Proteomics-Based Discovery of Koranimine, a Cyclic-Imine Natural Product

Bradley S. Evans[†], Ioanna Ntai[‡], Yunqiu Chen[‡], Sarah J. Robinson[‡] and Neil L. Kelleher[‡]*

[†]University of Illinois Urbana-Champaign, The Institute of Genomic Biology, 1206 W. Gregory Drive, Urbana, IL 61801 and the [‡]Department of Chemistry and the Chemistry of Life Processes Institute, Northwestern University, 2145 N. Sheridan Road, Evanston, IL 60208, United States

Corresponding Author email: n-kelleher@northwestern.edu

Supporting Information

1. Supplementary Figures and Legends

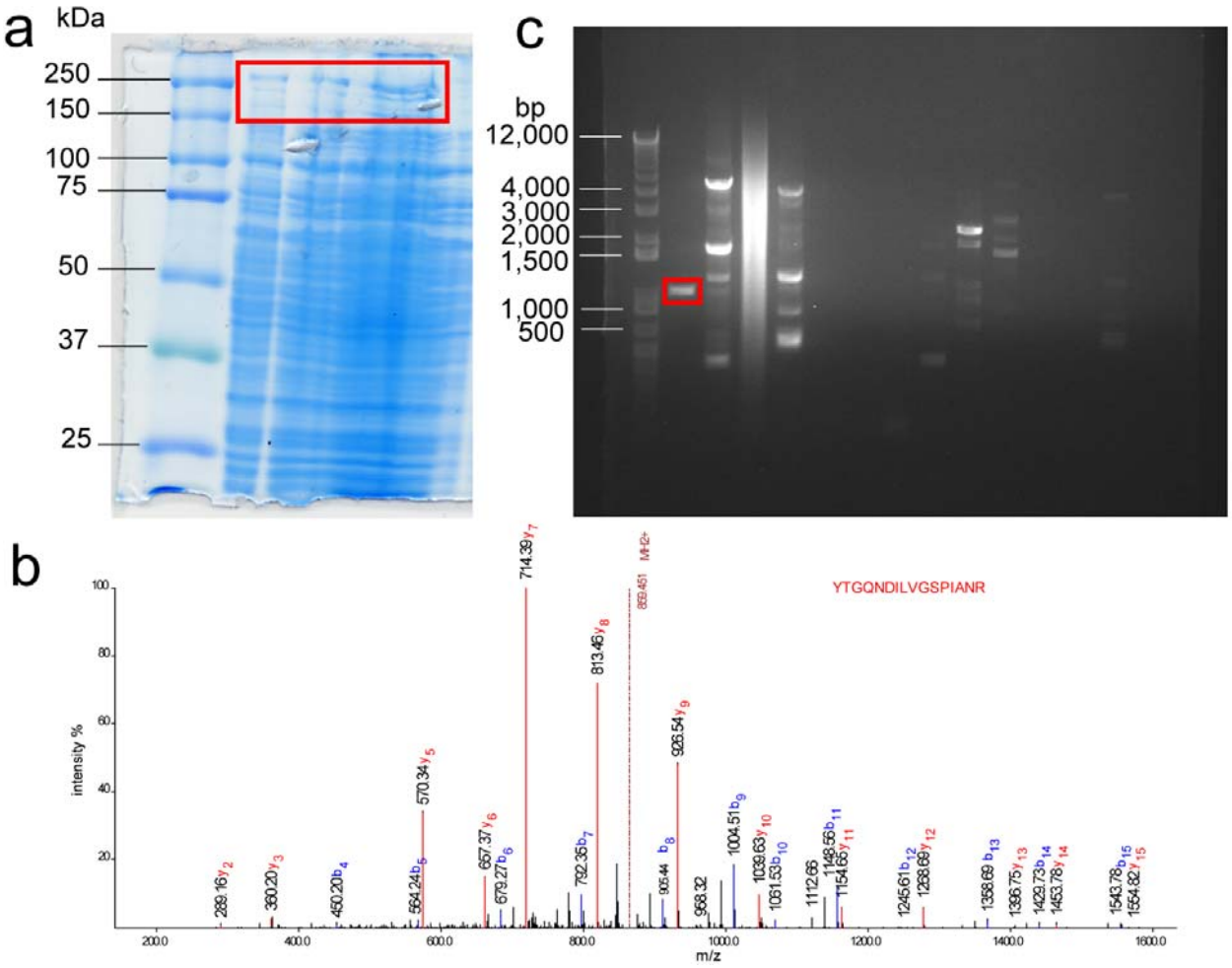

Figure S1. **PrISM-workflow for strain NK2003. a,** SDS-PAGE gel showing high-molecular weight protein bands analyzed by nanocapillary LC-MS/MS (highlighted in red box). **b,** Representative mass spectrum corresponding to a peptide sequence used in degenerate primer design. **c,** Agarose gel of PCRs conducted using peptide-based degenerate primers coupled to degenerate primers from conserved regions of adenylation domains; the highlighted band was excised and sequenced.

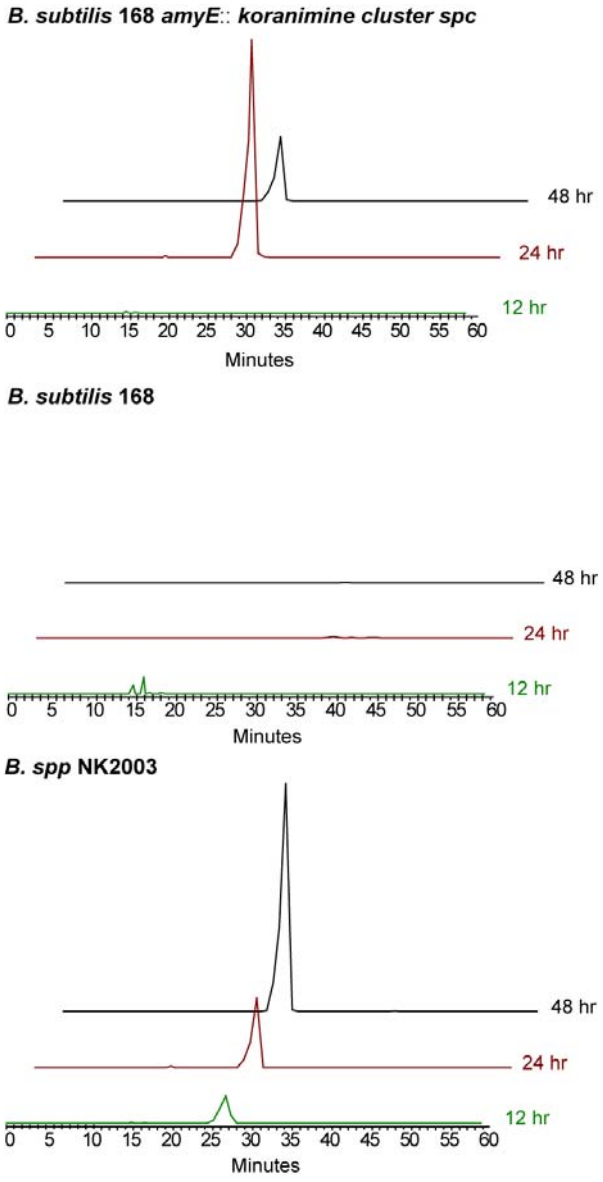

Figure S2. **Heterologous expression of koranimine gene cluster.** Selected ion chromatograms for *m/z* 804.502 for culture supernatant samples of indicated strain grown in Müeller-Hinton broth for indicated time at 30ºC and 250 rpm. There is a time-dependent increase in the peak at *m/z* 804.500 for the native producer (*Bacillus spp.* NK2003) and the heterologous host (*B. subtilis* 168 *amyE:: NE2-2 cluster spc*), but not the host background (*B. subtilis* 168).

