Title

Structure-guided function discovery of an NRPS-like glycine betaine reductase for choline biosynthesis in fungi

Journal

Proceedings of the National Academy of Sciences of the United States of America, 116(21)

ISSN

0027-8424

Authors

Hai, Yang
Huang, Arthur M
Tang, Yi

Publication Date

2019-05-21

DOI

10.1073/pnas.1903282116

Peer reviewed

Structure-guided function discovery of an NRPS-like glycine betaine reductase for choline biosynthesis in fungi

Yang Hai (海洋)[a], Arthur M. Huang[b], and Yi Tang[a,b,1]

[a]Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095; and [b]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095

Edited by Wilfred A. van der Donk, Howard Hughes Medical Institute and University of Illinois, Urbana–Champaign, Urbana, IL, and accepted by Editorial Board Member Stephen J. Benkovic April 10, 2019 (received for review February 27, 2019)

Nonribosomal peptide synthetases (NRPSs) and NRPS-like enzymes have diverse functions in primary and secondary metabolisms. By using a structure-guided approach, we uncovered the function of a NRPS-like enzyme with unusual domain architecture, catalyzing two sequential two-electron reductions of glycine betaine to choline. Structural analysis based on the homology model suggests cation-π interactions as the major substrate specificity determinant, which was verified using substrate analogs and inhibitors. Bioinformatic analysis indicates this NRPS-like glycine betaine reductase is highly conserved and widespread in kingdom fungi. Genetic knockout experiments confirmed its role in choline biosynthesis and maintaining glycine betaine homeostasis in fungi. Our findings demonstrate that the oxidative choline-glycine betaine degradation pathway can operate in a fully reversible fashion and provide insight in understanding fungal choline metabolism. The use of an NRPS-like enzyme for reductive choline formation is energetically efficient compared with known pathways. Our discovery also underscores the capabilities of the structure-guided approach in assigning functions of uncharacterized multidomain proteins, which can potentially aid functional discovery of new enzymes by genome mining.

glycine betaine reductase | NRPS-like | choline metabolism | biosynthesis | fungi

The nonribosomal peptide synthetase (NRPS)-like carboxylic acid reductases (CARs) catalyze ATP- and NADPH-dependent reduction of carboxylic acids to the corresponding aldehydes (1). These multidomain enzymes consist of a N-terminal adenylation (A) domain, which selects and activates a carboxylic acid substrate by adenylation, then transfers the acyl moiety to the phosphopantetheinyl arm from the thiolation (T) domain through a thioester linkage. The T domain delivers the acyl thioester to the C-terminal thioester reductase (R) domain where the NADPH-dependent thioester reduction occurs (Fig. 1*A*). Unlike some off-loading R domains in polyketide synthases and NRPS assembly lines, which catalyze nonprocessive [2 + 2] electron reduction to over-reduce the aldehyde product to alcohol (2, 3), the R domains in CARs catalyze a strict two-electron reduction, and the structural basis for such specificity was reported recently (4).

The members of the CAR family are functionally diverse and play indispensable roles in both primary and secondary metabolisms, such as the α-aminoadipate reductase Lys2, required for lysine biosynthesis in fungi (5); and tyrosine reductase HqlA, involved in biosynthesis of herquline A (Fig. 1*A*) (6). CARs also show substantial promise as green biocatalysts for the conversion of aromatic or aliphatic carboxylic acids into the corresponding aldehydes. Indeed, synthetic applications of CARs have been reported for the biocatalytic synthesis of vanillin (7), preparation of methyl branched aliphatic aldehydes (8), production of biofuels (9, 10), and other chemical commodities (11).

Prominent metabolic roles of CARs and their biocatalytic potential encouraged us to discover uncharacterized members with different functions by genome mining. During this effort, we focused on a fungal CAR-like protein with an unknown function. It is distinguished from all other CARs due to an extra C-terminal YdfG-like short-chain dehydrogenase/reductase domain. We named this protein ATRR and the corresponding gene *atrr* after its unusual domain architecture (A-T-R_1-R_2). The presence of two fused R domains in ATRR implies that it could catalyze two consecutive two-electron reductions of a carboxylic acid to yield an alcohol (Fig. 1*B*). Moreover, a genomic survey reveals that *atrr* genes are widespread but exclusive to eukaryotes, mostly in fungi and found in several protist and invertebrate species. In particular, ATRR orthologs are highly conserved in kingdom fungi: Sequence identities are greater than 60% from different species, which indicates a unified and conserved role of ATRR. Transcriptomics studies also suggest *atrr* genes are constitutively transcribed in different fungal species, including human opportunistic pathogen *Aspergillus fumigatus* (12, 13); and rice blast fungus *Magnaporthe oryzae* (14). These features motivated us to discover and characterize its enzymatic activity and biological function.

Results

Bioinformatic Analysis Provides Initial Insight. We first applied the "genomic enzymology" strategy to obtain clues regarding the ATRR function from the *atrr* genomic context since enzymes in a microbial metabolic pathway often are encoded by a gene cluster (15, 16). However, comparative genomic analysis of *atrr* orthologs in fungi reveals that there are no well-defined and conserved

Significance

Oxidation of choline to glycine betaine is a major pathway for choline degradation and glycine betaine biosynthesis. Although this oxidative process is found in all domains of life, enzymes involved in the reduction of glycine betaine to choline have not been found. Here, we found this missing link in choline metabolism through a structure-guided function discovery of a NRPS-like glycine betaine reductase, which catalyzes reduction of glycine betaine to choline via the intermediate glycine betaine aldehyde. Bioinformatic analysis suggests this enzyme is highly conserved and widespread in the fungal kingdom, and our discovery of this reductive choline biosynthesis pathway provides insight into fungal choline metabolism.

