Supplementary Information

Biochemical Characterization and Substrate Specificity of the Gene Cluster for Biosyntheses of K-252a and Its Analogs by *In Vitro* Heterologous Expression System of *Escherichia coli*

Hsien-Tai Chiu,* Yu-Chin Lin, Meng-Na Lee, Yi-Lin Chen, Mei-Sin Wang and Chia-Chun Lai

Department of Biological Science and Technology,
National Chiao Tung University, 75 Po-Ai Street,
Hsinchu 300, Taiwan.

CONTENTS

*To whom correspondence should be addressed at the Department of Biological Science and Technology, National Chiao Tung University, 75 Po-Ai Street, Hsinchu 300, Taiwan. Tel: +886-3-5131595; Fax: +886-3-5719605; E-mail: Chiu@mail.nctu.edu.tw.

Experimental Procedures (Supplementary)

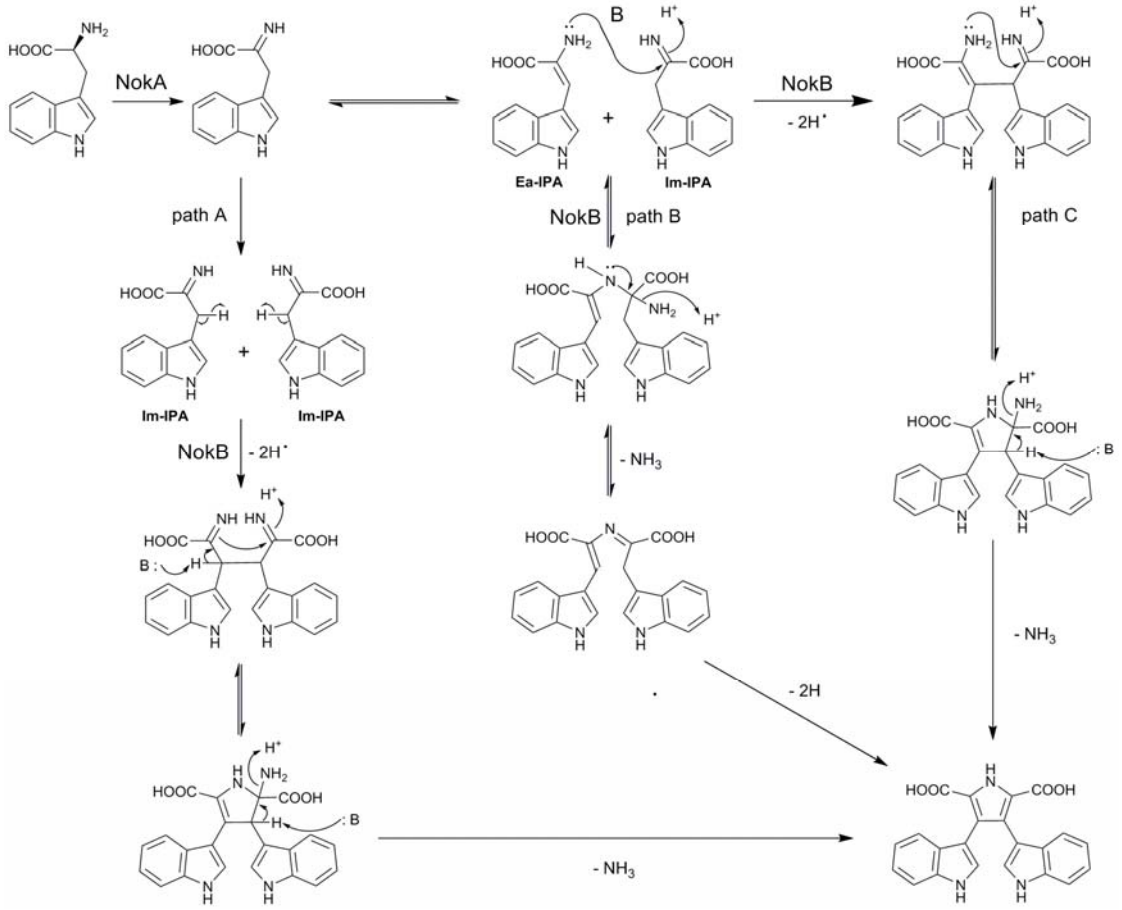
In vitro biosynthesis of K-252d by the *E. coli* expression system of NokL

