Characterization of the Heterodimeric MegBIIa:MegBIIb Aldo-Keto Reductase Involved in the Biosynthesis of L-Mycarose from *Micromonospora megalomicea*†

Salvador Peirú,‡ Eduardo Rodríguez,*,‡ Chau Q. Tran,§ John R. Carney,§ and Hugo Gramajo‡

Microbiology Division, Instituto de Biología Molecular y Celular de Rosario, Consejo Nacional de Investigaciones Científicas y Técnicas, Facultad de Ciencias, Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, S2002LRK Rosario, Argentina, and Kosan Biosciences, Inc., 3832 Bay Center Place, Hayward, California 94545

Received February 27, 2007; Revised Manuscript Received May 9, 2007

ABSTRACT: Two putative C3-ketoreductases, MegBIIa and MegBIIb (formerly MegBII and MegDVII, respectively), homologues to members of the family 12 of aldo-keto reductase (AKR12) superfamily of enzymes, were identified in the megalomicin gene cluster from *Micromonospora megalomicea*. Proteins from this family are involved in the metabolism of TDP-sugars by actinomycetes. MegBIIa was originally proposed to be involved in the L-mycarose biosynthetic pathway, while MegBIIb in the L-megosamine biosynthetic pathway. In this work we have investigated the role of these proteins in the biosynthesis of dTDP-L-mycarose. *In vivo* analysis of the dTDP-sugar intermediates indicated that neither MegBIIa nor its homologue, MegBIIb, was a fully active enzyme by itself. Surprisingly, C3-ketoreductase activity was observed only in the presence of both MegBIIa and MegBIIb, suggesting the formation of an active complex. Copurification and size exclusion chromatography experiments confirmed that MegBIIa and MegBIIb interact forming a 1:1 heterodimeric complex. Finally, a mycarose operon containing *megBIIa* and *megBIIb* together with the other biosynthetic genes of the L-mycarose pathway was constructed and tested by bioconversion experiments in *Escherichia coli*. High levels of mycarosyl-erythronolide B were produced under the condition tested, confirming the role of these two proteins in this metabolic pathway.

The aldo-keto reductase (AKR[1]) is a superfamily of proteins currently divided into 15 families (AKR1−AKR15) of NAD(P)(H)-dependent oxidoreductase enzymes which metabolize a broad range of substrates (*1, 2*). These enzymes share a characteristic $(\alpha/\beta)_8$ TIM barrel structure and four conserved amino acid residues in the active site (Asp, Tyr, Lys, and His). They bind the cofactor in an extended conformation without a Rossman fold, which also appears to be conserved. The catalytic mechanism involves the hydride transfer from NAD(P)(H) to the substrate carbonyl and protonation of the oxygen by the tyrosine residue acting as a general acid. The conserved tyrosine residue is involved in a hydrogen-bonding network with an underlying conserved pair of lysine and aspartic acid residues, and a conserved histidine residue was shown to direct substrate orientation at the active site (*3*) (*4*). They are mostly monomeric proteins of approximately 35 kDa in mass, with the exception of some members of the AKR2, AKR6, and AKR7 families, which have been shown to form multimers. AKRs have been identified in vertebrates, invertebrates, plants, protozoa, fungi, eubacteria, and archaebacteria, implying that it is an ancient superfamily of enzymes.

Family 12 of AKR currently consists of three members, TylC1 from the tylosin gene cluster of *Streptomyces fradiae*, EryBII from the erythromycin (*ery*) gene cluster of *Saccharopolyspora erythraea*, and AveBVIII from the avermectin biosynthetic gene cluster of *Streptomyces avermitilis (1*). Based on genetic and biochemical studies, TylC1 and EryBII were assigned for the biosynthesis of L-mycarose, while AveBVIII for the biosynthesis of L-oleandrose (*5*−*8*). The analysis of recently sequenced polyketide biosynthetic gene clusters from actinomycetes has revealed the presence of new reductases homologous to this AKR12 family. Thus, two putative reductases were identified in the megalomicin (*meg*) gene cluster of *Micromonospora megalomicea*, named MegBIIa and MegBIIb, showing high amino acid sequence homology to TylC1 and EryBII (Figure 1) (*9, 10*).

Megalomicins consist of a 14-membered macrolactone ring 6-deoxyerythronolide B (6-dEB) carrying three deoxysugar residues, L-mycarose, D-desosamine, and L-megosamine, with acetyl or propionyl groups at the 3‴ or 4‴ hydroxyls of the mycarose moiety (*11*). Mycarose is a 2,6-dideoxyhexose with a C-3 methyl branch, which is also found in erythromycin, tylosin, and a few other secondary metabolites (Figure 2). As with other 2,6-dideoxyhexoses, it is generated from

† This work was supported by ANPCyT Grants 01-06622 and 03397-1 and PIP 1005/98 CONICET to H.G.

* Corresponding author. E-mail: erodriguez@ibr.gov.ar. Phone: 54-341-4351235. Fax: 54-341-4390465.

‡ Universidad Nacional de Rosario.

§ Kosan Biosciences, Inc.

[1] Abbreviations: AKR, aldo-keto reductases; ery, erythromycin; meg, megalomicin; 6-dEB, 6-deoxyerythronolide B; MEB, mycarosyl-erythronolide B; dig-EB, digitoxosyl-erythronolide B; 6-dMEB, 6-deoxy-mycarosyl-erythronolide B; EB, erythronolide B; PKS, polyketide synthase.

10.1021/bi700396n CCC: $37.00
Published on Web 06/16/2007

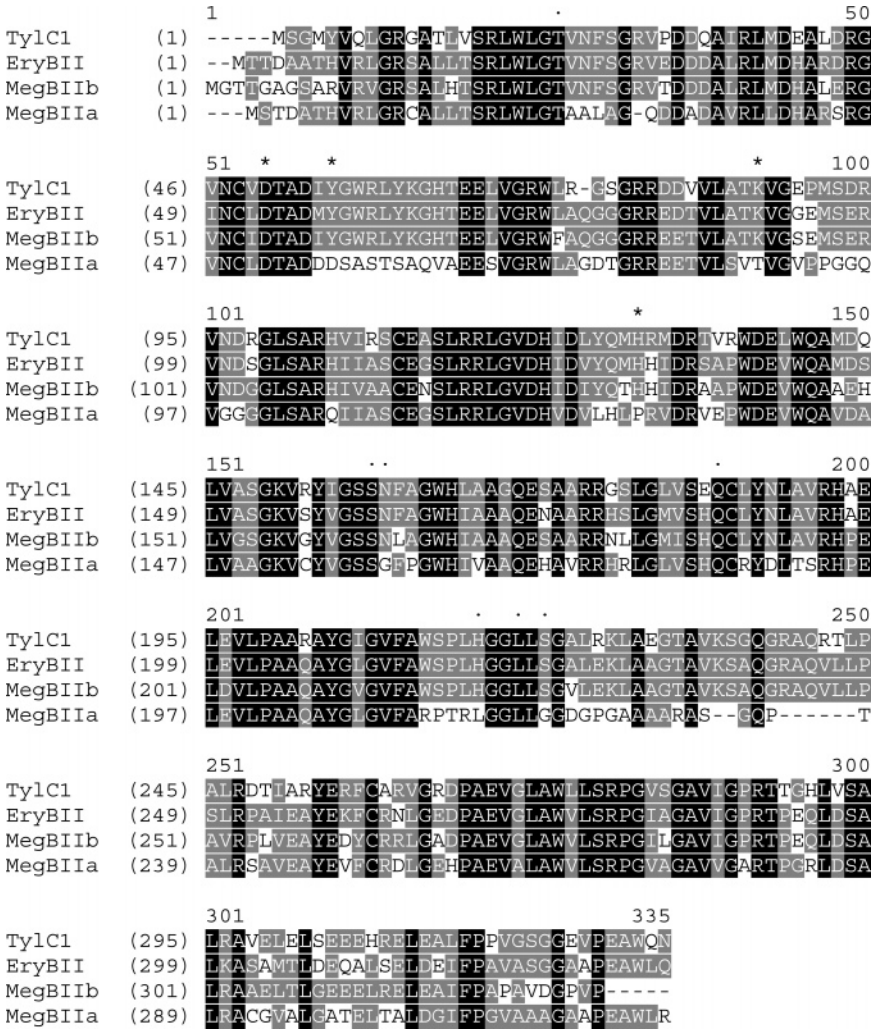

FIGURE 1: Sequence alignments of the AKR12 family involved in the biosynthesis of deoxysugars in actinomycetes. Asterisks indicate conserved residues of the catalytic tetrad Asp-55, Tyr-60, Lys-92, and His-133 based on MegBIIb sequence. Putative cofactor binding residues are indicated by dots.

