

Growth of complete ammonia oxidizers on guanidine

<https://doi.org/10.1038/s41586-024-07832-z>

Received: 22 January 2023

Accepted: 15 July 2024

Published online: 14 August 2024

Open access

 Check for updates

Marton Palatinszky¹, Craig W. Herbold^{1,10}, Christopher J. Sedlacek¹, Dominic Pühringer^{2,3}, Katharina Kitzinger¹, Andrew T. Giguere¹, Kenneth Wasmund^{1,4,11}, Per H. Nielsen⁴, Morten K. D. Dueholm⁴, Nico Jehmlich⁵, Richard Gruseck^{1,6}, Anton Legin⁷, Julius Kostan^{2,3}, Nesrete Krasnici^{2,3}, Claudia Schreiner^{2,3}, Johanna Palmetzhofer^{1,6}, Thilo Hofmann¹, Michael Zumstein¹, Kristina Djinović-Carugo^{2,3,8,9}, Holger Daims^{1,8} & Michael Wagner^{1,4,8}✉

Guanidine is a chemically stable nitrogen compound that is excreted in human urine and is widely used in manufacturing of plastics, as a flame retardant and as a component of propellants, and is well known as a protein denaturant in biochemistry^{1–3}. Guanidine occurs widely in nature and is used by several microorganisms as a nitrogen source, but microorganisms growing on guanidine as the only substrate have not yet been identified. Here we show that the complete ammonia oxidizer (comammox) *Nitrospira inopinata* and probably most other comammox microorganisms can grow on guanidine as the sole source of energy, reductant and nitrogen. Proteomics, enzyme kinetics and the crystal structure of a *N. inopinata* guanidinase homologue demonstrated that it is a bona fide guanidinase. Incubation experiments with comammox-containing agricultural soil and wastewater treatment plant microbiomes suggested that guanidine serves as substrate for nitrification in the environment. The identification of guanidine as a growth substrate for comammox shows an unexpected niche of these globally important nitrifiers and offers opportunities for their isolation.

Recently, microbial guanidine metabolism has received a lot of attention, as the identification of four different classes of riboswitches (RNA elements that bind metabolites or metal ions as ligands and regulate mRNA expression) selective for guanidine in the genomes of many bacteria in various phyla suggested that guanidine is an important metabolite of microorganisms. However, the microbial pathways for the formation and degradation of this compound and its ecological importance largely remain to be explored^{4–8}. During the past few years, bacterial guanidine production by at least three pathways has been demonstrated. Some bacteria encode an ethylene-forming enzyme that has an important role in bioethylene production and produces guanidine from arginine and 2-oxoglutarate⁹. Furthermore, during synthesis of the antibiotic naphthyridinomycin by *Streptomyces luitanensis*, the arginine-4,5-desaturase NapI leads to guanidine formation as a side reaction¹⁰. Moreover, bacteria can transform guanylurea to ammonia and guanidine¹¹. This is particularly important as guanylurea is one of the most common contaminants in nature formed by degradation of the type 2 diabetes drug metformin (a biguanidine and one of the most prescribed drugs globally) and of the fertilizer additive cyanoguanidine (dicyandiamide)^{11–13}. However, additional guanidine-forming mechanisms must exist in bacteria, as guanidine has been detected under nutrient-poor conditions in *Escherichia coli*

lacking the aforementioned pathways⁴. Notably, in plants and algae, guanidine is also produced by homoarginine-6-hydroxylases^{10,14}.

Our knowledge of guanidine degradation by microorganisms is also still in its infancy, but recent research has revealed two degradation pathways for guanidine that are used by bacteria to not only detoxify guanidine but also to use it as a nitrogen source and, ultimately, to produce three molecules of ammonia from one molecule of guanidine. A widespread sequential decomposition pathway (Fig. 1a) involves a biotin-containing and ATP-dependent guanidine carboxylase, a heteromeric carboxyguanidine deiminase and an allophanate hydrolase to convert guanidine to ammonia and CO₂^{15,16}. In 2021, a Ni-containing guanidinase, mediating the direct hydrolysis of free guanidine to urea (which is further converted by urease to ammonia and CO₂) and ammonia, was described in a *Synechocystis* strain (Fig. 1b). This guanidinase has a crucial role in detoxifying guanidine during bioethylene production and enables this cyanobacterium to tap the guanidine pool as a nitrogen source without spending ATP for its degradation^{17,18}.

Guanidine use by nitrifying microorganisms

Chemolithoautotrophic nitrifying microorganisms catalyse the aerobic oxidation of ammonia through nitrite to nitrate and therefore have

¹Centre for Microbiology and Environmental Systems Science, University of Vienna, Vienna, Austria. ²Department of Structural and Computational Biology, Center for Molecular Biology, University of Vienna, Vienna, Austria. ³Max Perutz Labs, Vienna Biocenter Campus (VBC), Vienna, Austria. ⁴Center for Microbial Communities, Department of Chemistry and Bioscience, Aalborg University, Aalborg, Denmark. ⁵Helmholtz-Centre for Environmental Research-UFZ, Department of Molecular Systems Biology, Leipzig, Germany. ⁶Doctoral School in Microbiology and Environmental Science, University of Vienna, Vienna, Austria. ⁷Institute of Inorganic Chemistry, Faculty of Chemistry, University of Vienna, Vienna, Austria. ⁸The Comammox Research Platform, University of Vienna, Vienna, Austria. ⁹European Molecular Biology Laboratory (EMBL), Grenoble, France. ¹⁰Present address: Te Kura Pūtaiao Kōiora (School of Biological Sciences), Te Whare Wānanga o Waitaha (University of Canterbury), Ōtautahi (Christchurch), Aotearoa New Zealand. ¹¹Present address: School of Biological Sciences, University of Portsmouth, Portsmouth, UK.

✉e-mail: michael.wagner@univie.ac.at

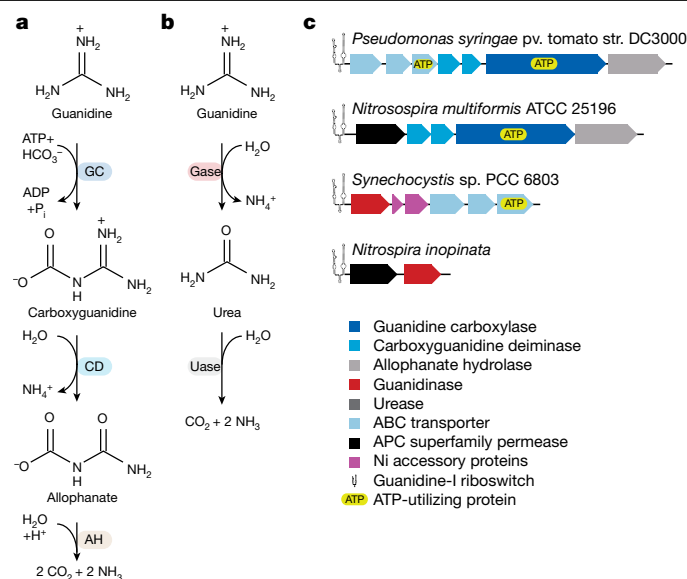


Fig. 1 | Pathways and genes involved in guanidine degradation. **a**, The guanidine carboxylase pathway. AH, allophanate hydrolase; CD, carboxyguanidine deiminase; Gase, guanidinase; GC, guanidine carboxylase; P_i , inorganic phosphate; Uase, urease. **b**, The guanidinase pathway. **c**, Arrangements of genes encoding proteins involved in guanidine degradation under the regulation of type I guanidine riboswitches in the comammox microorganism *N. inopinata*, the AOB *Nitrospira multiformis* and two bacterial model organisms for guanidine catabolism. Each of the four genomes also encodes urease (not displayed) at locations that are not regulated by guanidine riboswitches.

an essential role in global biogeochemical nitrogen cycling, lead to massive fertilizer loss in agriculture, contribute to the emission of the potent greenhouse gas and ozone-depleting substance nitrous oxide, and are essential for nutrient removal in wastewater treatment plants (WWTPs)^{19,20}. In addition to ammonia-oxidizing bacteria (AOB) and archaea (AOA), which convert ammonia to nitrite (that is subsequently oxidized to nitrate by nitrite-oxidizing bacteria), the recently identified complete ammonia oxidizers (comammox) of the genus *Nitrospira* use ammonia as a substrate and oxidize it through nitrite to nitrate^{21,22}. In contrast to canonical nitrite oxidizers of the genus *Nitrospira*^{23,24}, the recognized range of energy sources used by ammonia-oxidizing microorganisms is very narrow. In addition to ammonia, only urea and cyanate have been experimentally confirmed to support the growth of some ammonia-oxidizing microorganisms by serving as an ammonia source after being converted enzymatically by cytoplasmic ureases and cyanases, respectively^{25–27}, although meta-omic studies of comammox strains in WWTPs suggested a more pronounced metabolic versatility²⁸.

We wondered whether ammonia-oxidizing microorganisms might be able to exploit guanidine as a source of energy, reductant and nitrogen. We therefore screened all available genomes from this guild for genes possibly encoding enzymes involved in the two known pathways that would enable guanidine to be used as an ammonia source (Fig. 1 and Supplementary Table 1). We found genes for guanidine degradation in most betaproteobacterial AOB (135 out of 145 genomes) and comammox strains (76 out of 83 genomes) but none within the gammaproteobacterial AOB or the various members of the AOA. Notably, most ammonia oxidizers encoding enzymes involved in guanidine utilization also possess an amino acid/polyamine/organocation permease (APC superfamily) in the same genomic region, which is predicted to enable the import of guanidine without using ATP (possibly through proton symport; Fig. 1). By contrast, *Pseudomonas syringae* pv. tomato str. DC3000 and *Synechocystis* sp. PCC 6803 (the model organisms for the guanidine carboxylase and guanidinase pathways, respectively)

encode an ATP-dependent ABC transporter for guanidine transport in the neighbourhood of the genes encoding either guanidine utilization pathway^{15,17,18}. Furthermore, the betaproteobacterial AOB exclusively possess genes for the ATP-requiring guanidine carboxylase pathway. By contrast, only two comammox genomes encode this pathway, whereas the majority of comammox genomes instead are predicted to be equipped with the more energy-efficient pathway using a Ni-containing guanidinase (Extended Data Fig. 1). Most comammox genomes and metagenome-assembled genomes (MAGs; 75 out of 83) also encode a urease for conversion of urea formed by guanidinase to ammonia. Taken together, ammonia oxidizers equipped for guanidine metabolism do not rely on an ATP-consuming guanidine transporter and comammox microorganisms additionally use an ATP-independent guanidinase. Thus, these ammonia oxidizers and, in particular, comammox, if indeed capable of using guanidine, would be much more energy efficient than the previously characterized *Pseudomonas* and *Synechocystis* strains (that use guanidine as a nitrogen source) in converting guanidine to three molecules of ammonia.

With the distribution patterns of guanidine utilization genes in mind, we tested the guanidine metabolism ability using equally dense cultures of five AOB strains that possess the complete gene set of the pathway (one lacking only the urease), and the only described comammox pure culture *N. inopinata*. Notably, after incubation for 2 weeks with 50 μ M guanidine as the sole substrate, only *N. inopinata* was able to almost fully degrade guanidine as a sole substrate and produce nitrite and nitrate. Three of the AOB strains were able to convert a fraction of the added guanidine to nitrite but only in the concomitant presence of ammonium (Extended Data Fig. 2).

Consistent with a guanidine metabolism, a guanidine-I riboswitch (*ykkC-yxkD*) is found immediately upstream of the APC superfamily permease and guanidinase genes in *N. inopinata* and many other comammox microorganisms (Fig. 1c, Extended Data Fig. 1 and Supplementary Tables 2 and 3). In comammox organisms, ABC transporters and, unexpectedly, also Ni chaperones were absent from the genomic regions surrounding the guanidine-I riboswitch. This riboswitch class acts as a transcriptional suppressor, permitting transcription of downstream genes only in the presence of guanidine⁴. Phylogenetic analyses revealed that the comammox guanidinases belong to the ureohydrolase enzyme superfamily (Fig. 2a). Substrates converted by this enzyme family typically contain a guanidine moiety, which is hydrolysed to release urea. In the guanidinase from *Synechocystis* sp. PCC 6803, specificity to guanidine was dependent on two amino acids, threonine at position 97 and tryptophan at position 305 (ref. 18). The guanidinase from *N. inopinata* belongs to the same clade of this superfamily and also possesses threonine and tryptophan at the key sites for guanidine specificity (Fig. 2a,b). Comparison of the comammox guanidinase phylogenetic tree with the comammox ammonia monooxygenase tree revealed a clear separation between the comammox A and B groups, with several subclades of comammox group A also showing similar clustering and approximate branching order in the trees (Fig. 2c). These phylogenetic consistencies, and the very widespread distribution of putative guanidinases among comammox microorganisms, suggest that guanidine use is probably an ancient and persistent trait of comammox organisms. Together with the observation that most comammox strains contain the enzymatic repertoire for guanidine degradation, this strongly indicates that the ability for guanidine degradation confers a selective advantage to comammox strains and has therefore been retained in their genomes during evolution.

Growth of *N. inopinata* on guanidine

To further characterize guanidine use by *N. inopinata*, and to test whether this comammox organism can grow on guanidine as the sole source of energy, reductant and nitrogen, batch incubation experiments were performed. For quantification of the guanidine

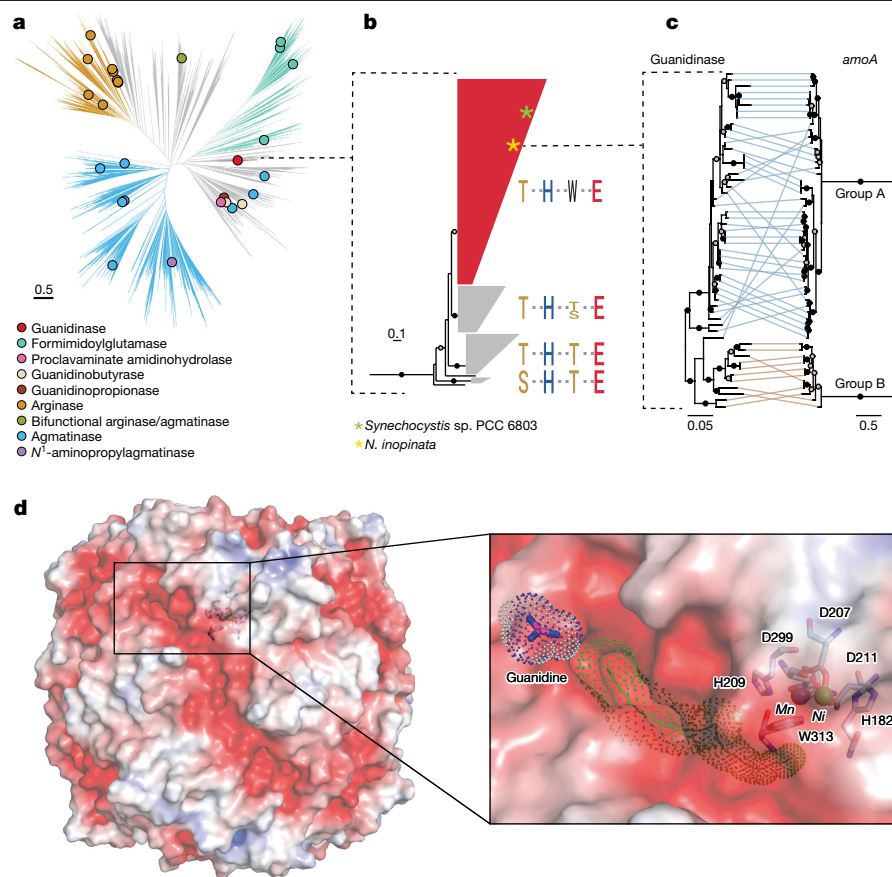


Fig. 2 | Phylogeny and structure of comammox guanidinases. a, Phylogeny of the ureohydrolase superfamily. The circles indicate functionally characterized members. Characterized *N*¹-aminopropylagmatinases were included in the indicated agmatinase clade. Branches in grey do not correspond to any known function or are not monophyletic for function. **b**, Simplified phylogeny of guanidinases, using biochemically characterized arginase family members as the outgroup (a full tree is shown in Extended Data Fig. 4a). Specific amino acid residues that might be important for guanidine catalysis according to previous studies^{15,17,18} (*N. inopinata* guanidinase positions Thr105, His222, Trp313 and Glu344; Supplementary Table 4) are indicated. **c**, Comparison of the

phylogenetic relationships of comammox guanidinases (amino acids) and ammonia monooxygenases (*amoA* nucleotides). Tree tips are connected for genes found in the same genome. **d**, Electrostatic surface representation of guanidinase (scaled from −5 kT/e (red) to +5 kT/e (blue)), with the suggested entry of a tunnel towards the active site highlighted (left). Right, magnification of the region indicated by a box. The tunnel as determined by CAVER⁵⁷ is shown as a green line, and its width is indicated by a dot representation. Active-site residues are shown as sticks, and nickel (Ni) and manganese (Mn) ions are shown as green and purple spheres, respectively. Guanidine is shown as sticks with corresponding van der Waals atomic radii indicated as dots.

concentrations, we adapted an analytical workflow based on derivatization of guanidine with benzoin²⁹. In addition to the previously reported derivatization product, our analysis revealed an additional product that showed a higher signal intensity. We optimized the derivatization protocol to increase the fraction of this product and thereby improve the analytical sensitivity, confirmed the specificity of the derivatization (Supplementary Fig. 1 and Supplementary Table 5) and then used this protocol (with calibration solutions containing 1–100 μM guanidine) for all measurements.

In batch incubation experiments with around 50 μM of guanidine, *N. inopinata* converted guanidine stoichiometrically to nitrite and nitrate (Fig. 3 and Extended Data Fig. 3f). No abiotic guanidine degradation was detectable with dead *N. inopinata* biomass (Fig. 3a). During incubation with guanidine, *N. inopinata* cell numbers increased significantly in comparison to the controls without guanidine (Fig. 3c). Furthermore, *N. inopinata* assimilated nitrogen and carbon from isotopically labelled guanidine and bicarbonate, clearly demonstrating chemolithoautotrophic growth of *N. inopinata* on guanidine as a sole substrate (Fig. 3f,g). *N. inopinata* growth on guanidine was slower compared with growth on ammonium with maximum growth rates (division rates) of 0.076 and 0.632 d^{−1}, respectively (Fig. 3c, Extended Data Fig. 3c and Supplementary Table 6), and nitrite/nitrate production in the presence of both ammonium and guanidine was slightly (but significantly)

slower compared with ammonium only (Extended Data Fig. 3g), yet the cell yield (that is, the number of cells formed per mol of combined nitrite and nitrate produced) did not differ between treatments receiving guanidine, ammonium or both substrates (Extended Data Fig. 3i). Notably, when both ammonium and guanidine were supplied, both were used simultaneously, rather than in a diauxic growth pattern (Extended Data Fig. 3d). Guanidine utilization rates per cell did not differ between treatments receiving only guanidine, or both ammonium and guanidine (Extended Data Fig. 3h). These observations imply that guanidine is used by comammox microorganisms whenever it becomes available, consistent with the ‘on/off’ type regulation of guanidinase expression by the guanidine-dependent riboswitch.

