

Supporting Information

Oja et al. 10.1073/pnas.1207407110

SI Text

Purification and Structural Elucidation of Alnumycins P1 (4a) and P2 (4b). For preparative production of alnumycins P1 (4a) and P2 (4b), 60 0.4-mL analytical-scale AlnA reactions consisting of 75 mM D-ribose-5-phosphate, 0.5 mM prealuminumycin (2), and 23 μ M AlnA were incubated for 3 h at 288 K. After incubation, the combined mixtures were acidified with 2.5% (vol/vol) HCOOH, centrifuged, and applied to Sep-Pak PLUS C18 cartridges (Millipore, Waters) preequilibrated with 0.1% (vol/vol) HCOOH and eluted with methanol. The reaction extracts in methanol were then applied to a Discovery HSC18 column (5 μ m, 5 cm \times 4.6 mm; Supelco) and eluted with 20 mM aqueous ammonium acetate pH 3.6 and a 70–100% (vol/vol) methanol gradient to yield 0.5 mg of alnumycin P (4), ~50% pure by HPLC and NMR. For all other HPLC columns tested, the chromatographic behavior of 4 was exceedingly poor as the sample appeared to be strongly retained on the columns with only minor amounts of material eluting over extended periods of time even when using 100% (vol/vol) organic solvent. Unfortunately, the compound degraded readily in aqueous environments complicating its further purification.

The structural elucidation of 4 followed from ESI⁻-HR-MS, which provided an accurate mass of 495.1059 amu for the M – 1 ion alluding to a molecular formula of C₂₂H₂₅O₁₁P (calculated for the M – 1 ion, 495.1062 amu). The presence of an intact prealuminumycin moiety was inferred by the observation of (i) a chromophore in UV-vis spectroscopy pertaining to such and (ii) various protons and carbons (Table S3) that were consistent to previous observations of the prealuminumycin moiety with regard to both chemical shift and multiplicity (1, 2). Because of the limited sample amount, direct observation of carbon was not possible, and only indirect observation of the carbons was amenable via heteronuclear single quantum coherence (HSQC) and heteronuclear multiple bond correlation (HMBC) spectroscopy.

The inclusion of phosphorous in 4 as indicated by HR-MS was readily confirmed by ³¹P NMR. The presence of a phosphate group combined with a sugar residue—thus a highly polar entity—was also strongly implied by the aberrant chromatographic behavior and extraction properties (viz. negligible solubility in chloroform and water and only ready solubility in methanol) exhibited by the compound. The HR-MS result also implied the presence of a ring-closed ribose unit attached to a prealuminumycin moiety with an attendant phosphate group, presumably attached to the ribose unit, to account fully for the molecular formula. Indeed, a ribose unit in the furanose form with both β (4a, major isomer, 83.9%) and α (4b, 16.1%) anomers present in the sample was clearly indicated by ¹H NMR, with chemical shifts and couplings for the signals of H-1' akin to the analogous H-1' signals observed in alnumycins C1 (3a) and C2 (3b) (2), respectively. Although only one distinct ³¹P NMR signal was forthcoming, its multiplicity was determined to be a triplet ($J \sim 5.8$ Hz) by coupled acquisition, as analysis of the ¹H{³¹P}-HMBC spectrum was compromised by the extensive overlap of the ¹H signals in the sugar residue. Thus, the positioning of the phosphate group at C-5' in both 4a and 4b was based on this coupling together with the knowledge that O-4' must be blocked due to the furanose ring formation and the chemical shifts and couplings, where observable, of H-1', H-2', and H-3' did not indicate phosphate attachment to the pertinent oxygens. Finally, the attachment of the furanose ring to C-8 of the prealuminumycin skeleton and the attachment of the prealuminumycin moiety to C-1' of the ribose unit were evident by the long-range coupling

between H-7 and H-1' by long-range correlation spectroscopy (COSY) and decoupling experiments.

Of note, neither 3a nor 3b was discerned to be present in the sample by ¹H NMR, nor was any open-chain form of the ribose unit in effect based on the integration of the ¹H signals, e.g., H-1' and H-7 with respect to H-4. Furthermore, additional ³¹P NMR signals were not observed, thereby precluding the presence of any free phosphate in the sample.

