Biochemical Characterization of the SgcA1 α-D-Glucopyranosyl-1-phosphate Thymidylyltransferase from the Enediyne Antitumor Antibiotic C-1027 Biosynthetic Pathway and Overexpression of *sgcA1* in *Streptomyces globisporus* to Improve C-1027 Production[⊥,#]

Jeffrey M. Murrell,[†] Wen Liu,[‡] and Ben Shen*,[‡,§]

Department of Chemistry, University of California, Davis, California 95616, and Division of Pharmaceutical Sciences and Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53705

Sequence analysis of the biosynthetic gene cluster for the enediyne antitumor antibiotic C-1027 from *Streptomyces globisporus* has previously suggested that the *sgcA1* gene encodes a α-D-glucopyranosyl-1-phosphate thymidylyltransferase (Glc-1-P-TT) catalyzing the first step in the biosynthesis of the 4-deoxy-4-(dimethylamino)-5,5-dimethyl-D-ribopyranose moiety by activating α-D-glucopyranosyl-1-phosphate (Glc-1-P) into deoxythymidine diphosphate-α-D-glucose (dTDP-Glc). Here we report the overexpression of *sgcA1* in *E. coli*, purification of the overproduced SgcA1 to homogenetity, biochemical and kinetic characterization of the purified SgcA1 as a Glc-1-P-TT, and yield improvement for C-1027 production by overexpression of *sgcA1* and its flanking gene in *S. globisporus*. These findings provide biochemical evidence supporting the genetics-based hypothesis for C-1027 biosynthesis, set the stage for further investigation of the deoxysugar biosynthetic pathway, and demonstrate the utility of sugar biosynthesis genes in natural product yield improvement via combinatorial biosynthesis methods. In contrast to the homotetrameric quaternary structure known for Glc-1-P-TT enzymes from primary metabolic pathways, Glc-1-P-TT enzymes such as SgcA1 from secondary metabolic pathways are monomeric in solution. Sequence differences between the two subclasses of Glc-1-P-TT enzymes were noted. The monomeric structural feature of the latter enzymes could be exploited in engineering Glc-1-P-TT enzymes with broad substrate specificity for structural diversity via the glycorandomization strategy.

The enediynes as a family contain some of the most powerful antitumor agents in existence today.[1–7] C-1027, produced by *Streptomyces globisporus*, is the most potent member of the enediyne family.[8–10] C-1027 belongs to the nine-membered enediyne core subcategory that are chromoproteins consisting of an apo-protein and the enediyne chromophore.[3,4,7] Other members of this subcategory include neocarzinostatin, kedarcidin, maduropeptin, and N199A2, with the latter as the only exception that was isolated as a chromophore alone. The enediynes exert their antitumor activities by undergoing an electronic rearrangement (Bergman or Myers rearrangement) to furnish a transient diradical that damages DNA by abstracting hydrogen atoms from the deoxyribose moiety on both strands. Subsequent reactions of the resultant deoxyribose carbon-centered radicals with molecular oxygen initiate a process that leads to both single-strand and double-strand DNA cleavage. Most of the enediynes also contain one or multiple deoxygenated and often functionalized sugars. While not directly involved in DNA cleavage, these sugar moieties imbue unusual hydrophobic and hydrophilic motifs to the enediynes, playing essential roles in various molecular recognition events for enediynes' in vivo biological activities.[1,3,5,6]

As a model system to study enediyne biosynthesis, we have recently cloned and sequenced the C-1027 biosynthetic gene cluster from *S. globisporus*.[11–14] Genetic analysis of the 85-kb C-1027 gene cluster revealed (1) an iterative modular polyketide synthase for enediyne core biosynthesis and (2) a convergent biosynthetic strategy for the C-1027 chromophore (**1**) from four building blocks, one of which is the 4-deoxy-4-(dimethylamino)-5,5,-dimethyl-D-ribopyranose (**2**) (Figure 1A).[12] Among the genes identified within the C-1027 gene cluster are seven deoxy aminosugar biosynthesis genes, *sgcA, sgcA1, sgcA2, sgcA3, sgcA4, sgcA5,* and *sgcA6,* that are loosely clustered in two loci near the upstream boundary (Figure 1B). On the basis of sequence analysis of the deduced gene products, we have previously proposed a biosynthetic pathway for **2** from α-D-glucopyranosyl-1-phosphate (Glc-1-P, **3**) and predicted that SgcA1 catalyzed the first step for **2** biosynthesis by activating **3** into nucleotide diphosphate (NDP)-α-D-glucose (NDP-Glc), thereby diverting the sugar precursor into secondary metabolite biosynthesis (Figure 1C).[12] We now report the heterologous expression of *sgcA1* in *E. coli* and functional confirmation of SgcA1 as a α-D-glucopyranosyl-1-phosphate thymidylyltransferase (Glc-1-P-TT). Biochemical and kinetic analysis of SgcA1 represents the first complete characterization of a Glc-1-P-TT involved in secondary metabolite biosynthesis in *Streptomyces*. Overproduction and characterization of SgcA1 also allow for the facile preparation of deoxy thymidine diphosphate-α-D-glucose (dTDP-Glc, **4**) in vitro, setting the stage for further investigation of the deoxysugar biosynthetic pathway. Overexpression of *sgcA1* in the *S. globisporus* wild-type strain results in a significant increase in C-1027 production, demonstrating the utility of sugar biosynthesis genes

Figure 1. (A) Structure of the C-1027 chromophore (**1**) and a proposed convergent assembly model for its biosynthesis. (B) Partial C-1027 biosynthetic gene cluster showing the biosynthesis genes (shaded in black) for the deoxy aminosugar moiety (**2**). (C) Proposed biosynthetic pathway for **2** from Glc-1-P (**3**). The SgcA1 Glc-1-P-TT-catalyzed formation of dTDP-Glc (**4**) from **3** is highlighted (boxed).

in natural product yield improvement via combinatorial biosynthesis methods.

Results

Overproduction and Purification of SgcA1. To biochemically characterize SgcA1 in vitro, we overexpressed *sgcA1* (pBS1027) in *E. coli* and purified the overproduced SgcA1 as an N-terminal His$_6$-tagged fusion protein to homogeneity in two steps using affinity and gel-filtration chromatography. Introduction of pBS1027 into *E. coli* BL-21 (DE3) under standard overexpression conditions recommended by the manufacturer resulted in an excellent overproduction of SgcA1 but in almost completely insoluble form, as judged by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Solubility of the overproduced SgcA1 was improved by lowering the fermentation temperature to 18 °C and increasing the incubation time (48 h). Under these conditions, approximately 5% of the overproduced SgcA1 was soluble, from which SgcA1 was partially purified (∼95% pure) by affinity chromatography using Ni-NTA. Further purification by gel-filtration chromatography on a TSK-Gel G300 SWXL column finally afforded pure SgcA1, as judged by SDS-PAGE, where it migrated as a single band with the expected size of 42 kDa (calculated 40 709 Da) (Figure 2). The purified SgcA1, when stored at −80 °C in 50 mM potassium phosphate buffer, pH 7.0, containing 1 mM DTT and 20% glycerol, is stable for several months with minimal loss of activity.

