

Widespread bacterial utilization of guanidine as nitrogen source

Malte Sinn¹ | Franziskus Hauth^{1,2} | Felina Lenkeit^{1,2} | Zasha Weinberg³ | Jörg S. Hartig ^{1,2}

¹Department of Chemistry, University of Konstanz, Konstanz, Germany

²Konstanz Research School Chemical Biology (KoRS-CB), University of Konstanz, Konstanz, Germany

³Bioinformatics Group, Department of Computer Science and Interdisciplinary Centre for Bioinformatics, Leipzig University, Leipzig, Germany

Correspondence

Jörg S. Hartig, Department of Chemistry, University of Konstanz, Universitätsstraße 10, 78457 Konstanz, Germany.
Email: joerg.hartig@uni-konstanz.de

Funding information

H2020 European Research Council, Grant/Award Number: ERC CoG 681777 'RiboDisc'; Deutsche Forschungsgemeinschaft, Grant/Award Number: WE6322/1-1

Abstract

Guanidine is sensed by at least four different classes of riboswitches that are widespread in bacteria. However, only very few insights into physiological roles of guanidine exist. Genes predominantly regulated by guanidine riboswitches are Gdx transporters exporting the compound from the bacterial cell. In addition, urea/guanidine carboxylases and associated hydrolases and ABC transporters are often found combined in guanidine-inducible operons. We noted that the associated ABC transporters are configured to function as importers, challenging the current view that riboswitches solely control the detoxification of guanidine in bacteria. We demonstrate that the carboxylase pathway enables utilization of guanidine as sole nitrogen source. We isolated three enterobacteria (*Raoultella terrigena*, *Klebsiella michiganensis*, and *Erwinia rhapontici*) that utilize guanidine efficiently as N-source. Proteome analyses show that the expression of a carboxylase, associated hydrolases and transport genes is strongly induced by guanidine. Finding two urea/guanidine carboxylase enzymes in *E. rhapontici*, we demonstrate that the riboswitch-controlled carboxylase displays specificity toward guanidine, whereas the other enzyme prefers urea. We characterize the distribution of riboswitch-associated carboxylases and Gdx exporters in bacterial habitats by analyzing available metagenome data. The findings represent a paradigm shift from riboswitch-controlled detoxification of guanidine to the uptake and assimilation of this enigmatic nitrogen-rich compound.

KEYWORDS

ABC transporter, bacterial metabolism, guanidine, guanidinium, nitrogen source, regulation of gene expression, riboswitch, urea carboxylase, urease

1 | INTRODUCTION

Although guanidine is widely used in the manufacturing of plastics, as reactant in certain explosives and propellants and as a chaotropic

reagent in biochemistry, a prominent role in nature has not been identified to date. The lack of knowledge about the physiological role and source of free guanidine in biological systems is in bright contrast to the widespread occurrence of guanidine riboswitches. There is a plethora of guanidine-containing compounds ranging from

Malte Sinn and Franziskus Hauth contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Molecular Microbiology* published by John Wiley & Sons Ltd.

arginine and creatinine to guanidine and large secondary metabolites such as streptomycin. Although all these guanidine compounds comprise a potential source of guanidine, the known catabolic routes proceed mostly by hydrolytic attack of the guanidine carbon atom, giving rise to urea. Despite some early reports about the biological formation of guanidine and its utilization (Kihara et al., 1955; Natelson & Sherwin, 1979), only few biotic reactions have been characterized that produces or breaks down guanidine to date. The best studied enzyme that catalyzes the production of guanidine is the ethylene-forming enzyme (EFE). Here, guanidine is formed via the δ -hydroxylation of arginine and subsequent loss of guanidine (Fukuda et al., 1992; Hausinger, 2004).

The idea that guanidine might have a prominent but so-far overlooked role in natural physiology or metabolism was sparked by a series of publications from Breaker and coworkers who described three classes of riboswitches that respond to guanidine (Nelson et al., 2017; Sherlock & Breaker, 2017; Sherlock et al., 2017). The recent discovery of a fourth class of guanidine riboswitches further supports this idea (Lenkeit et al., 2020; Salvail et al., 2020). The *ykkC* motif RNA was published as a riboswitch candidate in 2004 (Barrick et al., 2004; Battaglia et al., 2017; Reiss et al., 2017), and is common in various bacterial clades and associated with genes encoding transporters like multidrug efflux pumps, urea carboxylases, purine biosynthesis, and amino acid metabolism enzymes. Two additional riboswitch candidates, the mini-*ykkC* (Huang et al., 2017a; Weinberg et al., 2007) and the *ykkC*-III RNA motif (Huang et al., 2017b; Weinberg et al., 2017) were subsequently identified to be found upstream of several of these genes. Although the three motifs do not share sequence or structural characteristics in the ligand-binding domain, it was hypothesized that they sense the same ligand, based on the extensive overlap of the genetic context (Meyer et al., 2011; Sherlock & Breaker, 2020). The wide variety of associated genes hampered the search of the ligand based on the genetic context, but also indicated that its ligand participates in widespread metabolic reactions (Barrick et al., 2004; Meyer et al., 2011). All three *ykkC* motifs, now renamed as guanidine-I, -II, and -III riboswitches (Nelson et al., 2017; Sherlock & Breaker, 2017; Sherlock et al., 2017), and the recently published guanidine-IV (Lenkeit et al., 2020; Salvail et al., 2020) riboswitch were verified to respond selectively to guanidine.

A gene commonly controlled by guanidine riboswitches encodes certain representatives of "small multidrug resistance" (SMR) transporters such as *YkkCD*, *EmrE*, and *SugE*. Subsequent to the discovery of guanidine riboswitches, this specific family of SMR proteins has been demonstrated to function as selective guanidine exporters, termed *Gdx* (for guanidine exporters) (Kermani et al., 2018), see Figure 1a. When *SugE*-type genes occur under guanidine riboswitch control, we will refer to these as *Gdx* throughout the manuscript. Another gene product that is frequently controlled by guanidine riboswitches is annotated as urea carboxylase (*Uca*), see Figure 1a. It has been demonstrated already during the initial discovery of guanidine-I riboswitches that these carboxylases can use both urea and guanidine as substrates (Nelson et al., 2017). The

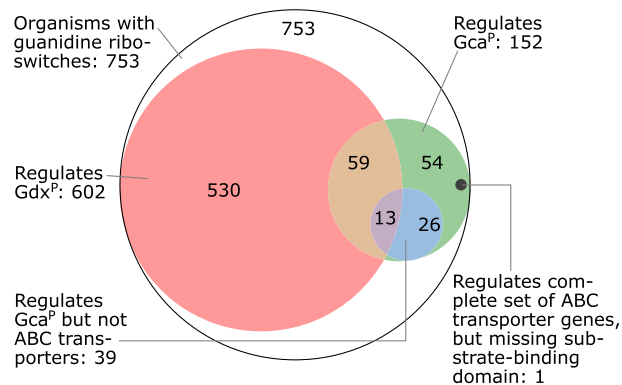


FIGURE 1 Association of guanidine-I, -II, -III, and -IV riboswitch classes with specific gene functions. The largest circle represents all 753 organisms with complete genomes that contain at least one predicted guanidine riboswitch. The sets indicate which organisms have guanidine riboswitches that regulate guanidine exporters *Gdx* (602 organisms) or guanidine carboxylase *Gca* genes (152 organisms), and also considers the regulation of ABC transporter genes by guanidine riboswitches. A total of 530 organisms only regulate *Gdx* by guanidine, 80 organisms only regulate *Gca* in this way, and 72 organisms do both. Of 113 riboswitch-regulated operons encoding ABC transporters, 112 contain a substrate-binding domain [Colour figure can be viewed at wileyonlinelibrary.com]

studied riboswitch-controlled carboxylase showed a 40-fold lower K_M for guanidine compared to urea but a comparable k_{cat} . Thus, the riboswitch-associated carboxylase annotated as urea carboxylase prefers guanidine as a substrate. We will refer to carboxylase enzymes that occur under guanidine riboswitch control as guanidine carboxylases (*Gca*). It was speculated that the carboxylation reaction initiates the degradation of guanidine for the purpose of detoxification (Nelson et al., 2017). When urea is carboxylated, the resulting product allophanate is further hydrolyzed by the enzyme allophanate hydrolase (Kanamori et al., 2004). Indeed, many guanidine riboswitch-regulated operons contain genes annotated as allophanate hydrolases. However, guanidine carboxylation results in carboxyguanidine, which might require different hydrolysis activities compared to allophanate. Two further genes (annotated as urea carboxylase-associated genes 1 and 2; *uca* 1 and 2) often associate with guanidine carboxylase enzymes. A recent report clarified that these two proteins comprise the subunits of a heterodimeric carboxyguanidine deiminase (*CgdAB*) enabling its hydrolysis to allophanate and ammonia (Schneider et al., 2020). The work also addresses the question of substrate specificity of the associated carboxylases and speculates about its involvement of the use of guanidine as nitrogen source.

