26. White HL, Cavallito CJ (1970) Choline acetyltransferase. Enzyme mechanism and mode of inhibition by a styrylpyridine analogue. *Biochim Biophys Acta* 206:343–358.
27. Scheibner KA, De Angelis J, Burley SK, Cole PA (2002) Investigation of the roles of catalytic residues in serotonin N-acetyltransferase. *J Biol Chem* 277:18118–18126.
28. Belshaw PJ, Walsh CT, Stachelhaus T (1999) Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* 284:486–489.
29. Gaudelli NM, Long DH, Townsend CA (2015) β-Lactam formation by a non-ribosomal peptide synthetase during antibiotic biosynthesis. *Nature* 520:383–387.
30. Patteson JB, Dunn ZD, Li B (2018) In vitro biosynthesis of the nonproteinogenic amino acid methoxyvinylglycine. *Angew Chem Int Ed Engl* 57:6780–6785.
31. Haslinger K, Peschke M, Brieke C, Maximowitsch E, Cryle MJ (2015) X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. *Nature* 521:105–109.
32. Kuranaga T, et al. (2018) Total synthesis of the nonribosomal peptide surugamide B and identification of a new offloading cyclase family. *Angew Chem Int Ed Engl* 57: 9447–9451.
33. Izoré T, et al. (2018) C-terminal condensation domain of Ebony. Protein Data Bank. Available at https://www.rcsb.org/structure/6DYM. Deposited July 3, 2018.
34. Izoré T, et al. (2018) C-terminal condensation domain of Ebony in complex with histamine. Protein Data Bank. Available at https://www.rcsb.org/structure/6DYN. Deposited July 3, 2018.
35. Izoré T, et al. (2018) C-terminal condensation domain of Ebony in complex with L-dopamine. Protein Data Bank. Available at https://www.rcsb.org/structure/6DYO. Deposited July 3, 2018.
36. Izoré T, et al. (2018) C-terminal condensation domain of Ebony in complex with carcinine. Protein Data Bank. Available at https://www.rcsb.org/structure/6DYR. Deposited July 3, 2018.
37. Izoré T, et al. (2018) C-terminal condensation domain of Ebony in complex with beta-alanyl-dopamine. Protein Data Bank. Available at https://www.rcsb.org/structure/6DYS. Deposited July 3, 2018.