Carnitine metabolism to trimethylamine by an unusual Rieske-type oxygenase from human microbiota

Yijun Zhu[*1], Eleanor Jameson[*1], Marialuisa Cosatti[†1], Hendrik Schäfer[*], Kumar Rajakumar[†], Timothy D.H. Bugg[‡], Yin Chen[*]

[*]School of Life Sciences, University of Warwick, CV4 7AL, United Kingdom

[†]Department of Infection, Immunity and Inflammation, University of Leicester, LE1 9HN, United Kingdom

[‡]Department of Chemistry, University of Warwick, CV4 7AL, United Kingdom

[1]These authors contribute equally to this work.

Correspondence to Dr Y. Chen, School of Life Sciences, University of Warwick, Coventry, CV4 7AL, United Kingdom.

Phone 00 44 24 76528976

Fax 00 44 24 76523568

E-mail: Y.chen.25@warwick.ac.uk

Abstract

Dietary intake of L-carnitine can promote cardiovascular diseases in humans through microbial production of trimethylamine (TMA) and its subsequent oxidation to trimethylamine *N*-oxide (TMAO) by hepatic flavin-containing monooxygenases. Although our microbiota are responsible for TMA formation from carnitine, the underpinning molecular and biochemical mechanisms remain unclear. In this study, using bioinformatics approaches, we first identified a two-component Rieske-type oxygenase/reductase (CntAB) and associated gene cluster proposed to be involved in carnitine metabolism in representative genomes of the human microbiota. CntA belongs to a group of previously uncharacterized Rieske-type proteins and has an unusual "bridging" glutamate but not the aspartate residue, which is believed to facilitate inter-subunit electron transfer between the Rieske centre and the catalytic mononuclear iron centre. Using *Acinetobacter baumannii* as the model, we then demonstrate that *cntAB* is essential in carnitine degradation to TMA. Heterologous overexpression of *cntAB* enables *Escherichia coli* to produce TMA, confirming that these genes are sufficient in TMA formation. Site-directed mutagenesis experiments have confirmed that this unusual "bridging glutamate" residue in CntA is essential in catalysis and neither mutant (E205D, E205A) is able to produce TMA. Together, our study reveals the molecular and biochemical mechanisms underpinning carnitine metabolism to TMA in human microbiota and assigns the role of this novel group of Rieske-type proteins in microbial carnitine metabolism.

Methylated amine metabolism | comparative genomics | gut microbiota

Significance

Metabolism of L-carnitine, a compound abundant in human diet, to trimethylamine by human microbiota has been shown to promote atherosclerosis and subsequent development of heart disease. However, the underpinning molecular and biochemical mechanisms remain unknown. In this study, we reveal for the first time that a novel Rieske type protein is responsible for carnitine transformation to trimethylamine from human microbiota. Knowledge gained in our study opens the opportunity not only to explore Rieske protein inhibitors in preventing trimethylamine formation in animal studies and clinical trials, but also for its use as a functional genetic marker to better understand human microbiota and their dynamics in our health and disease in future epidemiological studies and dietary interventions.

\ **body**

Introduction

It is increasingly clear that the human microbiota plays an essential role in our health and disease (1-5). Understanding the metabolic potential of the human microbiota and its interaction with and regulation by the host and the environment holds the key to unravel the dynamic relationship between ourselves and our associated microbes (4-6). Over the last decade, our knowledge of the human microbiota has been significantly improved thanks to the technological advances, including high throughput sequencing, development of powerful bioinformatics, and the use of germ-free animal models (7-13). We can now not only characterise the taxonomic composition, species richness and dynamics, but combine the genetic potential encoded in human microbiota and establish the core metabolic pathways enabled by direct sequencing of the human microbiome (14-15).

Advances in our understanding of the microbiome functions, however, do not match the pace of taxonomic characterization of species diversity and dynamics (16-18). Fully resolving the functional capacity encoded in the human microbiome and the dynamic effects on health and disease still remains a great challenge (11, 16-18). Direct sequencing of the human microbiome generates datasets dominated by genes encoding pathways for central metabolism, *e.g.* transcriptional and translational machinery, ATP synthesis, *etc.* (14-15), therefore contributing little to our understanding of the variable functional capacity encoded in human microbiota between individuals (7-8). Furthermore, a large fraction of the genes encoded in the human microbiome remain to be functionally characterized (14-15). Assigning functions encoded in the human microbiome using existing databases can be problematic. For example, the Pfam protein database currently contains over 25% of protein families with no assigned functions (release 26.0) (19).

Lack of functional characterization of key microbial functions in our microbiota is exemplified by very recent studies on cardiovascular diseases (20-23). These studies have shown that the human microbiota is responsible for the production of trimethylamine *N*-oxide (TMAO) which is believed to promote atherogenesis through its interaction with macrophages and lipid metabolism (20-23). TMAO is derived from the microbial metabolism of dietary quaternary amines, *e.g.* choline, L-carnitine, glycine betaine (GBT) and phosphatidylcholine, to trimethylamine (TMA), which is subsequently oxidized to TMAO by the host hepatic flavin monooxygenases (21-22). L-carnitine (hereafter as carnitine unless otherwise specified) is considered as an important nutrient for human health, playing a key role in mitochondrial fatty acid beta-oxidation (24-25). Although carnitine can be acquired through endogenous biosynthesis, our daily demand is largely met through dietary intake from carnitine-containing food (26). It is known that a significant proportion of dietary carnitine can be further metabolized by microbiota before absorption (20, 27). This microbial-mediated metabolic pathway not only diverts carnitine away from the host, causing conditional carnitine deficiency in certain human populations, but promotes TMAO formation and subsequent increased risk of atherosclerosis (20-21). However, the underlying genetic and biochemical mechanisms of carnitine-dependent TMA production in human microbiota have not been uncovered.

In this report, we describe the discovery of the genetic and biochemical mechanisms for TMA production from carnitine in representative human microbiota (*Gammaproteobacteria, Betaproteobacteria, Firmicutes*) through a synthesis of bioinformatics, genetic and biochemical approaches. These *in vivo* experiments with microbial isolates prove valuable for exploring the metabolic attributes of a particular group of microbes in carnitine metabolism

guarded by bioinformatics (10). This newly-discovered carnitine-to-TMA enzyme is composed of an oxygenase component (CntA) and a reductase component (CntB). CntA belongs to a previously uncharacterized group of Rieske-type proteins, which are best known for ring-hydroxylation of aromatic hydrocarbons (28). It has an unusual alteration of the so-called "bridging" aspartate (substituted by a glutamate residue, E205), which plays an essential role in electron transfer in catalysis in the Rieske-type protein family (29-32).

Results and discussions

Discovery of a carnitine utilization gene cluster and identification of a candidate carnitine oxygenase (cntAB)

We used the Human Microbiome Project (HMP) reference genomes (15) for mining for carnitine degrading enzymes guarded by the following hypotheses. Microbial conversion of carnitine to TMA has been studied in bacteria isolated from humans, *Acinetobacter calcoaceticus* (33) and *Serratia marcescens* (34). It is known that the cleavage of the carbon (C) – nitrogen (N) bond of carnitine produces TMA and a four-carbon (C4) molecule, which likely enters the central tricarboxylic acid cycle in the form of malate or succinate. We therefore hypothesize that the enzyme responsible for splitting the C-N bond is clustered in the genome with enzymes responsible for the synthesis of malate and/or succinate since the C4 unit is further used as a carbon source in these bacteria (33-34).

We further reasoned that a transporter is needed for microbial carnitine transport. Two types of bacterial carnitine transporters are known, a BCCT type permease (35) and an ABC type active transporter of the choline/betaine/carnitine family (36). We therefore used BLASTP algorithm to search the HMP reference genomes (754 were available as of February 2013) for

genes encoding either CaiT (a BCCT type carnitine-specific antiporter) or CaiX (the carnitine-specific substrate binding protein of the ABC transporter cassette). Our BLASTP search data revealed that many gammaproteobacterial HMP genomes contain CaiT but not CaiX homologs (**Table S1**). CaiT is also found in some *Betaproteobacteria* and *Firmicutes*.

