

Molecular Basis of Substrate Promiscuity for the SAM-Dependent *O*-Methyltransferase NcsB1, Involved in the Biosynthesis of the Enediye Antitumor Antibiotic Neocarzinostatin[†]

Heather A. Cooke,[§] Elizabeth L. Guenther,[§] Yinggang Luo,^{||} Ben Shen,^{||,‡} and Steven D. Bruner*,[§]

[§]Department of Chemistry, Merkert Chemistry Center, Boston College, Chestnut Hill, Massachusetts 02467, ^{||}Division of Pharmaceutical Sciences and [‡]Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53705

Received July 22, 2009; Revised Manuscript Received August 20, 2009

ABSTRACT: The small molecule component of chromoprotein enediye antitumor antibiotics is biosynthesized through a convergent route, incorporating amino acid, polyketide, and carbohydrate building blocks around a central enediye hydrocarbon core. The naphthoic acid moiety of the enediye neocarzinostatin plays key roles in the biological activity of the natural product by interacting with both the carrier protein and duplex DNA at the site of action. We have previously described the *in vitro* characterization of an *S*-adenosylmethionine-dependent *O*-methyltransferase (NcsB1) in the neocarzinostatin biosynthetic pathway [Luo, Y., Lin, S., Zhang, J., Cooke, H. A., Bruner, S. D., and Shen, B. (2008) *J. Biol. Chem.* 283, 14694–14702]. Here we provide a structural basis for NcsB1 activity, illustrating that the enzyme shares an overall architecture with a large family of *S*-adenosylmethionine-dependent proteins. In addition, NcsB1 represents the first enzyme to be structurally characterized in the biosynthetic pathway of neocarzinostatin. By cocrystallizing the enzyme with various combinations of the cofactor and substrate analogues, details of the active site structure have been established. Changes in subdomain orientation were observed via comparison of structures in the presence and absence of substrate, suggesting that reorientation of the enzyme is involved in binding of the substrate. In addition, residues important for substrate discrimination were predicted and probed through site-directed mutagenesis and *in vitro* biochemical characterization.

Enediye antitumor antibiotics are structurally complex natural products possessing remarkable cytotoxicity (1). Members of the nine-member subfamily of enediynes are composed of a small molecule chromophore and an apoprotein, which sequesters and stabilizes the reactive enediye and aids in delivery to target cells, leading to single- and double-stranded DNA cleavage (2). Biosynthetic gene clusters of multiple enediynes from actinomycetes have recently been sequenced and annotated, allowing detailed investigation into their complex biosynthesis (3–8). The pathways utilize diverse enzymology to construct the natural products, including the use of polyketide and nonribosomal peptide machinery. Neocarzinostatin (NCS,¹ **1**) is an archetypal example of the nine-membered enediynes, the studies of which have contributed to the general understanding of enediye biosynthesis and mode of action (4, 9–12). The naphthoate moiety of NCS functions by both binding to the apoprotein and intercalating target DNA, thereby positioning the enediye in the minor groove (9, 10, 13–15). Alterations of the peripheral moieties of NCS surrounding the enediye core

may be a way to control or mediate the reactivity and specificity of NCS (1).

The recent cloning and sequencing of the NCS biosynthetic pathway from *Streptomyces carzinostaticus* revealed a convergent path with two polyketide synthases playing central roles in building the carbon framework (4, 12, 16). The scaffold for the naphthoate moiety in NCS is constructed by an iterative type I polyketide synthase via a mechanism common to the biosynthesis of diverse aromatic products (17–19). Four additional gene products decorate and couple the naphthoic acid to the enediye core (Figure 1). The product of NcsB (**2**) is first hydroxylated at the C-7 position by a cytochrome P450 hydroxylase (NcsB3) to produce **3**. The resulting hydroxyl group is methylated by NcsB1, producing the methyl ether **4** present in the natural product. This functionalized naphthoic acid is activated via adenylation, converted to the aryl-CoA **5** by the CoA-ligase NcsB2, and finally attached to the enediye core by NcsB4 (a putative acyl-transferase) to produce the natural product (**1**). The assignment of the naphthoic acid biosynthetic pathway is based on bioinformatic analysis (4) and biochemical characterization of NcsB2 as a CoA-ligase (20) and NcsB1 as an *S*-adenosylmethionine (SAM)-dependent *O*-methyltransferase (*O*-MTase) (21). NcsB1 was shown to catalyze the third step in the biosynthesis of the naphthoate moiety of the NCS chromophore (**3** to **4**). Biochemical analysis also revealed a relaxed substrate flexibility toward substituted naphthoic acids, methylating a variety of 3-hydroxy-2-naphthoic acids with the reactive hydroxyl group located at either the 5 or 7 position on the ring. NcsB1 also was demonstrated to methylate 1,4-dihydroxy-2-naphthoic acid

[†]This work was supported in part by funds from the Damon Runyon Cancer Research Fund (to S.D.B.) and National Institutes of Health Grants CA78747 and CA113297 (to B.S.).

*To whom correspondence should be addressed. Telephone: (617) 552-2931. Fax: (617) 552-2705. E-mail: bruner@bc.edu.

[‡]Abbreviations: NCS, neocarzinostatin; SAH, *S*-adenosyl-L-homocysteine; SAM, *S*-adenosyl-L-methionine; CoA, coenzyme A; PP_i, pyrophosphate; MTase, methyltransferase; DHN, 1,4-dihydroxy-2-naphthoic acid; rmsd_{Cα}, root-mean-square difference at α -carbons; PDB, Protein Data Bank; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

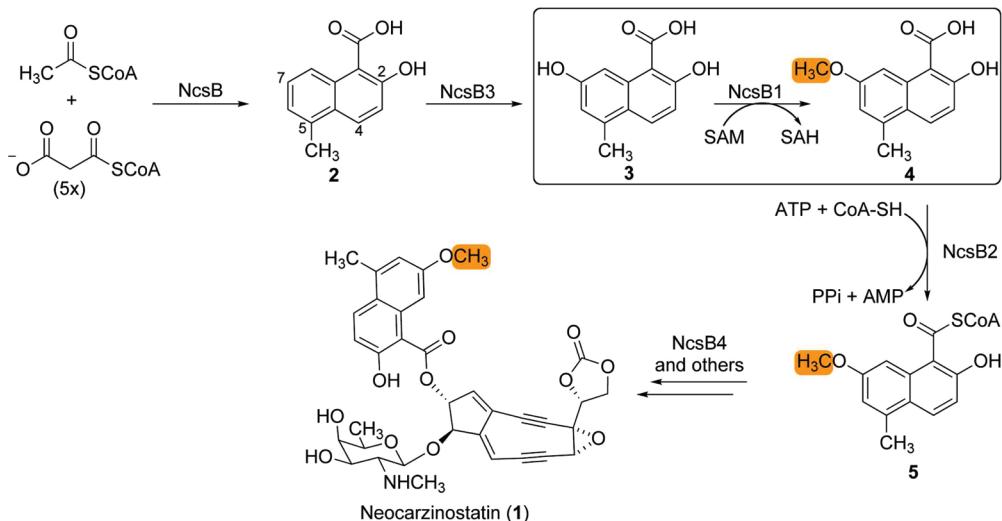


FIGURE 1: Biosynthesis of the naphthoic acid moiety and its attachment to the enediyne core, resulting in the natural product neocarzinostatin (**1**). The *O*-methyl transfer reaction catalyzed by NcsB1 is boxed (methyl group highlighted in orange).