Taken together, guanidine riboswitches predominantly induce *Gdx* transporters in order to export this compound from bacterial cells before it reaches problematic concentrations. However, other gene functions controlled by guanidine riboswitches enable the carboxylation and subsequent degradation. We show that this carboxylase pathway enables the utilization of guanidine as sole nitrogen source. We describe the isolation and characterization of guanidine-assimilating bacteria and demonstrate that carboxylase enzymes

have evolved that display selectivity for either urea or guanidine. In addition, we analyze metagenomics data in order to characterize habitats that are enriched for carboxylase pathway enzymes for the utilization of guanidine as nutrient.

2 | RESULTS

2.1 | Gene functions under control of guanidine riboswitches

Since the discovery of widespread riboswitches that induce gene expression in response to the presence of guanidine, the physiology of guanidine has remained a mystery. In order to shed more light on the physiology of guanidine in bacteria, we sought to investigate genes that are commonly associated with the four known riboswitch classes for guanidine. For a comprehensive analysis, we combined the information of the genetic context for all four known riboswitch classes in the available genomic data. Since certain gene functions under guanidine riboswitch control seem to occur in conserved operons that contain groups of highly associated genes, we grouped the gene categories into different pathways: The gene classes most frequently controlled by guanidine riboswitches are Gdx exporters that have been demonstrated experimentally to export guanidine, see Figure 1 and Supporting Information File 1. They occur as individual genes or tandem arrangements (reminiscent of their homo- or heterodimeric topology) and appear only sometimes associated with other genes. The second-largest group of genes associated with guanidine riboswitches comprise a pathway (Gca^P) that consists in most cases of ATP-binding cassette-type (ABC) transporters, accompanied by genes that encode a urea carboxylase, two different urea carboxylase-associated genes, and allophanate hydrolases. As it has been demonstrated for a representative before (Nelson et al., 2017) and detailed in the introduction, it is likely that most of the riboswitch-controlled carboxylase enzymes are in fact guanidine carboxylases. In addition, it could be speculated that the associated ABC transporters are specific for guanidine transport.

2.2 | Detoxification versus utilization

Subsequent to the discovery of widespread occurrence of guanidine riboswitches in bacteria, it has been believed that the controlled genes encode activities that facilitate either the export or the degradation of guanidine by carboxylation and subsequent hydrolysis for the purpose of detoxification (Battaglia & Ke, 2018; Nelson et al., 2017). However, we noted that organisms that regulate urea/guanidine carboxylase-containing pathways by guanidine riboswitches often (in 113 of 152 organisms, 74%) also regulate ABC-type transporters in this way (Figure 1). This suggests that the ABC-type transporters, which are not structurally related to Gdx exporters, are somehow functionally connected to the carboxylase pathway gene. If the associated ABC-type transporters are exporters, this

would mean that the guanidine-controlled carboxylase operons contain two different means to detoxify this compound, export as well as modification via carboxylation and subsequent hydrolysis. Moreover, in 52.2% of the organisms with carboxylase pathway-associated ABC-type transporters, a Gdx-type exporter is also found in the same genome (59 from 113 occurrences, see Figure 1). Hence, these organisms would encode two different transporters for the export of guanidine. Such a high frequency of redundancy would be surprising. Instead, the ABC transporter could work together with the carboxylase and imports its substrate guanidine.

In order to clarify the role of the ABC transporters associated with the carboxylase pathway, we analyzed the encoded genes more carefully. ABC transport systems can either import or export compounds (Locher, 2016; Wilkens, 2015). In bacteria, these two activities can be distinguished easily by the composition of the subunits: Importers are characterized by the presence of a periplasmic substrate-binding protein (in Gram-negative) or a lipid-anchored external protein (Gram-positive) that delivers the transported substrate to the outer side of the ABC-type channel (Berntsson et al., 2010; Maqbool et al., 2015). In the case of exporters, this domain is not necessary and is lacking. We further analyzed the guanidine riboswitch-controlled carboxylase operon-associated ABC transporters and found periplasmic substrate-binding domains in 112 of the 113 cases, see Figure 1. This finding implies that these carboxylase pathway-associated ATP transporters function rather as importers instead of detoxifying guanidine via export. The prospect of active, ATP hydrolysis-driven import into bacteria opens up the possibility that guanidine is utilized as a nutrient.

2.3 | Enrichment of guanidine utilizers

In order to investigate whether guanidine can serve as nutrient for bacterial growth, we sought to enrich microorganisms that are able to utilize this N-rich compound as nitrogen source. We collected water samples from the lake shore surface sediment of Lake Constance (Bodensee) and plated filtered and diluted samples on minimal media that contained glycerol as carbon source and guanidine as sole nitrogen source. With undiluted lake water samples, densely overgrown plates were observed at 25°C over 24 hr. In diluted samples, many colonies could be identified. From these colonies, three morphologically different ones were chosen for enrichment and subsequent characterization. Individual colonies were passaged several times on selective media in order to obtain homogenous strains. By 16S rRNA sequencing the three isolated bacteria were identified as strains of *Raoultella terrigena* (Schicklberger et al., 2015), *Erwinia rhapontici* (Huang et al., 2003), and *Klebsiella michiganensis* (Saha et al., 2013).

In order to characterize the utilization of guanidine as N-source, the three isolated strains were cultivated in liquid culture and growth was monitored. As comparison, the bacteria were grown on minimal medium with glycerol as C-source and supplemented guanidinium chloride (5 mM), urea (10 mM), or ammonium chloride (15 mM) as sole nitrogen source. All strains were able to grow efficiently on guanidine

(Figure 2). The bacteria grew slightly slower on guanidine compared to ammonia. Nevertheless, they reached the same maximal optical density indicating that all nitrogen atoms of guanidine can be assimilated. Interestingly, only *E. rhapontici* was able to grow on urea as N-source. This result is somewhat unexpected as the isolated *R. terrigena* and *K. michiganensis* strains both encode urease genes, see below. *E. rhapontici*, which is able to utilize urea efficiently, does not encode a urease, but rather two copies of annotated urea carboxylases, see below for a detailed characterization of the two enzymes. We speculated that the difference in growth on urea could be related to the difference of the two means of urea degradation. Urease is a Ni-dependent metallohydrolase (Boer et al., 2014). When we repeated the growth experiments with supplemented Ni²⁺, both *R. terrigena* and *K. michiganensis* grew to the same optical density (Figure S1). The inability of *R. terrigena* and *K. michiganensis* to utilize urea in absence of supplemented Ni demonstrates that the encoded riboswitch-controlled guanidine carboxylases (see next paragraph) do not also allow for the utilization of urea but seem to be specific for guanidine.