Binding of D-Ribulose-5-Phosphate to AlnA. The binding of D-ribulose-5-phosphate to AlnA was studied by transfer of bulk water magnetization to the ligand (3, 4). The NMR samples consisted of the following: 4 mM D-ribulose-5-phosphate and 428 (124) μ M AlnA in 12.5 mM H₃PO₄ (pH 6.85), 37.5 mM NaCl, 1.25 mM MgCl₂, 15% (vol/vol) glycerol, 5% (vol/vol) D₂O, and 1 mM DMSO to yield substrate to protein ratios of 9.3:1 (32:1) (Fig. S5) and 2 mM D-ribulose-5-phosphate and 33.6 μ M AlnA in a similar solution to yield a substrate to protein ratio of 60:1. To prevent the possible spontaneous oxidation of the substrates, the samples were bubbled for ~10 min with 99.999% (vol/vol) argon gas and stored under argon. Control samples lacking the protein were prepared in an otherwise identical fashion. All samples were prepared and stored on ice. The waterLOGSY spectra were acquired using the pulse sequence of Dalvit (3, 4) and with an Aq of 1.64 s, 3.8 s for the postacquisition delay (PAD), and a mixing time (τ_m) of 750 ms; suppression of the water signal was effected by excitation sculpting (5) using a 2-ms sinc-shaped pulse, whereas selective excitation of the water signal was effected by a 5-ms Gaussian-shaped pulse; spectra were processed with zero-filling ($\times 4$) and the application of 2 Hz of line broadening.

Crystallization of AlnA and AlnB. The proteins AlnA and AlnB were produced and purified as described previously (2). For AlnA, diffraction quality cubic crystals were obtained by mixing 2 μ L of AlnA (7 mg/mL) in 25 mM Hepes (pH 7.2), 75 mM NaCl, 5% glycerol, and 5 mM MgCl₂ together with 1 μ L of well solution containing 100 mM malonic acid–imidazole–boric acid (MIB) buffer (pH 7.2), 200 mM calcium acetate, and 20% (vol/vol) PEG3000 at 277 K. Thick rod shape crystals of AlnB were obtained by mixing 1 μ L of AlnB (13 mg/mL) in 50 mM Hepes (pH 7.2), 150 mM NaCl, 10 mM MgCl₂, and 10% (vol/vol) glycerol together with 2 μ L of well solution, which consisted of 30% (vol/vol) PEG1500 and 100 mM MIB buffer (pH 7.0) at 277 K. The native crystals were cryo-protected by quick soaking in mother liquor supplemented with 20% (vol/vol) (AlnA) or 25% (vol/vol) (AlnB) PEG600 and flash frozen at 100 K. The phosphate complex of AlnB was obtained by adding 10 mM sodium phosphate into the cryo solution. The D-ribose-5-phosphate complex of AlnA was obtained by soaking the crystal for 40 min in a cryo solution containing 100 mM MIB buffer (pH 7.2), 25% (vol/vol) PEG3000, 100 mM D-ribose-5-phosphate, and 20% (vol/vol) PEG400.

Generation of AlnA and AlnB Enzyme Variants. AlnA variants K86A, H130A, and K159A were prepared using the plasmid pBla7AB (1) as a template. The first round of PCR was accomplished with the primer AlnAmutfor with each reverse mutagenesis primer and AlnAmutrev with each forward mutagenesis primer (Table S2). PCR products isolated from a preparative agarose gel were used as templates for the second round of PCR with AlnAmutfor and -rev as a primer pair, and the product fragments were digested with Eco8II and PaeI and cloned into a similarly digested pBla7AB. For protein overexpression, the constructs were di-

gested with SgrAI and Bpu1102I and cloned into a similarly digested pBADHBdaA (2).

For generation of AlnA variants E29A, E29Q, D138A, and K159R, a similar four-primer approach was used with pBADHBdaA (2) as a template in the first round of PCR. The forward sequencing primer pBADseqfor was used with each reverse mutagenesis primer, and pBADseqrev with each forward mutagenesis primer (Table S2). The PCR products were isolated and used as templates for the second round of PCR with pBADseqfor and -rev as a primer pair. The products were then digested with NcoI and HindIII and cloned into a similarly digested pBADHBd (6).

