JB
Journals.ASM.org

Regulation and Evolution of Malonate and Propionate Catabolism in Proteobacteria

I. A. Suvorova,[a] D. A. Ravcheev,[a,b] and M. S. Gelfand[a,c]

Institute for Information Transmission Problems (the Kharkevich Institute), Moscow, Russia[a]; Sanford-Burnham Medical Research Institute, La Jolla, California, USA[b]; and Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russia[c]

Bacteria catabolize malonate via two pathways, encoded by the *mdc* and *mat* genes. In various bacteria, transcription of these genes is controlled by the GntR family transcription factors (TFs) MatR/MdcY and/or the LysR family transcription factor MdcR. Propionate is metabolized via the methylcitrate pathway, comprising enzymes encoded by the *prp* and *acn* genes. PrpR, the Fis family sigma 54-dependent transcription factor, is known to be a transcriptional activator of the *prp* genes. Here, we report a detailed comparative genomic analysis of malonate and propionate metabolism and its regulation in proteobacteria. We characterize genomic loci and gene regulation and identify binding motifs for four new TFs and also new regulon members, in particular, tripartite ATP-independent periplasmic (TRAP) transporters. We describe restructuring of the genomic loci and regulatory interactions during the evolution of proteobacteria.

Malonate can be used as a carbon source by a variety of bacteria, such as *Acinetobacter calcoaceticus*, *Klebsiella pneumoniae*, *Pseudomonas fluorescens*, and *Pseudomonas putida* (15). It is a competitive inhibitor of succinate dehydrogenase. Malonate was found in symbiotic legumes and in developing rat brains (15, 16). Malonate metabolism is essential in symbiotic nitrogen metabolism, since mutant bacteria with deleted malonate metabolic genes lose this symbiosis ability (13). In addition, malonate metabolic genes have been used for generation of the industrial strain of *Streptomyces* for the production of antibiotics (13).

Two groups of structural genes involved in malonate metabolism have been characterized. The first group was described for *Rhizobium leguminosarum* (16). It comprises three clustered malonate metabolic genes, *matA*, *matB*, and *matC*, and the divergently transcribed gene of the malonate regulator *matR* (Table 1). The MatR transcription factor (TF) belongs to the FadR subfamily of the GntR family.

Another system for malonate metabolism was described for *Acinetobacter calcoaceticus* KCCM 4090 (14, 15). This system is encoded by the operon of structural genes *mdcLMACDEGBH* and the divergently transcribed regulator gene *mdcY* (Table 1). Like MatR, MdcY is also a FadR family TF. Some *Gammaproteobacteria* have a different malonate transporter gene (compared to that of *A. calcoaceticus*), *mdcF* instead of *mdcLM* (19). The gene organization of the malonate operon of *A. calcoaceticus* is similar to that of *K. pneumoniae* and *P. putida* (14, 19), but these and many other *Gammaproteobacteria* have another TF gene adjacent to the malonate operon, encoding the MdcR transcription factor from the LysR family. MdcR activates expression of the *mdc* genes and represses its own transcription (14, 19).

Propionate can also serve as a single carbon source for many bacteria. Its metabolism is strongly connected to the malonate metabolic pathway and central metabolism, e.g., the tricarboxylic acid (TCA) cycle (Fig. 1). Propionate is converted to pyruvate via the methylcitrate pathway, enzymes for which are encoded by the *prpBCDE*, *acnB*, and *acnD* genes (3, 4). AcnB is a bifunctional enzyme that also belongs to the TCA cycle and the glyoxylate pathway (3, 4). Propionate also participates in the citramalate cycle comprising the products of the *mutB* (*mcm*), *meaB*, *pccBA*, and *mce* (*epi*) genes (11, 17) (Table 1). PrpR, a sigma 54-dependent TF belonging to the Fis family, is known to be a transcriptional activator of the *prp* genes in *Escherichia coli* and *Ralstonia eutropha* HF39 (3). Moreover, it is known that a GntR family regulator gene is colocalized with the *prp* cluster in the *Pseudomonas* spp. and *Vibrio cholerae*, and it is likely one more propionate regulator (3).

Here, we report a detailed comparative genomic analysis of malonate and propionate metabolism and its regulation in proteobacteria. We characterized genomic loci and gene regulation, identified binding motifs and new regulon members, and predicted novel TFs. We also describe rearrangements of genomic loci and regulatory interactions during the evolution of proteobacteria.

MATERIALS AND METHODS

Computational analysis of regulons. The genomic sequences of the analyzed proteobacteria were obtained from GenBank (2); the genomes are listed in Table S1 in the supplemental material. Orthologs of TFs in bacterial genomes were identified by PSI-BLAST (1) searches with default parameters and confirmed by construction of phylogenetic trees for identified homologs and by colocalization with genes of the corresponding metabolic pathways. Amino acid sequence alignment was performed using the MUSCLE package (8). Phylogenetic trees were constructed with the PHYLIP package, using the Protdist program for the calculation of distances and the maximum likelihood method with default parameters for the tree construction (9). For all bacterial species that had any of the TFs described in the introduction, the comparative genomics-based reconstruction of the malonate and propionate regulons was performed. Nucleotide positional weight matrices (profiles) for all TF-binding sites (see Table S2 in the supplemental material) were constructed by the SignalX program (18), using training sets of upstream regions of genes