We then simultaneously inspected the neighbourhood of *caiT* for genes involved in malate and/or succinate metabolism and this resulted in the identification of 5 groups of gene clusters in 39 HMP reference genomes (**Figure 1A)** with representative isolates from various sites of the body, including skin, airway, gastrointestinal tract and faeces **(Table S2**). We focused our analyses on *Acinetobacter* spp. since they are known to degrade carnitine to TMA (33). A close investigation of the *caiT* neighbourhood revealed the presence of genes likely to be involved in C4 metabolism (**Figure 1A**). These include genes coding for a malic semi-aldehyde dehydrogenase and a malate dehydrogenase, together channelling the C4 carbon into the central tricarboxylic acid cycle. A conserved *lysR* type transcriptional regulator is present in these gene clusters. Immediately downstream of the carnitine transporter c*aiT* is the gene encoding an acylcarnitine hydrolase which is required for growth and hydrolysis of acylcarnitine to carnitine (37). The functions of two other genes which are always present in the *caiT* gene cluster are not explicit, which we designate as *cntA* and *cntB* respectively.

Bioinformatic analyses predicted that CntB encodes a NAD(P)-dependent reductase, containing a flavin-binding domain as well as a plant-type ferredoxin [2Fe-2S] domain (**Figure 1B, Figure S1**). CntA is predicted to be a member of the Rieske-type protein family, which is characterized by two histidine and two cysteine residues coordinating the [2Fe-2S] cluster (28). Sequence analyses of CntA from HMP bacterial genomes showed conserved

Rieske motif and a catalytic mononuclear iron domain (**Figure 1B**). To date, most characterized microbial Rieske-type proteins are involved in the catabolism of ring-structured aromatic compounds (**Table S3**) (38). However, recent bioinfomatic and evolutional analyses suggest that Rieske-type proteins are much more diverse than previously thought (38-40). Indeed, eukaryotic Rieske-type terminal oxygenases have also been identified, *e.g.* choline monooxygenase (41-42) and recent studies have confirmed that Rieske-type proteins can carry out non-ring hydroxylation reactions, *e.g.* oxidative demethylation (43) and oxidative carbocyclization (44). Furthermore, many novel Rieske-type proteins have been identified with no assigned function, many of which originated from newly sequenced microbial genomes (38). Our phylogenetic analyses reveal that CntA forms a distinct group in the Rieske-type protein family which is more closely related to the eukaryotic choline monooxygenases (**Figure 1C**). Our bioinformatics and phylogenetic analyses therefore suggest that *cntAB* may encode a novel enzyme which catalyses the initial step of carnitine degradation to TMA (and a C4 compound) although C-N bond cleavage by Rieske and Rieske-type enzymes has not previously been reported. The predicted pathway for carnitine metabolism to TMA is shown in **Figure 1D**.

Deletion of either cntA or cntB abolishes TMA formation from carnitine in Acinetobacter baumannii

In order to test whether this gene cluster is indeed involved in carnitine transformation, we used *Acinetobacter baumannii* ATCC19606 as a model. The genetics for *A. baumannii* has been established and *Acinetobacter* spp. are known to degrade carnitine to TMA (33). This bacterium was cultivated in a defined medium with carnitine as the sole carbon source in order to establish whether or not it could produce TMA from carnitine. As predicted from the bioinformatics analyses, this strain could grow on carnitine as the sole carbon source (**Figure**

2). TMA production was observed in the culture supernatant supplemented with carnitine but not succinate. We then carried out marker-exchange mutagenesis to investigate if *cntA*/*cntB* genes are essential in carnitine-dependent TMA production in this strain. The results shown in **Figure 2** demonstrate that the mutants lacking either *cntA* or *cntB*, which was replaced by a gentamicin resistance gene cassette (*aacC1*), could no longer grow on carnitine as a sole carbon and energy source whereas the growth on succinate was not affected. Furthermore, when the mutants were complemented with the native copy of the carnitine degradation gene cluster, their ability to grow on carnitine was restored (Figure 2).

We performed further experiments to validate the mutants by quantifying TMA production and carnitine consumption from the wild type, the mutants and the complemented mutants of *A. baumannii* ATCC19606 (**Figure 3**). Because the mutants did not grow on carnitine alone, we supplemented the medium with carnitine and succinate. As predicted, both mutants lost the ability to catalyse TMA formation from carnitine although they could grow on succinate as the sole carbon and energy source, whereas the complemented mutants restored the ability to convert carnitine to TMA. Taken together, the experiments confirmed that *cntA* and *cntB* are essential in TMA formation from carnitine.

Heterologous expression of carnitine oxygenase (cntAB) in Escherichia coli

To complement our experiment *in vivo*, we performed further experiment *in vitro* by heterologous expression of *cntAB* in *Escherichia coli*. We cloned these two genes (*cntA, cntB*) from *A. baumannii* into an inducible T7 promoter-specific expression system to assess whether they are sufficient to perform carnitine-dependent TMA production heterologously. Using ion-exchange chromatography, we quantified TMA formation from carnitine using the supernatant of recombinant *E. coli* containing either over-expressed CntA or CntB and no

TMA production was seen (**Figure 4**). However, when the two genes were co-expressed, TMA was detected from carnitine degradation using the cell-free culture extracts.

In order to further confirm that CntAB is sufficient for carnitine-dependent TMA production, we purified CntA and CntB by affinity chromatography and successfully reconstituted the activity of carnitine-to-TMA degradation using the purified recombinant proteins (**Figures S2 & S3**). The identity of TMA produced from carnitine oxidation by CntAB was further confirmed by gas chromatography - mass spectrometry using authentic TMA standards (Figure S3). Overall, our experiments demonstrate that CntA and CntB are necessary and sufficient for *in vitro* carnitine degradation to TMA.

The unusual "bridging glutamate" residue is essential in carnitine oxidation

Phylogenetic analyses place the CntA protein within the Rieske-type protein family (**Figure 1C**). This is confirmed by the presence of the highly conserved Rieske sequence motif [-$CXHX_{15\text{-}17}CXXH$-] in CntA in HMP reference genomes (**Figure 5A**). A close investigation of the catalytic mononuclear iron centre revealed the conserved two-histidine-one-carboxylate facial triad motif in these CntA proteins (45). However, a key difference between CntA and other Rieske-type terminal oxygenases lies in the unusual substitution of the previously identified, highly conserved aspartate residue located immediately in front of the first histidine residue of the mononuclear iron centre (E205, **Figure 5A**). This glutamate residue is, however, strictly conserved in all CntA proteins identified from the 39 genomes of HMP reference strains. Although the aspartate-to-glutamate substitution is conservative, it is rather unusual. X-ray crystal structures of Rieske oxygenases (*e.g.* naphthalene 1,2-dioxygenase, biphenyl 2,3-dioxygenase, nitrobenzene 1,2-dioxygenase) showed that the mononuclear iron and the Rieske centre on the same subunit are too far (> 40 Å) to allow

direct electron transfer between them (46-49). The mononuclear iron centres, however, are located only ~ 12 Å from the Rieske centres on the adjacent subunits of the homotrimers (46-49). It is believed that inter-subunit electron transfer occurs between the ligating histidine residues which are coordinated by a so-called “bridging aspartate” residue (29-32) (**Figure S4)**, and substitutions of this aspartate residue to glutamate (D205E in naphthalene dioxygenase) severely diminished its catalytic activity (29).

To gain more insight into the role of this glutamate residue in CntA, we performed site-directed mutagenesis experiments, changing this glutamate residue to either aspartate (E205D) or alanine (E205A). The mutants were purified from recombinant *Escherichia coli* and activity assays were performed by quantifying NADH oxidation and TMA production. The results shown in **Figure 5B** demonstrate that this glutamate residue is essential in electron transfer between the Rieske centre and the mononuclear iron centre since electron transfer (as determined by coupling efficiency and TMA formation) was completely abolished in the mutants. Experiments using circular dichroism and native polyacrylamide gel electrophoresis demonstrated that site-directed mutagenesis of CntA neither altered its secondary structure nor the oligomeric state of the protein (Figure S5). Our data, therefore, indicate that in CntA, a glutamate, but not aspartate, residue in this position is crucial in catalysis since neither of the mutants (E205A; E205D) could catalyse carnitine degradation to TMA (**Figure 5B**). It is interesting to note that the glutamate residue is found in at least two other enzymes of the Rieske-type protein family, dimethylproline demethylase (Stc2) (43) and GBT demethylase (GbcA) (50), although no further data were available on the role of this residue in these proteins. Our study therefore suggests that caution needs to be taken when interpreting structure-function relationships of Rieske-type proteins using existing structures which are dominated by ring-hydroxylating oxygenases. Further X-ray structures are clearly warranted

to reveal the function-structure relationships of an expanded Rieske-type protein family as revealed by bioinformatics approaches (38-40).