(see Figure 5, 6) at the 4-hydroxyl group, which was surprising because of the large difference in the location of the reactive hydroxyl group from that on the natural substrate, necessitating a reorientation of the substrate in the active site. This observation led to the prediction of a model by which naphthoic acids flexibly bind the active site and in which the 1,2-hydroxy acid functionality was required for substrate binding and SAM-dependent methylation (21).

There are ~160 classes of SAM-dependent MTases comprised of thousands of identified members. These functionally diverse enzymes methylate a wide range of substrates, including nucleic acids for regulation of gene expression, DNA repair or protection against restriction enzymes, proteins for repair or control of signal transduction pathways, hormones and neurotransmitters, and biosynthetic intermediates to produce secondary metabolites. Structurally characterized MTases can be segregated into five general classes on the basis of specific structural features (22). Among the several examples of MTases that act on small molecules, catechol *O*-MTase isolated from rat liver was the first to be structurally characterized (23). This structure exemplifies the basic SAM-MTase core fold: a mixed α , β secondary structure with a β -sheet flanked on each side by three α -helices (24). A common feature of the active site of MTases is a glycine-rich region within the SAM binding pocket. Despite a low degree of overall sequence homology among family members, structurally characterized SAM-MTase complexes share this common fold and the biologically active unit is a homodimer.

The role of NcsB1 was assigned in the enediyne neocarzinostatin biosynthetic pathway on the basis of sequence homology to small molecule SAM-dependent *O*-MTases, including DnrK, an enzyme involved in the biosynthesis of the aromatic polyketide antibiotic daunorubicin from *Streptomyces peucetius* (44% sequence identity) (25). The X-ray crystal structure of DnrK revealed a homodimer with each monomer exhibiting the conserved small molecule MTase fold (26). RdmB from rhodomycin biosynthesis in *Streptomyces purpurascens* is also homologous to NcsB1, being 44% identical and 58% similar, but interestingly is a hydroxylase, not an *O*-MTase (27, 28). RdmB was used as a model for molecular replacement in determining the crystal structure of DnrK. In addition to DnrK, other bacterial small molecule *O*-MTase structures have recently been determined,

including examples requiring a cation for catalysis. Cation-dependent *O*-MTases include CmcI from cephamycin biosynthesis (29), LiOMT from the pathogenic bacterium *Leptospira interrogans* (30), and BcOMT2 from *Bacillus cereus* (31). A limited number of cation-independent bacterial *O*-MTase structures have been structurally characterized, including DnrK and RebM from rebeccamycin biosynthesis (32).

Here we report the structural basis for NcsB1 activity in a variety of cocomplexes with SAM or *S*-adenosyl-L-homocysteine (SAH) with or without 5-methyl-2-hydroxynaphthoic acid (**2**), product (**4**), or 1,4-dihydroxy-2-naphthoic acid (**6**), an alternate substrate. The structures reveal that NcsB1 shares an overall architecture common to the large MTase family. The active site binding pocket is able to accommodate the natural substrate and structurally diverse analogues, allowing efficient methylation with distinct regiospecificity. The specificity determinants of the naphthoate binding pocket were probed using site-directed mutagenesis and alternate substrates. On the basis of the results, residues that affect the substrate specificity of the enzyme were identified.

MATERIALS AND METHODS

Protein Expression and Purification. NcsB1 was overproduced as an N-terminal His₆-tagged fusion protein using expression plasmid pBS5039 in *Escherichia coli* BL21(DE3) cells as reported previously (21). Cells were grown in 1 L of Luria-Bertani medium at 37 °C and 150 rpm to an OD₆₀₀ of 0.5–0.8. Overexpression was induced with 50 μ M isopropyl β -D-thiogalactopyranoside at 18 °C for 16 h. Cells were collected by centrifugation (20 min at 2000g), resuspended in 25 mL of 20 mM Tris-HCl (pH 7.5) and 500 mM NaCl, and flash-frozen at -78 °C. Cell pellets were thawed and lysed via two passes through a French press cell disruptor at 1000 psi. The lysate was then clarified by centrifugation (20 min at 10000g), and the protein was purified batchwise using Ni-NTA resin (Qiagen, Valencia, CA). NcsB1 used in biochemical assays was dialyzed into 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, and 1 mM β -mercaptoethanol, then concentrated using an Amicon Ultra-4 concentrator (10 kDa molecular mass cutoff, GE Healthcare), and frozen with 40% supplemented glycerol at -25 °C.

Table 1: Data Collection and Refinement Statistics

	SAH	SAM and 2	SAH and 4	SAH and 6
space group	<i>C</i> 2 <i>2</i> ₁	<i>P</i> 6 ₅	<i>P</i> 6 ₅	<i>P</i> 6 ₅
unit cell parameters (Å)	<i>a</i> = 91.3, <i>b</i> = 161.6, <i>c</i> = 98.9	<i>a</i> = <i>b</i> = 109.1, <i>c</i> = 206.9	<i>a</i> = <i>b</i> = 108.4, <i>c</i> = 210.3	<i>a</i> = <i>b</i> = 108.0, <i>c</i> = 211.9
Data Collection				
resolution (Å)	25.0–2.08 (2.15–2.08)	50–2.6 (2.69–2.60)	30.0–2.69 (2.79–2.69)	50.0–3.0 (3.1–3.0)
wavelength (Å)	1	1	1	1
no. of reflections (measured/unique)	345090/44038	967889/42568	658394/38115	35878/27798
% completeness ^a	99.6 (99.6)	99.8 (100.0)	98.5 (100.0)	99.8 (100.0)
<i>R</i> _{merge} ^a	0.052 (0.349)	0.077 (0.441)	0.094 (0.49)	0.076 (0.523)
<i>I</i> / <i>σ</i> ^a	32.2 (4.4)	49.6 (6.4)	24.0 (7.0)	29.3 (4.5)
redundancy ^a	7.8 (7.4)	22.7 (22.9)	17.3 (17.8)	12.6 (12.8)
Refinement				
no. of reflections (total/test)	41442/2071	41357/2064	37165/3704	26923/2698
<i>R</i> _{work}	0.240	0.206	0.212	0.217
<i>R</i> _{test}	0.272	0.227	0.247	0.245
no. of amino acids (chains A/B)	325/303	325/325	328/328	328/328
no. of atoms				
protein (chains A/B)	2379/2193	2379/2379	2406/2406	2406/2406
ligands (cofactor/substrate/glycerol)	52/−/6	54/30/12	52/34/12	50/30/12
water molecules	305	249	163	95
B-factors				
(NcsB1/cofactor/substrate/water)	46.0/36.0/−/50.1	53.0/42.3/51.6/55.6	71.4/59.6/81.4/74.3	69.9/59.9/57.2/65.4
rmsd from ideal geometry				
bonds (Å)	0.006	0.006	0.006	0.0013
angles (deg)	1.2	1.3	1.2	1.7
Ramachandran plot by PROCHECK (%)				
core region	90.6	93.2	91.7	88.8
allowed region	9.0	6.2	7.9	10.4
generously allowed	0.2	0.5	0.2	0.7
disallowed	0.2	0.0	0.2	0.0
PDB entry	3I53	3I5U	3I58	3I64

^aOverall (highest-resolution shell).