Author contributions: Y.H. and Y.T. designed research; Y.H. and A.M.H. performed research; Y.H. analyzed data; and Y.H. and Y.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. W.A.v.d.D. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: yitang@ucla.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1903282116/-/DCSupplemental.

Published online May 6, 2019.

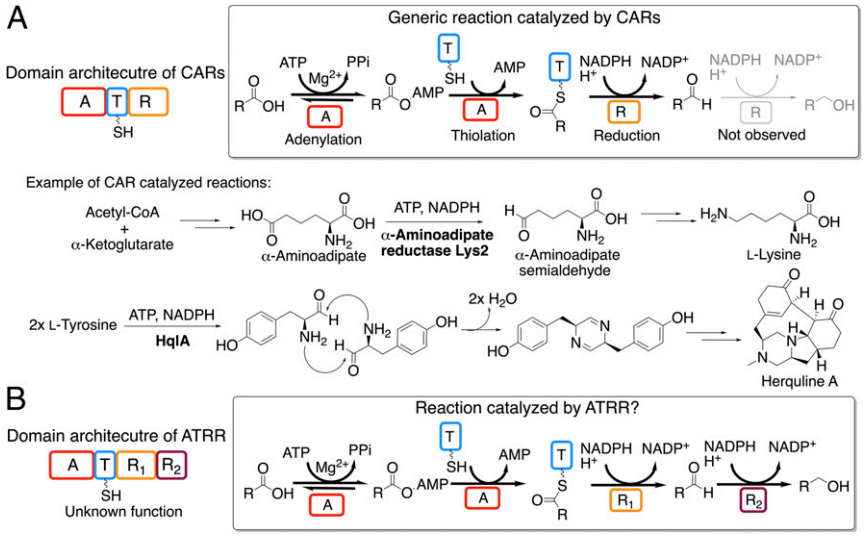

Fig. 1. Comparison between typical CARs and ATRR. (*A*) Domain architecture of typical CARs and CAR-catalyzed reactions. (*B*) Domain architecture of ATRR and the postulated reaction based on domain function.

gene neighborhoods. Even in some closely related species of the same genus (e.g., *Aspergillus*), the genome environment of *atrr* differs substantially (*SI Appendix*, Figs. S2 and S3). The diversified genomic context and lack of obvious co-occurrence with signature metabolic pathway genes prevented us from reliable functional assignment.

We then focused on characterization of the A domain of ATRR (ATRR-A), which selects a substrate and is, therefore, the entry point to ATRR catalysis. Because the substrate specificity of A domains is dictated by the amino acid residues lining the active site pocket (10 AA code), deciphering this "nonribosomal code" of ATRR-A would, in theory, allow us to predict its function (17, 18). Bioinformatic analysis showed the ATRR-A has a unique 10 AA code, and it is conserved among ATRR orthologs from different fungal species (*SI Appendix*, Fig. S4). However, substrate prediction by online servers was unsuccessful due to the lack of characterized homologs with significant similarity (19). Closer inspection reveals that the highly conserved aspartate residue (D235 in PheA), which typically interacts with the substrate α-amino group via a salt bridge (20, 21), is not conserved in ATRR-A (*SI Appendix*, Fig. S5). Loss of this key aspartate residue suggests that the substrate is not a typical α-amino acid or an unusual amino acid substrate accommodated in an unprecedented fashion as in the case of ClbH-A2 which activates *S*-adenosyl-L-methionine (SAM) (22). Indeed, most characterized A domains without D235 equivalent residues (Asp/Glu) are known to activate nonamino acid-type substrates, including aryl acid, keto acid, and hydroxy acid (*SI Appendix*, Fig. S5) (23–26). Based on this prediction, we screened the CAR activity of ATRR in vitro against a library of common carboxylic acids from primary metabolism as well as the 20 proteinogenic amino acids (*SI Appendix*, Fig. S6). The *Aspergillus nidulans* ATRR protein (Uniprot AN5318.2) was expressed from *Escherichia coli* BL21(DE3) and purified to homogeneity by affinity chromatography and size-exclusion chromatography (*SI Appendix*, Fig. S7). *Apo*-ATRR was enzymatically converted to *holo*-form by using phosphopantetheinyl transferase NpgA and CoA (*SI Appendix*, Fig. S8) (27). The reductase activity of *holo*-ATRR was examined by monitoring the oxidation of NADPH at 340 nm. ATRR showed no activity toward any of the substrates in this library.

Structure-Based Approach Predicts a Domain Substrate. We then used a structure-based approach to gain more insight about the substrate specificity of ATRR-A, a strategy that has led to many successful predictions for enzymes of unknown function (28, 29). We constructed an active site homology model of ATRR-A by using a TycA variant (PDB entry 5N82) (21) as the template (34% sequence identity to ATRR-A) and manually replacing the 10 AA code residues with the corresponding ones in *A. nidulans* ATRR-A (Fig. 2*A*). In comparison with the structure of the archetypical A domain PheA (17), three aromatic residues (F201, W204, and W270) stood out in the ATRR-A model. These residues could potentially form an aromatic cage to bind a quaternary ammonium group through cation-π interactions (30, 31). This quaternary ammonium group can be further electrostatically stabilized by a nearby anionic residue E304. Such an aromatic box is reminiscent of those observed in quaternary ammonium moiety binding proteins, such as periplasmic betaine binding protein Prox (32); *cis*-4-hydroxy-D-proline betaine epimerase HpbD, and demethylase HpbJ (28); histone trimethyllysine reader proteins (33); and acetylcholine esterase (30, 31).