Received 8 February 2012 Accepted 4 April 2012
Published ahead of print 13 April 2012
Address correspondence to M. S. Gelfand, gelfand@iitp.ru.
Supplemental material for this article may be found at http://jb.asm.org/.

doi:10.1128/JB.00163-12

TABLE 1 Genes involved in malonate and propionate catabolism

Gene	Function of product[a]
mdcL	Malonate transporter subunit
mdcM	Malonate transporter subunit
mdcA	Malonate/acetyl-CoA transferase, alpha subunit
mdcC	Acyl carrier protein, delta subunit
mdcD	Malonyl-CoA decarboxylase, beta subunit
mdcE	Protein involved in the stability of the enzyme complex or codecarboxylase, gamma subunit
mdcF	Malonate transporter
mdcG	Holo-ACP synthase
mdcB	Triphosphoribosyldephospho-CoA synthase, a protein involved in the formation of the prosthetic group precursor
mdcH	Malonyl-CoA/ACP transacylase
matA	Malonyl-CoA decarboxylase
matB	Malonyl-CoA synthetase
matC	Malonate transporter
matPQM	TRAP malonate transporter
matR (*mdcY*)	GntR family regulator of malonate metabolism
mdcR	LysR family regulator of malonate metabolism
prpB	2-Methylisocitrate lyase
prpC	2-Methylcitrate synthase
prpD	2-Methylcitrate dehydratase
prpE	Propionyl-CoA synthetase
prpF	Aconitase accessory protein
acnD	Aconitate hydratase
acnB	Bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase
acnA	Aconitate hydratase
prpR	Regulator of propionate metabolism
mutB	Metylmalonyl-CoA mutase
meaB	Putative auxiliary metallochaperone, involved in protection and assembly of methylmalonyl-CoA mutase (G3E family of P-loop GTPases)
pccA	Propionyl-CoA carboxylase, subunit alpha
pccB	Propionyl-CoA carboxylase, subunit beta
mce	Methylmalonyl-CoA epimerase

[a] CoA, coenzyme A; ACP, acyl carrier protein.

from the analyzed bacteria. Score thresholds for identification of sites were selected so that candidate sites upstream of functionally relevant genes were accepted, while the fraction of genes preceded by candidate sites did not exceed 5% in each studied genome. Under these conditions, for long, conserved motifs, the number of candidate sites per genome did not exceed 50.

Candidate binding sites were confirmed by phylogenetic footprinting (21). In summary, we manually analyzed alignments of upstream regions of orthologous genes and identified groups of consecutive conserved positions, relying on the assumption that binding sites are more conserved than adjacent intergenic regions. A computational search for candidate TF-binding sites in upstream gene regions (−350 to +50 nucleotides [nt] relative to the start codon) was performed using the Genome Explorer package (18). The threshold scores for all types of TF-binding sites are given in Table S2 in the supplemental material. Weaker sites (with scores 10% less than the threshold) were also taken into account if their positions were similar to positions of stronger sites upstream of orthologous genes and there were no stronger competing sites in the same intergenic region. New candidate members were assigned to a regulon if they were preceded (as a single gene or as a part of an operon) by candidate TF-binding sites in at least four genomes. The requirement that a gene should be preceded by a candidate site in at least four genomes was established based on empirical evidence. A stricter requirement would lead to the loss of known regulon members, such as MatR/MdcY-regulated *matC*, which is present in the studied regulon in only four bacteria. A weaker criterion of three genomes yields candidate regulon members with clearly irrelevant function. In most cases, including all nontrivial predictions, the actual number of orthologs preceded by candidate sites was considerably higher (e.g., tripartite ATP-independent periplasmic [TRAP] transporters) (see Table S3 in the supplemental material). Moreover, it was also required that candidate sites were observed in sufficiently distant species, so that the site conservation could not be explained by residual sequence similarity.

Genes were considered to belong to one operon if they were transcribed in the same direction, with intergenic distances not exceeding 200 nt. Motif logos were constructed using WebLogo (7).

RESULTS AND DISCUSSION

Phylogenetic distribution of malonate metabolism genes and their regulators between bacteria. (i) MatR/MdcY. MatR and MdcY have been experimentally identified in different bacteria (14, 16), and this seems to be the only reason for the different names. The amino acid sequences of MatR and MdcY are very similar (49% identity), these TFs are closely related according to phylogenetic analysis, and their previously predicted binding sites are also similar (14, 16, 20). Moreover, there is no evident phylogenetic separation between TFs whose genes are colocalized in genomes with either *mat* or *mdc* genes. Hence, this TF is referred to herein as MatR/MdcY.

An exhaustive BLAST search identified a number of MatR/MdcY TFs, mostly in the *Alphaproteobacteria* and several *Betaproteobacteria* and *Gammaproteobacteria* (see Table S3 and Fig. S1 in the supplemental material).

Phylogenetic footprinting of regions upstream of the *mat* and *mdc* genes revealed the MatR/MdcY binding motif with the consensus TTGTATACAA (14, 16, 20) (see Fig. S7A in the supplemental material). This motif coincides with the one previously predicted and confirmed by the DNase I footprint assay (14, 16).

