ISME

www.nature.com/ismej

ARTICLE

Novel D-glutamate catabolic pathway in marine *Proteobacteria* and halophilic archaea

Yang Yu[1], Peng Wang[2,3], Hai-Yan Cao[1], Zhao-Jie Teng[1], Yanping Zhu[1], Min Wang[2], Andrew McMinn[2,4], Yin Chen[2,5], Hua Xiang[6,7], Yu-Zhong Zhang[2,3,8]✉, Xiu-Lan Chen[1,3]✉ and Yu-Qiang Zhang[1]✉

D-glutamate (D-Glu) is an essential component of bacterial peptidoglycans, representing an important, yet overlooked, pool of organic matter in global oceans. However, little is known on D-Glu catabolism by marine microorganisms. Here, a novel catabolic pathway for D-Glu was identified using the marine bacterium *Pseudoalteromonas* sp. CF6-2 as the model. Two novel enzymes (DgcN, DgcA), together with a transcriptional regulator DgcR, are crucial for D-Glu catabolism in strain CF6-2. Genetic and biochemical data confirm that DgcN is a *N*-acetyltransferase which catalyzes the formation of *N*-acetyl-D-Glu from D-Glu. DgcA is a racemase that converts *N*-acetyl-D-Glu to *N*-acetyl-L-Glu, which is further hydrolyzed to L-Glu. DgcR positively regulates the transcription of *dgcN* and *dgcA*. Structural and biochemical analyses suggested that DgcN and its homologs, which use D-Glu as the acyl receptor, represent a new group of the general control non-repressible 5 (GCN5)-related *N*-acetyltransferases (GNAT) superfamily. DgcA and DgcN occur widely in marine bacteria (particularly *Rhodobacterales*) and halophilic archaea (*Halobacteria*) and are abundant in marine and hypersaline metagenome datasets. Thus, this study reveals a novel D-Glu catabolic pathway in ecologically important marine bacteria and halophilic archaea and helps better understand the catabolism and recycling of D-Glu in these ecosystems.

The ISME Journal (2023) 17:537–548; https://doi.org/10.1038/s41396-023-01364-6

INTRODUCTION

All proteinogenic amino acids, except for glycine, have D/L-enantiomers. Although almost all living organisms exclusively use L-amino acids (LAAs) instead of D-amino acids (DAAs) for protein synthesis, DAAs do occur widely in nature, including in bacterial peptidoglycan and other key bacterial macromolecules and metabolites, e.g. capsules, lipopolysaccharides, siderophores and antimicrobial peptides [1–5]. Archaeal cell-walls are not known to contain DAAs [6]; however DAAs are present in their membranes and proteins, and can be detected as free amino acids in certain species in culture [7]. In the global oceans, DAAs have been detected at nanomolar concentrations [8], representing an important, yet overlooked, group of organic matter in the marine microbial food web. However, it is puzzling that the catabolism and recycling of DAAs in the ocean are still largely unknown. Only a few reports on marine DAA-utilizing microorganisms are available. Kubota et al. isolated several bacteria that can utilize D-Val, D-Leu, D-Phe and/or D-Trp from deep-sea sediments, and showed that two strains can deaminate not only these four DAAs, but also D-Met, D-Pro, D-Lys, D-Arg and D-His to α-keto acids [9]. Yu et al. isolated several bacteria that utilize uncommon DAAs (D-Asp, D-Ser D-Leu, D-Met, D-Tyr, D-Thr and/or D-Phe) from the Arctic, and found that conversion of these seven DAAs to α-keto acids is the common pathway for DAA catabolism in these bacteria. Formation of α-keto acids from DAA involves several key enzymes, including DAA oxidases/dehydrogenases, DAA ammonia-lyases and DAA transaminases. In addition, conversion of D-Asp into L-Asp is another pathway, which is performed by Asp racemase from *Pseudoaltermonas elyakovii* [10]. In *Geobacillus kaustophilus*, *N*-succinyl amino acid racemase and DAA *N*-succinyltransferase irreversibly convert hydrophobic and basic DAAs (D-Ala, D-Val, D-Leu, D-Ile, D-Phe, D-Trp and D-Lys) to corresponding LAAs [11]. Ala racemase is involved in the catabolism of D-Ala in the archaeon *Methanococcus maripaludis* and the yeast *Schizosaccharomyces pombe* [12, 13].

D-glutamate (D-Glu) is an essential component of bacterial cell-wall peptidoglycans, and also occurs in several other bacterial macromolecules such as capsular polypeptide [14] and extracellular D-glutamyl polypeptide [15]. D-Glu has also been found in membranes in archaea, which are also known to release free DAAs [7]. D-Glu is relatively abundant in the oceans, with its concentration ranging from 0.3–62 nmol/L in Roskilde Fjord [8], and can make up to 6.2% of the total dissolved free amino acids in sediment pore-waters [16]. Therefore, the catabolism and recycling of D-Glu may contribute significantly to marine carbon and nitrogen cycles. While it is well established that D-Glu is a key precursor in the synthesis of cell-wall

[1]State Key Laboratory of Microbial Technology, Shandong University, Qingdao, China. [2]College of Marine Life Sciences & Frontiers Science Center for Deep Ocean Multispheres and Earth System, Ocean University of China, Qingdao, China. [3]Laboratory for Marine Biology and Biotechnology, Pilot National Laboratory for Marine Science and Technology, Qingdao, China. [4]Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, TAS, Australia. [5]School of Life Sciences, University of Warwick, Coventry, UK. [6]University of Chinese Academy of Sciences, Beijing, China. [7]State Key Laboratory of Microbial Resources, the Institute of Microbiology CAS, Beijing, China. [8]Marine Biotechnology Research Center, State Key Laboratory of Microbial Technology, Shandong University, Qingdao, China. ✉email: zhangyz@sdu.edu.cn; cxl0423@sdu.edu.cn; yqzh1989@sina.com

Received: 21 July 2022 Revised: 7 January 2023 Accepted: 11 January 2023
Published online: 23 January 2023

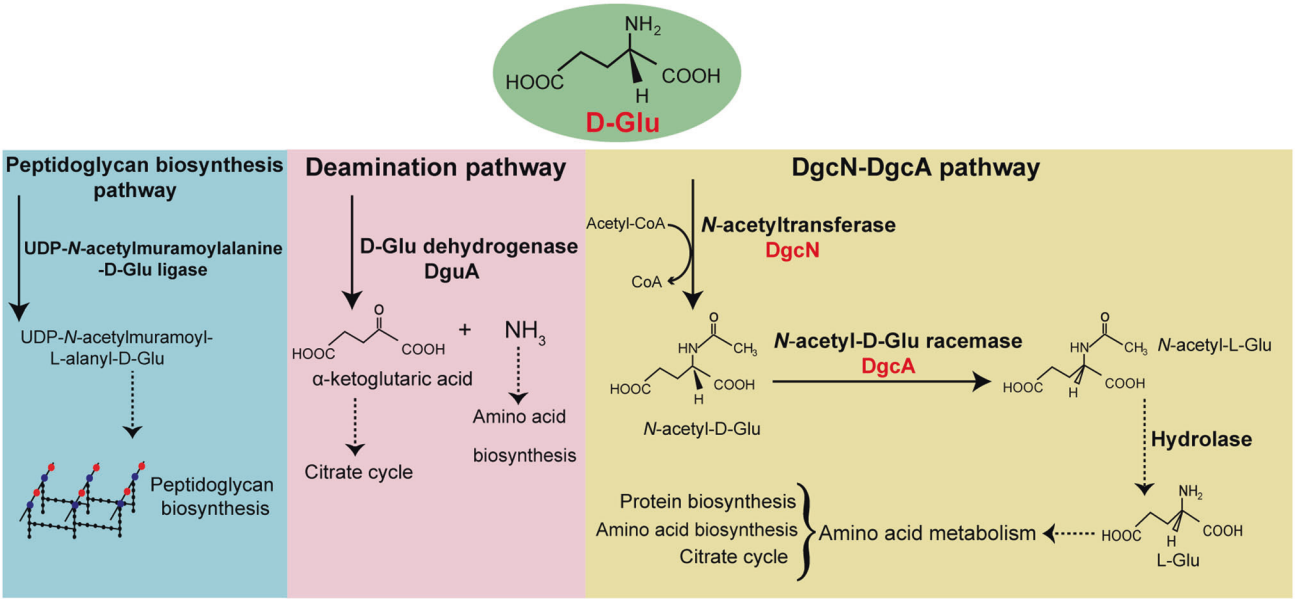

Fig. 1 Metabolic pathways for D-Glu in bacteria. The pathway involved in peptidoglycan biosynthesis [17] is shown in blue. The deamination pathway with a D-Glu dehydrogenase DguA as a key enzyme in the human pathogen *Pseudomonas aeruginosa* PAO1 [18] is shown in pink. The DgcN-DgcA pathway with an *N*-acetyltransferase and an *N*-acetyl-D-Glu racemase as key enzymes in *Pseudoalteromonas* sp. CF6-2 is shown in yellow. Key enzymes in the pathways are shown in bold.

peptidoglycans (Fig. 1) [17], its catabolism is poorly studied. Very few studies have investigated D-Glu catabolism in bacteria [9, 18] whilst no archaea is known to catabolize D-Glu. In the human pathogen *Pseudomonas aeruginosa* PAO1, D-Glu is deaminated by D-Glu dehydrogenase to α-ketoglutarate and ammonia, which are used as a carbon and a nitrogen source, respectively (Fig. 1) [18]. Similarly, some marine bacteria are able to deaminate D-Glu to α-ketoglutaric acid although D-Glu specific deaminase activities in these strains are rather weak, therefore begging the question whether the deamination pathway is important in ecologically abundant marine microorganisms [9].

We have previously shown that marine bacterium *Pseudoalteromonas* sp. CF6-2 can kill a variety of Gram-positive bacteria by attacking their cell-wall peptidoglycans and can use D-Glu for growth [19]. In this study, we show that strain CF6-2 employs a novel deamination-independent pathway to catabolize D-Glu (Fig. 1). In this new DgcN-DgcA pathway, D-Glu is catalyzed by the *N*-acetyltransferase DgcN to generate *N*-acetyl-D-Glu, which is then converted to *N*-acetyl-L-Glu by the *N*-acetyl-D-Glu racemase DgcA. *N*-acetyl-L-Glu is subsequently hydrolyzed into L-Glu, which serves as a carbon/nitrogen source for growth (Fig. 1). DgcR is a transcriptional regulator that positively regulates the genes involved in D-Glu catabolism. DgcN and its homologs represent a new group of the general control non-repressible 5 (GCN5)-related *N*-acetyltransferases (GNAT) superfamily based on the structural, biochemical and phylogenetic analyses. DgcA and DgcN occur widely in marine bacterial genomes and metagenomes, and are found in *Proteobacteria* (predominantly *Rhodobacterales*) and halophilic archaea (*Halobacteria*). The results thus reveal a novel D-Glu catabolic pathway which appears common in many ecologically important bacterial and archaeal groups, shedding light on the recycling and mineralization of D-Glu in these ecosystems.