In order to investigate the molecular basis for the observed utilization of guanidine, the genomes of the three strains were sequenced. The analysis of the resulting genome sequences *R. terrigena* JH1 (NZ_CP050508.1), *E. rhapontici* JH2 (JABANZ000000000.1), and *K. michiganensis* JH7 (JABANY000000000.1) utilizing the Type

(Strain) Genome Server (Meier-Kolthoff & Göker, 2019) confirmed the taxonomic classification based on the 16S RNA analysis. We next analyzed the occurrence of guanidine riboswitches and associated genes in the obtained genomes, see Table 1. All three strains contain a guanidine carboxylase operon under the control of a guanidine riboswitch. The organization of the genes is very similar among the bacterial strains with the ABC transporter genes and the carboxyguanidine deiminase genes (*cgdAB*) under control of a guanidine-I riboswitch followed by guanidine carboxylase (*gca*) and allophanate hydrolase genes (*atzF*) under the control of a guanidine-II riboswitch (see Figure 3). In *E. rhapontici* str. JH02 *atzF* is missing downstream of the riboswitch-associated *gca*. However, *atzF* is found at a different locus in conjunction with a second urea/guanidine carboxylase gene cluster. Both *R. terrigena* and *K. michiganensis* genomes contain a Gdx-type exporter under guanidine riboswitch control, whereas this activity is lacking in *E. rhapontici*.

2.4 | Analysis of guanidine-dependent gene expression

In order to investigate whether the observed utilization of guanidine as N-source is based on the activity of the riboswitch-controlled

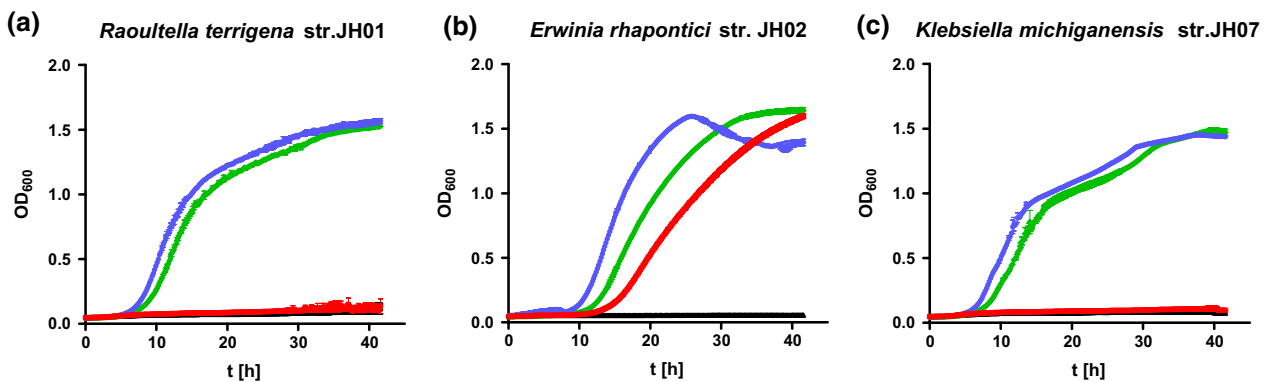


FIGURE 2 Bacterial utilization of guanidine as nitrogen source. Growth of isolated bacterial strains in minimal medium containing 5 mM guanidine (green), 10 mM urea (red), 15 mM ammonia (blue), and no nitrogen source (black). Values are means of biological triplicates with SD [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Guanidine riboswitches in isolate genomes

Organism	Genome accession	Riboswitch	Coordinates ^a	Regulated genes ^b
<i>Raoultella terrigena</i> str. JH1	CP050508.1	Guanidine-I	2977411–2977313 (-)	Abc-transporter ^c , <i>cgdA</i> , <i>cgdB</i>
		Guanidine-II	2973102–2973053 (-)	<i>gca</i> , <i>atzF</i>
		Guanidine-II	5159112–5159068 (-)	<i>gdx</i>
<i>Erwinia rhapontici</i> str. JH2	JABANZ000000000.1	Guanidine-I	402759–402857 (+)	Abc-transporter ^c , <i>cgdA</i> , <i>cgdB</i>
		Guanidine-II	406997–407043 (+)	<i>gca</i>
<i>Klebsiella michiganensis</i> str. JH7	JABANY010000008.1	Guanidine-I	84523–84621 (+)	Abc-transporter ^c , <i>cgdA</i> , <i>cgdB</i>
		Guanidine-II	88830–88878 (+)	<i>gca</i> , <i>atzF</i>
		Guanidine-II	205465–205421 (-)	<i>gdx</i>

^a(+) plus strand; (-) minus strand.

^b*gca*, guanidine carboxylase; *cgdA* and *cgdB*, carboxyguanidine deiminases A and B; *atzF*: allophanate hydrolase; *gdx*: guanidine exporter.

^cThree subunits present: Periplasmic substrate-binding, Transmembrane, and ATP-binding domains.

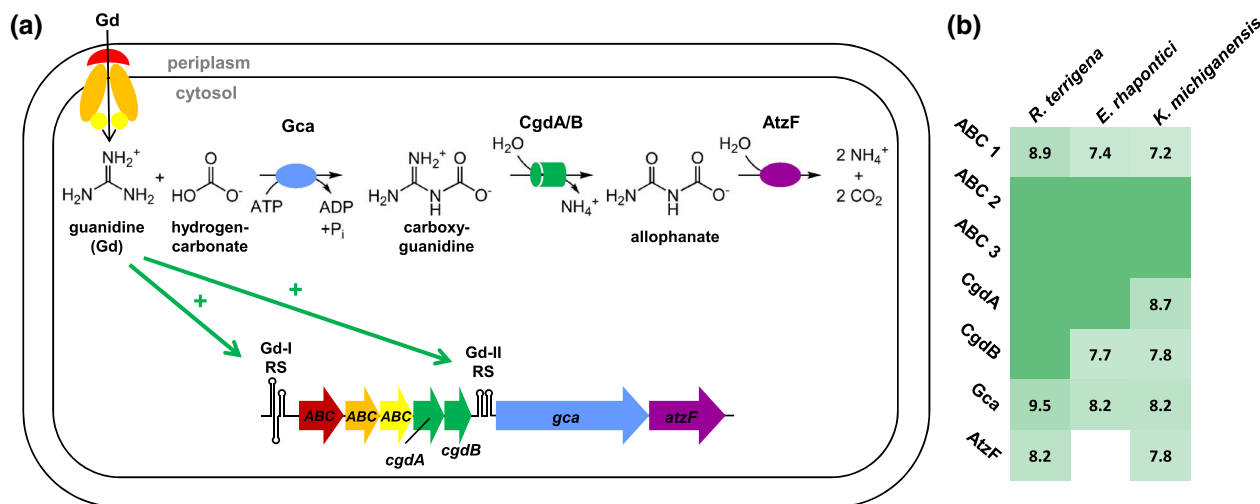


FIGURE 3 (a) Scheme of the investigated guanidine assimilation pathway and its genetic organization for the three isolates. Note that as an exception *atzF* of *Erwinia rhapontici* str. JH02 is encoded at a different genetic locus. Guanidine induces (green arrows) gene expression by binding to the riboswitches (Gd-I and Gd-II RS). (b) Proteome data showing the induction of the guanidine assimilation operons during growth on guanidine compared to growth on ammonia. Changes in the protein abundance are expressed as log₂ values. Dark green indicates that proteins were identified only when bacteria were grown on guanidine. ABC 1, periplasmic substrate-binding domain; ABC 2, transmembrane domain; ABC 3, ATP-binding cassette [Colour figure can be viewed at wileyonlinelibrary.com]

gene products, we analyzed the proteome of the three isolated strains in response to guanidine. Each of the isolated bacteria was grown in minimal medium that contained 1% of glycerol as carbon source and either 5 mM guanidinium chloride or 15 mM ammonium chloride as nitrogen source. Bacteria were harvested in late exponential phase and the whole proteome was determined by mass spectrometry. In all three bacteria the genes under the control of guanidine riboswitches (Table 1) were highly upregulated, see Figure 3 and Supporting Information File 2. We noticed that nitrogen metabolism-related genes are generally enhanced when bacteria utilize guanidine, see Supporting Information File 2. We speculate that this is due to the lack of a preferred nitrogen source.