The distribution of the *mdc* and *mat* operons in bacterial taxa is very flexible. Some bacteria have both *mat* and *mdc* genes, while others have either the *mdc* or *mat* operon. In most bacteria, the *matR* (*mdcY*) gene is colocalized with the regulated genes. Most bacteria possess only one malonate regulator of the MatR/MdcY type, although *Methylobacterium* spp. have two copies of this TF, one clustered with the *mdc* operon and the other one with tripartite ATP-independent periplasmic (TRAP) genes and, in some cases, like in *Methylobacterium* sp. strain 4-46, with TRAP and *matAB* genes.

(ii) TRAP transporters. The TRAP transporters are characterized by the usage of an electrochemical ion gradient as the driving force for solute accumulation. The best-characterized TRAP is the high-affinity C4-dicarboxylate transport (Dct) system formed by three proteins: extracytoplasmic solute receptor subunit (DctP) and small (DctQ) and large (DctM) integral membrane proteins (10, 12, 22). Representatives of the TRAP family are present in a wide range of eubacteria and archaea. Some organisms possess a single TRAP system (*E. coli*), while others have several TRAP transporters (*Pseudomonas aeruginosa* and *Bacillus halodurans*). Probable substrates of the TRAP transporters are L-xylulose, gluconate, mannonate, succinate, etc. (10, 12, 22).

Many *Alphaproteobacteria* and some *Beta-* and *Gammaproteobacteria* have C3-dicarboxylic acid TRAP genes colocalized with genes encoding TFs or enzymes of malonate metabolism, either forming an operon with them or preceded by their own candidate

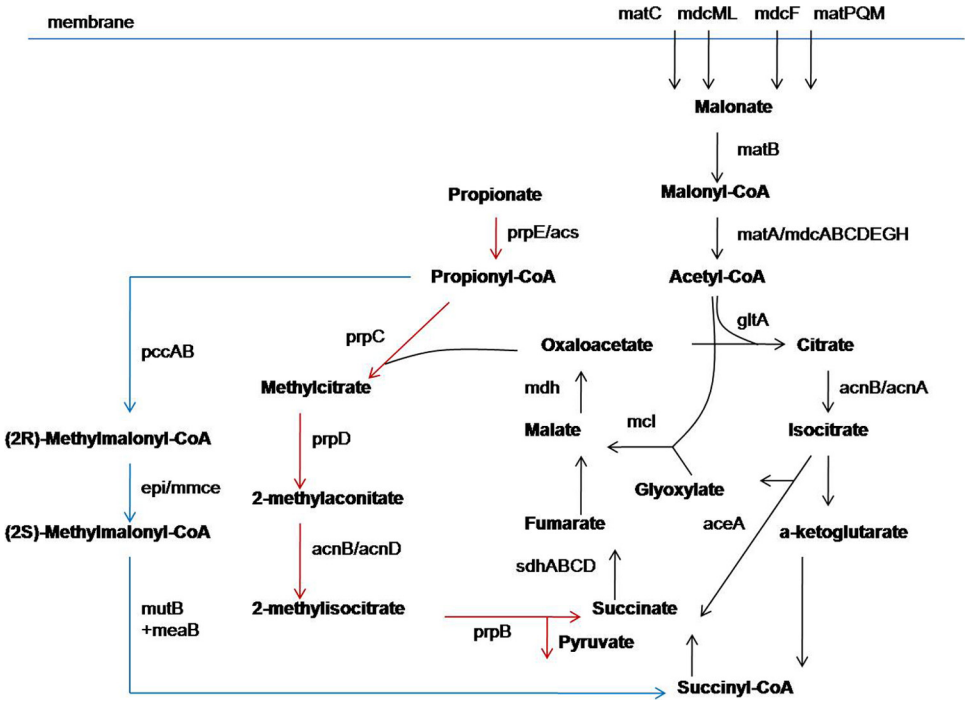

FIG 1 Malonate and propionate metabolism. The methylcitrate pathway is marked with red lines, the citramalate cycle with blue lines.

binding sites. In *Sinorhizobium meliloti*, these TRAP components were named *matPQM* (6). Null mutants for each of the *matPQM* genes were shown to be unable to grow on the minimal medium containing malonate as the sole carbon source (6). In this study, in most cases the presence of TRAP genes in the regulated operons correlated with the absence of known malonate transporters *mdcLM*, *mdcF*, and *matC*. Among completely sequenced genomes, orthologs of these TRAP transporters were found only in bacteria that had malonate metabolic genes. These observations suggest that these TRAP transporters are involved in malonate transport and belong to the MatR/MdcY regulon.

(iii) MdcR. MdcR orthologs were found in the *Gammaproteobacteria* and in the *Burkholderiales* among the *Betaproteobacteria*, but not in the *Alphaproteobacteria*. In the *Gammaproteobacteria*, MdcR, when present, almost always is the only malonate regulator, whereas in the *Betaproteobacteria*, it often accompanies other malonate regulators. A phylogenetic tree of all found MdcR TFs is shown in Fig. S2 in the supplemental material.