The *in vitro* biosynthesis of K-252d **12** with dTDP-D-glucose (TDP-Glc) or dTDP-4-keto-6-deoxy-glucose (KDG) by the *E. coli* expression system of NokL was carried out similarly as described for preparation of NokL enzymatic product with TDP-Rha **11** in **Materials and methods**. The time-dependent analysis of NokL reaction with TDP-Glc or KDG was also similarly as described therein. A typical reaction (15 μL) contained NDP-sugar (1.2 mM), K-252c (0.3 mM), $MgCl_2$ (12 mM), K_2HPO_4 (50 mM, pH 9.0) and the NokL cell-free crude extract (3 μL). The reaction was incubated at 30℃ for 24 h (for *in vitro* biosynthesis) or for different periods of time (for time-dependent analysis by HPLC). The reaction was subsequently terminated by quenching with an equal volume of ice-cold MeOH. The reaction mixture was analyzed by a 5C18-AR-II RP-HPLC column (4.6 x 250 mm, 5 μm, Cosmosil, Nacalai tesque) as described.

The pH-dependent profile of NokL activity

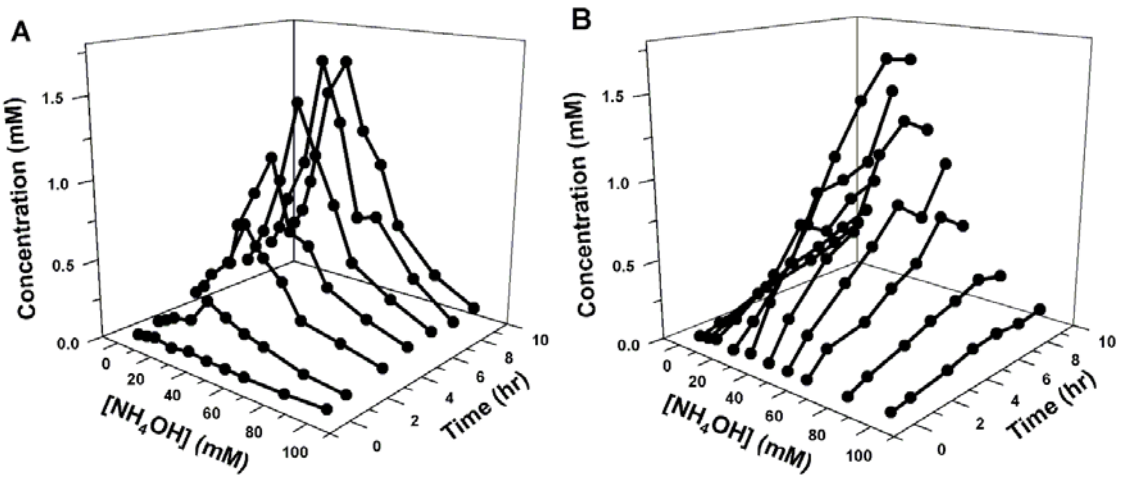
The optimum pH for *N*-glycosylation of NokL was determined by detecting the production of rhamnosyl-K252c (K-252d, **12**) at 30^{o}C for 4.5 h in potassium phosphate buffer (50 mM) set at a desired pH (range 6.0-12.0). The reactions were analyzed by RP-HPLC for NokL activity as described in **Materials and methods**. Consequently, NokL activity displayed an essential pH-dependence over pH 6.0-12.0. A maximal activity for NokL was observed at pH 9.0, accompanied with a gradual increase from pH 6.0 to pH 8.5 and a relatively dramatic drop while pH exceeded 9.0.

Scheme S1 Proposed Reaction Pathways of the NokB Catalysis



For the NokB-catalyzed reaction in the formation of CPA, three reaction paths can be proposed: Path A: As proposed for VioB in violacein biosynthesis, two molecules of Im-IPA may undergo benzylic carbon-carbon coupling, followed by tautomerization and subsequent intramolecular condensation to give CPA.[1] Path B: As proposed for RebD in rebeccamycin biosynthesis, CPA may be generated by condensation of Im-IPA with Ea-IPA to yield a cross-conjugated imine, followed by subsequent intramolecular oxidative coupling.[2] Path C: Im-IPA and Ea-IPA may first undergo oxidative coupling in the enzyme active-site, followed by intramolecular condensation to give CPA.