2.5 | Characterization of carboxylase enzymes

When we analyzed the genome and proteome data of *E. rhapontici*, we noticed that this organism encodes two different urea/guanidine carboxylase genes. The first carboxylase (WP_171149239.1) is not riboswitch-controlled, whereas the second (WP_171148480.1) is under control of a guanidine riboswitch, see Table 1. As both were upregulated upon growth on guanidine we were interested to understand their roles in guanidine utilization. Both urea/guanidine carboxylases were overexpressed as His-tagged versions in *E. coli* and purified via Ni-NTA in order to determine their substrate preferences. The carboxylation reaction consumes ATP stoichiometrically with regard to substrate turnover (Fan et al., 2012). We monitored the reaction as described before (Kanamori et al., 2004; Roon et al., 1972; Roon & Levenberg, 1972) by coupling the ATP-consuming carboxylation reaction to the reactions of pyruvate kinase and lactate dehydrogenase, which result in the oxidation of NADH. The decrease

of NADH can be monitored spectrophotometrically. By plotting the initial velocity over the substrate concentration the kinetic parameters K_M and k_{cat} could be obtained. The data for both carboxylase enzymes fitted well to Michaelis–Menten kinetics (Figure 4). The calculated parameters are shown in Table 2. The riboswitch-associated carboxylase (*Gca*, WP_171148480.1) prefers guanidine over urea with K_M values of 0.058 ± 0.005 mM and 5.3 ± 0.3 mM, respectively. k_{cat} for both substrates is almost the same, 5.9 ± 0.1 s⁻¹ for guanidine and 5.3 ± 0.3 s⁻¹ for urea. Interestingly, the non-riboswitch-associated carboxylase (*Uca*, WP_171149239.1) prefers urea ($K_M = 0.085 \pm 0.009$ mM) over guanidine ($K_M = 17 \pm 4$ mM) as substrate and k_{cat} is much lower compared to the guanidine carboxylase (with 0.13 ± 0.01 s⁻¹ for guanidine and 0.25 ± 0.01 s⁻¹ for urea). Thus, the substrate preference of the enzymes is interchanged. The urea carboxylase has a 375-fold higher specificity constant (k_{cat}/K_M) for urea compared to a 90-fold higher specificity constant of guanidine carboxylase for guanidine. As saturation was not reached for the respective poorer substrate, kinetic parameters for those substrates should be taken with care. However, saturation was almost reached as visualized by plotting the data on a linear scale (Figure S2). Since *E. rhapontici* encodes no urease enzyme, the non-riboswitch-associated urea carboxylase (WP_171149239.1) is likely responsible for the observed Ni-independent utilization of urea as N-source (Figure 2).

2.6 | Distribution of urea and guanidine carboxylase enzymes

Previously, it has been noticed that in the substrate-binding pocket of the then described urea carboxylase of *O. sagarensis* an aspartic acid

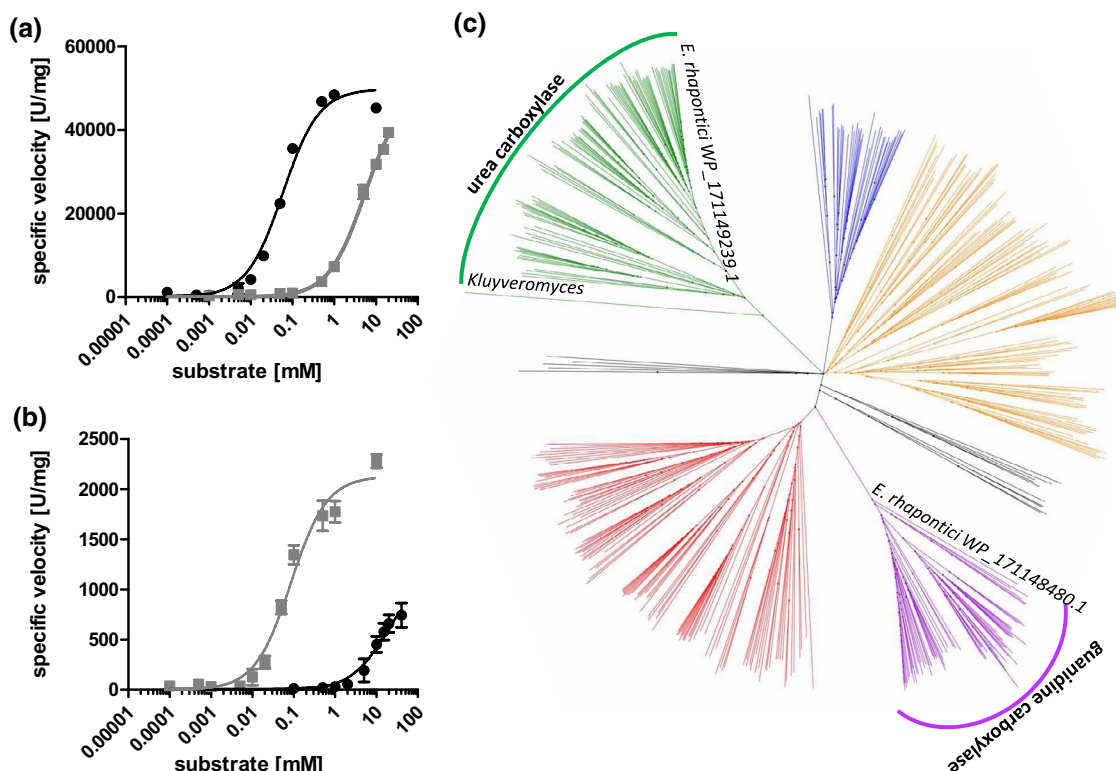


FIGURE 4 Substrate specificity of the two urea/guanidine carboxylase enzymes encoded by *E. rhapontici* and their phylogenetic relationship. (a) Characterization of the riboswitch-associated carboxylase WP_171148480.1 and (b) non-riboswitch-associated carboxylase WP_171149239.1 of *E. rhapontici* JH02. The initial velocity (1 U equals $\mu\text{moles NAD}^+$ generated per min) per mg of enzyme is plotted against substrate concentrations for guanidine (black dots) and urea (gray squares). Error bars represent SD from triplicates. (c) Neighbor-joining phylogeny of urea/guanidine carboxylases and their homologs. Homology search and multiple sequence alignment were performed with Consurf (Landau et al., 2005). The sequences of the urea carboxylase and guanidine carboxylase were aligned subsequently using Jalview (Waterhouse et al., 2009). Neighbor-joining phylogeny was performed with BLOSUM62 (Henikoff & Henikoff, 1992). The tree was illustrated with iTOL (Letunic & Bork, 2019) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Michaelis–Menten kinetic parameters

Enzyme/substrate	$k_{\text{cat}} [\text{s}^{-1}]$		$K_M [\text{mM}]$		$k_{\text{cat}}/K_M [\text{s}^{-1}\text{mM}^{-1}]$	
	guanidine	urea	guanidine	urea	guanidine	urea
urea carboxylase WP_171149239.1	0.13 ± 0.01	0.25 ± 0.01	17 ± 4	0.085 ± 0.009	0.008 ± 0.002	3.0 ± 0.2
guanidine carboxylase WP_171148480.1	5.9 ± 0.1	5.3 ± 0.3	0.058 ± 0.005	5.3 ± 0.3	100 ± 4	1.1 ± 0.04

Note: Kinetic parameters were obtained from a Michaelis–Menten fit to data in Figure 4.