Molecular Weight Determination. The molecular weight for the native form of SgcA1 was determined by gel-
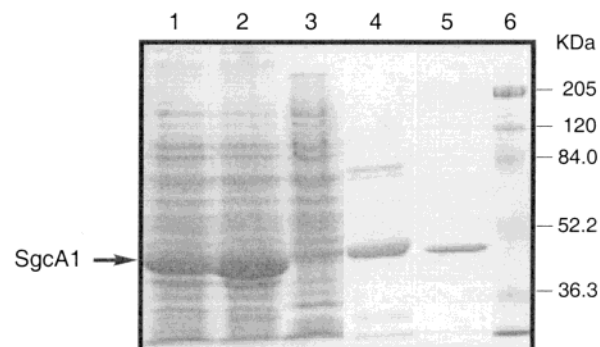
Figure 2. Overexpression and purification of SgcA1 as analyzed by SDS-PAGE: lane 1, total proteins; lane 2, insoluble proteins; lane 3, soluble proteins; lane 4, Ni-NTA-purified protein; lane 5, gel-filtration-chromatography-purified protein; lane 6, molecular weight markers.

filtration chromatography on a TSK-Gel G300 SWXL column, precalibrated with molecular weight standards. The purified SgcA1 was eluted as a single peak with an apparent molecular weight of 48 kDa.

In Vitro Assay of SgcA1 as a Glc-1-P-TT. We adopted an HPLC method using a Sphereclone SAX column with UV detection at 254 nm to assay SgcA1 as a Glc-1-P-TT by following the formation of **4** from dTTP and **3** (Figure 1C).[15−18] Under the optimized HPLC conditions, the substrate (dTTP with retention time at 15.8 min) and product (**4** with retention time at 8.3 min) were well resolved (Figure 3). The identity of **4** was initially established by coelution with an authentic standard and subsequently confirmed by electrospray-ionization mass spectrometry
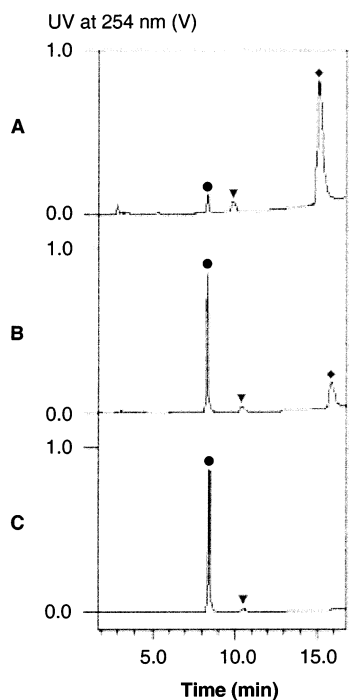
Figure 3. HPLC analysis of the formation of dTDP-Glc (**4**, ●) from Glc-1-P (**3**) and dTTP (◆). Samples were taken after reaction for 15 s (A), 15 min (B), and 1 h (C), respectively. Trace amount of dTTP is hydrolyzed into dTDP (▼) under the assay condition.

Table 1. Effect of Metal Ions on SgcA1 as α-D-Glucopyranosyl-1-phosphate Thymidylyltransferase

| metal ion (400 μM)	activity (%)^a
Mn^{2+}	100
Mg^{2+}	86.4 ± 2.4
Zn^{2+}	23.1 ± 1.1
Cu^{2+}	21.3 ± 0.7
Fe^{2+}	18.9 ± 0.4
Ni^{2+}	11.9 ± 0.6
Co^{2+}	6.4 ± 0.6
no metal (negative control)	0

^{a} Initial velocity relative to Mn^{2+}.

(ESIMS) and ^{1}H NMR analyses: **4** showed a $(M - H)^-$ ion at m/z 563, consistent with the molecular formula $C_{16}O_{16}$-$N_2P_2H_{26}$ (calculated 564) and ^{1}H NMR data identical to those of dTDP-α-D-glucose standard.

pH Dependence. SgcA1 displays a broad pH optimum of 6.0–9.0 for the Glc-1-P-TT activity and shows essentially no dependence on pH for its enzyme activity within the tested pH range of 6.0–9.0.

Metal Ion Requirements. SgcA1 shows an absolute requirement for a divalent metal ion for Glc-1-P-TT activity, as shown in Table 1. The relative reactivity of SgcA1 with various divalent metals was as follows: $Mn^{2+} > Mg^{2+} \gg Zn^{2+} > Cu^{2+} > Fe^{2+} > Ni^{2+} > Co^{2+}$. Although the enzyme is most effective with Mn^{2+} or Mg^{2+}, it demonstrates a small but measurable activity with all other divalent metals tested. Experiments that measure **4** production as a function of the $[Mn^{2+}]/[dTTP]$ ratio indicated that the M^{2+} and dTTP likely form a complex with a 1:1 ratio (Figure 4). This ratio is consistent with the $M^{2+}\cdot$UTP complex utilized by the α-D-glucopyranosyl-1-phosphate uridylyltransferase enzymes.[19−21] Recent X-ray structural analyses of several Glc-1-P-TT enzymes for L-rhamnose biosynthesis also indicate a 1:1 stoichiometry between nucleotide and Mg^{2+} for this family of enzymes.[18,22,23]

Substrate Specificity. The substrate specificity of the purified enzyme was examined for the production of a
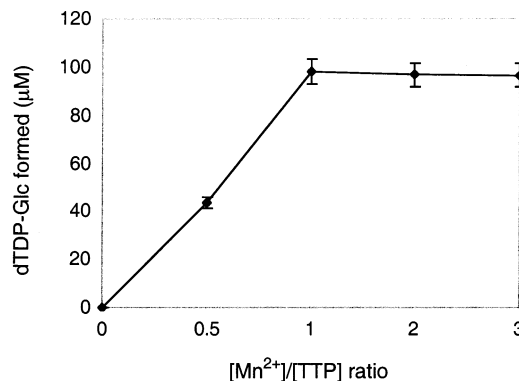
Figure 4. Effect of $[Mn^{2+}]/[dTTP]$ ratio on the SgcA1 Glc-1-P-TT activity.

Table 2. Substrate Specificity of SgcA1 as a α-D-Glucopyranosyl-1-phosphate Thymidylyltransferase

| substrate	activity (%)^a
Nucleotides	
dTTP	100
UTP	6.2 ± 0.7
dUTP	1.3 ± 0.4
ATP	0
GTP	0
CTP	0
ITP	0
D-Hexose-1-phosphates	
α-D-glucose-1-phosphate	100
β-D-glucose-1-phosphate	0
α-D-mannose-1-phosphate	0^b
α-D-galactose-1-phosphate	0
α-D-lactose-1-phosphate	0
α-D-maltose-1-phosphate	0

^{a} Initial velocity relative to dTTP and α-D-glucose-1-phosphate. ^{b} Small amount of dTDP-α-D-mannose was formed with a large excess of both substrates and SgcA1 and prolonged incubation time.

dNDP-D-hexose or an NDP-D-hexose using various dNTP or NTP and D-hexose-1-phosphates as substrates, respectively. As summarized in Table 2, while SgcA1 exhibits somewhat relaxed substrate specificity toward the nucleotides, accepting dTTP, UTP, and dUTP, it is highly specific toward the D-hexose-1-phosphates, accepting essentially only α-D-glucose-1-phosphate. A minute amount of dTDP-α-D-mannose was detected only when the assay was performed with a large excess of both substrates and SgcA1 and prolonged incubation time.

