BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Open Access on 08/22/2015

Investigating the Physiological Roles of Low-Efficiency D-Mannonate and D-Gluconate Dehydratases in the Enolase Superfamily: Pathways for the Catabolism of L-Gulonate and L-Idonate

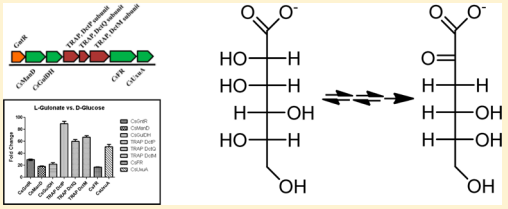
Daniel J. Wichelecki,[†] Jean Alyxa Ferolin Vendiola,[†] Amy M. Jones,[†] Nawar Al-Obaidi,[‡] Steven C. Almo,[‡] and John A. Gerlt*,[†]

[†]Departments of Biochemistry and Chemistry and Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States

[‡]Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, United States

Supporting Information

ABSTRACT: The sequence/function space in the D-mannonate dehydratase subgroup (ManD) of the enolase superfamily was investigated to determine how enzymatic function diverges as sequence identity decreases [Wichelecki, D. J., et al. (2014) *Biochemistry* 53, 2722−2731]. That study revealed that members of the ManD subgroup vary in substrate specificity and catalytic efficiency: high-efficiency (k_{cat}/K_M = 10^3−10^4 M^{-1} s^{-1}) for dehydration of D-mannonate, low-efficiency (k_{cat}/K_M = 10−10^2 M^{-1} s^{-1}) for dehydration of D-mannonate and/or D-gluconate, and no activity. Characterization of high-efficiency members revealed that these are ManDs in the D-glucuronate catabolic pathway {analogues of UxuA [Wichelecki, D. J., et al. (2014) *Biochemistry* 53, 4087−4089]}. However, the genomes of organisms that encode low-efficiency members of the ManDs subgroup encode UxuAs; therefore, these must have divergent physiological functions. In this study, we investigated the physiological functions of three low-efficiency members of the ManD subgroup and identified a novel physiologically relevant pathway for L-gulonate catabolism in *Chromohalobacter salexigens* DSM3043 as well as cryptic pathways for L-gulonate catabolism in *Escherichia coli* CFT073 and L-idonate catabolism in *Salmonella enterica* subsp. *enterica* serovar *Enteritidis* str. P125109. However, we could not identify physiological roles for the low-efficiency members of the ManD subgroup, allowing the suggestion that these pathways may be either evolutionary relics or the starting points for new metabolic potential.



As the protein databases continue to expand (release 2014_07 of UniProt contains 80370243 nonredundant sequences) and, in parallel, the homologous families identified by Pfam increase in size and functional diversity (Pfam release 27.0 describes 14831 families and 515 clans), the problem of assigning function on the basis of sequence identity and similarity is becoming increasingly more challenging. For the past 25 years, we have developed the enolase superfamily (ENS) as a paradigm to characterize and understand the structural bases for divergence of function in functionally diverse superfamilies.[1−4] The members of the ENS share not only a conserved structure, an (α+β) capping domain that determines substrate specificity and a modified TIM-barrel domain [(β/α)$_7$-β-barrel] that contains the acid/base residues that determine the reaction mechanism, but also a conserved mechanism that is initiated by abstraction of a proton from the carbon adjacent to a carboxylate group to generate a Mg^{2+}-stabilized enolate anion intermediate.[5,6] Our recent studies have focused on studying how function diverges as sequence diverges in homologous subgroups.[7]

The D-mannonate dehydratase subgroup (ManD) of the ENS recently was found to include members that dehydrate D-mannonate and/or D-gluconate to 2-keto-3-deoxy-D-gluconate (equivalently, 2-keto-3-deoxy-D-mannonate).[1,7] [The subgroup was named by the function of the first member that was functionally and structurally characterized, when (2007) the protein sequence databases were considerably smaller than they are today.[1]] Also, the members of the ManD subgroup display varying catalytic efficiencies: high-efficiency (specific for D-mannonate with a k_{cat}/K_M of 10^3−10^4 M^{-1} s^{-1}), low-efficiency (D-mannonate and/or D-gluconate with a k_{cat}/K_M of 10−10^2 M^{-1} s^{-1}), and no activity.[7] Although we recently reported the physiological role of the high-efficiency ManDs,[8] dehydration of D-mannonate in the D-glucuronate catabolic pathway in organisms that lack UxuA, the roles of the low-efficiency members of the ManD subgroup have not been determined.

In this article, we describe *in vitro* characterization of two metabolic pathways that are encoded by the genes that are proximal to those that encode low-efficiency members of the ManD subgroup. The pathways utilize two successive dehydrogenase reactions (oxidation followed by reduction) to catalyze epimerization of carbon 5 of a six-carbon acid sugar,

Received: July 7, 2014
Revised: August 11, 2014
Published: August 22, 2014

one for conversion of L-gulonate to D-mannonate (*Escherichia coli* CFT073 and *Chromohalobacter salexigens* DSM3043) and the second for conversion of L-idonate to D-gluconate (*Salmonella enterica* subsp. *enterica* serovar *Enteritidis* str. P125109). *In vivo* studies using genetic deletions (knockouts) and transcriptomics established that the L-gulonate pathway in *C. salexigens* DSM3043 allows growth on L-gulonate as a carbon source. However, we were unable to demonstrate a physiological requirement for the ManD (Uniprot entry Q1QT89; *Cs*ManD[9]) that catalyzes the dehydration of D-mannonate (k_{cat}/K_M = 5 M^{-1} s^{-1}) and D-gluconate (k_{cat}/K_M = 40 M^{-1} s^{-1}).[7] We also describe cryptic pathways for the catabolism of L-gulonate and L-idonate in *E. coli* CFT073 and *S. enterica* subsp. *enterica* serovar *Enteritidis* str. P125109, respectively, that are encoded by the same genome neighborhoods that encode the low-efficiency ManD and GlcD, respectively, that could participate in the pathways.