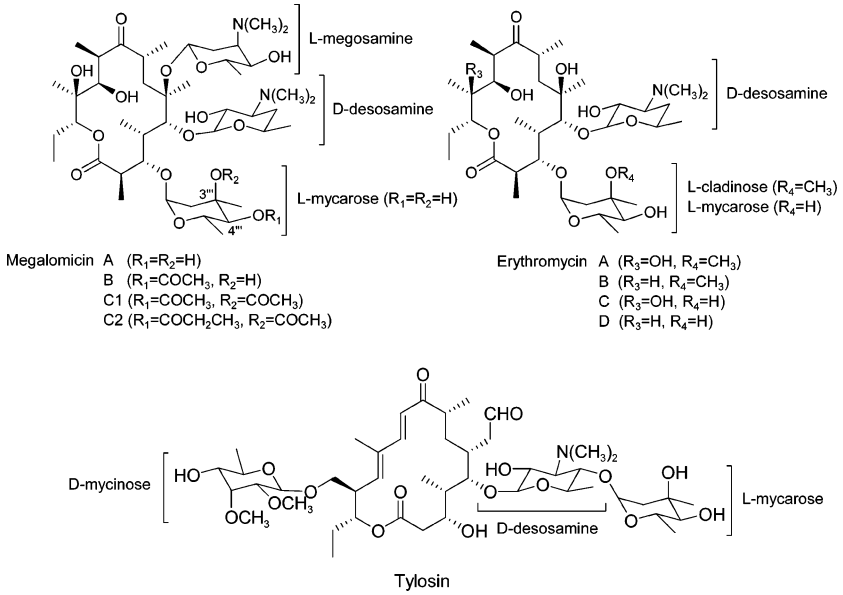

FIGURE 2: Structures of the polyketide macrolides megalomicins, erythromycins, and tylosin.

glucose-1-phosphate via a dTDP-4-keto-6-deoxy-D-allose intermediate. The formation of this intermediate requires the activation of D-glucose-1-phosphate into dTDP-D-glucose followed by removal of the oxygen function at C-6. The next modification is the loss of the hydroxyl group from C-2 by a 2,3-dehydration/3-reduction sequential mechanism which generates a 2,6-dideoxy-D-allose intermediate (*5*, *12*). The last steps involve a C-3 methylation, a C-5 epimerization,

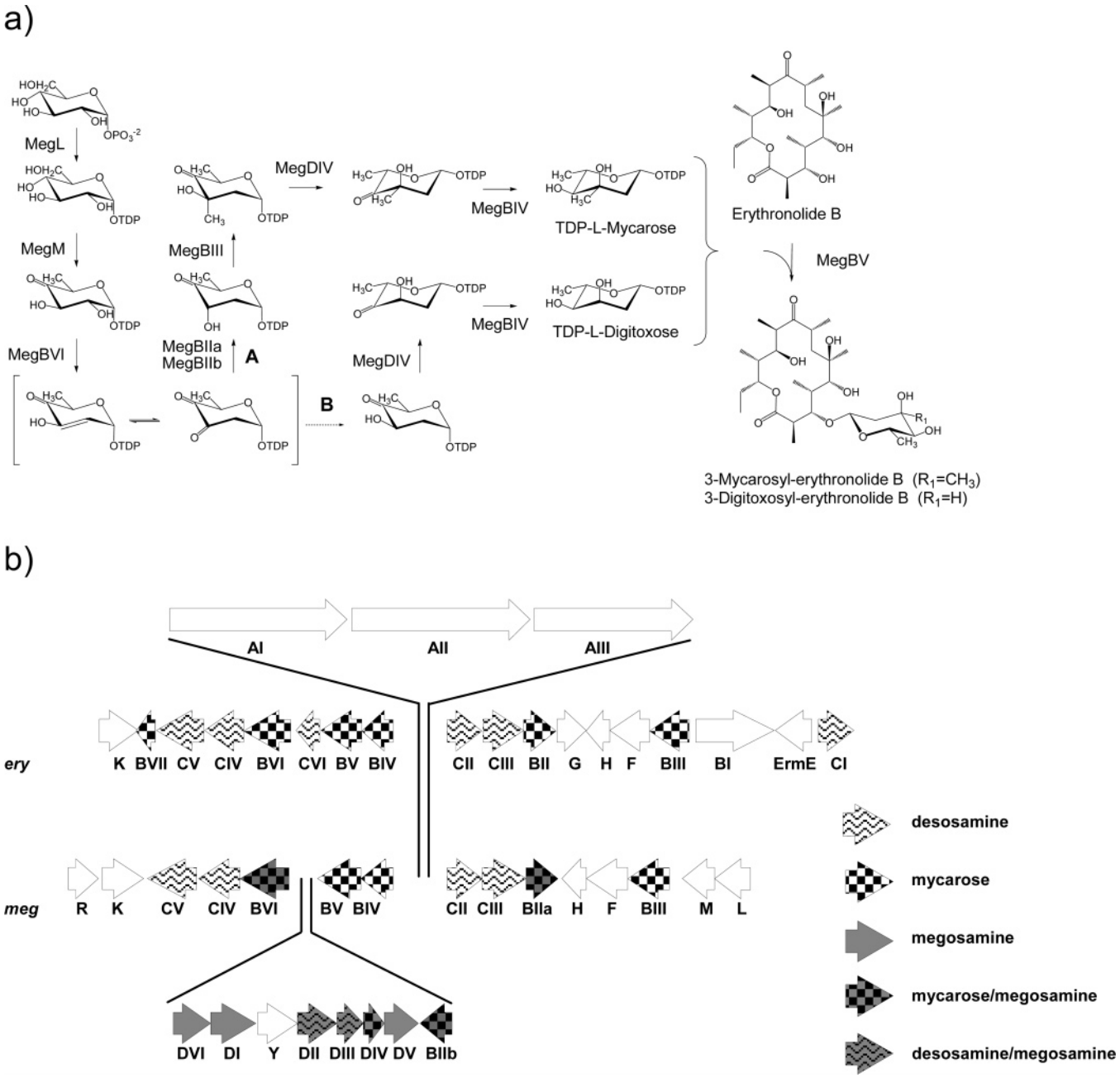

FIGURE 3: (a) Proposed pathways for the biosynthesis of MEB (route A) and dig-EB (route B) in *E. coli*. (b) Genetic organization of the *meg* and *ery* gene clusters.

and a C-4 ketoreduction to give dTDP-L-mycarose (*13*, *14*). Finally, the L-mycarose is transferred to the corresponding aglycone by a specific glycosyltransferase. Figure 3a, route A, depicts the proteins from the *meg* cluster assumed to be involved in the formation of L-mycarose. The role of these proteins was initially assigned on the basis of sequence similarities to other sugar biosynthetic genes, particularly those derived from the *ery* cluster (*9*). As mentioned above, the *meg* cluster contains two genes, *megBIIa* and *megBIIb*, encoding putative C3-ketoreductases highly homologous to EryBII. Based on the conserved genetic organization of both clusters, MegBIIa was originally assigned to the mycarose pathway, while MegBIIb was proposed to be involved in the megosamine pathway (Figure 3b) (*9*).