Kinetic Characterization. Double-reciprocal plots of initial velocity as a function of either constant concentrations of dTTP at various levels of **3** or constant concentrations of **3** at various levels of dTTP were used to determine the kinetic mechanism of this reaction.[15,16,18,23] As exemplified in Figure 5A, the double-reciprocal plots of initial velocity against **3** concentration at various levels of dTTP afforded intersecting lines converging at the y-intercept. This pattern is indicative of a sequential mechanism for a bimolecular reaction according to eq 1 (Experimental Section), and the reciprocal of the y-intercept value yields $V_{max} = 4.2 \pm 0.2$ μmol min^{-1} mg^{-1} and $k_{cat} = 2.9 \pm 0.2$ s^{-1}. Re-plotting of the slopes against the reciprocal of dTTP concentrations resulted in a straight line (Figure 5B), further supporting a sequential mechanism, and the y-intercept $[K_m(\text{for } \mathbf{3})/V_{max}]$ and x-intercept $[-1/K_m(\text{for dTTP})]$ values afford the apparent $K_m(\text{for } \mathbf{3}) = 0.33 \pm 0.04$ mM and $K_m(\text{for dTTP}) = 0.20 \pm 0.04$ mM, respectively.

Expression of *sgcA1* in *S. globisporus*. The effect of *sgcA1* on C-1027 production was tested in the wild-type *S.*
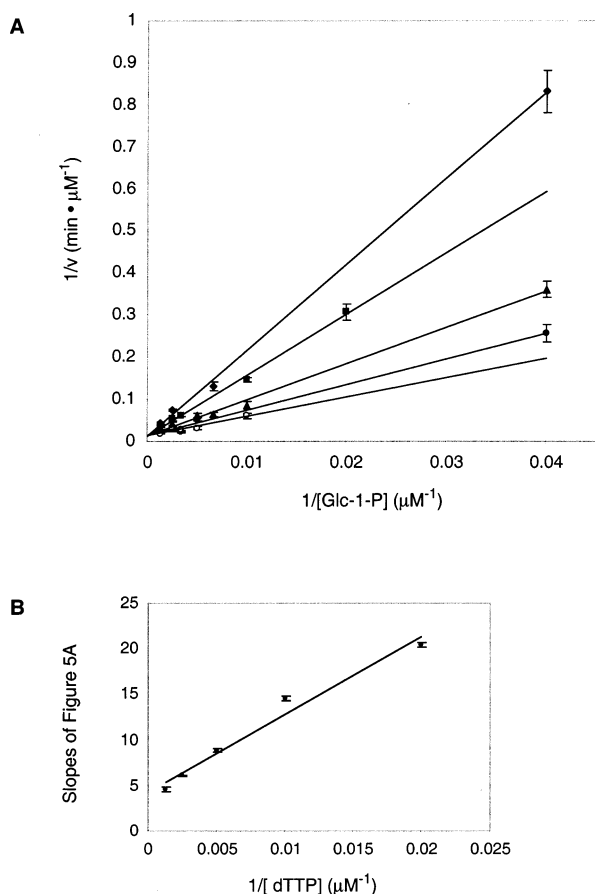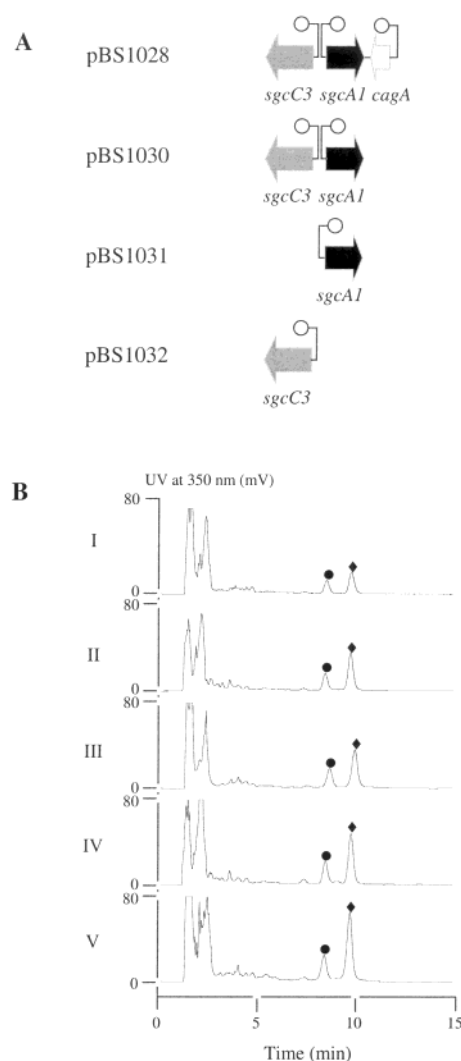
Figure 5. Determination of the steady-state kinetic parameters for the SgcA1-catalyzed formation of dTDP-Glc from Glc-1-P. (A) Double-reciprocal plots of initial velocity as a function of Glc-1-P concentration at varying dTTP concentration: 50 μM ($\blacklozenge$), 100 μM ($\blacksquare$), 200 μM ($\blacktriangle$), 400 μM ($\bullet$), and 800 μM ($\circ$) and (B) re-plotting of the slopes at varying dTTP concentration in part A as a function of the reciprocal of dTTP concentration.

Figure 6. Effect of overexpression of *sgcA1* and its flanking genes on C-1027 production in *S. globisporus*. (A) Constructs for expression of *sgcA1* and its flanking genes and (B) HPLC analysis of C-1027 production by *S. globisporus* wild-type (I) and recombinant strains harboring the overexpression plasmid pBS1031 (II), pBS1032 (III), pBS1030 (IV), or pBS1028 (V), respectively. $\blacklozenge$, C-1027 chromophore; $\bullet$, the aromatized C-1027 chromophore.

globisporus strain. Four different expression vectors were constructed, which harbored *sgcA1* alone (pBS1031), *sgcA1* in combination with its upstream gene *sgcC3* (pBS1030), *sgcA1* with both the upstream gene *sgcC3* and downstream gene *cagA* (pBS1028), and the *sgcC3* gene alone (pBS1032) (Figure 6A). These constructs and the pKC1139 vector as a control were each introduced by conjugation into *S. globisporus*, and the resultant *S. globisporus* recombinant strains were cultured and analyzed by HPLC for C-1027 production.[11,12] While no apparent difference for C-1027 production was observed between the *S. globisporus* wild-type and *S. globisporus* (pKC1139) control strains, a significant increase in C-1027 production was consistently observed for all strains overexpressing the *sgcC3*, *sgcA1*, or *cagA* gene, ranging from 2- to 4-fold. Overexpression of *sgcA1* alone (pBS1031) resulted in at least a 2-fold improvement for C-1027 production over the wild-type *S. globisporus* strain (Figure 6B).

Discussion

Glc-1-P-TT enzymes catalyze the synthesis of **4** from **3** and dTTP, activating glucose to initiate the biosynthesis of various primary and secondary metabolites. This family of enzymes could be grouped into two subclasses according to the pathways in which they participate. Those involved in the biosynthesis of L-rhamnose or other components of cell-surface structures include the RfbA enzymes from *E. coli* K12 (accession no. NP_416543) and *Salmonella enterica* LT2 (accession no. CAD02458) and the RmlA en-

zymes from *Streptococcus mutans* (accession no. BAA11247), *Aneurinibacillus thermoaerophilus* (accession no. AAO-06356), *Pseudomonas aeruginosa* (accession no. CAC82197), and *Mycobacterium tuberculosis* H37Rv (accession no. NP_214848). Those involved in the biosynthesis of secondary metabolites are exemplified by TylA1 for tylosin biosynthesis from *Streptomyces fradiae* (accession no. AAA21343), OleS for oleandomycin biosynthesis from *Streptomyces antibioticus* (accession no. AAD55453), MtmD for mithramycin biosynthesis from *Streptomyces argillaceus* (accession no. CAA07754), and UrdG for urdamycin biosynthesis from *Streptomyces fradiae* (accession no. AAL-06657) (Figure 7). Detailed biochemical characterizations of Glc-1-P-TT enzymes have primarily focused on those involved in the L-rhamnose biosynthetic pathway.[15,16,18,22,23]