contacts the substrate (Kanamori et al., 2004; Nelson et al., 2017). Already Nelson et al. hypothesized that the carboxylate in the side-chain of the aspartate is preferably binding the positively charged guanidinium rather than being protonated and binding urea, explaining the lower K_M of guanidinium. Interestingly, we noticed that another aspartate on the opposite site of the binding pocket contacting the biotin cofactor is replaced by asparagine in the urea carboxylase of *Kluyveromyces lactis* as it is also the case in the non-riboswitch-associated urea carboxylase (WP_171148480.1) of *E. rhapontici* studied above. We speculated whether this aspartate (Asn1330 of the *K. lactis* structure (Schneider et al., 2020) could be indicative for the substrate specificity of urea carboxylases. We performed a

homology search starting from *K. lactis* urea carboxylase followed by a multiple sequence alignment using the Consurf platform (Landau et al., 2005). We subsequently aligned the urea and guanidine carboxylase sequences from our strains and performed a neighbor-joining using BLOSUM62 (Henikoff & Henikoff, 1992). The homolog sequences clustered in five major and three minor clades (see Figure 4c). Remarkably, *K. lactis* and the non-riboswitch-associated urea carboxylase (WP_171149239.1) clustered in one clade (green), whereas the two studied riboswitch-associated guanidine carboxylases clustered in another clade (purple). All other clades comprise families of yet uncharacterized carboxylases. Hence, we conclude that urea/guanidine carboxylases can be divided in at least two

specific enzyme clades: urea carboxylases and guanidine carboxylases. Interestingly, the asparagine residue is only found in the binding pocket of the urea carboxylase (green) clade, whereas all other clades in Figure 4c comprise aspartate at that position. If the occurrence of the aspartate residue in the binding pocket is indeed indicative for guanidine specificity, it seems that guanidine carboxylation is the much more widespread activity of this class of enzymes compared to urea carboxylation. However, more representatives from other clades need to be tested in order to support such a conclusion.

2.7 | Distribution of guanidine-utilizing carboxylases in metagenomes/habitats

So far, we have demonstrated that the guanidine-controlled operons encoding ABC-type transporters and carboxylases, carboxyguanidine deiminases, and allophanate hydrolases enable the uptake and assimilation of guanidine. This result is contrasted by widespread occurrence of riboswitch-controlled Gdx-type exporters of guanidine. In order to shed more light on the physiology of guanidine utilization in nature, we aimed at investigating the occurrence of both pathways in certain bacterial habitats. When the occurrence of guanidine riboswitches in known organisms is taken into account, one notes that the switches are found widely distributed in many phyla of bacteria. An analysis is complicated by the fact that many bacteria are ubiquitously distributed and it is not always possible to determine whether a given bacterium has a predominantly water-, soil-, plant-, or animal-associated life style. In order to nevertheless extract information about the habitat where guanidine might play a pronounced role, we envisioned that metagenome data could be very helpful in order to connect the occurrence of a riboswitch-controlled activity to a given environment. Since for these data sets there is always a more or less specific sampling, and therefore, the isolated sequences are likely to be typical of the given habitat, we analyzed riboswitches that occur in metagenome data and correlated the frequency of occurrence of Gdx-type exporters and guanidine carboxylase-type utilization pathways (Gca^P) to the annotated habitat, see Figure 5 and Supporting Information File 3.

We selected some representative habitats such as human skin and gut metagenome data as well as environmental samples such as soil and aqueous habitats. We then identified Gdx-type exporters and Gca^P (related) genes under control of guanidine riboswitches. Next, we calculated a frequency of the occurrence of Gca^P in the metagenome data in relation to the sum of Gca^P and Gdx^P . By doing so, we wanted to gain insight into whether riboswitch-controlled genes in a given habitat prefer guanidine utilization via Gca^P or Gdx^P -mediated export of guanidine, in comparison to other environments. In Figure 5, the X-axis value is zero if all riboswitches in the environment regulate Gdx-type exporter genes, and one if all riboswitches regulate Gca^P -related genes. Environments are sorted from those that relatively favor guanidine carboxylase-mediated utilization (freshwater) to those that relatively favor Gdx^P -mediated export (human skin). Interestingly, it seems that the occurrence of the guanidine-utilizing carboxylase pathway negatively correlates with the nutrient- and nitrogen-richness of the respective habitat. Animal-based microbiomes thrive under nutrient-rich conditions, whereas aquatic habitats are often nitrogen-scarce environments (Young et al., 2016). We identified riboswitch-controlled Gca^P activity to be more prevalent in N-scarce environments such as fresh water, marine, and soil samples, whereas in N-rich habitats such as the human gut guanidine carboxylase pathway genes are found less frequently in comparison to Gdx exporters.

3 | DISCUSSION

Here, we show that guanidine is utilized by bacteria using riboswitch-controlled carboxylases and hydrolases. ABC-type importers are also often encoded under control of guanidine riboswitches, likely facilitating the efficient uptake of the N-rich compound. Given that the guanidine carboxylases are found widespread in bacteria, it seems likely that the utilization of guanidine as N-source is a common activity in many organisms. It seems that a single amino acid in the active site is responsible for determining the substrate specificity of the carboxylase reaction. This finding has also been reported recently when for the first time the role

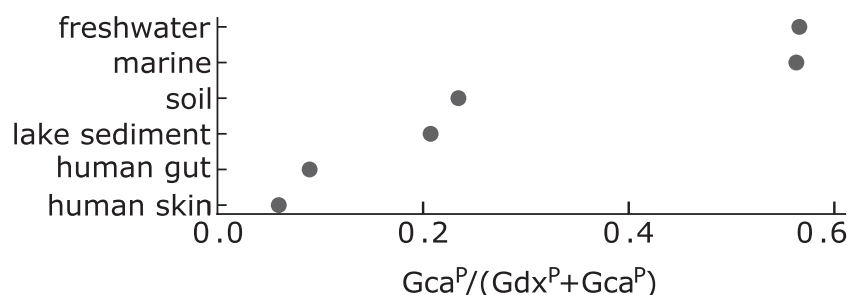


FIGURE 5 Occurrence of Gdx^P and Gca^P activities regulated by guanidine riboswitches in different environments. Gca^P : the number of occurrences of genes assigned to the guanidine carboxylase pathway that are regulated by guanidine riboswitches in data sets from the given environment. Gdx^P : the number of *gdx* genes controlled by guanidine riboswitches. For the selected environments, $Gca^P + Gdx^P$ is at least 285

of the associated *uca* genes was clarified as carboxyguanidine deiminases (Schneider et al., 2020).

The ATP-dependent carboxylation and subsequent hydrolysis of urea for its utilization has been first described in yeasts and algae in 1968 (Roon & Levenberg, 1968). However, it took until 2004 for the same activity to be described in bacteria (Kanamori et al., 2004). It has been noted before that the urease-mediated and the urea carboxylase/allophanate hydrolase-mediated reactions are two apparently redundant means of degrading urea (Hausinger, 2004). The co-occurrence of bacterial urease and urea carboxylase in one organism led Hausinger to speculate that one of the enzymes might catalyze an alternative reaction. As we have demonstrated, it seems that the majority of organisms encode enzymes that should show greater specificity toward guanidine than urea carboxylation, nevertheless carboxylases with higher specificity toward urea do exist, such as the examples in *E. rhapontici* (this study) or in *S. cerevisiae* and *C. albicans* (Schneider et al., 2020). Vice versa, considering that the two additional bacteria isolated in this work (*K. michiganensis* and *R. terrigena*) are not able to grow efficiently on urea without supplemented Ni for urease activation, it seems that the guanidine carboxylases that these two organisms encode are so specific that they are not able to hydrolyze urea in sufficient amounts in order to sustain growth in absence of a urease activity.

With regard to a possible redundancy of urea degradation in organisms that contain both urease and urea carboxylase it might be advantageous for certain bacteria specialized on the utilization of such compounds to invest in the maintenance of genes encoding both systems. Although the urease-dependent direct hydrolysis of urea is more straightforward and seems more energy-economic than the ATP-dependent detour via the carboxylated intermediate, it requires Ni as cofactor that might not always be available in sufficient amounts. Such a scenario was observed in our experiments when the two strains *R. terrigena* and *K. michiganensis* did not grow in minimal media with urea as sole N-source unless Ni was supplemented. In addition, several additional proteins responsible for the modification of the active site, Ni²⁺ loading, and Ni homeostasis are necessary for the activation of urease (Farrugia et al., 2013).