Heterologous expression of polyketide synthases (PKS) and tailoring enzymes involved in the modification of polyketide aglycones has become a routine technique for the production of glycosylated polyketides in bacteria (*15*−*18*). Recently, we were able to produce the antibiotic erythromycin C in *Escherichia coli* by expressing biosynthetic genes from the *ery* and *meg* gene clusters of *Sac. erythraea* and *M. megalomicea*, respectively (*10*). As an intermediate step, we obtained mycarosyl-erythronolide B (MEB) by feeding 6-dEB to an *E. coli* strain expressing eight proteins of the L-mycarose pathway (MegL, MegM, MegBIII, MegBIV, MegDIV, MegBV, MegBIIa, MegBVI) and the 6-dEB C-6 hydroxylase MegF. However, an unexpected product was also detected in those experiments. Mass spectrometry analysis suggested that this product was a MEB derivative lacking a methyl group in the neutral sugar (*10*). NMR studies allowed characterizing this compound as digitoxosyl-erythronolide B (dig-EB). Early genetic studies showed that *eryBII* mutant strains of *Sac. erythraea* accumulated erythronolide B (EB) as a major compound, as expected for *eryB*

Table 1: Relevant Plasmids and Strains Used in This Study

strain	relevant genotype	source or reference
DH5α	*lacZΔM15, recA1*	Promega
K207-3	F-*ompT hsdS$_B$(r$^-$m$^-$), gal dcm* (DE3), *panD::panDS25A,* Δ*prpRBCD::*T7prom-*sfp,* T7prom-*prpE, ygfG::*T7prom-*accA1*-T7prom-*pccB*	(*33*)
K506-121	K207-3 derivative, Δ*rmlC*	(*24*)

plasmid	description[a]	source or reference
pET24b	*E. coli* expression vector, ColE1 ori, *kan*	Novagen
pET28a	*E. coli* expression vector, ColE1 ori, *kan*	Novagen
pKOS431-39.1	*E. coli* expression vector, RSF1030 ori, *kan*	(*10*)
pLB29	*His-megBVI*, ColE1 ori, *kan*	(*10*)
pLB328	*FLAG-megBIIa*, ColE1 ori, *kan*	this work
pKOS506-45b	*His-megBIIb*, ColE1 ori, *kan*	this work
pKOS506-40b	*His-eryBIIb*, ColE1 ori, *kan*	this work
pLB339	*His-megBIIb-eryBII*, ColE1 ori, *kan*	this work
pLB338	*FLAG-megBIIa His-megBIIb*, ColE1 ori, *kan*	this work
pLB341	*His-megBVI, His-megBIIa*, ColE1 ori, *kan*	this work
pLB342	*His-megBVI, His-megBIIb*, ColE1 ori, *kan*	this work
pLB344	*His-megBVI, FLAG-megBIIa* and *His-megBIIb*, ColE1 ori, *kan*	this work
pKOS452-53e	*megDIV, megBIIa, megM, megL, megF, megBVI, megBIV, megBV,* and *megBIII*; RSF1030 ori, *kan*	(*10*)
pKOS506-33B	*megDIV, megM, megL, megF, megBVI, megBIV, megBV,* and *megBIII*; RSF1030 ori, *kan*	this work
pKOS506-48A	*megDIV, megBIIb, megM, megL, megF, megBVI, megBIV, megBV,* and *megBIII*; RSF1030 ori, *kan*	this work
pKOS506-85	*megDIV, megBIIa, megM, megL, megF, megBVI, megBIV, megBV, megBIII* and *megBIIb*; RSF1030 ori, *kan*	this work
pKOS506-112A	*megBIIb-eryBII, megDIV, megM, megL, megF, megBVI, megBIV, megBV,* and *megBIII*; RSF1030 ori, *kan*	this work
pGro7	P$_{BAD}$ *groES -groEL* in a pACYC184 backbone, *cat*	Takara

[a] Abbreviations: *kan*, kanamycin resistance gene; *cat*, chloramphenicol acetyltransferase gene. Genes in each operon are named according to their position in the operon.

mutants. However, the strain also produced bioactive minor species including 3″-C desmethyl-erythromycin A, C, and D, in addition to traces of erythromycin congeners (*7*). According to the authors, the presence of erythromycins and 3″-C desmethyl-erythromycin derivatives could be derived from endogenous reductase activities that would generate the two alternative C3 epimers of the dTDP-2,6-dideoxy-hexose intermediate. The 2,6-dideoxyhexose with an axial 3-OH group was shown to be substrate for the C3-methyltransferases TylC3 and EryBIII (*13*, *19*). Although EryBIII is unable to methylate the intermediate with an equatorial 3-OH, this compound can be subsequently reduced and transferred to the macrolactone ring to give the erythromycin derivatives, as found for the *eryBII* mutants (*7*, *19*).

Based on these results, the unexpected production of dig-EB in our previous experiments in *E. coli* could be due to a failure in the 3-ketoreductase activity of MegBIIa, which would allow endogenous reductases from the host to generate this undesired byproduct (Figure 3a, route B). To address this issue, we investigated the role of both putative reductases MegBIIa and MegBIIb from the *meg* cluster in the biosynthetic pathway of dTDP-L-mycarose.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Growth Conditions. The strains used in this study are presented in Table 1. Luria−Bertrani (LB) medium was used for growth of *E. coli* strains. The following antibiotics were added to the media when necessary: kanamycin (50 μg/mL) and chloramphenicol (20 μg/mL).

DNA Manipulations. DNA restriction enzymes were used as recommended by the manufacturer (New England Biolabs, USA). DNA manipulations were performed using standard protocols (Sambrook 1989). DNA fragments were purified from agarose gels with the Freeze 'N Squeeze DNA columns (Bio-Rad). Plasmids were prepared using a QIAprep Spin Miniprep Kit (Qiagen). Deep Vent DNA polymerase was used in all PCR reactions according to the supplier instructions (New England Biolabs, USA).

Plasmid Constructions. Each gene of the L-mycarose biosynthetic pathway was amplified by PCR from *M. megalomicea* genomic DNA, individually cloned and sequenced to confirm that they were free of errors, as previously described (*10*). Briefly, the 5′ primers used were designed to have an *Nde*I site overlapping the translational initiation codon, changing GTG start codons to ATG when necessary. The 3′ primers contained *Eco*RI and adjacent *Spe*I sites downstream from the stop codon. Sequences of primers were previously described, except those for cloning *megBIIb* and *eryBII* (*10*). For cloning *megBIIb*, the oligonucleotides 5′-TTTCTAGAGGAGCATATGGGACCACCGGGGCCG (upper) and 5′-GGAATTCACTAGTCTACGGCACCGGCCCGTC (lower) were used. For cloning *eryBII*, the oligonucleotides 5′-TTTCTAGAGGAGCATATGACCACCGACGCCGCG (upper) and 5′-GGAATTCACTAGTTCACTGCAACCAGGCTTC (lower) were used, using genomic DNA from *Sac. erythraea* as a template. PCRs were performed using a DNA Thermal Cycler 480 (Perkin-Elmer) with the following cycling parameters: 30 cycles of 30 s at 94 °C, 30 s of annealing at 56 °C and 80 s at 72 °C. The PCR products were digested with *Nde*I and *Eco*RI and cloned into identical sites of pET24b or pET28a vector in order to express proteins with their natural N terminus or as His-tag fusions, respectively. MegBIIa was obtained as a FLAG-

tag fusion by amplifying *megBIIa* gene with the following alternative upper primer: 5′ CCATATGGACTACAAGGACGACGATGACAAAAGCACCGACGCCAC 3′ and cloned as an *Nde*I/*Eco*RI fragment into pET24b vector to give pLB328. A hybrid *megBIIb-eryBII* gene was generated by replacing the last 421 bp from the 3′ end of *megBIIb* with the last 435 bp from the 3′end of *eryBII*. For this end, the fragment *Bsr*GI/*Spe*I from pKOS506-40b was cloned into pKOS506-45b digested with the same restriction enzymes, to give pLB339. *Bsr*GI restriction site is located in a conserved central region of both genes.

The general strategy for the construction of the mycarose operons was carried out as previously described (*10*). Briefly, the *Xba*I/*Eco*RI fragment from the pET-derived construct harboring the second gene of the operon was cloned into the *Spe*I/*Eco*RI sites of the pET-derived plasmid already containing the first gene of the final construction. This resulted in an intermediate vector harboring the first and second genes separated by a 43 bp sequence with an appropriately positioned ribosomal binding site. The process was repeated in a recursive fashion until the entire operon containing all genes in tandem was constructed. All the mycarose operons contained in pET vectors were removed as *Xba*I/*Eco*RI fragments and cloned into pKOS431-39.1 to give the final operons listed in Table 1.