Protein expression of *N. inopinata* grown on guanidine was compared to growth on ammonium. Pure cultures were repeatedly fed with either 50 μM guanidine or 150 μM ammonium after substrate depletion. After incubation for 3 weeks, three proteins showed significant differential expression between the guanidine and ammonium treatments: guanidinase (\log_{10} [fold change (FC)] = 1.2, adjusted P (P_{adj}) = 0.001), a putative fatty acid methyltransferase (FAMT, \log_{10} [FC] = 1.1, P_{adj} = 0.036) and a putative RNA-binding protein (RbpB, \log_{10} [FC] = 2.5, P_{adj} = 0.036) (Extended Data Fig. 4b). All three showed higher expression levels in the guanidine incubation than in the ammonium incubation, with guanidinase showing the highest significance

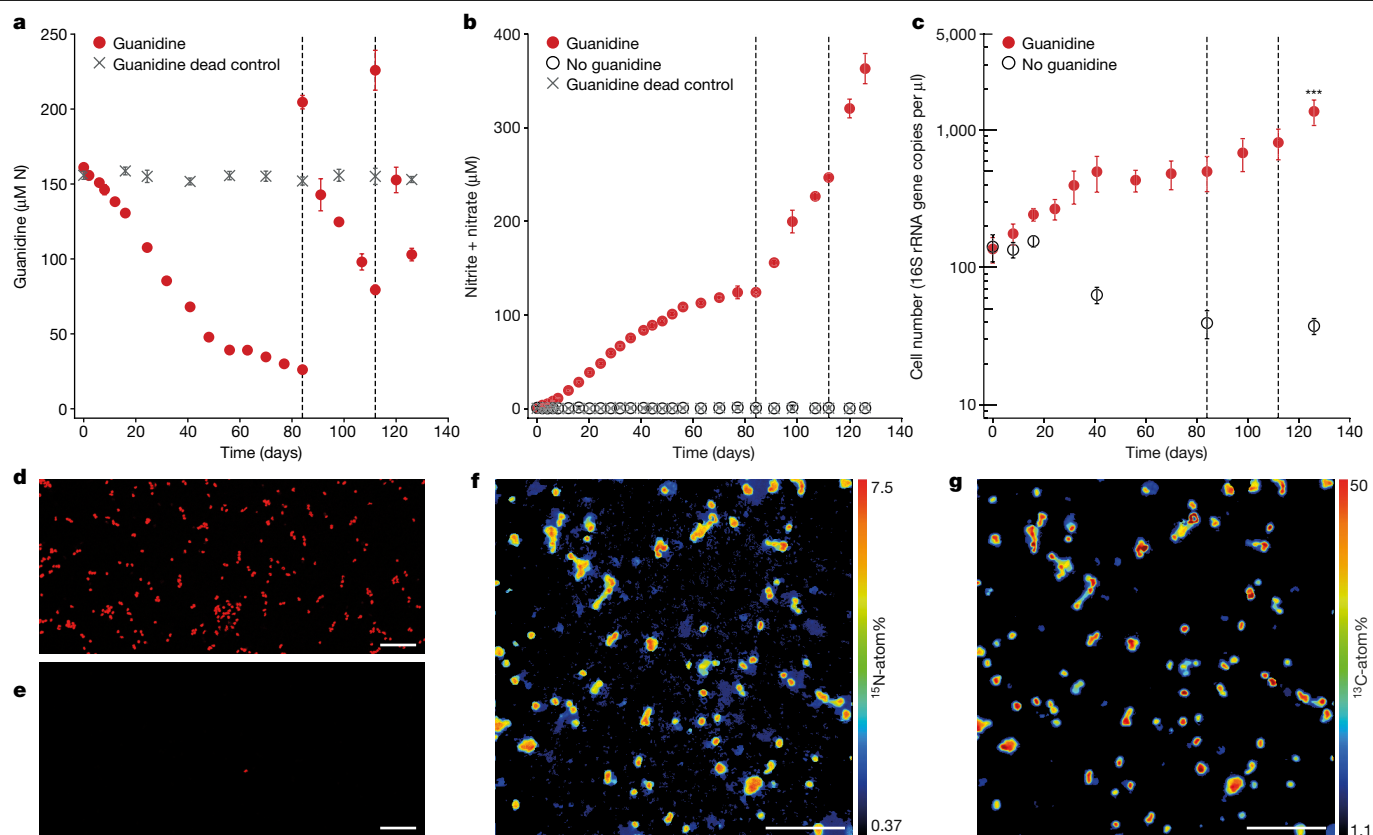


Fig. 3 | Growth of *N. inopinata* on guanidine as the sole source of energy, reductant and nitrogen. a, Biodegradation of guanidine over time. Around 50 μM (150 $\mu\text{M N}$) isotopically labelled guanidine was added to a washed culture of *N. inopinata* (after pre-incubation with guanidine and ammonium for 1 month) and incubated for 126 days. The control for abiotic guanidine decay was performed using autoclaved *N. inopinata* biomass. On days 84 and 112, additional spikes of around 50 μM of guanidine were added to living biomass incubations (dashed grey lines). Ammonium never increased above the level of detection (5 μM) and urea concentrations remained below 2.5 μM (Supplementary Table 6). **b**, NO_2^- and NO_3^- production (combined). At the end of the incubation, 78% of the total added guanidine nitrogen was oxidized to NO_2^- and NO_3^- . Nitrogen balances are shown in Extended Data Fig. 3f. **c**, qPCR analysis of 16S rRNA gene copy numbers. Statistical analysis was performed using Welch two-sample *t*-tests; ****P* = 0.0049, *t* = 10.348, d.f. = 4, comparing

N. inopinata cell numbers incubated with guanidine versus without guanidine at the 126 day timepoint. For **a–c**, data are mean \pm s.d. across five biological replicates. **d, e**, Representative images of DAPI-stained *N. inopinata* cells (red, 10 ml culture filtered onto 0.2 μm GTP filter) after 107 days of incubation with guanidine (**d**) and without guanidine (**e**). The same results were observed for all five biological replicates. **f, g**, Nitrogen (**f**) and corresponding carbon (**g**) isotopic enrichment of *N. inopinata* cells after 107 days of incubation with ^{15}N -guanidine and ^{13}C -bicarbonate as measured using nanoSIMS. The ^{13}C -enrichment is lowered by dilution of the ^{13}C -bicarbonate with CO_2 from the headspace air, CO_2 formation from pyruvate and breakdown of isotopically unlabelled guanidine in the medium. Data from one biological replicate are shown; a second replicate was measured with the same results. Scale bars, 10 μm (**d–g**).

(*P*_{adj} = 0.001). The APC permease, the putative guanidine transporter of *N. inopinata*, was detected in only one ammonium-treatment replicate, but was observed in 4 out of 5 guanidine-treatment replicates (Extended Data Fig. 4d). The APC permease and guanidinase are both under the control of the *ykkC-yxD* riboswitch (Fig. 1c) and, therefore, both were expected to respond to guanidine in the same manner. However, the APC superfamily permease as an integral membrane protein is difficult to detect using proteomic approaches despite our use of protocols to maximize the recovery of such proteins. Neither FAMT nor RbpB are in the vicinity of a guanidine-dependent riboswitch in *N. inopinata*. Both were detected as differentially expressed with marginal significance (*P*_{adj} = 0.036) and were not pursued further.

Comammox guanidinase characterization

The guanidinase enzyme of *N. inopinata* was expressed in *E. coli*, purified to homogeneity and characterized as a homohexamer with a molecular mass of 240 kDa (Extended Data Fig. 5a), resembling the recently discovered guanidinase of *Synechocystis* sp. PCC6803 (GdmH)¹⁸. Crystal structure analyses of the heterologously expressed *N. inopinata*

guanidinase (resolution of 1.58 Å) revealed that it exhibits the same α - β - α fold of the arginase subfamily as its cyanobacterial homologue¹⁸, but with each subunit (Fig. 2d and Extended Data Fig. 6) containing one nickel and one manganese ion with the overall structure being highly similar to the *Synechocystis* enzyme with an root mean squared deviation of 0.74 Å over an aligned length of 361 C α atoms. However, the *N. inopinata* guanidinase possesses an N-terminal extension, which is lacking in the cyanobacterial enzyme (Extended Data Fig. 6b,c). This extension stabilizes the tertiary and quaternary structure of the guanidinase through extensive interactions, consistent with its notably high thermostability and activity in a wide pH range (Extended Data Fig. 5b,d and Supplementary Table 7). On the other hand, in comparison to the cyanobacterial enzyme, the *N. inopinata* guanidinase lacks a C-terminal extension (Supplementary Table 8 and Extended Data Fig. 6f).

In the active site of the heterologously expressed *N. inopinata* guanidinase, one nickel and one manganese ion are coordinated through conserved histidine and aspartic acid residues (Fig. 2d, Extended Data Fig. 6 and Supplementary Table 9). The positions and identity of metal sites were determined by the diffraction anomalous dispersion signal (Extended Data Fig. 6h,i) and corroborated by inductively coupled

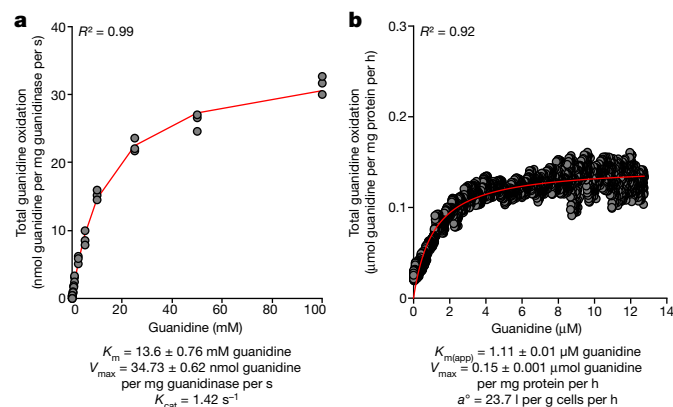


Fig. 4 | Guanidine oxidation kinetics of purified guanidinas and of *N. inopinata* cells. **a**, Kinetic characterization of the heterologously expressed *N. inopinata* guanidinas. A Michaelis–Menten model (red line) was fit to triplicate guanidine consumption rates by the guanidinas (pH 7.5, 37 °C) and used to determine the K_m and V_{max} . **b**, The whole-cell guanidine oxidation rates of *N. inopinata* were determined for cells that were pre-induced with guanidine for around 12 h with microsensor measurements (two additional biological replicates are shown in Extended Data Fig. 7a,b). A Michaelis–Menten model (red line) was used to determine the apparent half-saturation constant ($K_{m(app)}$), V_{max} and substrate specific affinity (a°) based on guanidine oxidation rates.

plasma mass spectrometry (ICP-MS) metal analysis (Extended Data Table 1). Another notable difference between the active sites of the *Synechocystis* and the *N. inopinata* enzymes is a tilt of *N. inopinata* guanidinas Trp313 towards the active-site nickel ion. The tilted Trp313 in *N. inopinata* guanidinas permits access to the active site from the surface, in contrast to Trp305 of the cyanobacterial enzyme, which blocks it (Extended Data Fig. 6g). The hexamer's ridges display a negative electrostatic potential, and each cavity marks the start of a 17 Å negatively charged tunnel towards the active site (Fig. 2d and Extended Data Fig. 6e). This is suggestive of electrostatic attraction of the positively charged substrate guanidinium (the predominant form of guanidine under physiological conditions).

Kinetic analysis of the purified guanidinas of *N. inopinata* revealed a substrate affinity (K_m) of 13.6 ± 0.76 mM for guanidine at pH 7.5 and a temperature of 37 °C (Fig. 4a) when expressed in the presence of 1 mM NiSO_4 , highly similar to the guanidinas of the cyanobacterium *Synechocystis* sp. PCC6803^{17,18}, although some kinetic properties were dependent on the nickel concentration used during heterologous expression (Extended Data Table 1). Moreover, the maximum reaction rate (V_{max}), enzyme turnover (K_{cat}) and enzymatic catalytic efficiency (K_{cat}/K_m) are all highly similar to the values obtained in a previous study¹⁷ after guanidinas overexpression in *Synechocystis*, but much lower than the reported values of the latter guanidinas overexpressed along with nickel-loading chaperones in *E. coli*¹⁸ (Extended Data Table 1). By contrast, the nickel-loading mechanism of *N. inopinata* is unclear as no such chaperones were found in the genome and, therefore, could not be co-expressed with the guanidinas in *E. coli*. Thus, it is tempting to speculate that the lower V_{max} and K_{cat} values of the heterologously expressed *N. inopinata* guanidinas reflect that the enzyme was not completely loaded with metals. This is consistent with an inferred average occupancy of 0.33 for nickel ions (0.66 per subunit) in the *N. inopinata* enzyme and with the ICP-MS data of the purified enzyme, which revealed a stoichiometry of 0.42 nickel ions per subunit (Extended Data Table 1). An alternative explanation for the low affinity for guanidine could be that the enzyme of *N. inopinata* converts guanidine as a side reaction and is actually specialized on another substrate (which would require in vivo its recognition by the riboswitch, which has a very high substrate specificity and has been shown to not recognize metabolites with a guanidino group³⁰). However, consistent with previous work on

the cyanobacterial guanidinas¹⁸, we could not detect urea formation by the comammox guanidinas from the putative alternative guanidine compounds methylguanidine, arginine, creatine, guanidino-butyrates and guanidino-propionate. Only for agmatine, a very minor urea production was detected that was much lower than the enzyme activity measured with guanidine as substrate (Extended Data Fig. 5e). This lack of activity for the alternative substrates aligns well with structural analysis, which indicates that the small entry channel can accommodate guanidine only when accounting for dynamic channel 'breathing'. The calculated molecular volumes for the potential alternative substrates are 127% to 287% larger than that of guanidine³¹.

We also determined the whole-cell kinetic properties of *N. inopinata* using guanidine through substrate-dependent O_2 consumption rates. This analysis reflects more accurately the in vivo ability of this organism to convert guanidine, as it assesses the entire pathway from guanidine uptake and degradation through ammonia and nitrite oxidation to terminal oxidase activity in the actual cellular environment. Guanidine oxidation rates increased with increasing guanidine concentrations in a Michaelis–Menten kinetic profile, which was used to model the apparent substrate affinity (mean $K_{m(app)}$, 1.34 ± 0.25 μM guanidine; $n = 3$) and maximum guanidine oxidation rate (mean V_{max} , 0.16 ± 0.02 μM guanidine per mg protein per h; $n = 3$) (Fig. 4b and Extended Data Fig. 7a,b). The whole-cell affinity for guanidine was much higher than the affinity of the purified heterologously expressed enzyme. This might, for example, reflect a more efficient metal loading of the enzyme in vivo, differences in post-translational modification in the natural host or crowding effects in the cytoplasm. At guanidine concentrations above around 400 μM, which probably do not occur in environments in which comammox thrive, and which are much higher than the guanidine concentrations used for the growth experiments, additional guanidine resulted in a decreased rate of respiratory activity. Guanidine inhibition at high concentrations was not specifically related to inhibition of the guanidinas (see the enzyme kinetic experiments) and also occurred when guanidine was added to ammonia-oxidizing *N. inopinata* cells (Extended Data Fig. 7c,d). The ability of *N. inopinata* to scavenge guanidine from a dilute pool, such as in an environmental setting, can be assessed by its substrate-specific affinity (a°), incorporating the whole-cell $K_{m(app)}$ and V_{max} (ref. 32). *N. inopinata* has a mean a° for guanidine of 21.6 l per g cells per h. In comparison, when pregrown under the same conditions, *N. inopinata* has a comparable mean a° of 74.5 l per g cells per h for urea (although a higher a° for this substrate has been reported recently³³) and an a° of 528 to 2,262 l per g cells per h for total ammonium (Extended Data Fig. 7e–h). This highlights the hierarchy of substrate acquisition and use by *N. inopinata*. Although the whole-cell affinity of *N. inopinata* for guanidine is by two orders of magnitude below its affinity for NH_3 , it still resembles or exceeds the whole-cell affinities for NH_3 of several terrestrial AOA from the *Nitrososphaerales* and *Nitrosocaldales* phylogenetic lineages and of many AOB³⁴.

Comammox use guanidine in WWTPs

Nitrification is an essential step for the efficient removal of nitrogen compounds in WWTPs. Comammox thrives in some but not all of these systems. As guanidine has been detected in human urine at concentrations between 2 and 20 μM (refs. 35,36), and we detected guanidine in the influent of a municipal WWTP (concentration, 0.5 μM; Extended Data Table 2), we hypothesized that comammox microorganisms in WWTPs degrade guanidine under competitive conditions and studied how they respond to guanidine pulses. For our experiments, we selected two municipal WWTPs (Ribe and Haderslev, Denmark), in which comammox *Nitrospira* were abundant nitrifiers (Extended Data Fig. 8a–d). A previous metagenomic analysis had reconstructed an abundant comammox MAG from each of the two WWTPs³⁷. Both comammox MAGs contained the guanidine-I riboswitch

upstream of the APC superfamily permease and guanidinase genes (Extended Data Fig. 1). After addition of 50 μM guanidine to starved biomass from these WWTPs, considerable guanidine degradation was observed, whereas a much slower degradation rate occurred in a control experiment with biomass from an Austrian WWTP in which comammox could not be detected by PCR³⁸ (Extended Data Fig. 8e–g). Activated sludge from the Ribe WWTP was additionally incubated with urea or ammonium and one replicate of each substrate amendment experiment was analysed by metatranscriptomics, using the different timepoints as replicates for comparison. At $T = 0$, baseline transcription of the guanidinase gene was 227 transcripts per million (TPM) in the Ribe sludge. In the Ribe WWTP, only the comammox guanidinase ($\log_2[\text{FC}] = 2.72$; $P_{\text{adj}} = 1.3 \times 10^{-8}$) and APC superfamily permease ($\log_2[\text{FC}] = 2.62$; $P_{\text{adj}} = 7.5 \times 10^{-11}$) were more highly transcribed in the guanidine-spiked incubation compared with in the ammonia-spiked incubation, demonstrating the functionality of the riboswitch. Three genes were more highly transcribed in the guanidine-spiked incubation compared with in the urea-spiked incubation: guanidinase ($\log_2[\text{FC}] = 2.63$; $P_{\text{adj}} = 8.2 \times 10^{-8}$), APC superfamily permease ($\log_2[\text{FC}] = 2.57$; $P_{\text{adj}} = 3.6 \times 10^{-10}$) and a CBS-domain-containing protein with unknown function (MBK9946601.1: $\log_2[\text{FC}] = 2.19$; $P_{\text{adj}} = 3.9 \times 10^{-3}$). High transcriptional levels of guanidinase and APC superfamily permease in the guanidine-spiked incubation were maintained throughout the experiment compared with the ammonia-spiked and urea-spiked incubations, and incubations with no experimental substrate addition (Fig. 5a), consistent with a high concentration of guanidine ($>40 \mu\text{M}$) over the time course (Fig. 5c and Extended Data Fig. 8f).

Similar metatranscriptomic experiments were performed with biomass from the Haderslev WWTP, but no ammonia- or urea-spiked incubations were included, and the guanidine-spiked incubation was compared to a no-substrate control. At $T = 0$, the baseline transcription of the guanidinase gene was 316 TPM in the Haderslev sludge. In total, 373 genes were found to be differentially transcribed, including the guanidinase ($\log_2[\text{FC}] = 1.91$; $P_{\text{adj}} = 1.3 \times 10^{-6}$). The APC superfamily permease transcripts increased at the early timepoints but were not found to have a significant difference between the guanidine-spiked incubation and the no-substrate control over the entire experiment ($\log_2[\text{FC}] = 1.58$; $P_{\text{adj}} = 0.09$). The guanidine concentration in the Haderslev replicate time series used for transcriptomics dropped substantially between 8 and 16 h of incubation; this decrease was accompanied by a decrease in the transcriptional activity of the guanidinase and the permease (Fig. 5b). These results corroborate the findings from the Ribe sludge experiment, in that the guanidine concentration appears to exert strong control over the transcription of the comammox guanidine permease and guanidinase in WWTP communities.

Guanidine metabolism in soil

As we detected guanidine at low concentrations in urine and faecal samples from cows, pigs, chicken and sheep (Extended Data Table 2), guanidine metabolism in a comammox-containing agricultural soil, which had been fertilized with 525 kg total N per ha per year from solid cattle manure, was also investigated. As expected, the soil microbiome showed strong nitrification activity and added ammonium (30 μg N per g dry-weight soil) was rapidly nitrified (<2 days). Soil nitrification (ammonium consumption and nitrate formation) was completely inhibited by acetylene (Fig. 6). Notably, guanidine added to the soil was nitrified to nitrate over 27 days whether provided as the sole nitrogen source (30 μg N per g dry-weight soil) or in combination with ammonium (each ammonium and guanidine, 15 μg N per g dry-weight soil). However, when nitrification was inhibited, guanidine persisted in the soil at a high concentration for more than 27 days. Although these data do not prove guanidine degradation by comammox, they highlight the large contribution of ammonia-oxidizing microorganisms (or potentially

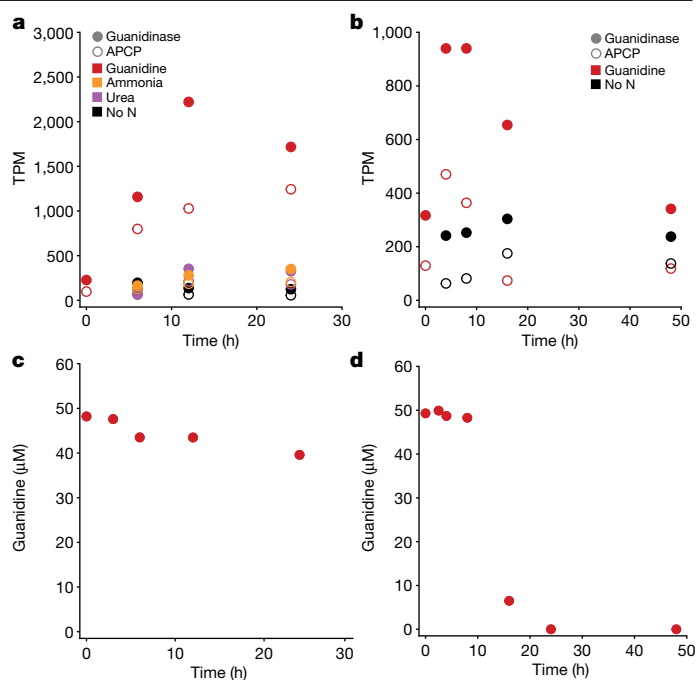


Fig. 5 | Metatranscriptomic response of nitrifying activated sludges to guanidine amendment. **a, b**, The transcriptional response of the guanidine APC superfamily permease (APCP) and guanidinase to substrate-spiked (guanidine, ammonium, urea) incubations in abundant comammox organisms from the Ribe (**a**) and Haderslev (**b**) WWTPs. Metatranscriptomic reads from the Ribe and Haderslev experiments were mapped to the respective Ribe and Haderslev comammox MAGs; transcriptional levels of the genes of interest are shown in TPM for all timepoints. **c, d**, Corresponding concentrations of guanidine in the replicate used for metatranscriptomics at time of sampling for Ribe (**c**) and Haderslev (**d**). Note that transcripts of the comammox guanidinases were also found at the start of the experiment, although the activated sludge samples had been starved during transport and storage, showing that transcription is either also occurring in the WWTP or that the sludge produced guanidine that was immediately consumed.

other acetylene-sensitive microorganisms) to the observed guanidine degradation (Fig. 6). In all treatments, quantitative PCR (qPCR)-based quantification of AOB, AOA, and comammox clade A and B using the functional marker gene ammonia monooxygenase subunit A did not detect growth of ammonia-oxidizing microorganisms over the course of 27 days ($P > 0.05$; Supplementary Table 10). This is not surprising, as conditions that support growth of ammonia oxidizers in soil vary widely among soils^{39–41}. Growth of AOB is typically observed at higher nitrogen fertilization treatments than those used in our study^{42,43}, while the few studies reporting on comammox growth in soil observed it in soils unamended with nitrogen, or after substantially higher urea-N additions than those used in this study⁴¹.