For crystallography, the Ni-NTA-purified protein was dialyzed into 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 2 mM CaCl₂, and 1 mM β-mercaptoethanol and concentrated to ~1 mL. The His₆ tag was cleaved by incubation with protease Factor Xa for 36 h at 4 °C (monitored by SDS-PAGE). The protein solution was then diluted to 5 mL and purified on a HiTrap-Q ion exchange column [0 to 1 M NaCl gradient in 50 mM Tris-HCl (pH 7.5) and 1 mM β-mercaptoethanol] followed by a Superdex 200 gel filtration column [20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM β-mercaptoethanol, and 10% glycerol]. The purified protein was concentrated to ~10 mg/mL for crystallography (a total yield of 4 mg/L of cells). Protein concentrations were determined using UV absorption at 280 nm and the Bradford assay.

Site-Directed Mutagenesis of ncsB1. The Quikchange II site-directed mutagenesis kit (Stratagene, La Jolla, CA) was used to generate mutant constructs of NcsB1 using the following DNA oligonucleotide primers (mutant codon underlined): Arg11Ala, 5'-GGC TGC ACA CAT CGG ATT GGC GGC GCT GGC CGA TCT GGC GAC-3' (forward) and 5'-GTC GCC AGA TCG GCC AGC GCC GGC AAT CCG ATG TGT GCA GCC-3' (reverse); Arg11Trp, 5'-GGC TGC ACA CAT CGG ATT GTG GGC GCT GGC CGA TCT GGC GAC-3' (forward) and 5'-GTC GCC AGA TCG GCC AGC GCC CAC AAT CCG ATG TGT GCA GCC-3' (reverse); Arg11Lys, 5'-GGC TGC ACA CAT CGG ATT GAA GGC GCT GGC CGA TCT GGC GAC-3' (forward) and 5'-GTC GCC AGA TCG GCC AGC

GCC TTC AAT CCG ATG TGT GCA GCC-3' (reverse); Tyr293Ile, 5'-GCG CAT GCT CAC CAT CTT CGG AGG CAA GGA ACG C-3' (forward) and 5'-GCG TTC CTT GCC TCC GAA GAT GGT GAG CAT GCG C-3' (reverse). The Quikchange protocol was followed as described with the addition of DMSO to the reaction to a final concentration of 5%. Mutations were confirmed by sequencing (Genewiz, South Plainfield, NJ), and mutant constructs were transformed into *E. coli* BL21-(DE3) cells, overexpressed, and purified as described above.

Crystallization and X-ray Data Collection. Purified NcsB1 was crystallized with a substituted naphthoic acid and/or SAM/SAH by the hanging drop vapor diffusion method at 20 °C with ~4 M sodium formate [7 M stock solution from Hampton Research (see Table S1 of the Supporting Information for crystallization conditions)]. Stock solutions of the small molecules for cocrystallization were in 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM β-mercaptoethanol, and 10% glycerol. Naphthoic acids were synthesized as described previously (20) or purchased from Aldrich. Typically, NcsB1 crystals appeared within 1 day and reached full size within 1 week. Mature crystals were transferred to a cryoprotectant solution (4.0 M sodium formate and 15% glycerol) and soaked briefly before being flash-frozen in liquid nitrogen. X-ray diffraction data were collected on beamline X25 at the National Synchrotron Light Source at Brookhaven National Laboratories with an ADSC Q315 CCD X-ray detector. Diffraction intensities were

indexed, integrated, and scaled using HKL2000 (33) as summarized in Table 1. The crystals belonged to space group $P6_5$ except for the NcsB1–SAH cocomplex, which belonged to space group $C222_1$.

Structure Determination. An initial solution for the NcsB1–SAM–2 ternary complex was obtained using the molecular replacement program PHASER (34), part of the CCP4 suite (35). A polyalanine dimer model of DnrK from *S. peucetius* (PDB entry 1TW2) truncated to residues 15–350 was used as the search model. Manual building of the NcsB1 model was performed using COOT (36), and the structure was refined using CNS (37, 38). This solution was used to solve phases in subsequent data sets in the $P6_5$ space group using molecular replacement. For the NcsB1–SAH cocomplex ($C222_1$ space group), a partial solution (chain A and the N-terminus of chain B) was found using PHASER. The remaining C-terminus of chain B was placed into the model using MOLREP (39). All final structures were subjected to multiple rounds of building and refinement until the R values converged. Noncrystallographic restraints were used for all $P6_5$ space group structures until the final stages of refinement. Ligands were fit into the electron density maps, and the PRODRG server (40) was used for generating topology and parameter files; SAM and SAH coordinates were obtained from the HICup server (41) and were subsequently altered to fit electron density. PyMOL (DeLano Scientific, San Carlos, CA) was used to generate graphic images. rmsd values were calculated via structural comparison using TopMatch (42, 43).

The NcsB1–SAH structure ($C222_1$ space group) was determined at a resolution of 2.08 Å and refined to R_{work} and R_{free} values of 24.0 and 27.2%, respectively (Table 1). The asymmetric unit consisted of two monomers. For chain A, there was sufficient electron density to build in residues 5–330 and the bound SAH molecule. For chain B, the final model consisted of residues 9–150, 159–274, 282–297, and 302–330, as insufficient electron density was available to build the 29 remaining residues. Electron density for the SAH molecule was well-defined, albeit with higher B factors for the SAH molecule in chain B. In addition, the dimer contained one glycerol molecule. The structure of the NcsB1 cocomplex with SAM and the substrate analogue (2) was determined at a resolution of 2.60 Å with R_{work} and R_{free} values of 20.6 and 22.7%, respectively. For both chains, N-terminal residues 1–4 could not be built, nor could residues 279–281, though sufficient electron density was available for both SAM and 2 to be built into each active site. NcsB1 with SAH and product (4) diffracted to 2.69 Å and was refined to R_{work} and R_{free} values of 21.2 and 24.7%, respectively. All residues of both chains could be built in as well as SAH, the product in the active site, and four glycerol molecules.

The final structure was a cocomplex of NcsB1 with SAH and an alternate substrate, 1,4-dihydroxynaphthoic acid (6). The crystal diffracted to 3.0 Å, likely a result of reduced crystal quality as 10 mM β -mercaptoethanol was required for crystal growth because of problematic oxidation of the substrate. The final structure refined to R_{work} and R_{free} values of 21.7 and 24.5%, respectively, and included residues 4–332, SAH, and 6 in the final model.

Kinetics of NcsB1 Mutants. To determine the kinetic parameters of all mutant constructs with the natural substrate, the previously reported biochemical assay was used with minor modifications (21). Reactions were conducted at 25 °C, and mixtures contained 10 μM enzyme, 2.5 mM SAM, 100 mM

phosphate buffer (pH 6.5), and varying concentrations of 3 ranging from 50 μM to 2.5 mM for a total volume of 50 μL . Reactions were quenched with trifluoroacetic acid to a final concentration of 16% after 15 min, except for the wild type (WT) and Arg11Ala, which were quenched at 10 min. Samples were centrifuged at 14000g for 2 min, and the supernatant was analyzed by high-performance liquid chromatography (HPLC) at 340 nm [eluent A, water with 0.1% trifluoroacetic acid; eluent B, acetonitrile; gradient, 10 to 60% acetonitrile from 0 to 12 min and 0% from 12 to 17 min at a flow rate of 1 mL/min on a VYDAC C18 Protein and Peptide column (4.6 mm × 250 mm)]. To obtain kinetic parameters, initial rates were plotted against substrate concentration and were fitted to the Michaelis–Menten equation using KaleidaGraph (Synergy Software, Reading, PA).