Growth Conditions, Protein Expression, and Preparation of Cell Extracts. For the expression of heterologous proteins, *E. coli* BL21 (DE3) strains harboring the appropriate plasmids were grown at 37 °C in shake flasks in LB medium in the presence of the corresponding antibiotics for plasmid maintenance. Overnight cultures were diluted 1:100 in fresh medium and grown to an A_{600} of 0.5 to 0.8 before the addition of IPTG to a final concentration of 0.5 mM. Induction was allowed to proceed for 14 h at 22 °C. The cells were harvested, resuspended in 20 mM Tris buffer pH 7.6, and disrupted by sonication. After centrifugation at 20,000 rpm for 20 min, the supernatants were analyzed by LC/MS/MS. Cell extracts for protein purifications were prepared by resuspending cell pellets in buffer A (50 mM Tris-HCl, pH 8, 150 mM NaCl, 0.5 mM dithiothreitol (DTT), 1 mM EDTA, 10% glycerol), and disrupting by sonication.

Protein Analysis. Cell extracts and purified proteins were analyzed by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) (*20*) using a Bio-Rad mini-gel device. The final acrylamide monomer concentrations were 12% (wt/vol) for the separating gel and 5% for the stacking gel. Coomassie brilliant blue was used to stain protein bands. Protein contents were determined by the method of Bradford (*21*) with bovine serum albumin as a standard. Western blotting was carried out as described by Nikolau et al. (*22*) by transferring proteins onto nitrocellulose membranes (Bio-Rad). His-tagged proteins were identified using a monoclonal anti-His_5 antibody (QIAGEN) as a primary antibody. FLAG-tagged proteins were identified using a monoclonal anti-FLAG antibody (Sigma) as a primary antibody. Antigenic polypeptides were visualized using an alkaline phosphatase-conjugated secondary antibody.

Protein Purification Protocols. Cell extracts obtained as described above were clarified by centrifugation at 20,000 rpm for 30 min. The supernatant was applied to a Ni^{2+}-NTA-agarose affinity column (QIAGEN) and equilibrated with the same buffer supplemented with 10 mM of imidazole. The column was subsequently washed and the His_6-tagged proteins eluted from the column using binding buffer containing increasing concentrations of imidazole (20, 40, 60, 80, and 100 mM). Fractions of the eluate were collected and proteins were analyzed by SDS−PAGE. The fractions containing purified proteins were dialyzed at 4 °C overnight against 50 mM Tris buffer, pH 7.6, 0.5 mM DTT, 1 mM EDTA, and 20% glycerol.

Interaction Assay. The physical interaction of His_6-MegBIIb with FLAG-MegBIIa was assayed by coexpressing the genes from plasmid pLB338 in *E. coli* BL21 (DE3). A culture of this strain was grown at 37 °C, and induction and protein purification were carried out as described above. The fractions from this column were evaluated by SDS−PAGE and Western blot as described above. The same assay was carried out with a cell extract obtained from a culture of *E. coli* BL21 (DE3) expressing only FLAG-MegBIIa, from plasmid pLB328, to check for the absence of nonspecific interactions with the resin.

Size-Exclusion Chromatography Analysis. The molecular mass of the native enzyme complexes were determined by size-exclusion chromatography on an AKTA basic high performance liquid chromatograph (Amersham) using a Superdex 75 HR 10/30 column. The column was equilibrated and eluted with 50 mM NaCl in 50 mM Tris buffer, pH 7.5, at a flow rate of 0.5 mL/min, and calibrated with the following protein standards: albumin (66 kDa); FapR (43 kDa); β-lactamase II (25 kDa); and lysozyme (14 kDa). The elution position of blue dextran was used as the Superdex S75 void volume (V_0 = 7.10 mL). Protein elution was detected by absorbance at 280 nm. All fractions were analyzed by SDS−PAGE and Western blotting as described above.

LC/MS/MS Analyses of Sugar Nucleotides. A system consisting of an Applied Biosystems API-3000 triple quadrupole mass spectrometer equipped with a Turbo-ionspray source, Agilent 1100 series HPLC pump, and HTC PAL autosampler was used for LC/MS/MS analyses. The mass spectrometer was operated in negative ion mode with the electrospray needle voltage (−3200), source temperature (400 °C), declustering potential (−46 V), focusing potential (−290 V), and collision energy (−34 eV) set to the values indicated. Nitrogen was used as the collision gas. For precursor scans, resolutions in Q1 and Q3 were set to unit and low, respectively, while for MRM experiments, low resolution was used in both quadrupoles. A Waters YMC ODS-A column (S-5, 2.0 × 250 mm) maintained at 37 °C was used for separations. Mobile phase A was 10 mM triethyl amine in H_2O (pH adjusted to 5.60 with acetic acid), and mobile phase B was MeCN. Typically 10 μL samples were injected and eluted with a linear gradient from 2 to 13% B over 10 min, then held at 13% for 2.5 min. The flow rate was 0.200 mL/min. The column eluate was introduced unsplit into the mass spectrometer source.

Bioconversion Experiments. E. coli K207-3 or K506-121 harboring pGro7 and the different expression plasmids were cultured overnight at 37 °C in LB with appropriate antibiotics, subcultured by 1:50 dilution in the same medium and grown to mid-log phase (0.4−0.6 OD_{600} nm/mL). Chaperones and sugar gene expression were induced by addition of 2

mg/mL L-arabinose and 0.5 mM IPTG, respectively, and cultures were supplemented with 50 μg/mL of 6-dEB. Cultures were grown at 22 °C for 2 days, centrifuged at 14000 rpm for 5 min, and culture broths were analyzed by LC/MS.

LC/MS for Bioconversion Studies. The LC/MS system consisted of an Applied Biosystems Mariner time-of-flight mass spectrometer operated in positive ion mode and configured with a Turbo-ionspray source, an Agilent 1100 HPLC pump, and a Gilson 215 sample handler. A Develosil ODS-UG-5 column (2.0 × 150 mm) at 60 °C was used for the chromatography. Mobile phase A was 5 mM NH_4OAc in water, and mobile phase B was 5 mM NH_4OAc in 4:1 (v:v) MeCN/MeOH. Samples (20 μL) were injected and eluted with a gradient from 35 to 100% B over 10 min at 0.250 mL/min. The eluate was delivered unsplit into the mass spectrometer source.

RESULTS

In Vivo Characterization of MegBIIa and MegBIIb. Alignment of the amino acid sequences of MegBIIa and MegBIIb with other members of the AKR12 family revealed that MegBIIa has several substitutions in conserved amino acids residues (Figure 1). The most relevant changes involve three amino acid residues (Tyr-60, Lys-92, and His-133 based on MegBIIb sequence) proposed to take part in the catalytic mechanism of these proteins (*23*). Moreover, MegBIIa also exhibited substitutions in amino acid residues implicated in the cofactor binding of the AKRs. On the other hand, MegBIIb showed end-to-end homology to the proteins of this family except in its C-terminal region (Figure 1). This sequence analysis suggests that MegBIIa might not be a fully active ketoreductase, supporting our hypothesis that the dig-EB formed in a recombinant *E. coli* strain expressing the putative L-mycarose pathway occurs because host reductases generate a 2,6-dideoxyhexose with an equatorial 3-OH group (Figure 3a, route B) (*10*).