Discussion

Comammox *Nitrospira* are to our knowledge the first organisms identified to grow with guanidine as their sole source of energy, reductant and nitrogen. It is well documented that functionally redundant microorganisms can coexist by partitioning low-concentration substrates even though they compete for one dominant substrate^{44,45}. Thus, it is tempting to speculate that the ability to use the low-concentration substrate guanidine provides an additional niche for comammox organisms in guanidine-containing environments and is probably important for their co-existence with other ammonia-oxidizing microorganisms, with which they compete for the dominant substrate ammonia. Co-existence instead of competitive exclusion enables functional redundancy of

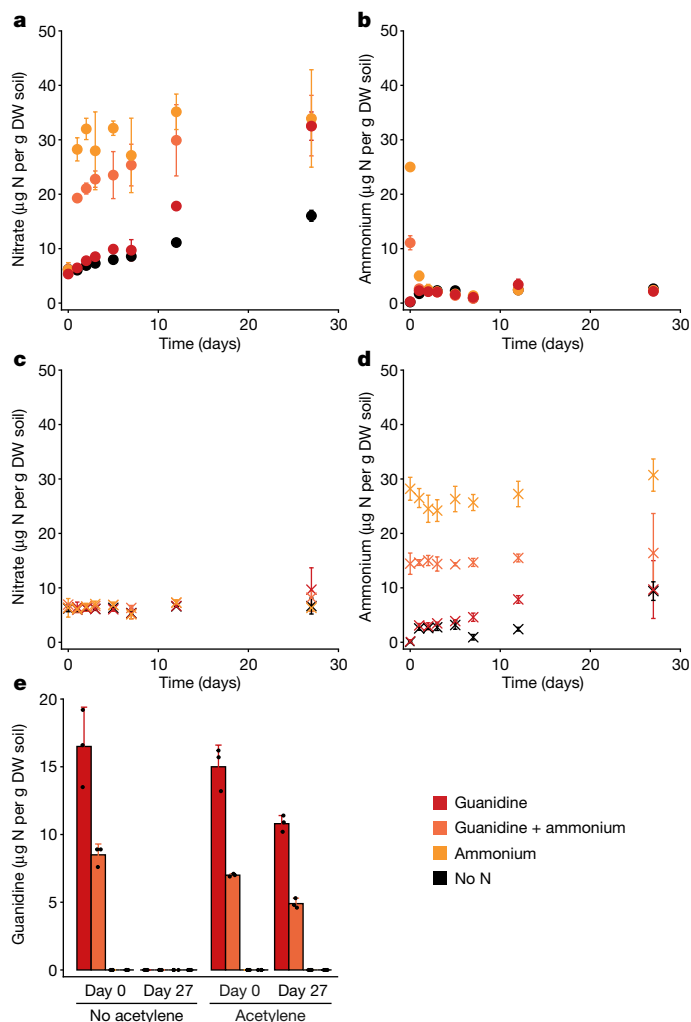


Fig. 6 | Soil nitrification activity in the presence and absence of externally added ammonium and guanidine. Nitrogen applications used in the soil microcosm incubations are comparable to those applied in field nitrogen fertilizer additions⁵⁸. The red symbols represent addition of 30 µg guanidine-N per g dry weight (DW) soil; the orange symbols represent addition of 15 µg guanidine-N + 15 µg NH_4^+ -N per g dry-weight soil; the yellow symbols represent addition of 30 µg NH_4^+ -N per g dry-weight soil; and the black symbols represent no N addition. The circles represent treatments with active nitrification, and the crosses represent treatments with nitrification inhibited with 0.02% acetylene. For **a–d**, data are mean \pm s.d. across three biological replicates. **a**, Net nitrate accumulation with active nitrification. In all treatments, urea and nitrite remained below the detection limit for the entire time course. **b**, Ammonium consumption with active nitrification. **c**, Nitrate and nitrite concentrations with nitrification inhibited with acetylene. **d**, Ammonium concentrations in the presence of acetylene. **e**, Guanidine concentrations (detection limit of 50 nM in the soil extract, corresponding to 14.1 ng N per g dry-weight soil) at day 0 and day 27 in the absence and presence of acetylene. Data are mean \pm s.d. across three biological replicates. The overlaid dots show values from biological triplicates.

ammonia oxidizers, is frequently observed in many environments and probably preserves their overall ecosystem service.

The specific adaptation of comammox to using guanidine for nitrification also opens interesting perspectives for research and applications. Guanidine is the first substrate that has the potential to be selective for any lineage of ammonia oxidizers and has been reported to inhibit the AOB *Nitrosomonas europaea* at 1 mM concentration^{46,47}. Thus, mineral cultivation media containing guanidine as the only substrate should be tested for purification of new comammox *Nitrospira*

strains from enrichments and environmental samples. Currently, only one comammox isolate (*N. inopinata*) and one high enrichment are available and have been partly physiologically characterized^{42,48–50}. In comparison to ammonia oxidation by AOB, complete nitrification by *N. inopinata* yields significantly less nitrous oxide (N_2O) as a by-product under oxic and hypoxic conditions⁴⁹. Future work is needed to examine whether the long-term application of guanidine-containing nitrogen fertilizers will influence the composition of the ammonia-oxidizer community in agricultural soils and the emissions of the greenhouse gas and ozone-depleting N_2O . Comammox *Nitrospira* could also have an important role for the removal of the widely used drug metformin in the environment. It is tempting to speculate that the presence of metformin—for example, in wastewater^{51,52}—opens together with other guanidine sources a niche for comammox in exposed habitats. In this case, the continuous release of metformin into the environment might influence the composition of nitrifying populations.

As we are just beginning to understand the environmental concentrations of guanidine, the importance of this nitrogen-rich compound for the biogeochemical nitrogen cycle cannot yet be quantified. For this purpose, a more sensitive analytical method for measuring in situ guanidine concentrations with high specificity is urgently needed. As shown previously for cyanate, even low standing concentrations of a nitrification substrate can be linked to relatively high production and consumption rates⁵³. Thus, future studies should aim to determine both the steady-state concentrations and the turnover rates of guanidine in pristine and human-affected ecosystems. Moreover, the diversity of natural and synthetic guanidinium compounds and their degradation pathways, such as of many neonicotinoid pesticides⁵⁴ and nitroguanidine, which is used in explosives^{55,56}, indicates that a multitude of hitherto mostly overlooked, biotic and abiotic guanidine-producing reactions may exist. Insights into the environmental dynamics of guanidine and its derivatives will close yet another gap in our picture of the nitrogen cycle and shed light onto the ecological roles of comammox *Nitrospira* and their enigmatic functional separation from the canonical nitrifiers.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-024-07832-z>.

- Lewis, C. A. & Wolfenden, R. The nonenzymatic decomposition of guanidines and amidines. *J. Am. Chem. Soc.* **136**, 130–136 (2014).
- Putnam, D. F. *Composition and Concentrative Properties of Human Urine* Contractor Report 1802 (NASA, 1971). ntrs.nasa.gov/citations/19710023044.
- Wishart, D. S. et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **46**, D608–D617 (2018).
- Nelson, J. W., Atilho, R. M., Sherlock, M. E., Stockbridge, R. B. & Breaker, R. R. Metabolism of free guanidine in bacteria is regulated by a widespread riboswitch class. *Mol. Cell* **65**, 220–230 (2017).
- Huang, L., Wang, J. & Lilley, D. M. J. The structure of the guanidine-II riboswitch. *Cell Chem. Biol.* **24**, 695–702 (2017).
- Huang, L., Wang, J., Wilson, T. J. & Lilley, D. M. J. Structure of the guanidine III riboswitch. *Cell Chem. Biol.* **24**, 1407–1415 (2017).
- Sherlock, M. E. & Breaker, R. R. Biochemical validation of a third guanidine riboswitch class in bacteria. *Biochemistry* **56**, 359–363 (2017).
- Salvail, H., Balaji, A., Yu, D., Roth, A. & Breaker, R. R. Biochemical validation of a fourth guanidine riboswitch class in bacteria. *Biochemistry* **59**, 4654–4662 (2020).
- Eckert, C. et al. Ethylene-forming enzyme and bioethylene production. *Biotechnol. Biofuels* **7**, 33 (2014).
- Dunham, N. P. et al. Two distinct mechanisms for C–C desaturation by iron(II)- and 2-(oxoglutarate)-dependent oxygenases: importance of α -heteroatom assistance. *J. Am. Chem. Soc.* **140**, 7116–7126 (2018).
- Tassoulas, L. J., Robinson, A., Martinez-Vaz, B., Aukema, K. G. & Wackett, L. P. Filling in the gaps in metformin biodegradation: a new enzyme and a metabolic pathway for guanylurea. *Appl. Environ. Microbiol.* **87**, e03003-20 (2021).
- Amberger, A. & Vilsmeier, K. Dicyandiamidabbau in Quarzsand und Böden. *Z. Pflanzenernähr. Bodenkd.* **142**, 778–785 (1979).

13. Hartig, J. et al. Metformin hydrolase is a recently evolved, nickel-dependent, heteromeric ureohydrolase. Preprint at Research Square <https://doi.org/10.21203/rs.3.rs-3956544/v1> (2024).
14. Funck, D., Sinn, M., Forlani, G. & Hartig, J. S. Guanidine production by plant homoarginine-6-hydroxylases. *eLife* **12**, RP91458 (2024).
15. Schneider, N. O. et al. Solving the conundrum: widespread proteins annotated for urea metabolism in bacteria are carboxyguanidine deiminases mediating nitrogen assimilation from guanidine. *Biochemistry* **59**, 3258–3270 (2020).
16. Sinn, M., Hauth, F., Lenkeit, F., Weinberg, Z. & Hartig, J. S. Widespread bacterial utilization of guanidine as nitrogen source. *Mol. Microbiol.* **116**, 200–210 (2021).
17. Wang, B. et al. A guanidine-degrading enzyme controls genomic stability of ethylene-producing cyanobacteria. *Nat. Commun.* **12**, 5150 (2021).
18. Funck, D. et al. Discovery of a Ni²⁺-dependent guanidine hydrolase in bacteria. *Nature* **603**, 515–521 (2022).
19. Ward, B. B., Arp, D. J. & Klotz, M. G. (eds) *Nitrification* (ASM, 2011); <https://doi.org/10.1128/9781555817145>.
20. Erisman, J. W., Sutton, M. A., Galloway, J., Klimont, Z. & Winiwarer, W. How a century of ammonia synthesis changed the world. *Nat. Geosci.* **1**, 636–639 (2008).
21. Daims, H. et al. Complete nitrification by *Nitrospira* bacteria. *Nature* **528**, 504–509 (2015).
22. van Kessel, M. A. H. J. et al. Complete nitrification by a single microorganism. *Nature* **528**, 555–559 (2015).
23. Koch, H. et al. Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus *Nitrospira*. *Proc. Natl Acad. Sci. USA* **112**, 11371–11376 (2015).
24. Leung, P. M. et al. A nitrite-oxidising bacterium constitutively consumes atmospheric hydrogen. *ISME J.* **16**, 2213–2219 (2022).
25. Koper, T. E., El-Sheikh, A. F., Norton, J. M. & Klotz, M. G. Urease-encoding genes in ammonia-oxidizing bacteria. *Appl. Environ. Microbiol.* **70**, 2342–2348 (2004).
26. Palatinszky, M. et al. Cyanate as an energy source for nitrifiers. *Nature* **524**, 105–108 (2015).
27. Kitzinger, K. et al. Cyanate and urea are substrates for nitrification by Thaumarchaeota in the marine environment. *Nat. Microbiol.* **4**, 234–243 (2019).
28. Yang, Y. et al. Activity and metabolic versatility of complete ammonia oxidizers in full-scale wastewater treatment systems. *mBio* **11**, e03175-19 (2020).
29. Kai, M., Miura, T., Kohashi, K. & Ohkura, Y. New method for the fluorimetric determination of guanidino compounds with benzoin. *Chem. Pharm. Bull.* **29**, 1115–1120 (1981).
30. Reiss, C. W., Xiong, Y. & Strobel, S. A. Structural basis for ligand binding to the guanidine-1 riboswitch. *Structure* **25**, 195–202 (2017).
31. Pedretti, A., Mazzolari, A., Gervasoni, S., Fumagalli, L. & Vistoli, G. The VEGA suite of programs: an versatile platform for cheminformatics and drug design projects. *Bioinformatics* **37**, 1174–1175 (2021).
32. Button, D. K. Nutrient uptake by microorganisms according to kinetic parameters from theory as related to cytoarchitecture. *Microbiol. Mol. Biol. Rev.* **62**, 636–645 (1998).
33. Qin, W. et al. Ammonia-oxidizing bacteria and archaea exhibit differential nitrogen source preferences. *Nat. Microbiol.* **9**, 524–536 (2024).
34. Jung, M.-Y. et al. Ammonia-oxidizing archaea possess a wide range of cellular ammonia affinities. *ISME J.* **16**, 272–283 (2022).
35. Stöckler, S., Marescau, B., De Deyn, P. P., Trijebels, J. M. & Hanefeld, F. Guanidino compounds in guanidinoacetate methyltransferase deficiency, a new inborn error of creatine synthesis. *Metab. Clin. Exp.* **46**, 1189–1193 (1997).
36. Gatti, R. & Gioia, M. G. Anisoin: a useful pre-chromatographic derivatization fluorogenic reagent for LC analysis of guanidino compounds. *J. Pharm. Biomed. Anal.* **42**, 11–16 (2006).
37. Singleton, C. M. et al. Connecting structure to function with the recovery of over 1000 high-quality metagenome-assembled genomes from activated sludge using long-read sequencing. *Nat. Commun.* **12**, 2009 (2021).
38. Pjevac, P. et al. AmoA-targeted polymerase chain reaction primers for the specific detection and quantification of comammox *Nitrospira* in the environment. *Front. Microbiol.* **8**, 1508 (2017).
39. Li, C., Hu, H.-W., Chen, Q.-L., Chen, D. & He, J.-Z. Comammox *Nitrospira* play an active role in nitrification of agricultural soils amended with nitrogen fertilizers. *Soil Biol. Biochem.* **138**, 107609 (2019).
40. He, S. et al. Ammonium concentration determines differential growth of comammox and canonical ammonia-oxidizing prokaryotes in soil microcosms. *Appl. Soil Ecol.* **157**, 103776 (2021).
41. Wang, Z. et al. Comammox *Nitrospira* clade B contributes to nitrification in soil. *Soil Biol. Biochem.* **135**, 392–395 (2019).
42. Taylor, A. E. et al. Use of aliphatic *n*-alkynes to discriminate soil nitrification activities of ammonia-oxidizing thaumarchaea and bacteria. *Appl. Environ. Microbiol.* **79**, 6544–6551 (2013).
43. Hink, L., Nicol, G. W. & Prosser, J. I. Archaea produce lower yields of N₂O than bacteria during aerobic ammonia oxidation in soil. *Environ. Microbiol.* **19**, 4829–4837 (2017).
44. Yu, X. A. et al. Low-level resource partitioning supports coexistence among functionally redundant bacteria during successional dynamics. *ISME J.* **18**, wrad013 (2024).
45. Brochet, S. et al. Niche partitioning facilitates coexistence of closely related honey bee gut bacteria. *eLife* **10**, e68583 (2021).
46. Lees, H. & Quastel, J. H. Biochemistry of nitrification in soil: 3. Nitrification of various organic nitrogen compounds. *Biochem. J.* **40**, 824–828 (1946).
47. Lees, H. The biochemistry of the nitrifying organisms. I. The ammonia oxidizing systems of *Nitrosomonas*. *Biochem. J.* **52**, 134–139 (1952).
48. Kits, K. D. et al. Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature* **549**, 269–272 (2017).
49. Kits, K. D. et al. Low yield and abiotic origin of N₂O formed by the complete nitrifier *Nitrospira inopinata*. *Nat. Commun.* **10**, 1836 (2019).
50. Sakoula, D. et al. Enrichment and physiological characterization of a novel comammox *Nitrospira* indicates ammonium inhibition of complete nitrification. *ISME J.* **15**, 1010–1024 (2021).
51. Balakrishnan, A., Sillanpää, M., Jacob, M. M. & Vo, D.-V. N. Metformin as an emerging concern in wastewater: occurrence, analysis and treatment methods. *Environ. Res.* **213**, 113613 (2022).
52. Oosterhuis, M., Sacher, F. & Ter Laak, T. L. Prediction of concentration levels of metformin and other high consumption pharmaceuticals in wastewater and regional surface water based on sales data. *Sci. Total Environ.* **442**, 380–388 (2013).
53. Mooshammer, M., Wanek, W., Jones, S. H., Richter, A. & Wagner, M. Cyanate is a low abundance but actively cycled nitrogen compound in soil. *Commun. Earth Environ.* **2**, 161 (2021).
54. Anjos, C. S., Lima, R. N. & Porto, A. L. M. An overview of neonicotinoids: biotransformation and biodegradation by microbiological processes. *Environ. Sci. Pollut. Res. Int.* **28**, 37082–37109 (2021).
55. Burrows, W. D., Schmidt, M. O., Chyrek, R. H. & Noss, C. I. *Photochemistry of Aqueous Nitroguanidine* Technical Report AD-A203 200 (US Army Biomedical Research and Development Laboratory, 1988).
56. Kaplan, D. L., Cornell, J. H. & Kaplan, A. M. Decomposition of nitroguanidine. *Environ. Sci. Technol.* **16**, 488–492 (1982).
57. Chovancova, E. et al. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput. Biol.* **8**, e1002708 (2012).
58. Adalibieke, W., Cui, X., Cai, H., You, L. & Zhou, F. Global crop-specific nitrogen fertilization dataset in 1961–2020. *Sci. Data* **10**, 617 (2023).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

Methods

Gene collection and functional classification

Supplementary Table 11 contains details for the collection and analyses parameters of each gene family and the following description is generalized. Predicted proteins from all publicly available genomes in GenBank as of 1 July 2022 (429,896 genomes) were screened with *hmmsearch*⁵⁹ using 'collection HMMs' for genes related to guanidine metabolism (Fig. 1). The resulting genes were used as query sequences against a combined HMM set of PFAMs, TIGRFAMs, NCBI FAMS and PANTHERFAMs for a set of acceptable 'cross-check HMMs'. Genes were further screened using specific *e*-value and coverage cut-offs. Cross-check HMM names were used to query UniProt and results were filtered for reviewed entries with evidence at the protein level to identify functionally characterized proteins and download them if not already present in the dataset. The portion of the protein sequence that was aligned to the 'collection HMM' was extracted and clustered using *usearch*⁶⁰ with specified *-id* and *-query_cov* values to identify centroids. HMM-based alignments of centroid sequences generated from the initial *hmmsearch* were used in *FastTree*⁶¹ to generate phylogenetic trees for each gene family of interest. For the APC superfamily permease and the allophanate hydrolase, all proteins that passed the *e*-value and coverage cut-offs were inferred to possess the expected function. Guanidinases were also required to possess threonine at *N. inopinata* position 105 (PF00491 HMM position 15), histidine at *N. inopinata* position 222 (PF00491 HMM position 134) and tryptophan at *N. inopinata* position 313 (PF00491 HMM position 223)¹⁸. Guanidine carboxylases were required to possess a conserved aspartic acid at *K. lactis* position 1,584 (TIGR02712 HMM position 956) and were further differentiated from urea carboxylase by having an aspartic acid at *K. lactis* position 1,330 (TIGR02712 HMM position 701)¹⁵. Carboxyguanidine deiminases were defined using the common ancestor of the two subunits (CgdA and CgdB) in the general tree for PF09347. This common ancestral node then gave rise to CgdA and CgdB as two monophyletic clades defined as such.

Riboswitch collection

Genomes from ammonia-oxidizing microorganisms were screened using *infernal* (v.1.1.3)⁶² using established RFAM covariance models for guanidine riboswitches I (RF00442), II (RF01068) and III (RF01763) and a model for the recently described guanidine IV riboswitch, which was constructed with *infernal* using 'GGAM-1-curated.sto'⁶³. Scaffold IDs, coordinates and orientation were recorded and cross-referenced against *gff* files to identify downstream genes. A gene was considered to be under the control of a riboswitch if it was in the same orientation as the riboswitch and the 5' end of the gene was within 1,000 nucleotides of the riboswitch. The inferred operon was then extended downstream until genes could be found with the opposite orientation.

Phylogenetic analyses

Ureohydrolase. HMM-based alignments of centroid sequences defined above were used in *FastTree*⁶¹ with the default parameters to generate a phylogenetic tree. The tree was midpoint rooted using the function *midpoint()* from the *phangorn* package, and functional clades were defined using the *getMRCA()* function within the *ape* package and visualized using the *ggtree* package in R.