Figure S1 [NH_4OH] and Time Correlation Plots of the Tandem NokA/NokB Enzymatic Activity



The tandem NokA/NokB reactions were carried out by incubation of the NokABCD proteins, encoded by pCY20, with L-tryptophan in the presence of the chaperone teams, encoded by pG-KJE7, as described in **Materials and methods**. The reactions were conducted at various concentrations (0, 5, 10, 20, 30, 40, 50, 60, 80 and 100 mM) of exogenous ammonium hydroxide for different periods of time (0, 1, 3, 4.5, 6, 7.5, and 9 h). Upon termination of the reactions, the reaction mixtures were subsequently subjected to RP-HPLC analysis as described for measurements of CPA (chromopyrrolic acid, **3**) production. (A) the CPA production v.s. [NH_4OH] plots at different, specific reaction times. (B) the CPA production v.s. time plots at different, fixed concentrations of ammonium hydroxide.

Figure S2 The 2D-NMR Spectra of Indole-3-Carboxaldehyde (ICA)

Figure S2a The gCOSY Spectrum of ICA (6)

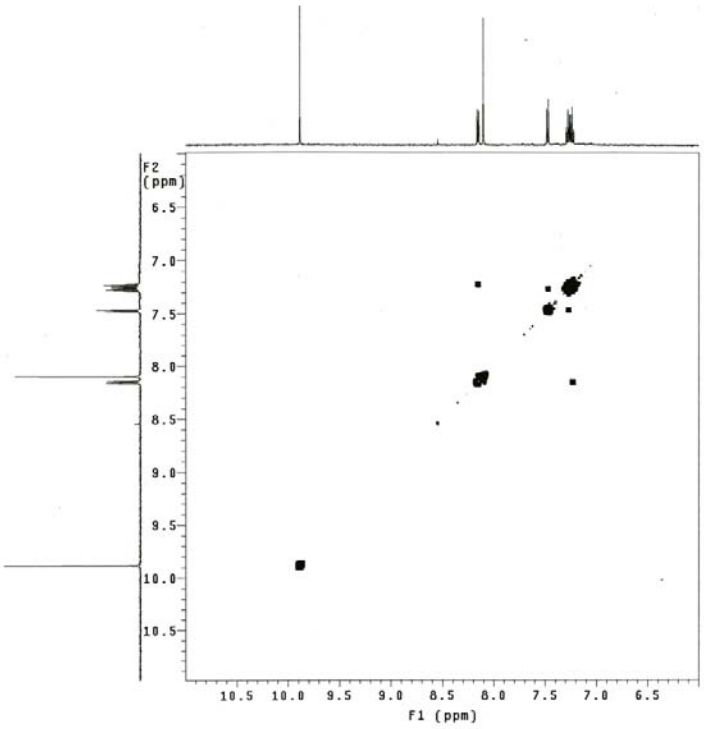


Figure S2b The gHMQC Spectrum of ICA (6)

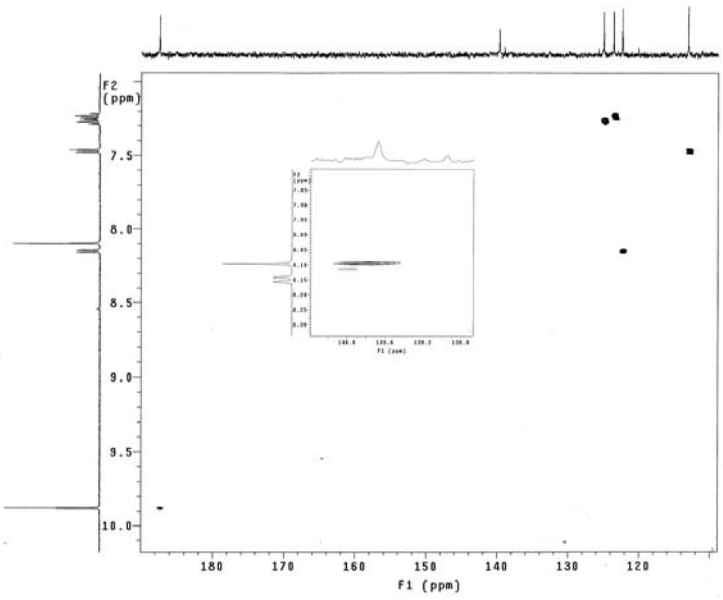


Figure S3 The 2D-NMR Spectra of 9,9'-Difluoro-Chromopyrrolic Acid (2F-CPA)