To test *in vivo* the C3-ketoreductase activity of MegBIIa and/or MegBIIb, we expressed these proteins and MegBVI in *E. coli* and analyzed the dTDP-sugar intermediates accumulated by an LC/MS/MS method recently described that allows detection of dTDP-sugars in the low ng mL^{-1} range (*24*). *E. coli* strain K506-121 was used for these experiments (*24*). This strain carries a deletion of the *rmlC* gene that results in the accumulation of high levels of dTDP-4-keto-6-deoxy-D-allose with parent/daughter pairs of *m*/*z* 545/321 and *m*/*z* 563/321, with the latter corresponding to its hydrated form (*25*) (Figure 4a). Three plasmids harboring *megBVI* (pLB29), or the combinations *megBVI-megBIIa* (pLB341) and *megBVI-megBIIb* (pLB342), were constructed and used to express the corresponding proteins in *E. coli*. Analysis of dTDP-sugars in cell-free extracts revealed that expression of *megBVI*, encoding a putative 2,3-dehydratase, resulted in the consumption of dTDP-4-keto-6-deoxy-D-allose. However, the expected product dTDP-3,4-diketo-2,6-dideoxy-D-allose was not detected (Figure 4b). Instability of this compound, which results in its degradation to maltol and dTDP, was previously observed for the 2,3-dehydratases Gra Orf27 of the granaticin gene cluster of *Streptomyces violaceoruber* and TylX3 of the tylosin gene cluster of *S. fradiae* (*5*, *12*). Expression of *megBVI* together with *megBIIa* or *megBIIb* was also unsuccessful at producing the expected

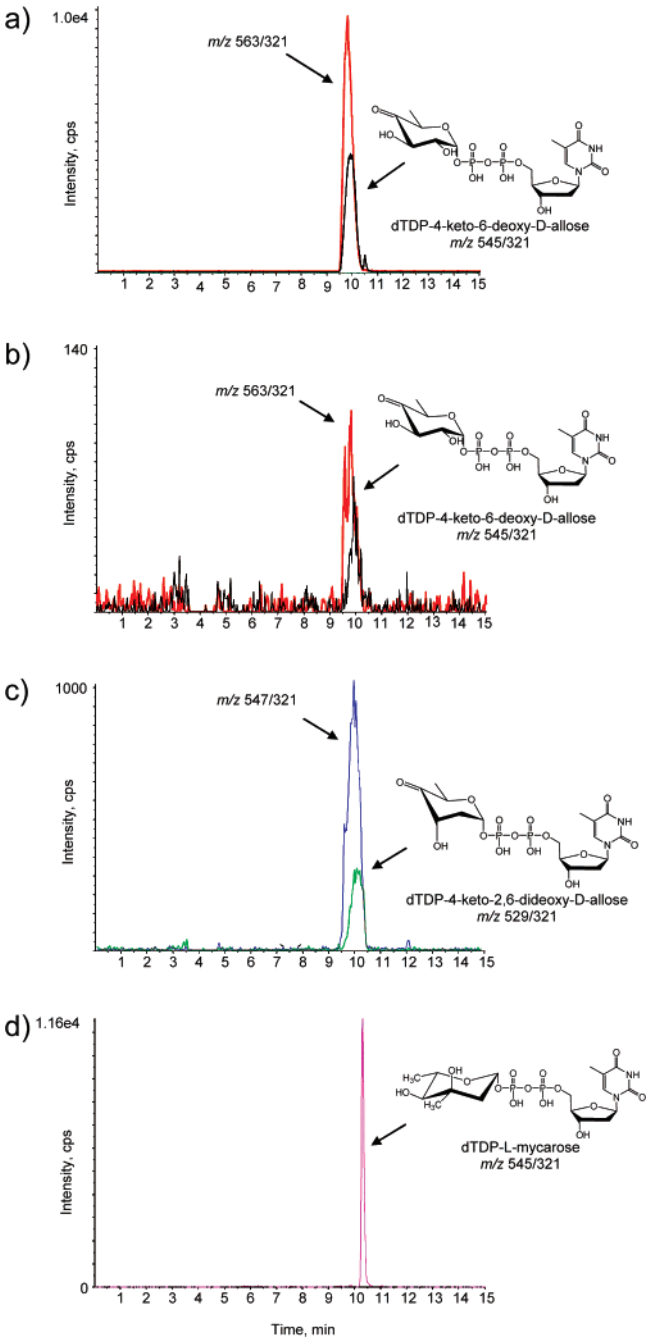

FIGURE 4: Multiple reaction monitoring analysis of dTDP-sugars in cell-free extracts of *E. coli* expressing proteins from the L-mycarose pathway: (a) control with an empty vector; (b) MegBVI; (c) MegBVI, MegBIIa, and MegBIIb; (d) MegBVI, MegBIIa, MegBIIb, MegBIII, MegDIV, and MegBIV. Each *m*/*z* parent/daughter pair is shown with specific colors.

compound dTDP-4-keto-2,6-dideoxy-D-allose. Instead, the results obtained were similar to the expression of *megBVI* alone (data not shown). These results suggest that neither MegBIIa nor MegBIIb is a functional C3-ketoreductase under the conditions assayed, as we previously proposed for MegBIIa based on the amino acid sequence alignments (Figure 1).

The lack of activity of MegBIIa and MegBIIb prompted us to test the coexpression of MegBIIa and MegBIIb, in an attempt to emulate the conditions found in the natural host. For this, a new plasmid harboring *megBVI* together with *megBIIa* and *megBIIb* was constructed (pLB344). Surprisingly, the LC/MS/MS analysis of cell-free extracts of the

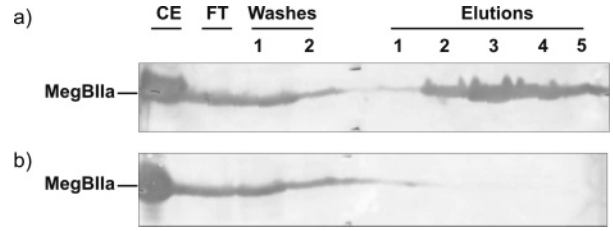

FIGURE 5: Evidence of MegBIIa−MegBIIb interaction by copurification from a Ni^{2+} affinity column. Cell extracts expressing His_6-MegBIIb and FLAG-MegBIIa (a) or FLAG-MegBIIa alone (b) were loaded on a Ni^{2+} affinity column, washed with buffer A, and eluted in the same buffer containing 20, 40, 60, 80, and 100 mM imidazole (elution fractions 1 to 5). Samples were run on a 12% SDS−PAGE gel, transferred to nitrocellulose, and probed with monoclonal anti-FLAG antibodies. CE: cell extract. FT: flowthrough.

strain expressing these three proteins revealed a peak with a parent/daughter pair of *m*/*z* 529/321, consistent with the production of dTDP-4-keto-2,6-dideoxy-D-allose. In addition, a peak with a parent/daughter pair of *m*/*z* 547/321 was also observed, consistent with the hydrated form of dTDP-4-keto-2,6-dideoxy-D-allose (*26*) (Figure 4c). This result suggests that both MegBIIa and MegBIIb are required to catalyze the 3-ketoreductase step that follows dehydration by MegBVI. Moreover, expression of the complete L-mycarose pathway (MegBVI, MegBIIa, MegBIIb, MegBIII, MegDIV, and MegBIV) resulted in the formation of a peak with a parent/daughter pair of *m*/*z* 545/321, consistent with the mass of dTDP-L-mycarose (Figure 4d). This peak was not detected in the absence of either MegBIIa or MegBIIb (data not shown), reinforcing the proposal that MegBIIa and MegBIIb are both needed for the biosynthesis of dTDP-L-mycarose in *M. megalomicea*.

Characterization of the MegBIIa−MegBIIb Dimeric Complex. The fact that MegBIIa and MegBIIb are both required to catalyze the 3-ketoreductase step in the L-mycarose pathway suggests that these proteins should interact with each other in order to form an active heterooligomeric complex. To test this hypothesis we studied the physical interaction between these two polypeptides. To this end, MegBIIa and MegBIIb were tagged with different N-terminal epitopes, coexpressed in *E. coli*, and used to perform copurification experiments. In this way, if a stable protein complex is formed, both proteins may copurify when passed through a column with affinity to one of the tags. For this purpose we used plasmid pLB338, where MegBIIa is expressed as a FLAG-tag fusion protein while MegBIIb is expressed with a His_6-tag. Cell-free extracts obtained from *E. coli* cultures expressing both tagged proteins were run through a Ni^{2+}-NTA-agarose affinity column. After extensive washes, the specifically bound proteins were eluted with buffer containing increasing concentration of imidazole. The outcome of the column was evaluated by SDS−PAGE and Western blotting using anti-tag monoclonal antibodies, confirming the presence of FLAG-MegBIIa (Figure 5a) and His_6-MegBIIb (data not shown). These results validate the hypothesis that these two proteins interact *in vivo* forming an active heterooligomeric complex. Absence of interactions of FLAG-MegBIIa with the Ni^{2+}-NTA column was verified using cell-free extracts of cultures expressing only this protein (Figure 5b).