The AlnB variants D17A, K119A, and K119R were prepared using the plasmid pBla7AB (1) as a template. The first round of PCR was accomplished with the primer pBla7ABfor together with each reverse mutagenesis primer and pBla7ABrev with each forward mutagenesis primer (Table S2). The PCR products isolated from a preparative agarose gel were used as templates for

the second round of PCR with pBla7ABfor and -rev as a primer pair. The product fragments were digested with PaeI and XbaI and cloned into a similarly digested pBla7AB and further as XbaI + HindIII digested insert fragments into a similarly digested pJ487 vector (7). The plasmid DNA isolated from *Streptomyces lividans* TK24 was then used as a PCR template with alnBBfor and alnBErev as a primer pair, and the BglII + EcoRI digested PCR products were finally cloned into a similarly digested pBADHBd.

The AlnB variants D15A, D15N, and Y79A were prepared using the protein production construct pBADHBdaB (2) as a template. The pBADseqfor sequencing primer was used with the reverse mutagenesis primers D15Arev and D15Nrev, whereas alnBBfor was used with the Y79Arev primer and the alnBErev primer with each forward mutagenesis primer (Table S2) for the first round of PCR. With the isolated PCR products as templates, the second round was accomplished using alnBBfor and alnBErev as primers. The PCR products were again digested with BglII and EcoRI and cloned into a similarly digested pBADHBd.

- Oja T, et al. (2008) Characterization of the alnumycin gene cluster reveals unusual gene products for pyran ring formation and dioxan biosynthesis. *Chem Biol* 15(10):1046–1057.
- Oja T, et al. (2012) Biosynthetic pathway toward carbohydrate-like moieties of alnumycins contains unusual steps for C-C bond formation and cleavage. *Proc Natl Acad Sci USA* 109(16):6024–6029.
- Dalvit C (1996) Homonuclear 1D and 2D NMR experiments for the observation of solvent-solute interactions. *J Magn Reson B* 112(3):282–288.
- Dalvit C, et al. (2000) Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *J Biomol NMR* 18(1):65–68.
- Hwang T-L, Shaka AJ (1995) Water suppression that works. Excitation sculpting using arbitrary waveforms and pulsed field gradients. *J Magn Reson A* 112:275–279.
- Kallio P, Sultana A, Niemi J, Mäntsälä P, Schneider G (2006) Crystal structure of the polyketide cyclase AknH with bound substrate and product analogue: Implications for catalytic mechanism and product stereoselectivity. *J Mol Biol* 357(1):210–220.
- Hopwood DA, et al. (1985) *Genetic Manipulation of Streptomyces: A Laboratory Manual* (John Innes Foundation, Norwich, UK).