RESULTS AND DISCUSSION

Crystallization and X-ray Structure Determination. The 34.5 kDa/332-residue NcsB1 was cocrystallized with various combinations of SAH, SAM, deshydroxy substrate analogue 2, and methylated product 4. Analogue 2 was used in crystallization experiments because of the oxidative instability of the substrate (3) over extended time periods. Two crystal forms were generated depending on the nature of the bound small molecules. When cocrystallized with a substrate analogue/product and SAH or SAM, the cocomplex crystallized in the $P6_5$ space group. In the absence of naphthoic acid, distinct crystals belonging to the $C222_1$ space group formed. The NcsB1–SAM–2 ternary complex was the initial structure determined using molecular replacement with a polyalanine dimer model of DnrK (44% identical) as the search model. This structure was then used as a search model for the alternate cocomplexes. Solution of the NcsB1–SAH complex ($C222_1$ space group) by molecular replacement was not straightforward, suggesting an alternate overall structure. The final structural model was obtained by performing rotation and translation searches separately on the N- and C-terminal domains. While a 0.6 Å rmsd_{Cα} was observed among the $P6_5$ structures, the NcsB1–SAH structure differed significantly from the others by an rmsd_{Cα} of 1.8 Å. In addition, the two subunits of the NcsB1–SAH structure were quite distinct, with an rmsd_{Cα} of 2.8 Å between the homodimer chains.

Overview of the Structures. The resulting dimer models of NcsB1 display a high degree of structural similarity with DnrK (rmsd_{Cα} = 1.8 Å). In the initial NcsB1–SAM–2 structure, and others in the $P6_5$ space group, the two monomeric subunits exhibit a high degree of symmetry (Figure 2A). Each subunit of the NcsB1 dimer is made up of three domains totaling 19 α -helices and eight β -sheets (Figure 2B). The N-terminal domain is largely α -helical, with just two β -strands (β 1 and β 2), and constitutes most of the residues involved in the homodimer interface (~5300 Å² total surface). Of note, α 5 (residues 52–65) has significant interactions with α 18 (residues 284–292) from the C-terminal domain of the paired dimer. The dimer interface consists of primarily hydrophobic residues, and a dimeric biological unit was supported by size-exclusion chromatography (Figure S1 of the Supporting Information). A middle domain acts as a hingelike region between the two larger terminal domains and consists of α -helices 9–11. The C-terminal domain exhibits a Rossman-like fold (β 3–9 flanked by α 12–19), which is conserved in structural homologues and makes up the majority of the SAM binding site (24). On the basis of the structural alignment using the DALI server, NcsB1 shares the highest

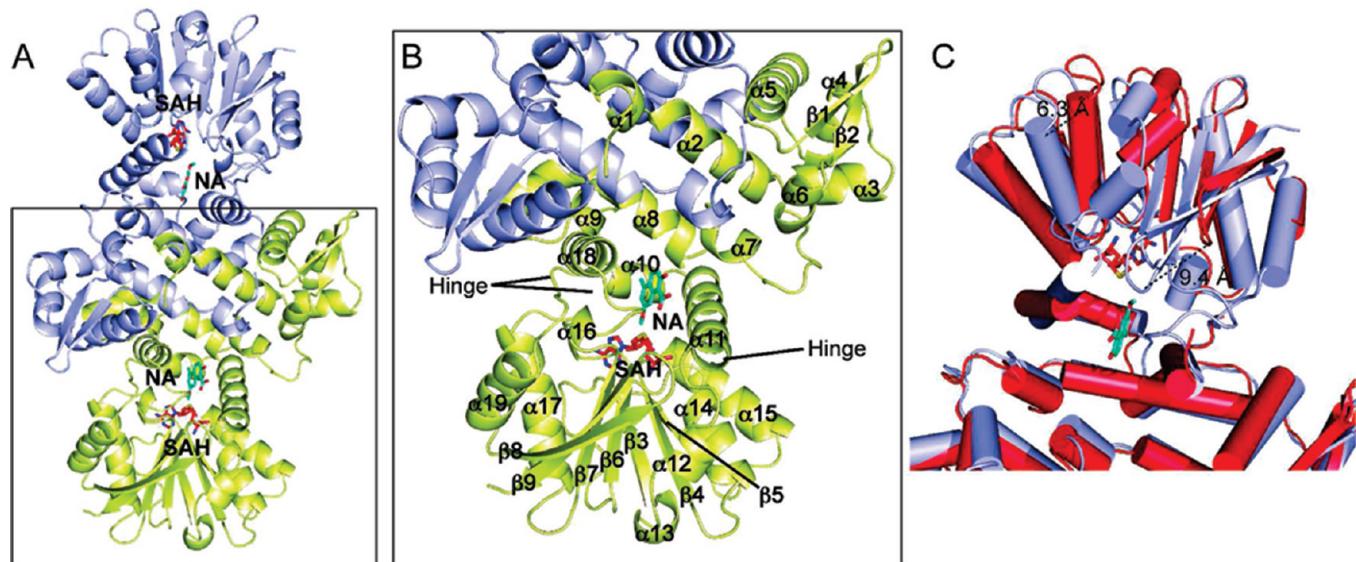


FIGURE 2: Cartoon representations of NcsB1. (A) Dimer of NcsB1 with the active site region indicated. Chain A is colored light blue and chain B yellow. Ligands are depicted in stick format with SAH colored red and naphthoic acid **4** cyan. (B) Close-up view of the monomer with secondary structural elements and hinge regions labeled. (C) Overlay of the NcsB1–SAH–**4** monomer (light blue) and the NcsB1–SAH monomer (red).

degree of structural similarity with SAM-dependent enzymes that act on small molecules (Figure 3 and Table S2) (44). Homologues include DnrK (an *O*-MTase) and RdmB (a hydroxylase), iso-flavone-*O*-MTase from alfalfa (45), a putative *O*-MTase from *Nostoc punctiforme*, a phenazine-specific *N*-MTase from *Pseudomonas aeruginosa* (46), and caffeic acid 3-*O*-MTase from alfalfa (47). Among the structures, an indicative glycine-rich region of the SAM binding pocket is conserved. These six structural homologues demonstrate the ability of the conserved scaffold to catalyze a variety of transformations.

In the NcsB1–SAH cocomplex structure, the C-terminal domain of chain B is rotated by ~20° compared to that in the NcsB1–SAH–**2** structure, displacing α13 by 6.8 Å and β8/β9 by 9.4 Å (Figure 2C). This movement results in a more open conformation of the active site, possibly functioning to allow the entrance of substrate. A similar domain displacement was seen in the DnrK ternary complex where two crystals with different space groups were also observed (26). Both DnrK crystals were bound to SAH and product, but the *B* factors were higher in one of the two subunits, suggesting a lower occupancy of the ligands. One of the crystal structures exhibited a significant difference between the two monomers, with an rmsd_{Cα} of 1.3 Å. Likewise, the RdmB structures exhibit a domain movement upon substrate binding, though in this cocomplex structure, both domains were displaced (27). The NcsB1–SAH cocomplex described here exhibits a larger difference, likely due to a complete lack of substrate analogue or product under the crystallization conditions. In addition, SAH bound in the open subunit (chain B) has higher relative *B* factors. Several regions in chain B exhibited weak or no electron density, possibly indicating areas of flexibility on which the domain can twist. These hinge regions include part of α11 (residues 150–159), the loop region between β7 and α18 (residues 274–282), and the loop between α18 and α19 (residues 297–301) (Figure 2B).