MATERIALS AND METHODS

Cloning, Expression, and Purification. The genes encoding the L-gulonate 5-dehydrogenase from *Halomonas elongata* DSM 2581 (*He*GulDH, Uniprot entry E1V4Y1, an orthologue of the L-gulonate 5-dehydrogenase from *C. salexigens* DSM3043, *Cs*GulDH), the L-idonate dehydrogenase from *Salmonella enteriditis* (*Se*IdoDH, Uniprot entry B5R538), and the D-gluconate dehydrogenase from *S. enteriditis* (*Se*GlucDH, Uniprot entry B5R540) were synthesized (Genscript) and codon-optimized for expression in *E. coli*; the genes were synthesized with 5′-NdeI and 3′-BamHI restriction sites and were received in vector pUC57. The pUC57 constructs were transformed into *E. coli* XLI Blue for subcloning and storage. Purified vector was digested with NdeI/BamHI (New England Biolabs) and ligated into NdeI/BamHI-digested pET15b (Novagen). The pET15b constructs were transformed into *E. coli* BL21(DE3) cells for expression. The proteins were purified from 1 L cultures using a chelating Sepharose Fast Flow (Amersham Biosciences) column charged with Ni^{2+} as previously described,[7] concentrated (*He*GulDH to 15 mg/mL, *Se*IdoDH to 8.1 mg/mL, and *Se*GlcDH to 4.7 mg/mL), flash-frozen using liquid nitrogen, and stored at −80 °C prior to use.

The gene encoding RspB, the L-idonate 5-dehydrogenase from *E. coli* CFT073 (Uniprot entry Q8FHC8), was amplified from *E. coli* CFT073 genomic DNA (ATCC 700928) using RspB_FOR and RspB_REV primers (Table S1 of the Supporting Information). The mixture for the polymerase chain reaction (PCR) (30 μL) contained 50 ng of template, 1 mM $MgCl_2$, 1× Pfx Amp Buffer, 0.33 mM dNTP, primers (0.33 μM each), and 1.25 units of Pfx polymerase (Invitrogen Platinum Pfx DNA Polymerase kit). Amplifications were performed according to the manufacturer's guidelines. The amplified product was digested with NdeI/BamHI (New England Biolabs) and ligated into NdeI/BamHI-digested pET15b (Novagen). The pET15b *RspB* construct was transformed into *E. coli* BL21(DE3) cells for expression. RspB was purified from a 1 L culture using a chelating Sepharose Fast Flow (Amersham Biosciences) column charged with Ni^{2+} as previously described.[7] The protein was concentrated to 4.2 mg/mL, flash-frozen using liquid nitrogen, and stored at −80 °C prior to use.

The gene encoding RspD, the D-mannonate 5-dehydrogenase from *E. coli* CFT073 (Uniprot entry Q8FHD0), was amplified from *E. coli* CFT073 genomic DNA (ATCC 700928) using RspD_FOR and RspD_REV primers (Table S1 of the Supporting Information). The mixture for the PCR (30 μL) contained 50 ng of template, 1 mM $MgCl_2$, 1× Pfx Amp Buffer, 0.33 mM dNTP, primers (0.33 μM each), and 1.25 units of Pfx polymerase (Invitrogen Platinum Pfx DNA Polymerase kit). Amplifications were performed according to the manufacturer's guidelines. The amplified product was digested with SacI/BamHI (New England Biolabs) and ligated into SacI/BamHI-digested pET17b (Novagen). The pET17b *RspD* construct was transformed into *E. coli* BL21(DE3) cells for expression. RspD was purified from a 1 L culture using DEAE Sepharose, Q-Sepharose, and phenyl Sepharose columns (all from Amersham Biosciences) as previously described.[7] The protein was concentrated to 13 mg/mL, flash-frozen using liquid nitrogen, and stored at −80 °C prior to use.

The gene encoding *Cs*GntR, the GntR transcriptional factor in the genome neighborhood of *Cs*ManD (Uniprot entry Q1QT90), was amplified from *C. salexigens* DSM3043 genomic DNA (ATCC BAA-138) using CsGntR_FOR and CsGntR_REV primers (Table S1 of the Supporting Information). The PCR was the same as that for RspD described above but with the *Cs*GntR primers. The amplified product was digested with NdeI and BamHI (New England Biolabs) and ligated into similarly digested pET15b. The pET15b *Cs*GntR construct was transformed into *E. coli* BL21(DE3) cells for expression. *Cs*GntR was purified from a 1 L culture using a chelating Sepharose Fast Flow (Amersham Biosciences) column charged with Ni^{2+} as previously described.[7] The protein was concentrated to 12 mg/mL, flash-frozen using liquid nitrogen, and stored at −80 °C prior to use.

*Cs*UxuA was obtained from the Albert Einstein College of Medicine as previously described.[7]

Screen for Oxidation Activity of *He*GulDH, *Se*GlcDH, *Se*IdoDH, RspB, and RspD. Reactions to test for oxidation activity were performed in acrylic, UV transparent 96-well plates (Corning Inc.) using a library of 72 acid sugars (Figure S1 of the Supporting Information). Reaction mixtures (60 μL) contained 50 mM HEPES (pH 7.9), 10 mM $MgCl_2$, 1 mM NAD^+, 1 μM enzyme, and 1 mM acid sugar substrate (blanks without enzyme). The plates were incubated at 30 °C for 16 h. The absorbancies were measured at 340 nm (ε = 6220 M^{-1} cm^{-1}) using a Tecan Infinite M200PRO plate reader. Products were verified via ^1H nuclear magnetic resonance.

Screen for the Dehydration Activity of *Cs*UxuA. Reactions to test for dehydration activity were performed in acrylic, UV transparent 96-well plates (Corning Inc.) using a library of 72 acid sugars (Figure S1 of the Supporting Information) as previously described.[7]

Kinetic Assays for *He*GulDH, *Se*IdoDH, RspB, and RspD. Oxidation of L-gulonate, L-idonate, or D-mannonate was monitored using a continuous spectrophotometric assay. For *He*GulDH, the assays (200 μL), at 25 °C, contained 50 mM potassium HEPES (pH 7.9), 5 mM $MnSO_4$, 3 mM NAD^+, and 200 nM *He*GulDH. For the remaining enzymes, the assays (200 μL), at 25 °C, contained 50 mM potassium HEPES (pH 7.9), 5 mM $MgCl_2$, 3 mM NAD^+, and either 200 nM RspB, 200 nM *Se*IdoDH, or 20 nM RspD (lower concentration because of the greater catalytic efficiency). Substrate concentrations varied from 12.5 μM to 30 mM depending on the K_M values of the enzymes. Oxidation was quantitated by measuring the decrease in absorbance at 340 nm (ε = 6220 M^{-1} cm^{-1}).

Kinetic Assays for *Se*GlcDH. Reduction of 5-keto-D-gluconate (fructuronate) was monitored using a continuous spectrophotometric assay. The assay (200 μL), at 25 °C,

dx.doi.org/10.1021/bi500837w | *Biochemistry* 2014, 53, 5692−5699

contained 50 mM potassium HEPES (pH 7.9), 5 mM $MgCl_2$, 200 μM NADH, and 200 nM *Se*GlcDH. The substrate concentration varied from 50 μM to 30 mM. Oxidation was quantitated by measuring the decrease in absorbance at 340 nm (ε = 6220 M^{-1} cm^{-1}).