MATERIALS AND METHODS

Bacterial strains, growth conditions and materials

Bacterial and archaeal strains used in this study are listed in Supplementary Table S1. Detailed protocols for the cultivation of bacterial and archaeal strains are described in Supplementary Materials and Methods.

Transcriptome analysis

Strain CF6-2 was cultured in marine LB medium at 20°C to an OD_{600} of 0.8. Cells were collected from the culture, washed three times with 3% sterile sea salt solution, and inoculated into the defined medium with 2 mM L/D-Glu, which were then cultivated at 20°C. Samples were taken at the initial (0 h) and the logarithmic growth (approximately 50 h) phases. Transcriptome sequencing and analysis were performed at Beijing Genome Institute (China). Upregulated genes were identified by p values <0.05 and an absolute fold-change threshold of >2.0.

Construction of mutants

The knockout mutants of the genes in the D-Glu gene cluster (DGC) of strain CF6-2, as well as their complementary strains, and the recombinant strain SM9913/pEV*dgcANR* were constructed with the method described previously [20, 21]. The detailed construction of mutants is shown in Supplementary Fig. S1, and detailed protocols for construction of mutants of strain CF6-2 and complemented strains are described in Supplementary Materials and Methods. The strains, plasmids and primers used in this study are listed in Supplementary Tables S2 and S3.

Real-time qPCR and reverse transcription PCR analysis

The wild-type (WT) strain CF6-2, Δ*dgcR* and Δ*dgcR*-pEV*dgcANR* were cultured with 2 mM D-Glu and sampled the same as in "Transcriptome analysis". Detailed protocols for extraction of total RNA and real-time qPCR analysis are described in Supplementary Materials and Methods.

Reverse transcription PCR (RT-PCR) was used to analyze co-transcriptional relationships between the key genes in DGC. The cDNA was formed by reverse-transcribed of the extracted RNA, and used as a template for PCR with five gene-specific primer pairs as shown in Supplementary Table S3. Each pair spanned the intergenic region between two open reading frames (ORFs).

Gene cloning, point mutation, protein expression and purification

Genes encoding DgcA, DgcN, DgcH and DgcR were cloned from the genomic DNA of strain CF6-2 via PCR and overexpressed in *Escherichia coli* BL21 (DE3) cells using the pET-30a (for DgcA) or pET-22b (for DgcN and DgcH) vector containing a His tag or the pMAL-c4X (for DgcR) vector containing a maltose binding protein (MBP) tag. Detailed protocols for protein expression and purification are described in Supplementary Materials and Methods.

To express the DgcA and DgcN proteins from marine *Rhodobacterales*, *Vibrionales* strains and halophilic archaeal strains, the genes were cloned from the genomic DNA via PCR and expressed in *E. coli* as described in the protein expression and purification section in Supplementary Materials and Methods. All point mutations in DgcN-25328 were introduced using the PCR-based method and verified by DNA sequencing. The circular dichroism spectra (CD) of DgcN-25328 and its mutants were monitored between 205 and 250 nm on a JASCO J-1500 Spectrometer (Japan).

Electrophoretic mobility shift assays (EMSAs)

EMSAs were conducted as reported previously [22]. DNA probes containing the upstream region of the genes *dgcR*, *dgcT*, *dgcP* and *dgcH* were labeled with 5′-Biotin. Detailed protocol for EMSAs is described in Supplementary Materials and Methods.

Enzyme assays

The *N*-acyltransferase activity of DgcN proteins was measured through a continuous spectrophotometric assay, based on detecting the release of 4-thiopyridine from 4,4′-dipyridyl disulfide [23]. The L-Ala-L/D-Glu epimerase activity and the *N*-acetylamino acid racemase activity of DgcA were assayed using the CD method previously described [24]. The *N*-acetyl-L-Glu hydrolase activity of DgcH was assayed by a modified colorimetric ninhydrin-based method [25]. Detailed protocols for these enzyme assays are described in Supplementary Materials and Methods.

Crystallization, data collection, structure determination and molecular docking simulations

The structure of DgcN-25328 was solved by molecular replacement using Phaser [26] with the protein DUF1611 structure (PDB code 2OBN) as the starting model. Detailed protocols for crystallization, data collection, structure determination and molecular docking simulations are described in Supplementary Materials and Methods.

Bioinformatic analysis

Marine bacterial genomes and archaeal genomes in the Integrated Microbial Genome database of the Joint Genome Institute (the IMG/JGI database) were probed using DgcA, DgcN and DgcR of strain *Pseudoalteromonas* sp. CF6-2 and DguA of strain *Pseudomonas aeruginosa* PAO1 as the queries with an *e*-value cut-off of e^{-50} (for marine bacterial genomes) and e^{-30} (for archaeal genomes). The *Tara* Oceans Microbiome and polar metagenome including 60 metagenomics seawater samples (NCBI BioProject accession no. PRJNA588686) [27], were probed for the distribution of DgcA and DgcN proteins in marine metagenomes using DgcA and DgcN as the queries sequence with an *e*-value cut-off of e^{-50}. The metagenomes from hypersaline environment in the IMG/JGI database were probed using DgcA-33500 and DgcN-33500 of strain ATCC33500 as the queries with an *e*-value cut-off of e^{-50}. Horizontal gene transfer events were predicted via RANGER-DTL 2.0 [28].

RESULTS AND DISCUSSION

Identification of the gene cluster involved in D-Glu catabolism in strain CF6-2

Strain CF6-2 can utilize either 2 mM L-Glu or D-Glu as the sole nitrogen source for growth (Fig. 2a), consistent to our previous report [19]. We also tested the ability of strain CF6-2 to utilize D-Glu at much lower, and perhaps more environmentally relevant, concentrations down to 50 μM. The results showed that strain CF6-2 can utilize D-Glu ranging from 2 mM to 50 μM as the sole nitrogen source for growth, although the growth with 50 μM D-Glu was quite weak (Fig. 2b).

To investigate the genes involved in D-Glu catabolism, transcriptomic analysis was performed on strain CF6-2 cultivated with either 2 mM L-Glu or D-Glu as the sole nitrogen source. The data revealed that 1267 and 962 genes were upregulated in the cells cultivated with D-Glu and L-Glu, respectively, and 876 genes were upregulated in the cells cultivated with D-Glu compared to those with L-Glu (Supplementary Table S4). Kyoto Encyclopedia of Genes and Genomes pathways analyses of these upregulated genes with D-Glu (compared to those with L-Glu) showed that most D-Glu induced genes are involved in amino acid metabolism, carbohydrate metabolism and energy metabolism (Supplementary Fig. S2). In addition, we noticed that the transcripts of genes orf01803-orf01809 (Fig. 2c) were significantly upregulated (fold change > 100) in the cells cultivated with D-Glu, but only one of these genes (orf01805) showed a weaker up-regulation (fold change: 2.96) in the cells with L-Glu (Table 1), maybe due to the structural similarity of L-Glu and D-Glu. Based on these data, it was postulated that this gene cluster is likely involved in D-Glu catabolism in strain CF6-2. The six genes in this gene cluster encode a putative L-Ala-D/L-Glu epimerase (orf01803), a protein of unknown function (orf01804), a LysR family transcriptional regulator (orf01805), a Na^+/H^+ antiporter NhaC (orf01806), a peptidase M14-like domain-containing protein (orf01808) and an amidohydrolase (orf01809), respectively (Table 1). Hereafter, this gene cluster was designated the D-Glu gene cluster (DGC), and the genes were *dgcA* (orf01803), *dgcN* (orf01804), *dgcR* (orf01805), *dgcT* (orf01806), *dgcP* (orf01808) and *dgcH* (orf01809) (Table 1).

The transcriptional levels of DGC genes were detected by RT-qPCR when strain CF6-2 was cultured with different concentrations (2 mM-50 μM) of D-Glu as the sole nitrogen source. The transcriptions of all the DGC genes were upregulated in the cells cultivated with all concentrations D-Glu (Fig. 2d).

To determine the role of the DGC genes in D-Glu catabolism, single gene deletion mutants (Δ*dgcA*, Δ*dgcN*, Δ*dgcR*, Δ*dgcT*, Δ*dgcP* and Δ*dgcH*) were constructed. While the growth of the mutant Δ*dgcT*, Δ*dgcP* or Δ*dgcH* was similar to that of the WT, the growth of the mutant Δ*dgcA*, Δ*dgcN* or Δ*dgcR* was significantly reduced (Fig. 3a, b, c, f), indicating that *dgcA*, *dgcN* and *dgcR* are essential in D-Glu catabolism. Moreover, *dgcA*, *dgcN* and *dgcR*, but not *dgcT*, *dgcP* nor *dgcH*, were co-transcribed (Fig. 3g).

To further investigate the functional relationships of *dgcA*, *dgcN* and *dgcR*, a double mutant Δ*dgcA*Δ*dgcN* and a triple mutant Δ*dgcA*Δ*dgcN*Δ*dgcR* were constructed. These mutants grew similarly on D-Glu compared to that of the single gene deletion mutants (Fig. 3a–e). Interestingly, the single gene deletion mutants as well as the double deletion and triple deletion mutants can only be complemented by expressing the *dgcANR* genes in these mutants (Fig. 3a–e), suggesting that all three genes are required to restore the growth of these mutants on D-Glu, possibly because the polar effect of mutation affects the expression of downstream genes.

To investigate whether these genes (*dgcA*, *dgcN* and *dgcR*) is sufficient for D-Glu catabolism, the plasmid pEV*dgcANR* carrying the native *dgcANR* genes of strain CF6-2 was transferred into strain *Pseudoalteromonas* sp. SM9913, a marine bacterium that is unable to use D-Glu as a sole nitrogen source for growth (Fig. 3h). Indeed, expression of *dgcANR* enabled strain SM9913 to grow on D-Glu (Fig. 3h), indicating that *dgcA*, *dgcN* and *dgcR* are not only involved in but also sufficient for D-Glu catabolism in strain CF6-2.