In most *Gammaproteobacteria*, MdcR controls a single operon, *mdcABCDEGHLM* or *mdcABCDEFGH*. In the *Betaproteobacteria*, the *mdc* genes have a diverse arrangement, usually as a single operon or sometimes in two operons (in *Burkholderia phytofirmans* PsJN and in *Ralstonia eutropha* JMP134), with *mdcLM* genes in a different locus. The *mdcLM* transporter genes are typical for the MdcR regulon. In several *Gammaproteobacteria* (e.g., *Citrobacter koseri* ATCC BAA-895 and *Enterobacter* sp. strain 638) they are replaced with *mdcF*, and in some *Betaproteobacteria* (*Delftia acidovorans* SPH-1 and *Methylibium petroleiphilum* PM1), they are replaced with *matC* (see Table S3 in the supplemental material).

The candidate MdcR binding motif was predicted by the phylogenetic footprinting of the *mdcA* upstream region. It is a 23-nt palindrome with the consensus sequence ATCATTACCCTgAggg TAATGAT (lowercase letters denote less-conserved nucleotides) (see Fig. S7B in the supplemental material). In most *Gamma-* and *Betaproteobacteria*, *mdcR* is not autoregulated. The exceptions are *Ralstonia pickettii* 12J and *Psychromonas ingrahamii* 37, where *mdcR* is located in a divergon with other *mdc* genes and thus shares the candidate binding site.

(iv) PrpR. PrpR is a transcriptional activator of the *prp* genes in some *Gammaproteobacteria* (*Enterobacteriales* and *Xanthomonadales*) and *Betaproteobacteria* (*Burkholderiales*). A phylogenetic tree of all found PrpR TFs is shown in Fig. S3 in the supplemental material.

All *Enterobacteriales* have propionate genes organized in the divergon *prpR prpBCDE*. In most *Xanthomonadales* and *Betaproteobacteria*, the propionate divergon has the *prpR prpBC-acnD-prpF* structure. Some *Betaproteobacteria* have diverse propionate regulon structure and several other TFs regulating propionate metabolism (see below and Table S3 in the supplemental material).

The candidate PrpR binding motif was identified using phylogenetic footprinting of the *prpB* and *prpR* upstream regions. The predicted binding motif is a 16-nt palindrome with the consensus sequence rTTTCAwwwwTGAAAy (lowercase letters denote less-conserved nucleotides) (see Fig. S8A in the supplemental material). PrpR is a sigma 54-dependent transcription activator, and indeed, candidate sigma 54 promoters were identified upstream of all propionate gene clusters belonging to the PrpR regulon.

New regulators of the malonate and propionate metabolism. (i) MlnR* (FadR subfamily of the GntR family). Some bacteria from several families of the *Betaproteobacteria* (*Alcaligenaceae*, *Burkholderiaceae*, *Comamonadaceae*, and *Rhodocyclaceae*) and *Gammaproteobacteria* (*Ectothiorhodospiraceae* and *Xanthomonadaceae*) have another GntR family TF adjacent to the malonate metabolism genes. This TF is related to MatR/MdcY but is not its ortholog (confirmed by PSI-BLAST and the phylogenetic tree; see

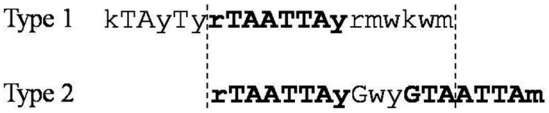

FIG 2 Common parts of the two types of MlnR* binding motifs. The common parts are marked with dashed lines; the repeated part in the type 1 and type 2 motifs is set in bold.

Fig. S1 in the supplemental material). This protein was named MlnR* (here and below, an asterisk denotes a newly given name).

The MlnR* TF was predicted to control malonate utilization and a part of the citramalate cycle. In some *Betaproteobacteria*, two paralogous copies of MlnR* were found. In that case, one copy is colocalized with the *matAB* genes, whereas the other one is clustered on the chromosome with genes of the citramalate cycle, *mutB*, *meaB*, *pccBA*, and *mce*. In bacteria having only one copy of this TF, the *mlnR** gene is clustered with both *mat* and the citramalate cycle genes (e.g., *Azoarcus* sp. strain BH72, *Dechloromonas aromatica*) or colocalized with either malonate metabolic genes (e.g., *Bordetella* spp., *Ralstonia metallidurans*, and *Cupriavidus taiwanensis*) or genes of the citramalate cycle (e.g., *Delftia acidovorans*, *Leptothrix cholodnii*, and *Polaromonas naphthalenivorans*). In *Xanthomonadaceae*, the *mlnR** gene is clustered with *mdc* genes and *matC*. *Alkalilimnicola ehrlichii* MLHE-1 (*Ectothiorhodospiraceae*) has the *mlnR** gene colocalized with *matAB* and the malonate TRAP genes *matPQM*, and such gene organization resembles the one in the MatR/MdcY regulon in many *Alphaproteobacteria*.

Phylogenetic footprinting of the *mlnR** upstream regions revealed two types of candidate binding motifs. The first one, with the consensus sequence TTATTCATAATTATGAATAA (type 1; see Fig. S7C in the supplemental material), was found in the *Betaproteobacteria*. The second type of the predicted MlnR* binding motif was found in the *Xanthomonadaceae* and *Alkalilimnicola ehrlichii* MLHE-1. Its consensus, ATAATTACGATGTAATTAC (type 2; see Fig. S7D), partially coincides, after a shift, with the betaproteobacterial type 1 motif. The type 1 and type 2 motifs contain a common, short palindromic motif, RTAATTAY, with two repeats in the type 2 motif and only one copy in the type 1 motif (Fig. 2). The type 1 and type 2 profiles cross-recognize some sites.