Kinetic Assay of *Cs*UxuA. Dehydration of D-mannonate was monitiored using a continuous, coupled-enzyme spectrophotometric assay as previously described.[7]

Quantitative RT-PCR of *Cs*ManD, RspABCD, and *Se*GlcD Genome Neighborhoods. *C. salexigens* DSM3043 was grown to an optical density (absorbance at 600 nm) of 0.4−0.5 in M9 minimal salts medium with additional NaCl (1.7 M NaCl, 6.1 mM Na_2HPO_4, 3.9 mM KH_2PO_4, 9.3 mM NH_4Cl, 0.5 mM $MgSO_4$, and 0.5 mM $CaCl_2$) and 10 mM D-mannonate, 10 mM L-gulonate, 10 mM D-glucuronate, or 10 mM D-glucose. Cells were pelleted at 15000 rpm, and the supernatant was removed. mRNA was purified from the cells using an RNeasy Mini Kit (Qiagen). The RNA was further purified using RNase-free DNase (Qiagen) following the manufacturer's protocol. The purity was verified using 30 μL PCR mixtures consisting of 50 ng of mRNA, 1 mM $MgCl_2$, 1× Pfx Amp Buffer, 2× PCR enhancer, 0.33 mM dNTP, 0.33 μM primers [CsRpoD_RTPCR_FOR and CsRpoD_RTPCR_REV (Table S1 of the Supporting Information)], and 1.25 units of Pfx polymerase (Invitrogen Platinum Pfx DNA Polymerase kit). The PCR mixtures were electrophoresed on an agarose gel to check for amplification. cDNA was prepared using Protoscript First Strand (New England Biolabs) and 1 μg of mRNA; the manufacturer's protocol was followed. The qRT-PCR was performed using the Light Cycler 480 SYBR Green I Master Kit (Roche) and a Light Cycler 480 II (Roche) according to the manufacturer's protocol. The qRT-PCR primers are listed in Table S1 of the Supporting Information. C_p values were analyzed using the Light Cycler 480 application and fold changes calculated in Microsoft Excel.

Salmonella enterica serovar *Enteritidis* str. P125109 was grown was grown to an optical density (absorbance at 600 nm) of 0.4−0.5 in M2 minimal salts medium (6.1 mM Na_2HPO_4, 3.9 mM KH_2PO_4, 9.3 mM NH_4Cl, 0.5 mM $MgSO_4$, 0.5 mM $CaCl_2$, and 10 μM $FeSO_4$) and either 10 mM L-gulonate or D-glucose or 1−100 mM L-idonate, D-gluconate, or D-glucose. The following steps are identical to those taken with *C. salexigens*. The primers for *Se*RspABCD and *Se*GlcD qRT-PCR are listed in Table S1 of the Supporting Information.

Knockout Construction in *C. salexigens* DSM3043. The knockouts of genes in *C. salexigens* DSM3043 were constructed using overlap extension and a suicide vector (pK19mobsacB, ATCC 87097).[10] The DNA ~600 bp upstream and downstream of the gene of interest (GOI) was cloned using overlap extension to create a single linear product that perfectly excises the GOI as previously described (primers listed in Table S1 of the Supporting Information).[8] This PCR product was digested with either EcoRI and HindIII or BamHI and HindIII (New England Biolabs) and ligated into similarly digested pK19mobsacB. The ligation products with confirmed sequences were transformed into *E. coli* WM6029 (obtained from W. Metcalf at the University of Illinois at Urbana-Champaign) for conjugation. *E. coli* WM6029 plus pK19mobsacB containing the GOI and *C. salexigens* DSM3043 were grown to an optical density (absorbance at 600 nm) of 0.4−0.5 in 3 mL of PYE medium[11] containing 500 mM NaCl and 50 μg/mL diaminopimelic acid (conjugation medium). The cells were pelleted at 4500 rpm for 5 min; the supernatant was removed, and the cells were resuspended in 3 mL of conjugation medium. Aliquots (250 μL) of both *E. coli* WM6029 and pK19mobsacB containing the GOI and *C. salexigens* DSM3043 were mixed in a 1.5 mL Eppendorf tube and pelleted at 4500 rpm for 5 min. The supernatant was removed, and the cells were resuspended in 100 μL of conjugation medium. The resuspended cells were inoculated onto a nitrocellulose filter, placed on a plate of conjugation medium with 1.2% agar, and incubated overnight at 30 °C. The overnight conjugation was resuspended in 1 mL of PYE medium containing 500 mM NaCl. An aliquot (40 μL) was plated onto PYE medium containing 500 mM NaCl, 4× Kn (200 μg/mL), and 1.2% agar for selection of single-crossover events. Single-crossover events were verified by colony PCR. Colonies with successful single-crossover events were subcultured onto a plate of PYE medium containing 500 mM NaCl and 1.2% agar. Single colonies were plated onto PYE medium containing 500 mM NaCl, 20% sucrose, and 1.2% agar. Colonies were probed for double-crossover events via colony PCR. Successful double-crossover events were verified by isolating the genomic DNA (Qiagen DNeasy Blood and Tissue Kit) and sequencing the knockout region.

Knockout Construction in *Salmonella enterica* serovar *Enteritidis* str. P125109. All knockouts were performed via the method described by Datsenko and Wanner for *E. coli*.[12] Primers are listed in Table S2 of the Supporting Information.

Growth Curves of Wild-Type (WT) and Knockout Strains in *C. salexigens* DSM3043. All strains were grown to an optical density (absorbance at 600 nm) of 0.4−0.5 in 3 mL of PYE medium containing 1.7 M NaCl (50 μg/mL streptomycin for conjugations). The cells were pelleted at 5000 rpm for 5 min and resuspended in 3 mL of M9 minimal salts medium containing 1.7 M NaCl. The cells were inoculated into triplicate 300 μL cultures of M9 minimal salts medium containing 1.7 M NaCl and 10 mM D-mannonate or L-gulonate. Growth curves were recorded using a Bioscreen C instrument (Growth Curves USA). Cells were grown at 37 °C for 120 h while being continuously shaken.

Fluorescence-Based Thermal Shift Assay for *Cs*GntR (ThermoFluor). The library for screening the specificity of *Cs*GntR contained 188 ligands in duplicate wells plus eight control wells (protein without compound) (Figure S2 of the Supporting Information). The assays were performed as previously described.[8] A hit is considered significant if the ΔT_M is ≥4 °C or 3 times the standard deviation of the T_M values for the control wells.