DgcR positively regulates the transcription of the DGC

The gene *dgcR* in the DGC was predicted to encode a transcriptional regulator of the LysR family, the most widespread transcriptional regulators in bacteria [29]. DgcR shares 100% coverage and 78.0% sequence identity with CysB of *Pseudoalteromonas arctica* NEC-BIFX-2020_0012, a LysR family transcriptional regulator that controls the expression of genes associated with cysteine biosynthesis [30]. The fact that *dgcR* co-transcribes with *dgcA* and *dgcN* suggests that DgcR likely functions as a transcriptional regulator in D-Glu catabolism in strain CF6-2. To support this, transcriptions of the DGC genes were compared between the WT, the Δ*dgcR* mutant and the complemented strain Δ*dgcR*-pEV*dgcANR*. The RT-qPCR results showed that the transcriptions of the DGC genes were significantly reduced in the Δ*dgcR* mutant but restored in the complemented strain Δ*dgcR*-pEV*dgcANR* (Fig. 4a), indicating

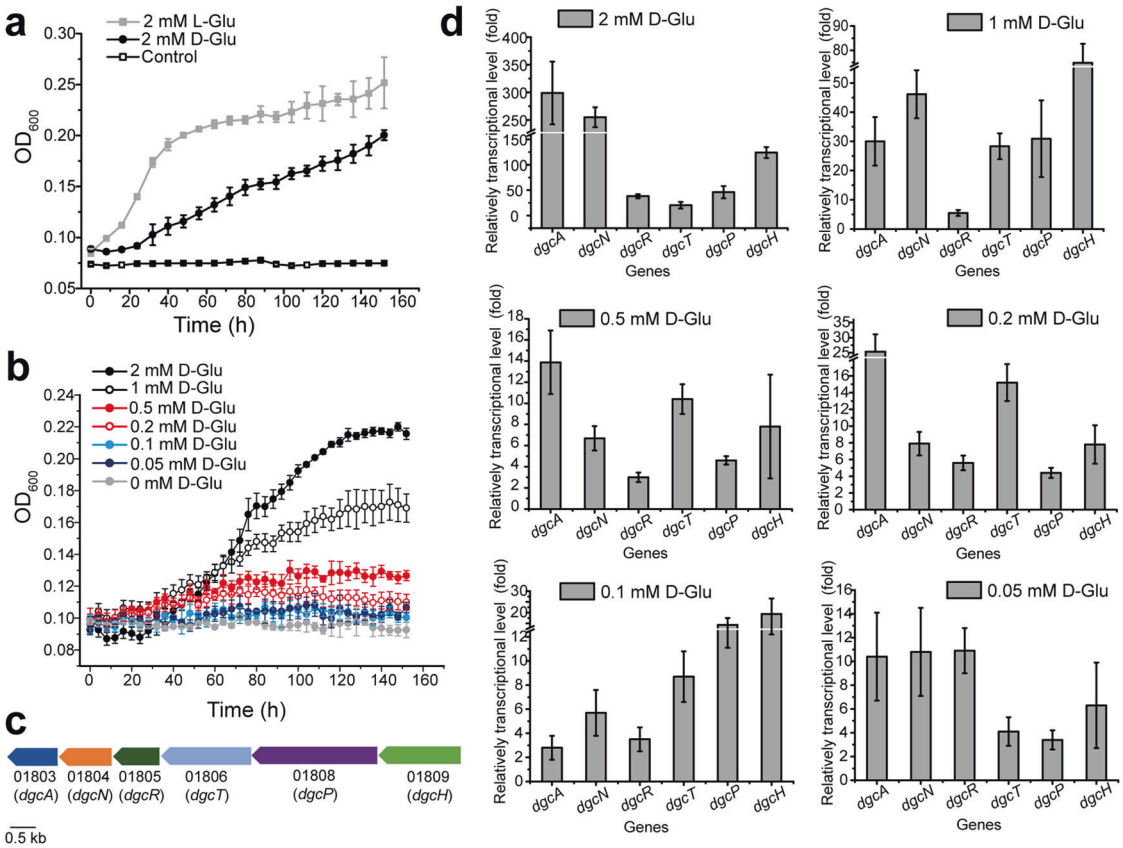

Fig. 2 Identification of the gene cluster responsible for D-Glu catabolism in strain CF6-2. a Growth curves of strain CF6-2 in the media containing 50 mM glucose and 2 mM L-Glu or D-Glu. The medium containing all components except for L-Glu/D-Glu was used as the control. **b** Growth curves of strain CF6-2 in the media containing 50 mM glucose and 0.05-2 mM D-Glu. The medium containing all components except for D-Glu was used as the control. **c** The gene cluster involved in D-Glu catabolism in strain CF6-2. **d** RT-qPCR analysis of the transcriptional levels of genes in the gene cluster involved in D-Glu catabolism in strain CF6-2 cultured with 0.05-2 mM D-Glu as a sole nitrogen source. The *recA* gene was used as the reference. The graphs in **a**, **b** and **d** show data from triplicate experiments (mean ± SD).

Table 1. Transcriptomics analysis of genes in the D-Glu gene cluster in strain CF6-2 cultured with L-Glu or D-Glu (2 mM) as a sole nitrogen source.

Gene no.	Gene name	Accession no.	Annotation	Medium	Fold change*	*p* value
orf01803	*dgcA*	ON505815	L-Ala-D/L-Glu epimerase	L-Glu	0.34	2.91E-13
				D-Glu	112.63	2.64E-14
orf01804	*dgcN*	ON505816	Protein of unknown function	L-Glu	0.25	1.08E-18
				D-Glu	175.12	9.17E-14
orf01805	*dgcR*	ON505817	LysR family transcriptional regulator	L-Glu	2.96	5.82E-09
				D-Glu	171.08	2.06E-12
orf01806	*dgcT*	ON505818	Na^+/H^+ antiporter NhaC	L-Glu	1.26	0.04444
				D-Glu	105.05	0
orf01808	*dgcP*	ON505819	Peptidase M14-like domain-containing protein	L-Glu	1.27	0.000369
				D-Glu	144.93	6.22E-13
orf01809	*dgcH*	ON505820	Amidohydrolase	L-Glu	0.66	2.54E-05
				D-Glu	125.91	1.35E-13

*The folds were calculated by comparing the logarithmic growth phase to 0 h, and upregulated genes were identified by *p* values <0.05 and an absolute fold-change threshold of >2.0.

that DgcR positively regulates the transcriptions of the genes involved in D-Glu catabolism.

To investigate the mechanism of which DgcR controls the expression of the DGC genes, *dgcR* was expressed in *E. coli* BL21 (DE3) as a recombinant fusion protein DgcR-MBP. To identify the DgcR binding sites upstream of the DGC genes *dgcR* (*dgcA* and *dgcN* co-transcribe with *dgcR*), *dgcT*, *dgcP* and *dgcH*, several DNA probes containing the upstream of these genes, which were divided into multiple segments (*dgcR*-1, 2, 3, 4; *dgcT*-1, 2, 3; *dgcP*-1, 2, 3; *dgcH*-1, 2, 3), were generated by PCR (Supplementary Table S5), and EMSA experiments were performed. The result showed that DgcR-MBP bound DNA probes *dgcR*-4, *dgcT*-1, *dgcT*-

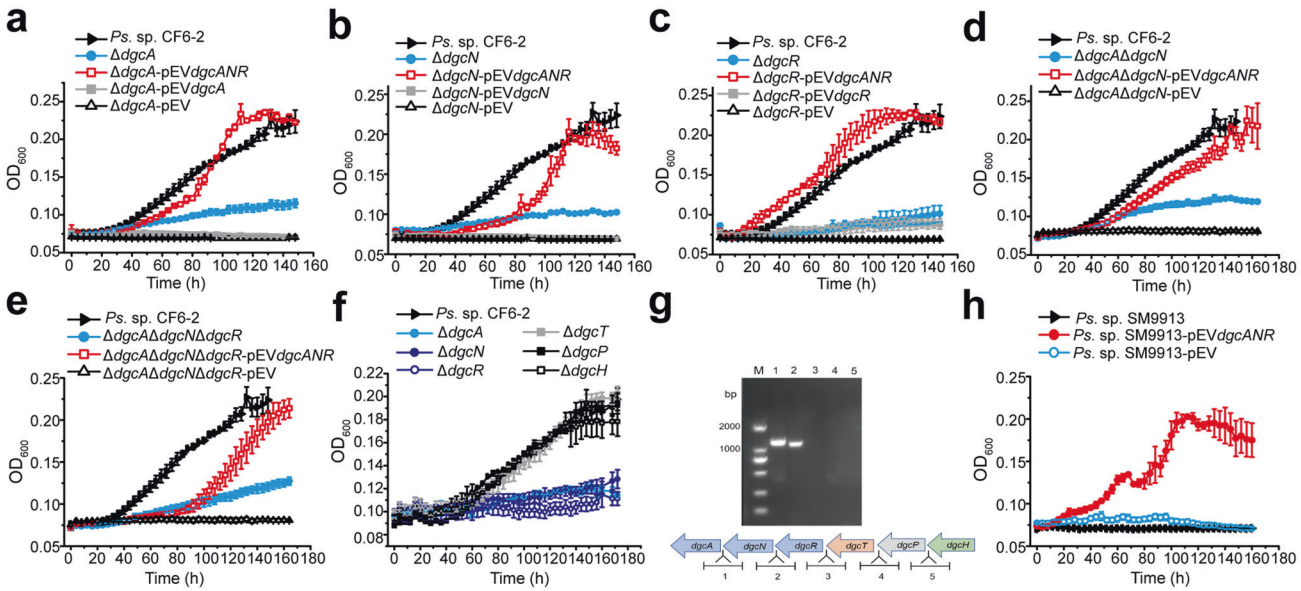

Fig. 3 Identification of the key genes in the DGC. The growth of single gene/multiple genes deletion mutants and complement strains cultured with 2 mM D-Glu as the sole nitrogen source. The single gene/multiple genes deletion mutants are Δ*dgcA* (**a**), Δ*dgcN* (**b**), Δ*dgcR* (**c**), Δ*dgcA*Δ*dgcN* (**d**) and Δ*dgcA*Δ*dgcN*Δ*dgcR* (**e**). **f** Growth curves of strains WT CF6-2, Δ*dgcA*, Δ*dgcN*, Δ*dgcR*, Δ*dgcT*, Δ*dgcP* and Δ*dgcH* in the medium containing 2 mM D-Glu as the sole nitrogen source. **g** RT-PCR analysis of co-transcription of the genes in the DGC in strain CF6-2. Cells cultivated with 2 mM D-Glu as the sole nitrogen source were used for RNA extraction and cDNA synthesis. Lanes 1 to 5 correspond to RT-PCR products of intergenic regions. **h** Growth curves of strains WT SM9913, SM9913-pEV*dgcANR* containing the plasmid pEV*dgcANR* and SM9913-pEV containing the plasmid pEV cultured with 2 mM D-Glu as the sole nitrogen source. The graphs show data from triplicate experiments (mean ± SD).