Guanidinase. The most recent common ancestor of all guanidinases (as defined above) was identified in the ureohydrolase tree using the *getMRCA()* function and the descendant centroids were collected using the *Descendants()* function, both from the *ape* package⁶⁴. All HMM alignments of ureohydrolases that were represented by centroids collected using the *Descendants()* function were additionally required to have covered PF00491 over 90% of its length reclustered using *usearch* (*-id* 0.9 *-query_cov* 0.9). The HMM-aligned portion of this sequence dataset was used to calculate phylogeny with *IQ-TREE*⁶⁵, using the

best model (LG+I+R5), and bipartition support was evaluated using ultrafast bootstraps. Logos for each resulting clade of guanidinases and close relatives were generated using the *ggseqlogo* package in R.

For co-phylogeny analyses of comammox guanidinases and ammonia monooxygenases, comammox genomes were screened for the presence of *amoA* and guanidinase genes. As most available genomes were MAGs, it was required that exactly one copy of each gene was identified per genome. This resulted in 54 genomes for analysis. The *AlignTranslation* and *AlignSeqs* functions from the *Decipher* package were used to align *amoA* nucleotide and guanidinase amino acid sequences, respectively. *IQ-TREE*⁶⁵ was used to identify the best models (*amoA*, TPM3u+F+I+R3; guanidinase protein, LG+I+G4), calculate trees and evaluate bipartition support with ultrafast bootstraps. Trees were visualized in R using a combination of the *cophylo* function from the *phytools* package and the *ggtree* package.

Guanidine quantification

The following chemicals were purchased from Sigma-Aldrich: guanidine hydrochloride ($\geq 99\%$, G3272), benzoin ($\geq 99\%$, 8.01776), potassium hydroxide ($\geq 85\%$, 1.05033), ethanol ($\geq 99.8\%$, 02851), formic acid ($\geq 98\%$, 5.43804), β -mercaptoethanol ($\geq 99\%$, 8.05740), L-arginine ($\geq 99.5\%$, 11009) and sodium sulfite ($\geq 98\%$, 239321). Hydrogen chloride solution (32%, 20254.321) and acetonitrile ($\geq 99.9\%$, 20060.320) were purchased from VWR. 2-Methoxyethanol ($\geq 99.5\%$, 10582945) was purchased from Thermo Fisher Scientific. MilliQ water was obtained from a water purification system (0.071 $\mu\text{S cm}^{-1}$; Elga Veolia, PURELAB Chorus). The derivatization protocol for guanidine was adapted from a previous study²⁹. In brief, 150 μl of an aqueous solution potentially containing guanidine was cooled to 0 °C in a 0.5 ml plastic tube (Eppendorf, Protein LoBind, 0030108434) and spiked with 75 μl of a benzoin solution in ethanol (4 mM), 75 μl of an aqueous solution containing both β -mercaptoethanol (0.1 M) and sodium sulfite (0.2 M), and 150 μl of an aqueous solution of potassium hydroxide (1.6 M). The resulting solution was mixed, heated in a bath of boiling water for 10 min and cooled in an ice bath for 2 min. Subsequently, 25 μl of an aqueous solution of hydrogen chloride (4.8 M) was added. The resulting solution was mixed and transferred to a 1.5 ml plastic tube (Eppendorf, 0030120086) and centrifuged at 10,000g for 2 min. Before analysis, the supernatant was diluted to obtain analyte concentrations in the optimal quantification range of the analytical instrument (that is, 0.05–5 μM). The predominant derivatization product (proposed structure in Supplementary Fig. 1c) was analysed using liquid chromatography (Agilent 1290 Infinity II) coupled to triple quadrupole mass spectrometry (Agilent, 6470) with a retention time of 3.73 min. We used the InfinityLab Poroshell 120 Bonus-RP (Agilent, 2.7 μm , 2.1 \times 150 mm) column for separation, an injection volume of 2 μl , a flow rate of 0.4 ml min⁻¹, a column compartment temperature of 40 °C and the following eluents: aqueous (A): MilliQ water with 0.1% (v/v) formic acid; organic (B): acetonitrile with 0.1% (v/v) formic acid. The eluent gradient was as follows: 0–1.5 min, 5% B; 1.5–4 min, 5–61% B; 4–4.5 min, 61–95% B; 4.5–7 min, 95% B; 7–8 min, 95–5% B; 8–10 min, 5% B. The source parameters were set as follows: positive mode electrospray ionization; gas temperature, 250 °C; gas flow, 10 l min⁻¹; nebulizer, 45 psi; sheath gas temperature, 280 °C; sheath gas flow, 11 l min⁻¹; capillary voltage, 3.5 kV; nozzle voltage, 0.5 kV. The following product ions of the derivatization product (*m/z* of parent ion: 252.2) were monitored: *m/z*: 182.1 (quantifier) and *m/z*: 104.1 (qualifier). The resulting chromatographs were integrated using MassHunter (Agilent, v.10.1). For absolute quantification, we used a series of guanidine solutions with a concentration range after dilution between 0.05 and 5 μM . For pure culture medium, activated sludge and soil extracts, calibration solutions were prepared in the respective matrix. Animal urine and faeces were quantified with calibration solutions in water. Accurate quantification for animal samples was confirmed by spiking 20 μM guanidine to an animal faeces sample with a recovery of >83%. Animal manure samples were freeze-dried, and

subsamples were dispersed in 2 M KCl solution (1 ml per 100 mg sample) and bead-beated for 15 min in a Lysing matrix A tube (MPBiomedicals), then centrifuged at 20,000g for 15 min. Wastewater treatment plant influent was quantified by standard addition. Calibration solutions were derivatized and analysed in the same way as and in parallel to the respective samples. For Orbitrap (high-resolution) MS analyses, we used liquid chromatography coupled to the Thermo QExactive mass spectrometer with the following parameters: positive electrospray ionization; capillary temperature, 275 °C; sheath gas, 15; aux gas, 10; sweep gas, 1; S-lens RF, 50.0; resolution, 140,000 (MS full-scan), 17,500 (MS/MS); NCE (stepped), 10, 20, 30. For growth experiment samples containing heavy-isotope-labelled guanidine, the total guanidine concentrations were inferred by assuming the measured isotopically unlabelled guanidine concentrations to correspond to 90% (we used 10% isotopically labelled guanidine).

Physiology experiments with *N. inopinata* and ammonia-oxidizing bacteria

The cells were grown in medium containing 54.4 mg l⁻¹ KH₂PO₄, 74.4 mg l⁻¹ KCl, 49.3 mg l⁻¹ MgSO₄·7 H₂O, 584 mg l⁻¹ NaCl, 147 mg l⁻¹ CaCl₂, 34.4 µg l⁻¹ MnSO₄·1H₂O, 50 µg l⁻¹ H₃BO₃, 70 µg l⁻¹ ZnCl₂, 72.6 µg l⁻¹ Na₂MoO₄·2H₂O, 1 mg l⁻¹ FeSO₄·7H₂O, 20.0 µg l⁻¹ CuCl₂·2H₂O, 80 µg l⁻¹ CoCl₂·6 H₂O, 3 µg l⁻¹ Na₂SeO₄·5H₂O, 4 µg l⁻¹ Na₂WO₄·2H₂O, 24 µg l⁻¹ NiCl₂·6 H₂O and 0.5 mM pyruvate. The medium was buffered by addition of 4 mM HEPES, with the pH set to 8. For regular culture maintenance, cultures were kept in closed Schott bottles at 37 °C without shaking in the dark. When indicated, guanidine hydrochloride was added from a filter-sterilized 0.1 M stock solution to a final concentration of 50 µM.

For comparing guanidine utilization by pure cultures of *N. inopinata* and AOB, all strains were induced in 1 l batch cultures for 6 weeks with 0.5 mM ammonium and 1 µM guanidine fed weekly. Subsequently, the same amount of biomass per culture as determined using the Pierce BCA protein quantification kit (Thermo Fisher Scientific; calculated final concentration in the incubation, 10 µg ml⁻¹) was collected, washed and resuspended in fresh medium in equal volumes and transferred to 96-well, flat bottom culture plates (Greiner Bio-One). In these plates the following incubations were done with either 50 µM guanidine only; 50 µM guanidine plus 150 µM ammonium; or 150 µM ammonium only for 14 days at 28 °C in the dark and without agitation (optimal growth conditions for the ammonia oxidizing organisms used, while 9 °C colder than the optimum for *N. inopinata*).

For growth experiments, *N. inopinata* pure culture cells pregrown on 10 µM guanidine and 0.5 mM ammonium (with weekly refeedings) for 1 month were collected by centrifugation (4,500g, 30 min), washed with N-free medium three times and resuspended in fresh medium. Aliquots of 200 ml were distributed into 250 ml serum bottles. Aliquots used as dead controls were autoclaved (120 °C, 20 min) before substrate additions. The following N substrates were added (always 150 µM N) to five replicate bottles each: (A) ¹⁵N-guanidine (10% ¹⁵N-guanidine hydrobromide, 90% guanidine hydrochloride); (B) guanidine (as guanidine hydrochloride); (C) ¹⁵N-guanidine (10% ¹⁵N-guanidine hydrobromide, 90% guanidine hydrochloride) and ammonium (each 150 µM N); (D) ammonium only; (E) no N addition (starved control); (F) dead (autoclaved) control with ¹⁵N-guanidine (10% ¹⁵N-guanidine hydrobromide, 90% guanidine hydrochloride). Moreover, all bottles received ¹³C-bicarbonate additions (¹³C-NaHCO₃; 1 mM final concentration, 99% ¹³C) to detect chemolithoautotrophic growth and 0.5 mM sodium pyruvate as a reactive-oxygen-species scavenger⁶⁶. Serum bottles were closed with sterile, HCl-cleaned blue butyl rubber stoppers (Chemglass) and incubated at 37 °C in the dark without agitation. Samples of 2 ml for the determination of cell numbers (using qPCR) and of N-compound concentrations were taken with sterile syringes and needles and replaced with air every 1 to 14 days (frequent sampling in the beginning of the experiment,

more spaced-out sampling after incubations containing ammonium were ended) over a time course of 126 days (12 days for treatments containing ammonium). Substrates were replenished after depletion. After 107 days of incubation, 10 ml samples were removed from treatments A, B, E and F, fixed with 3% formaldehyde (final concentration) for 30 min at room temperature, filtered onto polycarbonate filters (0.2 µm pore size, GTTP, 40 nm gold sputtered), washed with sterile 1× PBS, dried and stored frozen until further use. Cells were visualized after staining with 4',6'-diamidino-2-phenylindole (DAPI, 10 µg ml⁻¹, 5 min at room temperature) using a confocal laser-scanning microscope (inverted Leica TCS SP8X CLSM equipped with a 405 nm UV diode). At the end of the growth experiments, the absence of heterotrophic contaminants was confirmed by inoculation into heterotrophic growth medium (LB and TSY).

Ammonium, urea, nitrite and nitrate concentrations were measured by colorimetric protocols published previously⁶⁷. In brief, combined ammonia and ammonium concentrations were determined using the indophenol blue method. Nitrite concentrations were measured spectrophotometrically using the Griess method after reacting with sulfanilamide and *N*-1-naphthyl-ethylenediamine dihydrochloride. Nitrate was measured by the same method after reduction to nitrite with vanadium chloride. Urea concentrations were measured using the thiosemicarbazide-diacetylmonoxime method⁶⁸, according to a previous study¹⁸.

For quantification of *N. inopinata* cell numbers, qPCR was performed using the primers 515F/806R, targeting the V4 region of the 16S rRNA gene as described previously^{69,70}. Standards were generated from purified PCR products generated from *N. inopinata* genomic DNA as template. The standards were quantified according to the Qubit dsDNA HS Assay Kit instructions. Standards containing 10⁹ gene copies per µl were aliquoted and stored frozen at -20 °C until further use. Each standard aliquot was used and defrosted only once to freshly prepare tenfold serial dilutions (10⁸–10² gene copies per µl). The qPCR assays were performed as follows: the frozen culture aliquots were four times freeze-thawed for cell disruption. A total of 0.25 µM of each primer was used in a mixture of 10 µl SYBR Green Supermix (Bio-Rad), 2 µl cell lysate or standard, and water in a final volume of 22 µl per reaction. The qPCR cyclers (C1000-CFX96, Bio-Rad) settings were as follows: 95 °C for 15 min; 40 cycles of 95 °C for 30 s, 50 °C for 1 min and 72 °C for 45 s (plate read); and finishing with 72 °C for 2 min and a melting curve performance from 40 °C to 95 °C with an increase of 0.5 °C every 5 s. The efficiencies of the standard curves had an average of 86% and an R² of 0.999. Growth rates (division rates) were calculated as follows:

$$\nu(d^{-1}) = \log_2(N_{i+1}/N_i)/t \quad (1)$$

where ν is the rate of division (d⁻¹), N is the qPCR determined cell number at timepoint $i + 1$ and i , and t is the time interval between time point $i + 1$ and i in days.

For visualization of stable N and C isotope assimilation into *N. inopinata* cells from the supplied ¹⁵N-guanidine and ¹³C-bicarbonate, gold-sputtered filters containing cells from two replicate bottles (Treatment A, replicate A1 and A2) and a natural abundance (NA) control were glued onto antimony-doped silicon wafers (7.1 × 7.1 × 0.11 mm, Active Business Company) using superglue (Loctite). NanoSIMS measurements were performed on the NanoSIMS 50L instrument (Cameca) at the Large-Instrument Facility for Environmental and Isotope Mass Spectrometry at the University of Vienna. Before image acquisition, each analysis area was preconditioned by sequence of high and extreme low ion impact energy (EXLIE) Cs⁺ depositions as follows: high energy (16 keV) at 50 pA beam current to a fluence of 5 × 10¹⁴ ions cm⁻²; EXLIE (50 eV) at 400 pA beam current to a fluence of 5 × 10¹⁶ ions cm⁻²; high energy to an additional fluence of 2.5 × 10¹⁴ ions cm⁻². Data were acquired as multilayer image stacks by sequential scanning with a finely focused primary Cs⁺ ion beam (approximately 80 nm probe size at a

Article

2 pA beam current) over $45 \times 45 \mu\text{m}^2$ areas with 512×512 pixel image resolution. The primary ion beam dwell time varied between 1 ms (A1, 74 planes; NA, 50 planes) and 5 ms (A2, 21 planes) per pixel per cycle. The detectors of the multicollection assembly were positioned to enable parallel detection of $^{12}\text{C}_2^-$, $^{12}\text{C}^{13}\text{C}^-$, $^{12}\text{C}^{14}\text{N}^-$, $^{12}\text{C}^{15}\text{N}^-$, $^{31}\text{P}^-$ and $^{32}\text{S}^-$ secondary ions. Image data analysis was performed using the OpenMIMS ImageJ plugin (OpenMIMS v.3.0.5, ImageJ v.1.54f), where the acquired datasets were aligned, deadtime and QSA corrected, processed (for example, accumulation, stable isotope ratio calculation) and exported for visualization of ^{13}C and ^{15}N enrichment (as ^{13}C and ^{15}N atom%).

N. inopinata shotgun proteomics

For protein analysis, biomass was dissolved in lysis buffer (8 M urea, 2 M thiourea, 1 mM PMSF). Protein extraction was done by incubation at 95°C , while shaking at 1,400 rpm for 5 min. Subsequently, the samples were treated for 3 min in an ultrasonication water bath (Elmasonic S30 H). To the cell suspension, 6.75 μl 25 mM 1,4 dithiothreitol (in 20 mM ammonium bicarbonate) was added and incubated for 1 h at 60°C and 1,400 rpm shaking. Next, 150 μl 10 mM iodoacetamide (in 20 mM ammonium bicarbonate) was added and incubated for 30 min at 37°C with 1,400 rpm shaking in the dark. Finally, 200 μl of 20 mM ammonium bicarbonate was added and the protein lysates were proteolytically cleaved overnight at 37°C with trypsin (2.5 μl of 0.1 $\mu\text{g} \mu\text{l}^{-1}$ trypsin, Promega). The cleavage was stopped by adding 50 μl 10% formic acid. The peptide lysates were desalted using ZipTip $\mu\text{C}18$ tips (Merck Millipore). The peptide lysates were resuspended in 15 μl 0.1% formic acid and analysed using nanoliquid chromatography–MS (UltiMate 3000 RSLC-nano, Dionex, Thermo Fisher Scientific). MS analyses of eluted peptide lysates were performed on the Q Exactive HF mass spectrometer (Thermo Fisher Scientific) coupled with a TriVersa NanoMate (Advion). Peptide lysates were injected onto a trapping column (Acclaim PepMap 100 C18, 3 μm , nanoViper, 75 $\mu\text{m} \times 2 \text{ cm}$, Thermo Fisher Scientific) with 5 $\mu\text{l} \text{ min}^{-1}$ by using 98% water/2% acetonitrile with 0.5% trifluoroacetic acid, and separated on an analytical column (Acclaim PepMap 100 C18, 3 μm , nanoViper, 75 $\mu\text{m} \times 25 \text{ cm}$, Thermo Fisher Scientific) at a flow rate of 300 nL min^{-1} . Mobile phase was 0.1% formic acid in water (A) and 80% acetonitrile/0.08% formic acid in water (B). Full MS spectra (350–1,550 m/z) were acquired in the Orbitrap at a resolution of 120,000 with automatic gain control target value of 3×10^6 ions.

Acquired LC–MS data were analysed with the Proteome Discoverer (v.2.5, Thermo Fischer Scientific) using SEQUEST HT and INFERYS Rescoring. Protein identification was performed using a database constructed from predicted proteins of *N. inopinata* downloaded from MicroScope⁷¹ and common contaminating proteins. Searches were conducted with the following parameters: trypsin as enzyme specificity and two missed cleavages allowed. A peptide ion tolerance of 10 ppm and an MS/MS tolerance of 0.02 Da were used. As modifications, oxidation (methionine) and carbamidomethylation (cysteine) were selected. Peptides that scored $q > 1\%$ based on a decoy database and with a peptide rank of 1 were considered identified. Differential expression of proteins was evaluated using the DEqMS⁷². Normalized spectral abundance factors were also calculated for visualization purposes only.

Heterologous expression and purification of *N. inopinata* guanidinase

The guanidinase gene of *N. inopinata* was amplified with self-designed, specific PCR primers which already contained the vector-specific linker overhangs for Gibson cloning (5'-CTGGAAGTTCTGTTCACGGGGCCCATGGCGAAAAAGAGAACGTACC-3' and 5'-CCCCAGAACATCAGGTTAATGGCGTCAGCGTTTCTTCGATTGCC-3'), using high-fidelity Phusion Plus PCR Master Mix (Thermo Fisher Scientific). The purified product was cloned into the pCoofy4 (pETM44; His6-MBP) expression vector by using the GeneArt Gibson Assembly EX Cloning Kit (Invitrogen) according to the manufacturer's protocol. The sequence of the insert was verified by Sanger sequencing.

Cultures were grown at 37°C in auto-induction ZYP-5052 medium⁷³ supplemented with 0.5 μM , 20 μM , or 1 mM NiSO_4 for 5 h before cooling down at 4°C for 15 min, followed by overnight expression at 20°C . Cells were lysed in the presence of a protease inhibitor cocktail in 50 mM HEPES, 200 mM NaCl, 5% glycerol, pH 7.4 using a cell disruptor (Constant Systems) and centrifuged at 4°C and 45,000g for 30 min. Guanidinase fused N-terminally to a His-MBP-tag was purified by affinity chromatography using MBPTrap HP columns (Cytiva). Subsequently, the His-MBP-tag was cleaved overnight with HRV-3C protease added at a mass ratio of protease to protein of 1:50. Guanidinase was further purified by MBPTrap HP columns (Cytiva), followed by size-exclusion chromatography on the HiLoad Superdex 200 26/600pg column (Cytiva) equilibrated with 20 mM HEPES, 200 mM NaCl, 5% glycerol, pH 7.4. For the 20 μM nickel in expression batch, this nickel concentration was maintained in all buffers during purification.

The sample was concentrated to around 10 mg ml^{-1} by ultrafiltration by using Vivaspin centrifugal concentrators (Sartorius) and flash-frozen and stored at -80°C . Protein identity and purity were analysed using SDS–PAGE.

Size-exclusion chromatography combined with multiangle light scattering

SEC-MALS was performed using a Superdex 200 increase 10/300 GL (Cytiva) operated at 20°C on the 1260 Infinity HPLC system (Agilent Technologies) coupled to a miniDawn Treos MALS detector (Wyatt Technology). The samples were injected (80 μl at 1 mg ml^{-1}) onto a column extensively equilibrated with 20 mM HEPES, 150 mM NaCl, pH 7.4. Measurement was performed using BSA as a control. The protein concentration was measured with a RI-101 refractive index detector (Shodex) and the average molecular mass was calculated using the program Astra (Wyatt Technology). The first-order fit Zimm formalism was used for analysis of light-scattering data as a data process procedure in Astra, and a generic protein dn/dc value of 0.185 ml g^{-1} was used for guanidinase and BSA.

Protein T_m determination

The Prometheus NT.48 instrument (NanoTemper Technologies) was used to determine the melting temperatures (T_m). Before measurements, samples of the guanidinase expressed with 0.5 μM Ni^{2+} were centrifuged for 10 min at 16,000g at 4°C to remove any large aggregates. To identify the buffer/pH, at which the T_m of the protein was the highest, the protein was diluted using a DSF-buffer/pH screen containing different buffers and pH values⁷⁴. The capillaries were filled with 10 μl of sample and placed onto the sample holder. A temperature gradient of 1°C min^{-1} from 20 to 95°C was applied and the intrinsic protein fluorescence at 330 and 350 nm was recorded. Data were processed using MoltenProt⁷⁵, where the melting temperatures from the curves were estimated using the two-state reversible unfolding model.