SAM/SAH Binding Pocket. The SAM/SAH binding pocket is located at the C-terminus of the β-strands comprising part of the Rossman-like fold. Electron density for SAM or SAH is clearly defined in all four cocomplex structures, and the SAM binding interactions are largely conserved among small molecule

MTases. The adenine ring is involved in a hydrogen bonding interaction with Ser227, a π interaction with Phe228, and the binding pocket is lined with additional aromatic and/or hydrophobic residues, including Trp133, Trp248, Phe229, and Leu201 (Figure 4A). The ribosyl moiety is anchored by two hydrogen bonds between Asp200 and Ser143, and the homocysteine portion of SAM forms hydrogen bonding interactions with the side chains of Asp175, Ser242, and His153 and the backbone carbonyl of Gly177. The LDXGXGXG motif indicative of SAM-utilizing proteins is found in NcsB1 as VDVGGGSG and is located between β3 and α14. We previously reported that NcsB1 alkylates substrate **3** with a variety of SAM analogues, including *S*-ethyl and *S*-*n*-propyl, to the corresponding 7-alkyl ether with reasonable efficiency (21). The side chain closest to the methyl group on SAM is the sterically small Ala243 (3.5 Å distance). On the basis of the large size and flexibility of the substrate binding pocket, it is not surprising that NcsB1 could accommodate larger SAM analogues.

Naphthoic Acid Binding Pocket. The naphthoic acid binding site is located at the juncture of helices from all three subdomains, including α7, α10, α11, α16, and α18, and the pocket is lined primarily with hydrophobic/aromatic residues (Figure 4A). Three methionine side chain thioethers (Met150, -286, and -290) are present on either side of the naphthoate ring forming van der Waals interactions with the substrate. The carboxylate and 2-hydroxyl groups of the substrate interact with Arg11 directly and with Asp157 through an ordered water. This interaction represents the only hydrogen bonds between the enzyme and substrate and appears to anchor the naphthoic acid into the active site adjacent to the bound SAM. To probe the importance of Arg11 in enzyme activity, we used site-directed mutagenesis to alter this site and evaluate the activity. This residue was mutated to Ala and Lys to assess the requirement of this hydrogen bonding interaction for binding naphthoic acids. Unexpectedly, the Arg11Ala mutant still effectively methylated naphthoic acid **3** (Table 2), although the results showed a doubling of *K*_M that was countered by a similar increase in *k*_{cat}. The Arg11Lys mutant had 110% of the catalytic efficiency compared to the WT. This mutant likewise showed an increased rate of turnover. This observation suggests that the specific

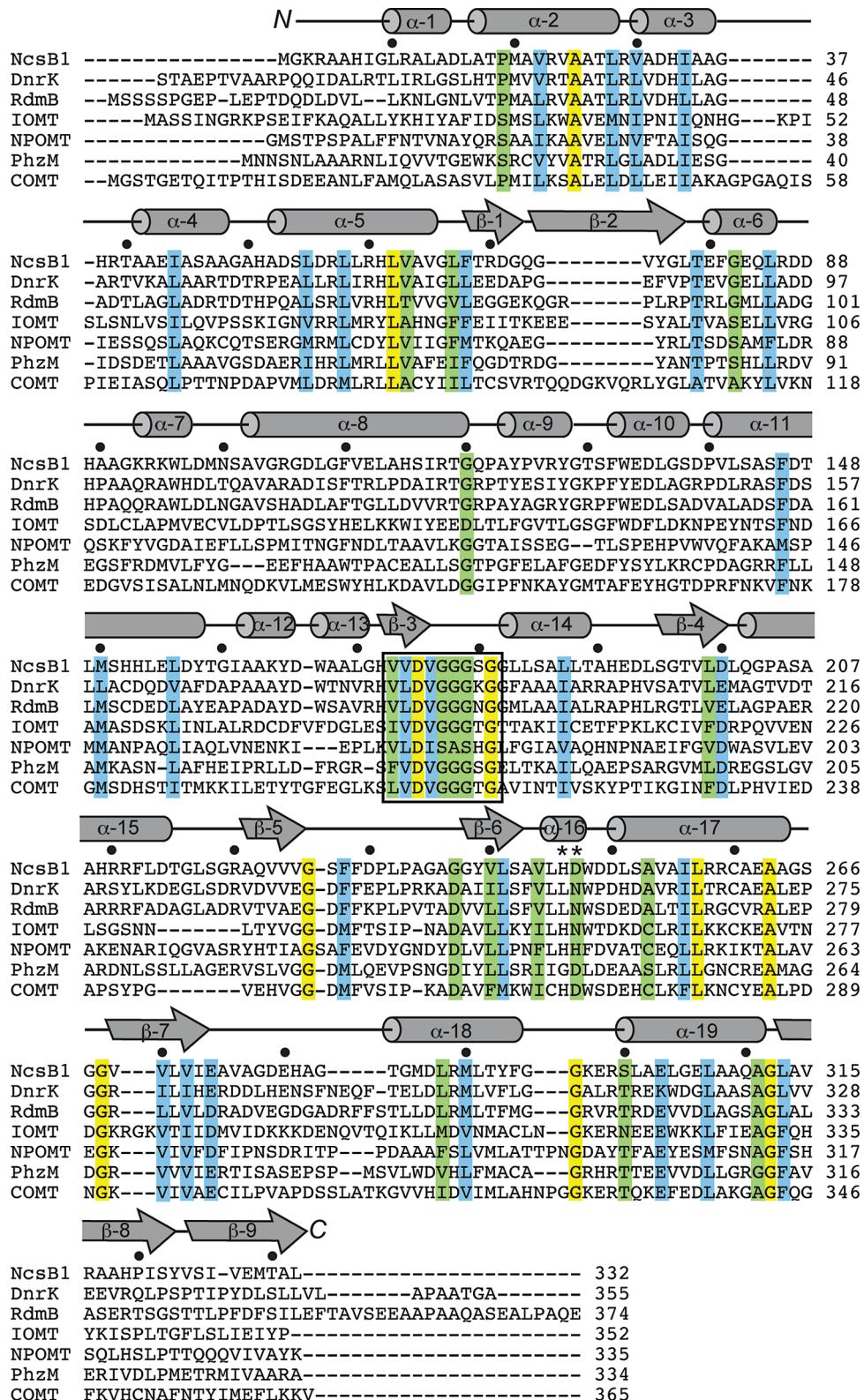


FIGURE 3: Sequence alignment of NcsB1 structural homologues as determined by the DALI structural alignment server. Abbreviations used with accession codes from GenBank in parentheses: DnrK, *S. peucetius* O-MTase (Q06528); RdmB, *S. purpurascens* hydroxylase (Q54527); IOMT, alfalfa O-MTase (O24529); NPOMT, *N. punctiforme* putative O-MTase (ZP_00112478); PhzM, *P. aeruginosa* N-MTase (Q9HW2); COMT, alfalfa O-MTase (P28002). Completely conserved residues are highlighted in yellow, highly conserved residues in blue, and conserved residues in green. Catalytic residues are denoted with asterisks, and the conserved glycine region is boxed.

hydrogen bond from Arg11 is not entirely crucial for substrate turnover and can be offset by a similar interaction with Lys or by an increased overall rate of catalytic turnover. In addition, simple benzoic acids lacking a hydroxyl group adjacent to the aryl acid [for example, 3-hydroxybenzoic acid (see Figure S4)] were

methylated by NcsB1, although at a significantly decreased rate, further suggesting that the 1,2-hydroxy acid is not absolutely required for enzyme-catalyzed chemistry.