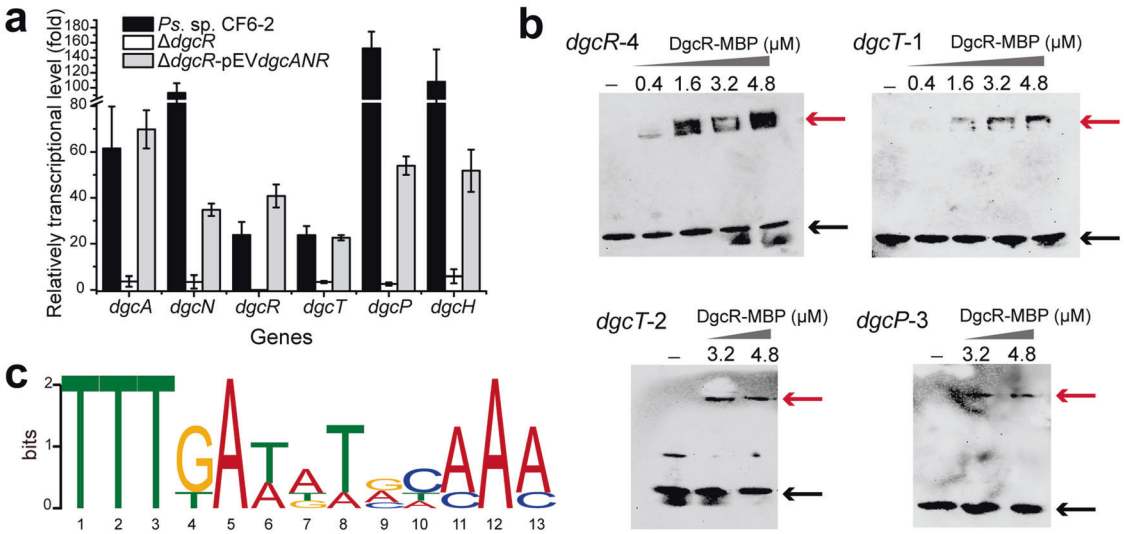

Fig. 4 Analysis of the regulatory mechanism of DgcR on the transcriptions of the DGC genes. a RT-qPCR analysis of the relative transcriptional levels of the genes in the DGC in WT strain CF6-2, the Δ*dgcR* mutant and the complemented strain Δ*dgcR*-pEV*dgcANR* cultured with 2 mM D-Glu as the sole nitrogen source. The *recA* gene was used as the reference, and the folds were calculated using the transcripts of gene *recA* and the 0 h sample for normalization. The graph shows data from triplicate experiments (mean ± SD). **b** EMSA analysis of the DgcR binding sites upstream of the DGC genes *dgcR*, *dgcT*, and *dgcP*. DNA probe (1 nM) was incubated with different amounts of DgcR-MBP. Black arrows indicate the free DNA probe, and red arrows indicate DgcR-MBP-DNA complex. **c** Multiple EM for Motif Elicitation analysis of the putative DgcR binding site. The sequences of DNA probes *dgcR*-4, *dgcT*-1, *dgcT*-2, *dgcP*-3 and *dgcH*-3 were used as inputs for MEME analysis. A 13-nt consensus sequence rich in AT (TTTGnAnTnCAAA) was identified.

2 and *dgcP*-3 strongly (Fig. 4b) and also bound DNA probe *dgcH*-3 weakly (data not shown), but the MBP protein itself did not bind these DNA probes (Supplementary Fig. S3). We then used the Multiple EM for Motif Elicitation (MEME; http://meme-suite.org/) toolbox to analyze the probes *dgcR*-4, *dgcT*-1, *dgcT*-2, *dgcP*-3 and *dgcH*-3, and identified a 13-nt consensus sequence with reverse palindrome (TTTGnAnTnCAAA) in all these probes (Fig. 4c and Supplementary Fig. S3). The 13-nt consensus sequence is rich in AT, a characteristic of the binding sites of LysR transcriptional regulatory family [31–33], suggesting that it is likely the binding site of DgcR. Together, these results suggested that, DgcR positively regulates the transcriptions of all the DGC genes via binding the consensus sequence upstream of *dgcR*, *dgcT*, *dgcP* and *dgcH*. Such a mechanism of transcriptional factors to regulate gene expression is common in bacteria [34, 35].

DgcN converts D-Glu to *N*-acetyl-D-Glu and DgcA converts *N*-acetyl-D-Glu to *N*-acetyl-L-Glu

We set out to analyze the functions of DgcA and DgcN in D-Glu catabolism in strain CF6-2. While *dgcN* is annotated as a protein of unknown function, *dgcA* is predicted to encode a L-Ala-D/L-Glu epimerase, a member of the muconate lactonizing enzyme (MLE) subgroup of the enolase superfamily. DgcA shares 44.5% sequence identity (91% coverage) with L-Ala-D/L-Glu epimerase YcjG of *E. coli* MG1655 [36] and 24.5% sequence identity (76% coverage) with the *N*-succinyl amino acid racemase of strain *G. kaustophilus* [11]. Since both L-Ala-D/L-Glu epimerase and *N*-succinyl amino acid racemase belong to the MLE subgroup, it suggests that DgcA may have *N*-acyl-amino acid racemase activity. Indeed, the genes *dgcA* and *dgcN* are locating next to each other in strain CF6-2, similar to the arrangement of the *N*-succinyl amino acid racemase and DAA *N*-succinyltransferase genes in *G. kaustophilus* (Fig. 5a). Thus, we speculated that DgcN of strain CF6-2 may have *N*-acyltransferase activity, although DgcN was predicted to be a protein of unknown function and shares no sequence identity with the DAA *N*-succinyltransferase of strain *G. kaustophilus*. To test this hypothesis, *dgcN* was expressed in *E. coli* BL21 (DE3). Recombinant DgcN showed *N*-acetyltransferase activity towards acetyl-CoA but not succinyl-CoA using D-Glu as the CoA acceptor (Fig. 5b), suggesting that DgcN catalyzes the *N*-acetylation of D-Glu to generate *N*-acetyl-D-Glu in vivo. DgcN also catalyzed the transfer of the acetyl moiety from acetyl-CoA to D-Gln, D-Asp and D-Asn, but displayed the highest activity to D-Glu (Fig. 5b).

Because *N*-acetyl-D-Glu can be formed by DgcN, DgcA may therefore catalyze the racemization of *N*-acetyl-D-Glu. Indeed, recombinant DgcA showed racemase activity towards *N*-acetyl-D-Glu (Fig. 5c), suggesting that DgcA likely catalyzes the racemization of *N*-acetyl-D-Glu to *N*-acetyl-L-Glu in vivo. Furthermore, DgcA, as a putative L-Ala-D/L-Glu epimerase, also showed racemase activity towards dipeptide L-Ala-D-Glu (Fig. 5c), a key constituent of peptidoglycan muropeptides, suggesting that DgcA may also contribute to the degradation of peptidoglycans.

DgcN and its homologs represent a new group in the GNAT superfamily

Although DgcN is annotated as a protein of unknown function, our results indicated that DgcN is an *N*-acetyltransferase responsible for *N*-acetylation of D-Glu. We next attempted to solve the crystal structure of DgcN, but this was not successful due to poor diffraction of the crystals. However, we solved the crystal structure of DgcN-25328, a homolog of DgcN from marine bacterium *Tritonibacter scottomollicae* DSM25328 [37]. DgcN-25328 is also able to convert D-Glu to *N*-acetyl-D-Glu and shares 45.2% sequence identity (99% coverage) with DgcN. The structure of DgcN-25328 was solved by molecular replacement using the unknown function protein DUF1611 structure (PDB code 2OBN) as the starting model. DgcN-25328 is composed of 332 residues and exists as a monomer in the crystal structure. It consists of 9 β-strands and 14 α-helices, which form two domains, an N-terminal domain (β0-α0-α1-α2-β1-α3-β2-α4-α5-β3) and a C-terminal domain (β4-α6-β5-α7-α8-β6-α9-β7-α10-β8-α11-α12-α13) (Fig. 6a). The N-terminal domain is a Rossman-like domain, in which the β-strands form a hydrogen-bonded β-sheet (Fig. 6a). The C-terminal domain contains an atypical fold (β5-α7-α8-β6-α9-β7-α10-β8-α11) (Fig. 6a, b) that is similar to, but distinct from, the conserved core fold (β-α-α-β-β-β-α-β-α) of the GNAT superfamily enzymes (Fig. 6b, Supplementary Fig. S4). First, the C-terminal fold of DgcN-25328 contains 4 β-strands and 5 α-helices, but the typical conserved core fold of a GNAT enzyme contains 5 β-strands and 4 α-helices. Second, the β-sheets of the DgcN-25328 fold are parallel, but those of the GNAT superfamily are antiparallel. Third, the location of the signature motif at pyrophosphate of acyl-CoA binding site, the

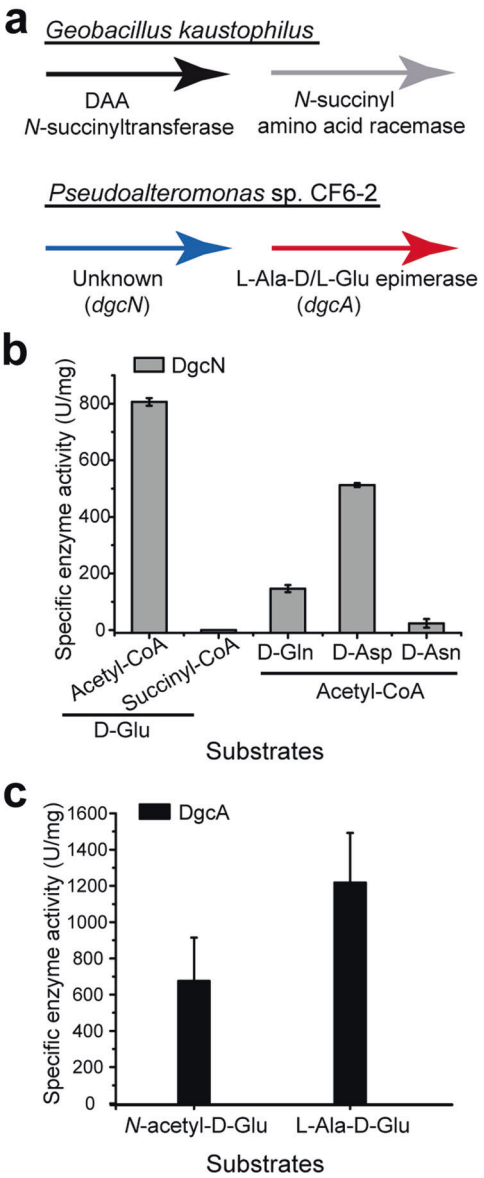

Fig. 5 Functional analysis of the key enzymes DgcA and DgcN in the DGC from strain CF6-2. a The operon context of the *N*-succinyl amino acid racemase and DAA *N*-succinyltransferase genes in *Geobacillus kaustophilus*, and genes *dgcA* and *dgcN* in strain CF6-2. **b** Specific activities of DgcN towards D-Glu and acetyl/succinyl-CoA, and towards acetyl-CoA and D-Gln/D-Asp/D-Asn. **c** Specific activities of DgcA towards *N*-acetyl-D-Glu and L-Ala-D-Glu detected by CD. Measurement of CD spectra were recorded between 190 nm and 250 nm by 1 nm carving, and were averaged from three scans. The graphs in **b**, **c** show data from triplicate experiments (mean ± SD).