To demonstrate that megBIIa and MegBIIb coeluted from the affinity column and to determine the oligomeric state of this complex, a size exclusion chromatography analysis was performed using a 60 mM imidazole elution fraction previously obtained from the Ni^{2+}-NTA-agarose affinity column (Figure 5a). The sample was run through a Superdex S75 column, where a single peak with a calculated molecular mass of approximately 65 kDa, consistent with a dimeric structure, was observed (Figure 6a). Samples corresponding to this peak were collected and analyzed by Western blotting with anti-tag antibodies, which confirmed the presence of both FLAG-MegBIIa and His_6-MegBIIb proteins and was consistent with a 1:1 stoichiometry for the MegBIIa−MegBIIb heterodimeric complex (Figure 6b). The individual expression and purification of each tagged protein resulted mainly in the formation of high molecular weight protein aggregates, as observed through size exclusion chromatography using either Superdex S75 or Superdex S200 columns (data not shown). These results indicate that the coexpression of MegBIIa and MegBIIb favors the formation of a properly folded heterodimeric complex over the high molecular weight aggregates.

Improvement of the C3-Ketoreductase Activity of the L-Mycarose Pathway Leads to an Efficient Bioconversion of MEB in E. coli. The nucleotide sugar analysis and interaction studies indicate that the MegBIIa−MegBIIb complex is needed for a proficient biosynthesis of dTDP-L-mycarose in *M. megalomicea*. In order to test the efficiency of the complete L-mycarose pathway through bioconversion experiments, we constructed a series of vectors harboring the genes needed to convert 6-dEB into MEB. The expression vector pKOS506-85 was constructed by cloning the genes *megL*, *megM*, *megBVI*, *megBIIa*, *megBIIb*, *megBIII*, *megDIV*, *megBIV*, and *megF* as an operon. In addition, two other plasmids were made containing the same set of genes but without *megBIIa* (pKOS506-48A), or without both m*egBIIa* and *megBIIb* genes (pKOS506-33B). The expression of these sets of genes was tested by feeding 6-dEB to cultures of *E. coli* strains, expressing the different operons, and analyzing the product profiles after the bioconversion by LC−MS. Analysis of the cultures from strains expressing pKOS506-48A or pKOS506-33B showed very similar profiles, as it was previously observed with an operon without *megBIIb* (pKOS452-53e) (Figure 7a,b) (*10*). In all cases, both MEB and dig-EB were detected. These results demonstrate that neither MegBIIa nor MegBIIb contributes to the MEB production in those operons, and confirm that endogenous reductase activities should generate the two alternative C3 epimers of the dTDP-2,6-dideoxyhexose intermediate that results in the formation of MEB and dig-EB. Although we predict the presence of host reductase activities, the accumulation of either dTDP-4-keto-2,6-dideoxy-D-allose isomers was undetectable in a strain expressing only MegBVI by the dTDP-sugar analysis shown before (Figure 4b). This could be due to a lower sensitivity of the method, or further modifications of the intermediate generated.

Finally, the bioconversion of 6-dEB in cultures of an *E. coli* strain expressing the entire pathway from plasmid pKOS506-85, including both MegBIIa and MegBIIb, showed a high production of MEB concomitant with a reduction in the EB accumulation, while no dig-EB was detected (Figure 7c). It is worth mentioning that we also detected a compound with a fragmentation profile corresponding to 6-deoxy-mycarosyl-erythronolide B (6-dMEB), demonstrating an

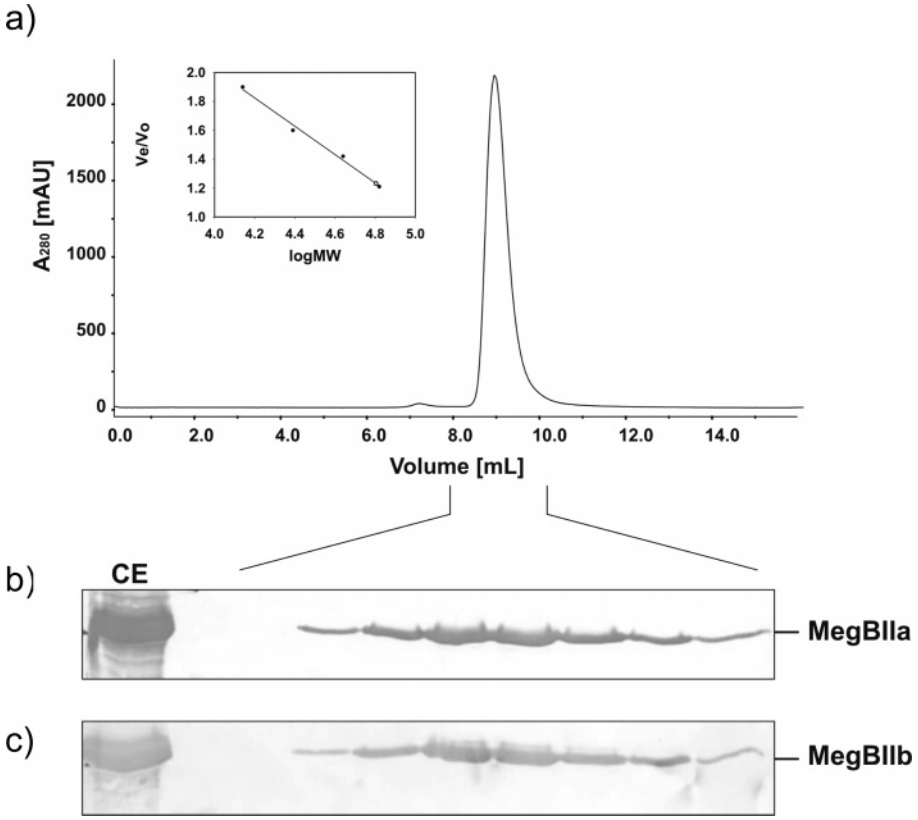

FIGURE 6: (a) Size-exclusion chromatography analysis of MegBIIa−MegBIIb complex. Fraction 3 collected from the Ni^{2+} affinity column (Figure 5a) was run on a Superdex 75 HR 10/30 column. The inset shows molecular mass determination for MegBIIa−MegBIIb complex (open circle) using a calibration obtained with molecular mass standards (full circles). V_e is the elution volume of the sample or the standards. The void volume of the column ($V_0 = 7.10$) was estimated with blue dextran. The calculated molecular weight for the MegBIIa−MegBIIb complex is 65 kDa. Elution fractions were collected from 8 to 10 mL which correspond to the dimeric complex. These samples were run on a 12% SDS−PAGE gel, transferred to nitrocellulose, and probed with monoclonal anti-FLAG (b) or anti-His_5 antibodies (c). CE: cell extract.

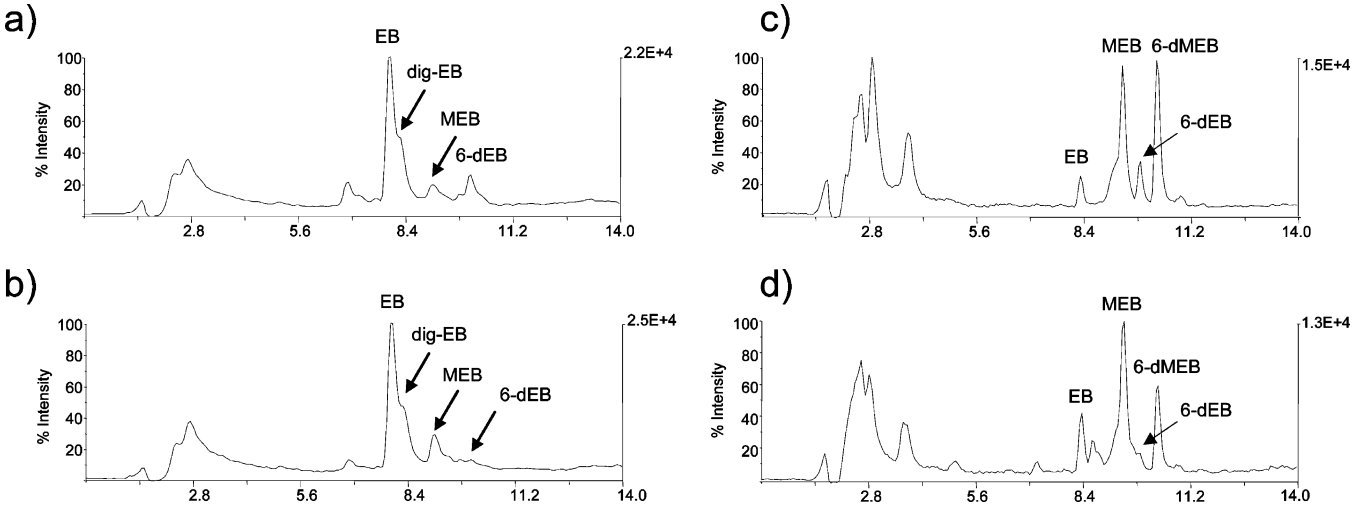

FIGURE 7: LC−MS analysis of the bioconversion of 6dEB by an *E. coli* strain expressing the L-mycarose pathway from different plasmids: (a) pKOS452-53e, (b) pKOS-506-33B, (c) pKOS506-85, (d) pKOS506-112A.