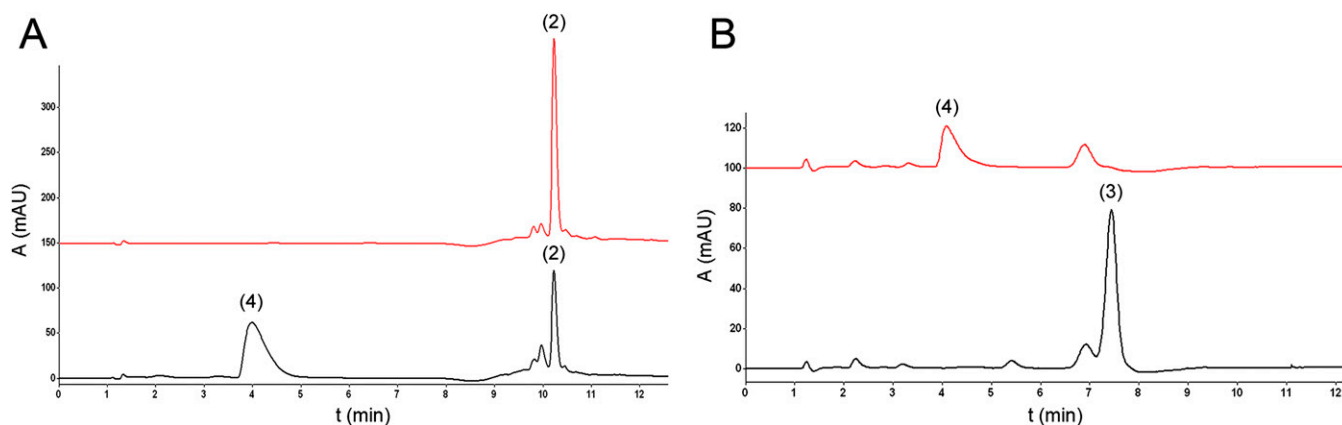


Fig. S1. Enzymatic synthesis of (A) alnumycin P (4) from 100 mM D-ribose-5-phosphate and 1 mM prealuminum (2) by 15 μ M AlnA (3 h incubation) and (B) enzymatic conversion of 4 to alnumycin C (3) by 4 μ M AlnB (1-h incubation). In both cases, a parallel control reaction with boiled enzyme is shown in red. All HPLC traces shown are recorded at 470 nm.