"P-loop" (L215-I216-E217-G218-Q219-G220), of DgcN-25328 connecting β6 and α9 is different from that of the typical P-loop in the GNAT superfamily. The GNAT enzymes catalyze the transfer of an acyl moiety from acyl-CoA to the amino group of a large array of substrates ranging from small molecules to macromolecules such as aminoglycosides [38], aryl alkylamine [39], glucosamine-6-phosphate [40], protein *N*-termini and lysine residues of histones [41] and spermidine/spermine [42]. Considering that both DgcN and DgcN-25328 have *N*-acetyltransferase activity on D-Glu and DgcN-25328 has a similar core structure to the GNAT enzymes, DgcN and DgcN-25328 were suggested to belong to the GNAT superfamily.

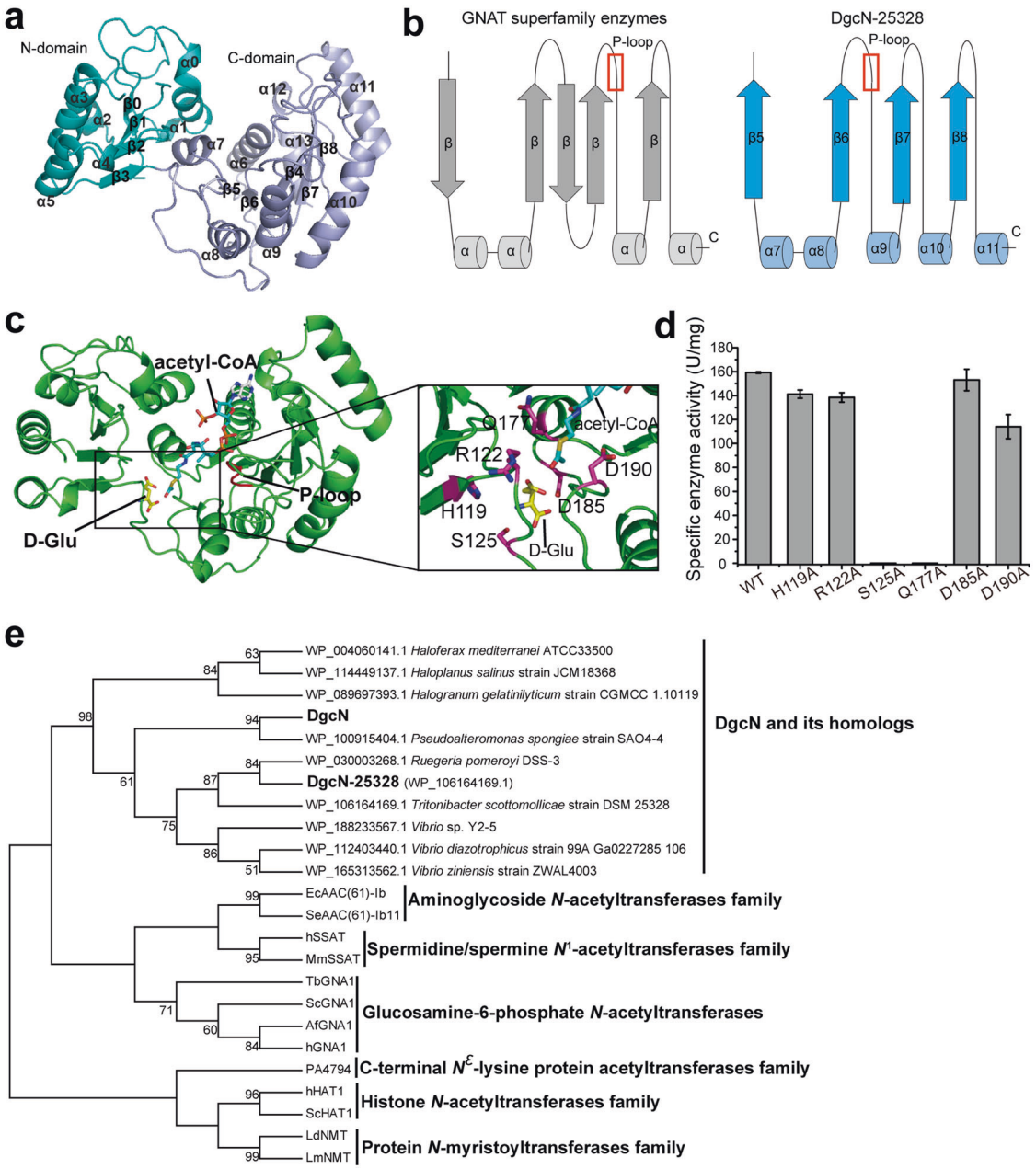

Fig. 6 Structural and phylogenetic analyses of DgcN-25328. a The overall structure of DgcN-25328. The N-terminal and C-terminal domains are colored in blue and purple, respectively. **b** Topologies of the core folds of DgcN-25328 and the general GNAT enzymes. The locations of the P-loops are framed by red boxes. **c** The structure of DgcN-25328 docked with acetyl-CoA and D-Glu. The pyrophosphate of acetyl-CoA binding site "P-loop" is colored in red, and the substrates acetyl-CoA and D-Glu are shown as blue and yellow sticks, respectively. The residues, His119, Arg122, Ser125, Gln177, Asp185 and Asp190, close to the acetyl group of acetyl-CoA and D-Glu are shown as magenta sticks. **d** Specific enzyme activities of DgcN-25328 and its mutants. The graph shows data from triplicate experiments (mean ± SD). **e** Phylogenetic tree of DgcN and other acyltransferases in the GNAT superfamily. The tree was built using the Maximum-likelihood method with Jones-Taylor-Thornton (JTT) matrix-based model from 131 amino acid positions. Bootstrap analysis of 1000 replicates was conducted, and values above 50% are shown.

We attempted to solve the structure of DgcN-25328 in complex with acetyl-CoA or D-Glu to analyze the key amino acid residues involved in substrate recognition and catalysis, however this was not successful. We therefore docked acetyl-CoA and D-Glu into the structure of DgcN-25328. In this structure, the pyrophosphate of acetyl-CoA was predicted to be bound at the "P-loop" (Fig. 6c), just like that in the aminoglycoside *N*-acetyltransferase EcAAC(6')-Ib of *Escherichia coli* [43]. Residues His119, Arg122, Ser125, Gln177, Asp185 and Asp190 are close to the acetyl group of acetyl-CoA and D-Glu (Fig. 6c), suggesting that they may be involved in substrate recognition and catalysis. Mutations of Ser125 and Gln177 to Ala almost completely abolished the enzymatic activity of DgcN-25328, but the mutants His119Ala, Arg122Ala, Asp185Ala and Asp190Ala still maintained more than 70% activities (Fig. 6d). CD spectra of these mutants were indistinguishable from that of WT DgcN-25328 (Supplementary Fig. S5), indicating that the decrease in the enzymatic activities of the mutants were caused by residue replacement rather than structural alteration of the enzyme. These structural and mutational analyses suggested

that amino acid residues Ser125 and Gln177 play a crucial role in the catalytic process of DgcN-25328.

The GNAT superfamily enzymes are coarsely classified into many groups based on the differences of acyl acceptors [44]. D-Glu is the acetyl acceptor of DgcN, representing a new acetyl acceptor that has not been shown for other members of the GNAT superfamily. Moreover, in the constructed phylogenetic tree, DgcN, together with DgcN-25328 and their homologs, forms a separate branch in the GNAT superfamily (Fig. 6e). These data suggest that DgcN and its homologs represent a new group of the GNAT superfamily.

N-acetyl-L-Glu may be hydrolyzed into L-Glu by DgcH

We suspected that *N*-acetyl-L-Glu is converted to L-Glu which is used by strain CF6-2 as a nitrogen source. To identify the enzyme to hydrolyze *N*-acetyl-L-Glu into L-Glu, we tried to express the three non-specific genes *dgcT*, *dgcP* and *dgcH*. However, only gene *dgcH* was successfully expressed in *E. coli* and the recombinant DgcH protein was purified. The gene *dgcH* encodes a putative *N*-acyl-amino acid-amidohydrolase belonging to the metal-dependent amidohydrolase superfamily, which catalyze the hydrolysis of *N*-acyl-amino acids to produce the corresponding amino acids [45]. DgcH shares 29.1% sequence identity (89% coverage) with the *N*-acyl-amino acids-amidohydrolase of *Alcaligenes faecalis* DA1, which has a high stereospecificity to *N*-acyl-DAAs [46]. Similarly, recombinant DgcH had the ability to hydrolyze *N*-acetyl-L-Glu to generate L-Glu (8.4 ± 1.7 U/mg), suggesting that the gene *dgcH* in the DGC may be responsible for the hydrolysis of *N*-acetyl-L-Glu into L-Glu. However, the mutant Δ*dgcH* was still able to utilize D-Glu just like the WT. We thus believed that there is some redundancy in *N*-acetyl-L-Glu hydrolysis in strain CF6-2 and that it is likely that some other gene(s) might have compensated the loss of gene *dgcH* in the mutant Δ*dgcH*. Although genes *dgcT* and *dgcP* in the DGC are non-specific for D-Glu catabolism of strain CF6-2, they may also play a role in D-Glu catabolism. The gene *dgcT* was predicted to encode a Na^+/H^+ antiporter NhaC, which, therefore, may be involved in D-Glu transport. The gene *dgcP* was predicted to encode a peptidase M14-like domain-containing protein that can cleave *N*-acetyl-L-Asp into L-Asp and acetate [47], and thus *dgcP* may be involved in *N*-acetyl-L-Glu hydrolysis. Because the heterogenous expression of the genes *dgcT* and *dgcP* was unsuccessful, the functions of these genes still need further study.

Diversity and distribution of marine bacteria catabolizing D-Glu via the DgcN-DgcA pathway

Based on the above results, the D-Glu catabolism pathway in strain CF6-2 is proposed (Fig. 7a). After entering the cell, D-Glu is first converted by DgcN into *N*-acetyl-D-Glu, which is then racemized by DgcA into *N*-acetyl-L-Glu. The produced *N*-acetyl-L-Glu is subsequently hydrolyzed into L-Glu, which is further metabolized as a nitrogen source. DgcR positively regulates the transcription of *dgcA* and *dgcN*. Because DgcA and DgcN are two key functional enzymes in this D-Glu catabolism pathway, this pathway is named the DgcN-DgcA pathway in this study.