Figure S3a The gCOSY Spectrum of 2F-CPA (10)

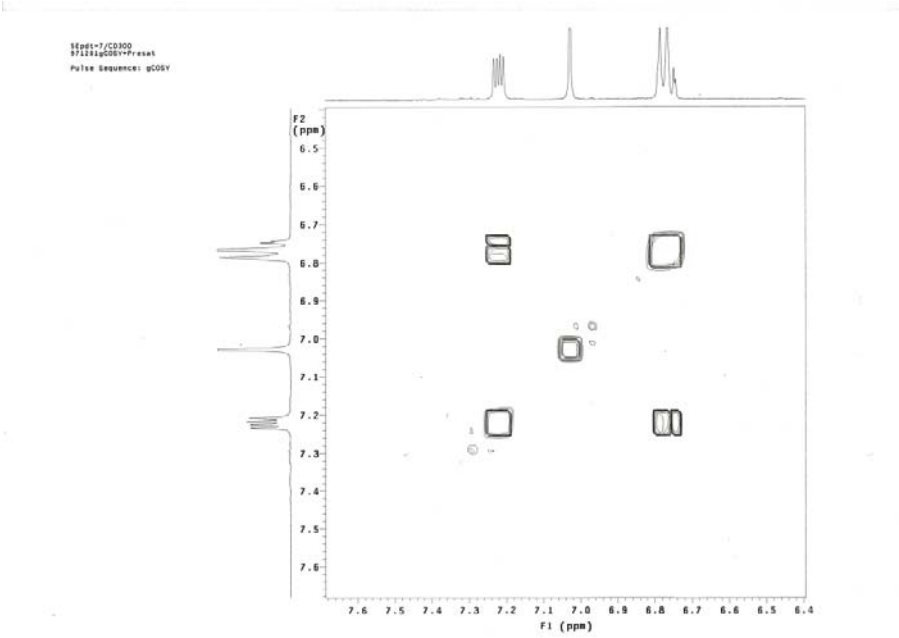


Figure S3b The gHMQC Spectrum of 2F-CPA (10)

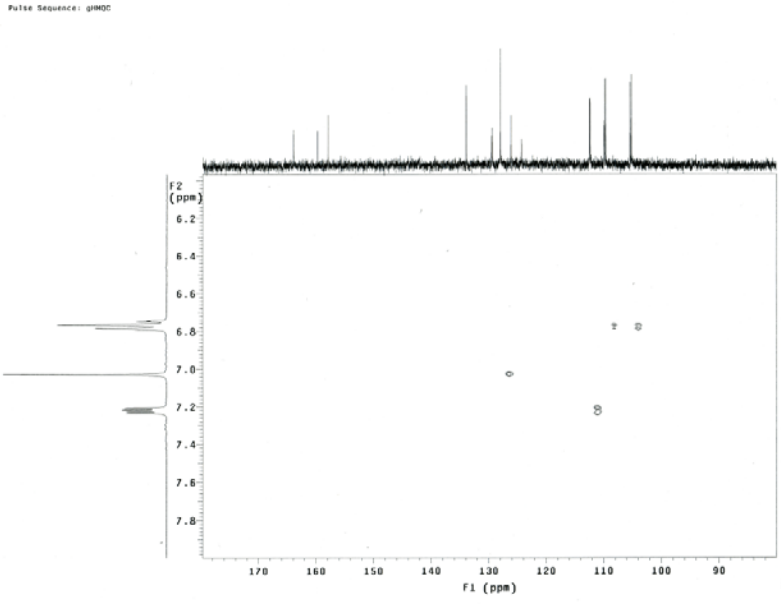


Figure S4 The 2D-NMR Spectra of 5-Fluoro-Indole-3-Carboxaldehyde (F-ICA)