Here, we have shown that guanidine utilization is widespread and predominantly found in bacterial organisms living in nutrient-scarce environments. The three isolated bacteria seem to have specialized on the utilization of alternative nitrogen sources since all of them possess activities for the assimilation of guanidine and urea. Additionally, two of the isolates (*R. terrigena* and *K. michiganensis*) possess genes necessary for N₂ fixation, a rather rare feature among enterobacteria. Guanidine utilization is carried out via carboxylation and subsequent hydrolysis. The carboxylase enzymes appear to be specific for guanidine, although urea-specific homologs also exist, sometimes even in the same organism as we have found for *E. rhapontici*. Interestingly, *R. terrigena* and *K. michiganensis* also contain Gdx-type guanidine exporters. Similar to the Gca pathway enzymes Gdx is also under control of a guanidine-dependent on-riboswitch. We speculate that at low nitrogen concentrations guanidine is utilized via the Gca activities

and at high nitrogen and guanidine concentrations the Gdx exporter is getting rid of excess guanidine. Such a scenario could be facilitated by placing the expression of the Gca pathway under control of a nitrogen limitation-responsive control mechanism. We see some evidence in the proteome data where nitrogen limitation activities are upregulated in general when guanidine is offered as sole nitrogen source. We have further presented a bioinformatics method in order to assign a given biochemical activity to certain habitats by surveying the occurrence and frequency of certain genes in metagenome data from specific habitats. However, the presented findings pose again the intriguing question of the source of guanidine in nature. Given the widespread occurrence of guanidine-sensing riboswitches as well as guanidine-transporting and metabolizing activities, it seems very likely that so-far overlooked, widespread guanidine-producing abiotic or biotic reactions exist in nature.

4 | METHODS

4.1 | Media

Minimal medium contained 8.5 g/L Na₂HPO₄·2H₂O, 3 g/L KH₂PO₄, 0.5 g/L NaCl, 1 g/L NH₄Cl, 2 mM MgCl₂, 100 μM CaCl₂ and was supplemented with trace elements (0.1 mM EDTA, 0.03 mM FeCl₃, 6.2 μM ZnCl₂, 0.76 μM CuCl₂, 0.42 μM CoCl₂, 1.62 μM H₃BO₃; 0.08 μM MnCl₂) and vitamins (0.1 mg/L cyanocobalamin, 0.08 mg/L 4-aminobenzoic acid, 0.02 mg/L D-(+)-biotin, 0.2 mg/L niacin, 0.1 mg/L Ca-D-(+)-pantothenic acid, 0.3 mg/L pyridoxamine hydrochloride, 0.2 mg/L thiamine hydrochloride). About 1% (v/v) glycerol was added as carbon source. As nitrogen source either guanidine (5 mM), urea (10 mM), or NH₄Cl (15 mM) were added. For the overexpression of urea and guanidine carboxylase cells were grown in LB supplemented with kanamycin as selection marker.

4.2 | Enrichment of guanidine-utilizing bacteria

Environmental sample was taken in late September 2019 from the lake shore sediment of the Lake of Constance (47°41'44.2''N 9°11'35.1''E). Sediment was rinsed with lake water and filtered. The sample was sequentially diluted to obtain single colonies for selection. Dilutions were streaked out on selective minimal media agar plates supplemented with 5 mM guanidine as the sole nitrogen source and 1% v/v glycerol as carbon source. Plates were incubated at RT for 96 hr. Single colonies were picked and transferred onto a new plate. Selection was repeated until homogenous colonies were obtained.

4.3 | Growth analysis

Isolated strains were grown in minimal medium with 5 mM guanidine, 10 mM urea, or 15 mM NH₄Cl as the sole carbon source, respectively.

Growth was monitored in a 96-well plate in biological triplicates. The medium was inoculated to $OD_{600} = 0.0005$. The growth medium was covered with M20 silicon oil to allow gas exchange but avoid evaporation. Plates were shaken at 200 rpm at 30°C in a TECAN reader and OD_{600} was measured every 10 min until stationary phase was reached.

4.4 | Proteome data

Cells were grown in minimal media with the respective nitrogen source at 30°C. Same amounts of cells were harvested after approximately 10 hr. Cells were lysed by sonification with a Branson Sonifier in 1x PBS. Total protein amount was determined with the BCA Kit from Thermo Scientific according to the manufacturer's protocol. A 50 µg total protein were send for proteome analysis.

4.5 | Construction and expression of urea and guanidine carboxylase

The full length gene of urea (*uca1*, WP_171149239.1) and guanidine (*uca2*, WP_171148480.1) carboxylase from *E. rhapsontici* was cloned into a pet28a vector bearing a N-terminal 6xHisTag, following standard Phusion PCR and Gibson assembly (New England Biolabs). *E. rhapsontici* template genomic DNA was extracted using a DNA easy kit (Qiagen), following the manufactures protocol for bacterial gDNA isolation. PCR primer sequences used can be found in Table S1. The resulting plasmids were transformed into electro competent *E. coli* XL10 cells via electroporation and grown over night on selective medium. Plasmids were extracted from single colonies (Zyppy Plasmid MiniPrep kit), verified by sequencing (GATC) and transformed into *E. coli* BL21 (DE3) cells.

4.6 | Enzyme expression

Starter cultures of single colony *E. coli* BL21 (DE3) cells harboring the pet28a-His6-*uca1* or *uca2* plasmid were grown overnight at 37°C, diluted 1:100 in lysogeny broth containing 50 µg/ml kanamycin and cultivated at 37°C, 200 rpm, to an OD_{600} of approximately 0.5. After induction with 0.5 mM isopropyl-β-D-thiogalactopyranoside the cultures were incubated at 18°C for approximately 16 hr. Cells were harvested by centrifugation and stored at -20°C. Active enzyme was purified by standard nickel metal affinity chromatography. In short, frozen pellet was resuspended in lysis buffer (20 mM Tris, 20 mM imidazole, 200 mM NaCl, EDTA-free protease inhibitor (cOmplete, Roche), 0.02 mg/ml lysozyme, pH 8.0) and left on ice for 30 min. Cell lysis was completed by two cycles of sonication (1.5 s on/off, 3 min, 20% amplitude) on ice, after which the lysate was centrifuged to remove cell debris and insoluble material. The supernatant fraction was filtered through a 0.2 µm filter and loaded onto high performance Ni-NTA-resin at a flow rate of 1 ml/min. The resin was washed with buffer A (20 mM Tris, 20 mM imidazole, 200 mM NaCl, pH 8.0) and the protein eluted with buffer B (20 mM Tris, 500 mM

imidazole, 200 mM NaCl, pH 8.0). The protein elution was buffer exchanged into buffer C (20 mM Tris, 200 mM NaCl, pH 8.0) using PD10-desalting columns (GE Healthcare), concentrated by Amicon centrifugal filters (Merck), snap frozen in liquid nitrogen and stored at -80°C upon further usage. SDS-page samples were taken from all steps to monitor enzyme purification.

4.7 | Enzymatic assay

Carboxylation activity (ATP cleavage) was measured by monitoring the coupled activity of pyruvate kinase and lactate dehydrogenase, as described before (Kanamori et al., 2004). The enzyme assay was performed in 20 mM Tris-HCl (pH8.0), 200 mM NaCl, 2mM DTT, containing Uca or Gca (FINAL conc.), 8 mM MgCl₂, 8 mM NaHCO₃, different concentrations of guanidine or urea, 1 mM phosphoenolpyruvate, 0.8 mM NADH, and 5 U/ml pyruvate kinase/lactate dehydrogenase enzyme mix (Sigma). The reaction was started by the addition of 1 mM ATP at 25°C and the decrease in NADH absorbance (340 nm) monitored in 96-well plates using a TECAN spark plate reader at a final assay volume of 50 µl. Technical triplicates were run for each substrate concentration and the slope of absorbance decrease over time was used to evaluate the rate of substrate carboxylation. Velocity was defined such as that 1 U of activity equalizes 1 µmol NAD⁺ formation per minute. Kinetic parameters were determined by fitting a plot of velocity versus substrate concentration using GraphPad Prism 6 assuming Michaelis-Menten kinetics.