elevated capacity to produce dTDP-L-mycarose by this system (Figure 7c). This result confirms, *in vivo*, that the MegBIIa−MegBIIb heterodimeric complex works efficiently to catalyze the C3-ketoreduction reaction on the L-mycarose pathway, avoiding the interference from host enzymes that results in the generation of byproducts.

Hybrid Protein Studies. Site-directed mutagenesis of mammalian AKR has revealed that the C-terminal region is important in determining the substrate specificity of the enzyme. Protein sequence alignments of MegBIIb with the other proteins of the family revealed that MegBIIb has an incomplete C-terminal domain (Figure 1). To analyze whether the amino acid sequence change in the C-terminal region of MegBIIb protein could be responsible for the lack of activity of this enzyme, we replaced the C-terminal half of MegBIIb by the C-terminal half of EryBII to generate a hybrid protein MegBIIb-EryBII. A new mycarose operon containing this hybrid gene was constructed (pKOS506-

112B) and tested in a bioconversion experiment as mentioned before. Fermentation analysis of an *E. coli* strain expressing this operon showed efficient mycarosylation of 6dEB (Figure 7d), comparable to the coexpression of MegBIIa and MegBIIb from pKOS506-85, with no dig-EB shunt product formation, suggesting that the C-terminal domain of EryBII is important to restore the activity of MegBIIb.

DISCUSSION

Polyketides are one of the most interesting subgroups of bioactive natural products because of their remarkable diversity of chemical structures and biological activities. The enormous structural diversity of these compounds is dictated by the PKS itself and by post-PKS modifications, such as glycosylations. The establishment of deoxysugars as vital components for the efficacy and specificity of biologically active glycoconjugates suggests that altering and/or exchanging these crucial sugar structures may enhance or vary the biological characteristics of their parent molecules. Such an approach has become an appealing strategy for the development of a new generation of antimicrobial/antitumor agents (*17*, *18*). For this reason, extensive efforts have been made to investigate the relevant genetics, enzymology, and mechanistic features of the biosynthetic pathways leading to these sugars. Generally, the assignment of the enzymatic activity for each of the proteins involved in the sugar biosynthetic pathways has been made on the basis of similarities with related enzymes in databases, and through analysis of the compounds accumulated by mutants affected in selected deoxysugar genes (*6*, *27*, *28*). However, the role of each protein in a particular pathway needs to be validated.

A key step in the formation of L-mycarose is the elimination of the oxygen function at C-2 by a 2,3-dehydretase/3-ketoreductase sequential mechanism, as described for TylX3/TylC1 from the tylosin cluster (*5*). Here we report the characterization of two 3-ketoreductases from the *meg* gene cluster of *M. megalomicea*, MegBIIa and MegBIIb, homologues to the AKR12 family. The analysis of the C3-ketoreductase activity of MegBIIa and MegBIIb was carried out through detection of the dTDP-sugar intermediates accumulated (*in vivo*) in *E. coli*. This analysis indicated that neither MegBIIa nor MegBIIb is fully active alone. Surprisingly, C3-ketoreductase activity was observed in the presence of both MegBIIa and MegBIIb, suggesting that both proteins are part of an active complex. Copurification and size exclusion chromatography experiments confirmed that MegBIIa and MegBIIb interact forming a 1:1 heterodimeric complex, the first bacterial dimeric AKR complex characterized so far.

Homodimeric AKR structures have been found in some members of family 2 and 7 of AKR (*29*−*31*). Xylose reductase (AKR2) from *Candida tedius* dimerizes by utilizing residues from α4, loop 4, α5, α6, and the C-terminal loop. Helices α5 and α6 also play a major role in the dimeric interface of rat aflatoxin reductase (AKR7A1). However, this enzyme has a truncated C-terminal loop and absence of interaction in loop 4 (*30*). The structure and sequence comparison between XR and AKR7A1 suggest that these proteins utilize nonhomologous interactions to dimerize in a similar manner (*32*). Nothing is currently known about the nature of the dimeric interface or whether the formation of this interface is a prerequisite for activity. Examination of the structure suggests that the dimeric interface is probably necessary for structural stability of the protein rather than for catalysis. Sequence alignments revealed that MegBIIb has an altered C-terminal domain compared to other proteins of the AKR12 family. The significance of this C-terminal region in the activity and/or dimerization of MegBIIb was tested by generation of hybrid proteins. Replacement of the C-terminal half of MegBIIb with the C-terminal half of EryBII generated a functional MegBIIb-EryBII hybrid protein, as observed by bioconversion experiments. Size exclusion chromatography revealed that this hybrid protein can form a dimeric structure (data not shown), suggesting that the C-terminal domain of MegBIIb lacks amino acid residues important for self-dimerization. The substitution of several conserved amino acid residues which participate in the catalysis as well as those involved in the cofactor binding (*2*) makes MegBIIa catalytically inactive, even after forming a dimeric complex with MegBIIb. Therefore, the fact that MegBIIb can only be fully active after forming a dimeric complex with MegBIIa suggests that MegBIIa should work as a scaffold required to maintain structural stability of MegBIIb in the active dimeric complex. This is supported by the finding that individually expressed and purified MegBIIb is unable to form an active dimer, with only monomeric and high molecular weight aggregate forms being observed. Although the interaction between these proteins has been established, further structural analysis is required to establish the basis of the complex formation and activation.

Although there are great advantages to using a heterologous expression system for combinatorial biosynthesis, it is important to note that endogenous enzymes might interfere with the heterologously expressed pathway. Several NDP-sugar pathways exist in *E. coli* which can interfere with the dTDP-sugar pathway expressed (*24*). In this work we found that expression of five proteins of the L-mycarose pathway without MegBIIa and MegBIIb generates both MEB and dig-EB in bioconversion experiments, indicating that endogenous reductase activities are responsible for generating MEB and dig-EB.

Finally, reconstitution of the L-mycarose pathway in *E. coli* confirmed that both MegBIIa and MegBIIb are required for a fully functional pathway. Expression of both MegBIIa and MegBIIb with the other mycarose biosynthetic enzymes produces significantly increased amounts of MEB, in addition to 6-dMEB, thus confirming the role of the MegBIIa−MegBIIb heterodimeric complex in the L-mycarose pathway from *M. megalomicea*. Previous experiments have shown that introduction of a set of genes comprising *megDI* to *megBIIb* from *M. megalomicea* into *Sac. erythraea* results in the production of megalomicins (*9*). It is clear that these genes include those required for the biosynthesis of dTDP-megosamine and its transfer to the macrolactone. Two alternative pathways for the biosynthesis of dTDP-megosamine have been proposed, with one of them including a reduction step catalyzed by MegBIIb (*9*). Now the role of both MegBIIa and MegBIIb for the megosamine pathway remains to be established.

ACKNOWLEDGMENT

We thank Mark Burlingame, Kosan Biosciences, for isolating the 6-dEB used in these studies.