RESULTS AND DISCUSSION

***In Vitro* Pathway for *Cs*ManD.** The genome neighborhood for *Cs*ManD suggests a pathway for D-glucuronate metabolism that involves dehydration of D-mannonate (Figure 1).[13] Specifically, the genome proximal fructuronate reductase that reduces fructuronate to D-mannonate and UxuA[14] that dehydrates D-mannonate to 2-keto-3-deoxy-D-mannonate (not a member of the ENS) are involved in D-glucuronate metabolism in many eubacteria.[8] However, *C. salexigens* DSM3043 uses an alternate pathway for D-glucuronate metabolism initiated by oxidation of D-glucuronate to D-glucaro-1,5-lactone[15−17] (Figure 2). The genome neighborhood encoding the predicted uronate dehydrogenase (Udh, Uniprot entry Q1QUN7) also encodes D-glucarolactone cycloisomerase (Gci, Uniprot entry Q1QUN5), a member of the amidohydrolase superfamily (presumably D-galactarolactone isomerase,

dx.doi.org/10.1021/bi500837w | *Biochemistry* 2014, 53, 5692−5699

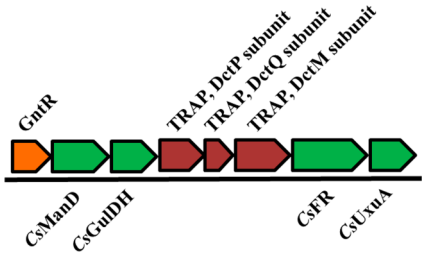


Figure 1. Genome neighborhood for *Cs*ManD in *C. salexigens* DSM3043. Carbohydrate metabolism genes are colored green. TRAP transporters are colored red. The GntR transcriptional regulator is colored orange.

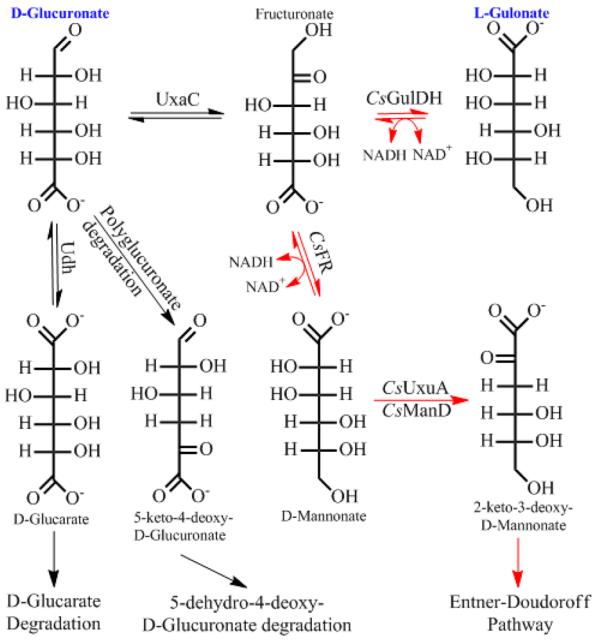


Figure 2. Catabolic pathways for D-glucuronate in eubacteria. The pathway hypothesized for degradation of L-gulonate in *C. salexigens* has red arrows: *Cs*GulDH (alcohol dehydrogenase), *Cs*FR (fructuronate reductase), *Cs*ManD (D-mannonate dehydratase), and UxuA (D-mannonate dehydratase). Starting compounds are labeled in blue (D-glucuronate and L-gulonate).

Uniprot entry Q1QUN6), D-glucarate dehydratase (GlucD, Uniprot entry Q1QUM5), and 5-dehydro-4-deoxy-D-glucarate dehydratase/decarboxylase (Uniprot entry Q1QUM4) (Figure S3 of the Supporting Information). These enzymes are involved in an alternate pathway for D-glucuronate degradation.[18,19]

The physiological relevance of these enzymes and genes was probed using RT-PCR, comparing growth on D-glucuronate to that on D-glucose. The genes encoding all of these enzymes are upregulated; however, the genome neighborhood encoding *Cs*ManD is not upregulated (Figure S4 of the Supporting Information). Further studies are needed to determine whether *C. salexigens* DSM3043 utilizes the newly discovered Gci-based catabolic pathway[18,19] to metabolize D-glucuronate or the previously described pathway in which D-glucarolactone is hydrolyzed to D-glucarate that is dehydrated by GlucD;[15−17] the Gci and "amidohydrolase" in *C. salexigens* DSM3043 are members of the same Pfam families as those characterized by Bouvier and co-workers,[18] PF02746 and PF01979, respectively. Irrespective of the precise pathway, we conclude that *C. salexigens* DSM3043 metabolizes D-glucuronate via conversion to D-glucarolactone/D-glucarate and not D-mannonate.

In addition to fructuronate reductase, a second alcohol dehydrogenase is encoded by the *Cs*ManD genome neighborhood. Its presence suggested that the genome neighborhood allows another acid sugar to be utilized as a carbon source, with the oxidation product an intermediate in the ManD-utilizing D-glucuronate catabolic pathway; e.g., oxidation at carbon 5 of L-gulonate produces fructuronate (Figure 2). Therefore, we hypothesized that the genome neighborhood of *Cs*ManD is responsible for L-gulonate metabolism. A pathway for L-gulonate catabolism in *E. coli* K-12 was suggested by Cooper in 1980, but the gene or enzyme responsible for oxidation of L-gulonate was not identified.[20] Furthermore, Cooper was able to see growth of *E. coli* K-12 only after spontaneous mutations had occurred.[20]

Functional Assignment within the *Cs*ManD Genome Neighborhood: *Cs*UxuA. The functions of the proteins encoded by the genome neighborhood must be determined to assign a physiological role to *Cs*ManD. *Cs*UxuA (Uniprot entry Q1QT83) was screened for dehydration activity using a library of 72 acid sugars (Figure S1 of the Supporting Information). The only hit was D-mannonate that was dehydrated to 2-keto-3-deoxy-D-mannonate with a catalytic efficiency of 1.5×10^3 M^{-1} s^{-1} (Table 1). Therefore, it was unclear whether UxuA, *Cs*ManD, or both were responsible for D-mannonate dehydration in this pathway.