We, therefore, speculated that the carboxylic acid substrate of ATRR could be a small betaine, such as the physiologically abundant carnitine, glycine betaine, etc. To test this hypothesis, we assayed ATRR activity with a panel of naturally occurring betaines (Fig. 2*B*). Gratifyingly, rapid oxidation of NADPH was observed only in the presence of the smallest betaine, glycine betaine. Consumption of NADPH indicates that glycine betaine was reduced. In accordance with our early hypothesis about ATRR function (Fig. 1*B*), the predicted corresponding two sequential two-electron reduction product choline was indeed accumulated in the reaction mixture (*SI Appendix*, Fig. S9). We then determined the apparent steady-state kinetic constants for ATRR glycine betaine reductase activity (*SI Appendix*, Fig. S10 and Table S1). The overall k_{cat} value (~1 s^{-1}) is modest but similar to that of previously characterized CARs (4), whereas the K_m value (~2 mM) matches the intracellular level of glycine betaine (~3 mM) in fungi (34). Therefore, glycine betaine is a physiologically relevant substrate for ATRR.

Mechanistic Study of ATRR. We next studied the catalytic mechanism of the glycine betaine reductase function of ATRR. By using a hydroxylamine-trapping assay designed for testing adenylation activity (35), we confirmed that glycine betaine was indeed adenylated by ATRR-A (Fig. 3*A*). Furthermore, similar to all CARs, the glycine betaine reductase activity of ATRR

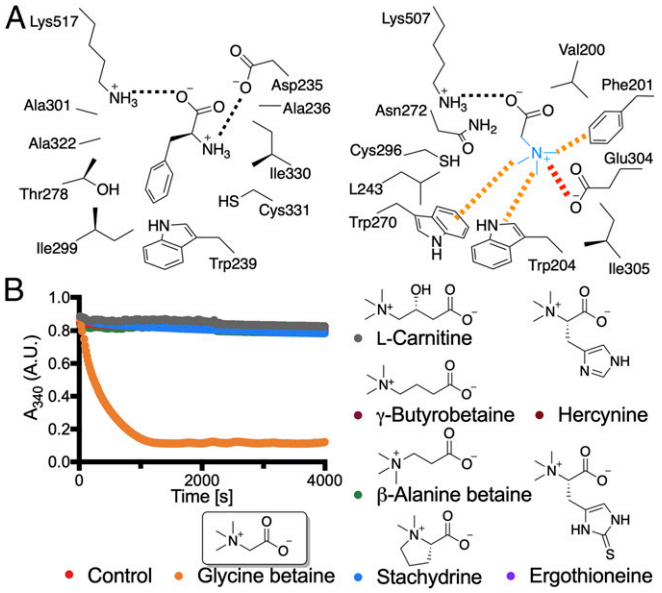

Fig. 2. Structure-based prediction of the ATRR substrate. (*A*) Comparison of the homology model of *A. nidulans* ATRR-A domain with the structure of PheA in complex with its cognate substrate L-phenylalanine. The conserved salt bridges recognizing the α-carboxylate and α-amino groups are shown as black dashes. The proposed cation-π and electrostatic interactions favoring binding of betaine-like substrates in the ATRR A domain are shown as orange and red dashes, respectively. The simplest betaine, glycine betaine, was shown to indicate the plausible binding mode. (*B*) CAR activity of ATRR against a panel of betaine substrates.

requires phosphopantetheinylation of the T domain, suggesting the corresponding glycine betainoyl thioester intermediate is critical for reduction (*SI Appendix*, Fig. S11). We verified the formation of this intermediate using matrix-assisted laser desorption/ionization–time-of-flight mass spectrometry with the stand-alone ATRR-T domain, which can be acylated by ATRR-A in the presence of glycine betaine and ATP (*SI Appendix*, Fig. S8). The expected on-pathway intermediate glycine betaine aldehyde (hydrate) was also observed during the reaction course by continuous trapping with phenylhydrazine (*SI Appendix*, Fig. S12). Direct reduction of glycine betaine aldehyde to choline by ATRR was also observed, confirming its role as an intermediate in the stepwise reduction (*SI Appendix*, Fig. S9). Kinetic assay showed that glycine betaine aldehyde reduction by ATRR is more efficient: The catalytic efficiency ($k_{cat}/K_m = 3.0 \times 10^4\ M^{-1}\ s^{-1}$) is 66-fold greater than that of the overall carboxylic acid reduction, whereas the turnover number ($k_{cat} = 7\ s^{-1}$) is sixfold higher. Rapid reduction of glycine betaine aldehyde shows that the aldehyde reduction step is not rate limiting thereby avoiding accumulation of the reactive aldehyde species.

We next characterized the individual function of R_1 and R_2 domains. To isolate the catalytic activity of the individual domains, we inactivated each reductase domain by substituting the conserved catalytic tyrosine residue with phenylalanine (Y811F in R_1 and Y1178F in R_2). In agreement with our two consecutive two-electron reductions hypothesis (Fig. 1*B*), mutation of R_1 essentially abolished overall carboxylic acid reductase activity, whereas aldehyde reductase activity remained intact (*SI Appendix*, Table S1). On the contrary, mutation of R_2 did not substantially affect carboxylic acid reductase activity but resulted in the accumulation of glycine betaine aldehyde intermediate (*SI Appendix*, Fig. S12). The Y1178F mutation in R_2 also led to a greater than 150-fold loss of aldehyde reductase activity (catalytic efficiency). (*SI Appendix*, Table S1). To further confirm that the remaining aldehyde reductase activity was not contributed from R_1, we introduced another mutation at R_2 (G1036A) to compromise binding of the cosubstrate NADPH. As expected, double mutation (G1036A/Y1178F) in R_2 further decreased the aldehyde reductase activity by 4,000-fold, whereas the carboxylic acid reductase activity was not compromised. The role of R_2 as a dedicated aldehyde reductase was also supported by the robust reduction of betaine aldehyde to choline using the excised stand-alone R_2 domain (*SI Appendix*, Table S1). To summarize, the thioester reductase domain R_1 strictly carries out a single two-electron reduction of glycine betainoyl thioester to glycine betaine aldehyde, which is then reduced to choline at the aldehyde reductase domain R_2.