4.8 | Phylogenetic analysis of urea and guanidine carboxylases

Homology search and multiple sequence alignment were performed with ConSurf (Landau et al., 2005) platform based on *K. lactis* crystal structure (PDB: 3VA7). Multiple sequence alignment was built using MAFFT on homologs collected from UNIREF90 database using BLAST/PSI-BLAST (three iterations, E-value 0.0001). Jalview (Waterhouse et al., 2009) was used for the alignment of the sequences of the guanidine carboxylases and urea carboxylase from our strains. Subsequently, a phylogenetic tree was generated based on neighbor joining with BLOSUM62 (Henikoff & Henikoff, 1992). The phylogenetic tree was illustrated with iTOL (Letunic & Bork, 2019).

4.9 | Analysis of guanidine operons in genomes and metagenomes

All complete bacterial genomes in version 87 of the RefSeq nucleotide database (NCBI Resource Coordinators, 2015) were analyzed. Complete genomes were defined as those whose accession begins with "NC_." The guanidine-I, -II, -III riboswitches were searched using the standard procedure in Rfam [cite <https://pubmed.ncbi>.

nlm.nih.gov/33211869]. For guanidine-I, -II, and -III, we used Rfam entries RF00442, RF1068, and RF01763, respectively. Guanidine-IV riboswitch locations were taken from our recent publication (Lenkeit et al., 2020). Guanidine-I riboswitches are highly similar to riboswitches with other ligand specificities (Sherlock & Breaker, 2020). To extract only guanidine-I riboswitches, we looked at the two nucleotides immediately following the conserved CAC sequence (Nelson et al., 2017). We accepted only sequences in which these nucleotides were GG. To reduce false positive riboswitch predictions, we also eliminated guanidine-II sequences unless they had the tetramer ACGR in both hairpins. We also enumerated guanidine-III nucleotides that were at least 97% conserved and not predicted to form a Watson-Crick base pair, and eliminated sequences that deviated from these conserved nucleotides.

If the distance between a riboswitch and the first downstream gene was at most 700 nucleotides away, and the gene is encoded in the same direction, we assumed that the gene was regulated by the riboswitch. Subsequent genes were presumed to be co-transcribed if they were also encoding in this strand, and were located no more than 500 nucleotides from the previous gene. These maximum distances are conservatively high to ensure all relevant genes would be found.

Genes were functionally classified based on conserved protein domains in version 32.0 of the Pfam database (Mistry et al., 2021). In addition to assigning Gdx-type exporters and Gca-type carboxylase pathways, a third activity (termed Agmat for Agmatinase-like proteins) was included in the analysis since it is also often controlled by guanidine riboswitches, see Supporting Information File 1 and Table S2. However, since its function is unclear and it is not connected to the Gdx and Gca activities, we have not further pursued the occurrence and function of the Agmat pathway. We manually defined a mapping between guanidine-associated gene functions and Pfam entries (Table S2: Functions-and-PFAM). We also defined a mapping between gene functions and pathways. Riboswitches were deemed to control a pathway when they appeared to regulate at least one gene function that is unambiguously associated with that pathway, that is, are not components of ABC transporters. Metagenomes were downloaded from various sources, predominantly IMG/M (Chen et al., 2019) and GenBank (NCBI Resource Coordinators, 2015), and we classified them into environmental categories based on available metadata. Due to inconsistencies in metadata, environmental categories were largely created manually. The locations of genes were predicted by MetaProdigal (Hyatt et al., 2012). In Figure 5, we counted the number of pathway-specific and riboswitch-regulated genes to arrive at numbers for the two pathways (Gdx exporter and Gca carboxylase). Otherwise, our annotations of riboswitches and genes in metagenomes was the same as with RefSeq.

ACKNOWLEDGMENT

JSH acknowledges funding from the ERC CoG "RiboDisc." ZW acknowledges funding from the DFG (WE6322/1-1). We thank Stephanie Gurses and Astrid Joachimi for excellent technical assistance. The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Genomes of *R. terrigena* JH1 (NZ_CP050508.1), *E. rhapsodici* JH2 (JABANZ000000000.1), and *K. michiganensis* JH7 (JABANY000000000.1) are available at NCBI. Proteome and bioinformatics data are available in the supplementary files of this article.

ORCID

Jörg S. Hartig  <https://orcid.org/0000-0001-6601-7217>

REFERENCES

- Barrick, J.E., Corbino, K.A., Winkler, W.C., Nahvi, A., Mandal, M., Collins, J. et al. (2004) New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 6421–6426. <https://doi.org/10.1073/pnas.0308014101>.
- Battaglia, R.A. & Ke, A. (2018) Guanidine-sensing riboswitches: How do they work and what do they regulate? Wiley interdisciplinary reviews. *RNA*, e1482.
- Battaglia, R.A., Price, I.R. & Ke, A. (2017) Structural basis for guanidine sensing by the ykkC family of riboswitches. *RNA*, 23, 578–585.
- Berntsson, R.-P.-A., Smits, S.H.J., Schmitt, L., Slotboom, D.-J. & Poolman, B. (2010) A structural classification of substrate-binding proteins. *FEBS Letters*, 584, 2606–2617. <https://doi.org/10.1016/j.febslet.2010.04.043>.
- Boer, J.L., Mulrooney, S.B. & Hausinger, R.P. (2014) Nickel-dependent metalloenzymes. *Archives of Biochemistry and Biophysics*, 544, 142–152. <https://doi.org/10.1016/j.abb.2013.09.002>.
- Chen, I.-M.-A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J. et al. (2019) IMG/M vol 5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Research*, 47, D666–D677.
- Fan, C., Chou, C.-Y., Tong, L. & Xiang, S. (2012) Crystal structure of urea carboxylase provides insights into the carboxyltransfer reaction. *The Journal of Biological Chemistry*, 287, 9389–9398. <https://doi.org/10.1074/jbc.M111.319475>.
- Farrugia, M.A., Macomber, L. & Hausinger, R.P. (2013) Biosynthesis of the urease metallocenter. *The Journal of Biological Chemistry*, 288, 13178–13185. <https://doi.org/10.1074/jbc.R112.446526>.
- Fukuda, H., Ogawa, T., Tazaki, M., Nagahama, K., Fujii, T., Tanase, S. et al. (1992) Two reactions are simultaneously catalyzed by a single enzyme: The arginine-dependent simultaneous formation of two products, ethylene and succinate, from 2-oxoglutarate by an enzyme from *Pseudomonas syringae*. *Biochemical and Biophysical Research Communications*, 188, 483–489. [https://doi.org/10.1016/0006-291X\(92\)91081-Z](https://doi.org/10.1016/0006-291X(92)91081-Z).
- Hausinger, R.P. (2004) Metabolic versatility of prokaryotes for urea decomposition. *Journal of Bacteriology*, 186. <https://pubmed.ncbi.nlm.nih.gov/15090490/>
- Henikoff, S. & Henikoff, J.G. (1992) Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America*, 89, 10915–10919. <https://doi.org/10.1073/pnas.89.22.10915>.
- Huang, H.C., Erickson, R.S., Yanke, L.J., Hsieh, T.F. & Morrall, R.A.A. (2003) First report of pink seed of lentil and chickpea caused by *Erwinia rhapsodici* in Canada. *Plant Disease*, 87, 1398.
- Huang, L., Wang, J. & Lilley, D.M.J. (2017a) The Structure of the guanidine-II riboswitch. *Cell Chemical Biology*, 24, 695–702.e2. <https://doi.org/10.1016/j.chembiol.2017.05.014>.
- Huang, L., Wang, J., Wilson, T.J. & Lilley, D.M.J. (2017b) Structure of the guanidine III riboswitch. *Cell Chemical Biology*, 24, 1407–1415.e2. <https://doi.org/10.1016/j.chembiol.2017.08.021>.
- Hyatt, D., LoCascio, P.F., Hauser, L.J. & Uberbacher, E.C. (2012) Gene and translation initiation site prediction in metagenomic sequences.