Table 1. Kinetics of Enzymes Encoded by the *Cs*ManD Operon

protein	substrate	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
*Cs*ManD[7]	D-mannonate	0.02 ± 0.0005	5.0×10^0
*Cs*UxuA	D-mannonate	1.6 ± 0.1	1.5×10^3
*He*GulDH (*Cs*GulDH)	L-gulonate	1.8 ± 0.13	6.0×10^2
RspD (*Cs*FR)	D-mannonate	30 ± 1.8	2.8×10^5

Functional Assignment within the *Cs*ManD Genome Neighborhood: *Cs*GulDH. The alcohol dehydrogenase (*Cs*GulDH, Uniprot entry Q1QT88) in the *Cs*ManD genome neighborhood was not amenable to purification, despite several cloning strategies. A presumably orthologous alcohol dehydrogenase from the closely related *H. elongata* DSM 2581 (*He*GulDH, Uniprot entry E1V4Y1) is 76% identical and 86% similar to *Cs*GulDH. Additionally, the genome neighborhood of *He*GulDH is identical to the *Cs*GulDH neighborhood. The purified *He*GulDH was screened for oxidation activity using the library of acid sugars (Figure S1 of the Supporting Information). Multiple hits were observed, with L-gulonate exhibiting the most efficient oxidation [$k_{cat}/K_M = 6 \times 10^2$ M^{-1} s^{-1} (Table 1)]. All three L-gulonate dehydrogenases discussed herein (*Cs*GulDH, *He*GulDH, and RspB) are in Pfam family PF08240.

Functional Assignment within the *Cs*ManD Genome Neighborhood: *Cs*FR. The fructuronate reductase (*Cs*FR, Uniprot entry Q1QT84) in the *Cs*ManD genome neighborhood also was not amenable to purification. *Cs*FR is 43% identical and 59% similar to a D-mannonate oxidoreductase in *E. coli* CFT073 (RspD, Uniprot entry Q8FHD0); both proteins belong to Pfam PF01232. RspD was purified and screened for oxidation activity with the acid sugar library described

dx.doi.org/10.1021/bi500837w | *Biochemistry* 2014, 53, 5692−5699

previously (Figure S1 of the Supporting Information). RspD exhibited specific oxidation of D-mannonate to fructuronate with a catalytic efficiency of 3×10^5 M^{-1} s^{-1} (Tables 1 and 2).

Table 2. Kinetics of Enzymes Encoded by the RspABD Operon

protein	substrate	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
RspA[7]	D-mannonate	0.02 ± 0.001	1.0×10
RspB	L-gulonate	3.9 ± 0.5	1.4×10^3
RspD	5-keto-D-mannonate	30 ± 1.8	2.8×10^5

Although the level of sequence similarity is not high, the similar genome contexts (*vide infra*) and enzymatic activities of *Cs*GulDH and *Cs*FR, and *He*GulDH and RspD, allow confident transfer of function between the two proteins, respectively. Considering their *in vitro* activities, these proteins provide support for the L-gulonate catabolic pathway proposed by Cooper.[20] Therefore, we pursued *in vivo* validation to corroborate the *in vitro* data and confirm the pathway.

Transcript Analysis of the *Cs*ManD Genome Neighborhood. Transcript analysis via qRT-PCR of mRNA revealed that all of genes in the neighborhood of *Cs*ManD (Figure 1) were upregulated ~20−300-fold during growth on both D-mannonate and L-gulonate relative to that on D-glucose (Figure 3). Mutual upregulation in combination with the proximity and identical orientation implies these genes are in an operon, but Northern blots were not performed for confirmation. These findings support the claim that the *Cs*ManD gene cluster is responsible for L-gulonate metabolism via D-mannonate.

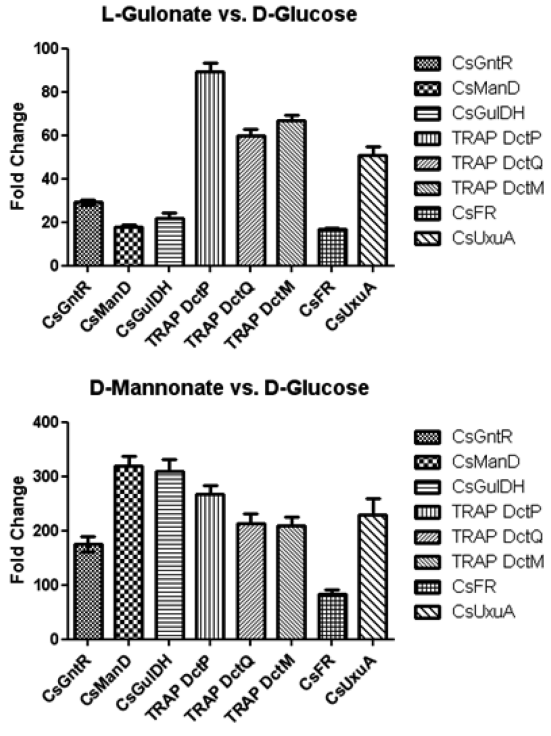


Figure 3. qRT-PCR data for genes in the *Cs*ManD operon for cells grown on L-gulonate (top) or D-mannonate (bottom) vs growth on D-glucose. Cells were grown as described in Materials and Methods to an optical density of 0.4−0.5 at 600 nm (early log phase). Upregulation is observed for all genes on both L-gulonate and D-mannonate.

ThermoFluor for *Cs*GntR. The GntR transcriptional regulator for the L-gulonate utilization operon (*Cs*GntR, Uniprot entry Q1QT90) was expressed, purified, and analyzed by ThermoFluor using a library of 188 ligands (sugars, amino acids, and various metabolites) (Figure S2 of the Supporting Information). Both D-mannonate (ΔT_m of 7 °C) and L-gulonate (ΔT_m of 2 °C) were among the top 10 hits (Table S2 of the Supporting Information). The ligand specificity of the GntR provides further support of the involvement of the operon in growth on L-gulonate and D-mannonate.

Knockout Studies in the *Cs*ManD Genome Neighborhood. Gene deletions or knockouts (KOs) were constructed to further probe the proposed L-gulonate pathway. The following KO strains were constructed: Δ*Cs*GulDH, Δ*Cs*FR, Δ*Cs*, Δ*Cs*ManD, Δ(*Cs*ManD,*Cs*UxuA), and Δ(*Cs*FR,*Cs*UxuA). All of the genes were removed from the genome via homologous recombination. Their growth phenotypes using either D-mannonate or L-gulonate as the sole carbon source were recorded (Figure 4).