C**oncluding remarks**

The last decade has witnessed unparalleled progress in research on the human microbiota and its complex and dynamic relationship to our health and disease. It has become increasingly evident that understanding the functional capacity encoded in the human microbiota and its variation between individuals is necessary for future personalized healthcare and targeted medication. A call for a community response to work collaboratively on the functional annotation of uncharacterized proteins was made almost a decade ago (51), and yet proteins of unknown function in prokaryotic and eukaryotic genomes are still increasingly accumulating (18-19). Indeed, a large proportion of the proteins encoded in the human microbiome have not yet been functionally characterized (14-15), leaving a serious gap in our knowledge for complete understanding of the role of our microbiota. On the other hand, a number of specific activities of human microbiota have been discovered, with the corresponding encoding genes and biochemical mechanisms remaining unknown (reviewed in 4, 11).

The key role played by human microbiota in TMA production has been known for more than a century, but we just began to understand the underpinning molecular and biochemical mechanisms (52). The vital importance of the microbiota-mediated metabolic pathway for TMA formation has been convincingly highlighted by several recent studies, linking increased levels of plasma TMAO, the metabolite produced by hepatic oxidation of TMA, with elevated risk of atherosclerosis and associated acute cardiovascular diseases in humans (20-23). Demystifying the genetic and biochemical mechanisms in TMA formation can yield

novel targets for future diagnosis and form baseline knowledge for personalized therapeutic strategies targeting an individual's microbiota. In this study, we report the discovery of a novel Rieske protein involved in TMA formation from carnitine in human microbiota. Coincidently, Rieske proteins have been extensively studied over the last few decades and many crystal structures of Rieske proteins are readily available. Our discovery now offers the opportunity to explore previously studied Rieske protein inhibitors in preventing TMA formation in animal studies and clinical trials (53). Furthermore, knowledge gained in our study and others (52) now offers the use of functional genetic markers (e.g. *cntA*, *cutC*), in addition to the taxonomic ribosomal RNA markers, to better understand our microbiota and their dynamics in human health and disease in large scale epidemiological and dietary intervention studies. What's more, beyond the specifics of this study, one can also envisage that the synthesis of genetic, biochemical and bioinformatics approaches can be a valuable tool for future functional genomic studies of human microbiota, not only addressing emerging issues in microbiota-host cross-talk, but also starting to fill in the gap in our knowledge of proteins of unknown functions in human microbiome.

Materials and methods

Bioinformatic identification of the carnitine oxygenase gene cluster. We used sequenced microbial genomes from the HMP as the database for mining the genes involved in carnitine-to-TMA degradation. HMP references genomes were selected and analysed through the IMG program on the Joint Genome Institute website (http://www.hmpdacc-resources.org/cgi-bin/imgm_hmp/main.cgi). Searching for homologues encoding either the carnitine-specific ABC transporter or the BCCT type permease was carried out through the BLASTP algorithm (E value -50) using the following queries sequences CaiX (PA5388) (36), CaiT (CAA52110)

(54). No CaiX homologues were found in the HMP reference genomes, including *Serratia* spp. and *Acinetobacter* spp. which have been previously shown to catalyse carnitine degradation to TMA (33-34). 122 close homologues (E≤-50) of CaiT were found in 91 unique genomes of *Gammaproteobacteria*, *Betaproteobacteria* and *Firmicutes*, including *Acinetobacter* species. Because cleavage of the trimethylammonium molecule from carnitine results in the release of a C4 carbon, the neighbourhood of *caiT* was manually inspected for genes encoding enzymes involved in the metabolism of C4 molecules, including malate, succinate, fumarate and oxaloacetate. Sequence alignment and phylogenetic analyses were performed as described previously (55).

Marker exchange mutagenesis of *cntA/cntB* in *A. baumannii*. Cultivation of *A. baumannii* was carried out in a defined medium supplemented with either succinate or carnitine (or both) as the sole carbon source. Targeted deletion of *cntA*/*cntB* was carried out by marker exchange mutagenesis as described previously (56). The mutants (Δ*cntA*; Δ*cntB*) were complimented by plasmid pKR706 by cloning the native carnitine oxygenase gene cluster into the vector pMQ300 (see supporting information). Bacterial strains, plasmids and primers used in this study are listed in **Table S4**).

Heterologous overexpression of *cntA*/*cntB*, site-directed mutagenesis and characterization of the CntA mutants. The *cntA* gene was amplified from *A. baumannii* and inserted into the vector pET28a (Novagen). Co-expression of *cntA*/*cntB* was achieved by insertion into pCOLADuet-1 under the *BamHI*/*HindIII* and the *NdeI*/*KpnI* sites, respectively. The CntA mutants (E205D, E205A) were chemically synthesized by GenScript and inserted into the expression vector pET28a under the *NdeI*/*HindIII* sites. The resulting plasmids were then transformed into the expression host *E. coli* BLR(DE3) pLysS (Merck Biosciences).

Cultivation of recombinant *E. coli*, protein induction by isopropyl β-D-1-thiogalactopyranoside (IPTG) and further purification using His-tag affinity chromatography were detailed in supplementary information. Enzyme assays were performed at room temperature (~ 22°C) by quantifying NADH oxidation and carnitine-dependent TMA production. A 1-ml enzyme assay mixture contained 10 mM HEPES buffer (pH 7.6), 60 μg purified CntA and CntB, respectively, 0.25 mM carnitine and 0.25 mM NADH. Coupling efficiency was determined as the ratio of the total amount of TMA formed to the amount of NADH consumed.

Analytical methods. Carnitine and TMA were quantified by a cation-exchange ion chromatography (Metrohm 881 Compact IC Pro) equipped with a Metrosep C4/250 mm separation column and a conductivity detector (Metrohm). NADH oxidation was quantified using a Shimadzu UV-VIS 1800 spectrophotometer by following decrease of absorbance at 340 nm ($\varepsilon = 6.2\ mM^{-1} \cdot cm^{-1}$).

Acknowledgements: This work is supported by Royal Society (RG2011/R1) and partially by Natural Environment Research Council, United Kingdom (NE/I027061/1).

References

1. Ley RE, Turnbaugh PJ, Klein S, Gordon JI (2006) Human gut microbes associated with obesity. *Nature* 444:1022-1023.
2. Turnbaugh PJ, et al. (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444:1027-1031.
3. Gill SR, et al. (2006) Metagenomic analysis of the human distal gut microbiome. *Science* 312:1355-1359.
4. Nicholson JK, et al. (2012) Host-gut microbiota metabolic interactions. *Science* 336:1262-1267.
5. Koren O, et al. (2011) Human oral, gut, and plaque microbiota in patients with atherosclerosis. *Proc Natl Acad Sci USA* 108 suppl 1:4592-4598.
6. Wu GD, et al. (2011) Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334:105-108.
7. Lozupone CA, Strombaugh JI, Gordon JI, Jansson JK, Knight R. (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* 489:220-230.
8. Tremaroli V, Bäckhed F. (2012) Functional interactions between the gut microbiota and the host metabolism. *Nature* 489:242-249.
9. Weinstock GM. (2012) Genomic approaches to studying the human microbiota. *Nature* 489:250-256.
10. Spor A, Koren O, Lay R (2011) Unravelling the effects of the environment and host genotype on the gut microbiome. *Nature Rev Microbiol* 9:279-290.
11. Marchesi JR. (2011) Human distal gut microbiome. *Environ Microbiol* 13:3088-3102.
12. Arumugam M, et al. (2011) Enterotypes of the human gut microbiome. *Nature* 473:174-180.

13. Yatsunenko T, et al. (2012) Human gut microbiome viewed across age and geography. *Nature* 486:222-227.
14. Qin J, et al. (2012) A metagenome-wide associated study of gut microbiota in type 2 diabetes. *Nature* 490:55-60.
15. Human Microbiome Project Consortium (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486:207-214.
16. Proctor LM (2011) The human microbiome project in 2011 and beyond. *Cell Host Microbe* 10:287-291.
17. Morgan XC, Segata N, Huttenhower C (2013) Biodiversity and functional genomics in the human microbiome. *Trends Genet* 29:51-58.
18. Galperin MY, Koonin EV (2010) From complete genome sequence to 'complete' understanding? *Trends Biotech* 28:398-406.
19. Punta M. et al. (2012) The Pfam protein family database. *Nucleic Acids Res* 40:D290-301.
20. Koeth RA, et al. (2013) Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med* 19(5):576-585.
21. Bennett BJ, et al. (2013) Trimethylamine-*N*-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation. *Cell Metab* 17(1):49-60.
22. Wang Z, et al. (2011) Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472:57-63.
23. Tang WH et al. (2013) Intestinal microbial metabolism of phosphatidycholine and cardiovascular risk. *N Engl J Med* 368:1575-1584.
24. Friedman S, Fraenkel G. (1955) Reversible enzymatic acetylation of carnitine. *Arch Biochem Biophys* 59:491-501.