Figure S4a The gCOSY Spectrum of F-ICA

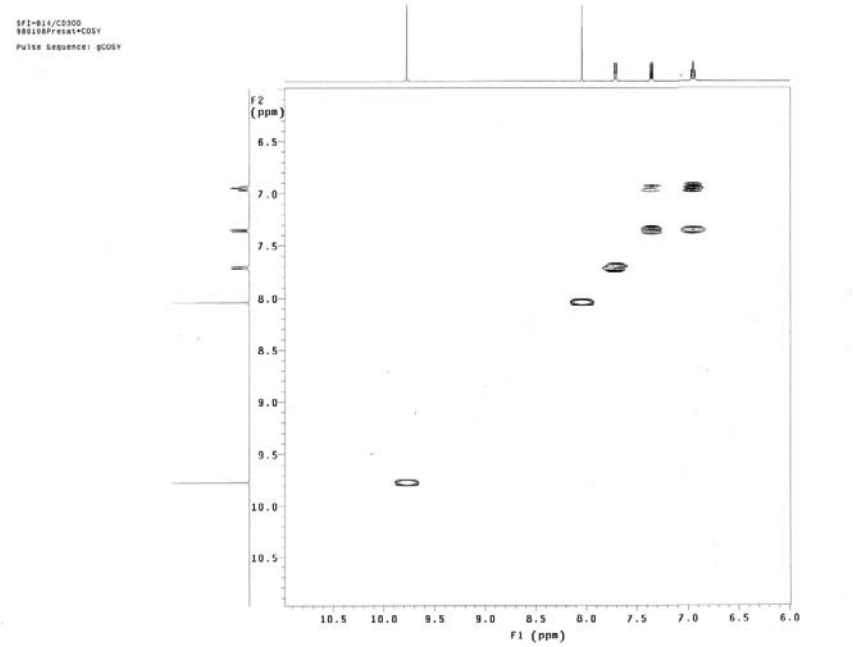


Figure S4b The gHMQC Spectrum of F-ICA

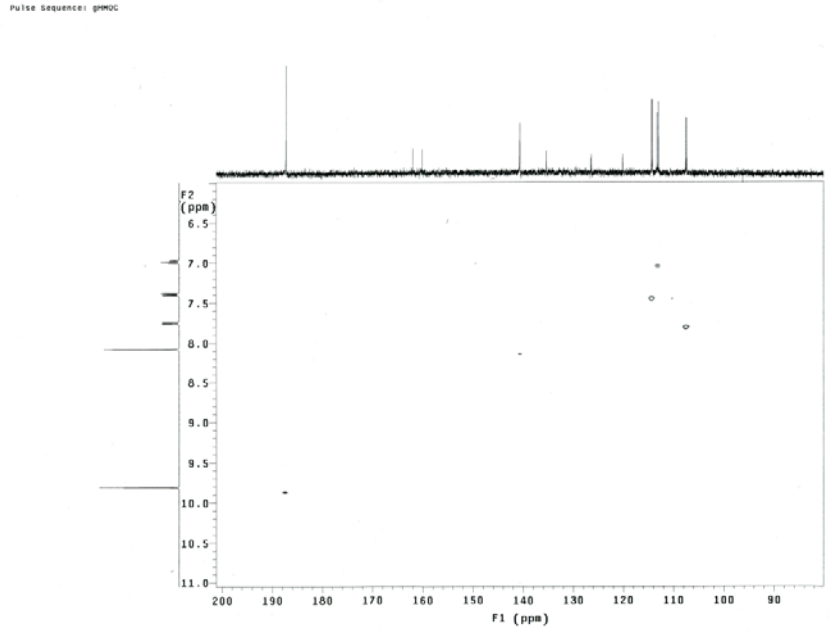
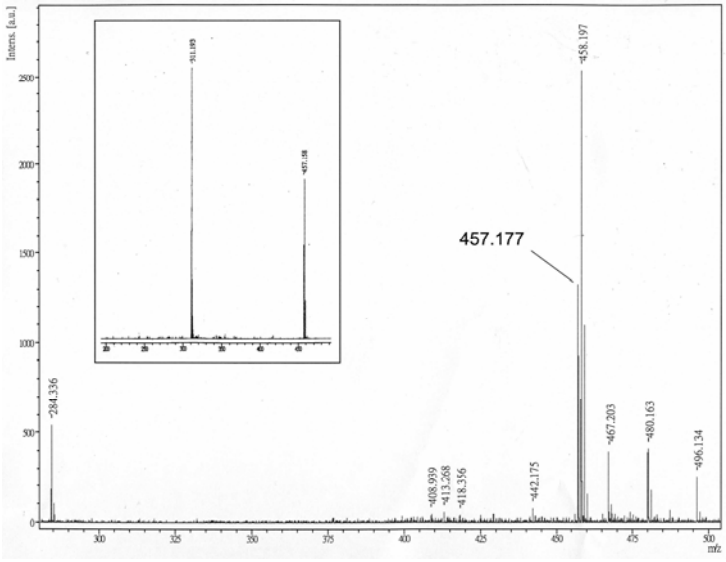
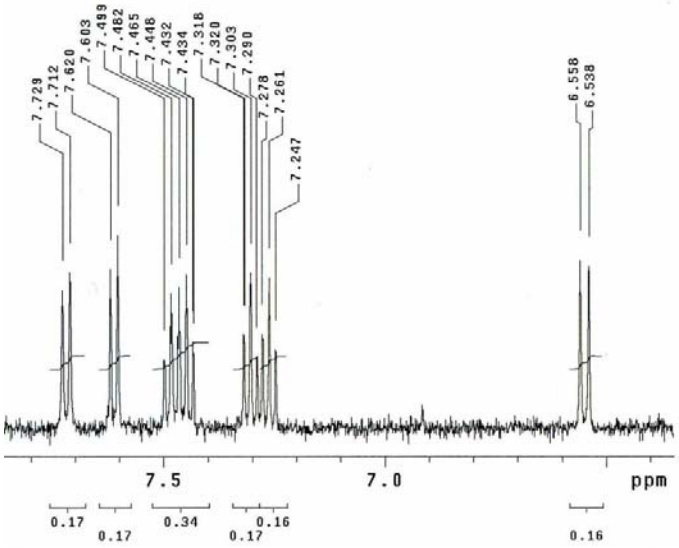