MS for heterologous expression experiments

Protein identity and purity were verified by intact protein mass spectrometry. A total of 40 ng of the sample was injected into a column on the LC–MS system: Dionex nano HPLC, Waters XBridge C4, flow rate 250 $\mu\text{l min}^{-1}$ step gradient 12–40–80% ACN Synapt G2Si, resolution mode. Reconstruction of average mass was done with MaxEnt1 software⁷⁶.

Metal determination by ICP-MS

To quantify Ni^{2+} and Mn^{2+} concentrations of the purified guanidinase, the samples were acid-digested and measured using ICP-MS. For acid digestion, HCl 30% (Supelco Suprapur, 100318, Merck), HNO_3 65% (3-fold subboiled, provided in analytically pure quality; 1.00441.1000, Merck) were used. H_2O_2 (31%, ROTIPURAN Ultra, HN69.1) was purchased from Carl Roth. Deionized water was produced with 0.075 $\mu\text{S cm}^{-1}$ using an Elga Veolia, PURELAB Chorus 3 RO. 180 μl of the sample was pipetted

into 7 ml PFA vials (Savillex), corresponding to a total sample amount of between 2 and 2.5 mg. Subsequently, 0.5 ml HCl and 1.5 ml HNO₃ were added. After closing the vials gas tight, they were heated to 120 °C on a hot plate (Savillex). The temperature was kept constant for 12 h. After the samples had cooled down to room temperature, a total of 500 µl of H₂O₂ was added in 50 µl steps. Vials were closed again and heated at 120 °C for 12 h. Subsequently, vials were opened and the samples were brought to dryness at 120 °C. After cooling, the digestions were dissolved in 2 ml HNO₃ and brought again to dryness at 140 °C. Finally, the digestions were dissolved in 1 ml HNO₃ and 2 ml deionized water. Vials were closed and heated again at 120 °C for 12 h to ensure complete dissolution. The digestions were then quantitatively transferred to 15 ml centrifuge tubes (polypropylene, metal free) and filled up to 10 ml with deionized water. Twofold dilutions of the digestions were measured with an Agilent 7900 Single Quad ICP-MS instrument (Agilent Technologies) in no-gas mode. The operation parameters for the plasma were set to the following values: RF power: 1,550 W; RF matching, 1.80 V; sample depth, 10 mm; nebulizer gas flow, 0.8 l min⁻¹; makeup/dilution gas, 0.4 l min⁻¹. The parameters for data acquisition were as follows: acquisition mode, spectrum; sweeps/replicate, 80; replicates, 3; integration time/mass, 0.1 sec. External calibration standards with an element concentration of 0.025 to 25 µg l⁻¹ were used for quantification. The limit of quantification values achieved for Mn²⁺ and Ni²⁺ were ≤0.17 µg l⁻¹ and ≤0.61 µg l⁻¹, respectively. The limit of detection (LOD) was ≤0.05 µg l⁻¹ for Mn²⁺ and ≤0.18 µg l⁻¹ for Ni²⁺. The measured concentrations of the diluted digestions ranged from 1.9 to 21.6 µg l⁻¹ for Mn²⁺ and from <LOD to 22.8 µg l⁻¹ for Ni²⁺. To exclude any contamination by the buffer used, this was also digested and measured. Here the concentrations ranged from <LOD to 0.07 µg l⁻¹ for Mn²⁺ and from <LOD to 0.29 µg l⁻¹ for Ni²⁺. Given that the concentrations of metals in all of the analysed samples were either significantly above the limit of quantification or below the LOD, the presence of these metals in the buffer solution was deemed not to have a relevant impact on the overall results.

Protein crystallization

For initial screening, guanidinase expressed with 0.5 µM Ni²⁺ was concentrated to 12.3 mg ml⁻¹ using the Amicon ultra centrifugal filter unit with 30 kDa MWCO and crystallized in MRC two-well crystallization plates with 50 µl of mother liquor set up using the TTP Labtech Mosquito pipetting robot system using the drop ratios 150 nl:200 nl and 200 nl:200 nl (protein:reservoir). Initial screens were performed using JCSG + HT, Index Screen, Morpheus Screen, PACT Premier screen and Crystal Screen at room temperature. Several hits were obtained from Crystal Screen and the condition F3, containing 0.5 M (NH₄)₂SO₄, 0.1 M Na₃ citrate pH 5.6 and 1 M Li₂SO₄ was used as a template for optimization screening by varying the (NH₄)₂SO₄ and Li₂SO₄ concentrations. The best crystals were obtained at 1 M (NH₄)₂SO₄ and 0.5 M to 0.7 M Li₂SO₄.

X-ray data collection, model building and refinement

Crystals were cryo-protected using 20% glycerol, flash-frozen in liquid nitrogen and diffraction datasets collected at beamline ID30B at the European Synchrotron Radiation Facility (ESRF, France) under cryogenic conditions. The collected datasets were processed with XDS and converted to the mtz file format using XDSCONV⁷⁷. The phase problem was solved with Phaser-MR⁷⁸, using its AlphaFold⁷⁹ prediction as a search model. The structure was further refined in iterative cycles of the manual model building using COOT⁸⁰ and maximum-likelihood refinement using the PHENIX software suite⁸¹. The final stages of refinement used the automated addition of hydrogens, and TLS refinement with one TLS group per chain. The models were validated with MolProbity⁸² and PDBREDO⁸³. Figures were created using PyMOL (The PyMOL Molecular Graphics System, v.2.0, Schrödinger) (Supplementary Table 12). Anomalous datasets were collected at ID30B at a wavelength of 1.8929 Å, close to the manganese anomalous scattering absorption

edge, and at a wavelength of 1.4825 Å, close to the nickel anomalous scattering absorption edge. The anomalous datasets were processed as described above and the obtained mtz files were refined using the finalized model obtained from the native dataset. The anomalous maps obtained from refinement were averaged using phenix.ncs_average supplying the refined pdb structure file and the corresponding anomalous map in ccp4 file format. Averaged anomalous maps were visualized using PyMol.

Substrate-dependent oxygen uptake measurements

Whole-cell substrate oxidation kinetics were determined from oxygen-uptake measurements as previously described^{34,48,84}. Here oxygen-uptake measurements were performed using a microrespirometry (MR) system equipped with a four-channel MicroOptode meter (Opto-F4 UniAmp) and O₂ MicroOptodes. Real-time O₂ concentration monitoring was supported through SensorTrace Rate software (Unisense).

N. inopinata biomass was cultivated in batch cultures in the same growth medium as described above and ammonium (1 mM) or urea (0.5 mM) as sole substrates. Ammonium and guanidine were also used as co-substrates and here ammonium grown cultures (1 mM) were supplemented with guanidine (10–20 µM) around 12 h before MR experiments to induce the expression of the guanidine transporter and the guanidinase. In all cases, active *N. inopinata* biomass was collected (3,000g, 6 min, 20 °C) from substrate replete cultures, washed and resuspended in identical but substrate-free medium, and incubated in a recirculating water bath (>30 min, 37 °C). Samples were taken for chemical analysis to ensure the absence of detectable ammonium, nitrite, nitrate and urea before MR experiments. All chemical species were determined photometrically as described above.

MR experiments were conducted in a glass MR chamber (~2 ml) containing a glass-coated magnetic stir bar, on an MR2 stirring rack (350 rpm), in a recirculating water bath (37 °C). MR chambers were over-filled with concentrated biomass to ensure the absence of a gaseous headspace, closed with an MR injection lid and submerged in the water bath. An O₂ MicroOptode was inserted into each MR chamber and left to equilibrate (~1 h), before a stable background signal was determined (15–30 min). The background rate of oxygen depletion was subtracted from all subsequent rate determinations in each MR chamber. A Hamilton syringe (10 µl; Hamilton) was used for all substrate (ammonium, urea, guanidine) injections. Both single- and multiple-trace oxygen uptake measurements were performed.

For single-trace measurements, a single-substrate injection was performed, and the oxygen uptake was recorded until complete substrate depletion in the presence of excess O₂ (>30 µM O₂). The single-injection scheme was used to determine the molar ratio of urea and guanidine consumed per O₂. The whole-cell kinetics of *N. inopinata* with urea and guanidine as substrates, respectively, were performed with single-injection traces. Here a single injection of urea (~20 µM) or guanidine (~20 µM) into the MR chamber was performed. Moreover, the whole-cell kinetics of total ammonium oxidation in *N. inopinata* precultivated with urea (0.5 mM) or ammonium plus guanidine (1 mM and ~20 µM) was determined with single-trace measurements. Here a single injection of NH₄Cl (~25 µM) was performed. In all cases, the experiments were halted after complete substrate depletion in the presence of excess O₂ (>30 µM O₂). Nitrate was the only detectable end product in all MR chambers used for whole-cell urea and guanidine kinetic calculations.

Multiple-trace measurements were used to determine the inhibitory effect of guanidine on the rate of maximum ammonium oxidation in *N. inopinata*. The maximum rate of ammonium oxidation was achieved with an initial injection of NH₄Cl (250–500 µM). Subsequently, several injections of varying guanidine concentrations were performed, and discrete slopes of oxygen depletion were calculated after each injection (~2–5 min).

In all cases, for both single and multiple injections, MR chamber contents were immediately centrifuged (19,000g, 15 min, 20 °C) after the measurements and the cell pellets and supernatant were stored separately for protein and chemical analysis, respectively (−20 °C). For protein analysis, the total protein content was determined photometrically using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). The chemical analyses (ammonium, nitrite, nitrate and urea) were performed as described above.

In vitro enzyme activity assay of guanidinase

Guanidine degradation by the heterologously expressed and purified guanidinase (in the presence of different Ni^{2+} concentrations; see above) was measured at 37 °C, pH 7.5, in a buffer containing 20 mM Tris-HCl and 50 mM NaCl by measuring urea production over 25 min of incubation. The measurements of the enzyme expressed in the presence of 1 mM Ni^{2+} were done in the presence of 1 mM Ni^{2+} . Kinetics were calculated from measurements at 50, 100, 250, 500, 1,000, 2,500, 5,000, 10,000, 25,000, 50,000 and 100,000 μM guanidine starting concentrations. For screening alternative substrate use, the guanidinase expressed in the presence of 1 mM Ni^{2+} was used. Then, 10 mM of methylguanidine, agmatine, arginine, creatine, guanidinobutyrate and guanidinopropionate each were incubated with the purified guanidinase enzyme or BSA at 37 °C for 30 min in three or six replicates. Guanidinase pH dependence was screened at 37 °C with incubations at pH 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10, 10.5 and 11 (set by addition of HCl or NaOH). Temperature dependence was screened at pH 7.5 with incubations at 14, 20, 28, 37, 46, 50, 55, 60, 65, 70, 80 and 90 °C. These incubations were done in triplicates.

Calculation of cellular substrate oxidation kinetic properties

The cellular kinetic properties of total ammonium, urea and guanidine oxidation were calculated from single-trace substrate-dependent oxygen uptake measurements. The substrate oxidation rates were calculated from oxygen uptake measurements using a substrate-to-oxygen consumption ratio. For total ammonium oxidation, a substrate-to-oxygen ratio of 1:2 was used. Single-trace experiments were used to confirm the substrate-to-oxygen ratio for urea (3.9 ± 0.31 , $n = 3$) and guanidine (6.17 ± 0.24 , $n = 4$) oxidation. Thus, for total urea oxidation and total guanidine oxidation, substrate-to-oxygen ratios of 1:4 and 1:6 were used, respectively. All substrate oxidation rates were normalized to total cellular protein in each MR chamber. In the case of total ammonium oxidation, the $K_{m(\text{app})}$ for unprotonated NH_3 was calculated based on the $K_{m(\text{app})}$ for total ammonium, incubation temperature, pH and salinity⁸⁵.

The cellular kinetic properties of total ammonium, urea and guanidine oxidation were determined with a Michaelis–Menten model fit to the data using equation (2) where V is the reaction rate (μM per mg protein per h), V_{max} is the maximum reaction rate (μM per mg protein per h), S is the total substrate concentration (μM), and $K_{m(\text{app})}$ is the reaction half saturation concentration (μM). An unconstrained nonlinear least-squares regression analysis was used to estimate the $K_{m(\text{app})}$ and V_{max} values^{86,87}.

$$V = (V_{\text{max}} \times [S]) \times (K_{m(\text{app})} + [S])^{-1} \quad (2)$$

The reaction half-inhibition concentration for total ammonium oxidation (K_i , μM), inhibition by guanidine, was also determined. The K_i was determined graphically with a Dixon plot analysis⁸⁸. Inverse total ammonium oxidation rates were plotted against total guanidine concentration. Total ammonium oxidation rates resulting in a linear trend were used for these analyses. Linear best fit trendlines from each biological replicate were used to determine intersection focal points and estimate K_i values. Furthermore, a linear regression of the percentage of the total ammonium oxidation rate at varying guanidine concentrations was used to determine the K_i .

The specific substrate affinity (a° ; litres per g wet cells per h) of ammonium, urea and guanidine oxidation was calculated using equation (3). The factor of 5.7 g wet cell weight per g of protein was used^{32,48,89}.

$$a^\circ = (V_{\text{max}} \times 5.7^{-1}) \times K_{m(\text{app})}^{-1} \quad (3)$$

WWTP community structure analyses

The Ribe WWTP (GPS coordinates: 55.33, 8.74) has biological N and P removal (enhanced biological phosphorus removal) and treats municipal wastewater with 20% industrial contribution (organic loading) corresponding to a total of 25,000 person equivalents. It is designed with recirculation and has return sludge sidestream hydrolysis. It does not have primary settling. Suspended solids around the time of sampling were $\sim 3.1 \text{ g l}^{-1}$. The Haderslev WWTP (GPS coordinates: 55.25, 9.51) has biological N and P removal and treats municipal wastewater with 5% industrial contribution corresponding to a total of 100,000 person equivalents. It is designed with alternating conditions and includes side stream hydrolysis. It does not have primary settling. Suspended solids around the time of sampling were $\sim 3.2 \text{ g l}^{-1}$. The Klosterneuburg WWTP (GPS coordinates: 48.29, 16.34) treats municipal wastewater corresponding to a total of 50,000 person equivalents with a two-stage, biological hybrid process. Suspended solids around the time of sampling were $\sim 4.4 \text{ g l}^{-1}$.

For characterizing the community structure, amplicon sequencing of the V1 to V3 regions of bacterial 16S rRNA genes was performed on samples from the Ribe and Haderslev WWTPs from the MiDAS BioBank collection. Applied PCR primers were 27F (5'-AGAGTTTGATCCTGG CTCAG-3') and 534R (5'-ATTACCGCGGCTGCTGG-3') with barcodes and Illumina adapters (IDT). PCR reactions (25 μl) were run in duplicate for each sample, using 1 \times PCR BIO Ultra Mix (PCR Biosystems), 400 nM of both the forward and reverse primer, and 10 ng template DNA. The PCR conditions were 95 °C for 2 min; followed by 20 cycles of 95 °C for 20 s, 56 °C for 30 s and 72 °C for 60 s; and a final elongation step at 72 °C for 5 min. The PCR products were purified using 0.8 \times CleanNGS beads and eluted in 25 μl nuclease-free water. The amplicon libraries were pooled separately in equimolar concentrations, diluted to 4 nM and paired-end sequenced (2 \times 300 bp) on the Illumina MiSeq sequencer using v3 chemistry (Illumina). A 20% phage PhiX control library was added to mitigate low-diversity library effects. The forward and reverse sequence reads were merged using the software *usearch*⁶⁰ with the *-fastq_mergepairs* command, filtered to remove phiX sequences using *usearch -filter_phix* and quality filtered using *usearch -fastq_filter* with parameter *-fastq_maxee* set to 1.0. Dereplication was performed by *usearch -fastx_uniques* with the option *-sizeout*, and amplicon sequence variants (ASVs) were resolved using the *usearch -unoise3* command. An ASV table was created by mapping the quality-filtered reads to the ASVs using the *usearch -otutab* command with the *-zotus* and *-strand* plus options. Taxonomy was assigned to ASVs using the *usearch -sintax* command with the parameters *-strand* both and *-sintax_cutoff* 0.8. The absence of comammox organisms in the sample from Klosterneuburg used for guanidine degradation measurements was confirmed by PCR using comammox clade A and clade B specific primer sets³⁸. Ribe and Haderslev sample DNA in the same concentration were used as positive controls.

Substrate incubation experiments with biomass from WWTPs

Activated sludge samples were collected from the aerated tanks of the Ribe and Haderslev WWTPs on 22 October 2021. Four litres of sludge from each WWTP were scooped into large sterile plastic bottles. The samples were transported to the laboratory on the same day, and were stored in the dark at ambient temperature, that is, ranging from 4 to 10 °C, until the incubations were started. The incubations with sludge from Ribe were started on the same day as collection, and the incubation with samples from Haderslev were started on the

day after collection. Before each incubation, the sludge was diluted approximately 1:4 as follows: the sludge was allowed to completely settle (1 h), then 1.5 l of the clear supernatant was gently collected to a new sterile flask without disturbing the flocs and, finally, 0.5 l of the remaining sludge was fully resuspended and added to the 1.5 l of supernatant. Well-mixed aliquots of 100 ml of the diluted sludge were then distributed to 200 ml sterile glass microcosms and covered with aluminium foil to enable gas exchange with the atmosphere. Substrates were added to the following final concentrations: guanidine, 50 μ M; ammonia, 150 μ M; and urea, 75 μ M. These different concentrations were chosen to account for the number of amino groups among the molecules. No substrate controls were also included. The samples were incubated at 23 °C with shaking at 100 rpm. All substrate and control treatments were performed in triplicate. Microcosms were subsampled immediately before and after initial substrate additions at T0. Additional subsamples from the Ribe incubation series were taken at 3, 6, 12 and 24 h. Additional subsamples from the Haderslev incubation series were taken at 2.5, 4, 8, 16, 24 and 48 h. Subsamples for metatranscriptomics were immediately flash-frozen with liquid-N₂ and stored at -80 °C until processing. Parallel samples (1 ml) for chemical analyses were centrifuged at 12,000g for 5 min, and the supernatant was taken and frozen immediately at -80 °C.

RNA extraction and purification

Total nucleic acids were extracted from activated sludge samples (500 μ l), which were thawed on ice and centrifuged (5 min, maximum speed, 4 °C), using the RNeasy PowerMicrobiome Kit (Qiagen) according to the manufacturer's instructions with the addition of phenol:chloroform:isoamyl alcohol (25:25:1) and β -mercaptoethanol (10 μ l ml⁻¹ final concentration). Bead beating (40 s at 6 m s⁻¹, four times with 2 min interval on ice) on the Fastprep FP120 (MP Biomedicals) system was performed for cell lysis instead of vortexing to improve lysis of bacteria with rigid cell walls. The total nucleic acid extracts were subjected to DNase treatment to remove DNA contaminants using the TURBO DNA-free kit (Invitrogen), and further cleaned up and concentrated with RNeasy XP beads (Beckman Coulter) before rRNA depletion. The integrity and quality of the purified total RNA were assessed on a Tapestation 2200 (Agilent) with the Agilent RNA ScreenTape (Agilent) system, and the concentration was measured using the Qubit RNA BR Assay Kit (Thermo Fisher Scientific). The average RNA integrity number was above 7.0 for all of the samples.

rRNA depletion, library preparation and sequencing

Total RNA was rRNA-depleted using the NEBNext rRNA Depletion Kit for Bacteria (New England Biolabs) with 100–300 ng total RNA as input. The NEBNext Ultra II Directional RNA Library Prep Kit (New England Biolabs) was used to prepare cDNA sequencing libraries according to the manufacturer's instructions. The libraries were pooled in equimolar concentration and 2.0 nM was sequenced on an S4 flow cell on the NovaSeq 6000 platform (Illumina) using the v1.5 300 cycle kit (Illumina, 20012863).

Identification of differentially transcribed genes

rRNA-depleted reads were adapter-screened, quality-filtered and mapped to published MAGs using bbmap v.38.92. Adapter removal and quality filtering was conducted using bbdduk (ktrim=r k=21 mink=11 hdist=2 minlen=119 qtrim=r trimq=15). Metatranscriptomic reads from WWTP Ribe were mapped to genome accession GCA_016722055.1 and reads from WWTP Haderslev were mapped to genome accession GCA_016712165.1, which were the dominant comammox MAGs in the respective WWTP³⁷. Both mappings were carried out using bbmap (minid=0.98 idfilter=0.98 ambiguous=toss pairedonly=t killbadpairs=t mappedonly=t bamscript=bs.sh) to produce bamfiles. Counts for each gene were calculated using bedtools coverage (-counts) using BAM files from bbmap and GFF files downloaded from GenBank for each

genome. Counts for each coding gene were examined for potential outliers, which identified MBK8278324.1 and MBK9947797.1 as potentially misannotated small RNAs and were removed from subsequent calculations. Differential transcription was evaluated by treating different timepoints as replicates and comparing treatments as factors using DESeq2⁹⁰. TPM was calculated and used for visualization purposes only.