393 25. Fritz IB, McEwen B (1957) Effects of carnitine on fatty-acid oxidation by muscle.
394 *Science* 129(3345):334-335.
395 26. Rebouche CJ (1992) Carnitine function and requirements during the life cycle.
396 *FASEB J* 6(15):3379-3386.
397 27. Rebouche CJ, Chenard CA (1991) Metabolic fate of dietary carnitine in human adults:
398 identification and quantification of urinary and fecal metabolites. *J Nutr* 121(4):539-
399 546.
400 28. Ferraro DJ, Gakhar L, Ramaswamy S (2005) Rieske business: structure-function of
401 Rieske non-heme oxygenases. *Biochem Biophys Res Commun* 338:175-190.
402 29. Parales RE, Parales JV, Gibson DT (1999) Aspartate 205 in the catalytic domain of
403 naphthalene dioxygenase is essential for activity. *J Bacteriol* 181:1831-1837.
404 30. Pinto A, Tarasev M, Ballou DP (2006) Substitutions of the “bridging” aspartate 178
405 result in profound changes in the reactivity of the Rieske center of phthalate
406 dioxygenase. *Biochemistry* 45:9032-9041.
407 31. Jiang H, et al. (1996) Site-directed mutagenesis of conserved amino acids in the alpha
408 subunit of toluene dioxygenase: potential mononuclear non-heme iron coordination
409 sites. *J Bacteriol* 178:3133-3139.
410 32. Beharry ZM, et al. (2003) Histidine ligand protonation and redox potential in the
411 rieske dioxygenases: role of a conserved aspartate in anthranilate 1,2-dioxygenase.
412 *Biochemistry* 42:13625-13636.
413 33. Seim H. et al. (1982) Splitting of the C-N bond in carnitine by an enzyme
414 (trimethylamine forming) from membranes of *Acinetobacter calcoaceticus*. *FEMS*
415 *Microbiol Lett* 15:165-167.
416 34. Unemoto T. et al. (1966) Formation of trimethylamine from DL-carnitine by *Serratia*
417 *marcescens*. *Biochim Biophys Acta* 121(1):220-222.

418 35. Ziegler C, et al. (2010) The BCCT family of carriers: from physiology to crystal
419 structure. *Mol Microbiol* 78(1):13-34.
420 36. Chen C. et al. (2010) The ATP-binding cassette transporter Cbc
421 (choline/betaine/carnitine) recruits multiple substrate-binding proteins with strong
422 specificity for distinct quaternary ammonium compounds. *Mol Microbiol* 75(1): 29-45.
423 37. Meadows JA, Wargo MJ (2013) Characterization of *Pseudomonas aeruginosa* growth
424 on O-acylcarnitines and identification of a short-chain acylcarnitine hydrolase. *Appl*
425 *Environ Microbiol* 79:3355-3363.
426 38. Capyk JK, Eltis LD (2012) Phylogenetic analysis reveals the surprising diversity of an
427 oxygenase class. *J Biol Inorg Chem* 17:425-436.
428 39. Schmidt CL, Shaw L (2001) A comprehensive phylogenetic analysis of Rieske and
429 Rieske-type iron-sulfur proteins. *J Bioenerg Biomembr* 33:9-26.
430 40. Kweon O, et al. (2008) A new classification system for bacterial Rieske non-heme
431 iron aromatic ring-hydroxylating oxygenases. *BMC Biochem* 9:11
432 41. Rathinasabapathi B, et al. (1997) Choline monooxygenase, an unusual iron-sulfur
433 enzyme catalyzing the first step of glycine betaine synthesis in plants: prosthetic
434 group characterization and cDNA cloning. *Proc Natl Acad Sci USA* 94:3454-3458.
435 42. Mitsuya S, et al. (2011) Isolation and characterization of a novel peroxisomal choline
436 monooxygenase in barley. *Planta* 234:1215-1226
437 43. Daughtry KD, et al. (2012) Quaternary ammonium oxidative demethylation: X-ray
438 crystallographic, resonance Raman, and UV-visible spectroscopic analysis of a
439 Rieske-type demethylase. *J Am Chem Soc* 134:2823-2834
440 44. Sydor PK, et al. (2011) Regio- and stereodivergent antibiotic oxidative
441 carbocyclizations catalysed by Rieske oxygenase-like enzymes. *Nat Chem* 3:388-392.

45. Hegg EL, Que L Jr (1997) The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur J Biochem* 250:625-629.
46. Karlsson A, et al. (2003) Crystal structure of naphthalene dioxygenase: side-on binding of dioxygen to iron. *Science* 299: 1039-1042.
47. Ferraro DJ, et al. (2007) Structural investigations of the ferredoxin and terminal oxygenase components of the biphenyl 2,3-dioxygenase from *Sphingobium yanoikuyae* B1. *BMC Struct Biol* 7:10
48. Friemann R, et al. (2005) Structural insight into the dioxygenase of nitroarene compounds: the crystal structure of nitrobenzene dioxygenase. *J Mol Biol* 348:1139-1151.
49. Martins BM, Svetlitchnaia T, Dobbek H (2005) 2-Oxoquinoline 8-monooxygenase oxygenase component: active site modulation by Rieske-[2Fe-2S] center oxidation/reduction. *Structure* 13:817-824.
50. Wargo MJ, Szwergold BS, Hogan DA (2008) Identification of two gene cluster and a transcriptional regulator required for *Pseudomonas aeruginosa* glycine betaine catabolism. *J Bacteriol* 190:2690-2699.
51. Roberts RJ (2004) Identifying protein function – a call for community action. *PLoS Biol* 2(3)E42.
52. Craciun S, Balskus EP (2012) Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *PNAS* 109:21307-21312.
53. Link TA, Haase U, Brandt U, von Jagow G. (1993) What information do inhibitors provide about the structure of the hydroquinone oxidation site of ubihydroquinone: cytochrome c oxidoreductase? *J Bioenerg Biomembr* 25:221-232.

54. Eichler K. et al. (1994) Molecular characterization of the *cai* operon necessary for carnitine metabolism in *Escherichia coli*. *Mol Microbiol* 13:775-786.
55. Chen Y, et al. (2001) Bacterial flavin-containing monooxygenase is trimethylamine monooxygenase. *Proc Natl Acad Sci USA* 108:17791-17796.
56. Kochar M. et al. (2012) Deletion of TnAbaR23 results in both expected and unexpected antibiogram changes in a multidrug-resistant *Acinetobacter baumannii* strain. *Antimicrob Agents Chemother* 56(4):1845-1853.
57. Nam JW, et al. (2001) New classification system for oxygenase components involved in ring-hydroxylating oxygenations. *Biosci Biotechnol Biochem* 65:254-263.

Figure legends

Figure 1 Discovery of the putative Rieske-type protein from human microbiota in carnitine-dependent trimethylamine formation. **A**: Putative carnitine-to-trimethylamine gene cluster in representative genomes of human microbiota. CaiT, carnitine transporter; CntA, A Rieske-type oxygenase protein; CntB, a predicted reductase with a plant-type ferridoxin domain. **B**: Analyses of conserved domains in CntA and CntB. FAD: flavin adenine dinucleotide; NAD^+: nicotinamide adenine dinucleotide. **C**: An unrooted phylogenetic tree (~ 305 amino acids) of CntA, microbial Rieske-type terminal oxygenases (group I-IV) and Eukaryotic Rieske-type choline monooxygenases. Microbial Rieske-type terminal oxygenases (group I-IV) are classified based on the nomenclature system of Nam and colleagues (57). These sequences are identified by a unique GenBank or PDB accession number followed by the gene name; other sequences are identified by the name of the species. Bootstrap values greater than 50 are shown (100 replicates). The scale bar represents 1 substitution per 10 amino acids. GBT, glycine betaine. **D**: Predicted pathway of carnitine catabolism *via* CntA/B.