(ii) GntR and LysR family TFs in *Burkholderia* spp. Bacteria from the genus *Burkholderia* have the *mdc* genes but lack orthologs of MatR/MdcY. MdcR orthologs among *Burkholderia* spp. are present only in *Burkholderia multivorans* ATCC 17616, *Burkholderia* sp. strain 383, and *Burkholderia phytofirmans* PsJN; moreover, MdcR binding sites upstream of the malonate metabolic genes are found only in *Burkholderia phytofirmans* PsJN. To predict other possible regulators of malonate metabolism in *Burkholderia* spp., we searched for TF genes colocalized with the *mdc* genes. Indeed, some *Burkholderia* spp. (see Table S3 in the supplemental material) have a gene encoding a different GntR family TF near the *mdc* genes. This TF is distantly related to MatR/MdcY and is not its ortholog (confirmed by PSI-BLAST and the phylogenetic tree; see Fig. S1 in the supplemental material). Other *Burkholderia* spp. contain a gene encoding a LysR family TF in the *mdc* locus. Orthologs of these GntR and LysR family TFs are absent in other genomes studied in this work. An attempt to find binding motifs of these regulators by phylogenetic footprinting failed because of the high degree of similarity of the entire intergenic regions of these closely related species.

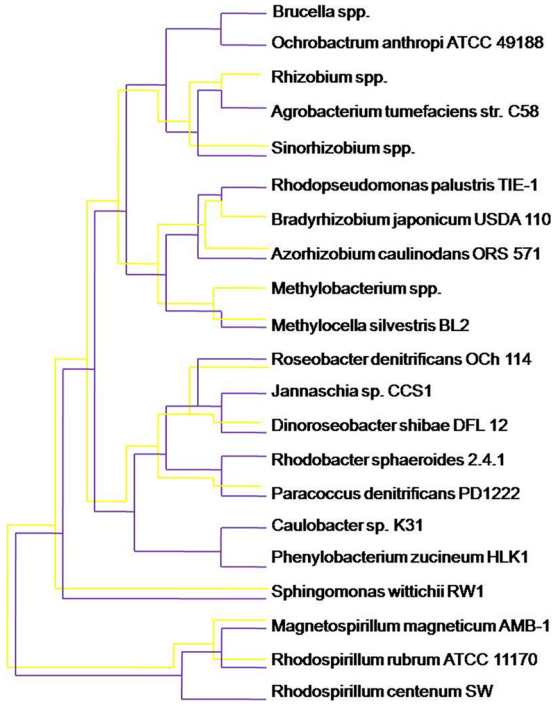

FIG 3 Regulation in *Alphaproteobacteria*. Line color denotes the mode of regulation, using the following color code: yellow, MatR/MdcY, and violet, PrpQ*.

(iii) PrpR* (FadR subfamily of the GntR family). Multiple representatives of the *Gammaproteobacteria* (*Alteromonadales*, *Oceanospirillales*, *Pseudomonadales*, and *Vibrionales*), some *Betaproteobacteria* (*Burkholderiales*), and even *Deltaproteobacteria* (*Geobacter metallireducens* GS-15) have a GntR family, FadR subfamily TF as a regulator of propionate metabolism. This TF is hereafter referred to as PrpR*. A phylogenetic tree of all found PrpR* TFs is shown in Fig. S4 in the supplemental material.

Most *Alteromonadales* and *Pseudomonadales* have propionate metabolic genes organized in the *prpR*BC-acnD-prpFD* operon, while some bacteria from these taxa have shorter operons, *prpR*BC-acnD-prpF* or *prpR*BC-acnD*. Most *Vibrionales* have the *prpR*BC-acnD-prpFE* operon, while *Oceanospirillales* have the *prpR*BCD* operon. A similar operon organization, *prpR*BDC*, was observed in *Geobacter metallireducens* GS-15.

Geobacter metallireducens is the only member of the *Deltaproteobacteria* that has any of the regulators studied in this work. Even the congeners *Geobacter sulfurreducens* and *Geobacter uraniireducens*, though having the *prp* metabolic genes, lack orthologous transcription factors and thus were not considered. Among other *Delta-* and *Epsilonproteobacteria*, only *Helicobacter hepaticus* has a complete propionate metabolic pathway (*prpEBCD* genes), but it lacks any known or predicted propionate TFs.

In the *Betaproteobacteria*, PrpR* not only is rare but also seems to be not the main propionate regulator, even if it is present in a genome, as most propionate genes are regulated by a HutC-type TF (SdhR*; see below and Table S3 in the supplemental material). Among the *Betaproteobacteria*, PrpR* controls propionate metabolic genes only in *Verminephrobacter eiseniae* EF01-2 and *Bordetella* spp.