Substrate Specificity and Inhibition of ATRR. Establishing the function of ATRR as glycine betaine reductase enabled us to further characterize its substrate specificity. We first studied ATRR specificity toward glycine betaine by assaying both A and carboxylic acid reductase activity of ATRR against different glycine betaine analogs (Fig. 3 *A* and *B*). The isosteric neutral analog 3,3-dimethylbutyrate completely abolished the activity. In contrast, the tertiary sulfonium analog sulfobetaine, which also bears a positive charge on its head group but lacks one methyl group

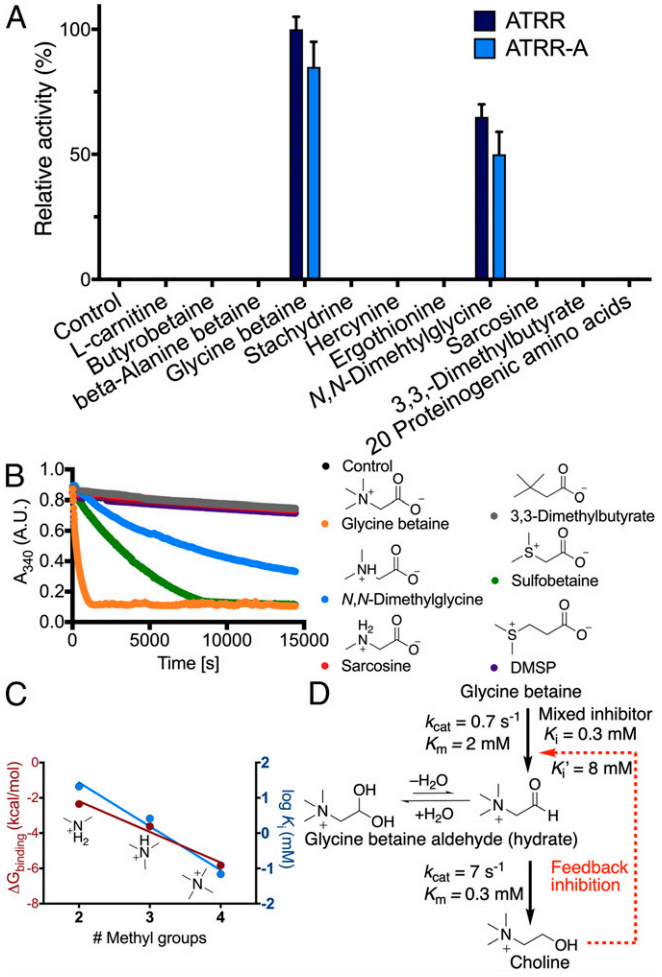

Fig. 3. ATRR substrate specificity and inhibition studies. (*A*) Substrate specificity of ATRR-A profiled with different substrates. The A activity was monitored here with the hydroxylamine capture assay. (*B*) CAR activity of ATRR against a panel of glycine betaine analogs. (*C*) Dependence of binding free energy change and inhibitor disassociation constant K_i on the number of methyl groups of ammonium inhibitors against glycine betaine reductase activity. Note that, for the mixed inhibitor trimethylammonium ion, K_i' is not taken into account. (*D*) Feedback inhibition from choline could regulate ATRR enzyme activity.

compared with glycine betaine, was partially active (an order of magnitude slower). Such comparison clearly demonstrates that the formal positive charge carried by the substrate is critical for activity and, thus, supports our predicted cation-π interactions conferred by ATRR-A during substrate recognition. The decreased activity of ATRR toward sulfobetaine suggests the importance of the trimethyl moiety for function. The important contribution of the trimethyl moiety of glycine betaine to ATRR catalysis was also revealed by carboxylic acid reduction assays with dimethylglycine and monomethylglycine (sarcosine), which are physiologically relevant analogs as they are intermediates in the glycine betaine degradation pathway (36). Progressively decreasing the number of *N*-methyl groups of glycine betaine resulted in a significant loss of carboxylic acid reductase activity: the reduction of dimethylglycine was two orders of magnitude slower than that of glycine betaine, whereas no reduction could be observed with sarcosine (Fig. 3 *A* and *B* and *SI Appendix*, Table S1). Since the ammonium groups of both dimethylglycine (pK_a = 9.89) and sarcosine (pK_a = 10.1) were thought to be protonated under our experimental condition (pH 7.5–8.0) (37), the decrease in activity per methyl group reflects the additive contribution of each methyl substitution to catalysis. ATRR-A was also shown to be strictly size selective in that β-alanine betaine and dimethylsulfoniopropionate, each containing one additional methylene group to glycine betaine and sulfobetaine, respectively, were not substrates of ATRR (Fig. 3 *A* and *B*). Collectively, these results show that ATRR is highly specific to glycine betaine and such specificity is likely achieved by favoring binding of its head group and distinguishing different extents of methylation.