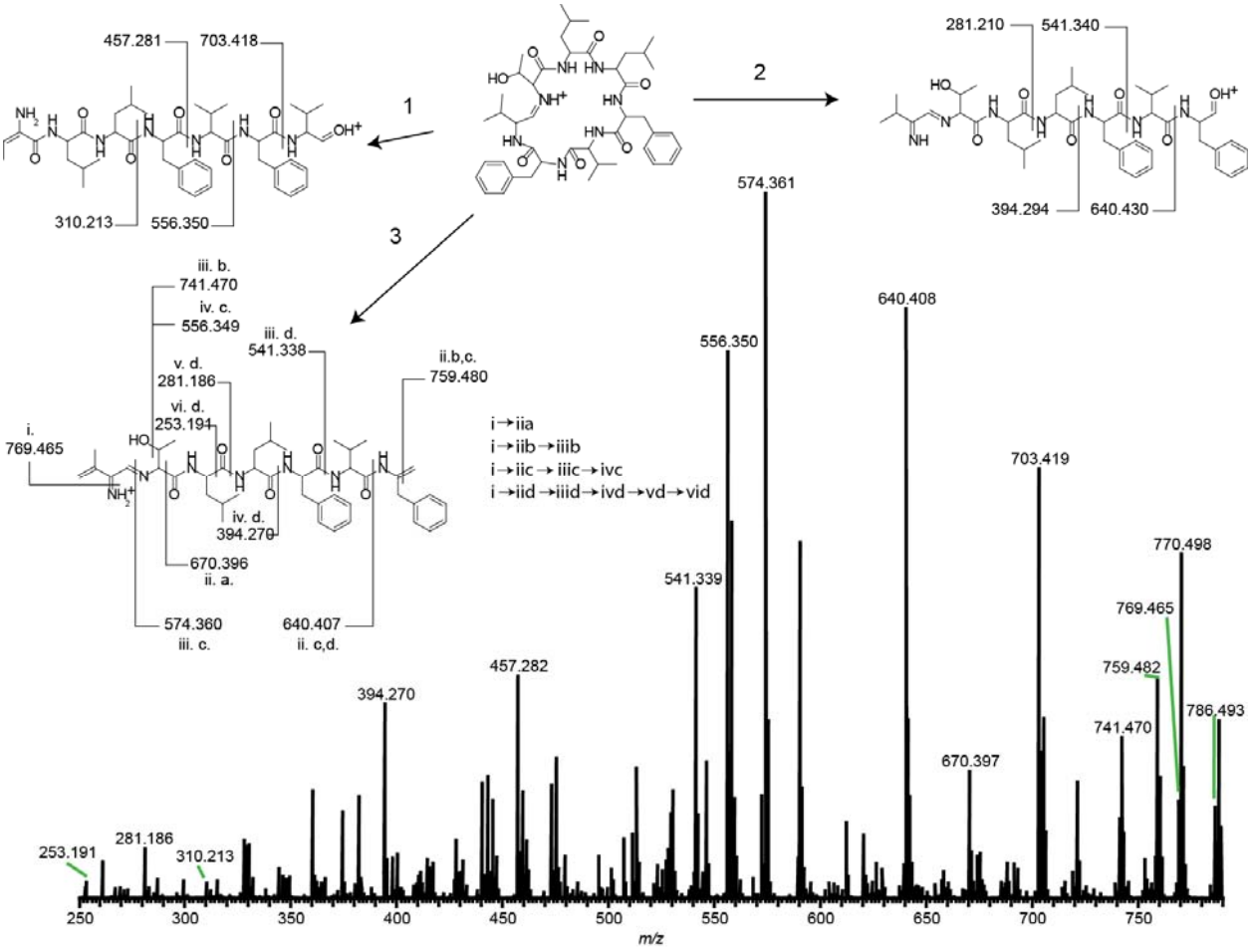

Figure S3. MS^2 for determination and confirmation of koranimine structure. Three major fragmentation pathways are observed. Ion structures are drawn as insets with theoretical fragment ion masses indicated below each cleavage site. Fragmentation subpathways are denoted with a Roman numeral (for the cleavage) followed by a letter (for the subpathway). For example, fragment ion iii. d at *m*/*z* 541.339 (theo. *m*/*z* 541.338) is the resultant ion of ring opening through pathway 3 and cleavage between Phe and Val to generate a b-type ion in a series of b-type ions, ii. d, iv. d, v. d, and vi. d. However, fragment ion iii. c is the apparent result of ring opening through pathway 3 and has lost the "N"-terminal imino-dehydrobutyrate-like residue (formerly a Val) as well as the "C"-terminal phenylpropylene amine (formerly a Phe).

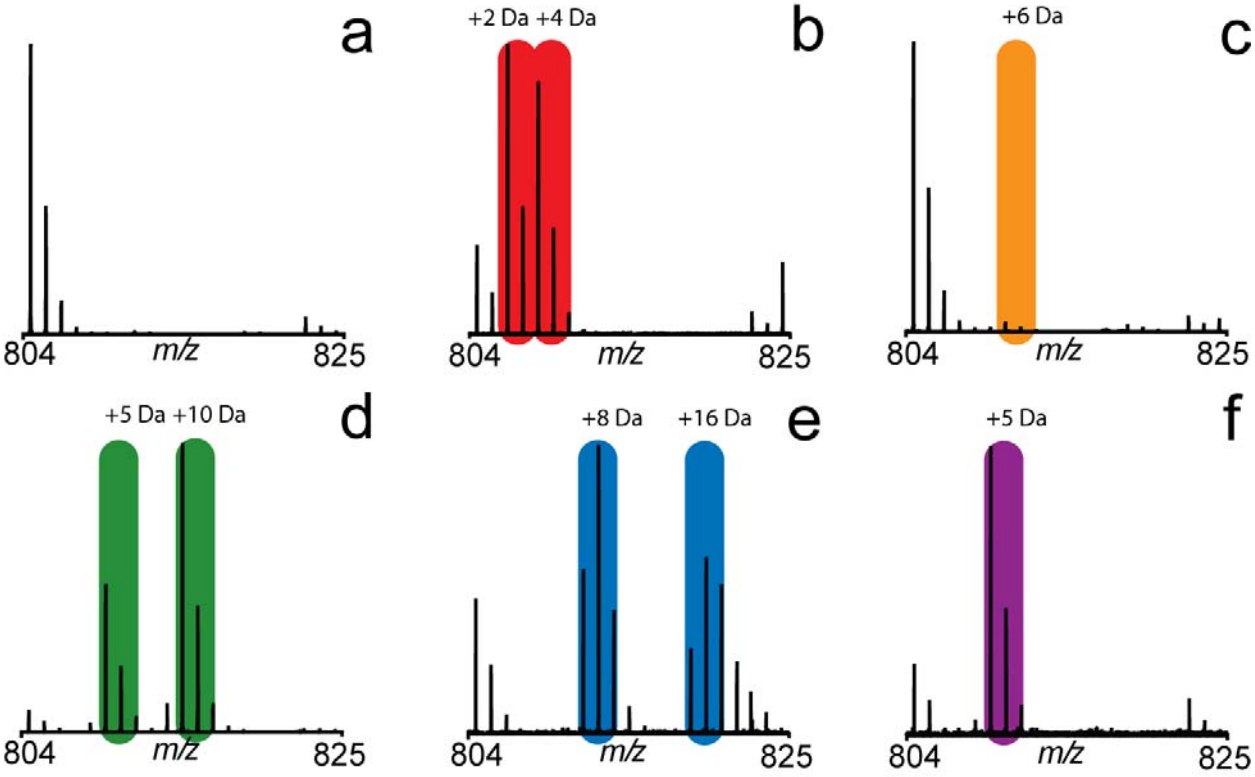

Figure S4. **Summary of metabolic labeling of koranimine with amino acid precursors containing stable isotopes. a,** unlabeled, **b,** $^{13}C_2$ leucine labeled, **c,** $^{13}C_6$ isoleucine labeled, **d,** D_5 phenylalanine labeled, **e,** D_8 valine labeled, **f,** $^{13}C_4$, ^{15}N threonine labeled. Mass shifts resulting from stable isotope incorporation are highlighted.