The predicted PrpR* binding motif, identified by phylogenetic

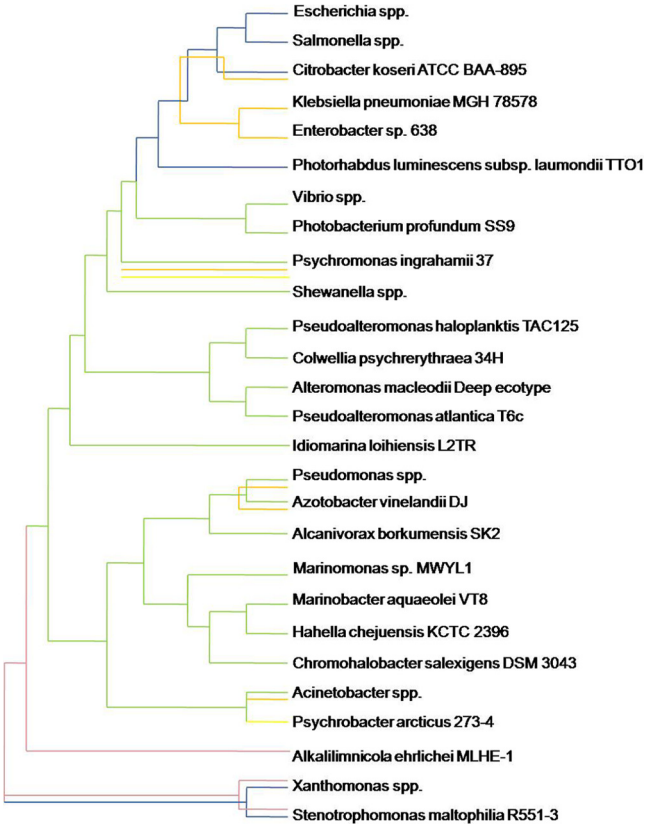

FIG 4 Regulation in *Gammaproteobacteria*. Line color denotes the mode of regulation, using the following color code: yellow, MatR/MdcY; orange, MdcR; pink, MlnR*; dark blue, PrpR; and green, PrpR*.

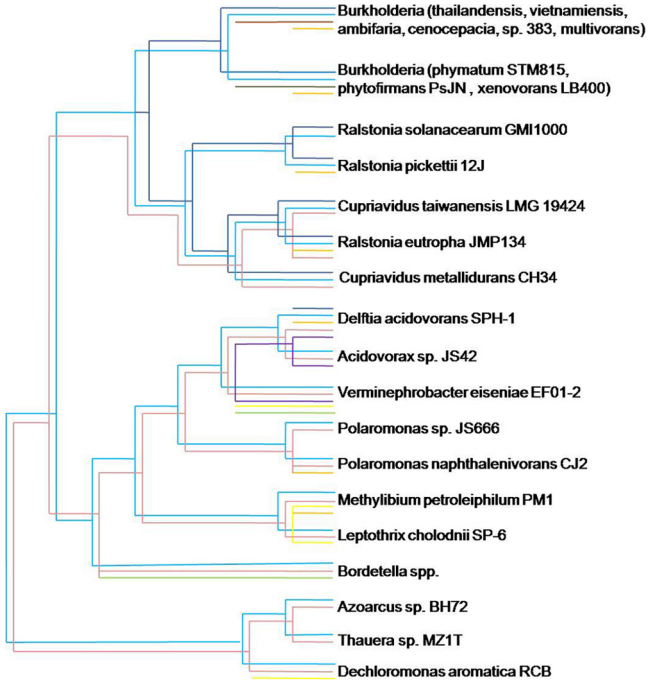

FIG 5 Regulation in *Betaproteobacteria*. Line color denotes the mode of regulation, using the following color code: yellow, MatR/MdcY; orange, MdcR; pink, MlnR*; light blue, SdhR*; dark blue, PrpR; green, PrpR*; violet, PrpQ*; light brown, LysR family TF (*Burkholderia*); and dark green, GntR family TF (*Burkholderia*).

footprinting, is a 12-nt palindrome with the consensus sequence ATTGTCGACAAT (see Fig. S8B in the supplemental material).

(iv) PrpQ* (XRE family). Most *Alpha-* and some *Betaproteobacteria* have a gene encoding an XRE family TF that is likely the transcriptional regulator of propionate metabolism. This TF is hereafter referred to as PrpQ*. A phylogenetic tree of all found PrpQ* TFs is shown in Fig. S5 in the supplemental material.

The candidate PrpQ* binding motif was identified by phylogenetic footprinting of the upstream regions of the *prpQ** and *pccBA* operons. The motif is a short (8-nt) palindrome with the consensus sequence TTTGCrAA, often present in multiple copies upstream of regulated genes (see Fig. S8C in the supplemental material). In *Betaproteobacteria*, the reconstructed PrpQ* regulon includes genes for a part of the methylcitrate pathway (see Table S3 in the supplemental material). In *Alphaproteobacteria*, PrpQ* binding sites were found upstream of genes encoding enzymes of the citramalate and/or methylcitrate pathways, such as the *prp*, *acnD*, *pccBA*, and *mutB* genes (see Table S3).

(v) SdhR* (HutC subfamily of the GntR family). Multiple *Betaproteobacteria* (*Burkholderiales*) have a GntR family, HutC subfamily TF that, according to the colocalization, may control expression of genes encoding enzymes involved in the TCA cycle, such as succinate dehydrogenase (*sdhABCD*), citrate synthase (*gltA*), bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase (*acnB*), malate dehydrogenase (*mdh*), and some other catabolic genes, in particular *ygfY*, functionally connected to *sdhB*, and *tam*, encoding *trans*-aconitate 2-methyltransferase; this enzyme catalyzes monomethyl esterification of *trans*-aconitate at high velocities and affinities and *cis*-aconitate, isocitrate, and citrate at lower velocities and affinities (5). This TF is hereafter referred to as SdhR*. A phylogenetic tree of all found SdhR* TFs is shown in Fig. S6 in the supplemental material.