To corroborate our findings on substrate specificity, we studied the inhibition of ATRR by using different ammonium-containing inhibitors (*SI Appendix*, Fig. S13 and Table S2). The tetramethylammonium ion, which mimics the head group of glycine betaine, was found to be a potent competitive inhibitor of ATRR-A (K_i = 69 μM), whereas the potency decreased by three orders of magnitude with dimethylammonium (K_i = 21 mM). As illustrated in Fig. 3*C*, a monotonic increase in inhibitor affinity is correlated with the addition of each methyl group yielding an incremental free energy relationship: Addition of each methyl group contributes ∼1.7 kcal/mol toward inhibitor binding. Therefore, our results reinforced the idea that the head group plays a major role in ATRR-A substrate recognition, and ATRR-A favors binding of tetra-alkylated quaternary ammonium moiety, presumably through a combination of cation-π interactions, hydrophobic effect, and maximized van der Waals interactions. Unlike ATRR-A, R_2 is not potently inhibited by the head-group mimic tetramethylammonium ion (mixed inhibitor, K_i = 5, K_i' = 18 nM), suggesting the head group is not the dominant determinant for R_2 substrate specificity. Nonetheless, the positive charge on the substrate is still essential for R_2 catalysis as the neutral isosteric analog 3,3-dimethylbutyraldehyde (DMBA) showed a 13,500-fold loss of catalytic efficiency (*SI Appendix*, Table S1). It is noteworthy to mention that DMBA exists mainly as the aldehyde form (≥65%) in aqueous solution favoring reductase-catalyzed NADPH-dependent reduction (38), whereas glycine betaine aldehyde predominantly exists as the hydrated *gem*-diol form (99%), which must then undergo dehydration to unveil the aldehyde functional group for reduction.

Since the product choline consists of the same head group, we next tested whether choline could act as an inhibitor competing with glycine betaine and glycine betaine aldehyde. As a result, choline was shown to be a mixed inhibitor for both glycine betaine reductase and aldehyde reductase activity but more potent (K_i = 0.27, K_i' = 6.7 mM) in competition against glycine betaine in the first reduction step. Therefore, choline could act as a feedback inhibitor to regulate ATRR enzymatic activity (Fig. 3*D*). The lowered binding affinity of choline to R_2 (K_i = 4 mM) favors the release of choline after glycine betaine aldehyde reduction to avoid direct product inhibition.

Physiological Role of ATRR in Fungi. Our discovery of ATRR as a glycine betaine reductase strongly indicates that it is the missing link in choline-glycine betaine metabolism (Fig. 4). Oxidation of choline to glycine betaine initiates the choline catabolic process. This common choline degradation pathway has been found in all domains of life, but the reverse process has only been implicated in fungi. A collection of experimental evidence, including genetics, metabolite quantification by NMR, and radioactive labeling studies, suggest glycine betaine can be reduced to choline in vivo and thereby can be an alternative source of choline for fungi (39, 40). However, the underlying biochemical mechanism

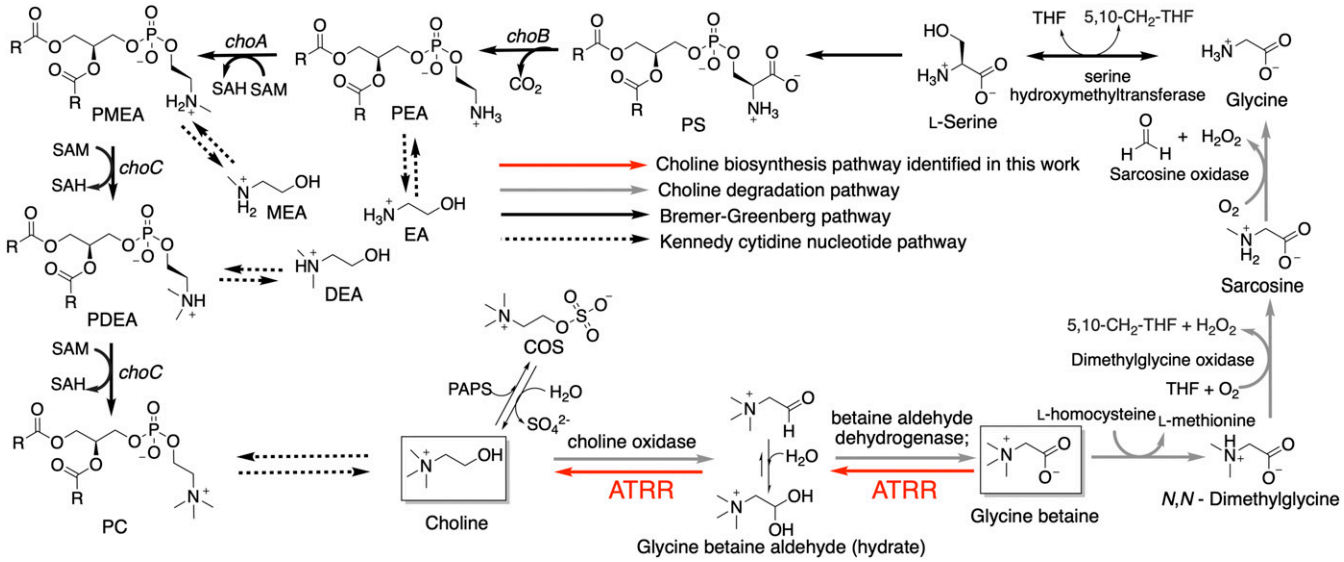

Fig. 4. Choline-glycine betaine metabolism in fungi. DEA, dimethylethanolamine; EA, ethanolamine; MEA, monomethylethanolamine; PAPS, 3′-phosphoadenosine-5′-phosphosulfate; PDEA, phosphatidyldimethylethanolamine; PEA, phosphatidylethanolamine; PMEA, phosphatidylmonomethylethanolamine; PS, phosphatidylserine; R, fatty acyl chain; THF, tetrahydrofolate.

and genetic basis have remained elusive. A plausible mechanism for this conversion was postulated based on the reduction of L-aspartate to L-homoserine in which the carboxy group is first activated via phosphorylation by a kinase. This is followed by conversion to a thioester that is reduced into aldehyde by dehydrogenase and then reduced to alcohol by a separate reductase. The biosynthetic logic of ATRR discovered in this paper also uses a carboxylic activation strategy for reduction but is more concise as all steps are performed by one enzyme, and the key thioester intermediate is channeled through the phosphopantetheine arm of a T domain.