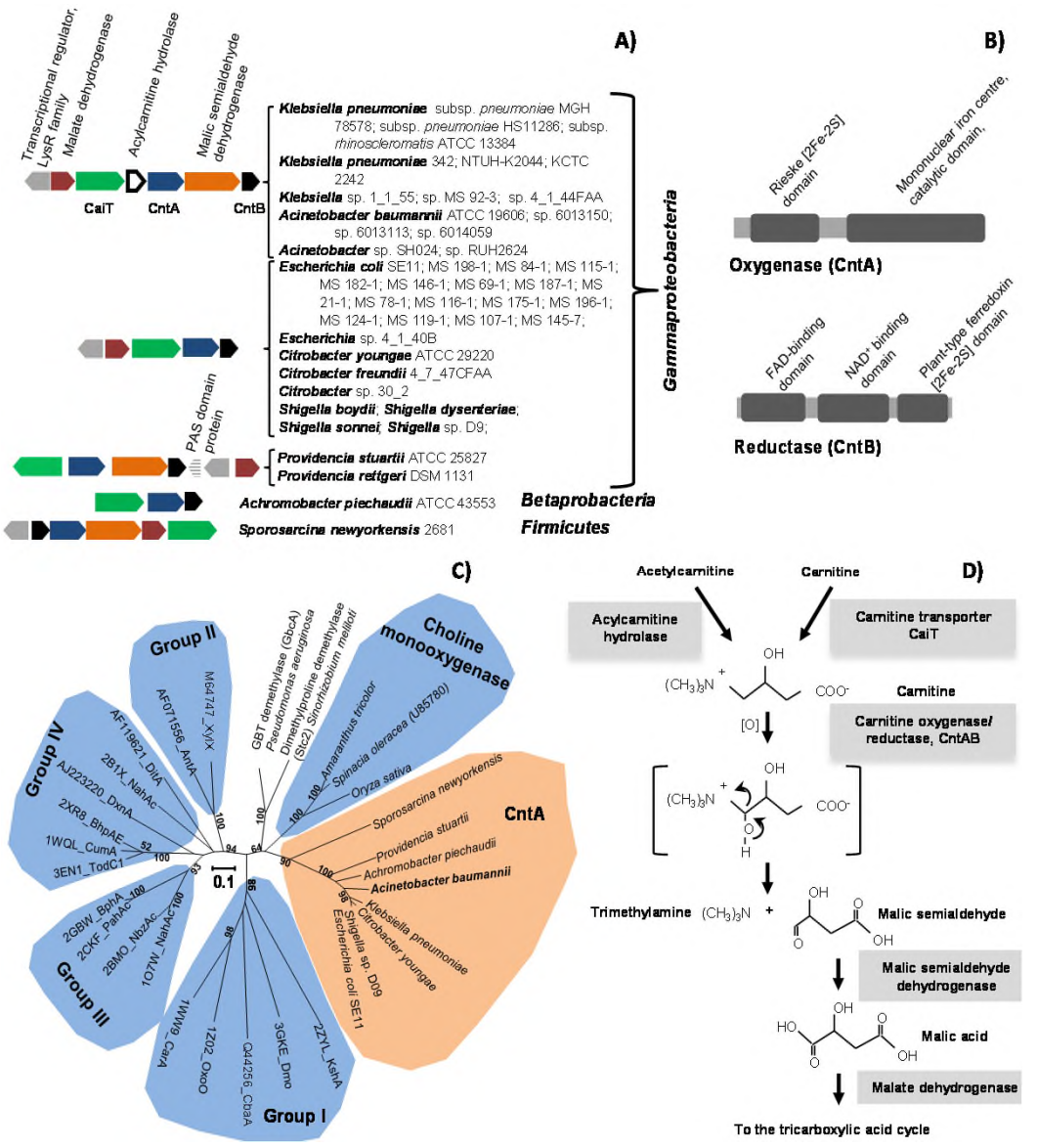
Figure 2: Growth of *Acinetobacter baumannii* ATCC19606 wild type, mutants (Δ*cntA*::*aacC1*, Δ*cntB*::*aacC1*) and complemented mutants with plasmid pKR706 on carnitine (10 mM) or succinate (20 mM) as the sole carbon and energy source. The error bars represent standard deviation from experiments run in triplicate.

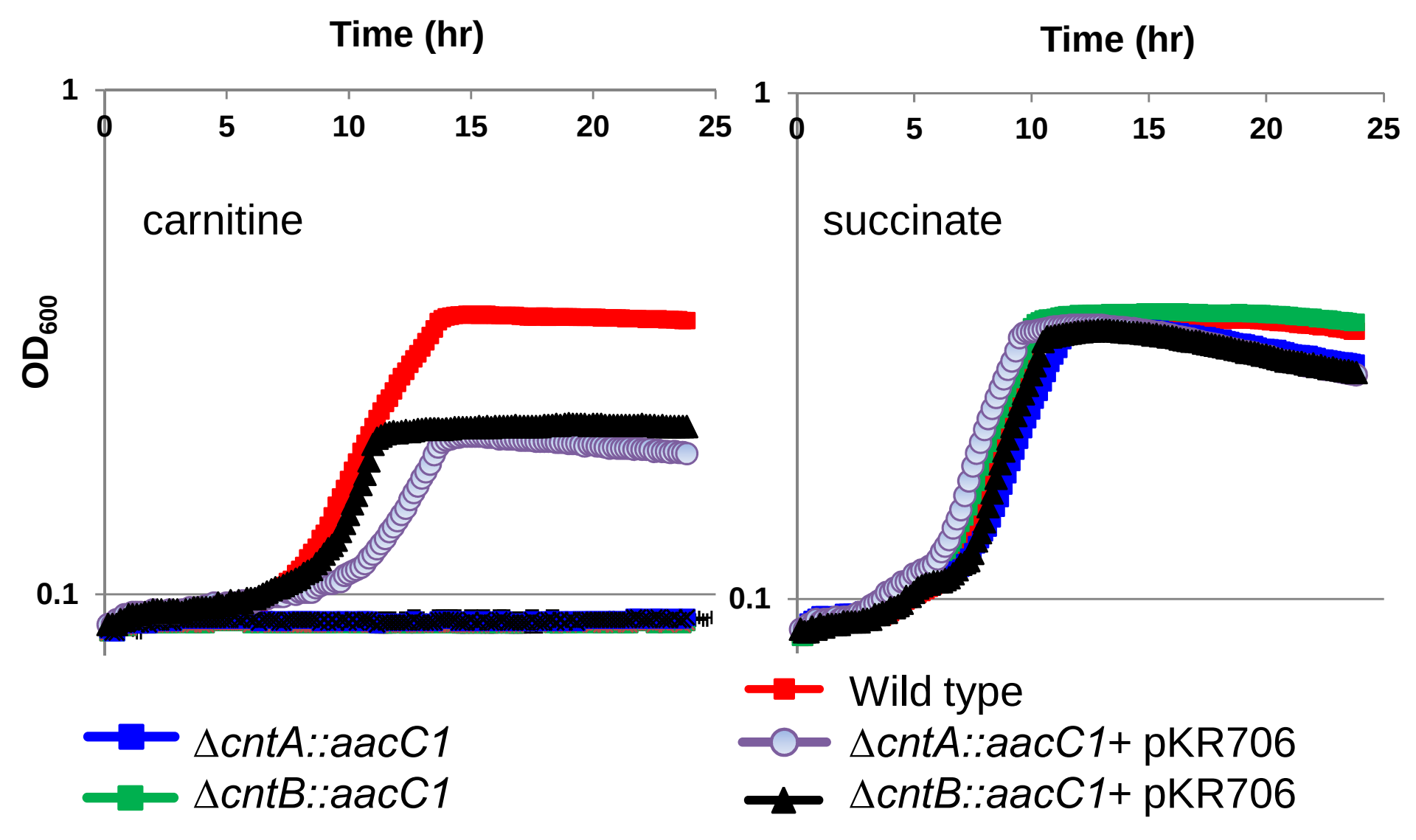
Figure 3: Quantification of trimethylamine (TMA) and carnitine of wild type, mutants (Δ*cntA*::*aacC1*, Δ*cntB*::*aacC1*) and complemented mutants with plasmid pKR706 in the culture medium supplemented with carnitine and succinate. The error bars represent standard deviation from experiments run in triplicate.