Phylogenetic footprinting of *sdhR**, *mdh*, and *acnB* upstream regions revealed a conserved area that is likely the SdhR* binding site. The predicted binding motif is a 26-nt palindrome with the consensus sequence TCTTATGTCTTATATAAGACATAAGA (see Fig. S8D in the supplemental material). This motif has an intrinsic symmetry, comprising two direct repeats of complementary motifs TCTTAT and ATAAGA in the distal and proximal parts of the palindrome, respectively.

In many *Betaproteobacteria* (see Table S3 in the supplemental material), SdhR* also controls a number of genes of the propionate utilization pathway (*prpB*, *prpC*, *prpD*, *acnA*, and *prpF*) and the glyoxylate shunt (malyl coenzyme A [malyl-CoA] lyase [*mcl*] and isocitrate lyase [*aceA*]).

Conclusions. The diversity of the organization of the malonate and propionate regulons in different taxonomic groups naturally leads to the question of the evolution of these regulatory systems.

The most consistent group is the *Alphaproteobacteria*. The mode of malonate and propionate regulation is the same among these bacteria: the malonate metabolic genes are regulated by MatR/MdcY, and propionate utilization is under PrpQ* regulation. Both these TFs are typical for the *Alphaproteobacteria* and are present in only a few *Beta-* and *Gammaproteobacteria*, seemingly as a result of horizontal gene transfer (Fig. 3).

Indeed, among the *Gammaproteobacteria*, the horizontal

TABLE 2 Distribution of malonate and propionate transcription factors among various classes of proteobacteria

Class of proteobacteria	No. of genera with TF						
	MatR/ MdcY	MlnR*	MdcR	PrpR	PrpR*	PrpQ*	SdhR*
Alpha	12	0	0	0	0	17	0
Beta	4	11	5	3	2	3	12
Gamma	2	3	7	6	16	0	0
Delta	0	0	0	0	1	0	0

transfer of *matR* (*mdcY*) occurred in the *Pseudomonadales* and *Psychromonadaceae* (confirmed by the phylogenetic tree [Fig. 4]; see also Fig. S1 in the supplemental material). *Ectothiorhodospiraceae* and *Xanthomonadaceae* have the GntR family TF named MlnR* as a malonate regulator, which is typical for the *Betaproteobacteria* (see below). It is likely they inherited it from their common ancestor with the *Betaproteobacteria* (see Fig. S1 in the supplemental material). In other *Gammaproteobacteria*, the malonate metabolism genes are regulated by MdcR, and this TF seems to be the original malonate regulator in this taxonomic group (Fig. 4). Propionate metabolism in the *Gammaproteobacteria* is mostly controlled by a GntR family TF named here PrpR*, but some *Enterobacteriales* and all analyzed *Xanthomonadales* use a Fis family TF, PrpR, as the propionate regulator. This TF is also present in the *Betaproteobacteria*, and according to the phylogenetic tree (see Fig. S3 in the supplemental material), either *Beta-* and *Gammaproteobacteria* inherited it from their common ancestor or the common ancestor of the *Enterobacteriales* and *Xanthomonadales* had it transferred from some ancient *Betaproteobacteria*. Both scenarios involve multiple losses of PrpR in a variety of lineages. The GntR family TF PrpR* is widespread among the *Gammaproteobacteria* (Fig. 4). This TF is also present in several *Betaproteobacteria* and the deltaproteobacterium *Geobacter metallireducens* GS-15. According to the branch localization in the phylogenetic tree, *G. metallireducens* got PrpR* as a result of horizontal gene transfer from a genome close to the ancestor of the *Alteromonadales*. It seems that some *Betaproteobacteria* had PrpR* transferred from the *Pseudomonadales* (for example, *Bordetella petrii* DSM 12804) (see Fig. S4 in the supplemental material), while some others inherited it from their common ancestor with the *Gammaproteobacteria* (e.g., *Bordetella bronchiseptica* RB50 and *Verminephrobacter eiseniae* EF01-2) (see Fig. S4). An interesting variant is seen in *Bordetella avium* 197N, which has two *prpR** genes, one likely obtained from the *Pseudomonadales* and the other inherited from the common ancestor with the *Gammaproteobacteria*. It is plausible that in the *Gammaproteobacteria* we observe an intermediate stage of replacement of PrpR by PrpR* (Fig. 4).

The most diverse regulation is seen in the *Betaproteobacteria* (Fig. 5 and Table 2). Overall, bacteria from this group possess at least five malonate regulators, including MatR/MdcY, which they presumably got from the *Alphaproteobacteria* via horizontal gene transfer (confirmed by the phylogenetic tree; see Fig. S1 in the supplemental material); MdcR, horizontally transferred (according to the phylogenetic tree; see Fig. S2 in the supplemental material) from the *Gammaproteobacteria* or inherited from the common ancestor of *Beta-* and *Gammaproteobacteria*; and LysR and GntR family regulators, found only in *Burkholderia* spp. However, most *Betaproteobacteria* contain a GntR family TF, MlnR*. Propionate utilization in the *Betaproteobacteria* is also under diverse regulation. Only few *Betaproteobacteria*, three closely related *Comamonadaceae* species, have PrpQ* as a result of the horizontal transfer from the *Alphaproteobacteria* (confirmed by the phylogenetic tree; see Fig. S5 in the supplemental material). Most *Betaproteobacteria* have SdhR*, either as the only propionate regulator or accompanied by PrpR or sometimes PrpR* (Fig. 5). It is interesting that most *Betaproteobacteria* that got MlnR* lack the Fis family TF PrpR, possibly due to the fact that these TFs control alternative pathways of propionate conversion (see Table S3 in the supplemental material).