To test whether ATRR plays a role in choline biosynthesis in fungi, we deleted the *atrr* gene in both the wild-type (*wt**) *A. nidulans* and the *A. nidulans* Δ*choA* strains by using the split-marker method (*SI Appendix*, Figs. S14 and S15) (41). Disruption of the *choA* gene encoding the phosphatidylethanolamine *N*-methyltransferase blocks the Bremer-Greenberg methylation pathway from PEA and renders the Kennedy pathway from free choline as the only option for biosynthesis of the essential phospholipid phosphatidylcholine (PC) (Fig. 4) (40). Deletion of the *atrr* gene would abolish conversion of glycine betaine to choline, but both the Bremer-Greenberg pathway and the Kennedy pathway remain functional. Accordingly, we expect no growth defects of the Δ*atrr* strain, whereas the Δ*choA* strain would show choline-/betaine-dependent growth. Without the ability to convert glycine betaine to choline, the Δ*atrr*Δ*choA* double knockout strain is expected to be truly choline-auxotrophic even in the presence of glycine betaine. Indeed, when cultured on solid minimal medium (MM), no major growth and morphological difference was observed between the Δ*atrr* and the *wt** strains. On the contrary, neither the Δ*choA* nor the Δ*atrr*Δ*choA* strain was able to grow because of obstructed PC biosynthesis (Fig. 5*A*). The vegetative growth of the Δ*choA* strain can be stimulated by supplementation of either choline or glycine betaine (2–20 μM), although full restoration of the conidiation requires higher concentrations (>100 μM). In contrast, the Δ*atrr*Δ*choA* strain exhibits choline-auxotroph character: Glycine betaine can no longer replace choline to restore growth and asexual development under the choline-limiting condition. These results confirm that ATRR rescues the phenotype of Δ*choA* by converting glycine betaine to choline. Therefore, ATRR mediated reduction of glycine betaine is a "shortcut" pathway to choline.

Unexpectedly, although no significant difference was observed between the Δ*atrr* and the *wt** strains growing at low concentration (2–20 μM) of exogenous glycine betaine, the hyphae growth and conidiation of the Δ*atrr* strain were severely impaired at high level glycine betaine (>100 μM, Fig. 5*A*), which indicates ATRR function is essential to *A. nidulans* under such a condition. To interrogate the role of glycine betaine in the absence of ATRR, we cosupplemented choline to the medium. With an increasing amount of choline (>20 μM) added, the growth of the Δ*atrr* strain can be fully reverted to the normal state even in the presence of a millimolar level of glycine betaine (Fig. 5*B*). Similar but not identical results were also observed when PC and L-methionine (10 mM) were cosupplemented. These chemical complementation results suggest the impaired growth caused by the high level of exogenous glycine betaine is presumably due to the impeded Bremer-Greenberg methylation pathway, which again leads to deficient PC biosynthesis. High levels of methionine could increase SAM biosynthesis via methionine catabolism, which may further up-regulate and drive the methylation pathway against glycine betaine. In comparison, cosupplementation of L-homocysteine (as racemic L- and D-mixtures) to favor glycine betaine degradation did not alleviate the impaired growth of the Δ*atrr* strain, which suggests the glycine betaine degradation pathway is likely not playing a major role in this phenotype. Nevertheless, inhibition of the Bremer-Greenberg methylation pathway by glycine betaine necessitates the function

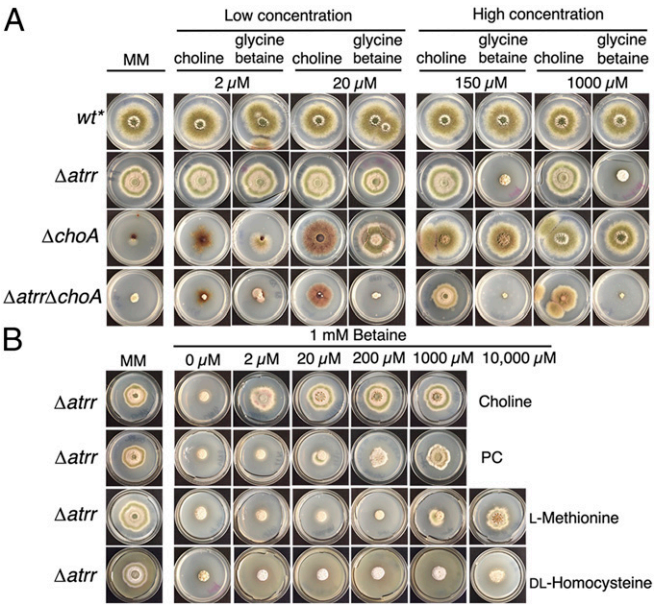

Fig. 5. Disruption of the *A. nidulans atrr* gene impairs conversion of glycine betaine to choline. (*A*) Growth of *A. nidulans* gene deletion mutants on solid MM at 37 °C for 3 d. The medium was supplemented with different concentrations of glycine betaine and choline. (*B*) Chemical complementation restores growth of the *A. nidulans* Δ*atrr* strain at a high concentration of glycine betaine (1 mM).

of ATRR in fungi, which not only can supply choline for PC biosynthesis via the Kennedy pathway, but also can help to "detoxify" high levels of glycine betaine. Therefore, our results confirmed an important physiological role of ATRR in choline and PC metabolism in *A. nidulans* and reveal a previously unknown effect of glycine betaine to *A. nidulans* growth.