In order to understand the ecological significance of the DgcN-DgcA pathway for D-Glu catabolism in marine microorganisms, the diversity and distribution of DgcA, DgcN and DgcR in the IMG/JGI database, the polar Microbiome database and the *Tara* Oceans Microbiome database were investigated. Of the 1956 genomes of marine bacteria in the IMG database, 11 contain all three genes in an apparent operon (DgcANR) and 531 contain both DgcA and DgcN sequences (Supplementary Table S6). The DgcAN sequences are predominantly found in the phylum *Proteobacteria* (99.0%), mainly affiliated with order *Rhodobacterales* (50.8%) of class *Alphaproteobacteria* and order *Vibrionales* (35.2%) of class *Gammaproteobacteria* (Fig. 7b). An in-depth analysis of the DgcAN sequences of *Rhodobacterales* showed that the *Roseobacter* clade, one of the major marine bacterial clades, is particularly abundant (66.6%).

The predominance of DgcA and DgcN sequences in *Rhodobacterales* was also observed in the polar and *Tara* Oceans Microbiome databases (Supplementary Table S7). According to the depth of sampling sites, polar and *Tara* samples were divided into four groups, Polar-Surface (0–100 m), Polar-Deep (300–3800 m), *Tara*-Surface (5–188 m) and *Tara*-Deep (250–1000 m). Of *Tara*-Surface, *Tara*-Deep, Polar-Surface and Polar-Deep samples, 31.5%, 7.5%, 38.2% and 19.3% of DgcA sequences, and 24.1%, 5.1%, 36.8% and 22.2% of DgcN sequences, respectively, are from *Rhodobacterales* (Fig. 7c). DgcA and DgcN sequences from *Vibrionales* are much less abundant in these metagenomes, likely reflecting the low abundance of *Vibrionales* in the microbial community (Fig. 7c). Together, bioinformatics investigations of marine bacterial genomes and metagenomes showed that DgcAN sequences, predominated by *Rhodobacterales*, are widely distributed in the global oceans (Supplementary Fig. S6).

To investigate whether bacteria containing the *dgcAN* genes can metabolize D-Glu via the DgcN-DgcA pathway, we tested five additional marine bacterial strains containing *dgcAN* (three were *Rhodobacterales*, and two were *Vibrionales*), among which the *Vibrionales* strains contain the gene *dgcR* in the *dgcANR* operon, and the *Rhodobacterales* strains do not. Indeed, all five strains (*Epibacterium scottomollicae* DSM25328, *Ruegeria pomeroyi* DSS-3, *Oceanicella* sp. SM1341, *Vibrio atlanticus* SM1925 and *Vibro tasmaniensis* SM1924) can grow on 2 mM D-Glu (Fig. 7d). To further support the role of DgcN-DgcA in D-Glu catabolism in these bacteria, we expressed and purified recombinant DgcA and DgcN of all five strains and carried out enzymatic activity assays. The recombinant DgcN proteins from all five strains had *N*-acetyltransferase activity towards acetyl-CoA and D-Glu (Fig. 7e). DgcA proteins were only successfully expressed from three out of the five strains (DSM25328, DSS-3 and SM1341) and they all had racemase activity towards *N*-acetyl-D-Glu (Fig. 7f). Together, our results indicated that these marine bacteria can utilize D-Glu most likely via the DgcN-DgcA pathway.

It has been reported that the human pathogen *Pseudomonas aeruginosa* PAO1 can catabolize D-Glu via the deamination pathway using the D-Glu dehydrogenase DguA (Fig. 1) [18].To investigate whether the marine bacteria having the DgcN-DgcA pathway also contain the deamination pathway, we examined the distribution of D-Glu dehydrogenase DguA in the 531 DgcAN-containing bacterial genomes. Of these genomes, 267 contain both DgcAN and DguA sequences (Supplementary Table S8), suggesting that these bacteria may be able to catabolize D-Glu via both the DgcN-DgcA and deamination pathways. These bacteria predominantly belong to orders *Vibrionales* (61.8%) of class *Gammaproteobacteria*, and *Rhodobacterales* (26.6%) of class *Alphaproteobacteria* (Supplementary Fig. S7).

Halophilic archaea catabolize D-Glu via the DgcN-DgcA pathway

A somewhat unexpected finding from genomic analysis of the distribution of DgcN-DgcA is that a significant number of DgcN-DgcA sequences were affiliated with archaea, despite the fact that no archaeal strains is known to metabolize D-Glu. To uncover the genetic potential of archaeal catabolism of D-Glu via the DgcN-DgcA pathway, all archaeal genomes in the IMG/JGI database were analyzed. Of the 1622 archaeal genomes, 223 contain DgcAN although DgcR is not found (Supplementary Table S9). Most of the archaeal strains containing DgcAN are halophilic archaea isolated from hypersaline environment, majority of which are from class *Halobacteria* (99.1%) of phylum *Euryarchaeota*. These are represented by five families *Haloferacaceae* (25.6%), *Natrialbaceae* (23.8%), *Halobacteriaceae* (16.6%), *Halorubraceae* (15.7%) and *Haloarculaceae* (10.8%) (Fig. 8a). These data suggest that halophilic archaea may also catabolize D-Glu via the DgcN-DgcA pathway.

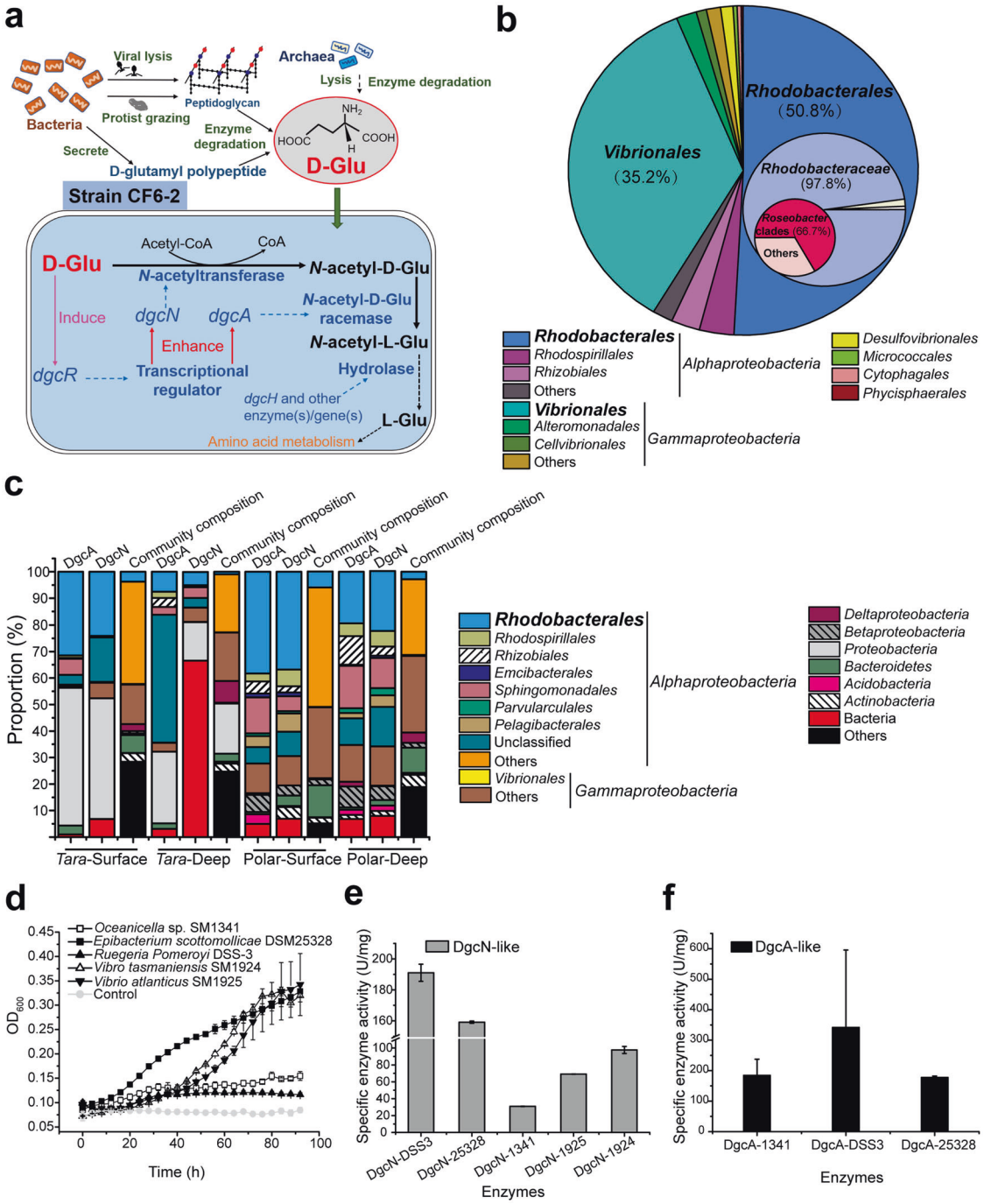

Fig. 7 Analyses of the diversity and distribution of DgcAN sequences and validation of the functions of the DgcA and DgcN homologs in marine bacteria. a D-Glu catabolism via the DgcN-DgcA pathway in strain CF6-2. Key enzymes are marked with blue color. The solid black arrows represent the reactions catalyzed by the key enzymes. The transcription of the gene *dgcR* can be induced by D-Glu, and DgcR positively regulates the transcription of genes *dgcA* and *dgcN*. When transported into the cell of strain CF6-2, D-Glu is catalyzed by DgcN and DgcA to generate *N*-acetyl-L-Glu, which is subsequently hydrolyzed by DgcH and other enzyme(s) to generate L-Glu. **b** Statistical analysis of marine bacteria containing both DgcA and DgcN sequences from marine bacterial genomes in the IMG/JGI database. **c** Relative percentage of DgcA and DgcN-sequences as well as microbial community composition of polar and *Tara* samples retrieved from the metagenomes in the polar and *Tara* Oceans Microbiome database. Polar samples were divided into two groups, Polar-Surface (0–100 m) and Polar-Deep (300-3800 m). *Tara* samples were divided into two groups, *Tara*-Surface (5-188 m) and *Tara*-Deep (250–1000 m). **d** Growth of 3 marine *Rhodobacterales* and 2 marine *Vibrionales* strains containing the *dgcAN* genes. The medium containing all components except for D-Glu was used as the control. **e** Detection of the *N*-acetyltransferase activity towards acetyl-CoA and D-Glu of the DgcN homologs. DgcN-DSS-3, DgcN-25328, DgcN-1341, DgcN-1925 and DgcN-1924 are from bacterial strains DSS-3, DSM25328, SM1341, SM1925 and SM1924 respectively. **f** Detection of the racemization activity to convert *N*-acetyl-D-Glu to *N*-acetyl-L-Glu of the DgcA homologs. DgcA-25328, DgcA-DSS-3 and DgcA-1341 are from bacterial strains DSM25328, DSS-3 and SM1341, respectively. The graphs in **d**, **e** and **f** show data from triplicate experiments (mean ± SD).