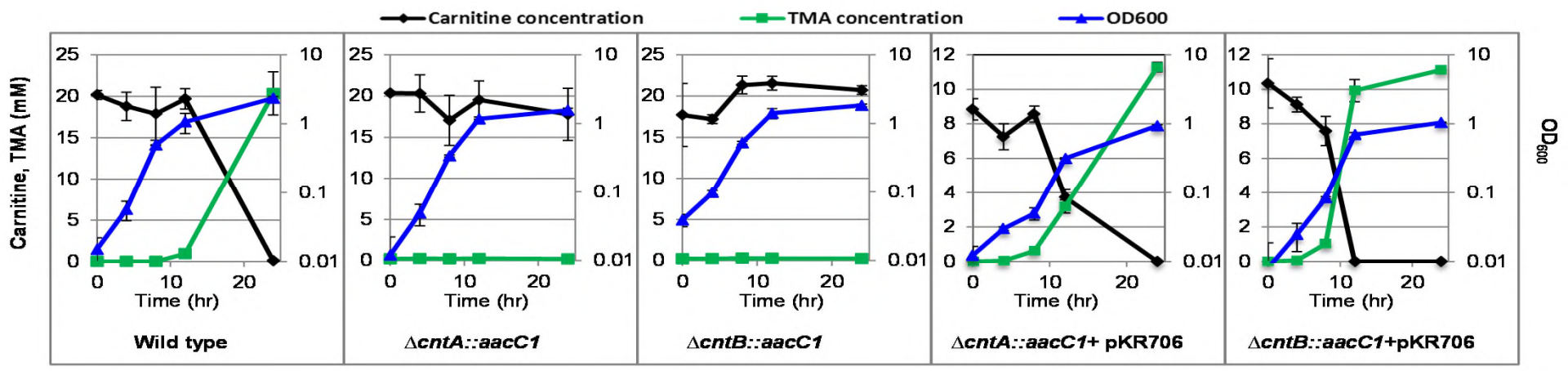
Figure 4 Quantification of trimethylamine (TMA) formation *in vitro* using the supernatant of recombinant *Escherichia coli* containing overexpressed CntA, CntB or CntAB, respectively. The error bars represent standard deviation from experiments run in triplicate.

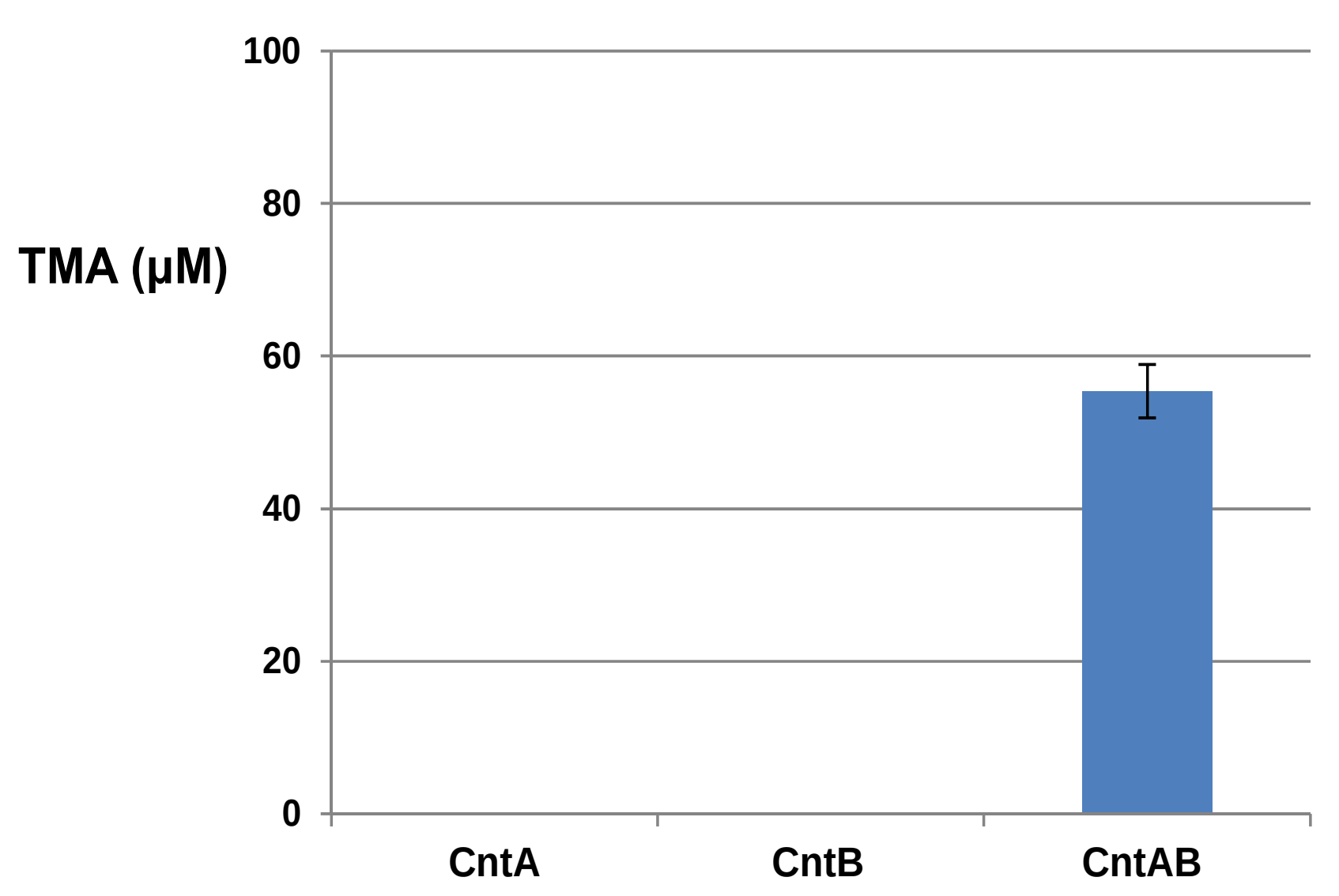
Figure 5 The unusual substitution of ″bridging″ aspartate to ″bridging″ glutamate (E205 in CntA) abolished catalytic activity for carnitine-dependent TMA formation. **A**: Multiple sequence alignment of CntA from representative human microbiota and related Rieske-type proteins. The two boxes highlight the conserved Rieske domain and the mononuclear iron centre domain. The arrow indicates the unusual but conserved glutamate residue in CntA. Each sequence has a unique identifiable label as shown in the legend of Figure 1C. **B**: Quantification of NADH oxidation and TMA production *in vitro* using purified CntA and site-directed mutants (E205A, E205D) in combination with purified CntB. Coupling efficiency is determined as the ratio of the total amount of TMA formed to the amount of

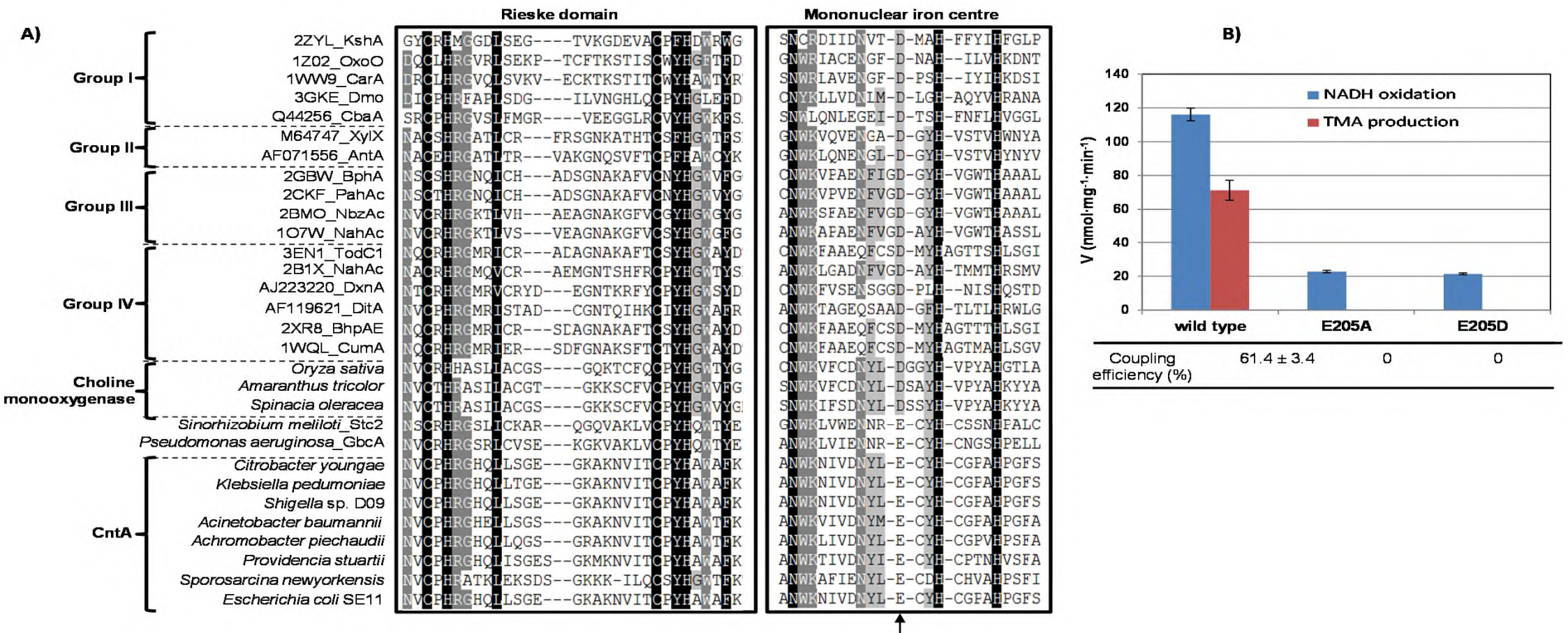
NADH consumed. The error bars indicate standard deviation from experiments run in triplicate.