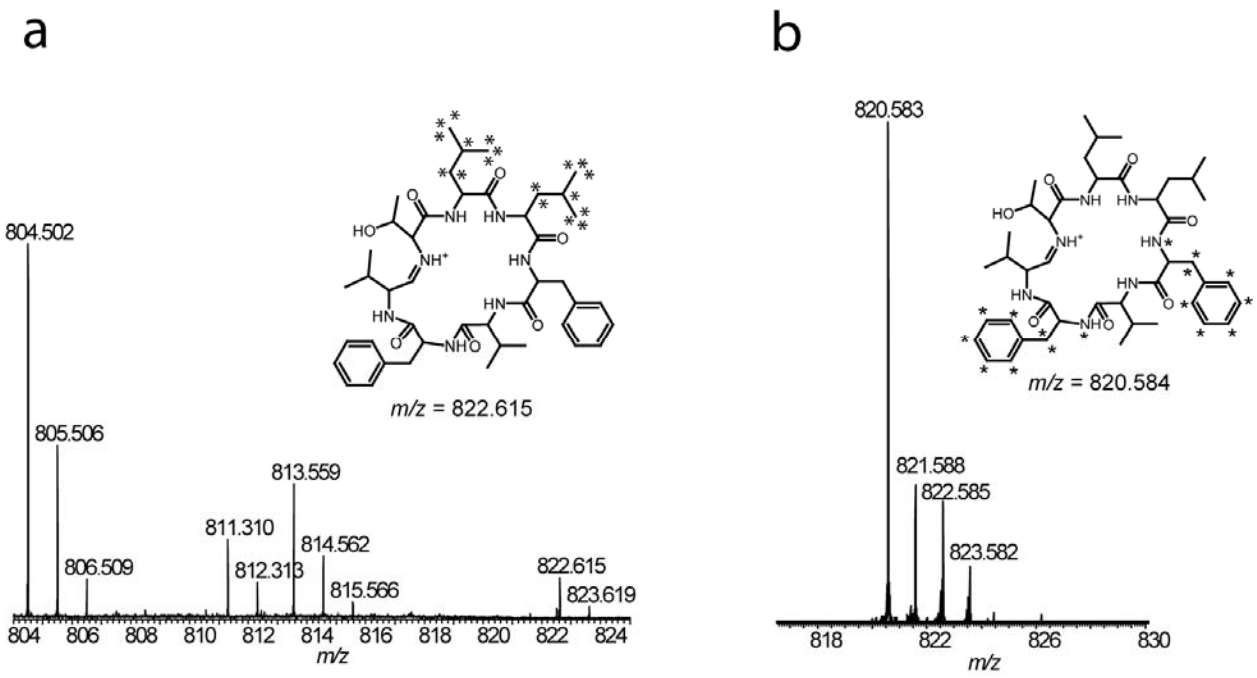

Figure S5. Stable isotope labeling of koranimine with D_{10}-Leucine and D_8-^{15}N-Phenylalanine. a, Washout of deuterium at the α-positions of the two leucine residues is consistent with L to D epimerization of leucine by the single epimerization domain in KorA. The peak at 804.502 represents unlabeled koranimine, the peak at 813.559 represents one labeled leucine incorporated and epimerized and the peak at 822.615 represents two labeled leucines incorporated and epimerized. (The peak at 811.310 is a coeluting impurity not related to koranimine) **b,** Washout of deuterium at the α-position is consistent with L to D epimerization of phenylalanine by the KorB and KorC epimerization domains. Deuterium labeled positions indicated by asterisks.

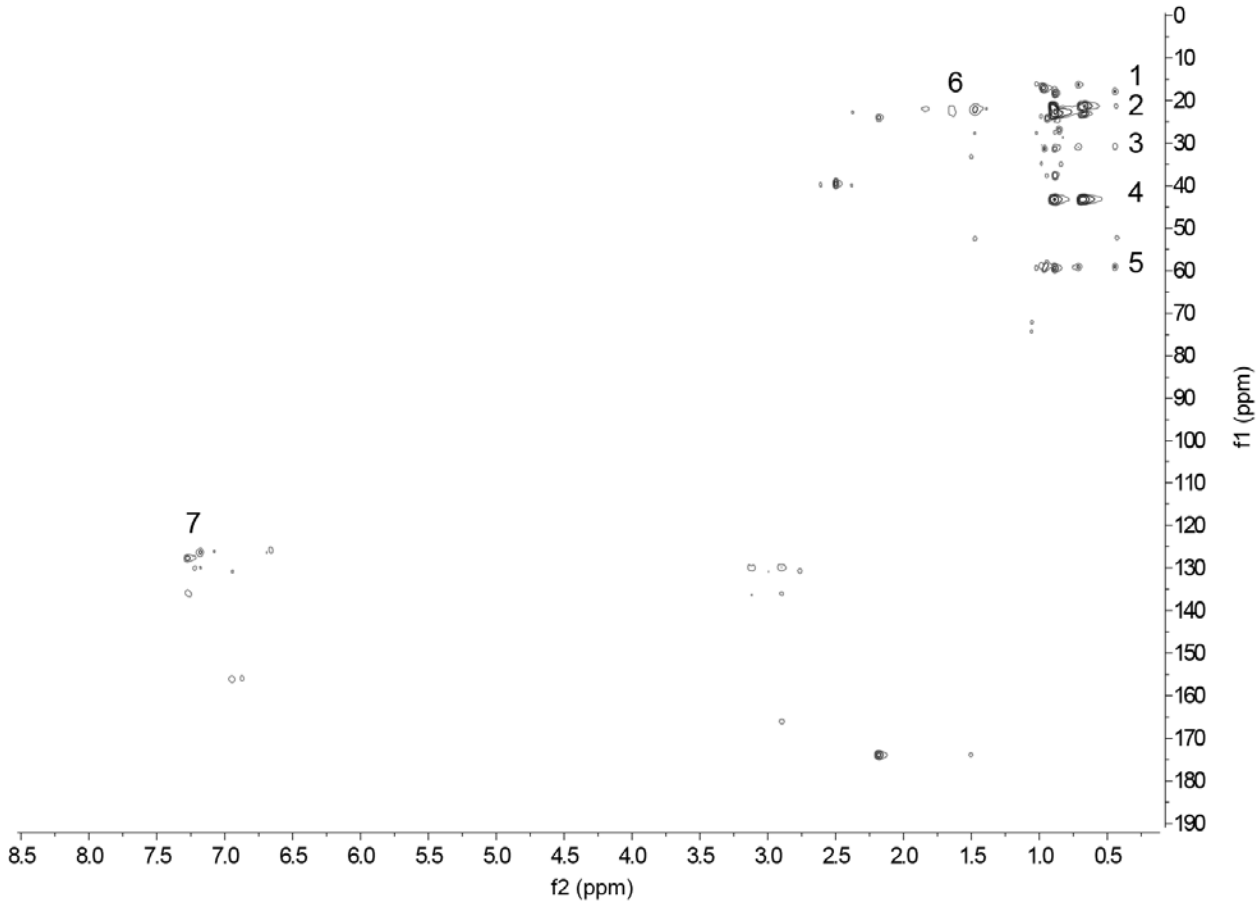

Figure S6. gHMBC spectrum of koranimine in d_6-DMSO. Signature peaks observed include: 1. Val CH_3 to Val CH_3; 2. Leu CH_3 to Leu CH_2 and CH_3; 3. Val CH_3 to Val CH; 4. Leu CH_3 to Leu CH_2; 5. Val CH_3 to Val CH; 6. Leu CH/CH_2 to Leu CH_3; 7. Phe CH – CH/C.