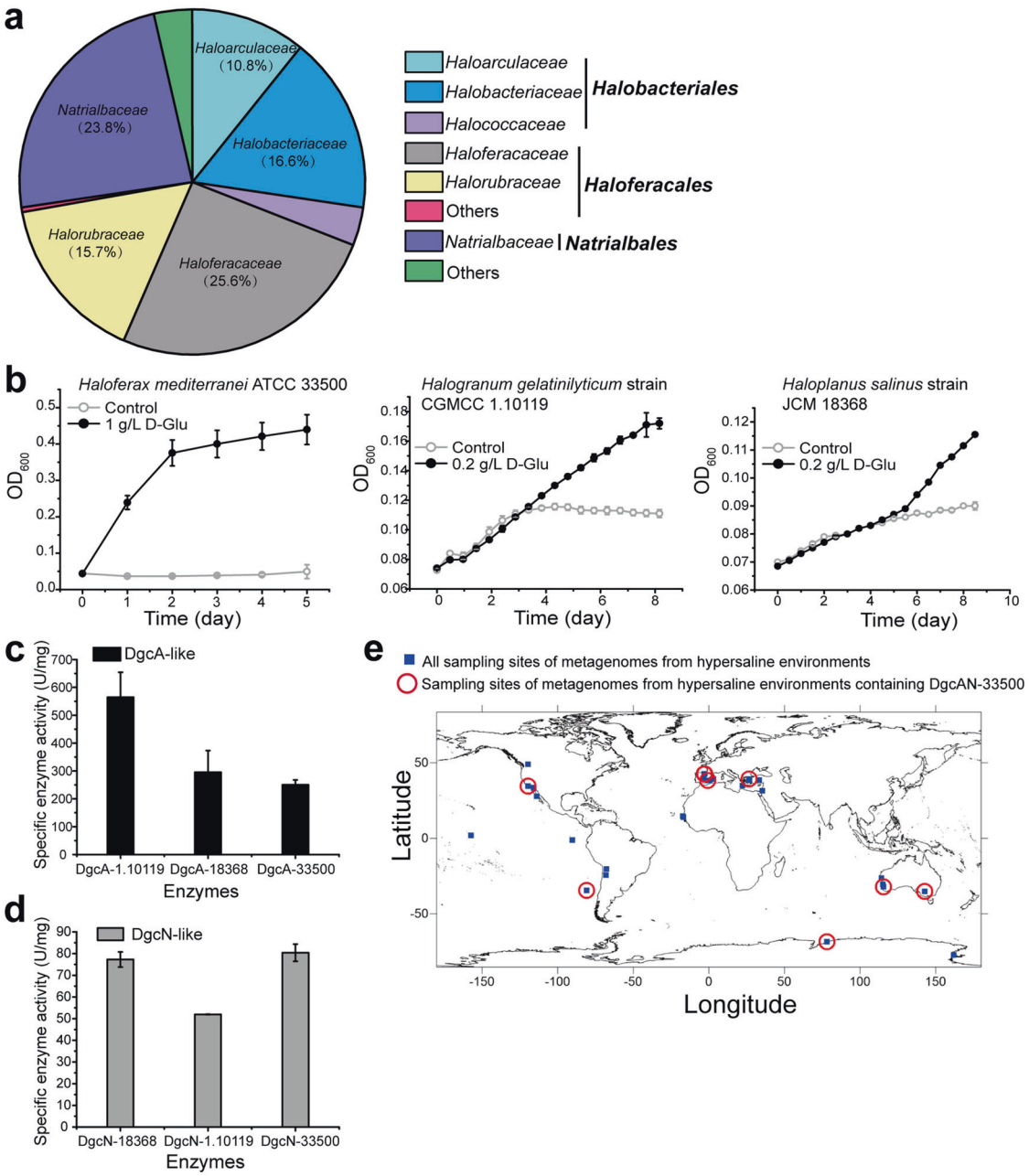

Fig. 8 Analyses of the diversity of archaea containing DgcAN homologs and validation of DgcA and DgcN enzyme activities from halophilic archaeal strains. a Statistical analysis of archaeal strains containing DgcAN sequences from the archaeal genomes in the IMG/JGI database. **b** Growth of 3 archaeal strains containing the *dgcAN* homologs. The medium containing all components except for D-Glu was used as the control. **c** Detection of the racemization activity of the DgcA homologs to convert *N*-acetyl-D-Glu to *N*-acetyl-L-Glu. **d** Detection of the *N*-acetyltransferase activity of the DgcN homologs towards acetyl-CoA and D-Glu. In c and d, DgcN-33500, DgcN-18368 and DgcN-1.10119 are from archaeal strains ATCC33500, JCM18368 and CGMCC1.10119, respectively. The graphs in b, c and d show data from triplicate experiments (mean ± SD). **e** The environmental distribution of archaeal DgcA and DgcN sequences in the hypersaline environment of the IMG/JGI database. Stations were plotted using Surfer version 12 (Golden Software LLC, USA).

We subsequently tested 3 halophilic archaeal strains (*Haloferax mediterranei* ATCC33500, *Halogranum gelatinilyticum* CGMCC1.10119, *Haloplanus salinus* JCM18368) for their ability to catabolize D-Glu. These three strains belong to class *Halobacteria* and are all predicted to encode DgcN-DgcA in their genomes. Indeed, all three strains can utilize D-Glu (Fig. 8b). Moreover, we successfully expressed and purified the recombinant DgcA and DgcN proteins from these three archaeal strains and all recombinant proteins had corresponding racemase activity (Fig. 8c), and *N*-acetyltransferase activity, respectively (Fig. 8d). These results indicated that halophilic archaeal strains containing the *dgcAN* genes can catabolize D-Glu most likely via the DgcN-DgcA pathway.

We further investigated the distribution of the archaeal DgcAN sequences in hypersaline environments. The IMG/JGI metagenomes from hypersaline environments were probed using DgcA-33500 and DgcN-33500 of strain ATCC33500 as the queries. Of the 48 sampling sites (297 samples) of metagenomes from hypersaline environments in the IMG database, 12 sampling sites (42 samples) contain both archaeal DgcA and DgcN sequences (Supplementary Table S10). As shown in Fig. 8e, these sampling

sites are distributed across the globe, including a deep lake in Antarctica, Ayvalik Saltern in Turkey, Lake Tyrrell in Australia, and Pena Hueca in Spain, showing that archaeal DgcAN sequences are distributed widely in global hypersaline environments. In addition, we further investigated the distribution of D-Glu dehydrogenase DguA sequences in the 223 DgcAN-containing archaeal genomes. However, no DguA sequence was found.

Although previous studies have confirmed the ability of archaea to use DAAs [13, 48–51], to the best of our knowledge, our results in this study is the first to demonstrate that halophilic archaea can catabolize D-Glu via this newly discovered DgcN-DgcA pathway. Interestingly, while DgcR was shown to be a positive regulator for the DgcN-DgcA pathway in strain CF6-2, homolog of *dgcR* adjacent to *dgcAN* was not found in either most bacteria or archaea. However, we do notice the presence of other common transcriptional factors genes near the *dgcAN* genes such as LuxR and AraC family transcriptional regulators in *dgcR*$^-$ bacterial genomes, which are likely involved in the transcriptional regulation of *dgcAN* in these bacteria. Phylogenetic analysis suggested that no horizontal gene transfer of genes *dgcAN* occurred between marine bacteria and halophilic archaea (Supplementary Fig. S8), implying that the functional genes (*dgcA* and *dgcN*) may have evolved independently in bacteria and archaea. The regulation of the D-Glu catabolism pathway in *dgcR*$^-$ genomes of marine bacteria and halophilic archaea awaits further investigation.

CONCLUSION

In this study, a novel deamination-independent D-Glu catabolic pathway, named the DgcN-DgcA pathway, was identified. The DgcN-DgcA pathway occurs widely in marine bacterial genomes and metagenomes, mostly adopted by *Rhodobacteraceae*, especially in the *Roseobacter* clade, in marine bacteria and *Halobacteria* in archaea. The results provide mechanistic insight into D-Glu catabolism in marine bacteria and halophilic archaea, shedding light on the recycling and mineralization of D-Glu in marine and hypersaline ecosystems.

DATA AVAILABILITY

All the RNA-seq read data have been deposited in NCBI's sequence read archive (SRA) under project accession number PRJNA829977. The structure of DgcN-25328 has been deposited in the PDB under the accession code 7XRJ.