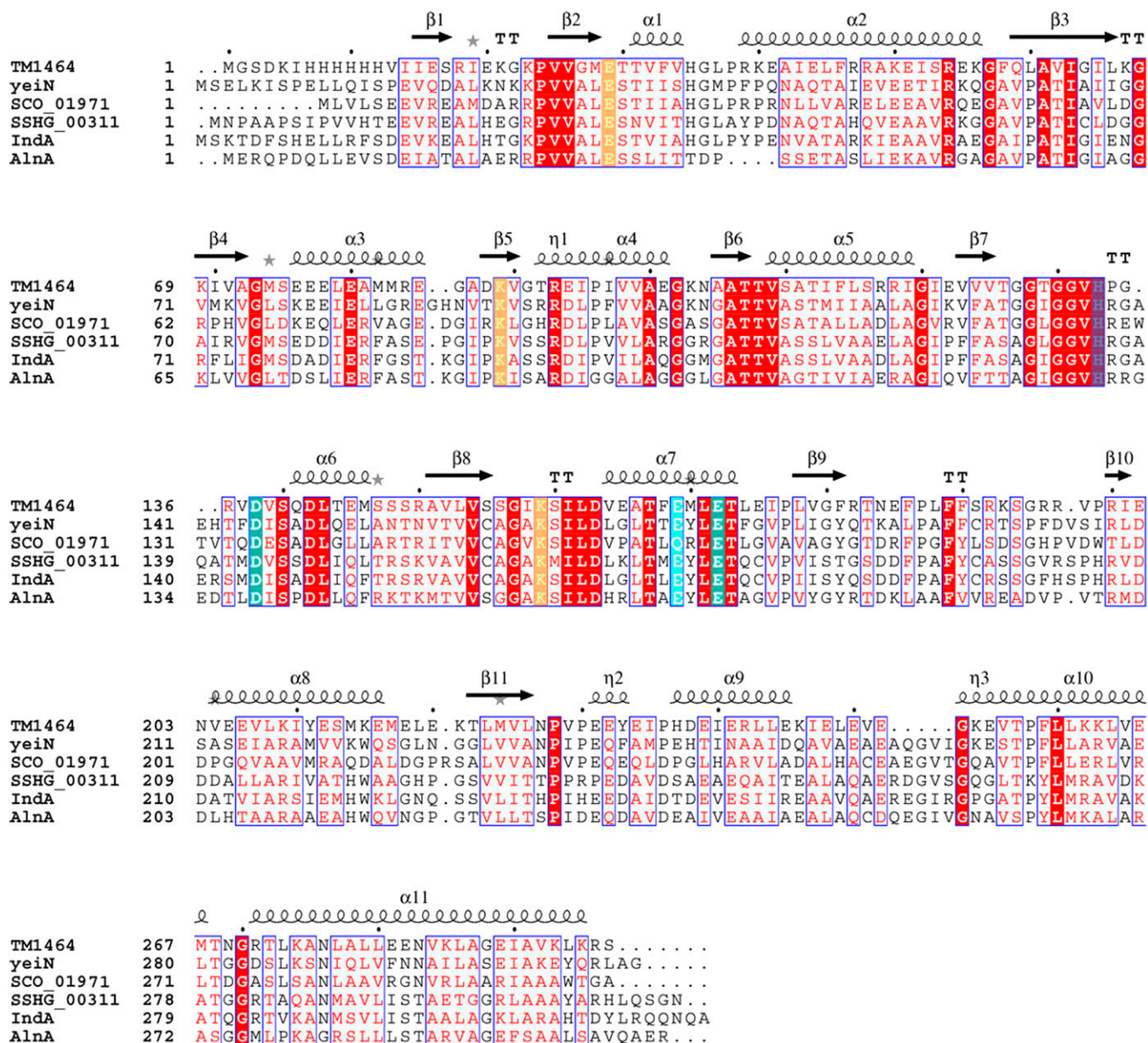


Fig. S7. Structure-based multiple sequence alignment of AlnA with related homologous enzymes. The proteins are pseudouridine glycosidase TM1464, *Thermotoga maritima*; pseudouridine glycosidase YeiN, *Escherichia coli*; protein of unknown function SCO_01971, *S. coelicolor*; protein of unknown function SSHG_00311, *S. albus* and protein putatively involved in the biosynthesis of the pigment indigoidine IndA, *Erwinia chrysanthemi*. Amino acid residues that are conserved (red line), involved in catalysis (orange line), coordinating to the metal ion (cyan line), or the phosphate group of the cosubstrate (purple line) are shown.

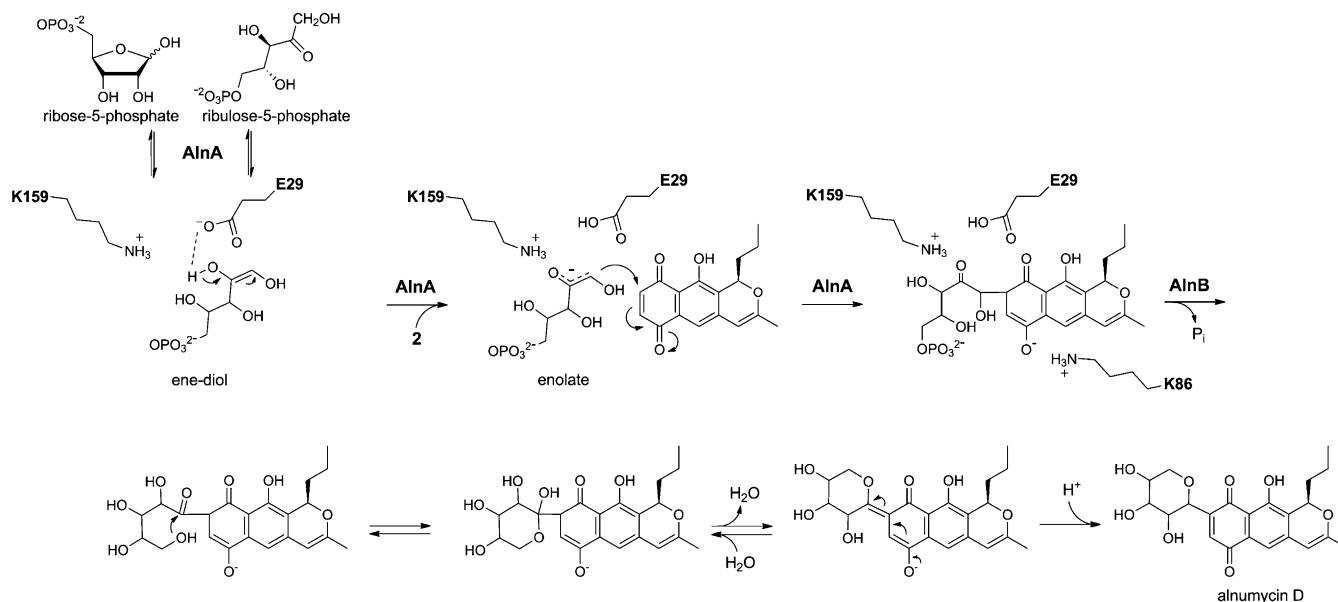


Fig. S8. Mechanistic model for formation of the shunt product alnumycin D by AlnA and AlnB. It is plausible that after dephosphorylation by AlnB, the reaction may be nonenzymatic.