Discussion

Our study here was driven by a highly conserved NRPS-like enzyme in fungi with an unusual domain architecture. Successful elucidation of its unknown function as glycine betaine reductase serendipitously led to the resolution of a long-standing question in choline-glycine betaine metabolism, how glycine betaine was converted back to choline. Now, our study provides the molecular basis: ATRR catalyzes ATP- and NAPDH-dependent reductions of glycine betaine by a catalytic mechanism similar to that of well-known NRPS-dependent CARs, generating a glycine betaine aldehyde intermediate, which is further reduced at the R_2 domain unique to ATRR, yielding, as a final product, choline. Orthologs of ATRR are found to be prevalent in fungi and exhibit high sequence identities with characterized *A. nidulans* ATRR (*SI Appendix*, Figs. S2 and S3), which suggests that this efficient single-enzyme catalyzed reductive choline biosynthesis may be widespread in fungi. Harboring this enzyme permits direct reutilization of endogenously stored glycine betaine (~3 mM) for on-demand biosynthesis of choline and choline derivatives, including phospholipid PC (essential role in maintaining membrane integrity and functionality) and choline-*O*-sulfate (a mean for intracellular sulfate storage) (42). Compared with the high metabolic cost of the de novo choline biosynthetic pathway (the Bremer-Greenberg pathway) at the expense of three consecutive methylation steps (costing 12 ATP molecules per methylation event) (43), regeneration of choline from glycine betaine is much more cost-effective (costing 1 ATP and 2 NADPH). More importantly, this shortcut pathway also acts as an "emergency safeguard" pathway to supply choline from its

"reserve form" glycine betaine for PC biosynthesis when the Bremer-Greenberg methylation pathway is obstructed.

Uncovering the ATRR function was rooted in understanding the biochemical logic of NRPS-like CARs. We reasoned that the A domain must play a gatekeeping role in selecting and activating the carboxylic acid substrate; and identification of the substrate would thereby reveal the function of ATRR. Although our initial search for the substrate, guided by the 10 AA code, was unsuccessful, a key insight of the A domain substrate specificity was provided by our interpretation of the structure-function relationship based on a simple active site homology model, which prioritized a pool of potential substrates, and the physiologically relevant substrate glycine betaine was identified eventually by experimental enzymatic activity screening. Our results expanded the biochemistry scope of the NRPS family: ATRR-A is shown to activate and thioesterify a betaine-type substrate; and convincing evidence supports our hypothesis that cation-π interactions play an important role in substrate recognition at the A domain. Our findings support the notion that cation-π interaction is a fundamental and universal noncovalent interaction for substrate/ligand recognition in biology (30, 31).

To summarize, our study demonstrated that the structure-guided approach can be a general strategy for functional discovery of uncharacterized NRPS-like enzymes, and orphan NRPS assembly lines from the ever-increasing size of genomic data, which may aid in the discovery of new metabolic pathways and natural products (16). In principle, such a process can be greatly facilitated by docking a library of carboxylic acid metabolites to high quality homology models of A domains.

Materials and Methods

Materials and methods are summarized in *SI Appendix*, including *SI Appendix*, Figs. S1–S16 and Tables S1–S3.

ACKNOWLEDGMENTS. We thank Dr. Mengbin Chen, Dr. Matthew Jenner for helpful discussions, and Abigail Yohannes for helping with bioinformatic analysis. This work was supported by the NIH (Grant 1R35GM118056) (to Y.T.). Y.H. is a Life Sciences Research Foundation fellow sponsored by the Mark Foundation for Cancer Research.