NcsB1 Residues Involved in Methyltransferase Chemistry. Roles for residues in the active site relevant for catalysis

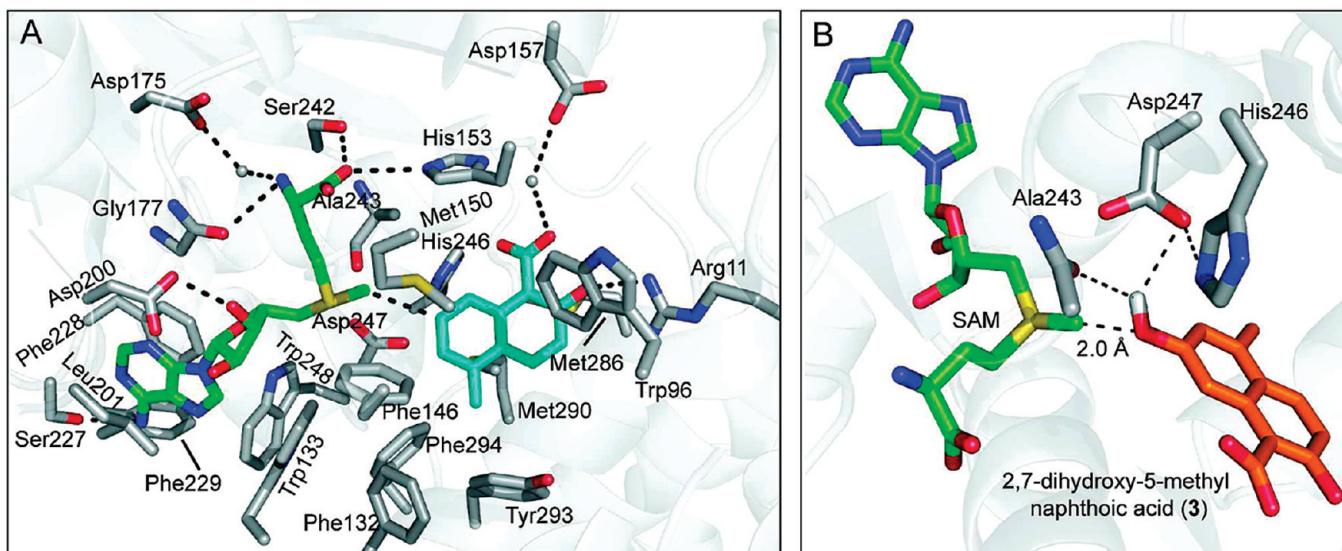


FIGURE 4: Active site of the NcsB1–SAM–2 complex. (A) Substrate and SAM binding pocket. SAH is colored green and naphthoic acid **2** cyan; residues in the substrate and SAM binding pockets are colored gray with pertinent hydrogen bonds indicated. (B) Close-up of the active site in the NcsB1–SAM–2 structure with the 7-hydroxy group modeled onto **2**, with the likely catalytic residues colored gray. SAM is colored green and naphthoic acid **3** orange. The methyl group on SAM is 2.0 Å from the hydroxyl group on **3**.

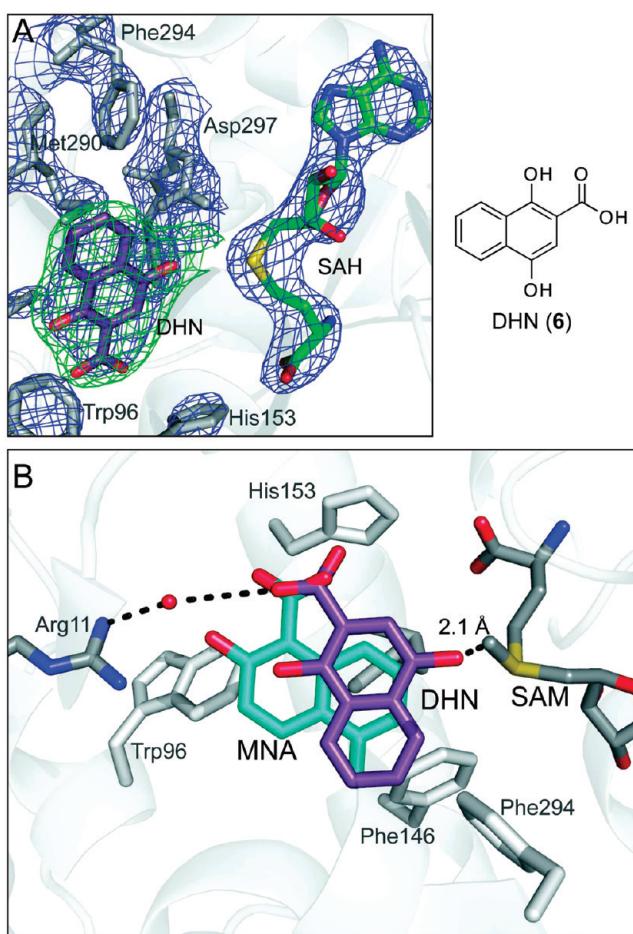


FIGURE 5: Binding of an alternate substrate, DHN (**6**). (A) Electron density map of the NcsB1–SAH–**6** active site, generated without **6**: $2F_o - F_c$ map (blue, 1.5σ) and $F_o - F_c$ map (green, 3σ). (B) Overlay of **2** and **6** in the NcsB1 active site. Naphthoic acid **6** is hydrogen bonded to Arg11 via a water, and the 4-hydroxyl group is sufficiently close to SAM (2.1 Å, SAM from the NcsB1–SAM–**2** structure) for methyl transfer to occur.

could include activation of the phenol for nucleophilic attack via acid–base chemistry. Histidine residues specifically have been

Table 2: Kinetics of NcsB1 and Mutant Constructs

protein–substrate	K_M (μM)	k_{cat} (min ⁻¹)	k_{cat}/K_M (relative to WT/ 3)
WT– 3 (<i>I8</i>)	206 ± 49	0.69 ± 0.05	1
Tyr293Ile– 3	649 ± 9	0.76 ± 0.04	0.35
Arg11Trp– 3	400 ± 68	0.60 ± 0.03	0.45
Arg11Ala– 3	419 ± 61	1.24 ± 0.06	0.88
Arg11Lys– 3	319 ± 40	1.18 ± 0.05	1.10

implicated to play this role in RebM, a carbohydrate *O*-MTase in rebeccamycin biosynthesis, where mutation of two histidine residues in the substrate binding pocket to alanines led to a complete loss of activity, while single mutations showed a marked decrease in activity (32). In NcsB1, three residues are in the proximity of the substrate phenol; the diad of His246 and Asp247 side chains and the backbone carbonyl of Ala243 may function to aid in activation and/or proton shuttling (Figure 4B).

Binding of 1,4-Dihydroxy-2-naphthoic Acid in the Active Site of NcsB1. Previous exploration of the substrate specificity of NcsB1 suggested a large degree of flexibility for naphthoate binding (21). A model was proposed that involved anchoring of the substrate by the 1,2-hydroxy acid motif and subsequent methylation of exposed phenols in the proximity of bound SAM. To probe this hypothesis, we determined the structure of NcsB1 bound to 1,4-dihydroxy-2-naphthoic acid (DHN, **6**). Methylation of this substrate would necessitate binding in a conformation distinct from the natural substrate while retaining the interaction with the enzyme through the 1,2-hydroxy acid. In the cocrystal structure, the proximity and orientation of **6** to SAH and important hydrogen bonding partners were readily apparent (Figure 5A). The substrate is bound in an appropriate position for regioselective methylation by SAM with an alternate orientation of the bound naphthoic acid. The methyl group on SAM is ~2.1 Å from the 5-hydroxyl group of **6**, and the carboxylate is hydrogen bonded to Arg11 via an ordered water (Figure 5B).