We initiated the present study of SgcA1 as a model to garner an understanding of the role and mechanism of a Glc-1-P-TT in secondary metabolite biosynthesis. We have previously proposed SgcA1 as a Glc-1-P-TT on the basis of its significant sequence homology to various known Glc-1-P-TT enzymes such as RfbA, RmlA, TylA1, OleS, UrdG, and MtmD (Figure 7). We now overexpressed *sgcA1* in *E. coli*, purified the overproduced SgcA1 to homogeneity

```
                              1                                                                          75
RfbA  (CAD02458)     -----MKTRKGIILAGGSGTRLYPVTMAVSKQLLPIYDKPMIYYPLSTLMLAGIRDILIISTPQDTPRFQQLLGD
RfbA  (NP_416543)    -----MKMRKGIILAGGSGTRLYPVTMAVSKQLLPIYDKPMIYYPLSTLMLAGIRDILIISTPQDTPRFQQLLGD
RmlA  (CAC82197)     -----MK-RKGIILAGGSGTRLHPATLAISKQLLPVYDKPMIYYPLSTLMLAGIREILIISTPQDTPRFQQLLGD
RmlA  (AAO06356)     --------MKGIILAGGSGTRLYPLTRAISKQLLPVYDKPMIYYPLSVLMLAGIRDILVISTPEDTPRFEQLLGD
RmlA  (BAA11247)     --------MKGIILAGGSGTRLYPLTRAASKQLMPVYDKPMIYYPLSTLMLAGIKDILIISTPQDLPRFKELLQD
RmlA  (NP_214848)    --------MRGIILAGGSGTRLYPITMGISKQLLPVYDKPMIYYPLTTLMMAGIRDIQLITTPHDAPGFHRLLGD
TylA1 (AAA21343)     MNDRPRRAMKGIILAGGSGTRLRPLTGTLSKQLLPVYDKPMIYYPLSVLMLAGIREIQIISSKDHLDLFRSLLGE
OleS  (AAD55453)     --------MKALVLAGGSGTRLRPITHTSAKQLVAVANKPVLFYGLEAIAAAGITDVGLIVG-DTAGEVPRAVGD
MtmD  (CAA07754)     --------MKALVLSGGSGTRLRPITHTSPKQLVPVANKPVLFYALEDIAQAGITDVGIVVG-ETAREIEDAVGD
UrdG  (AAF00210)     --------MKALVLAGGSGTRLRPFSYSMPKQLIPIANTPVLVHVLRNLRELGVTEAGVIVG-NRGPEISAVLGD
SgcA1 (AAL06657)     --------MKALVLSGGSGTRLRPISYAMPKQLVPIAGKPVLEYVLDNIRNLDIKEVAIVVG-DWAQEIIEAMGD
Consensus                    KgIILAGGSGTRL P T   sKQLLpV dkPMI Y L  L   gI D  II   q     f    L D

                              76                                                                        150
RfbA  (CAD02458)     GSQWGLNLQYKVQPSPDGLAQAFIIGEEFIGNNDCALVLGDNIFYGHDLPKLMEAAVNKESGATVFAYHVNDPER
RfbA  (NP_416543)    GSQWGLNLQYKVQPSPDGLAQAFIIGEEFIGGDCALVLGDNIFYGHDLPKLMEAAVNKESGATVFAYHVNDPER
RmlA  (CAC82197)     GSNWGLDLQYAVQPSPDGLAQAFLIGESFIGNDLSALVLGDNLYYGHDFHELLGSASQRQTGASVFAYHVLDPER
RmlA  (AAO06356)     GSELGISLSYAIQPSPDGLAEAFIIGEDFIGQDNVALILGDNIFYGQGFTPLLQRAVNRKKGATIFANVVKDPER
RmlA  (BAA11247)     GSEFGIKLSYAEQPSPDGLAQAFIIGEEFIGDDHVALILGDNIYYGPGLSRMLQKAASKESGATVFGYQVKDPER
RmlA  (NP_214848)    GAHLGVNISYATQDQPDGLAQAFVIGANHIGADSVALVLGDNIFYGPGLGTSLKRFQS-ISGGAIFAYWVANPSA
TylA1 (AAA21343)     GDRLGLSISYAEQREPRGIAEAFLIGARHIGGDDAALILGDNVFHGPGFSSVLTGTVARLDGCELFGYPVKDAHR
OleS  (AAD55453)     GAKFGLDITYIEQSRPLGLAHAVLIAHTYLGDDDFVMYLGDNIFVGGIDDLVRTFRDGRRPAARILLTHVSDPSG
MtmD  (CAA07754)     GSKFGLQVTYLPQEKPLGLAHAVLIARDYLGDDDFVMYLGDNFVVGGIAGLVDEFRADR-PDARILLTPVPDPRS
UrdG  (AAF00210)     GAEFGMRVTYIPQDAPRGLAHTVAITRDFLCDDDFVMYLGDNMLPDGVVDMRGGIHRPS-AGGQVVVYKVPDPRS
SgcA1 (AAL06657)     GSRFGLRLTYIRQEQPLGIAHCVKLARDFLDEDDFVLYLGDIMLDGDLSAQAGHFLHTR-PAARIVVRQVPDPRA
Consensus            G q GL L Y  Q  P GLAq f I   fI  D  aL LGD i  g              Vf    V Dp

                              151                                                                       225
RfbA  (CAD02458)     YGVVEFDQNGTAVSLEEKPLQPKSNYAVTGLYFYDNSVVEMAKNLKPSSRGELEITDINRIYMDQGRLSVAMMGR
RfbA  (NP_416543)    YGVVEFDKNGTAISLEEKPLEPKSNYAVTGLYFYDNDVVQMAKNLKPSARGELEITDINRIYLEQGRLSVAMMGR
RmlA  (CAC82197)     YGVVEFDQGGKAISLEEKPLEPKSNYAVTGLYFYDQQVVDIARDLKPSPRGELEITDVNRAYLERGQLSVEIMGR
RmlA  (AAO06356)     FGVVEFDENNKVLSIEEKPVYPKSNFAVTGLYFYDNDVVKIAKQIEPSARGELEITDVNKIYLERGDLNVELLGR
RmlA  (BAA11247)     FGVVEFDNDRNAISIEEKPEHPKSHYAVTGLYFYDNSVVDIAKNIKPSPRGELEITDVNKAYLDRGDLSVEVMER
RmlA  (NP_214848)    YGVVEFGAEGMALSLEEKPVTPKSNYAVPGLYFYDNDVIEIARGLKKSARGEYEITEVNQVYLNQGRLAVEVLAR
TylA1 (AAA21343)     YGVGEIDSGGRLLSLEEKPRRPLE-PGRHRLYLYTNDVVEIARTISPSARGELEITDVNKVYLEQGR-AAHGAGA
OleS  (AAD55453)     FGVAELDDDGRVVGLEEKPRHPKSDLALVGVYFFTPAIHEAVRAIEPSWRGELEITHAIQHLIDNG-ADIQSMVI
MtmD  (CAA07754)     FGIAELDGEGRVVRLEEKPTHPKSDMAVIGVYLFTPLVHEAVRAIEPSWRGELEITHAIQWLIDNR-HDVRSTVI
UrdG  (AAF00210)     FGVAELGPQGEVLRLVEKPRDPRSDMALVGVYFFTAAIHDAVAAIEPSARGELEITDAIQWLVSSG-ADVRASQY
SgcA1 (AAL06657)     FGVIELDGEGRVLRLVEKPREPRSDLAAVGVYFFTADVHRAVDAISPSRRGELEITDAIQWLLEQG-LPVEAGRY
Consensus            YGV EFd    V L EKP  P s  a   LYFY   V    a Ik S RGE EITd R  Ld

                              226                                                                       300
RfbA  (CAD02458)     GYAWLDTGTHQSLIEASNFIATIEERQGL-----------KVSCPEEIAFRKNFINAQQVIKLAGPLSKNDYGKY
RfbA  (NP_416543)    GYAWLDTGTHQSLIEASNFIATIEERQGL-----------KVSCPEEIAFRKGFIDVEQVRKLAVPLIKNNYGQY
RmlA  (CAC82197)     GYAWLDTGTHDSLLEAGQFIATLENRQGL-----------KVACPEEIAYRQKWIDAAQLEKLAAPLAKNGYGQY
RmlA  (AAO06356)     GFAWLDTGTHESLLEASQFIQTIEKRQGL-----------KVACLEEIAYYKGYINKEQLIRLAEPLKKNEYGAY
RmlA  (BAA11247)     GFAWLDTGTHESLLEAAQYIETVQRMQNL-----------QVANLEEIAYRMGYITADQVRELAQPLKKNEYGQY
RmlA  (NP_214848)    GTAWLDTGTFDSLLDAADFVRTLERRQGL-----------KVSIPEEVAWRMGWIDDEQLVQRARALVKSGYGNY
TylA1 (AAA21343)     VVAWLDMGTHDSLLQAGQYVQLLEQRQGE┌----------┐RIACIEEIAMRMGFISAEQCYRLGQELRSSSYGSY
OleS  (AAD55453)     EGYWKDTGNVADMLEVNRTVLEDLEPRIEGTVDEHTVVIGRVVVGEGARVTNSRIMGPAIIGAGPEISDSYIGPF
MtmD  (CAA07754)     HGYWKDTGNVMDMLEVNRTVLEGVEPCTEGPVDAESEVIGRVRIEAGARVSNSRIVGPAVIGSGTVISGSYIGPY
UrdG  (AAF00210)     DGYWKDTGNVEDVLECNRHLLDGLGARMDGTVDAASVLTGEVVIEPGAKVVRSRIMGAAIIGAGTVVQDSHVGPH
SgcA1 (AAL06657)     TDYWKDTGRVEDVVECNRRMLGRLALQVSGEVDPESELVGAVVVEEGARVTRSRVVGPAVIGAGTVVEDSQIGPY
Consensus                 W D G  sLIE    I               V               I        a L n G y

                              301                                                                       366
RfbA  (CAD02458)     LLKMVK------------GL------------------------------------------------
RfbA  (NP_416543)    LYKMTK------------DSN-----------------------------------------------
RmlA  (CAC82197)     LKRLLT------------ETVY----------------------------------------------
RmlA  (AAO06356)     LLELAEKGSLQTGSINTKEAVLYG--------------------------------------------
RmlA  (BAA11247)     LLRLIG------------EA------------------------------------------------
RmlA  (NP_214848)    LLELLERN----------------------------------------------------------
TylA1 (AAA21343)     IIDVAMRG┌------AAADSRAQ---------------------------------------------
OleS  (AAD55453)     TSVGDNCR│ITGSEMEFSIMLAESAITGVRRIEGSLIGRNVQVTQSLHAPNAHRFVLGDHSKVEIQS
MtmD  (CAA07754)     TSVSEGCR│IEDSEIEFSIVLRDSRVRGVQRVRGSLIGRSVTVAPAPRKPAGTRLVLGDHSEGQISS
UrdG  (AAF00210)     ASIGRRCT│VTDSRVENSIALDEASVTGVRGLRNSLIGRSASVGGTGQENERYRLVVGDHTRVEVAA
SgcA1 (AAL06657)     ASIGRRCT└VRASRLSDSIVLDDASILAVSGLHGSLIGRGARIAPGARGEARHRLVVGDHVQIEIAA┘
Consensus
```