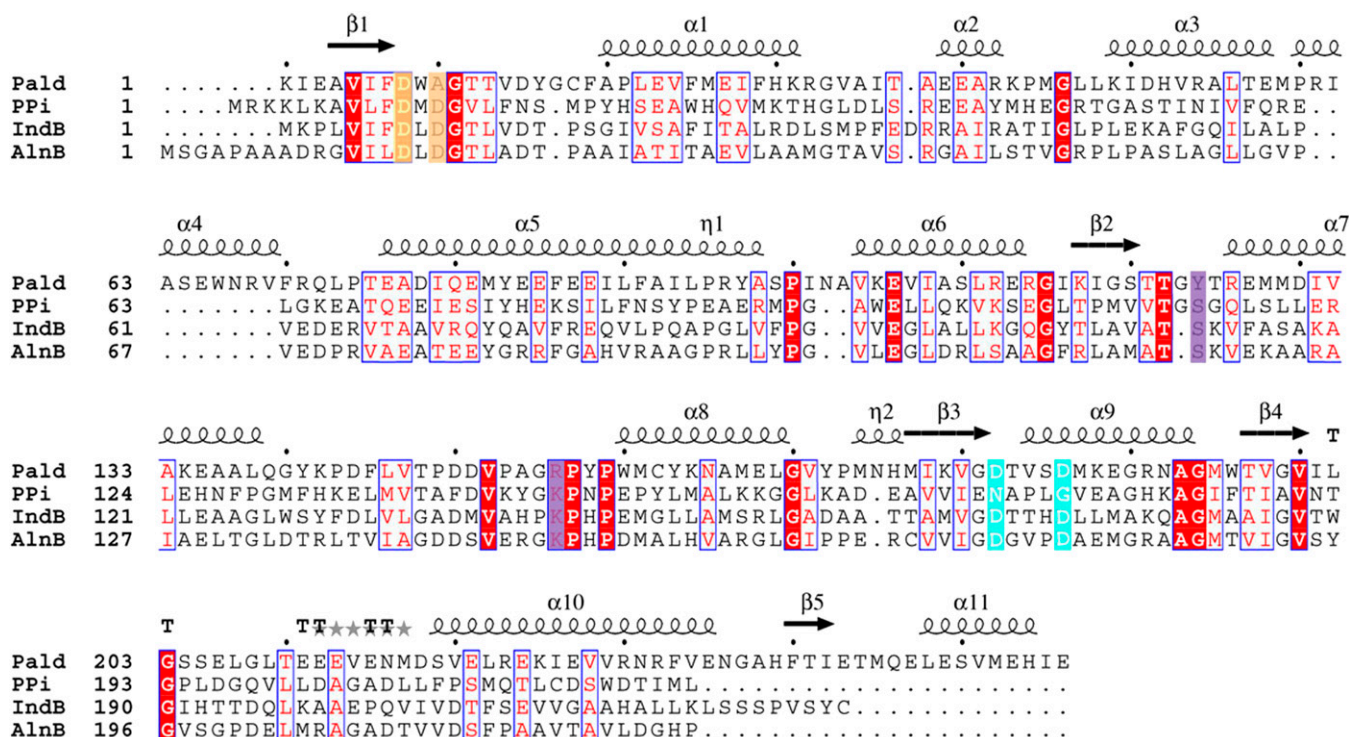


Fig. S9. Structure-based multiple sequence alignment of AlnB with related subfamily I members. The proteins included are the phosphonoacetaldehyde hydrolase Pald, *Bacillus cereus*; pyrophosphatase (PPI) BT2127, *Bacteroides thetaiotaomicron*; protein putatively involved in the biosynthesis of the pigment indigoidine IndB, *Erwinia chrysanthemi*. Amino acid residues that are conserved (red line), involved in catalysis (orange line), coordinating to the metal ion (cyan line), or the phosphate group of the cosubstrate (purple line) are shown. Only one of the catalytic aspartate residues is conserved in Pald (1).

1. Zhang G, et al. (2004) Investigation of metal ion binding in phosphonoacetaldehyde hydrolase identifies sequence markers for metal-activated enzymes of the HAD enzyme superfamily. *Biochemistry* 43(17):4990–4997.