Probing the Determinants of the Substrate Specificity for NcsB1. The biosynthetic gene cluster for the related enediyne C-1027 contains a gene (*sgcD4*) that is highly homologous to

ncsB1 (56% identical, 67% similar), and the enzyme is proposed to methylate a benzoxazolinate (Figure S3) (6, 48). Of nine residues in NcsB1 identified to play a role in substrate binding, four have identical counterparts in SgcD4 (Trp96, Phe146, Met286, and Met290), two are similar (Met150_{NcsB1} is a Leu in SgcD4, and Phe294_{NcsB1} is a Tyr), and two residues are distinct (Arg11_{NcsB1} is a Trp and Tyr293_{NcsB1} is an Ile). To probe the substrate specificity determinants of NcsB1, we introduced two rational mutations at the residues that significantly differed structurally between NcsB1 and SgcD4: Tyr293Ile and Arg11Trp. These two additional mutant constructs were assayed against the natural substrate (**3**) (Table 2). Both mutants methylated **3**, though they showed significant increases in K_M compared to that of the wild-type enzyme. A larger decrease in catalytic efficiency was seen with the Tyr293Ile mutant, which had a K_M approximately 3 times higher than that of the wild type. The K_M was doubled for Arg11Trp. The ability of Arg11 mutants, especially the Trp and Ala mutants, to methylate naphthoic acid suggests that the aromatic binding pocket of NcsB1 is substantially more crucial to binding substrate than the hydrogen bond from Arg11. The reduced aromatic character of the Tyr293Ile mutant supports this theory. We next tested a simple substrate analogue of the benzoxazolinate, 2,5-dihydroxybenzoic acid. In this simple model compound, methylation occurs at the 5 position, although with significantly lower efficiency (Figure S3). 2,5-Dihydroxybenzoic acid was assayed with the panel of mutants in Table 2, and indeed, alternate specificity was observed as compared to wild-type NcsB1 (Table S3). For example, although the relative measured activities were low, contrary to that of the wild-type enzyme, the Arg11Trp mutant has higher relative activity toward 2,5-dihydroxybenzoic acid and the Arg11Lys was less efficient. These results implicate Arg11 in substrate discrimination for these related methyltransferases.

In summary, we have determined the crystal structures of NcsB1 in two conformations with SAH/SAM bound, with and without a substrate analogue or product. These structures revealed a large displacement of the C-terminal domain, a movement that likely opens up the active site for naphthoate binding. Additionally, the ternary complex structure of 1,4-dihydroxynaphthoic acid and SAH bound to NcsB1 was determined and exhibited a rotation of this alternate substrate in the binding pocket, allowing for methylation of the hydroxyl group at the 4 position. These results led us to probe substrate binding using active site mutants, demonstrating altered substrate specificity and revealing the importance of key residues in substrate binding.

ACKNOWLEDGMENT

We thank Tim Montavon for critical reading of the manuscript and members of the Bruner and Shen groups for helpful discussions.

SUPPORTING INFORMATION AVAILABLE

Crystallization conditions (Table S1), determination of the biologically active unit of NcsB1 (Figure S1), a table containing results from the DALI structural homologue search (Table S2), electron density maps (Figure S2), alignment of NcsB1 with SgcD4 (Figure S3), and detailed methods and results for benzoic acid substrates (Figure S4 and Table S3). This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Shen, B., Liu, W., and Nonaka, K. (2003) Enediyne natural products: Biosynthesis and prospect towards engineering novel antitumor agents. *Curr. Med. Chem.* **10**, 2317–2325.
- Maeda, H., Aikawa, S., and Yamashita, A. (1975) Subcellular fate of protein antibiotic neocarzinostatin in culture of a lymphoid cell line from Burkitt's lymphoma. *Cancer Res.* **35**, 554–559.
- Liu, W., Ahlert, J., Gao, Q., Wendt-Pienkowski, E., Shen, B., and Thorson, J. S. (2003) Rapid PCR amplification of minimal enediyne polyketide synthase cassettes leads to a predictive familial classification model. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11959–11963.
- Liu, W., Nonaka, K., Nie, L., Zhang, J., Christenson, S. D., Bae, J., Van Lanen, S. G., Zazopoulos, E., Farnet, C. M., Yang, C. F., and Shen, B. (2005) The neocarzinostatin biosynthetic gene cluster from *Streptomyces carzinostaticus* ATCC 15944 involving two iterative type I polyketide synthases. *Chem. Biol.* **12**, 293–302.
- Udwary, D. W., Zeigler, L., Asolkar, R. N., Singan, V., Lapidus, A., Fenical, W., Jensen, P. R., and Moore, B. S. (2007) Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 10376–10381.
- Liu, W., Christenson, S. D., Standage, S., and Shen, B. (2002) Biosynthesis of the enediyne antitumor antibiotic C-1027. *Science* **297**, 1170–1173.
- Van Lanen, S. G., Oh, T. J., Liu, W., Wendt-Pienkowski, E., and Shen, B. (2007) Characterization of the maduropeptin biosynthetic gene cluster from *Actinomadura madurae* ATCC 39144 supporting a unifying paradigm for enediyne biosynthesis. *J. Am. Chem. Soc.* **129**, 13082–13094.
- Ahlert, J., Shepard, E., Lomovskaya, N., Zazopoulos, E., Staffa, A., Bachmann, B. O., Huang, K., Fonstein, L., Czisny, A., Whitwam, R. E., Farnet, C. M., and Thorson, J. S. (2002) The calicheamicin gene cluster and its iterative type I enediyne PKS. *Science* **297**, 1173–1176.
- Povirk, L. F., Dattagupta, N., Warf, B. C., and Goldberg, I. H. (1981) Neocarzinostatin chromophore binds to deoxyribonucleic acid by intercalation. *Biochemistry* **20**, 4007–4014.
- Lee, S. H., and Goldberg, I. H. (1989) Sequence-specific, strand-selective, and directional binding of neocarzinostatin chromophore to oligodeoxyribonucleotides. *Biochemistry* **28**, 1019–1026.
- Kappen, L. S., Ellenberger, T. E., and Goldberg, I. H. (1987) Mechanism and base specificity of DNA breakage in intact cells by neocarzinostatin. *Biochemistry* **26**, 384–390.
- Zhang, J., Van Lanen, S. G., Ju, J., Liu, W., Dorrestein, P. C., Li, W., Kelleher, N. L., and Shen, B. (2008) A phosphopantetheinylating polyketide synthase producing a linear polyene to initiate enediyne antitumor antibiotic biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1460–1465.
- Urbaniak, M. D., Bingham, J. P., Hartley, J. A., Woolfson, D. N., and Caddick, S. (2004) Design and synthesis of a nitrogen mustard derivative stabilized by apo-neocarzinostatin. *J. Med. Chem.* **47**, 4710–4715.
- Caddick, S., Musckett, F. W., Stoneman, R. G., and Woolfson, D. N. (2006) Synthetic ligands for apo-neocarzinostatin. *J. Am. Chem. Soc.* **128**, 4204–4205.
- Baker, J. R., Woolfson, D. N., Musckett, F. W., Stoneman, R. G., Urbaniak, M. D., and Caddick, S. (2007) Protein-small molecule interactions in neocarzinostatin, the prototypical enediyne chromoprotein antibiotic. *ChemBioChem* **8**, 704–717.
- Schapit, B., Oh, T. J., Lamichhane, R., Liou, K., Lee, H. C., Kim, C. G., and Sohng, J. K. (2004) Neocarzinostatin naphthoate synthase: An unique iterative type I PKS from neocarzinostatin producer *Streptomyces carzinostaticus*. *FEBS Lett.* **566**, 201–206.
- Weitnauer, G., Muhlenweg, A., Trefzer, A., Hoffmeister, D., Sussmuth, R. D., Jung, G., Welzel, K., Vente, A., Girreser, U., and Bechthold, A. (2001) Biosynthesis of the orthosomycin antibiotic avilamycin A: Deductions from the molecular analysis of the avi biosynthetic gene cluster of *Streptomyces viridochromogenes* Tu57 and production of new antibiotics. *Chem. Biol.* **8**, 569–581.
- Jia, X. Y., Tian, Z. H., Shao, L., Qu, X. D., Zhao, Q. F., Tang, J., Tang, G. L., and Liu, W. (2006) Genetic characterization of the chlorothricin gene cluster as a model for spiroteretone antibiotic biosynthesis. *Chem. Biol.* **13**, 575–585.
- Zhao, Q., He, Q., Ding, W., Tang, M., Kang, Q., Yu, Y., Deng, W., Zhang, Q., Fang, J., Tang, G., and Liu, W. (2008) Characterization of the azinomycin B biosynthetic gene cluster revealing a different iterative type I polyketide synthase for naphthoate biosynthesis. *Chem. Biol.* **15**, 693–705.
- Cooke, H. A., Zhang, J., Griffin, M. A., Nonaka, K., Van Lanen, S. G., Shen, B., and Bruner, S. D. (2007) Characterization of NcsB2 as a promiscuous naphthoic acid/coenzyme A ligase integral to the