Figure 7. Amino acid sequence alignment of representative Glc-1-P-TT enzymes from primary metabolic pathways (RfbA and RmlA) and secondary metabolic pathways (TylA1, OleS, MtmD, UrdG, SgcA1) with their accession numbers given in parentheses. The conserved amino acid residues are highlighted in bold, and the major sequence differences at the C-terminal regions between the two subclasses of enzymes are highlighted by rectangular boxes.

(Figure 2), and demonstrated that SgcA1, as a Glc-1-P-TT, catalyzes the synthesis of **4** from **3** and dTTP in vitro (Figure 3). These results confirm that SgcA1 participates in the biosynthesis of **2** from **3** by catalyzing the formation of **4** as the first biosynthetic intermediate, setting the stage for future characterization of **2** biosynthesis in vitro (Figure

1C). Increased production of C-1027 upon overexpression of *sgcA1* in *S. globisporus* demonstrates the feasibility of improving enediyne production by rational manipulation of genes governing deoxysugar biosynthesis (Figure 6).

SgcA1 absolutely requires a divalent metal for the Glc-1-P-TT activity, with Mn^{2+} being the preferred metal (Table

1 and Figure 4), and exhibits a broad pH optimum between pH 6.0 and 9.0. These properties are well conserved among Glc-1-P-TT enzymes characterized to date, although the divalent metal preference varies from Mn^{2+} to Mg^{2+} and even Co^{2+}.[15-18,22,23] SgcA1 shows somewhat relaxed substrate specificity toward the nucleotides, accepting dTTP, UTP, and dUTP as substrates, but it is highly specific toward the D-hexose-1-phosphate, accepting only **3** (Table 2). This is in contrast to the RfbA Glc-1-P-TT from *S. enterica*, which shows a broad substrate specificity, accepting a variety of D-hexose-1-phosphate substrates.[18,24,25] This strict D-hexose-1-phosphate specificity was also observed for ChlD involved in chlorothricin biosynthesis,[17] the only other Glc-1-P-TT from a secondary metabolic pathway whose substrate specificity has been determined biochemically.

Although an early study of the RfbA Glc-1-P-TT from *S. enterica* suggested a "ping-pong bi-bi" mechanism,[15] recent X-ray structural analysis of several Glc-1-P-TT enzymes from *E. coli*,[23] *P. aeruginosa*,[22] and *S. enterica*[18] and reevaluation of their kinetic mechanisms have unambiguously established a sequential ordered bi-bi mechanism for this family of enzymes. SgcA1 also proceeds through a rapid equilibrium sequential ordered bi-bi mechanism (Figure 5), suggesting that Glc-1-P-TT enzymes involved in primary and secondary metabolite biosynthesis apparently share a common kinetic mechanism.[15,18,22,23]