Soil incubations

Soil was collected from a long-term fertilization experiment managed by the Austrian Agency for Health and Food Safety located at the Ritzlhof field experiment (48° 11' 17.9" N 14° 15' 16.5" E) in May 2023. The soil is classified as a Cambisol and has been fertilized since 1991 with solid cattle manure at an application rate of 525 kg N per ha per year⁹¹. Soil incubations were conducted in 125 ml Wheaton bottles capped with grey butyl stoppers. In brief, 30 g soil was added to each replicate ($n = 3$) bottle, and amended with 820 μ l water, ammonium, guanidine or ammonium + guanidine for a final concentration of 30 μ g N per g dry-weight soil. Soils were incubated at 23 °C and sampled at 0, 1, 2, 3, 5, 7, 12 and 27 days. Acetylene (0.02%, v/v) was used to inhibit all lithotrophic ammonia oxidation. Acetylene was supplied by adding 0.3 ml of 10% acetylene gas to sealed bottles. Bottles were opened every 1–3 days, and acetylene was resupplied. For chemical analyses, around 2 g soil was extracted in water and 2 M KCl and extracts were frozen at -20 °C until analysis. Nitrate and nitrite were quantified in water extracts and ammonium, urea, and guanidine were quantified in KCl extracts as described above. Approximately 1 g soil was sampled for molecular analysis and was frozen at -80 °C until analysis. DNA extracts were performed using the ZymoBIOMICS DNA/RNA Miniprep Kit according to the manufacturer's instructions. AOB, AOA and comammox clade A and B *amoA* qPCRs were carried out as previously described^{38,92,93}.

Statistics

Statistical analysis on chemical, protein and qPCR data from physiological experiments and WWTP sample incubations were performed using two-tailed *t*-tests in SigmaPlot v.14.5 and R. No statistical methods were used to predetermine sample size, and blinding and randomization of samples were not used.

Inclusion and ethics statement

All collaborators on this study fulfil the criteria for authorship required by Nature journals, they have been included as authors as their work was essential in designing and performing the study. The roles and responsibilities were agreed among collaborators ahead of the research. No living animals or animal-derived material were used in this study, except dropped animal manure and urine. Animals were not forced to excrete.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The crystal structures have been deposited in the Protein Data Bank under accession code 9FEK. The anomalous scattering datasets are available at the European Synchrotron Radiation Facility⁹⁴ (<https://data.esrf.fr/doi/10.1515/ESRF-DC-1801440672>). WWTP metatranscriptome reads have been deposited under BioProject ID PRJNA1118285. The MS proteomics data have been deposited at the ProteomeXchange Consortium via the PRIDE⁹⁵ partner repository under dataset identifier PXD038826. A high-resolution version of Extended Data Fig. 4a is available at Figshare⁹⁶ (<https://doi.org/10.6084/m9.figshare.26139127>). Source data are provided with this paper.

59. Mistry, J., Finn, R. D., Eddy, S. R., Bateman, A. & Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* **41**, e121 (2013).
60. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
61. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
62. Nawrocki, E. P. & Eddy, S. R. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933–2935 (2013).
63. Lenkeit, F., Eckert, I., Hartig, J. S. & Weinberg, Z. Discovery and characterization of a fourth class of guanidine riboswitches. *Nucleic Acids Res.* **48**, 12889–12899 (2020).
64. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
65. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
66. Kim, J.-G. et al. Hydrogen peroxide detoxification is a key mechanism for growth of ammonia-oxidizing archaea. *Proc. Natl Acad. Sci. USA* **113**, 7888–7893 (2016).
67. Hood-Nowotny, R., Umana, N. H.-N., Inselbacher, E., Oswald-Lachouani, P. & Wanek, W. Alternative methods for measuring inorganic, organic, and total dissolved nitrogen in soil. *Soil Sci. Soc. Am. J.* **74**, 1018 (2010).
68. Geyer, J. W. & Dabich, D. Rapid method for determination of arginase activity in tissue homogenates. *Anal. Biochem.* **39**, 412–417 (1971).
69. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* **18**, 1403–1414 (2016).
70. Apprill, A., McNally, S., Parsons, R. & Weber, L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* **75**, 129–137 (2015).
71. Vallenet, D. et al. MicroScope: a platform for microbial genome annotation and comparative genomics. *Database* **2009**, bap021 (2009).
72. Zhu, Y. et al. DEqMS: a method for accurate variance estimation in differential protein expression analysis. *Mol. Cell. Proteom.* **19**, 1047–1057 (2020).
73. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
74. Mlynek, G., Kostan, J., Leeb, S. & Djinić-Carugo, K. Tailored suits fit better: customized protein crystallization screens. *Cryst. Growth Des.* **20**, 984–994 (2020).
75. Kotov, V. et al. In-depth interrogation of protein thermal unfolding data with MoltenProt. *Protein Sci.* **30**, 201–217 (2021).
76. Ferrige, A. G. et al. Disentangling electrospray spectra with maximum entropy. *Rapid Commun. Mass Spectrom.* **6**, 707–711 (1992).
77. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
78. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
79. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
80. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
81. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
82. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
83. Joosten, R. P. et al. PDB_REDO: automated re-refinement of X-ray structure models in the PDB. *J. Appl. Crystallogr.* **42**, 376–384 (2009).
84. Martens-Habbena, W. & Stahl, D. A. Nitrogen metabolism and kinetics of ammonia-oxidizing archaea. *Methods Enzymol.* **496**, 465–487 (2011).
85. Clegg, S. L. & Whitfield, M. A chemical model of seawater including dissolved ammonia and the stoichiometric dissociation constant of ammonia in estuarine water and seawater from –2 to 40 °C. *Geochim. Cosmochim. Acta* **59**, 2403–2421 (1995).
86. Kemmer, G. & Keller, S. Nonlinear least-squares data fitting in Excel spreadsheets. *Nat. Protoc.* **5**, 267–281 (2010).
87. Gokulakrishnan, S. & Gummadi, S. N. Kinetics of cell growth and caffeine utilization by *Pseudomonas* sp. GSC 1182. *Process Biochem.* **41**, 1417–1421 (2006).
88. Dixon, M. The determination of enzyme inhibitor constants. *Biochem. J.* **55**, 170–171 (1953).
89. Martens-Habbena, W., Berube, P. M., Urakawa, H., de la Torre, J. R. & Stahl, D. A. Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* **461**, 976–979 (2009).
90. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
91. Hendricks, S. et al. Agricultural management affects active carbon and nitrogen mineralisation potential in soils. *Z. Pflanzenernähr. Bodenkd.* **185**, 513–528 (2022).
92. Alves, R. J. E. et al. Nitrification rates in Arctic soils are associated with functionally distinct populations of ammonia-oxidizing archaea. *ISME J.* **7**, 1620–1631 (2013).
93. Rothauwe, J. H., Witzel, K. P. & Liesack, W. The ammonia monooxygenase structural gene *amoA* as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing populations. *Appl. Environ. Microbiol.* **63**, 4704–4712 (1997).
94. Puehringer, D. & McCarthy, A. *Anomalous Datacollection of Guanidinase from Nitrospira inopinata* Ver. 1 [Dataset] (European Synchrotron Radiation Facility, 2024); <https://doi.org/10.1515/ESRF-DC-1801440672>.
95. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).
96. Palatinszky, M. et al. Data for ‘Growth of complete ammonia oxidizers on guanidine’. *Figshare* <https://doi.org/10.6084/m9.figshare.26139127> (2024).
97. Landau, M. et al. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* **33**, W299–W302 (2005).

Acknowledgements We thank the staff of the ESRF and EMBL Grenoble, in particular A. McCarthy, for assistance and support in using beamline(s) ID30B under proposal number MX2591, the staff at the Mass Spectrometry Facility at Max Perutz Labs for proteomics analyses using the Vienna BioCenter Core Facilities (VBCF) instrument pool; V. Pevala for thermostability analysis of *N. inopinata* guanidinase; S. Hofbauer and P. Furtmüller as well as A. Sponga for their help with the kinetic modelling; M. Stockhausen for ICP-MS analysis; A. Schintlmeister for support with the NanoSIMS analyses; A. Spiegel and T. Sandén for providing access to the soil sample; and E. Selberherr and G. Terler for providing access to animal manure samples. N.J. thanks M. von Bergen and the funding of the UFZ for the ProMetheus platform for proteomics. This work was made possible in part by the OpenMIMS software, the development of which is funded by the NIH/NIBIB National Resource for Imaging Mass Spectrometry, NIH/NIBIB 5P41 EB001974-10. We are grateful for support by the Wittgenstein grant of the Austrian Science Fund FWF (Z383-B to M.W.), the FWF Young Investigators Research Grant program (ZK74; to C.J.S. and A.T.G.), the FWF grant P30570-B29 (to H.D.), the Comammox Research Platform of the University of Vienna (to H.D., M.W. and K.D.-C.), the Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology (to K.D.-C.), the Vienna Science and Technology Fund Chemical Biology project LS17-008 (to K.D.-C.), and the Austrian-Slovak Interreg Project B301 StruBioMol (to K.D.-C.). M.K.D.D. and P.H.N. were supported by the Villum Foundation (grant 16558). This research was funded in part by the Austrian Research Fund (FWF) (<https://doi.org/10.55776/COE7>) (to M.W., K.K., H.D., K.D.-C., M.Z. and T.H.). For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this publication.

Author contributions Conceptualization: M.W., H.D., K.D.-C., P.H.N., K.K. and A.T.G. Methodology: R.G. and M.Z. (guanidine measurements). Data curation: C.W.H., K.D.-C. and D.P. Formal analysis: D.P., C.W.H., C.J.S., M.K.D.D., M.P., A.T.G. and K.K. Investigation: M.P. (AOB activity, enzyme kinetics, enzyme specificity, comammox growth experiments), C.J.S. (whole-cell kinetics), N.J. (comammox proteome), D.P. (guanidinase structure and metal analyses), N.K. and J.K. (heterologous expression and protein purification), K.W. (WWTP experiments), C.S. (protein purification), J.P. (cultivation, qPCR), T.H. (guanidinase metal content), M.K.D.D. (WWTP amplicon analyses), A.L. (nanoSIMS analyses), K.K. (growth experiments, nanoSIMS), A.T.G. (soil incubations) and C.W.H. (bioinformatics analyses). Visualization: C.W.H., D.P., J.K., M.P. and K.K. Funding acquisition: M.W., H.D., C.J.S., K.D.-C., P.H.N., M.Z., K.K., T.H. and A.T.G. Project administration: M.W. Supervision: M.W., H.D., K.D.-C., D.P., M.Z. and K.K. Writing—original draft: M.W., C.W.H., M.P., C.J.S., H.D., D.P., K.D.-C. and K.K. All of the authors contributed to reviewing and editing the manuscript.

Competing interests The authors declare no competing interests.

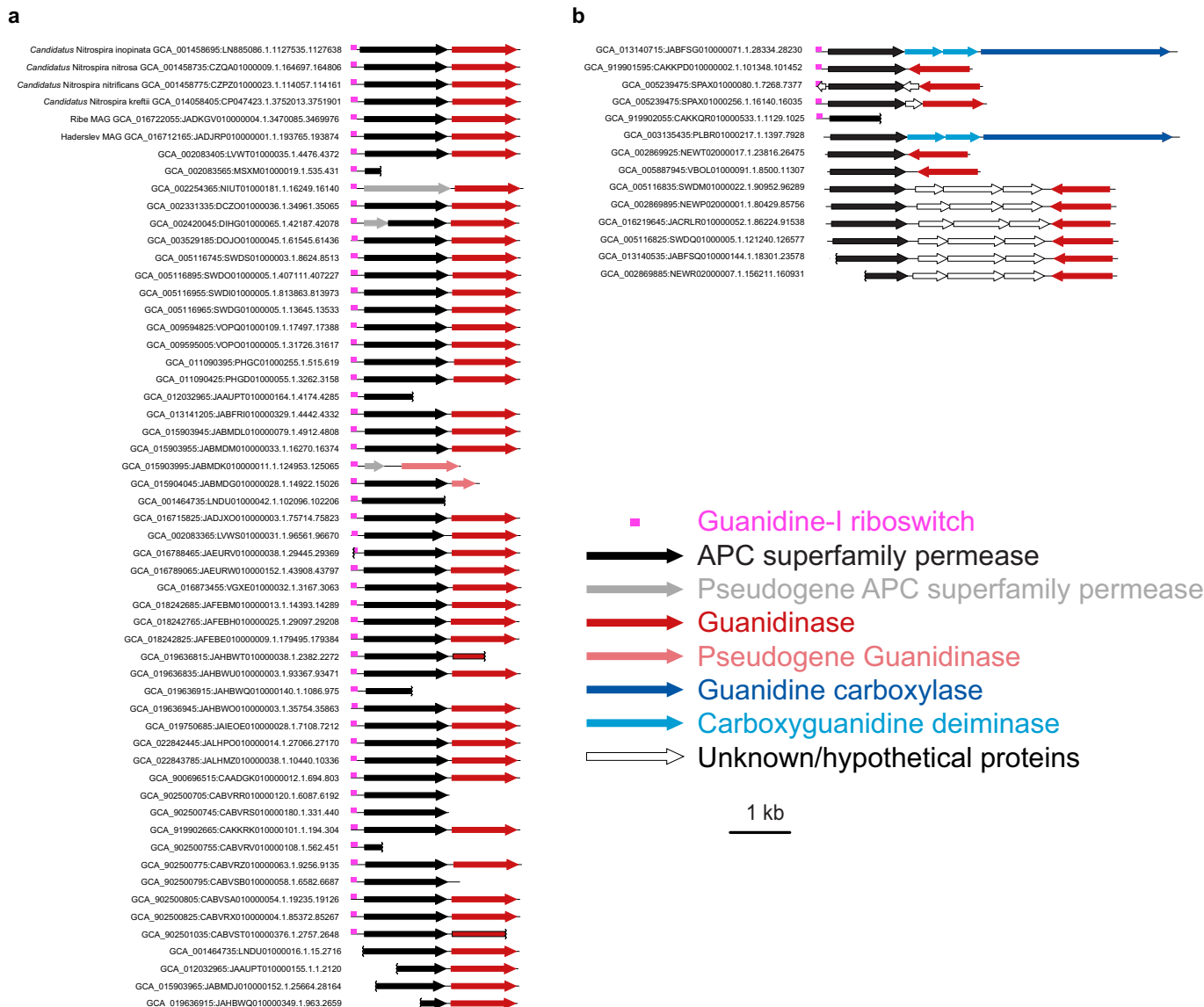
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-024-07832-z>.

Correspondence and requests for materials should be addressed to Michael Wagner.

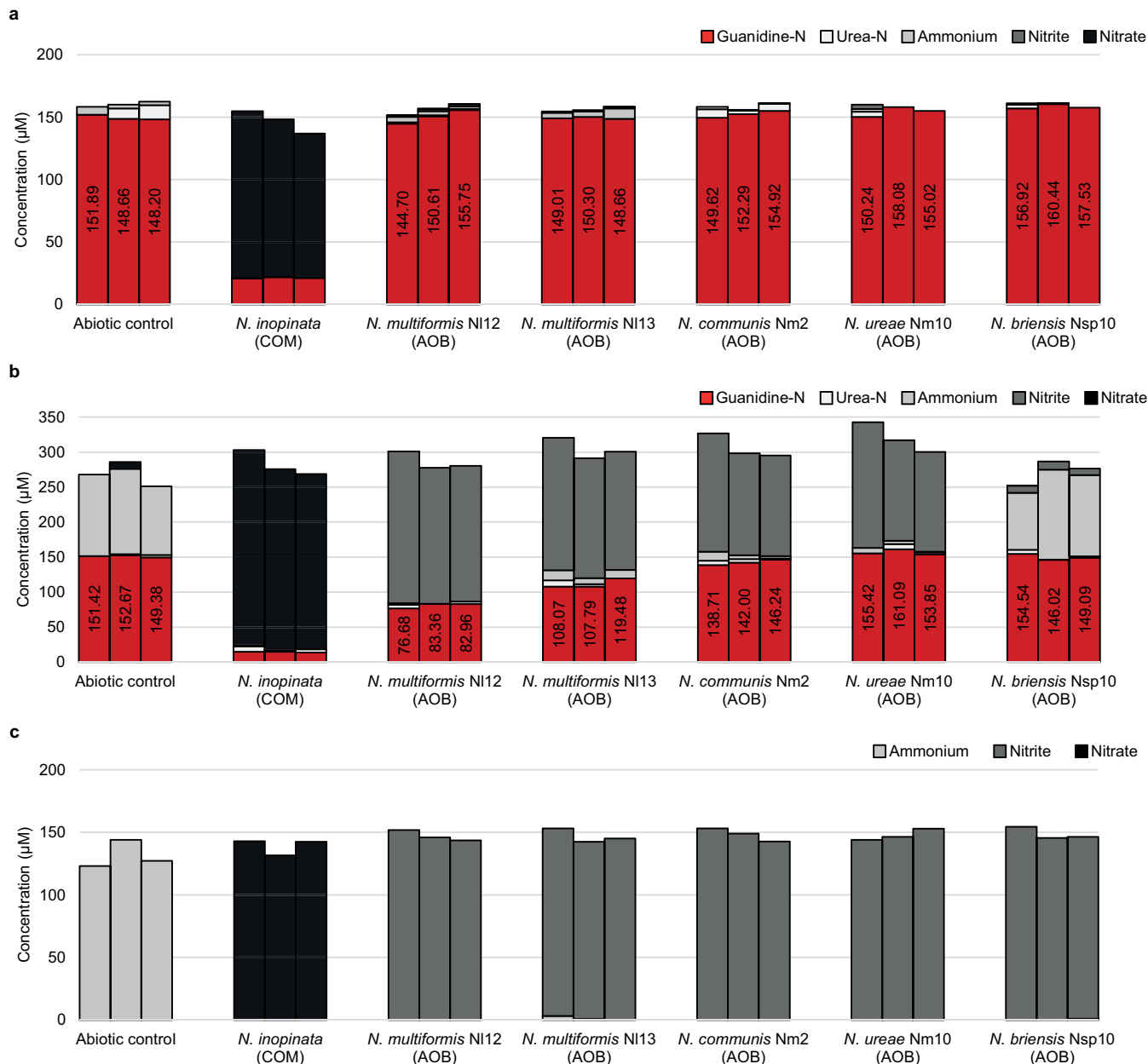
Peer review information Nature thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



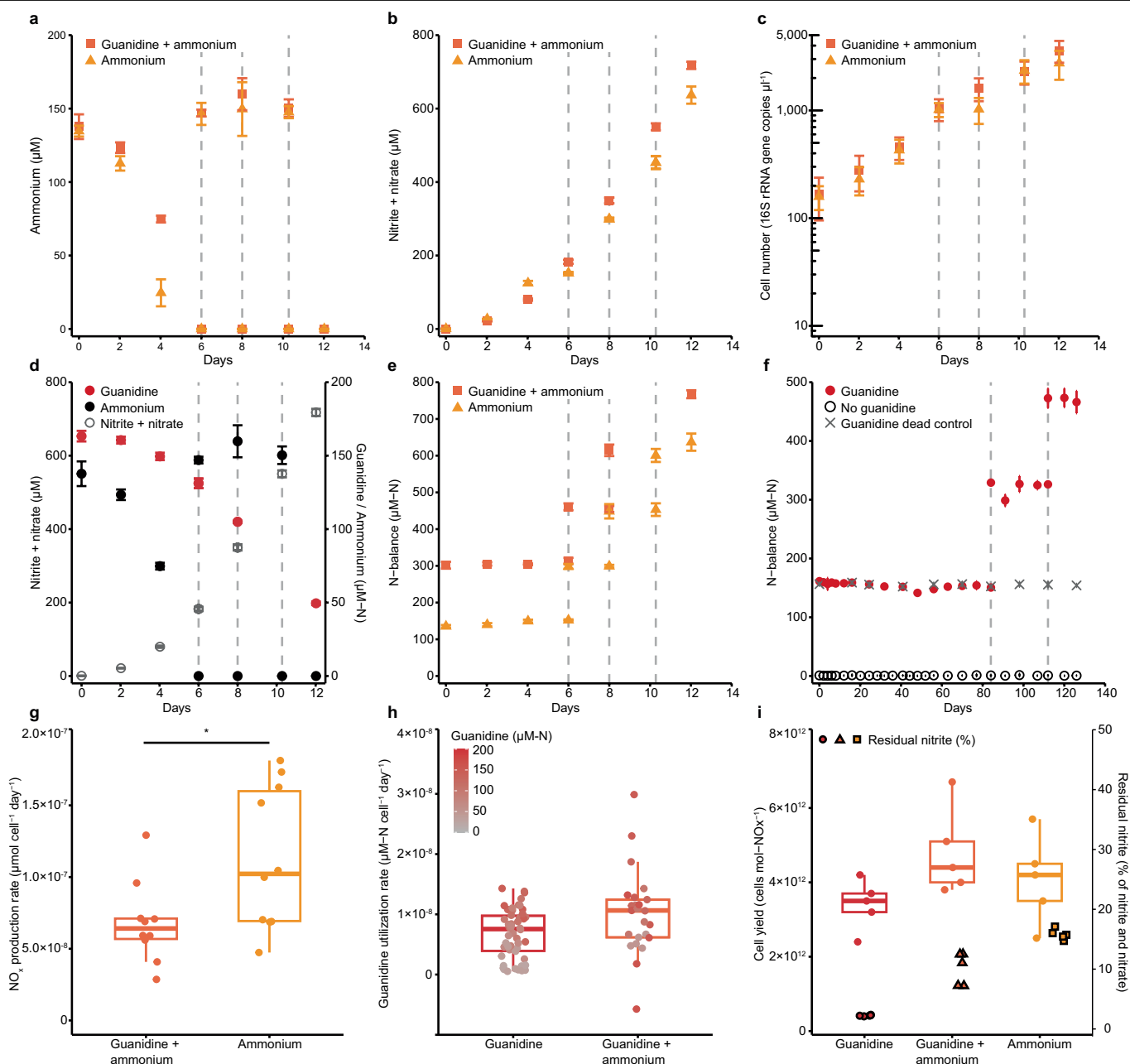
Extended Data Fig. 1 | Co-occurrence of guanidine-I riboswitch, APC superfamily permease, and guanidinase genes on genome scaffolds from comammox *Nitrospira* group A (a) and B (b). We identified a guanidine-I riboswitch in 67% (56/83) of the screened comammox genomes and found that the presence of the riboswitch was more widespread in comammox group A genomes (52/63; including the two metagenome-assembled comammox genomes from the wastewater treatment plants Ribe and Haderslev) than comammox group B genomes (4/20). The majority (75/83) of screened comammox genomes encoded urease. All comammox genomes with a guanidine-I riboswitch (56/56) encode an APC superfamily permease immediately downstream from the guanidine-I riboswitch, and the majority (46/56) encode guanidinase immediately downstream from the permease. It should be noted that the majority of comammox genomes are MAGs that are incomplete, preventing the triad of a guanidine-I riboswitch, an APC transporter, and a guanidinase from being located together in every genome. More specifically, for nine cases in which guanidinase is not found immediately downstream from the permease, the scaffold edge can be found within the permease or immediately downstream from the permease. In the final case, the genes immediately downstream from the riboswitch and permease are genes

involved in the guanidine carboxylase pathway. Labels indicate the genome and scaffold accession IDs as well as genomic coordinates. For scaffold neighbourhoods with a riboswitch, the genomic coordinates indicate the coordinates of the riboswitch. For entries without a riboswitch, the accession IDs and end-point coordinates of the displayed genomic segment region are indicated. All genome regions visualized have been oriented so that the riboswitch and permease would be transcribed from left to right. Grey (permease) and light red (guanidinase) indicate genes that contain extensive frameshifts and therefore could be pseudogenes. We also screened all comammox genomes for the presence of the other three known guanidine riboswitches (mini-*ykkC*, *ykkC*-III and IV⁶³) and examined whether genes involved in guanidine degradation were located nearby (Supplementary Table 3). Two additional guanidine-responsive riboswitches, the mini-*ykkC* and *ykkC*-III riboswitches, were identified in five and two comammox genomes, respectively, and no genes known to be involved in guanidine degradation, transportation nor nickel incorporation were found in the vicinity of these. More generally, no obvious, dedicated nickel loading protein could be detected in any of the comammox genomes.