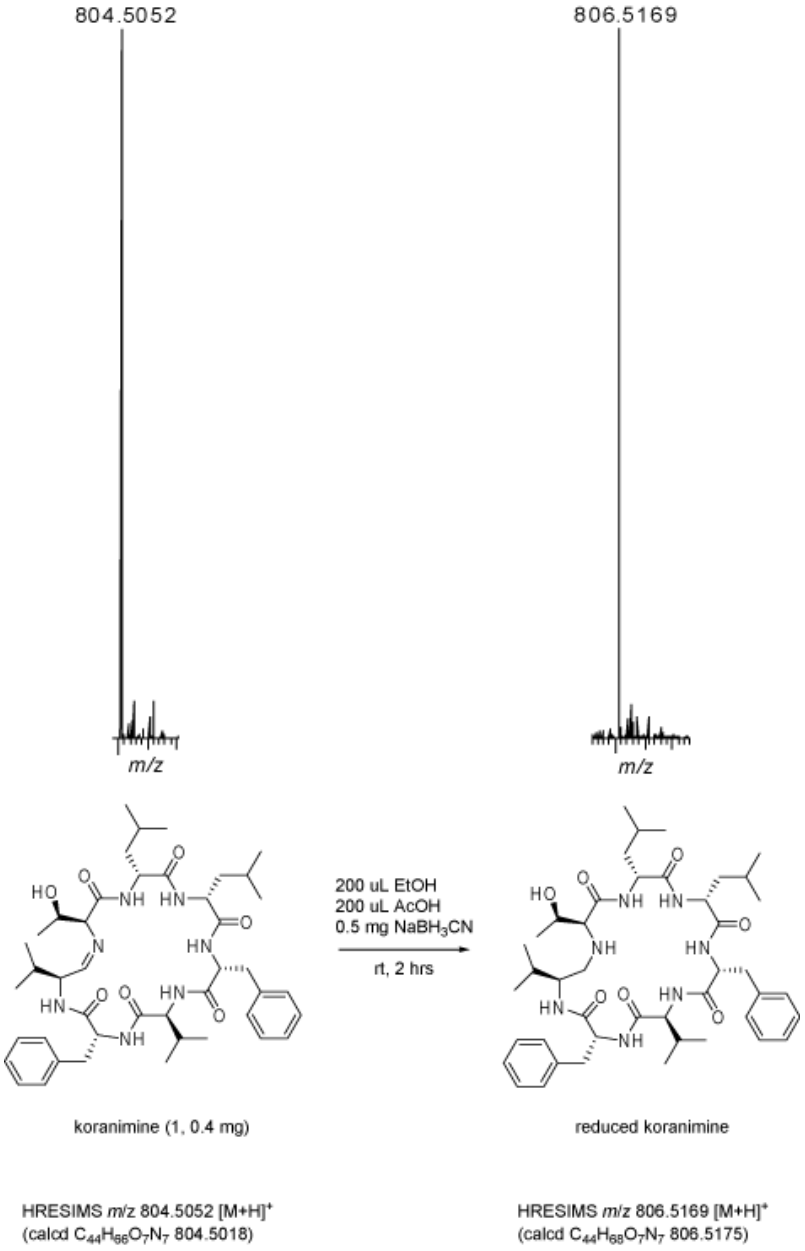

Figure S7. Chemical reduction of koranimine. Quantitative $NaCNBH_3$ reduction of the koranimine's imine to the corresponding secondary amine followed by mass spectrometry confirms assignment of the imine functionality introduced by the terminal reductase domain.

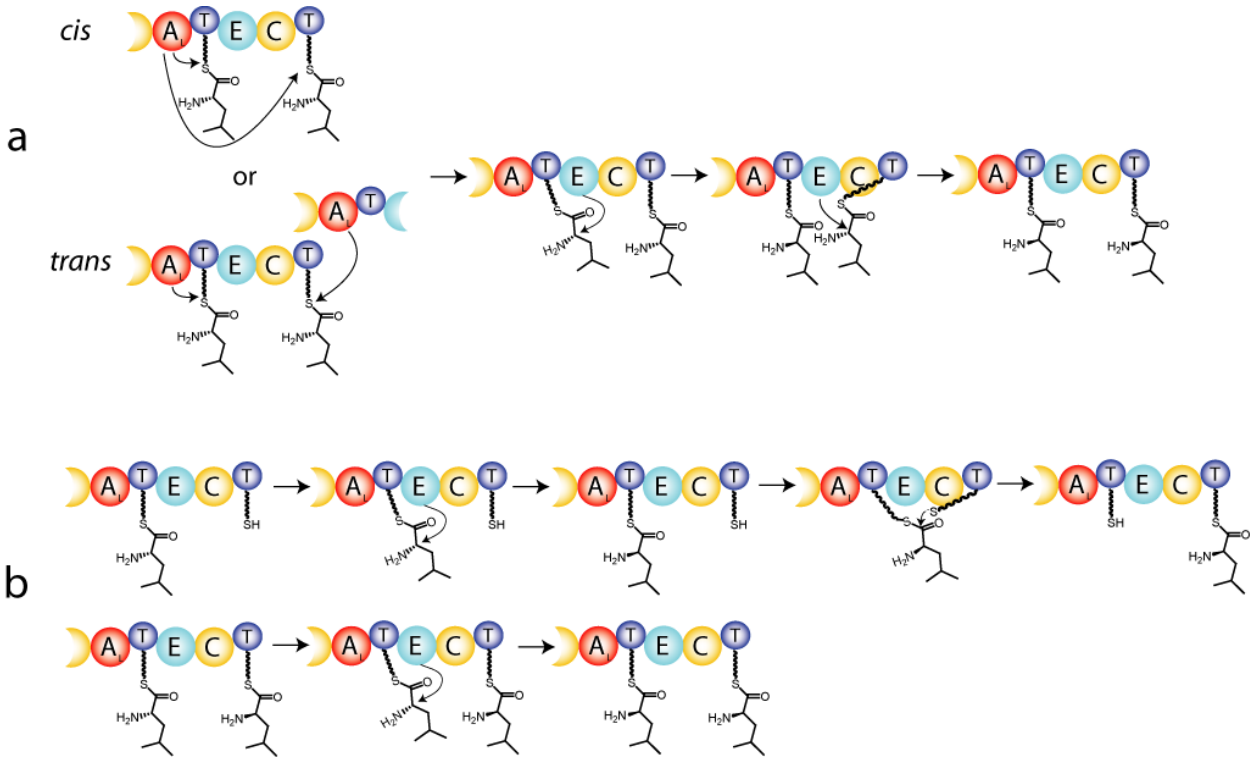

Figure S8. Two possible mechanisms for generation of two D-Leu residues by single adenylation and epimerization domains of the KorA enzyme. (a) cis- or trans-aminoacylation of KorA-T2 and KorA-T3 by KorA-A2 followed by epimerization of the L-Leu-acyl-S-KorA-T2 and L-Leu-acyl-S-KorA-T3 intermediates. (b) Sequential adenylation KorA-T2 and epimerization of L-Leu-acyl-S-KorA-T2 followed by transthioesterification of the now D-Leu-acyl-S intermediate to KorA-T3 allowing for a second round of adenylation and epimerization at KorA-T2.