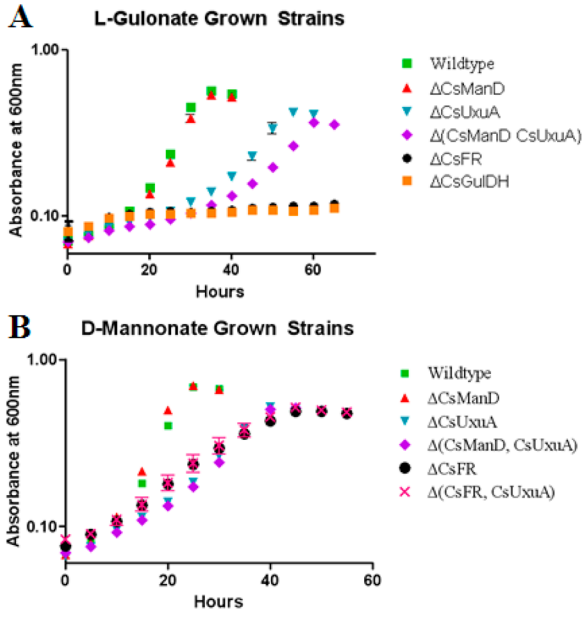


Figure 4. Growth curves for wild-type *C. salexigens* DSM3043 and various knockout strains. Growth was recorded in M9 minimal medium with 1.7 M NaCl and either L-gulonate (A) or D-mannonate (B) as the sole carbon source. The cultures were grown in triplicate.

For both D-mannonate and L-gulonate, the wild-type strain and the Δ*Cs*ManD KO had identical growth curves, indicating *Cs*ManD is not required for the catabolism of either carbon source. In contrast, the Δ*Cs*UxuA KO resulted in slower growth with both carbon sources, suggesting *Cs*UxuA is the predominant ManD for both L-gulonate and D-mannonate catabolism. That the Δ(*Cs*ManD,*Cs*UxuA) double KO was able to grow on D-mannonate and L-gulonate implicates a slower, as yet unidentified pathway for D-mannonate catabolism. Interestingly, the Δ(*Cs*ManD,*Cs*UxuA) double KO exhibited growth retardation (relative to either single KO) when the species was grown on L-gulonate, suggesting that *Cs*ManD plays a role in the alternate D-mannonate metabolic pathway when L-gulonate is the carbon source.

Most importantly, the Δ*Cs*GulDH, Δ*Cs*FR, and Δ(*Cs*FR,*Cs*UxuA) knockout strains did not grow on L-gulonate.

dx.doi.org/10.1021/bi500837w | *Biochemistry* 2014, 53, 5692−5699

This provides convincing evidence that *Cs*GulDH, *Cs*FR, and *Cs*UxuA are responsible for the catabolism of L-gulonate. For D-mannonate, the growth rates for Δ*Cs*FR are very similar to that for Δ*Cs*UxuA, which likely indicates a polar effect on *Cs*UxuA when *Cs*FR is knocked out (the genes encoding *Cs*FR and *Cs*UxuA are separated by 9 bp). Furthermore, the growth of the Δ(*Cs*FR,*Cs*UxuA) double KO on D-mannonate confirms that the *Cs*UxuA deletion strain does not grow by conversion of D-mannonate to fructuronate/L-gulonate and then funneling into an undiscovered catabolic pathway. The unidentified pathway utilizes genes not included in the *Cs*ManD genome neighborhood.

Other *in Vitro* Pathways Implicated by Low-Efficiency ManDs: RspA. An *in vitro* pathway for the catabolism of L-gulonate was derived from the genome neighborhood of a low-efficiency ManD discovered in *E. coli* CFT073. This pathway is encoded by four genes: *RspA*,[7] a low-efficiency ManD (Uniprot entry Q8FHC7); *RspB*, annotated as an alcohol dehydrogenase (Uniprot entry Q8FHC8); *RspC*, annotated as a hypothetical metabolite transporter (Uniprot entry Q8FHC9); and *RspD*, annotated as an oxidoreductase (Uniprot entry Q8FHD0) (Figure 5). The genome neighborhood does not encode a

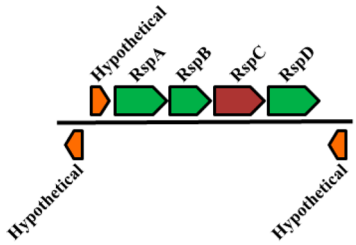


Figure 5. Genome neighborhood of the gene cluster in *E. coli* CFT073. RspA (low-efficiency D-mannonate dehydratase, Uniprot entry Q8FHC7), RspB (L-gulonate dehydrogenase, Uniprot entry Q8FHC8), and RspD (fructuronate reductase, Uniprot entry Q8FHD0) are colored green. The hypothetical metabolite transporter (RspC, Uniprot entry Q8FHC9) is colored red. Hypothetical genes of unknown function are colored orange.

UxuA, although one is encoded elsewhere in the genome. RspA (ManD), RspB (L-gulonate 5-dehydrogenase), and RspD (D-mannonate 5-dehydrogenase) were found to catalyze the *in vitro* conversion of L-gulonate to 2-keto-3-deoxy-D-gluconate (Table 2). Although RspA, RspB, and RspD convert L-gulonate to 2-keto-3-deoxy-D-mannonate/2-keto-3-deoxy-D-gluconate *in vitro*, no upregulation or KO phenotype for growth on L-gulonate was observed in either *E. coli* K-12 or *S. enterica* serovar *Enteritidis* str. P125109 where identical gene clusters are found: *E. coli* K-12 does not grow on L-gulonate until spontaneous mutations occur,[20] and while the *RspABCD* gene cluster in *S. enterica* serovar *Enteritidis* str. P125109 is negligibly upregulated (Figure S5 of the Supporting Information), knockouts of *RspA* and *RspB* grow to turbidity overnight on L-gulonate.

Other *in Vitro* Pathways Implicated by Low-Efficiency ManDs: *Se*GlcD. Bausch and co-workers discovered a novel pathway for L-idonate metabolism in *E. coli* K-12 in which L-idonate is converted to D-gluconate via oxidation and subsequent reduction of the C5 hydroxyl group.[21] D-Gluconate is then phosphorylated and dehydrated to form 2-keto-3-deoxy-D-gluconate 6-phosphate, which is cleaved by an aldolase to pyruvate and glyceraldehyde 3-phosphate. Interestingly, *S. enterica* subsp. *enterica* serovar *Enteritidis* str. P125109 encodes this L-idonate catabolic pathway as well as a similar pathway that is encoded by genes proximal to that encoding a low-efficiency GlcD (Uniprot entry B5R541, *Se*GlcD; k_{cat}/K_M = 80 M^{-1} s^{-1})[7] (Figures 6 and 7).

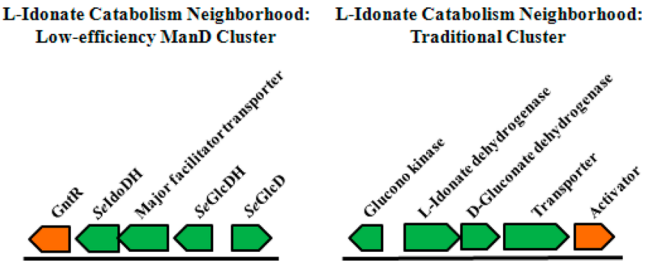


Figure 6. Genome neighborhoods of the previously described L-idonate catabolic pathway and the low-efficiency ManD-containing L-idonate catabolic pathway present in *S. enterica* subsp. *enterica* serovar *Enteritidis* str. P125109.