Table S1. X-ray data collection and crystallographic refinement statistics

Data collection	AlnA (native)	AlnA-5RP	AlnB (native)	AlnB + Pi
PDB entry				
X-ray source	ESRF ID23-1	ESRF ID23-2	ESRF BM14	ESRF ID14-2
Space group	I4 ₁ 32	I4 ₁ 32	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit cell (Å)	<i>a</i> = <i>b</i> = <i>c</i> = 168.9	<i>a</i> = <i>b</i> = <i>c</i> = 168.8	48.0 62.9 63.2	48.3 63.2 63.5
Resolution (Å)	83.9–2.1 (2.21–2.10)	50–3.15 (3.32–3.15)	38.1–1.25 (1.32–1.25)	38.4–1.5 (1.58–1.50)
Wavelength (Å)	0.979	0.8726	1.072	0.933
No. of unique reflections	23665	7 414	46971	31272
Multiplicity	6.0 (6.1)	7.7 (7.9)	7.8 (5.8)	3.0 (3.8)
Completeness (%)	99.4 (100.0)	100.0 (100.0)	89.1 (53.4)	98.4 (90.1)
Mean < <i>I</i> >/<σ _{<i>I</i>} >	12.5 (2.8)	10.1 (2.5)	25.2 (7.5)	11.2 (3.8)
R-sym (%)*	7.8 (56.5)	15.4 (80.7)	4.7 (18.1)	11.4 (21.1)
Wilson B-factor (Å ²)	37.4	72.3	9.7	12.9
Refinement				
R-factor (all reflections) (%)	17.2	22.0	11.9	14.4
R-free (%) [†]	20.0	24.8	15.3	18.6
No. of atoms	2362	2211	2024	1879
No. of water molecules	136	2	360	234
No. of other molecules	2 CA, CL, SO ₄ , 2 PEG	5RP, 2 CA, CL, EPE	MG, BO ₃	MG, BO ₃ , PO ₄
RMSD bond lengths (Å)	0.022	0.004	0.020	0.011
RMSD bond angles (°)	1.668	0.873	1.842	1.515
Average B-factor (Å²)				
All atoms	47.9	76.9	14.0	14.4
Protein	47.8	77.1	10.6	12.6
Water/other molecules				
	47.3/CA 45.0/ CL 72.5/ SO4 27.5/PEG 70.7	58.1/5RP 67.2/ CA 76.4/ CL 71.7/EPE 62.6	28.0/MG 6.1/ BO ₃ 8.0	28.7/MG 5.6/ BO ₃ 9.3/PO ₄ 16.6
Ramachandran plot				
Favored regions (%)	294 (96.1)	295 (98.3)	199 (98.5)	197 (99.0)
Allowed regions (%)	11 (3.6)	5 (1.7)	3 (1.5)	2 (1.0)
Outliers (%)	1 (0.3)	0	0	0

Values in parenthesis are for the highest resolution shell. BO₃, boric acid; CA, calcium; CL, chloride; EPE, Hepes; MG, magnesium; PEG, polyethylene glycol; PO₄, phosphate; SO₄, sulfate; 5RP, D-ribulose-5-phosphate.

*R-sym = $|\sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / (\sum_h \sum_i I_i(h))$, where $I_i(h)$ is the *i*th measurement of reflection *h* and $\langle I(h) \rangle$ is the weighted mean of all measurements of *h*.

[†]5% of the reflections were used in the R-free calculations.