Overall, we have reconstructed the malonate and propionate regulons, described their flexible and diverse regulation in proteobacteria, found new TFs that control malonate and propionate metabolism, and identified their candidate binding sites by positional and sequence analyses. The comparative genomic analysis also yielded new candidate members of the malonate regulon, namely, TRAP transporters.

ACKNOWLEDGMENTS

This study was partially supported by The Russian Foundation of Basic Research (grants 09-04-92745 and 10-04-00431), the Russian Academy of Sciences (via the Molecular and Cellular Biology program), and state contracts 14.740.11.0003 and 07.514.11.4007.

REFERENCES

1. **Altschul SF, et al.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389–3402.
2. **Benson DA, et al.** 1999. GenBank. Nucleic Acids Res. **27**:12–17.
3. **Brämer CO, Steinbuchel A.** 2001. The methylcitric acid pathway in Ralstonia eutropha: new genes identified involved in propionate metabolism. Microbiology **147**:2203–2214.
4. **Brock M, Maerker C, Schutz A, Volker U, Buckel W.** 2002. Oxidation of propionate to pyruvate in Escherichia coli. Involvement of methylcitrate dehydratase and aconitase. Eur. J. Biochem. **269**:6184–6194.
5. **Cai H, Clarke S.** 1999. A novel methyltransferase catalyzes the methyl esterification of trans-aconitate in Escherichia coli. J. Biol. Chem. **274**:13470–13479.
6. **Chen AM, et al.** 2010. Identification of a TRAP transporter for malonate transport and its expression regulated by GtrA from Sinorhizobium meliloti. Res. Microbiol. **16**:556–564.
7. **Crooks GE, Hon G, Chandonia JM, Brenner SE.** 2004. WebLogo: a sequence logo generator. Genome Res. **14**:1188–1190.
8. **Edgar RC.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5**:113.
9. **Felsenstein J.** 1996. Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. Methods Enzymol. **266**:418–427.
10. **Forward JA, Behrendt MC, Wyborn NR, Cross R, Kelly DJ.** 1997. TRAP transporters: a new family of periplasmic solute transport systems encoded by the *dctPQM* genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. J. Bacteriol. **179**:5482–5493.
11. **Hubbard PA, et al.** 2007. Crystal structure and mutagenesis of the metallochaperone MeaB. Insight into the causes of methylmalonic aciduria. J. Biol. Chem. **282**:31308–31316.
12. **Kelly DJ, Thomas GH.** 2001. The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. FEMS Microbiol. Rev. **25**:405–424.
13. **Kim YS.** 2002. Malonate metabolism: biochemistry, molecular biology, physiology, and industrial application. J. Biochem. Mol. Biol. **35**:443–451.
14. **Koo JH, Cho IH, Kim YS.** 2000. The malonate decarboxylase operon of *Acinetobacter calcoaceticus* KCCM 40902 is regulated by malonate and the transcriptional repressor MdcY. J. Bacteriol. **182**:6382–6390.
15. **Koo JH, Kim YS.** 1999. Functional evaluation of the genes involved in

malonate decarboxylation by Acinetobacter calcoaceticus. Eur. J. Biochem. **266**:683–690.
16. **Lee HY, An JH, Kim YS.** 2000. Identification and characterization of a novel transcriptional regulator, MatR, for malonate metabolism in Rhizobium leguminosarum bv. Trifolii. Eur. J. Biochem. **267**:7224–7229.
17. **Meister M, Saum S, Alber BE, Fuchs G.** 2005. L-Malyl-coenzyme A/beta-methylmalyl-coenzyme A lyase is involved in acetate assimilation of the isocitrate lyase-negative bacterium *Rhodobacter capsulatus*. J. Bacteriol. **187**:1415–1425.
18. **Mironov AA, Vinokurova NP, Gel'fand MS.** 2000. Software for analyzing bacterial genomes. Mol. Biol. (Mosk.) **34**:253–262.
19. **Peng HL, Shiou SR, Chang HY.** 1999. Characterization of *mdcR*, a regulatory gene of the malonate catabolic system in *Klebsiella pneumoniae*. J. Bacteriol. **181**:2302–2306.
20. **Rigali S, Derouaux A, Giannotta F, Dusart J.** 2002. Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR, and YtrA subfamilies. J. Biol. Chem. **277**:12507–12515.
21. **Rodionov DA.** 2007. Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. Chem. Rev. **107**:3467–3497.
22. **Thomas GH, Southworth T, Leon-Kempis MR, Leech A, Kelly DJ.** 2006. Novel ligands for the extracellular solute receptors of two bacterial TRAP transporters. Microbiology **152**:187–198.