REFERENCES

1. Park JT, Strominger JL. Mode of action of penicillin-biochemical basis for the mechanism of action of penicillin and for its selective toxicity. Science. 1957;125:99–101.
2. Yang CC, Leong J. Structure of pseudobactin 7sr1, a siderophore from a plant-deleterious *Pseudomonas*. Biochemistry. 1984;23:3534–40.
3. Ghuysen JM. Data on the structure of disaccharide-peptide complexes liberated from the wall of *Micrococcus lysodeikticus* by the action of beta(1-4)*N*-acetylhexosaminidases. Biochim Biophys Acta. 1961;47:561–8.
4. Ashiuchi M, Misono H. Biochemistry and molecular genetics of poly-gamma-glutamate synthesis. Appl Microbiol Biotechnol. 2002;59:9–14.
5. Jorgensen NO, Stepanaukas R, Pedersen AG, Hansen M, Nybroe O. Occurrence and degradation of peptidoglycan in aquatic environments. FEMS Microbiol Ecol. 2003;46:269–80.
6. Kandler O, Konig H. Cell wall polymers in Archaea (Archaebacteria). Cell Mol Life Sci. 1998;54:305–8.
7. Nagata Y, Tanaka K, Iida T, Kera Y, Yamada R, Nakajima Y, et al. Occurrence of D-amino acids in a few archaea and dehydrogenase activities in hyperthermophile *Pyrobaculum islandicum*. Biochim Biophys Acta 1999;1435:160–6.
8. Jorgensen NOG, Middelboe M. Occurrence and bacterial cycling of D amino acid isomers in an estuarine environment. Biogeochemistry. 2006;81:77–94.
9. Kubota T, Kobayashi T, Nunoura T, Maruyama F, Deguchi S. Enantioselective utilization of D-amino acids by deep-sea microorganisms. Front Microbiol. 2016;7:511.
10. Yu Y, Yang J, Zheng LY, Sheng Q, Li CY, Wang M, et al. Diversity of D-amino acid utilizing bacteria from Kongsfjorden, Arctic and the metabolic pathways for seven D-amino acids. Front Microbiol. 2020;10:2983.
11. Sakai A, Xiang DF, Xu C, Song L, Yew WS, Raushel FM, et al. Evolution of enzymatic activities in the enolase superfamily: *N*-succinylamino acid racemase and a new pathway for the irreversible conversion of D- to L-amino acids. Biochemistry. 2006;45:4455–62.
12. Uo T, Yoshimura T, Tanaka N, Takegawa K, Esaki N. Functional characterization of alanine racemase from *Schizosaccharomyces pombe*: a eucaryotic counterpart to bacterial alanine racemase. J Bacteriol. 2001;183:2226–33.
13. Moore BC, Leigh JA. Markerless mutagenesis in *Methanococcus maripaludis* demonstrates roles for alanine dehydrogenase, alanine racemase, and alanine permease. J Bacteriol. 2005;187:972–9.
14. Troy FA. Chemistry and biosynthesis of the poly(-D-glutamyl) capsule in *Bacillus licheniformis*. II. Characterization and structural properties of the enzymatically synthesized polymer. J Biol Chem. 1973;248:316–24.
15. Thorne CB, Gómez CG, Noyes HE, Housewright RD. Production of glutamyl polypeptide by *Bacillus subtilis*. J Bacteriol. 1954;68:307.
16. Pedersen AGU, Thomsen TR, Lomstein BA, Jorgensen NOG. Bacterial influence on amino acid enantiomerization in a coastal marine sediment. Limnol Oceanogr. 2001;46:1358–69.
17. Van Heijenoort J. Recent advances in the formation of the bacterial peptidoglycan monomer unit. Nat Prod Rep. 2001;18:503–19.
18. He WQ, Li GQ, Yang CK, Lu CD. Functional characterization of the *dguRABC* locus for D-Glu and D-Gln utilization in *Pseudomonas aeruginosa* PAO1. Microbiology. 2014;160:2331–40.
19. Tang BL, Yang J, Chen XL, Wang P, Zhao HL, Su HN, et al. A predator-prey interaction between a marine *Pseudoalteromonas* sp. and Gram-positive bacteria. Nat Commun. 2020;11:285.
20. Yu ZC, Tang BL, Zhao DL, Pang X, Qin QL, Zhou BC, et al. Development of a cold-adapted *Pseudoalteromonas* expression system for the *Pseudoalteromonas* proteins intractable for the *Escherichia coli* system. PLoS ONE. 2015;10:e0137384.
21. Yu ZC, Zhao DL, Ran LY, Mi ZH, Wu ZY, Pang X, et al. Development of a genetic system for the deep-sea psychrophilic bacterium *Pseudoalteromonas* sp. SM9913. Micro Cell Fact. 2014;13:13.
22. Zhu Y, Zhang P, Lu T, Wang X, Li A, Lu Y, et al. Impact of MtrA on phosphate metabolism genes and the response to altered phosphate conditions in *Streptomyces*. Environ Microbiol. 2021;23:6907–23.
23. Williams JW, Northrop DB. Kinetic mechanisms of gentamicin acetyltransferase I. Antibiotic-dependent shift from rapid to nonrapid equilibrium random mechanisms. J Biol Chem. 1978;253:5902–7.
24. Noda M, Matoba Y, Kumagai T, Sugiyama M. A novel assay method for an amino acid racemase reaction based on circular dichroism. Biochem J. 2005;389:491–6.
25. Doi E, Shibata D, Matoba T. Modified colorimetric ninhydrin methods for peptidase assay. Anal Biochem. 1981;118:173–84.
26. Mccoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007;40:658–74.
27. Cao SN, Zhang WP, Ding W, Wang M, Fan S, Yang B, et al. Structure and function of the Arctic and Antarctic marine microbiota as revealed by metagenomics. Microbiome. 2020;8:47.
28. Bansal M, Alm E, Kellis M. Efficient algorithms for the reconciliation problem with gene duplication, horizontal transfer and loss. Bioinformatics 2012;28:i283–91.
29. Henikoff S, Haughn GW, Calvo JM, Wallace JC. A large family of bacterial activator proteins. Proc Natl Acad Sci USA. 1988;85:6602–6.
30. Tyrrell R, Verschueren KH, Dodson EJ, Murshudov GN, Addy C, Wilkinson AJ. The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. Structure. 1997;5:1017–32.
31. Liu XX, Liu L, Song X, Wang GQ, Xiong ZQ, Xia YJ, et al. Determination of the regulatory network and function of the LysR-type transcriptional regulator of *Lactiplantibacillus plantarum*, LpLttR. Micro Cell Fact. 2022;21:65.
32. Eisfeld J, Kraus A, Ronge C, Jagst M, Brandenburg VB, Narberhaus F. A LysR-type transcriptional regulator controls the expression of numerous small RNAs in *Agrobacterium tumefaciens*. Mol Microbiol. 2021;116:126–39.
33. Liu XX, Xiong ZQ, Wang GQ, Wang LF, Xia YJ, Song X, et al. LysR family regulator LttR controls production of conjugated linoleic acid in *Lactobacillus plantarum* by directly activating the *cla* operon. Appl Environ Microbiol. 2021;87:e02798–20.
34. Schwartz CJ, Giel JL, Patschkowski T, Luther C, Ruzicka FJ, Beinert H, et al. IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins. Proc Natl Acad Sci USA. 2001;98:14895–900.
35. Ke Z, Zhou Y, Jiang W, Zhang M, Wang H, Ren Y, et al. McbG, a LysR family transcriptional regulator activates the *mcbBCDEF* gene cluster involved in the

upstream pathway of carbaryl degradation in *Pseudomonas* sp. XWY-1. Appl Environ Microbiol. 2021;87:e02970–20.
36. Schmidt DMZ, Hubbard BK, Gerlt JA. Evolution of enzymatic activities in the enolase superfamily: functional assignment of unknown proteins in *Bacillus subtilis* and *Escherichia coli* as L-Ala-D/L-Glu epimerases. Biochemistry. 2001;40:15707–15.
37. Vandecandelaere I, Nercessian O, Segaert E, Achouak W, Faimali M, Vandamme P. *Ruegeria scottomollicae* sp. nov., isolated from a marine electroactive biofilm. Int J Syst Evol Microbiol. 2008;58:2726–33.
38. Vetting MW, Hegde SS, Javid-Majd F, Blanchard JS, Roderick SL. Aminoglycoside 2′-*N*-acetyltransferase from *Mycobacterium tuberculosis* in complex with coenzyme A and aminoglycoside substrates. Nat Struct Biol. 2002;9:653–8.
39. Klein DC. Arylalkylamine *N*-acetyltransferase: "the timezyme". J Biol Chem. 2007;282:4233–7.
40. Mio T, Yamada-Okabe T, Arisawa M, Yamada-Okabe H. *Saccharomyces cerevisiae* GNA1, an essential gene encoding a novel acetyltransferase involved in UDP-*N*-acetylglucosamine synthesis. J Biol Chem. 1999;274:424–9.
41. Lee KK, Workman JL. Histone acetyltransferase complexes: one size doesn't fit all. Nat Rev Mol Cell Biol. 2007;8:284–95.
42. Forouhar F, Lee IN, Vujcic J, Vujcic S, Shen JW, Vorobiev SM, et al. Structural and functional evidence for *Bacillus subtilis* PaiA as a novel *N*-1-spermidine/spermine acetyltransferase. J Biol Chem. 2005;280:40328–36.
43. Vetting MW, Park CH, Hegde SS, Jacoby GA, Hooper DC, Blanchard JS. Mechanistic and structural analysis of aminoglycoside *N*-acetyltransferase AAC(6′)-Ib and its bifunctional, fluoroquinolone-active AAC(6′)-Ib-cr variant. Biochemistry. 2008;47:9825–35.
44. Krtenic B, Drazic A, Arnesen T, Reuter N. Classification and phylogeny for the annotation of novel eukaryotic GNAT acetyltransferases. PLoS Comput Biol. 2020;16:e1007988.
45. Arima J, Isoda Y, Hatanaka T, Mori N. Recombinant production and characterization of an *N*-acyl-D-amino acid amidohydrolase from *Streptomyces* sp. 64E6. World J Microbiol Biotechnol. 2013;29:899–906.
46. Tsai YC, Tseng CP, Hsiao KM, Chen LY. Production and purification of D-Aminoacylase from *Alcaligenes denitrificans* and taxonomic study of the strain. Appl Environ Microbiol. 1988;54:984–9.
47. Mendes MI, Smith DE, Pop A, Lennertz P, Fernandez Ojeda MR, Kanhai WA, et al. Clinically distinct phenotypes of Canavan disease correlate with residual aspartoacylase enzyme activity. Hum Mutat. 2017;38:524–31.
48. Kawakami R, Ohmori T, Sakuraba H, Ohshima T. Identification of a novel amino acid racemase from a hyperthermophilic archaeon *Pyrococcus horikoshii* OT-3 induced by D-amino acids. Amino Acids. 2015;47:1579–87.
49. Matsumoto M, Homma H, Long ZQ, Imai K, Iida T, Maruyama T, et al. Occurrence of free D-amino acids and aspartate racemases in hyperthermophilic archaea. J Bacteriol. 1999;181:6560–3.
50. Kawakami R, Ohshida T, Sakuraba H, Ohshima T. A novel PLP-dependent alanine/serine racemase from the hyperthermophilic archaeon *Pyrococcus horikoshii* OT-3. Front Microbiol. 2018;9:1481.
51. Satomura T, Sakuraba H, Suye S, Ohshima T. Dye-linked D-amino acid dehydrogenases: biochemical characteristics and applications in biotechnology. Appl Microbiol Biotechnol. 2015;99:9337–47.

ACKNOWLEDGEMENTS

We would like to thank Xiaoju Li and Haiyan Sui from Life Science General Research Technology Platform of SKLMT (State Key Laboratory of Microbial Technology, Shandong University) for the assistance in X-ray diffraction experiment. This work was supported by the National Science Foundation of China (U2006205, 91851205, 31961133016, U1706207), Major Scientific and Technological Innovation Project (MSTIP) of Shandong Province (2019JZZY010817), and the Program of Shandong for Taishan Scholars (tspd20181203).

AUTHOR CONTRIBUTIONS

YY performed all experiments. PW and H-YC helped in protein structure determination and molecular docking simulations. Z-JT helped in bioinformatic analysis. YZ helped in EMSAs experiments. YY and X-LC wrote the manuscript. MW, AM and YC revised the manuscript. HX helped in experiments. Y-ZZ, X-LC and Y-QZ designed the experiments and directed the study.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41396-023-01364-6.

Correspondence and requests for materials should be addressed to Yu-Zhong Zhang, Xiu-Lan Chen or Yu-Qiang Zhang.