REFERENCES

1. Jez, J. M., and Penning, T. M. (2001) The aldo-keto reductase (AKR) superfamily: an update, *Chem. Biol. Interact. 130*−*132*, 499−525.
2. Hyndman, D., Bauman, D. R., Heredia, V. V., and Penning, T. M. (2003) The aldo-keto reductase superfamily homepage, *Chem.-Biol. Interact. 143*−*144*, 621−631.
3. Bohren, K. M., Grimshaw, C. E., Lai, C. J., Harrison, D. H., Ringe, D., Petsko, G. A., and Gabbay, K. H. (1994) Tyrosine-48 is the proton donor and histidine-110 directs substrate stereochemical selectivity in the reduction reaction of human aldose reductase: enzyme kinetics and crystal structure of the Y48H mutant enzyme, *Biochemistry 33*, 2021−2032.
4. Barski, O. A., Gabbay, K. H., Grimshaw, C. E., and Bohren, K. M. (1995) Mechanism of human aldehyde reductase: characterization of the active site pocket, *Biochemistry 34*, 11264−11275.
5. Chen, H., Agnihotri, G., Guo, Z., Que, N. L.S., Chen, X. H., and Liu, H.-w. (1999) Biosynthesis of Mycarose: Isolation and Characterization of Enzymes Involved in the C-2 Deoxygenation, *J. Am. Chem. Soc. 121*, 8124−8125.
6. Summers, R.G., Donadio, S., Staver, M.J., Wendt-Pienkowski, E., Hutchinson, C.R., and Katz, L. (1997). Sequencing and mutagenesis of genes from the erythromycin biosynthetic gene cluster of Saccharopolyspora erythraea that are involved in L-mycarose and D-desosamine production, *Microbiology 143* (Part 10), 3251−3262.
7. Salah-Bey, K., Doumith, M., Michel, J. M., Haydock, S., Cortes, J., Leadlay, P. F., and Raynal, M. C. (1998) Targeted gene inactivation for the elucidation of deoxysugar biosynthesis in the erythromycin producer Saccharopolyspora erythraea, *Mol. Gen. Genet. 257*, 542−553.
8. Ikeda, H., Nonomiya, T., Usami, M., Ohta, T., and Omura, S. (1999) Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in Streptomyces avermitilis, *Proc. Natl. Acad. Sci. U.S.A. 96*, 9509−9514.
9. Volchegursky, Y., Hu, Z., Katz, L., and McDaniel, R. (2000) Biosynthesis of the anti-parasitic agent megalomicin: transformation of erythromycin to megalomicin in Saccharopolyspora erythraea, *Mol. Microbiol. 37*, 752−762.
10. Peiru, S., Menzella, H. G., Rodriguez, E., Carney, J., and Gramajo, H. (2005) Production of the potent antibacterial polyketide erythromycin C in Escherichia coli, *Appl. Environ. Microbiol. 71*, 2539−2547.
11. Weinstein, M. J., Wagman, G. H., Marquez, J. A., Testa, R. T., Oden, E., and Waitz, J. A. (1969) Megalomicin, a new macrolide antibiotic complex produced by Micromonospora, *J. Antibiot. (Tokyo) 22*, 253−258.
12. Draeger, G., Park, S.-H., and Floss, H. G. (1999) Mechanism of the 2-Deoxygenation Step in the Biosynthesis of the Deoxyhexose Moieties of the Antibiotics Granaticin and Oleandomycin, *J. Am. Chem. Soc. 121*, 2611−2612.
13. Chen, H., Zhao, Z., Hallis, T. M., Guo, Z., and Liu, H.-w. (2001) Insights into the Branched-Chain Formation of Mycarose: Methylation Catalyzed by an (S)-Adenosylmethionine-Dependent Methyltransferase, *Angew. Chem., Int. Ed. Engl. 40*, 607−610.
14. Takahashi, H., Liu, Y. N., Chen, H., and Liu, H. W. (2005) Biosynthesis of TDP-l-mycarose: the specificity of a single enzyme governs the outcome of the pathway, *J. Am. Chem. Soc. 127*, 9340−9341.
15. Rodriguez, E., and McDaniel, R. (2001) Combinatorial biosynthesis of antimicrobials and other natural products, *Curr. Opin. Microbiol. 4*, 526−534.
16. Wenzel, S. C., and Muller, R. (2005) Recent developments towards the heterologous expression of complex bacterial natural product biosynthetic pathways, *Curr. Opin. Biotechnol. 16*, 594−606.
17. Luzhetskyy, A., and Bechthold, A. (2005) It works: combinatorial biosynthesis for generating novel glycosylated compounds, *Mol. Microbiol. 58*, 3−5.
18. Salas, J. A., and Mendez, C. (2005) Biosynthesis pathways for deoxysugars in antibiotic-producing actinomycetes: isolation, characterization and generation of novel glycosylated derivatives, *J. Mol. Microbiol. Biotechnol. 9*, 77−85.
19. Lombo, F., Gibson, M., Greenwell, L., Brana, A. F., Rohr, J., Salas, J. A., and Mendez, C. (2004) Engineering biosynthetic pathways for deoxysugars: branched-chain sugar pathways and derivatives from the antitumor tetracenomycin, *Chem. Biol. 11*, 1709−1718.
20. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680−685.
21. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248−254.
22. Nikolau, B. J., Wurtele, E. S., and Stumpf, P. K. (1985) Use of streptavidin to detect biotin-containing proteins in plants, *Anal. Biochem. 149*, 448−453.
23. Jez, J. M., Flynn, T. G., and Penning, T. M. (1997) A new nomenclature for the aldo-keto reductase superfamily, *Biochem. Pharmacol. 54*, 639−647.
24. Rodriguez, E., Peiru, S., Carney, J. R., and Gramajo, H. (2006) In vivo characterization of the dTDP-D-desosamine pathway of the megalomicin gene cluster from Micromonospora megalomicea, *Microbiology 152*, 667−673.
25. Stein, A., Kula, M. R., Elling, L., Verseck, S., and Klaffke, W. (1995) Synthesis of dTDP-6-Deoxy-4-ketoglucose and Analogues with Native and Recombinant dTDP-Glucose-4,6-dehydratase, *Angew. Chem., Int. Ed. Engl. 34*, 1748−1749.
26. Amann, S., Drager, G., Rupprath, C., Kirschning, A., and Elling, L. (2001) (Chemo)enzymatic synthesis of dTDP-activated 2,6-dideoxysugars as building blocks of polyketide antibiotics, *Carbohydr. Res. 335*, 23−32.
27. Gaisser, S., Bohm, G. A., Cortes, J., and Leadlay, P. F. (1997) Analysis of seven genes from the eryAI-eryK region of the erythromycin biosynthetic gene cluster in Saccharopolyspora erythraea, *Mol. Gen. Genet. 256*, 239−251.
28. Gaisser, S., Bohm, G. A., Doumith, M., Raynal, M. C., Dhillon, N., Cortes, J., and Leadlay, P. F. (1998) Analysis of eryBI, eryBIII and eryBVII from the erythromycin biosynthetic gene cluster in Saccharopolyspora erythraea, *Mol. Gen. Genet. 258*, 78−88.
29. Kavanagh, K. L., Klimacek, M., Nidetzky, B., and Wilson, D. K. (2002) The structure of apo and holo forms of xylose reductase, a dimeric aldo-keto reductase from Candida tenuis, *Biochemistry 41*, 8785−8795.
30. Kozma, E., Brown, E., Ellis, E. M., and Lapthorn, A. J. (2002) The crystal structure of rat liver AKR7A1. A dimeric member of the aldo-keto reductase superfamily, *J. Biol. Chem. 277*, 16285−16293.
31. Zhu, X., Lapthorn, A. J., and Ellis, E. M. (2006) Crystal structure of mouse succinic semialdehyde reductase AKR7A5: structural basis for substrate specificity, *Biochemistry 45*, 1562−1570.
32. Wilson, D. K., Kavanagh, K. L., Klimacek, M., and Nidetzky, B. (2003) The xylose reductase (AKR2B5) structure: homology and divergence from other aldo-keto reductases and opportunities for protein engineering, *Chem. Biol. Interact. 143-144*, 515−521.
33. Murli, S., Kennedy, J., Dayem, L. C., Carney, J. R., and Kealey, J. T. (2003) Metabolic engineering of Escherichia coli for improved 6-deoxyerythronolide B production, *J. Ind. Microbiol. Biotechnol. 30*, 500−509.

BI700396N