While the emerging experimental data clearly suggest that Glc-1-P-TT enzymes involved in primary and secondary metabolite biosynthesis share common chemical and kinetic mechanisms, they appear to differ in quaternary structure. Recent X-ray structures of three Glc-1-P-TT enzymes involved in L-rhamnose biosynthesis confirmed that they are homotetramers.[18,22,23] In contrast, the SgcA1 Glc-1-P-TT reported in this work as well as the ChlD Glc-1-P-TT involved in chlorothricin biosynthesis,[17] the only two Glc-1-P-TT enzymes from secondary metabolic pathways whose quaternary structures have been determined, are monomeric in solution. This quaternary structural difference is consistent with sequence analysis of Glc-1-P-TT enzymes from both primary and secondary metabolic pathways using data obtained from the X-ray structures of three Glc-1-P-TT enzymes involved in L-rhamnose biosynthesis.[18,22,23] While all Glc-1-P-TT enzymes share the same catalytic residues regardless of their quaternary structures, several key differences exist (Figure 7). The most obvious difference occurs in the C-terminus of the proteins. Those from secondary metabolic pathways, such as OleS, MtmD, UrdG, and SgcA1, contain 11 extra amino acids around amino acid 250 and an elongated C-terminus of approximately 50 amino acids (Figure 7, boxed). The homology in this region is very weak compared to the rest of the proteins. These sequence variations may contribute to their quaternary structural differences because the X-ray structures of the Glc-1-P-TT enzymes involved in L-rhamnose biosynthesis revealed that the C-terminal region is the dimerization subdomain. The interface between the subunits of the tetrameric enzymes contains a nonspecific secondary dTTP binding site, postulated to be involved in allosteric regulation by dTDP-L-rhamnose. The C-terminus of the monomeric proteins lacks conserved residues involved in this secondary dTTP binding site, indicating differences in the regulatory mechanism of these two subclasses of proteins depending on the metabolic fate of their products. It is interesting to note that the TylA1 Glc-1-P-TT from the tylosin biosynthetic pathway contains many features conserved within the Glc-1-P-TT enzymes involved in L-rhamnose biosynthesis including the characteristic C-terminal domain. TylA1 may be an evolutionary link between the tetrameric proteins involved in primary metabolite biosynthesis and monomeric proteins involved in secondary metabolite biosynthesis. Very recently, the RfbA Glc-1-P-TT from *S. enterica* LT2 has been successfully engineered to accept a broad range of hexopyranosyl phosphates for the preparation of "unnatural" UDP- and TDP-nucleotide sugars for natural product structural diversity via the glycorandomization strategy.[18,24,25] It is tempting to propose that the monomeric Glc-1-P-TT enzymes from secondary metabolic pathways such as SgcA1 may hold equal or even greater potential than their counterparts from primary metabolic pathways in engineering of Glc-1-P-TT enzymes with a broad substrate specificity, given their structural simplicity.

Experimental Section

Strains, Plasmids, Biochemicals, and Chemicals. *Escherichia coli* DH5α and *E. coli* BL-21(DE3) (Novagen, Madison, WI) were used as hosts for subcloning and heterologous expression, respectively, *E. coli* S17-1[26] was used as the donor strain for conjugation, and *S. globisporus* wild-type strain has been described previously.[11,12] Litmus 28 was from New England Biolabs (Beverly, MA), pET28a(+) was from Novagen, and pBS1004,[11] pBS1008,[11] and pKC1139[26] were described previously. Ampicillin, kanamycin, nucleotide triphosphates, hexose-1-phosphates, and metal salts were from Sigma (St. Louis, MO) and used directly without further purification. Unless specified, common chemicals, molecular weight standards, Ni-NTA resin, restriction enzymes, DNA ligase, and other materials for recombinant DNA procedures were purchased from standard commercial sources and used as provided.

DNA Isolation, Sequencing, Manipulation, and PCR Conditions. Plasmid preparation was carried out using commercial kits (Qiagen, Santa Clarita, CA). Automated DNA sequencing was carried out on an ABI Prism 377 DNA sequencer with an ABI Prism dye terminator cycle sequencing ready reaction kit and Ampli*Taq* DNA polymerase FS (Perkin-Elmer/ABI, Foster City, CA). Sequencing service was provided by Davis Sequencing Inc. (Davis, CA), and the data were analyzed by ABI Prism Sequencing 2.1.1 software and the Genetics Computer Group software (Madison, WI). General procedures for restriction digests, ligations, genetic manipulations in *E. coli* and *S. globisporus*, conjugation between *E. coli* S17-1 and *S. globisporus*, and selection of exoconjugants were performed according to standard protocols[27,28] or previously described methods.[11,12] PCR primers were synthesized at the Protein Structure Laboratory, University of California at Davis. PCR was carried out on a Gene Amp PCR System 2400 (Perkin-Elmer/ABI) with *Pfu* polymerase and buffer from Stratagene (La Jolla, CA). The PCR mixture consisted of 5 ng of *S. globisporus* cosmid DNA as a template, 25 pmol of each of the primers, 25 μM deoxynucleoside triphosphates, 5% dimethyl sulfoxide, 2 U of *Pfu* polymerase, and 1× buffer in a final volume of 100 μL. The PCR temperature program was as follows: initial denaturation at 97 °C for 5 min, 30 cycles of 45 s at 97 °C, 45 s at 60 °C, and 4 min at 72 °C for extension, and an additional 10 min at 72 °C for termination.

Overexpression of *sgcA1* in *E. coli* and SgcA1 Purification. The following primers were used to amplify *sgcA1* from pBS1004: 5′-CG**GGATCC***CATATG*AAGGCACTTGTACTGTC-3′ (the engineered *Bam*HI and *Nde*I sites are in bold and italic, respectively) and 5′-T*GAATTC*TCAGGCCGCGATCTCGAT-3′ (the engineered *Eco*RI site is italic). The resultant product was cloned as a 1.1 kb *Bam*HI-*Eco*RI fragment into the same sites of Litmus 28 to yield pBS1026 and sequenced to confirm PCR fidelity. The *sgcA1* gene was then moved as a 1.1 kb *Nde*I-*Eco*RI from pBS1026 into the same sites of pET28a(+) to afford pBS1027. The latter resulted in the production of SgcA1 as an N-terminal His₆-tagged fusion protein.

pBS1027 was introduced into *E. coli* BL21(DE3) for *sgcA1* overexpression following the standard conditions recommended by the manufacturer (Novagen). The cells were initially grown at 37 °C until the cell density reached an OD_{600} of approximately 0.6, and *sgcA1* expression was induced by the addition of 0.1 mM IPTG followed by additional incubation at 30 °C for 3−4 h. To improve SgcA1 solubility, the incubation temperature after IPTG induction was lowered to 16−18 °C, and cultures were grown overnight.

To purify the overproduced SgcA1 protein, cells were harvested by centrifugation (6000*g*, 10 min, 4 °C), resuspended in 50 mM potassium phosphate buffer, pH 8.0, containing 300 mM NaCl and 10 mM imidazole, and lysed by the addition of lysozyme (1 mg/mL). This mixture was shaken for 15 min at room temperature and then subjected to sonication (6 × 10 s on ice with 30 s intervals, Microson Ultrasonic Cell Disruptor XL 2000, Misonix, Farmingdale, NY). The cell-free extract was isolated by centrifugation (12000*g*, 15 min, 4 °C). The soluble fraction of the overproduced SgcA1 was mixed with Ni-NTA resin for 1 h and purified according to the protocol provided by the manufacturer (Qiagen) to afford an approximately 95% pure SgcA1 preparation. The latter was further purified to homogeneity by gel-filtration chromatography on a TSK-Gel G3000 SWXL column (5 *μ*m, 7.8 × 300 mm, TOSOH Biosciences, Montgomeryville, PA), eluted with 50 mM potassium phosphate buffer, pH 8.0, containing 300 mM NaCl. The final pure SgcA1 protein was dialyzed into 50 mM potassium phosphate buffer, pH 7.0, containing 1 mM DTT and 20% glycerol and stored at −80 °C until used. Protein concentration was determined by the Coomassie brilliant blue G-250 dye-binding method of Bradford.[29]

Molecular Weight Determination. The molecular weight of the native SgcA1 was determined by gel-filtration chromatography on the TSK-Gel G3000 SWXL column, eluted with 50 mM potassium phosphate buffer, pH 8.0, containing 300 mM NaCl. The column was calibrated with blue dextran [molecular weight 2×10^6 as the void volume (V_o) standard], and *β*-amylase (200 000), alcohol dehydrogenase (150 000), albumin (66 000), carbonic anhydrase (29 000), and cytochrome C (12 400) as the molecular weight standards. The apparent molecular mass of the SgcA1 enzyme was estimated from a plot of the elution volume (V_e) divided by V_o against the logarithm of the molecular weight of protein standards.