- biosynthesis of the enediyne antitumor antibiotic neocarzinostatin. *J. Am. Chem. Soc.* 129, 7728–7729.
21. Luo, Y., Lin, S., Zhang, J., Cooke, H. A., Bruner, S. D., and Shen, B. (2008) Regiospecific O-methylation of naphthoic acids catalyzed by NcsB1, an O-methyltransferase involved in the biosynthesis of the enediyne antitumor antibiotic neocarzinostatin. *J. Biol. Chem.* 283, 14694–14702.
 22. Schubert, H. L., Blumenthal, R. M., and Cheng, X. (2003) Many paths to methyltransfer: A chronicle of convergence. *Trends Biochem. Sci.* 28, 329–335.
 23. Vidgren, J., Svensson, L. A., and Liljas, A. (1994) Crystal structure of catechol O-methyltransferase. *Nature* 368, 354–358.
 24. Martin, J. L., and McMillan, F. M. (2002) SAM (dependent) I AM: The *S*-adenosylmethionine-dependent methyltransferase fold. *Curr. Opin. Struct. Biol.* 12, 783–793.
 25. Madduri, K., Torti, F., Colombo, A. L., and Hutchinson, C. R. (1993) Cloning and sequencing of a gene encoding carminomycin 4-O-methyltransferase from *Streptomyces peucetius* and its expression in *Escherichia coli*. *J. Bacteriol.* 175, 3900–3904.
 26. Jansson, A., Koskinen, H., Mantsala, P., Niemi, J., and Schneider, G. (2004) Crystal structure of a ternary complex of DnrK, a methyltransferase in daunorubicin biosynthesis, with bound products. *J. Biol. Chem.* 279, 41149–41156.
 27. Jansson, A., Koskinen, H., Erola, A., Wang, J., Mantsala, P., Schneider, G., and Niemi, J. (2005) Aclacinomycin 10-hydroxylase is a novel substrate-assisted hydroxylase requiring *S*-adenosyl-L-methionine as cofactor. *J. Biol. Chem.* 280, 3636–3644.
 28. Jansson, A., Niemi, J., Lindqvist, Y., Mantsala, P., and Schneider, G. (2003) Crystal structure of aclacinomycin-10-hydroxylase, a *S*-adenosyl-L-methionine-dependent methyltransferase homolog involved in anthracycline biosynthesis in *Streptomyces purpurascens*. *J. Mol. Biol.* 334, 269–280.
 29. Oster, L. M., Lester, D. R., Terwisscha van Scheltinga, A., Svenda, M., van Lun, M., Genereux, C., and Andersson, I. (2006) Insights into cephamycin biosynthesis: The crystal structure of CmcI from *Streptomyces clavuligerus*. *J. Mol. Biol.* 358, 546–558.
 30. Hou, X., Wang, Y., Zhou, Z., Bao, S., Lin, Y., and Gong, W. (2007) Crystal structure of SAM-dependent O-methyltransferase from pathogenic bacterium *Leptospira interrogans*. *J. Struct. Biol.* 159, 523–528.
 31. Cho, J. H., Park, Y., Ahn, J. H., Lim, Y., and Rhee, S. (2008) Structural and functional insights into O-methyltransferase from *Bacillus cereus*. *J. Mol. Biol.* 382, 987–997.
 32. Singh, S., McCoy, J. G., Zhang, C., Bingman, C. A., Phillips, G. N., Jr., and Thorson, J. S. (2008) Structure and mechanism of the rebeccamycin sugar 4'-O-methyltransferase RebM. *J. Biol. Chem.* 283, 22628–22636.
 33. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326.
 34. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674.
 35. Bailey, S. (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D50*, 760–763.
 36. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
 37. Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR system. *Nat. Proc.* 2, 2728–2733.
 38. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905–921.
 39. Vagin, A., and Teplyakov, A. (1997) MOLREP: An automated program for molecular replacement. *J. Appl. Crystallogr.* 30, 1022–1025.
 40. Schuttelkopf, A. W., and van Aalten, D. M. F. (2004) PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D60*, 1355–1363.
 41. Kleywegt, G. J. (2007) Crystallographic refinement of ligand complexes. *Acta Crystallogr. D63*, 94–100.
 42. Sippl, M. J., and Wiederstein, M. (2008) A note on difficult structure alignment problems. *Bioinformatics* 24, 426–427.
 43. Sippl, M. J. (2008) On distance and similarity in fold space. *Bioinformatics* 24, 872–873.
 44. Holm, L., Kaarainen, S., Rosenstrom, P., and Schenkel, A. (2008) Searching protein structure databases with DaliLite. *Bioinformatics* 24, 2780–2781.
 45. Zubietta, C., He, X. Z., Dixon, R. A., and Noel, J. P. (2001) Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant O-methyltransferases. *Nat. Struct. Biol.* 8, 271–279.
 46. Parsons, J. F., Greenhagen, B. T., Shi, K., Calabrese, K., Robinson, H., and Ladner, J. E. (2007) Structural and functional analysis of the pyocyanin biosynthetic protein PhzM from *Pseudomonas aeruginosa*. *Biochemistry* 46, 1821–1828.
 47. Zubietta, C., Kota, P., Ferrer, J. L., Dixon, R. A., and Noel, J. P. (2002) Structural basis for the modulation of lignin monomer methylation by caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase. *Plant Cell* 14, 1265–1277.
 48. Van Lanen, S. G., Lin, S., and Shen, B. (2008) Biosynthesis of the enediyne antitumor antibiotic C-1027 involves a new branching point in chorismate metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 105, 494–499.