In the latter pathway, L-idonate is converted to D-gluconate via two dehydrogenases as previously described (*Se*IdoDH and *Se*GlcDH, Uniprot entries B5R538 and B5R540, respectively), but the subsequent steps of phosphorylation and dehydration apparently have shifted order: D-gluconate is dehydrated by *Se*GlcD to 2-keto-3-deoxy-D-gluconate, which can be catabolized by either the nonphosphorylated or phosphorylated Entner−Doudoroff pathway (Table 3). We have been unable to determine a physiological role for this pathway; it is a cryptic pathway, i.e., growth on L-idonate or D-gluconate at concentrations as high as 100 mM did not result in upregulation of the genes (Figures S6 and S7 of the Supporting Information).

As for the pathways for catabolism of L-gulonate using (1) *Cs*GulDH, *Cs*FR, *Cs*UxuA, and *Cs*ManD or (2) RspA, RspB, and RspD, the pathway for catabolism of L-idonate using *Se*IdoDH, *Se*GlcDH, and *Se*GlcD demonstrates the successive use of two dehydrogenases to epimerize the configuration of carbon 5. Although all three pathways are found via genome neighorhoods involving the low-efficiency ManDs, the physiological purposes, if any, of these dehydratases have yet to be discovered.

Conclusion. Our investigations into the role of the low-efficiency *Cs*ManD discovered a novel pathway for the catabolism of L-gulonate in *C. salexigens* DSM3043 via oxidation or reduction of carbon 5 (Figure 2). This pathway was proposed more than 30 years ago for *E. coli* K-12 when it was observed that spontaneous mutations allowed growth on L-gulonate, although the enzyme or gene for L-gulonate oxidation was not identified.[20] On the basis of sequence homology to *Cs*GulDH, *He*GulDH, and RspB, we propose that the likely L-gulonate 5-dehydrogenase in *E. coli* K-12 for which Cooper described the activity is Uniprot entry P38105 (44% identical to *Cs*GulDH, 46% identical to *He*GulDH, and 98% identical to RspB).

We have determined that low-efficiency D-gluconate and D-mannonate dehydratases need not be required for growth, although their genome neighborhoods encode enzymes sufficient to constitute catabolic pathways for L-idonate and L-gulonate, respectively. Perhaps, these "cryptic" pathways have been silenced as a result of the lack of selective pressure or, alternatively, are evolving to meet new metabolic needs.

dx.doi.org/10.1021/bi500837w | *Biochemistry* 2014, 53, 5692−5699

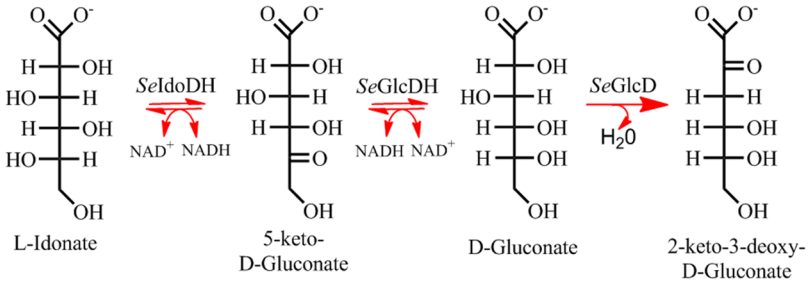


Figure 7. *In vitro* catabolic pathway for the consumption of L-idonate present in *S. enterica* subsp. *enterica* serovar *Enteritidis* str. P125109.

Table 3. Kinetics of Enzymes Encoded by the *Se*GlucD Operon

protein	substrate	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
*Se*IdoDH	L-idonate	1.2 ± 0.09	4.8×10^2
*Se*GlcDH	5-keto-D-gluconate	3.4 ± 0.5	3.8×10^3
*Se*GlcD[7]	D-gluconate	0.05 ± 0.003	8.0×10

Enzyme promiscuity has long been hypothesized to be the starting point of evolution.[22] A similar view, which is supported by laboratory evolution, is that the emergence of new functions is a gradual process in which the evolution of a low-efficiency, promiscuous activity is the most favored option rather than starting "from scratch".[23] This ideology supports the proposal that low-efficiency GlcDs are in evolutionary flux toward a new function if one assumes that the high-efficiency ManD activity is the progenitor function in the ManD subgroup. These hypotheses imply that low-efficiency ManDs are such because they are under no selective pressure: they may be derived from high-efficiency ManDs and are becoming less efficient, or they are relics that have not yet evolved to high-efficiency forms because of a lack of selective pressure. It is possible that we have not yet discovered the true physiological role of the low-efficiency ManDs and GlcDs. Finally, assessing the physiological roles of these pathways also may be complicated because in Nature *E. coli* and *S. enteriditis* are members of bacterial communities; i.e., these pathways may be activated by unknown metabolites produced by other members of the community.

■ ASSOCIATED CONTENT

Supporting Information

Acid sugar library, ThermoFluor library, and the ThermoFluor hit list for *Cs*GntR. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

This work describes characterization of *in vitro* enzymatic activities of proteins that have been deposited as UniProt entries E1V4Y1, Q1QT83, and Q8FHD0. Proteins deposited as UniProt entries Q1QT84 and Q1QT88 are also characterized by homology and conserved genome context.

■ AUTHOR INFORMATION

Corresponding Author

*Institute for Genomic Biology, University of Illinois, 1206 W. Gregory Dr., Urbana, IL 61801. E-mail: j-gerlt@illinois.edu. Phone: (217) 244-7414. Fax: (217) 333-0508.

Funding

This research was supported by a program project grant and three cooperative agreements from the National Institutes of Health (P01GM071790, U54GM093342, U54GM074945, and U54GM094662).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Drs. Katie Whalen and Brian San Francisco for assistance with manuscript preparation.

■ ABBREVIATIONS

*Cs*GntR, GntR from *C. salexigens* DSM3043; *Cs*GulDH, 5-dehydrogenase from *C. salexigens* DSM3043; *Cs*FR, fructuronate reductase from *C. salexigens* DSM3043; *Cs*ManD, mannonate dehydratase from *C. salexigens* DSM3043; *Cs*UxuA, UxuA from *C. salexigens* DSM3043; ENS, enolase superfamily; Gci, D-galactarolactone cycloisomerase; GlcD, D-gluconate dehydratase; GlucD, glucarate dehydratase; *He*GulDH, L-gulonate 5-dehydrogenase from *H. elongata* DSM 2581 L-gulonate; KO, knockout; ManD, D-mannonate dehydratase; *Se*IdoDH, L-idonate dehydrogenase from *S. enteriditis*; *Se*GlucDH, D-gluconate dehydrogenase from *S. enteriditis*; TRAP, tripartite ATP-independent periplasmic transporter; Udh, uronate dehydrogenase.