A)
Transcriptional regulator, LysR family
Malate dehydrogenase
Acylcarnitine hydrolase
Malic semialdehyde dehydrogenase
CaiT
CntA
CntB
Klebsiella pneumoniae subsp. pneumoniae MGH 78578; subsp. pneumoniae HS11286; subsp. rhinoscleromatis ATCC 13384
Klebsiella pneumoniae 342; NTUH-K2044; KCTC 2242
Klebsiella sp. 1_1_55; sp. MS 92-3; sp. 4_1_44FAA
Acinetobacter baumannii ATCC 19606; sp. 6013150; sp. 6013113; sp. 6014059
Acinetobacter sp. SH024; sp. RUH2624
Escherichia coli SE11; MS 198-1; MS 84-1; MS 115-1; MS 182-1; MS 146-1; MS 69-1; MS 187-1; MS 21-1; MS 78-1; MS 116-1; MS 175-1; MS 196-1; MS 124-1; MS 119-1; MS 107-1; MS 145-7;
Escherichia sp. 4_1_40B
Citrobacter youngae ATCC 29220
Citrobacter freundii 4_7_47CFAA
Citrobacter sp. 30_2
Shigella boydii; Shigella dysenteriae; Shigella sonnei; Shigella sp. D9;
PAS domain protein
Providencia stuartii ATCC 25827
Providencia rettgeri DSM 1131
Gammaproteobacteria
Achromobacter piechaudii ATCC 43553
Betaprobacteria
Sporosarcina newyorkensis 2681
Firmicutes
B)
Rieske [2Fe-2S] domain
Mononuclear iron centre, catalytic domain,
Oxygenase (CntA)
FAD-binding domain
NAD+ binding domain
Plant-type ferredoxin [2Fe-2S] domain
Reductase (CntB)
C)
Choline monooxygenase
Group II
Group IV
Group III
Group I
CntA
GBT demethylase (GbcA)
Pseudomonas aeruginosa
Dimethylproline demethylase (StcD) Sinorhizobium meliloti
Amaranthus tricolor
Spinacia oleracea (U85780)
Oryza sativa
Sporosarcina newyorkensis
Providencia stuartii
Achromobacter piechaudii
Acinetobacter baumannii
Klebsiella pneumoniae
Citrobacter youngae
Shigella sp. D9
Escherichia coli SE11
MG4T1_XylX
AF071556_AmiA
AF119821_DllA
2B1X_NahAc
AJ223220_DxnA
2XR8_BhpAE
1WQL_CumA
3EN1_TodC1
2GBW_BphA
2CKF_PahAc
2BMO_NbzAc
1O7W_NahAc
1WM9_CarA
1ZQ2_OxoO
Q44256_ChaA
3GKE_Dmo
2ZYL_KshA
0.1
D)
Acetylcarnitine
Carnitine
Acylcarnitine hydrolase
Carnitine transporter CaiT
Carnitine
Carnitine oxygenase/ reductase, CntAB
[O]
Trimethylamine
Malic semialdehyde
Malic semialdehyde dehydrogenase
Malic acid
Malate dehydrogenase
To the tricarboxylic acid cycle


Time (hr)
1
0
5
10
15
20
25
carnitine
OD600
0.1
Time (hr)
1
0
5
10
15
20
25
succinate
0.1
Wild type
ΔcntA::aacC1
ΔcntA::aacC1+ pKR706
ΔcntB::aacC1
ΔcntB::aacC1+ pKR706


Carnitine concentration
TMA concentration
OD600
Carnitine, TMA (mM)
OD600
Time (hr)
Wild type
ΔcntA::aacC1
ΔcntB::aacC1
ΔcntA::aacC1+ pKR706
ΔcntB::aacC1+pKR706


TMA (µM)
100
80
60
40
20
0
CntA
CntB
CntAB


A)
Rieske domain
Mononuclear iron centre
Group I
2ZYL_KshA
1Z02_OxoO
1WW9_CarA
3GKE_Dmo
Q44256_CbaA
Group II
M64747_XylX
AF071556_AntA
Group III
2GBW_BphA
2CKF_PahAc
2BMO_NbzAc
1O7W_NahAc
Group IV
3EN1_TodC1
2B1X_NahAc
AJ223220_DxnA
AF119621_DitA
2XR8_BhpAE
1WQL_CumA
Choline monooxygenase
Oryza sativa
Amaranthus tricolor
Spinacia oleracea
Sinorhizobium meliloti_Stc2
Pseudomonas aeruginosa_GbcA
CntA
Citrobacter youngae
Klebsiella pedumoniae
Shigella sp. D09
Acinetobacter baumannii
Achromobacter piechaudii
Providencia stuartii
Sporosarcina newyorkensis
Escherichia coli SE11
B)
V (nmol·mg-1·min-1)
NADH oxidation
TMA production
wild type
E205A
E205D
Coupling efficiency (%)
61.4 ± 3.4
0
0

Supplementary material

Cultivation of *Acinetobacter baumannii* ATCC19606 and mutants

Either a defined medium or Luria broth (LB) was used to cultivate the wild type or the mutants (Δ*cntA*::*aacC1*, Δ*cntB*::*aacC1*). The defined medium contained NH_4Cl (1 g l^{-1}), NaCl (0.5 g l^{-1}), KH_2PO_4 (3 g l^{-1}), $Na_2HPO_4 \cdot 7H_2O$ (12.8 g l^{-1}), $MgSO_4 \cdot 7H_2O$ (0.5 g l^{-1}), $CaCl_2 \cdot 2H_2O$ (0.15 g l^{-1}), $Na_2MoO_4 \cdot 2H_2O$ (0.5 mg l^{-1}), $FeCl_3$ (50 µM), and a mix of the following vitamins or supplements, including biotin (0.4 mg l^{-1}), folic acid (0.4 mg l^{-1}), pyridoxine hydrochloride (2 mg l^{-1}), thiamine hydrochloride (1 mg l^{-1}), riboflavin (1 mg l^{-1}), nicotinic acid (1 mg l^{-1}), pantothenic acid (1 mg l^{-1}), vitamin B12 (20 mg l^{-1}), 4-aminobenzoic acid (1 mg l^{-1}) and lipoic acid (1 mg l^{-1}). The Luria broth (LB) medium contained tryptone (10 g l^{-1}), yeast extract (5 g l^{-1}) and NaCl (5 g l^{-1}) and the pH was adjusted using NaOH (0.2 µM). Luria agar (LA) was made as LB and agar was added (15 g l^{-1}).

Growth curves of *A. baumannii* and the mutants on carnitine or succinate

To characterize growth of the ATCC 19606 wild type strain (*Salmonella* Genetic Stock Centre, University of Calgary) and the mutants, carnitine or succinate was added to the defined medium to a final concentration of 20 mM as the sole carbon and energy sources. The wild type strain and the mutants were cultivated in 5 ml LB for 8 h at 37°C while shaking (200 rpm). The cultures were then diluted 1:50 to inoculate the defined medium containing 20 mM succinate. These cultures were then incubated at 37°C for 15-16 h and the pellet was collected by centrifugation at 3,000×*g* for 10 min. Cell pellets were then washed twice with sterile distilled water and re-suspended in 5 ml of water. Finally 5 ml of the defined medium with either succinate or carnitine was inoculated with 100 µl of washed culture and distributed into 8 wells each in a 96-well tissue culture plate (sterile, F-bottom, with lid; Cellstar, Greiner bio-one), which was then incubated at 37°C and the optical density

at 600 nm was recorded every 10 min using a Multiskan GO spectrophotometer (Thermo Scientific).