Enzyme Assaays. The Glc-1-P-TT activity of SgcA1 was determined by monitoring the conversion of **3** and dTTP to **4** and pyrophosphate (PP_i) by HPLC according to a previously described method with modifications.[15−18] A typical assay solution (50 *μ*L), containing 400 *μ*M $MnCl_2$, 200 *μ*M **3**, 200 *μ*M dTTP, and variable amounts of enzyme (0.05−50 *μ*g) in 50 mM potassium buffer, pH 7.0, was incubated at 37 °C for 10 min. Aliquots were boiled for 2 min to stop the reaction and centrifuged (14000*g*, 5 min) to pellet the denatured enzyme. The supernatant was collected and stored at −20 °C until analysis by HPLC. Samples (20 *μ*L) were analyzed by HPLC on an Sphereclone SAX column (5 *μ*m, 4.6 × 250 mm, Phenomenex, Torrance, CA), developed with a linear gradient of 60 to 240 mM potassium phosphate buffer, pH 5.0, over 15 min at a flow rate of 1.0 mL/min with UV detection at 254 nm. The identity of **4** was first established by HPLC coelution with an authentic standard. The assay mixture was then subjected to a semipreparative Sphereclone SAX column (5 *μ*m, 10 × 250 mm, Phenomenex, Torrance, CA), eluted with a linear gradient from 0 to 250 mM ammonium formate buffer, pH 3.5, in 20 min at a flow rate of 2.0 mL/min, to isolate **4** for ESIMS and ¹H NMR analyses. The purified **4** was adjusted to pH 7.0 using NH_4OH, subsequently lyophilized, and stored at −20 °C until needed. ESIMS was performed on a LCQ Decca instrument (Thermo-Finningan, San Jose, CA) in a negative ion mode, giving a $(M - H)^-$ ion at *m/z* 563. ¹H NMR was carried out on a Varian Unity Inova 400 spectrometer (400 MHz, D_2O): nucleotide base, *δ* 1.91 (3H, br s, Me-5″), *δ* 7.73 (1H, d, *J* = 1.4 Hz, H-6″); ribose, *δ* 6.33 (1H, t, *J* = 7.0 Hz, H-1′), *δ* 2.36 (2H, m, H's-2′), *δ* 4.61 (1H, m, H-3′), *δ* 4.16 (3H, m, H-4′ and H's-5′); hexose, *δ* 5.58 (1H, dd, *J* = 7.4 and

3.6 Hz, H-1), *δ* 3.4−3.6 (2H, m, H-2 and H-4), *δ* 3.7−3.9 (4H, m, H-3, H-5, H's-6).

pH Dependence. A series of assays were carried out in buffers with 0.5 pH unit increments according to the conditions outlined in the Enzyme Assay section. Potassium phosphate buffer (50 mM) was used for measurements ranging from pH 6.0 to 7.5, and Tris-HCl buffer (50 mM) was used for measurements ranging from pH 7.5 to 9.0.

Metal Ion Requirement and Substrate Specificity. To study the metal ion requirement for SgcA1, $MgCl_2$, $ZnCl_2$, $CuCl_2$, $FeCl_2$, $NiCl_2$, or $CoCl_2$, was substituted for $MnCl_2$, respectively, according to conditions outlined in the Enzyme Assays section. Determination of the optimal metal ion and nucleotide ratio was carried out under the same conditions as outlined in the Enzyme Assays section but with variable ratios of Mn^{2+} to dTTP. The dTTP concentration was kept at a constant 200 mM while the Mn^{2+} concentration was varied from 0 to 600 mM, yielding $[Mn^{2+}]/[dTTP]$ ratios of 0, 0.5, 1, 2, and 3. To investigate the nucleotide specificity for SgcA1, UTP, dUTP, ATP, GTP, CTP, or ITP was substituted for dTTP, respectively, according to conditions outlined in the Enzyme Assays section. To examine the hexose-1-phosphate specificity, *β*-D-glucose-1-phosphate, *α*-D-mannose-1-phosphate, *α*-D-galactose-1-phosphate, *α*-D-maltose-1-phosphate, or *α*-D-lactose-1-phosphate was substituted for **3**, respectively, according to conditions outlined in the Enzyme Assays section.

Kinetics. Measurements of kinetic data were performed using the HPLC methods as outlined in the Enzyme Assays section. The assay solution consisted of 50 mM KH_2PO_4, pH 7.0, 1.8 mM $MnCl_2$, 50−800 *μ*M dTTP, 25−800 *μ*M **3**, and 1.8 *μ*g of enzyme in 100 *μ*L total volume. Excess Mn^{2+} at this concentration does not inhibit the enzyme. HPLC analysis on the Sphereclone SAX column was expedited by a linear gradient of 180−240 mM KH_2PO_4, pH 5.0, in 8 min at a flow rate of 1.0 mL/min with UV detection at 254 nm. Under the latter conditions, dTTP and **4** were eluted with retention times of 8.5 and 4.9 min, respectively. The experiments were designed with a variation of one substrate in the presence of a constant concentration of the second substrate. Determination of the kinetic mechanism and constants was carried out as described by Segel.[30] The experimental data were plotted according to

$$\nu = V_{max} [A][B]/(K_A K_B + K_B[A] + [A][B]) \qquad (1)$$

where [A] and [B] represent the concentration of dTTP and **3**, respectively, for a rapid equilibrium sequential ordered bi-bi mechanism to estimate the apparent Michaelis−Menten constants.

Overexpression of *sgcA1* in *S. globisporus* and C-1027 Production, Isolation, and Analysis. Sequence analysis of the C-1027 biosynthetic gene cluster has previously identified[12] promoter regions upstream of the *cagA*, *sgcA1*, or *sgcC3* gene, respectively, and four different expression constructs were made in pKC1139 for expression in *S. globisporus*. pBS1028 was constructed by cloning a 4.2 kb *Eco*RI-*Hin*dIII fragment containing *sgcC3*, *sgcA1*, and *cagA* and their promoter regions, from pBS1008 into the same sites of pKC1139. To make pBS1030 for *sgcC3* and *sgcA1* expression, pBS1008 was first digested with *Xho*I and *Xba*I to remove the *cagA* gene. The resultant 6.5 kb *Xho*I-*Xba*I fragment was purified, blunt-ended by treatment with mung bean nuclease, and self-ligated to yield pBS1029. A 3.2 kb *Eco*RI-*Hin*dIII fragment containing *sgcC3*, *sgcA1,* and their promoter regions was then cloned from pBS1029 into the same sites of pKC1139 to yield pBS1030. To express *sgcA1* alone, the *sgcA1* gene and its promoter region was amplified by PCR from pBS1004 with the following pair of primers: 5′-AGA*GAATTC*CAGAGGTTGCTCCTG-3′ (the engineered *Eco*R1 site is italic) and 5′-GCGCTTGCGCA*G-GATCC*GGGG-3′ (the native *Bam*HI site is italic). The resultant product was sequenced to confirm PCR fidelity and cloned as a 2.2 kb *Eco*RI-*Bam*HI fragment into the same sites of pKC1139 to afford pBS1031. To express *sgcC3* alone as a control, the *sgcC3* gene and its promoter region was amplified by PCR with the following pair of primers: 5′-AGA*GAAT-*