- Bioinformatics*, 28, 2223–2230. <https://doi.org/10.1093/bioinformatics/bts429>.
- Kanamori, T., Kanou, N., Atomi, H. & Imanaka, T. (2004) Enzymatic characterization of a prokaryotic urea carboxylase. *Journal of Bacteriology*, 186, 2532–2539. <https://doi.org/10.1128/JB.186.9.2532-2539.2004>.
- Kermani, A.A., Macdonald, C.B., Gundepudi, R. & Stockbridge, R.B. (2018) Guanidinium export is the primal function of SMR family transporters. *Proceedings of the National Academy of Sciences*, 115, 3060–3065. <https://www.pnas.org/content/115/12/3060.short>. <https://doi.org/10.1073/pnas.1719187115>.
- Kihara, H., Prescott, J.M. & Snell, E.E. (1955) The bacterial cleavage of canavanine to homoserine and guanidine. *The Journal of Biological Chemistry*, 217, 497–503. [https://doi.org/10.1016/S0021-9258\(19\)57198-8](https://doi.org/10.1016/S0021-9258(19)57198-8).
- Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T. et al. (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Research*, 33, W299–302. <https://doi.org/10.1093/nar/gki370>.
- Lenkeit, F., Eckert, I., Hartig, J.S. & Weinberg, Z. (2020) Discovery and characterization of a fourth class of guanidine riboswitches. *Nucleic Acids Research*, 48, 12889–12899. <https://doi.org/10.1093/nar/gkaa1102/6006198>.
- Letunic, I. & Bork, P. (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, 47, W256–W259. <https://doi.org/10.1093/nar/gkz239>.
- Locher, K.P. (2016) Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nature Structural & Molecular Biology*, 23, 487–493. <https://doi.org/10.1038/nsmb.3216>.
- Maqbool, A., Horler, R.S.P., Muller, A., Wilkinson, A.J., Wilson, K.S. & Thomas, G.H. (2015) The substrate-binding protein in bacterial ABC transporters: dissecting roles in the evolution of substrate specificity. *Biochemical Society Transactions*, 43, 1011–1017. <https://doi.org/10.1042/BST20150135>.
- Meier-Kolthoff, J.P. & Göker, M. (2019) TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications*, 10, 2182. <https://doi.org/10.1038/s41467-019-10210-3>.
- Meyer, M.M., Hammond, M.C., Salinas, Y., Roth, A., Sudarsan, N. & Breaker, R.R. (2011) Challenges of ligand identification for riboswitch candidates. *RNA Biology*, 8, 5–10. <https://doi.org/10.4161/rna.8.1.13865>.
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G., Sonnhammer, E.L.L. et al. (2021) Pfam: The protein families database in 2021. *Nucleic Acids Research*, 49(D1), D412–D419. <https://doi.org/10.1093/nar/gkaa913>.
- Natelson, S. & Sherwin, J.E. (1979) Proposed mechanism for urea nitrogen re-utilization: relationship between urea and proposed guanidine cycles. *Clinical Chemistry*, 25, 1343–1344. <https://doi.org/10.1093/clinchem/25.7.1343>.
- NCBI Resource Coordinators. (2015) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 43, D6–D17. <https://pubmed.ncbi.nlm.nih.gov/25398906/>.
- Nelson, J.W., Atilho, R.M., Sherlock, M.E., Stockbridge, R.B. & Breaker, R.R. (2017) Metabolism of free guanidine in bacteria is regulated by a widespread riboswitch class. *Molecular Cell*, 65, 220–230. <https://doi.org/10.1016/j.molcel.2016.11.019>.
- Reiss, C.W., Xiong, Y. & Strobel, S.A. (2017) Structural basis for ligand binding to the guanidine-I riboswitch. *Structure*, 25, 195–202. <https://www.sciencedirect.com/science/article/pii/S0969212616303628>.
- Roon, R.J., Hampshire, J. & Levenberg, B. (1972) Urea amidolyase. The involvement of biotin in urea cleavage. *The Journal of Biological Chemistry*, 247, 7539–7545. [https://doi.org/10.1016/S0021-9258\(19\)44559-6](https://doi.org/10.1016/S0021-9258(19)44559-6).
- Roon, R.J. & Levenberg, B. (1968) An adenosine triphosphate-dependent, avidin-sensitive enzymatic cleavage of urea in yeast and green algae. *The Journal of Biological Chemistry*, 243, 5213–5215. [https://doi.org/10.1016/S0021-9258\(18\)92012-0](https://doi.org/10.1016/S0021-9258(18)92012-0).
- Roon, R.J. & Levenberg, B. (1972) Urea amidolyase. I. Properties of the enzyme from *Candida utilis*. *The Journal of Biological Chemistry*, 247, 4107–4113. [https://doi.org/10.1016/S0021-9258\(19\)45046-1](https://doi.org/10.1016/S0021-9258(19)45046-1).
- Saha, R., Farrance, C.E., Verghese, B., Hong, S. & Donofrio, R.S. (2013) *Klebsiella michiganensis* sp. nov., a new bacterium isolated from a tooth brush holder. *Current Microbiology*, 66, 72–78. <https://doi.org/10.1007/s00284-012-0245-x>.
- Salvail, H., Balaji, A., Yu, D., Roth, A. & Breaker, R.R. (2020) Biochemical validation of a fourth guanidine riboswitch class in bacteria. *Biochemistry*, 59, 4654–4662.
- Schicklberger, M., Shapiro, N., Loqué, D., Woyke, T. & Chakraborty, R. (2015) Draft genome sequence of *Raoultella terrigena* R1Gly, a diazotrophic endophyte. *Genome Announcements*, 3, e00607-15. <https://pubmed.ncbi.nlm.nih.gov/26067957/>.
- Schneider, N.O., Tassoulas, L.J., Zeng, D., Laseke, A.J., Reiter, N.J., Wackett, L.P. et al. (2020) Solving the conundrum: widespread proteins annotated for urea metabolism in bacteria are carboxyguanidine deiminases mediating nitrogen assimilation from guanidine. *Biochemistry*, 59, 3258–3270. <https://doi.org/10.1021/acs.biochem.0c00537>.
- Sherlock, M.E. & Breaker, R.R. (2017) Biochemical validation of a third guanidine riboswitch class in bacteria. *Biochemistry*, 56, 359–363. <https://doi.org/10.1021/acs.biochem.6b01271>.
- Sherlock, M.E. & Breaker, R.R. (2020) Former orphan riboswitches reveal unexplored areas of bacterial metabolism, signaling, and gene control processes. *RNA*, 26, 675–693. <https://doi.org/10.1261/rna.074997.120>.
- Sherlock, M.E., Malkowski, S.N. & Breaker, R.R. (2017) Biochemical validation of a second guanidine riboswitch class in bacteria. *Biochemistry*, 56, 352–358. <https://doi.org/10.1021/acs.biochem.6b01270>.
- Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. & Barton, G.J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25, 1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>.
- Weinberg, Z., Barrick, J.E., Yao, Z., Roth, A., Kim, J.N., Gore, J. et al. (2007) Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Research*, 35, 4809–4819. <https://doi.org/10.1093/nar/gkm487>.
- Weinberg, Z., Nelson, J.W., Lünse, C.E., Sherlock, M.E. & Breaker, R.R. (2017) Bioinformatic analysis of riboswitch structures uncovers variant classes with altered ligand specificity. *PNAS*, 114, E2077–E2085. <https://doi.org/10.1073/pnas.1619581114>.
- Wilkins, S. (2015) Structure and mechanism of ABC transporters. *F1000Prime Reports*, 7, 14. <https://doi.org/10.12703/P7-14>.
- Young, E.B. & Berges, J.A. (2016) Nitrogen stress in the marine environment: From scarcity to surfeit. In: Solan, M., & Whiteley, N. (Eds) *Stressors in the marine environment*. Oxford University Press, pp. 93–116.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Sinn M, Hauth F, Lenkeit F, Weinberg Z, Hartig JS. Widespread bacterial utilization of guanidine as nitrogen source. *Mol Microbiol*. 2021;116:200–210. <https://doi.org/10.1111/mmi.14702>