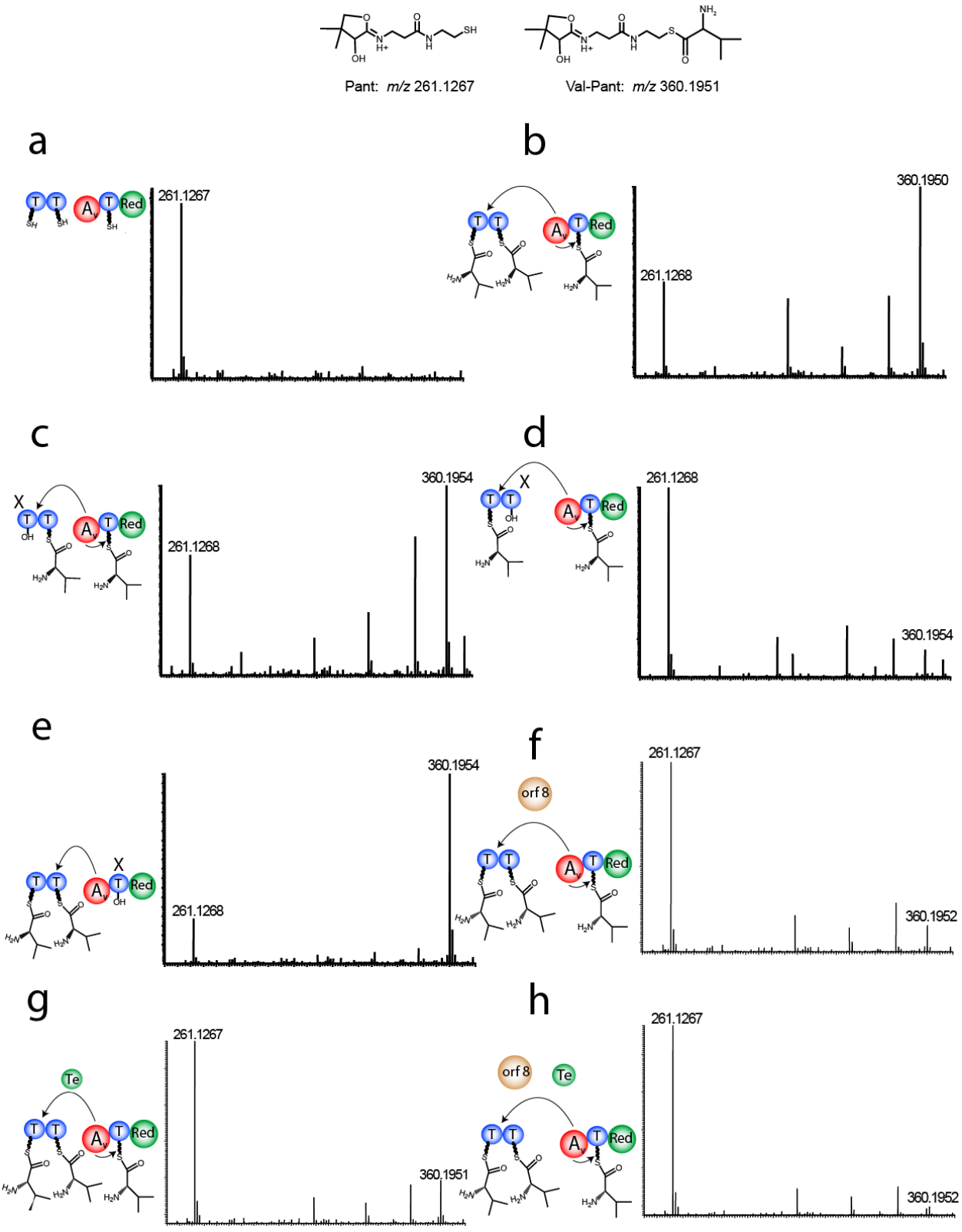
Pant: m/z 261.1267
Val-Pant: m/z 360.1951
a
261.1267
b
360.1950
261.1268
c
360.1954
261.1268
d
261.1268
360.1954
e
360.1954
261.1268
f
orf 8
261.1267
360.1952
g
261.1267
Te
360.1951
h
orf 8
Te
261.1267
360.1952

Figure S9. *In vitro* experiments showing long-range trans-aminoacylation of KorC-TT by KorD. KorD, alone is sufficient to transaminoacylate KorC-TT at either KorC-TT-T_1, KorC-TT-T_2 or both KorC-TT-T_1,T_2 without the need for covalent tethering to KorD or "shuttling" via KorTE of the α/β hydrolase, orf 8. **a,** PPANT ejection assay of holo KorD and KorC-TT. **b,** PPANT ejection assay of KorD and aminoacylated KorC-TT. **c,** PPANT ejection assay of KorD and KorC-TT-T_1KO. **d,** PPANT ejection assay of KorD and KorC-TT-T_2KO. **e,** PPANT ejection assay of KorD-TKO and KorC-TT. **f,** PPANT ejection assay of KorD and KorC-TT with addition of orf 8 protein. **g,** PPANT ejection assay of KorD and KorC-TT with addition of KorTE. **h,** PPANT ejection assay of KorD and KorC-TT with addition of orf 8 protein and KorTE. "X" indicates a domain inactivated by mutagenesis.

2. Methods

Chemicals, media and reagents: All chemicals were from Sigma unless otherwise noted. Restriction enzymes, DNA ligase and DNA polymerase were from New England Biolabs. Growth media were from BD Biosciences. Oligonucleotides were purchased from Integrated DNA Technologies and listed in Supplementary Table 1. Stable-isotope labeled amino acids were from Cambridge Isotopes.

Microorganisms: Bacilli were isolated as previously described[10]. *Escherichia coli* DH5α λpir and WM 4489 were a gift from William Metcalf. *Bacillus subtilis* 168 was a gift from George Ordal. *Saccharomyces cerevisiae* HZ 848 was a gift from Huimin Zhao.

PrISM-based screening: The use of SDS-PAGE to survey conditions and strains was published elsewhere[10]. *Bacillus* isolate NK2003 grown in Mueller-Hinton broth for 8 hours at 30 ºC produced the high molecular weight bands of Figure S1a. High molecular weight protein bands were subjected to in-gel trypsin digestion followed by nano-capillary LC-MS/MS on a 7T LTQ-FT (Thermo-Fisher Scientific) equipped with an Eksigent 1D nano HPLC system mounted with a 100 mm X 75 μm C_{18} picofrit column (New Objective) and a Picoview 500 nanospray source (New Objective). Solvents were water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). Peptides were eluted with a linear gradient from 0% B to 65 %B over 45 minutes. MS/MS data were acquired using data-dependent mode; the top five peaks from the FTMS full scan (*m/z* 400-1800) were selected for fragmentation and analyzed using ion trap detection. MS/MS data files were analyzed using OMSSA[23] and searched against a custom protein database containing all bacterial proteins and all NRPS and PKS proteins from GenBank. Peptides identified as coming from NRPSs were used for primer design as described above. We

assume peptides from a single gel band identified as coming from multiple different NRPSs according to the database search, are in fact the same NRPS that contains peptides identical or highly similar to peptides from the proteins present in our database, but recombined in a way that was not (before sequencing of the koranimine gene cluster) observed before**.** Each of the 106 identified peptides were from unique proteins from distinct gene clusters in disparate organisms; therefore degenerate primers were paired with NRPS adenylation domain degenerate primers for PCR. Twelve peptides were selected for back-translation and degenerate primer design. These twelve peptides were chosen because the tandem mass spectra that identified them showed matching ions for all or nearly all amino acids in the respective sequences. This high peptide sequence coverage is essential for the design of robust PCR primers from amino acid sequences. The proteomic data associated with this manuscript may be downloaded from ProteomeCommons.org Tranche using the following hash: 7LVejkLnLo2iGnfUwDlypaRckZYFVmvkl+P7EYVIguo7j5EVL9BNjQWV5JjBqHtMNHGLarjgGSUleZQlvlDO1o8pWhIAAAAAAAAG4w==.