1. Winkler M (2018) Carboxylic acid reductase enzymes (CARs). *Curr Opin Chem Biol* 43: 23–29.
2. Read JA, Walsh CT (2007) The lyngbyatoxin biosynthetic assembly line: Chain release by four-electron reduction of a dipeptidyl thioester to the corresponding alcohol. *J Am Chem Soc* 129:15762–15763.
3. Chhabra A, et al. (2012) Nonprocessive [2 + 2]e- off-loading reductase domains from mycobacterial nonribosomal peptide synthetases. *Proc Natl Acad Sci USA* 109:5681–5686.
4. Gahloth D, et al. (2017) Structures of carboxylic acid reductase reveal domain dynamics underlying catalysis. *Nat Chem Biol* 13:975–981.
5. Ehmann DE, Gehring AM, Walsh CT (1999) Lysine biosynthesis in *Saccharomyces cerevisiae*: Mechanism of α-aminoadipate reductase (Lys2) involves posttranslational phosphopantetheinylation by Lys5. *Biochemistry* 38:6171–6177.
6. Yu X, et al. (2016) Biosynthesis of strained piperazine alkaloid: Uncovering the concise pathway of herquline A. *J Am Chem Soc* 138:13529–13532.
7. Frost JW, Li K (1998) Synthesis of vanillin from glucose. *J Am Chem Soc* 120:10545–10546.
8. Fraatz MA, et al. (2018) Biotechnological production of methyl-branched aldehydes. *J Agric Food Chem* 66:2387–2392.
9. Akhtar MK, Turner NJ, Jones PR (2013) Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. *Proc Natl Acad Sci USA* 110:87–92.
10. Sheppard MJ, Kunjapur AM, Prather KLJ (2016) Modular and selective biosynthesis of gasoline-range alkanes. *Metab Eng* 33:28–40.
11. Carroll AL, Desai SH, Atsumi S (2016) Microbial production of scent and flavor compounds. *Curr Opin Biotechnol* 37:8–15.
12. Watkins TN, et al. (January 18, 2018) Comparative transcriptomics of *Aspergillus fumigatus* strains upon exposure to human airway epithelial cells. *Microb Genom*10.1099/mgen.0.000154.
13. Takahashi H, et al. (2017) Global gene expression reveals stress-responsive genes in *Aspergillus fumigatus* mycelia. *BMC Genomics* 18:942.
14. Dong Y, et al. (2015) Global genome and transcriptome analyses of Magnaporthe oryzae epidemic isolate 98-06 uncover novel effectors and pathogenicity-related genes, revealing gene gain and lose dynamics in genome evolution. *PLoS Pathog* 11:e1004801.
15. Rokas A, Wisecaver JH, Lind AL (2018) The birth, evolution and death of metabolic gene clusters in fungi. *Nat Rev Microbiol* 16:731–744.
16. Gerlt JA (2017) Genomic enzymology: Web tools for leveraging protein family sequence-function space and genome context to discover novel functions. *Biochemistry* 56:4293–4308.
17. Stachelhaus T, Mootz HD, Marahiel MA (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol* 6:493–505.
18. Challis GL, Ravel J, Townsend CA (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem Biol* 7:211–224.
19. Bachmann BO, Ravel J (2009) In silico prediction of microbial secondary metabolic pathways from DNA sequence data. *Methods Enzymol* 458:181–217.
20. Miyanaga A, Cieślak J, Shinohara Y, Kudo F, Eguchi T (2014) The crystal structure of the adenylation enzyme VinN reveals a unique β-amino acid recognition mechanism. *J Biol Chem* 289:31448–31457.
21. Niquille DL, et al. (2018) Nonribosomal biosynthesis of backbone-modified peptides. *Nat Chem* 10:282–287.
22. Zha L, et al. (2017) Colibactin assembly line enzymes use *S*-adenosylmethionine to build a cyclopropane ring. *Nat Chem Biol* 13:1063–1065.
23. May JJ, Kessler N, Marahiel MA, Stubbs MT (2002) Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. *Proc Natl Acad Sci USA* 99:12120–12125.
24. Balibar CJ, Howard-Jones AR, Walsh CT (2007) Terrequinone A biosynthesis through L-tryptophan oxidation, dimerization and bisprenylation. *Nat Chem Biol* 3:584–592.
25. Ames BD, Walsh CT (2010) Anthranilate-activating modules from fungal nonribosomal peptide assembly lines. *Biochemistry* 49:3351–3365.
26. Wackler B, et al. (2011) Ralfuranone biosynthesis in Ralstonia solanacearum suggests functional divergence in the quinone synthetase family of enzymes. *Chem Biol* 18: 354–360.
27. Hai Y, Tang Y (2018) Biosynthesis of long-chain *N*-acyl amide by a truncated polyketide synthase-nonribosomal peptide synthetase hybrid megasynthase in fungi. *J Am Chem Soc* 140:1271–1274.
28. Hermann JC, et al. (2007) Structure-based activity prediction for an enzyme of unknown function. *Nature* 448:775–779.
29. Zhao S, et al. (2013) Discovery of new enzymes and metabolic pathways by using structure and genome context. *Nature* 502:698–702.
30. Ma JC, Dougherty DA (1997) The cation-π interaction. *Chem Rev* 97:1303–1324.
31. Dougherty DA (2013) The cation-π interaction. *Acc Chem Res* 46:885–893.
32. Schiefner A, et al. (2004) Cation-π interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligand-binding protein ProX from *Escherichia coli*. *J Biol Chem* 279:5588–5596.
33. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ (2007) How chromatin-binding modules interpret histone modifications: Lessons from professional pocket pickers. *Nat Struct Mol Biol* 14:1025–1040.
34. Park Y-I, Gander JE (1998) Choline derivatives involved in osmotolerance of *Penicillium fellutanum*. *Appl Environ Microbiol* 64:273–278.
35. Kadi N, Challis GL (2009) Chapter 17. Siderophore biosynthesis a substrate specificity assay for nonribosomal peptide synthetase-independent siderophore synthetases involving trapping of acyl-adenylate intermediates with hydroxylamine. *Methods Enzymol* 458:431–457.
36. Mori N, Sano M, Tani Y, Yama H (1980) Purification and properties of sarcosine oxidase from *Cylindrocarpon didymium* M-1. *Agric Biol Chem* 44:1391–1397.
37. Rumble JR, ed (2018–2019) *CRC Handbook of Chemistry and Physics* (CRC Press, Boca Raton, FL), 99th Ed.
38. Fan F, Germann MW, Gadda G (2006) Mechanistic studies of choline oxidase with betaine aldehyde and its isosteric analogue 3,3-dimethylbutyraldehyde. *Biochemistry* 45:1979–1986.
39. Park Y-I, Buszko ML, Gander JE (1999) Glycine betaine: Reserve form of choline in *Penicillium fellutanum* in low-sulfate medium. *Appl Environ Microbiol* 65:1340–1342.
40. Kacprzak MM, Lewandowska I, Matthews RG, Paszewski A (2003) Transcriptional regulation of methionine synthase by homocysteine and choline in *Aspergillus nidulans*. *Biochem J* 376:517–524.
41. Turgeon BG, Condon B, Liu J, Zhang N (2010) Protoplast transformation of filamentous fungi. *Methods Mol Biol* 638:3–19.
42. Markham P, Robson GD, Bainbridge BW, Trinci AP (1993) Choline: Its role in the growth of filamentous fungi and the regulation of mycelial morphology. *FEMS Microbiol Rev* 10:287–300.
43. Atkinson DE (1977) *Cellular Energy Metabolism and Its Regulation* (Academic, New York), pp 74–75.