■ REFERENCES

(1) Rakus, J. F., Fedorov, A. A., Fedorov, E. V., Glasner, M. E., Vick, J. E., Babbitt, P. C., Almo, S. C., and Gerlt, J. A. (2007) Evolution of enzymatic activities in the enolase superfamily: D-Mannonate dehydratase from *Novosphingobium aromaticivorans*. *Biochemistry 46* (45), 12896−12908.

(2) Yew, W. S., Fedorov, A. A., Fedorov, E. V., Rakus, J. F., Pierce, R. W., Almo, S. C., and Gerlt, J. A. (2006) Evolution of enzymatic activities in the enolase superfamily: L-Fuconate dehydratase from *Xanthomonas campestris*. *Biochemistry 45*, 14582−14597.

(3) Gulick, A. M., Hubbard, B. K., Gerlt, J. A., and Rayment, I. (2000) Evolution of enzymatic activities in the enolase superfamily: Crystallographic and mutagenesis studies of the reaction catalyzed by D-glucarate dehydratase from *Escherichia coli*. *Biochemistry 39*, 4590−4602.

(4) Babbitt, P. C., Mrachko, G. T., Hasson, M. S., Huisman, G. W., Kolter, R., Ringe, D., Petsko, G. A., Kenyon, G. L., and Gerlt, J. A. (1995) A functionally diverse enzyme superfamily that abstracts the α protons of carboxylic acids. *Science 267*, 1159−1161.

(5) Gerlt, J. A., and Babbitt, P. C. (2001) Divergent evolution of enzymatic function: Mechanistically diverse superfamilies and functionally distinct suprafamilies. *Annu. Rev. Biochem. 70*, 209−246.

(6) Gerlt, J. A., Babbit, P. C., and Rayment, I. (2005) Divergent evolution in the enolase superfamily: The interplay of mechanism and specificity. *Arch. Biochem. Biophys. 433*, 59−70.

(7) Wichelecki, D. J., Balthazor, B. M., Chau, A. A., Vetting, M. W., Fedorov, A. A., Fedorov, E. V., Lukk, T., Patskovsky, Y. V., Stead, M. B., Hillerich, B. S., Seidel, R. D., Almo, S. C., and Gerlt, J. A. (2014) Discovery of function in the enolase superfamily: D-Mannonate and D-

gluconate dehydratases in the D-mannonate dehydratase subgroup. *Biochemistry 53*, 2722−2731.
(8) Wichelecki, D. J., Graff, D. C., Al-Obaidi, N., Almo, S. C., and Gerlt, J. A. (2014) Identification of the physiological role of a high efficiency enolase superfamily member mannonate dehydratase in *Caulobacter crescentus* NA1000. *Biochemistry 53*, 4087−4089.
(9) Arahal, D. R., Garcia, M. T., Vargas, C., Canovas, D., Nieto, J. J., and Ventosa, A. (2001) *Chromohalobacter salexigens* sp. nov., a moderately halophilic species that includes *Halomonas elongata* DSM 3043 and ATCC 33174. *Int. J. Syst. Evol. Microbiol. 51*, 1457−1462.
(10) Toporowski, M. C., Nomellini, J. F., Awram, P., and Smit, J. (2004) Two Outer Membrane Proteins are Required for Maximal Type I Secretion of the *Caulobacter crescentus* S-Layer Protein. *J. Bacteriol. 186*, 8000−8009.
(11) Hottes, A. K., Meewan, M., Yang, D., Arana, N., Romero, P., McAdams, H. H., and Stephens, C. (2003) Transcriptional Profiling of *Caulobacter crescentus* during Growth on Complex and Minimal Media. *J. Bacteriol. 186*, 1448−1461.
(12) Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A. 97*, 6640−6645.
(13) Ashwel, G. (1962) Enzymes of glucuronic and galacturonic acid metabolism in bacteria. *Methods Enzymol. 5*, 190−208.
(14) Zhang, Q., Gao, F., Peng, H., Cheng, H., Liu, Y., Tang, J., Thompson, J., Wei, G., Zhang, J., Du, Y., Yan, J., and Gao, G. F. (2009) Crystal Structures of *Streptococcus suis* Mannonate Dehydratase (ManD) and Its Complex with Substrate: Genetic and Biochemical Evidence for a Catalytic Mechanism. *J. Bacteriol. 191*, 5832−5837.
(15) Chang, Y. F., and Feingold, D. S. (1969) Hexuronic acid dehydrogenase of *Agrobacterium tumefaciens*. *J. Bacteriol. 99*, 667−673.
(16) Wagner, G., and Hollmann, S. (1976) Uronic acid dehydrogenase from *Pseudomonas syringae*. Purification and properties. *Eur. J. Biochem. 61*, 589−596.
(17) Blumenthal, H. J., and Fish, D. C. (1963) Bacterial conversion of D-glucarate to glycerate and pyruvate. *Biochem. Biophys. Res. Commun. 11*, 239−243.
(18) Bouvier, J. T., Groninger-Poe, F. P., Vetting, M., Almo, S. C., and Gerlt, J. A. (2014) Galactaro δ-lactone isomerase: Lactone isomerization by a member of the amidohydrolase superfamily. *Biochemistry 53*, 614−616.
(19) Andberg, M., Maaheimo, H., Boer, H., Penttilä, M., Koivula, A., and Richard, P. (2012) Characterization of a novel *Agrobacterium tumefaciens* galactarolactone cycloisomerase enzyme for direct conversion of D-galactarolactone to 3-deoxy-2-keto-L-threo-hexarate. *J. Biol. Chem. 287*, 17662−17671.
(20) Cooper, R. A. (1980) The pathway for L-gulonate catabolism in *Escherichia coli* K-12 and *Salmonella typhimurium* LT-2. *FEBS Lett. 115*, 63−67.
(21) Bausch, C., Peekhaus, N., Utz, C., Blais, T., Murray, E., Lowary, T., and Conway, T. (1998) Sequence analysis of the GntII (subsidiary) system for gluconate metabolism reveals a novel pathway for L-idonic acid catabolism in *Escherichia coli*. *J. Bacteriol. 180*, 3704−3710.
(22) Jenson, R. A. (1976) Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol. 30*, 409−425.
(23) Peisajovich, S. G., and Tawfik, D. S. (2007) Protein engineers turned evolutionists. *Nat. Methods 4*, 991−994.