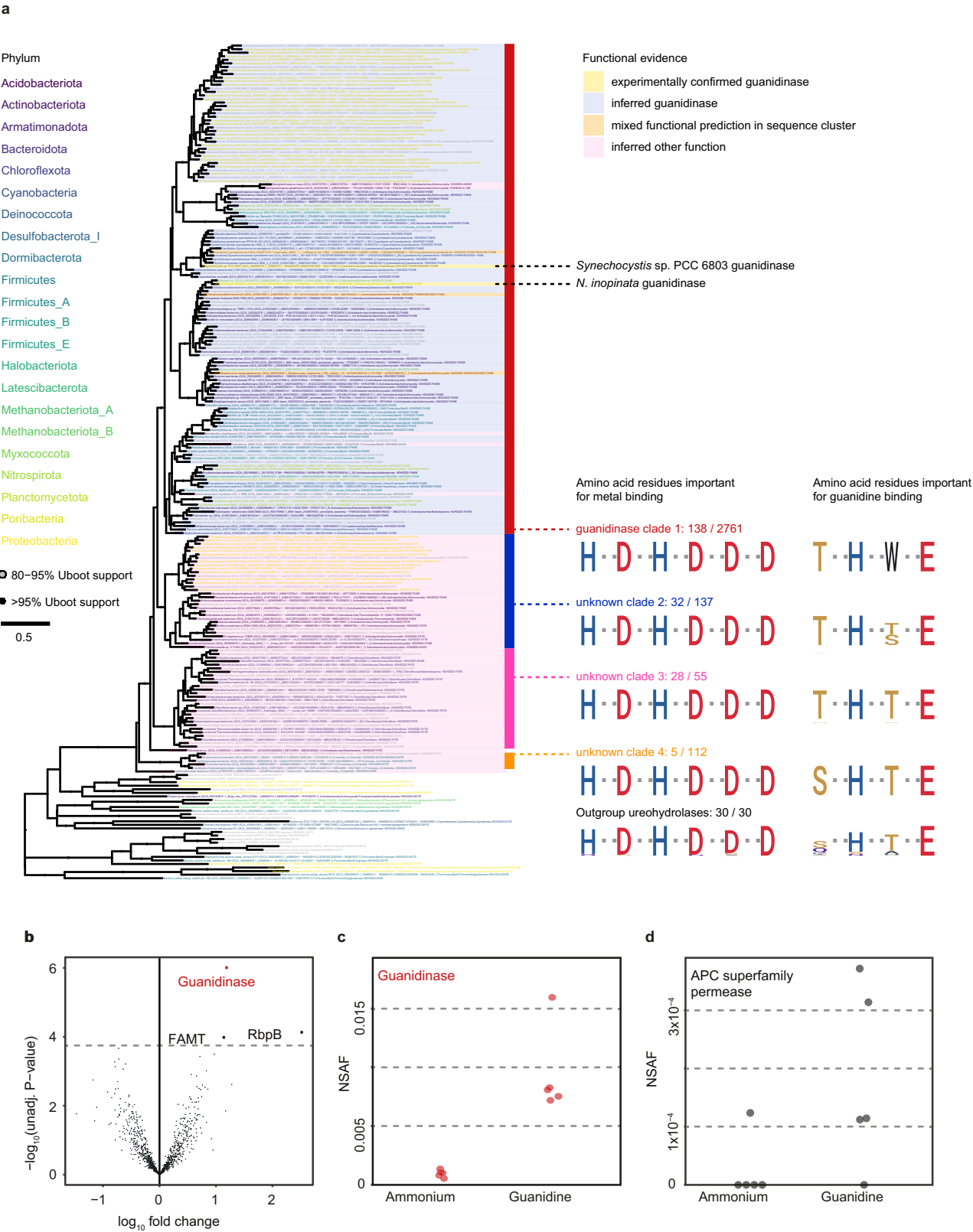
Extended Data Fig. 2 | Potential of guanidine utilization by selected ammonia-oxidizing microbes. Nitrogen balances for individual replicate cultures after 2 weeks of incubation with (a) 50 µM guanidine; (b) 50 µM guanidine plus 150 µM ammonium; and (c) 150 µM ammonium as the only substrate (activity control). Genomes of all five AOB strains harbour the complete genetic repertoire for guanidine transportation and utilization (guanidine/organocation transporter; guanidine carboxylase with signature amino acids^{15,16}; carboxyguanidine deiminase; allophanate hydrolase; urease); only *Nitrosomonas communis* Nm2 lacks urease. *N. inopinata* possesses a guanidine/organocation transporter, a guanidinase, and a urease, and is

among the tested organisms the only strain that was able to utilize guanidine in the absence of ammonium. In the presence of ammonium, three AOB strains were also able to significantly degrade guanidine (two-tailed student's *t*-test, d.f. = 4 for all): *N. multiformis* NI12 ($t = 24.932$, $P = 0.0000154$), *N. multiformis* NI13 ($t = 9.482$, $P = 0.000690$), and *N. communis* Nm2 ($t = 3.837$, $P = 0.0185$). All incubations were done in triplicates. Starting biomass was set equal among all strains, replicates, and treatments (with equal protein biomass density of the AOB strains as the comammox culture which showed rapid guanidine consumption).



Extended Data Fig. 3 | *N. inopinata* growth experiments. (a) Ammonium consumption over time in treatments containing either ammonium and guanidine, or ammonium only. 150 μM N for both guanidine and ammonium were added to a washed culture of *N. inopinata* (after pre-incubation with guanidine and ammonium for several months) and incubated for 12 days. On days 6, 8 and 10, additional spikes of around 150 μM of ammonium were added to the incubations (dashed grey lines). (b) Nitrite and nitrate production (combined). (c) 16S rRNA gene copy numbers measured by qPCR. (d) Concomitant utilization of guanidine and ammonium in treatments containing both substrates. (e, f) N balance for treatments receiving guanidine and ammonium, and ammonium only (e), and for guanidine, no guanidine (starved) and guanidine dead control (f). N-balances include guanidine (where added), ammonium, nitrite and nitrate. Urea concentrations were not included as they were always <2.5 μM . (g) Nitrite and nitrate (combined) production rate per cell and day in treatments receiving guanidine and ammonium, or ammonium only. Rates were calculated for time intervals before complete ammonium depletion (incl. day 4), by normalizing the difference in NO_x concentration to average cell numbers between time points and the duration of the time interval. * *N. inopinata* NO_x production when incubated with guanidine and ammonium versus ammonium only, Welch two sample *t*-test:

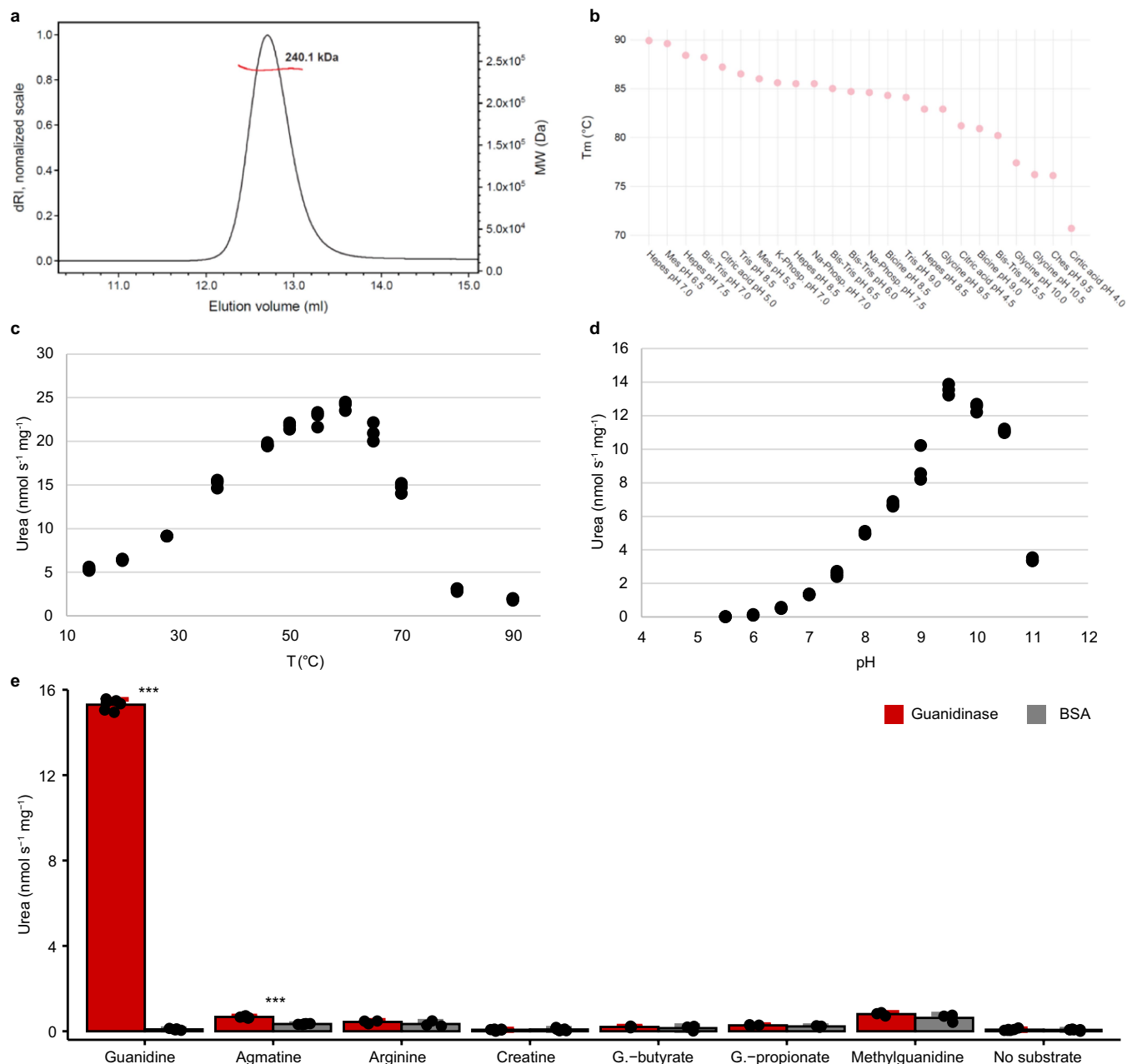
$t = -2.4714$, d.f. = 14.174, $P = 0.02673$. (h) Guanidine utilization rate per cell and day in treatments receiving only guanidine, or guanidine and ammonium. Rates were calculated across all time intervals as guanidine concentrations never dropped to 0, by normalizing the difference in guanidine concentration to average cell numbers between time points and the duration of the time interval. Colour gradient indicates the average guanidine concentration between time intervals. Higher utilization rates coincided with higher guanidine concentrations. No significant differences were found between treatments (Welch two sample *t*-test). (i) Cell yield per mol nitrite and nitrate produced across treatments, calculated between beginning and end of the incubation (12 days for treatments receiving guanidine and ammonium, or ammonium only; 126 days for the treatment receiving guanidine only). Note that all treatments contained residual nitrite (given as %), which may affect the overall energy conserved, and thus the cell yield per mol nitrite and nitrate produced. No significant differences were found between treatments (anova). All experiments were done in five biological replicates, datapoints show means, error bars standard deviations. Boxplots in (g-i) depict the 25–75% quantile range, with the centre line depicting the median (50% quantile); whiskers encompass data points within 1.5x the interquartile range.



Extended Data Fig. 4 | See next page for caption.

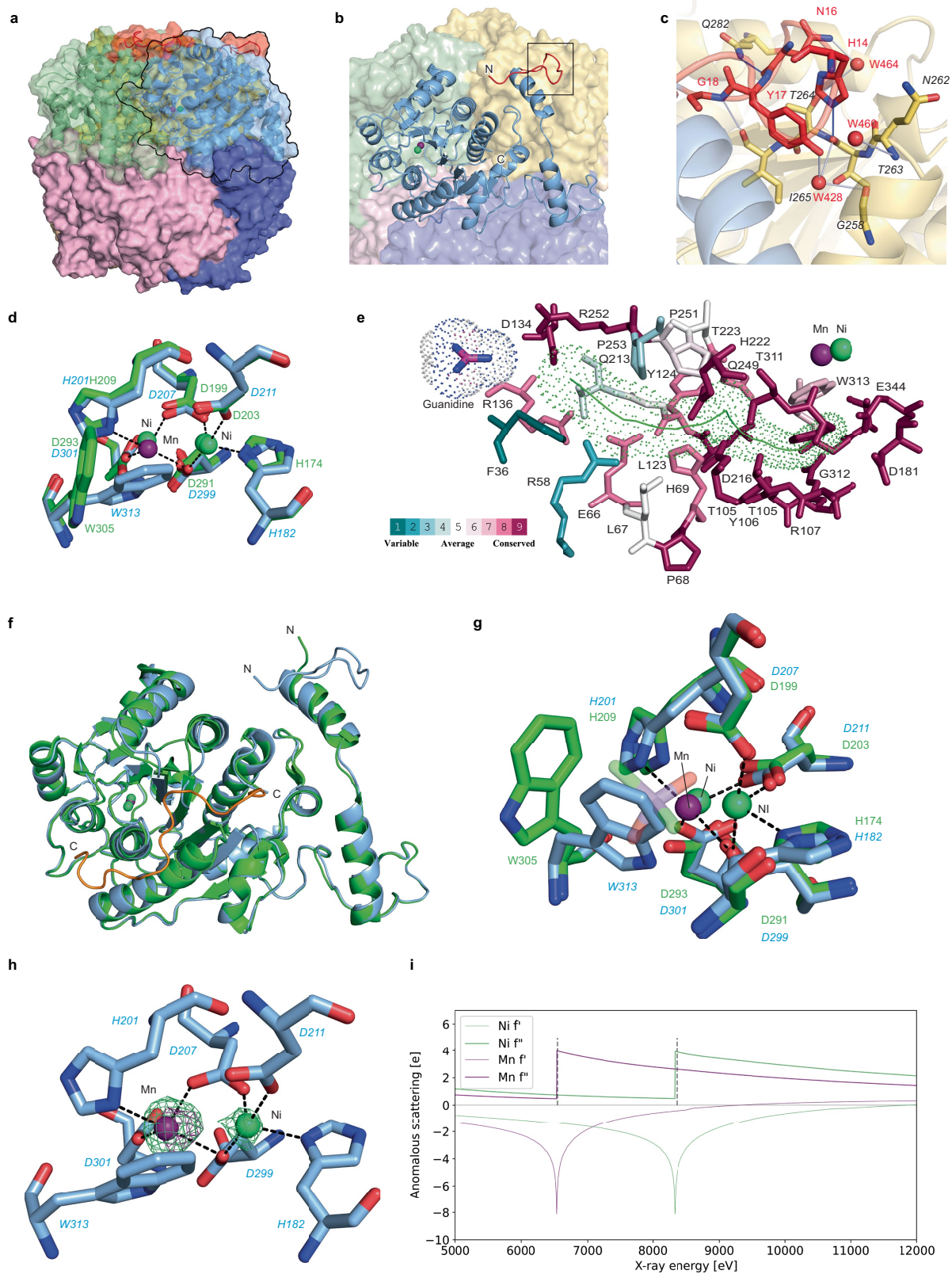
Extended Data Fig. 4 | Phylogeny of guanidinases and proteomics of *N. inopinata* grown on guanidine and ammonium. (a) Phylogeny of guanidinases according to an HMM alignment based on pfam model PF0049, using other biochemically characterized ureohydrolase family members as outgroup. The tree has been split into four general clades for simplification, which are indicated with coloured side bars. Specific residues that coordinated with nickel/manganese ions (*N. inopinata* guanidinase positions H182, D207, H209, D211, D299, D301) or have been proposed by Funck et al.¹⁸ to be important for guanidine hydrolysis (*N. inopinata* guanidinase positions T105, H222, W313, E344) of guanidinase are indicated as a logo for each of the general clades as well as the outgroup (for details see Supplementary Tables 4 and 9). Each shown tip label is a centroid representative of a cluster of sequences (90% identity over 90% of HMM alignment). The logo title indicates an arbitrary clade name n/N where n is how many representatives in the shown tree belong to the clade and N is the total number of sequences represented by the centroids in the tree. Tip labels are coloured according to the phylum-level classification of the centroid representative. Guanidinase-specific residues were used to classify each centroid representative as “inferred guanidinase” (all members possess

equivalent of T105 and W313), inferred not guanidinase (no members possess equivalent of both T105 and W313). “Mixed” refers to centroids that represent a cluster with mixed inferred function. The two experimentally characterized guanidinases from *Synechocystis* sp. PCC 6803 and *N. inopinata* are indicated. A high-resolution version of the phylogenetic tree can be found under <https://doi.org/10.6084/m9.figshare.26139127>. (b) Volcano plot of the differential expression of proteins in *N. inopinata* according to DEqMS⁷². Each point represents a single protein. Positive fold-change corresponds to higher expression levels during growth on guanidine, and negative fold-change corresponds to higher expression during growth on ammonium. Horizontal dashed lines show the significance threshold for *P*-values according to the Benjamini-Hochberg procedure with a false discovery rate of 0.05. Guanidinase (CUQ66148.1), RbpB (CUQ66942.1) and FAMT (CUQ67192.1) showed significant differential expression ($\log_{10}[\text{FC}] = 1.2, P_{\text{adj}} = 0.001$, $\log_{10}[\text{FC}] = 2.5, P_{\text{adj}} = 0.036$, and $\log_{10}[\text{FC}] = 1.1, P_{\text{adj}} = 0.036$, respectively). Protein expression of the (c) guanidinase and (d) APC superfamily permease in terms of normalized spectral abundance factor (NSAF). All experiments were performed in five replicates. Dots in (b) and (c) show data for individual replicates.



Extended Data Fig. 5 | Oligomeric state thermostability, pH and temperature dependence and possible alternative substrates of the heterologously expressed *N. inopinata* guanidinase. (a) Size exclusion chromatography coupled to multi-angle-light scattering (SEC-MALS) elution profile of the guanidinase. The red line crossing the SEC profile indicates the molecular mass of the protein. Guanidinase forms hexamers in solution as the calculated molecular mass of a subunit is 41.66 kDa. (b) Melting temperatures (T_m) of guanidinase in various buffer/pH conditions estimated using the two-state reversible unfolding model implemented in MoltenProt⁷⁵. 24 buffer/pH conditions with the highest T_m are shown. (c) Temperature dependence of heterologously expressed *N. inopinata* guanidinase at pH 7, 10 mM guanidine

starting concentration. (d) pH dependence of heterologously expressed *N. inopinata* guanidinase at 37 °C, 1 mM guanidine starting concentration. (e) Urea production by heterologously expressed *N. inopinata* guanidinase from guanidine and different guanidino compounds. 10 mM of substrates were incubated with guanidinase or BSA at 37 °C. Bars show the means, error bars the standard deviation of 3 replicates (for arginine, guanidinobutyrate, guanidinopropionate, methylguanidine) or 6 replicates (guanidine, except BSA control 5 replicates; agmatine, creatine, no substrate). *** Urea production from guanidino compounds when incubated with guanidinase versus bovine serum albumin (BSA). Guanidine one-sided t -test: $t = 133.92635$, d.f. = 9, $P = 1.83 \times 10^{-16}$. Agmatine one-sided t -test: $t = 17.44749$, d.f. = 10, $P = 4.06 \times 10^{-9}$.