*TC*AGCAGGATCTGGTCG-3′ (the engineered *Eco*R1 site is italic) and 5′-CGC*GGATCC*AAGTCGTACACGGGC-3′ (the engineered *Bam*HI site is italic). The resultant product was sequenced to confirm PCR fidelity and cloned as a 1.9 kb *Eco*RI-*Bam*HI fragment into the same sites of pKC1139 to afford pBS1032. The expressions of *sgcA1*, *sgcC3*, and *cagA* in pBS1028, pBS1030, pBS1031, and pBS1032 are all under the control of their respective native promoters.

pBS1028, pBS1030, pBS1031, and pBS1032, as well as pKC1139 as a control, were introduced into *S. globisporus* by conjugation as described previously.[11,12] Selection for apramycin resistance resulted in the isolation of the recombinant *S. globisporus* strains harboring various expression vectors. Fermentation of *S. globisporus* wild-type and recombinant strains for C-1027 production and C-1027 isolation and analysis were carried out as previously reported.[11,12] Briefly, *S. globisporus* fermentations (typically cultured at 28 °C in a shaker incubator at 250 rpm for 96−120 h) were filtered through cheesecloth to remove the mycelia, and the broth was adjusted to pH 4.0 with 0.1 N HCl and centrifuged to remove any precipitate. To the resultant supernatant was added $(NH_4)_2SO_4$ to 50% saturation, and the precipitated C-1027 chromoprotein was collected by centrifugation and dissolved in 0.1 M potassium phosphate, pH 8.0. The latter was extracted with ethyl acetate (EtOAc), and the combined EtOAc extracts were concentrated in a vacuum, redissolved in methanol, and stored at −80 °C. The isolated yield from the wild-type strain for the chromoprotein complex is between 50 and 100 mg/L, and that for the chromophore varies between 1 and 5.0 mg/L. For yield comparison, C-1027 chromophore was isolated from 50 mL fermentations of the *S. globisporus* wild-type and recombinant strains and was then dissolved in methanol (120 μL) and analyzed (30 μL per injection) by HPLC for quantitation. HPLC was carried out on a Prodigy ODS-2 column (5 μm, 4.6 × 150 mm, Phenomenex, Torrance, CA), eluted isocratically with 20 mM potassion phosphate (pH 6.86)/CH_3CN (50:50, v/v) at a flow rate of 1.0 mL/min with UV detection at 350 nm. The relative yield for C-1027 production was estimated to be at least 200% (pBS1032), 200% (pBS1031), 300% (pBS1030), and 400% (pBS1028) in comparison with the wild-type control harboring pKC1139 (100%).

References and Notes

(1) Danishefsky, S. J.; Shair, M. D. *J. Org. Chem.* **1996**, *61*, 16−44.
(2) Smith, A. L.; Nicolaou, K. C. *J. Med. Chem.* **1996**, *39*, 2103−2117.
(3) Xi, Z.; Goldberg, I. H. In *Comprehensive Natural Products Chemistry*; Kool, E. T., Ed.; Elsevier: New York, 1999; Vol. 7, pp 553−592.
(4) Thorson, J. S.; Shen, B.; Whitwam, R. E.; Liu, W.; Li, Y. *Bioorg. Chem.* **1999**, *27*, 172−188.
(5) Thorson, J. S.; Sievers, E. L.; Ahlert, J.; Shepard, E.; Whitwam, R. E.; Onwueme, K. C.; Ruppen, M. *Curr. Pharm. Des.* **2000**, *6*, 1841−1879.
(6) Jones, G. B.; Fouad, F. S. *Curr. Pharm. Des.* **2002**, *8*, 2415−2440.
(7) Shen, B.; Liu, W.; Nonaka, K. *Curr. Med. Chem.* **2003**, *10*, 2317−2325.
(8) Otani, T.; Miami, Y.; Marunaka, T.; Zhang, R.; Xie, M.-Y. *J. Antibiot.* **1988**, *41*, 1580−1585.
(9) Zhen, Y. S.; Ming, X. Y.; Yu, B.; Otani, T.; Saito, H.; Yamada, Y. *J. Antibiot.* **1989**, *42*, 1294−1298.
(10) Brukner, I. *Curr. Opin. Oncol. Endocr. Met. Invest. Drug.* **2000**, *2*, 344−352.
(11) Liu, W.; Shen, B. *Antimicrob. Agents Chemother.* **2000**, *44*, 382−392.
(12) Liu, W.; Christenson, S. D.; Standage, S.; Shen, B. *Science* **2002**, *297*, 1170−1173.
(13) Zazopoulos, E.; Huang, K.; Staffa, A.; Liu, W.; Bachmann, B. O.; Nonaka, K.; Ahlert, J.; Thorson, J. S.; Shen. B.; Farnet, C. M. *Nat. Biotechnol.* **2003**, *21*, 187−190.
(14) Liu, W.; Ahlert, J.; Gao, Q.; Wendt-Pienkowski, E.; Shen, B.; Thorson, J. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11959−11963.
(15) Lindquist, L.; Kaiser, R.; Reeves, P. R.; Lindberg, A. A. *Eur. J. Biochem.* **1993**, *211*, 763−770.
(16) Graninger, M.; Nidetzky, B.; Heinrichs, D. E.; Whitfield, C. *J. Biol. Chem.* **1999**, *274*, 25069−25077.
(17) Yoo, J.-C.; Lee, E.-H.; Han, J.-M.; Bang, H.-J.; Sohng J.-K. *J. Biochem. Mol. Biol.* **1999**, *32*, 363−369.
(18) Barton, W. A.; Lesniak, J.; Biggins, J. B.; Jeffrey, P. D.; Jiang, J. Q.; Rajashankar, K. R.; Thorson, J. S.; Nikolov, D. B. *Nat. Struct. Biol.* **2001**, *8*, 545−551.
(19) Tsuboi, K. K.; Fukunaga, K.; Petricciani, J. C. *J. Biol. Chem.* **1969**, *244*, 1008−1015.
(20) Gustafson, G. L.; Gander, J. E. *J. Biol. Chem.* **1972**, *247*, 1387−1397.
(21) Nakano, K.; Omura, Y.; Tagaya, M.; Fukui, T. *J. Biochem.* **1989**, *106*, 528−532.
(22) Blankenfeldt, W.; Asuncion, M.; Lam, J. S.; Naismith, J. H. *EMBO J.* **2000**, *19*, 6652−6663.
(23) Zuccotti, S.; Zanardi, D.; Rosano, C.; Sturla, L.; Tonetti, M.; Bolognesi, M. *J. Mol. Biol.* **2001**, *313*, 831−843.
(24) Jiang, J. Q.; Biggins, J. B.; Thorson, J. S. *J. Am. Chem. Soc.* **2000**, *122*, 6803−6804.
(25) Jiang, J. Q.; Biggins, J. B.; Thorson, J. S. *Angew. Chem., Int. Ed.* **2001**, *40*, 1502−1505.
(26) Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Rao, R. N.; Schoner, B. E. *Gene* **1992**, *116*, 43−49.
(27) Sambrook, J.; Russell, D. W.; Irwin, N. *Molecular Cloning: A Laboratory Manual*, 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2000.
(28) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces Genetics*; The John Innes Foundation: Norwich, UK, 2000.
(29) Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248−254.
(30) Segel, I. H. *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*; Wiley: New York, 1975.