Plasmids: Fosmid pJK0050 was a gift from William Metcalf. Plasmid pBEX was constructed by PCR from pQE60 (Qiagen), pET-Duet (Novagen), pMQ118 (Presque Isle Cultures, Erie, PA) pQE60-EGFP (unpublished) and pAIN750 (a gift from George Ordal). Left-side vector flanking sequence was amplified from pQE-60 with primers pQE60-F and pQE60-‘amyE-R to give PCR product pQE-60 spacer-F. ‘amyE was amplified from pAIN750 using primers pQE60-‘amyE-F and ‘amyE-LacI-R to give PCR product ’amyE. LacI was amplified from pET-Duet using primers ‘amyE-LacI-F and LacI-GFP-R to give PCR product LacI. pT5-EGFP-His_6 was amplified from pQE60-EGFP using primers LacI-GFP-F and GFP-spc-amyE’-R to give PCR product GFP. Spc-amyE’ was amplified from pAIN 750 using primers GFP-spc-

amyE'-F and Spc-amyE'-Cen-R to give PCR product spc-amyE'. Cen/ARS-URA3 was amplified from pMQ118 using primers Spc-amyE'-Cen-F and Cen-pQE60-R to give PCR product Cen/ARS-URA3. Right-side vector flanking sequence was amplified from pQE-60 using primers Cen-pQE60-F and pQE60-R to give PCR product pQE60 spacer-R. PCR products from the above reactions were gel purified using standard procedures. Fragment pQE60-'amyE-LacI was constructed by OE-PCR using PCR products pQE60 spacer-F, 'amyE and LacI in a three way OE extension reaction, followed by a final amplification using pQE-60-F and LacI-GFP-R. Fragment LacI-GFP was constructed by OE-PCR using PCR products LacI and GFP in an extension reaction followed by final amplification using primers 'amyE-LacI-F and GFP-spc-amyE'-R. Fragment GFP-Spc-amyE' was constructed by OE-PCR using PCR products GFP and spc-amyE' in an extension reaction followed by final amplification using primers LacI-GFP-F and Spc-amyE-Cen-R. Fragment Spc-amyE'-Cen/ARS-URA3-pQE60 was constructed by OE-PCR using PCR products spc-amyE', Cen/ARS-URA3 and pQE60 spacer-R in a three way extension reaction followed by final amplification using primers GFP-spc-amyE'-F and pQE60-R. Fragments pQE60-'amyE-LacI, LacI-GFP, GFP-Spc-amyE', spc-amyE'-Cen/ARS-URA3-pQE60 and NcoI/BamHI digested pQE60 were assembled in *S. cerevisiae* HZ 848 by homologous recombination as previously described[24] to make plasmid pBEX. This vector is an *E. coli- S. cerevisiae* shuttle vector capable of plasmid based IPTG inducible expression of genes cloned between PT5 and His_6 in *E. coli* as well as chromosomal based IPTG inducible expression of genes cloned between PT5 and His_6 in *B. subtilis* when integrated into the neutral *amyE* locus. Additionally, genes or entire pathways may be expressed under control of native promoters in *B. subtilis* when the insert is cloned between the 'amyE and spc-amyE' fragments.

Degenerate primer design: Primers were designed from peptides identified as coming from NRPS proteins. Peptide sequences were reverse translated using the web-based Reverse Translate Tool (http://www.bioinformatics.org/sms2/rev_trans.html) using the *Bacillus thuringiensis serovar kurstaki* codon usage table. Both the most likely and highly degenerate reverse translations were used to design primers. None of the identified peptides were from the same protein therefore degenerate primers were paired with NRPS adenylation domain degenerate primers[25] that were modified to use the *Bacillus thuringiensis serovar kurstaki* codon usage table. 64 PCRs were performed using the peptide based most likely and highly degenerate primers paired with the NRPS A3 and A7 degenerate primers. One reaction gave a single, reproducible band at the expected size of 700 bp. The 700 bp band was purified from the gel and sequenced using the primers that amplified it. BLAST analysis of the sequence of the 700 bp band indicated that it coded for the N-terminal portion of an NRPS adenylation domain. The sequence of the 700 bp PCR product was used to construct probe primers for screening a fosmid library of bacillus isolate NK2003 genomic DNA clones.

Fosmid library construction, screening, and cluster sequencing: A fosmid library of *Bacillus* isolate NK2003 was prepared then screened by PCR essentially as described[26]. The fosmid library was screened using primers NE2-2 probe-F and NE2-2 probe-R. Seven clones containing the probe sequence were isolated from the library screening. Sequencing was performed using a Roche 454 FLX-Titanium sequencer at the UIUC DNA Core facility on a mixture of the 7 overlapping fosmids. Sequence reads were assembled using the GS *de novo* sequencing assembler (Roche). Automatic annotation was performed using RAST[27]. Targeted PCR followed by Sanger sequencing on an ABI 3730xl capillary system was used to close gaps

and resolve ambiguities. The koranimine gene cluster has been deposited in GenBank under the accession number (JF828091).

Heterologous expression of koranimine gene cluster: The full 30 kb gene cluster was heterologously expressed in *B. subtilis* under the control of the native promoters and integrated into the neutral amyE locus. To construct the expression cassette, overlapping fosmid inserts were released from the vector by NotI/NcoI/MluI digestion. The released inserts were purified from an agarose gel using standard procedures. The released inserts were co-transformed into HZ 848 along with NcoI/BamHI digested and gel purified pBEX, essentially as previously described[24] for assembly by homologous recombination. Yeast clones were screened by PCR for correct inserts before the plasmid was transformed into DH5α λ pir for amplification. The plasmid was transformed into *B. subtilis* by natural transformation[28] and screened for proper insertion by PCR and assay for amylase activity. *B. subtilis* clones grown up in the same conditions as the native producer were assayed by LC-MS/MS for production as described above.

Targeted metabolite discovery and structure elucidation. Small molecule MS was performed on culture supernatant from bacillus isolate NK2003 grown in the same conditions as for protein analysis, or on semi-purified compound for direct infusion experiments. For LC-MS/MS experiments a 7 or 12 T LTQ-FT (Thermo-Fisher Scientific) equipped with a Surveyor (Thermo-Fisher Scientific) or Agilent 1200 autosampler and LC was used. Samples were run through a 4.6 X 150 mm Jupiter C_{18} column (Phenomenex) at 400 μL/min. using water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B) on a linear gradient from 0-100% B over 30 minutes. Data were collected in a data dependent fashion, selecting the top

three peaks in the FTMS full scan for fragmentation and detection by FTMS/MS. For direct infusion, semi-purified or Zip-Tip desalted samples were dissolved in 80% acetonitrile 20% water with 0.1% formic acid and infused using a Nano-Mate 100 nanospray source (Advion). Direct infusion data were collected manually using the LTQ-Tune function of Xcalibur (Thermo-Fisher Scientific).

Stable isotope feeding studies: Amino acids were added (1 mM final concentration) as required to mannitol supplemented minimal medium (1% mannitol, 0.1% caseamino acids, 0.1% yeast extract and M9 salts). Single colonies were picked into 5 mL media and grown at 30º C and 250 rpm. 1 mL samples were pulled from cultures at 24, 36 and 48 hours, cells removed by centrifugation then the supernatant was filtered through a 0.45 μm filter and stored at -20º C until analyzed. Culture supernatants were analyzed by direct infusion or by LC-MS as described above.