Figure S5 The MALDI-MS/MS Characterization of Rhamnosyl-K252c



The MALDI-TOF analysis of rhamnosyl-K252c **12** ($C_{26}H_{23}N_3O_5$, calcd. M.W., 457.163) gave a pair of parent ion peaks of $[M]^+$ (m/z 457.177) and $[M+H]^+$ (m/z 458.197). The MALDI-MS/MS fragmentation of rhamnosyl-K252c gave rise to a parent ion peak of 311.193 (m/z) corresponding to K-252c (calcd. M.W., 311.337), as shown in the inset.

Figure S6 Characteristic Region of the Rhamnosyl-K252c ^{1}H-NMR Spectrum



The ^{1}H-NMR spectrum showed that characteristic signals of rhamnosyl-K252c (K-252d, **12**) focused on the down-field region and the chemical shifts were identical to those reported by Yasuzawa *et al*.[3] The proton signal (δ_H=6.54) of the anomeric carbon in the L-rhamnose moiety showed a large coupling constant of ca. 10 Hz.

Figure S7 Structures of Aglycones and TDP-sugars Tested in This Work

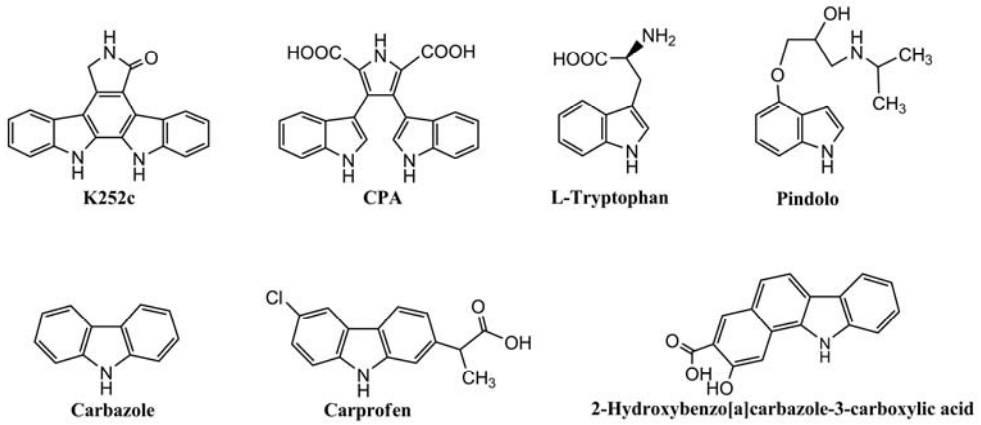


Figure S7a Aglycone Analogues: All the acceptors tested were commercially available, except that CPA is a biosynthetic product made by biotransformation of L-Tryptophan with the *E. coli* harboring the NokABCD expression plasmid.