Table S2. Oligonucleotide primers used in this study for cloning and mutagenesis

Primer	Sequence (5'→3')
pBADseqfor	CGCAACTCTCTACTGTTTCTCC
pBADseqrev	GCGTTCTGATTTAATCTGTATCAGG
pBla7ABfor	TACCTGATGAAGGCGCTCGCCA
pBla7ABrev	CTAAAGGGAACAAAAGCTGGAGCTCC
alnAmutfor	GTTTCTCTTGAGGACCGCTGCGGC
alnAmutrev	CGCCTTGGGCAGCATGC
alnBBfor	CAG AGATCT AGCGGCGCCCCCGC
alnBErev	CAG GAATTC GTTACGGGTGTCCGTCCAG
alnAE29Afor	GTGGTGGCGCTGGCGTCTCGCTCATC
alnAE29Arev	GATGAGCGAGGACGCCAGCCACCAC
alnAE29Qfor	GTGGTGGCGCTGCAGTCTCGCTCATC
alnAE29Qrev	GATGAGCGAGGACTGCAGCGCCACCAC
alnAD138Afor	GAGGACACCCTCGCCATCTCCCCGAC
alnAD138Arev	GTCGGGGAGATGGCGAGGGTGTCTC
alnAK159Rfor	CTCCGCGGCGCGAGAAGCATCCTGGACC
alnAK159Rrev	GGTCCAGGATGCTTCTCGCGCCCGGAG
alnAK86Afor	GGGCATTCCGGCGATCAGCGCGC
alnAK86Arev	GCGCGCTGATCGCCGGAATGCC
alnAH130Afor	CGGCGGTGTGGCCCGAGGGGGG
alnAH130Arev	CGCCCTGCGGGCCACACCGCCG
alnAK159Afor	CGGCGGCGGGCGAGCATCCTGG
alnAK159Arev	CCAGGATGCTCGCCGCGCCGCG
alnBD15Afor	CCTCGCGCTCGACGGCACACTC
alnBD15Arev	GAGTGTGCCGTCGAGCGCGAGG
alnBD15Nfor	CCTCAACCTCGACGGCACACTC
alnBD15Nrev	GAGTGTGCCGTCGAGGTTGAGG
alnBD17Afor	CCTCGACCTCGCGGCACACTC
alnBD17Arev	GAGTGTGCCGCGAGGTCGAGG
alnBY79Afor	GACCGAGGAGGCGGGGCGGG
alnBY79Arev	CGCCGCCCCGCTCCTCGGTC
alnBK119Afor	CGACCTCGGCGGTGAGAGGCC
alnBK119Arev	GGCCTTCTCGACCCCGAGGTGC
alnBK119Rfor	CGACCTCGAGGTCGAGAGGCC
alnBK119Rrev	GGCCTTCTCGACCCCTCGAGGTGC

Recognition sites of the restriction endonucleases BglII and EcoRI used in cloning are shown in bold.

Table S3. Partial ¹H and ¹³C NMR assignments of alnumycins P1 (4a) and P2 (4b) in CD₃OD at 25 °C

Position	Alnumycin P1 (4a)				Alnumycin P2 (4b)			
	δ/ppm		Mult.	<i>J</i> _{H,H} /Hz (spin partner)	δ/ppm		Mult.	<i>J</i> _{H,H} /Hz (spin partner)
	¹³ C	¹ H			¹³ C	¹ H		
1	74.1	5.63	app dd	9.39 (H-11a); 3.56 (H-11b)	ol	ol		
3	159.5	—			ol	—		
4	101.0	5.74	qt	0.88 (H-14)	ol	ol		
4a	123.1	—			ol	—		
5	108.6	7.01	br s		ol	ol		
5a	133.2	—			135.4	—		
6	186.7	—			no	—		
7	135.4	7.11	d	1.66 (H-1')	135.8	6.93	d	1.90 (H-1')
8	151.3	—			no	—		
9	190.2	—			no	—		
9a	no	—			no	—		
10	no	—			no	—		
10a	114.9	—			ol	—		
11	35.9	2.02, 1.50	m		ol	ol		
12	19.3	1.57, 1.49	m		ol	ol		
13	14.1	0.97	app t	7.17 (2 × H-12)	ol	ol		
14	20.3	1.95	d	0.88 (H-4)	ol	ol		
1'	80.7	5.01	dd	3.48 (H-2'); 1.66 (H-7)	79.3	5.20	dd	3.63 (H-2'); 1.90 (H-7)
2'	77.7	4.10	ol		74.7	4.49	dd	4.58 (H-3'); 3.63 (H-1')
3'	72.5	4.09	ol		73.9	4.42	dd	8.16 (H-4'); 4.58 (H-2')
4'	83.4	4.10	ol		82.2	4.05	ol	
5'	66.5	4.14, 4.08	ol		65.7	4.14, 4.08	ol	
P	4.75	—		ho t, ~ 5.8	ol	—		ol

The chemical shifts of ¹H and ¹³C nuclei are reported relative to TMS as an internal standard (δ = 0 ppm for both ¹H and ¹³C) and externally to 90% H₃PO₄ in D₂O for ³¹P. app, apparent; br, broad or broadened; d, doublet; dd, doublet of doublets; ho, higher order; ol, overlapped; m, multiplet; no, not observed; qt, quartet; s, singlet; t, triplet.