Isolation of koranimine (1): One colony (NK2003) was used to inoculate 1 mL of Müller-Hinton broth (MHB) at 30 °C, 220 rpm for 5 h. 250 μL of the 1 mL culture was used to inoculate four 50 mL starter cultures that were grown for 18 h at 30 °C, 220 rpm. The four starter cultures inoculated 20 L of MHB (1%) and grew for 48 h at 30 °C, 180 rpm. The EtOAc soluble material was dried and solubilized in CH_2Cl_2 and EtOAc (0.8541 g) and separated by RP-HPLC using a 5 – 100% linear gradient of 1:1 IPA:ACN – H_2O with 0.1% TFA at 42 °C on a 5 μm, C18, 250 x 10 mm column (Phenomenex) to afford **1** (1.1 mg).

Cloning, mutation, heterologous expression and purification of enzymes: KorD and KorTE were cloned into pET28a and expressed as C-terminal his-tag fusions. KorD was amplified from NK2003 gDNA using primers KorD-F and KorD-R, cleaned up and digested

with NcoI and XhoI. Gel-purified insert was ligated to similarly digested and purified vector. KorTE was amplified from NK2003 gDNA with primers KorTE-F and KorTE-R . The insert was cleaned up and digested with NcoI and BamHI, gel purified and ligated to similarly digested and purified vector. Orf 8was cloned into pET28a and expressed as a C-terminal his-tag fusion. Orf 8 was amplified from NK2003 gDNA using primers lipase-F and lipase-R, cleaned up and digested with NcoI and BglII. Gel purified insert was ligated to similarly digested and purified vector. KorC-TT was cloned and expressed as an N-terminal his-tag fusion in pET15b. KorC-TT was amplified from NK2003 gDNA using primers KorC-TT-F and KorC-TT-R. The insert was cleaned up and digested with NdeI and EcoRI, gel purified and ligated to similarly digested and purified vector. A mutant of KorD that has the phosphopantetheine attachment site mutated to alanine preventing formation of the holo-enzyme was constructed by OE-PCR using the primer pairs KorD-F/KorD-TKO-R and KorD-TKO-F/KorD-R to amplify the mutant gene fragments with NK2003gDNA as template. The mutant amplicons were gel purified and used in an overlap extension reaction and KorD-TKO was subsequently amplified using primers KorD-F and KorD-R. Cloning of the KorD-TKO proceeded as with the WT from this point. Mutants of KorC-TT were constructed so that either the first or second phosphopantetheine attachment site was changed to alanine. KorC-TT T_1KO and T_2KO were constructed using the same protocol as for KorD except using the primer pairs KorC-TT-F and (KorC-TT-T1KO-R/KorC-TT-T2KO-R) and (KorC-TT-T1KO-F/KorC-TT-T2KO-F) and KorCTT-R with NK2003 gDNA as template. For expression in *E. coli*, expression plasmids were transformed into BL21 (DE3), grown in 2xYT supplemented with appropriate antibiotics at 37º C and 250 rpm until OD_{600} reached 0.8. Expression was induced by addition of IPTG to a final concentration of 1 mM and the temperature was shifted to 18º C for overnight expression. For expression in *B. subtilis*, clones

containing the expression sequence on chromosomal DNA were grown in 2xYT supplemented with spectinomycin and grown at 37º C and 250 rpm until OD_{600} reached 0.2. Expression was induced by addition of IPTG to a final concentration of 1 mM and the temperature was shifted to 18º C for overnight expression. For all enzyme preparations, cells were harvested by centrifugation resuspended in lysis buffer (25 mM HEPES, 200 mM NaCl, 10 mM $MgCl_2$, 35 mM imidazole, 10% glycerol pH = 7.5) lysed by lysozyme/DNase treatment followed by sonication. Soluble fraction was collected by centrifugation and enzymes were purified by Ni^{2+} affinity chromatography on a 1 mL His-Trap column (GE Healthcare) on an Akta-purifier FPLC (GE Healthcare). The column was washed with 120 mL wash buffer (25 mM HEPES, 200 mM NaCl, 10 mM $MgCl_2$, 35 mM imidazole, 10% glycerol pH = 7.5) and eluted with elution buffer (25 mM HEPES, 200 mM NaCl, 10 mM $MgCl_2$, 250 mM imidazole, 10% glycerol pH = 7.5). Eluted protein fractions were desalted on a 5 mL Hi-Trap desalting column into 25 mM HEPES, 200 mM NaCl, 10 mM $MgCl_2$, 10% glycerol pH = 7.5 assayed by SDS-PAGE and BCA assay then frozen at -80º C until used.

In vitro characterization of enzymes: All enzyme reaction assays were carried out in 10 µL volume. Reactions were performed in buffer containing 50 mM Tris-HCl pH = 7.5, 10 mM $MgCl_2$, 200 µM CoA and 2 mM valine or leucine, as appropriate. Enzymes were converted to holo form by addition of 35 pmol sfp followed by incubation at room temperature for 30 minutes. The following amounts of enzymes were used for PPant ejection assay (when appropriate): 250 pmol KorC-TT (WT or single or double T-domain knockout), 50 pmol KorD (WT or T-domain knockout), 15 pmol lipase, 15 pmol KorTE. Aminoacylation was carried out by addition of 100 nmol ATP to holo enzyme mixtures followed by incubation at room temperature for one hour. Reactions were quenched by addition of 10 µL of 10% formic acid.

Quenched enzyme reactions were centrifuged to remove particulates before loading onto a capillary HPLC column. PPant ejection assays were performed on a 12T LTQ-FT (Thermo-Fisher Scientific) equipped with an Eksigent 2D nanoLC system and a Picoview 500 nano spray source on a 10 cm C_4 picofrit column (New Objective) and 2 cm C_4 homemade trap. PPant ejection assays monitored the intact KorC-TT di-domain construct. Solvents were water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). Proteins were eluted using a linear gradient from 15-80% B over 30 minutes. MS scans were alternating intact (FTMS 100,000 resolution 1 microscan SID = 0, *m/z* 500-1800) and PPant ejection scans (FTMS 25,000 resolution 1 microscan SID = 75, *m/z* 250-370). Data shown for PPant ejection assays are the summed PPant ejection scans under the KorC-TT peak in the chromatogram.

Supplemental References

(23) Geer, L. Y.; Markey, S. P.; Kowalak, J. A.; Wagner, L.; Xu, M.; Maynard, D. M.; Yang, X. Y.; Shi, W. Y.; Bryant, S. H. *J. Proteome Res.* **2004,** *3*, 958-964.

(24) Shao, Z.; Luo, Y.; Zhao, H. *Molecular Biosystems*, DOI:10.1039/C0MB00338G.

(25) Ayuso-Sacido, A. and Genilloud, O. *Microb. Ecol.* **2005,** *49*, 10-24.

(26) Woodyer, R. D; Shao, Z. Y; Thomas, P. M.; Kelleher, N. L.; Blodgett, J. A. V.; Metcalf, W. W.; Van der Donk, W. A.;Zhao, H. M. *Chem. Biol.* **2006,** *13*, 1171-1182.

(27) Aziz, R. K.; Bartels, D.; Best, A. A.; DeJongh, M.; Disz, T.; Edwards, R. A.; Formsma, K.; Gerdes, S.; Glass, E. M.; Kubal, M.; Meyer, F.; Olsen, G. J.; Olson, R.; Osterman, A. L.; Overbeek, R. A.; McNeil, L. K.; Paarmann, D.; Paczian, T.; Parrello, B.; Pusch, G. D.; Reich, C.; Stevens, R.; Vassieva, O.; Vonstein, V.; Wilke, A.; Zagnitko, O. *BMC Genomics* **2008,** doi:10.1186/1471-2164-9-75.