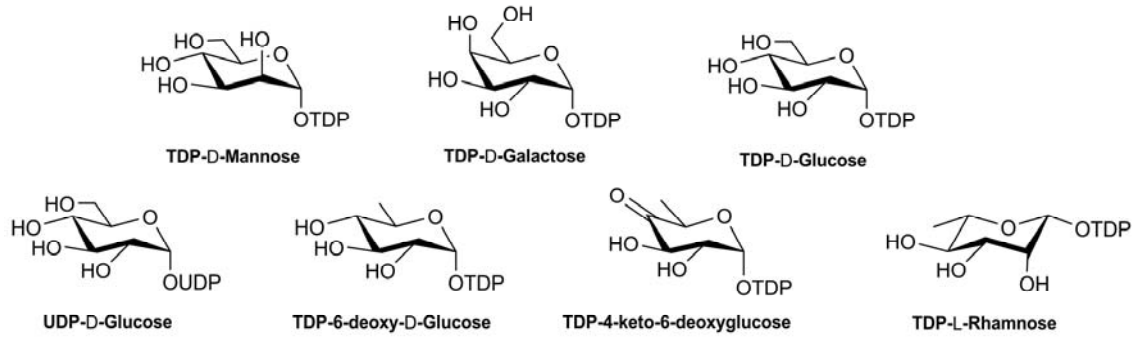


Figure S7b TDP-sugar Analogues: TDP-L-rhamnose was generated as described earlier in Experimental Procedures. TDP-D-mannose and TDP-D-galactose were generated by incubation of α-D-mannose-1-phosphate and α-D-galactose-1-phosphate, respectively, with *Saccharopolyspora spinosa* Gtt in the presence of thymidine triphosphate. TDP-6-deoxy-D-glucose was obtained by catalytic action of *Streptomyces peucetius* DnmV on TDP-4-keto-6-deoxy-D-glucose (KDG) with exogenous NADPH, where KDG was made from enzymatic conversion of TDP-D-glucose with *S. spinosa* Gdh. Gtt (TDP-D-glucose synthase)[4] and DnmV (TDP-4-keto-2,3,6-trideoxyhexulose reductase)[5] were cloned in pET expression vectors, overexpressed in *E. coli*, and purified to near homogeneity for above enzymatic conversions (H.-T. Chiu, unpublished data).

Table S1 Sequences of Primers used for Gene Cloning by PCR

Primer Name	Primer sequence (5'→3' direction)	Restriction site	Gene
NKLNdF1 (forward)	CATATGTTGGCACACGTTCTGATCG	*Nde*I	*nokL*
NKLXR1 (reverse)	CTCGAGCAGCGGGCCTCCGGTG	*Xho*I	*nokL*

restriction sites are underlined in sequence.

References (Supplementary Information)

S1. C. J. Balibar and C. T. Walsh, *Biochemistry*, 2006, **45**, 15444-15457.

S2. A. R. Howard-Jones and C. T. Walsh, *Biochemistry*, 2005, **44**, 15652-15663.

S3. T. Yasuzawa, T. Iida, M. Yoshida and N. Hirayama, *J. Antibiot.* (*Tokyo*), 1986, **39**, 1072-1078.

S4. Y.-L. Chen, Y.-H. Chen, Y.-C. Lin, K.-C. Tsai and H.-T. Chiu, *J. Biol. Chem.*, 2009, **284**, 7352-7363.

S5. S. L. Otten, M. A. Gallo, K. Madduri, X. Liu and C. R. Hutchinson, *J. Bacteriol.*, 1997, **179**, 4446-4450; H.-T. Chiu, Y.-L. Chen, Y.-C. Lin and M.-N. Lee, Functional Characterization and Substrate Specificity of the TDP-4-keto-hexose Reductase DnmV from Daunorubicin Biosynthetic Pathway, *manuscript in preparation*.