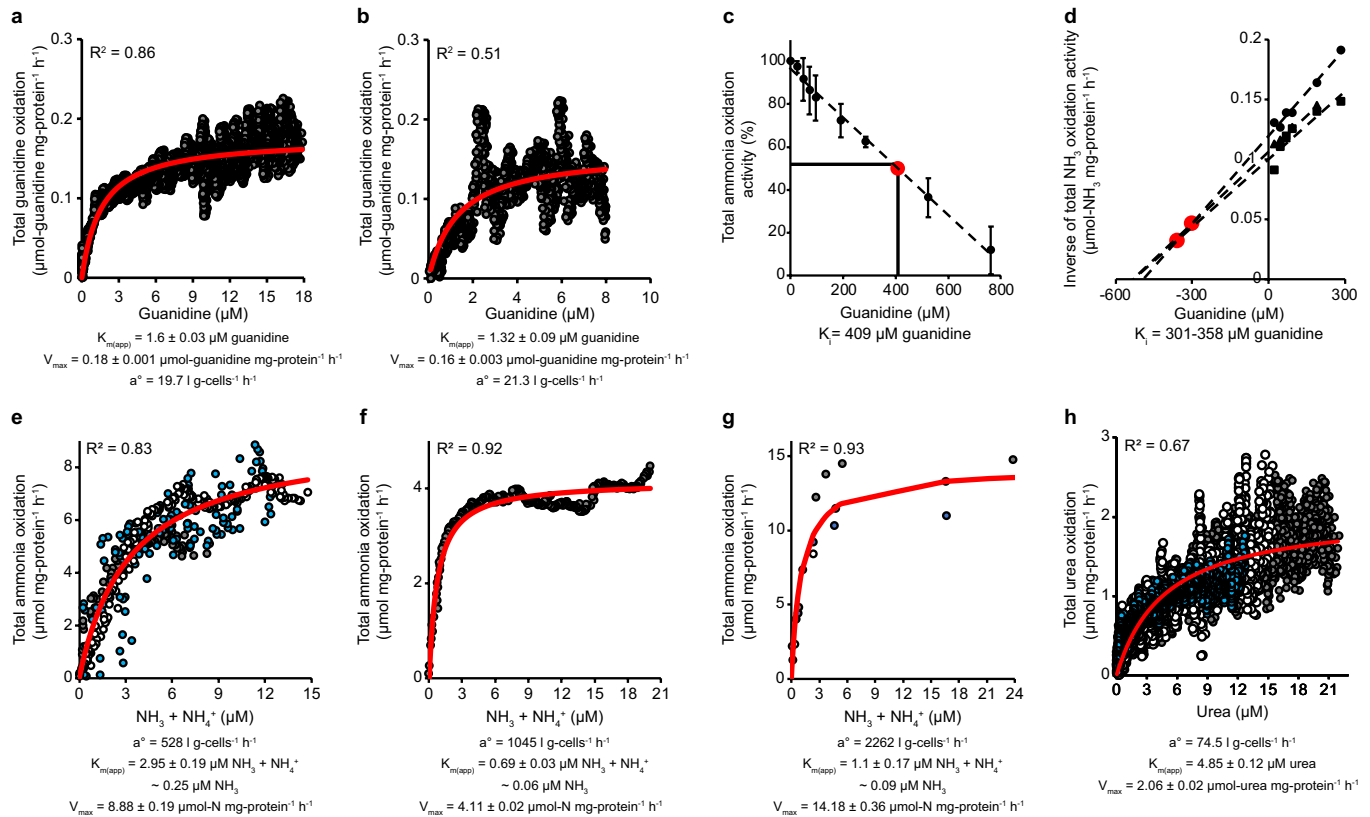


Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Structure of the *N. inopinata* guanidinase and comparison to the *Synechocystis* guanidinase. (a) Overview of the

N. inopinata guanidinase hexamer shown with surface representation, subunits are individually coloured. Nickel ions are shown as green, manganese ions as purple spheres. The N-terminal extension of 14 amino acids which partly protrudes together with the first N-terminal helix over the neighbouring subunit is highlighted in red (PDB ID 9FEK). (b) Cartoon representation of the *N. inopinata* guanidinase subunit with the colour code for the N-terminal extension and metal ions as in (a). (c) Close up view of the interaction of the *N. inopinata* N-terminal extension (red cartoon and sticks) with the residues of the neighbouring subunit (yellow cartoon and sticks) through three hydrogen bonds and four water bridges (Supplementary Table 7). Coordinated waters are shown as red spheres. (d) Superposition of the active sites of guanidinase from *N. inopinata* (green) and the *Synechocystis* GdmH (7OII) (cyan). Nickel ions are shown as green, manganese ions as purple spheres. (e) Model of guanidine and the tunnel shown as in Fig. 2d. Residues lining the tunnel are shown as sticks and colour coded according to conservation value, as provided by ConSurf⁹⁷. The tunnel has an average width of 1.4 Å and is flanked by highly conserved residues. The width of the tunnel corresponds to the size of a water molecule and would not allow the passage of the larger guanidinium cation. Hence, some molecular plasticity/flexibility is needed to allow minor rearrangements leading to the required widening of the tunnel. (f) Subunit comparison between *Synechocystis* (green) and *N. inopinata* guanidinases (cyan), with

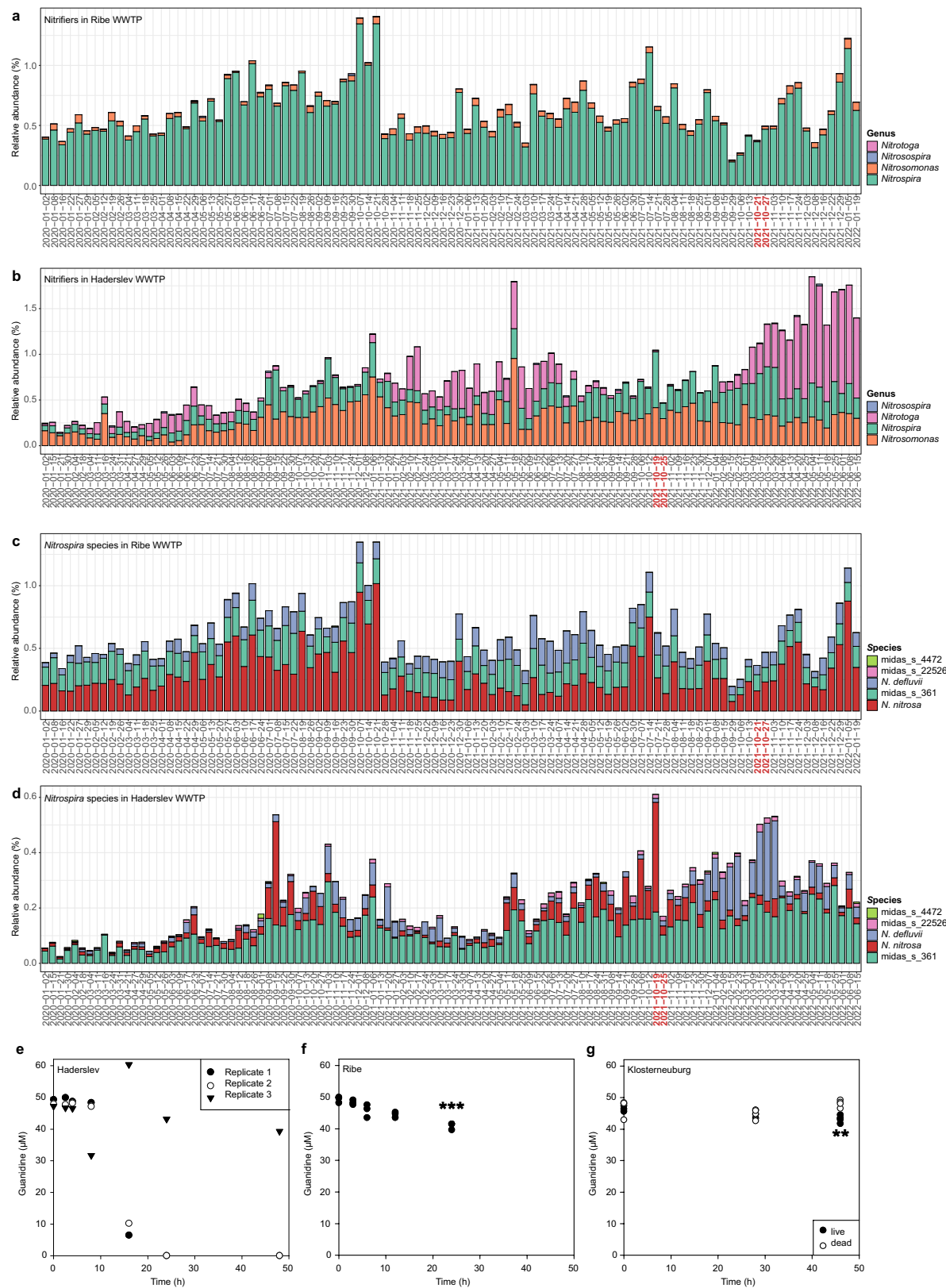
C-terminal extension of *Synechocystis* (residues 369–386 of 7OII) highlighted in orange, which is engaged in a series of interactions with the same subunit on the solvent-exposed side of the hexamer (Supplementary Table 8), and which is absent from the *N. inopinata* guanidinase. Nickel ions are shown as green, manganese ions as purple spheres. (g) Superposition of the active sites of *Synechocystis* (green, PDB ID 7OII) and *N. inopinata* guanidinases (light blue, PDB ID 9FEK). Particularly noteworthy is the tilt of *N. inopinata* guanidinase Trp313 [χ_1 (N, C α , C β , C γ) = 51.3°] towards the active site nickel/manganese ions, compared to the corresponding Trp305 in *Synechocystis* GdmH 7OII [χ_1 (N, C α , C β , C γ) = –80.5°] which is flipped out, creating a cavity occupied by ethylene glycol or a cacodylate ion (PDB ID 7ESR). Nickel ions are shown as green, manganese ions as purple spheres. Cacodylate ion from *Synechocystis* guanidinase (PDB 7OII) is shown as transparent sticks. (h) Active site of *N. inopinata* guanidinase (light blue) with averaged anomalous difference fourier maps contoured at 2 rmsd. The nickel ion is shown as green sphere, the nickel anomalous map is shown as a green mesh. The manganese ion is shown as purple sphere, the manganese anomalous map is shown as a purple mesh. (i) Anomalous scattering signal vs x-ray energy plotted for manganese (purple) and nickel (green) ions using data obtained from http://skuld.bmsc.washington.edu/scatter/AS_form.html, accessed on 20.02.2024. X-ray energies used to collect datasets of anomalous maps from C are indicated by dashed lines.



Extended Data Fig. 7 | Extended kinetics parameters of *N. inopinata*.

(a) and (b) Additional replicates for whole-cell guanidine oxidation rates of *N. inopinata* pre-induced with guanidine for -12 h. A Michaelis-Menten model (red line) was used to determine the $K_{m(\text{app})}$ and V_{max} . (c) The effect of guanidine on the rate of total ammonia oxidation (points are means \pm s.d., $n = 3$). A linear regression was fit (dashed line), and the 50% inhibition rate is marked in red. (d) A Dixon reciprocal plot for the inhibition of total ammonia oxidation by guanidine ($n = 3$; circles, triangles, and squares; the same biological replicates from panel (c) were used). Ammonia oxidation rates determined at guanidine concentrations that formed a linear inhibition trend were used to fit a linear regression for each replicate (dashed lines). Regression intersection points (red dots) were used to determine the inhibition constant value ($-K_i$). (e-h) Michaelis-Menten plots for *N. inopinata* pre-grown in medium containing ammonium and guanidine (-1 mM and -15 μM , respectively for -12 h, (e)), urea

(0.5 mM for >1 week (f and h)), or ammonium (-1 mM for >1 week (g)). Data presented in panel (g) was previously published and is shown for comparison⁴⁸. Total ammonia (e, f, and g) and urea oxidation rates (h) were determined from O_2 consumption rates and normalized to total protein concentration. A Michaelis-Menten model (red line) was used to determine the $K_{m(\text{app})}$ and V_{max} . The $K_{m(\text{app})}$ for unprotonated ammonia was calculated from total ammonium kinetics. In all cases, biological replicates are shown in white, blue, and grey. It should be noted that recently whole-cell urea oxidation kinetics of *N. inopinata* were also reported in another publication and differ from those shown here³³. Qin et al., reports a $K_{m(\text{app})}$ of -0.14 μM urea, a V_{max} of -8 μM urea $\text{mg protein}^{-1} \text{ h}^{-1}$, and an a° of -10,000 $\text{l g cells}^{-1} \text{ h}^{-1}$. In addition, no difference in the a° for NH_3 or urea for *N. inopinata* was reported. While the cause of these discrepancies is unknown, it is possible that this reflects differences related to concentrating cultures or lack thereof before microrespiration measurements.



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Community structure of nitrifying bacteria in the WWTPs Ribe and Haderslev, and Guanidine degradation in WWTPs with and without comammox microbes. (a-d) V1-V3 16 S rRNA gene amplicon sequencing revealed that, over a period of two years, (a) the nitrifying community in WWTP Ribe was dominated by members of the genus *Nitrospira*, while co-occurring AOB related to *Nitrosomonas* and *Nitrosospira* were less abundant. (b) The nitrifying community in WWTP Haderslev was more diverse and contained, in addition to the consistently abundant *Nitrospira*, also relatively large populations of *Nitrosomonas* and nitrite-oxidizing bacteria from the genus *Nitrotoga*, whereas *Nitrosospira*-related AOB were rare. The *Nitrospira* populations in the two WWTPs consisted of organisms closely related with the comammox species *N. nitrosa*, the canonical nitrite oxidizer *N. defluvii*, and yet uncharacterized *Nitrospira*-like bacteria referred to as 'midas_s_361', 'midas_s_4472' (both WWTPs), and 'midas_s_22526' (only WWTP Haderslev) (c and d). Samples taken shortly before and after the sampling of activated sludge for metatranscriptomic analyses from both plants are labelled in red. The partial 16 S rRNA gene sequence midas_s_361 shows >99% identity to a 16 S rRNA gene from a lab-scale reactor *Nitrospira* sp. UBC3 MAG (GCA_022226955), which possesses no genes for ammonia oxidation. The midas_s_4472 16 S rRNA

and midas_s_22526 16 S rRNA gene sequence are related to the canonical nitrite oxidizers *N. marina* and *N. defluvii*, respectively. Thus, most likely these three partial 16 S rRNA gene sequences do not represent comammox organisms. (e-g) 50 μ M guanidine was added to the (e) Haderslev, (f) Ribe (both with comammox), and (g) Klosterneuburg (no comammox) activated sludge samples and guanidine concentrations were determined over time. Dead controls showed no significant changes in guanidine concentrations, suggesting that chemical degradation or sorption of guanidine did not influence the results. The activated sludge biomass in the incubations was normalized by Total Suspended Solid levels and subsequently confirmed by BCA protein content measurements. Relative guanidine degradation rates normalized by protein content: Haderslev active replicates: 12.9 μ M per h; Ribe: 0.95 μ M per h; Klosterneuburg living: 0.19 μ M per h. All incubations were done in triplicates. Two-tailed *t*-test of significance: panel (b): *** between 0 h and 24 h timepoints: $t = 10.418$, d.f. = 4, $P = 0.000480$; panel (c): ** Living biomass between 0 h and 48 h timepoints: $t = 4.602$, d.f. = 6, $P = 0.00369$; Dead biomass between 0 h and 48 h timepoints: no significant difference. We have no good hypothesis, why one replicate of the Haderslev sample showed different results (e).

Extended Data Table 1 | Comparison of kinetic properties and metal loading of purified guanidinase enzymes from *N. inopinata* and *Synechocystis*

Source of guanidinase enzyme (kinetics pH)	K _m (mM)	V _{max} (nmol s ⁻¹ mg ⁻¹)	K _{cat} (s ⁻¹)	K _{cat} /K _M (s ⁻¹ mM ⁻¹)	Ni loading (atom / subunit)	Mn loading (atom / subunit)	Reference
<i>N. inopinata</i> ; 0.5 μM Ni in expr. (pH=7.5)	11.2 ±0.51	0.56 ±0.42	0.013	0.0012	0.26 ±0.004	0.18 ±0.006	This study
<i>N. inopinata</i> ; 20 μM Ni in expr. (pH=7.5)	16.1 ±0.40	2.02 ±0.02	0.083	0.0052	0.35 ±0.016	0.55 ±0.025	This study
<i>N. inopinata</i> ; 1 mM Ni in expr. (pH=7.5)	13.6 ±0.76	34.73 ±0.62	1.42	0.1044	0.42 ±0.006	0.03 ±0.001	This study
<i>Synechocystis</i> PCC6803 (pH=7.5)	5.3 ±1.2	0.91 ±0.23	0.040	0.0075			Wang <i>et al.</i> , 2021 ¹⁷
<i>Synechocystis</i> PCC6803 (pH=8)	7.8 ±1.8	84.8 ± 6.79	3.830	0.4337	1.9	1	Funck <i>et al.</i> , 2022 ¹⁸

The *Synechocystis* enzyme was either overexpressed in *Synechocystis*^{15,17218} or in *E. coli*¹⁸. In the latter heterologous expression experiment, Ni-loading chaperons of *Synechocystis* were co-overexpressed. For the *N. inopinata* guanidinase, three different nickel concentrations (indicated after the source species name) were used during expression. For the 20 μM nickel expression batch, nickel was present during all purification steps in the same concentration, and only removed for the ICP-MS analysis of metal loading. This was not possible for the 1 mM Ni²⁺ treatment as it resulted in protein denaturation during purification.

Extended Data Table 2 | Guanidine concentrations in animal and WWTP samples

Sample	Biological Replicates	Concentration ($\mu\text{mol kg}^{-1}$ (dry weight) or $\mu\text{mol l}^{-1}$)
Samples measured in this study		
Cow Urine*	5	13.6 ± 4.5
Sheep Urine	2	6.50 - 6.62
Cow Feces*	3	150 ± 17
Sheep Feces	4	69 ± 22
Chicken Feces	4	108 ± 46
Pig Feces	4	143 ± 40
WWTP Klosterneuburg Influent	1	0.5
Literature data		
Human Urine ^{36,37}	22	2.17 - 19.8
Rosettes of <i>Arabidopsis</i> ¹⁴	4	100-200
Legume Seeds ¹⁴	3	2400

Concentrations are reported as mean \pm s.d. or as a range of biological replicates in $\mu\text{mol per kg}$ (dry weight) for faecal and plant samples or $\mu\text{mol per l}$ for urine and wastewater samples^{14,35,36}. Based on the guanidine concentrations measured in cow faeces and urine samples (*), we roughly extrapolated to the amount of guanidine that is globally excreted by cows. Using the measured water content of our cow faeces samples (mean: 57% (w/w)) and the molar mass of guanidine, we calculated a guanidine concentration per faeces wet mass of 3.8 mg/kg. With an estimated global cow population of 940 million and estimated amounts per cow and day for feces (30 kg) and urine (13 L), we obtained > 100 tons of guanidine globally excreted by cows per day. Cow, sheep, pig and chicken samples were collected at the University of Veterinary Medicine Vienna.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	hmmsearch, infernal v1.1.3, XDS, SensorTrace Rate, bbmap v38.92, C1000-CFX96 BioRad, NovaSeq 6000
Data analysis	SigmaPlot v14.5, MS Excel 2016, IQ-TREE2, Proteome Discoverer v.2.5, Astra, MoltenProt, MaxEnt1, XDSCONV, Phaser-MR, AlphaFold, COOT, PHENIX, MolProbity, PDBREDO, PyMol v 2.0, DESeq2, R (packages: DEqMS, phangorn, ape, ggtree, ggseqlogo, Decipher.phytools), usearch, FastTree2, OpenMIMS v 3.0.5, Fiji ImageJ 1.54f, bbmap v38.92, bbdut, bedtools

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

- All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
 - A description of any restrictions on data availability
 - For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The crystal structures have been deposited in the Protein Data Bank under Code: 9FEK. WWTP metatranscriptome reads have been deposited under BioProject ID PRJNA1118285. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE 94 partner repository with the

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Reporting on race, ethnicity, or other socially relevant groupings

Population characteristics

Recruitment

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

Research sample

Sampling strategy

Data collection

Timing and spatial scale

Data exclusions

Reproducibility

Randomization

Blinding

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions

Field conditions	15°16.5'E under mild weather conditions.
Location	Ribe, Denmark (GPS: 55.33, 8.74), Haderslev, Denmark (GPS: 55.25, 9.51), Klosterneuburg, Austria (GPS: 48.29, 16.34), Agricultural soil 48°11'17.9"N 14°15'16.5"E
Access & import/export	Sampling of the wastewater treatment plants happened under the supervision of the respective facility management personnel. Living cultures did not cross any borders, they were incubated and sub-sampled in local laboratories. Austrian agricultural soil was sampled and all experiments were performed in Austria, which does not require specific permits.
Disturbance	Sampling the tanks of the wastewater treatment plants or the agricultural soils (longterm field experiment from AGES Austrian Agency for Health and Food Safety) did not cause disturbance to their microbial communities.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Plants

Seed stocks	not applicable
Novel plant genotypes	not applicable
Authentication	not applicable