(28) Yasbin, R. E.; Wilson, G. A.; Young, F. E. *J. Bacteriol.* **1975,** *121*, 296-304.

3. Supplemental Table

Table S1. Oligonucleotides Used (Restriction Sites Underlined)

Name	Sequence
pQE60 spacer-F	ACCCACTCGTGCACCCAAC
pQE60-‘amyE-R	TTGAATCGTTTTGCAAACATACTTTTCGGGGAAATGTGCG
pQE60-‘amyE-F	CGCACATTTCCCCGAAAAGTATGTTTGCAAAACGATTCAA
‘amyE-LacI-R	GGTTGAGGCCGTTGAGCACCGATCAGACCAGTTTTTAATTTGTGTG
‘amyE-LacI-F	CACACAAATTAAAAACTGGTCTGATCGGTGCTCAACGGCCTCAACC
LacI-Gfp-R	TTTCTTAGACGTCAGGTGGCTTCCGCGTTTCCAGACTTTA
LacI-Gfp-F	TAAAGTCTGGAAACGCGGAAGCCACCTGACGTCTAAGAAA
Gfp-spc-amyE’-R	GCAGGAATTCGATAAGCTTAAGGTTTTAGCTTCCTTAGCTCCTGAAAATC
Gfp-spc-amyE’-F	GATTTTCAGGAGCTAAGGAAGCTAAAACCTTAAGCTTATCGAATTCCTGC
Spc-amyE’-Cen-R	CATGACATTAACCTATAAAAAAGGCCTCTCAATGGGGAAGAGAACCGC
Spc-amyE’-Cen-F	GCGGTTCTCTTCCCCATTGAGAGGCCTTTTTTATAGGTTAATGTCATG
Cen-pQE60-R	CGGTATTTCACACCGCATATGCTAGTAGACGAGTCCATGTGCTGG
Cen-pQE60-F	CCAGCACATGGACTCGTCTACTAGCATATGCGGTGTGAAATACCG
pQE60-spacer-R	TGCGTTATCCCCTGATTCTGT
A-Domain A3F	GCWTATSYSATKTATACDTCWGG
A-Domain A3R	CCWGTGGTATAMATSRSATAWGC
A-Domain A7R	STGRTCHCCSGTYYKGTA
A-Domain A7F	TACMRRACSGGDGAYCAS
Peptide 1F	AGATTAGCACCATTAAGTGCAGCAGAAAGA
Peptide 1R	TCTTTCTGCTGCACTTAATGGTGCTAATCT
Peptide 1FD	MGNYTDGCNCCNYTAWSHGCDGCNGARMGW
Peptide 1RD	WCKYTCNGCHGCDSWTARNGGNGCHARNCK
Peptide2-F	TTAGTAGAAGTATTACAACCAGATAGA
Peptide2-R	TCTATCTGGTTGTAATACTTCTACTAA
Peptide 2D-F	YTDGTDGARGTDYTDCARCCDGAYMGD
Peptide 2D-R	HCKRTCHGGYTGHARHACYTCHACHAR
Peptide 3-F	AGTGGTGAATGTGATTTAGC
Peptide 3-R	GCTAAATCACATTCACCACT
Peptide 3D-F	WSHGGNGARTGYGATYTDGC
Peptide 3D-R	GCHARATCTCAYTCNCCDSW
Peptide 4-F	ATGATTACAGGTATTATGGG
Peptide 4-R	CCCATAATACCTGTAATCAT
Peptide 4D-F	ATGATHACDGGNATHATGGG
Peptide 4D-R	CCCATDATNCCHGTDATCAT
Peptide 5-F	GGTTGGTTTACAAGTTTATATCC
Peptide 5-R	GGATATAAACTTGTAAACCAACC
Peptide 5D-F	GGNTGGTTTACNWSNYTDTAYCC
Peptide 5D-R	GGRTAHARNSWNGTAAACCANCC
Peptide 6-F	GATATGATTGTAGGTATTTTTGG
Peptide 6-R	CCAAAAATACCTACAATCATATC
Peptide 6D-F	GATATGATHGTNGGNATHTTTGG
Peptide 6D-R	CCAAADATNCCNACDATCATATC

Table S1 (continued). Oligonucleotides Used (Restriction Sites Underlined)

Name	Sequence
Peptide 7-F	AATATGTATGGTATTACAGAAACAACAGTACATGT
Peptide 7-R	TGTACATGACAACAAAGACATTATGGTATGTATAA
Peptide 7D-F	AATATGTATGGNATYACNGARACNACNGTNCATGT
Peptide 7D-R	ACATGNACNGTNGTYTCNGTRATNCCATACATATT
Peptide 8-F	GGTTGGTTTACAAGTTTATATCC
Peptide 8-R	GGATATAAACTTGTAAACCAACC
Peptide 8D-F	GGNTGGTTTACNWSNYTDTAYCC
Peptide 8D-R	GGRATHARNSWNGTAAACCANCC
Peptide 9-F	GGTCAAAATGATATTTTAGTAGG
Peptide 9-R	CCTACTAAAATATCATTTTGACC
Peptide 9D-F	GGNCAAAATGATATHYTTGTNGG
Peptide 9D-R	CCNACAARDATATCATTTTGNCC
Peptide 10-F	ATGGGTGAATTATATATTGG
Peptide 10-R	CCAATATATAATTCACCCAT
Peptide 10D-F	ATGGGNGAAYTATATATHGG
Peptide 10D-R	CCDATATATARTTCNCCCTA
Peptide 11-F	ATTGTAGAAGTAGTACAACCAGAAAGA
Peptide 11-R	TCTTTCTGGTTGTACTACTTCTACAAT
Peptide 11D-F	ATHGTNGARGTNGTNCAACCNGAAMGN
Peptide 11D-R	NCWTTCNGGTTGNACNACYTCNACDTA
Peptide 12-F	GAAAGATTTGTAGCAGATCCATTTGG
Peptide 12-R	CCAAATGGATCTGCTACAAATCTTTC
Peptide 12D-F	GAAMGNTTTGTNGCNGATCCNTTTGG
Peptide 12D-R	CCAAANGGATCNGCNACAAANCWTTC
NE2-2-Probe-F	ACTTGAGGAGACATTCGAAGAGA
NE2-2-Probe-R	CAAGCAATAGGAAATGGACGA
KorD-F	AAATATAACCATGGCGACGAATAAATTGATGTCAATAACT
KorD-R	AAATAATACTCGAGTTAGTGATGGTGATGGTGATGTTTCAGCATGAACTTCATGAAG
KorTE-F	ATAATAATCCATGGAAGGGCTGAAGAAGAC
KorTE-R	AATTTATAGGATCCAATCAGAATTTCACAACGACTTCTT
Lipase-F	AATAAAATCCATGGAAACAACGACCAAATTTTATAAGAAAAA
Lipase-R	AATTTATAAGATCTTTCACTATGTTCCTTTAAGAATTGAATC
KorC-TT-F	ATAATATACATATGAAGATCGCATCCACATCCG
KorC-TT-R	ATTATAATGGATCCTGACATTTCATTAATGATGAAGGC
pBEX-Up-F	GAATCACCGATACGCGAGCG
pBEX-Down-R	CTTTCATTTCCATAAAACTAAAGTAAGTGTAAACCTATTCATTG
KorD TKO-F	AATGGAGGCCATGCCATTTTAGCTACA
KorD TKO-R	TGTAGCTAAAATGGCATGGCCTCCATT
KorC T1KO-F	GGCGGGCATGCGATAACAGCC
KorC T1KO-R	GGCTGTTATCGCATGCCCGCC
KorC T2KO-F	GGCGGACACGCGATTTTAGCCA
KorC T2KO-R	TGGCTAAAATCGCGTGTCCGCC