Cultivation of *A. baumannii* and the mutants for the quantification of TMA and carnitine by ion chromatography

To quantify the production of TMA in the wild type and the mutants, cultures were set up in triplicate in Corning cell culture flasks (surface area 75 cm^2) with 50 ml of the defined medium supplemented with both carnitine (20 mM) and succinate (20 mM). Inoculum was prepared as described above. The flasks were incubated vertically at 37°C while shaking continuously (200 rpm). A 2-ml sample was taken from each flask at T= 0 h, 4 h, 8 h, 12 h and 24 h. Optical density at 600 nm was recorded immediately for every sample at each time point. The samples were then passed through 0.22 μm filters to remove microbial cells and the supernatants were diluted 1:10 in double distilled water, which was then quantified for TMA and carnitine using a cation-exchange ion chromatography (Metrohm 881 Compact IC Pro) equipped with a Metrosep C4/250 mm separation column and a conductivity detector (Metrohm).

Marker-exchange mutagenesis of *cntA* and *cntB*

In order to obtain targeted deletion of *cntA* and *cntB* in *Acinetobacter baumannii* ATCC 19606, we employed the mutagenesis strategy developed by van Aarten and Rajakumar (2011), requiring the creation of a suicide plasmid containing the upstream and the downstream flanking regions of the targeted gene separated by a gentamicin antibiotic cassette, followed by transfer of the construct to *A. baumannii*.

Upstream and downstream flanking regions of *cntA* or *cntB* (~ 1 kb) were amplified from the genomic DNA of ATCC 19606 using primers indicated in ***Table S4*** and the gentamicin resistant gene cassette (*aacC1*) was amplified from pUC18R6K-mini-Tn*7*T-Gm (Choi *et al*., 2005). PCR amplification was performed using Phusion High-Fidelity DNA Polymerase (Thermo Scientific) according to manufacturer's specifications. To join upstream and downstream flanking regions with the gentamicin resistant gene cassette, spliced overlap extension-PCR (SOE-PCR) was performed (Choi and Schweizer, 2005) using KOD Hot Start DNA Polymerase (Merck). 10 ng of each flanking regions and the gentamicin resistant gene cassette were mixed and used as the template. SOE-PCR was run without primers for 5 cycles and then for 30 cycles after the primers were added (PR2768 and PR2771 for *cntA*; PR2774 and PR2777 for *cntB*) to generate a ~3 kb PCR product. Both SOE-PCR products and the vector (pJTOOL-3) were digested with NotI (NEB) for 4h at 37°C and the bands were purified from an agarose gel. Ligation was performed using T4 DNA ligase (Promega) using 50 ng of digested vector and 2:1 for *cntA* or 4:1 for *cntB* insert: vector molar ratio. Ligation was carried out for 7 h at 15 °C and 10 µl of ligation mix was transformed into chemically competent *E. coli* CC118λ*pir* which was selected on LA with 30µg ml^{-1} chloramphenicol and 15µg ml^{-1} gentamicin. Plasmids with insert were selected and tested by digestion to confirm the presence and the orientation of the insert. Finally, selected plasmids were transferred to chemically competent *E. coli* S17-1λ*pir*.

Conjugation was employed to efficiently transfer the suicide plasmid to ATCC 19606. The donor *E. coli* strains and the recipient *A. baumannii* strain ATCC 19606 were grown overnight in LB and 100 µl were used to inoculate 10 ml LB. Optical density at 600 nm was measured at regular intervals and growth was suspended when OD_{600} reached 0.3 ~ 0.4. Bacterial culture (5 ml) were collected by centrifugation at 3,000×*g* for 10 min and washed

with 1 ml LB twice and re-suspended in 1ml LB after the final wash. 200 µl of donor strains were mixed with 200 µl of the recipient, centrifuged at 16,000×*g* for 2 min, re-suspended with 100 µl of 10% (w/v) sterile glycerol, spread on LA and incubated at 37°C for 24 h. Bacterial cultures were collected in 1 ml of 10% (w/v) glycerol, serially diluted and plated on Simmons citrate agar (Oxoid) with 80 µg ml^{-1} gentamicin. Selected colonies were screened using colony-PCR to detect the insertion of the suicide plasmid at either the upstream flanking region or the downstream flanking region. Selection for double-crossover deletion mutants and the loss of plasmid backbone were carried out by sucrose counter-selection due to the presence of *sacB* gene. Single crossover mutants were cultivated in LB medium overnight and plated out on LA with 6% sucrose, which was subsequently incubated overnight at 37°C. Screen for double crossover deletion mutants was carried out by colony PCR. Primers used for screening are listed in ***Table S4***.

Complementation of the Δ*cntA* and Δ*cntB* mutants of *A. baumannii*

A ~7.8 kb DNA fragment comprising the carnitine oxygenase/reductase gene cluster (Figure 1C) was amplified from the genomic DNA of strain ATCC 19606 using the primers PR3103/PR3104 (*Table S4*) with a proof-reading DNA polymerase (Phusion® High-Fidelity DNA polymerase, Thermo Scientific). The PCR product was A-tailed using the GoTaq Polymerase (Promega), purified and cloned into the pGEM-Teasy cloning vector (Promega). The insert was amplified from vector pGEM-Teasy using the primers PR3120/PR3121 with the Phusion® High-Fidelity DNA Polymerase (Thermo Scientific), and the PCR product (30 µg) was subsequently digested with *Bam*HI-HF (NEB) and *Kpn*I-HF (NEB) for 4h at 37°C. The digested product was run in a 1% (w/v) agarose gel and the ~8 kb band was excised from the gel and purified. The pMQ300 vector (Kalidova *et al*., 2011) was prepared similarly by digesting 5 µg of the vector followed by gel purification. Thirty five ng of the purified

pMQ300 vector were ligated with the gel-purified insert in a 1:1 molar ratio for 6h at 15°C using the T4 DNA Ligase (Promega). Ten µl of the ligation product were chemically transformed into the competent cells of *E. coli* DH5α and the colonies were selected on LA supplemented with 140 µg ml^{-1} hygromycin. The resulting plasmid pKR706, containing the carnitine oxygenase/reductase gene cluster was transferred to chemically competent *E. coli* S17-1λ*pir*, which was used as the donor for the conjugation with the Δ*cntA::aacC1* and the Δ*cntB::aacC1* mutant, respectively. Complemented mutants containing pKR706 were selected on LA plates supplemented with 280 µg ml^{-1} hygromycin and 30 µg ml^{-1} chloramphenicol.

Growth of *Escherichia coli* SE11 and *Citrobacter freundii* M3 on carnitine

To characterize the growth of these strains on carnitine, the bacteria were inoculated to the defined medium as described above and carnitine was added as the sole carbon and energy source (final concentration, 5 mM), and the cultures were incubated at 37 °C while shaking continuously (200 rpm). Optical density at 600 nm was recorded and 2 ml samples were withdrawn from the culture at 0, 4, 8, 12 and 24 h. The culture supernatant was collected by passing through 0.22 µm filters before quantification for TMA and carnitine by cation-exchange ion chromatography as described above.

Heterologous overexpression of *cntA*/*cntB*, site-directed mutagenesis and characterization of the CntA mutants. The *cntA* and *cntB* genes were amplified from *A. baumannii* and inserted into the expression vector pCOLADuet-1 (Novagen) under the *BamHI*/*HindIII* sites or the vector pET28a (Novagen) under the *NdeI*/*HindIII* sites. Co-expression of *cntA*/*cntB* was achieved by insertion into pCOLADuet-1 under the *BamHI*/*HindIII* and the *NdeI*/*KpnI* sites, respectively. The CntA mutants (E205D, E205A)

were chemically synthesized by GenScript and inserted into the expression vector pET28a under the *NdeI*/*HindIII* sites. The resulting plasmids were then transformed into the expression host *E. coli* BLR(DE3) pLysS (Merck Biosciences). For protein overexpression, *E. coli* cells were grown at 37 °C to an OD_{600} of 0.5, and isopropyl β-D-1-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.2 mM and the induction of protein expression was then carried out at 18 °C before harvesting. Proteins were purified by His-tag affinity purification from recombinant *E. coli* after induction with IPTG according to the manufacturer's protocol (Merck). The activity of the reconstituted enzyme was assayed at room temperature (~ 22°C) by quantifying NADH oxidation and carnitine-dependent TMA production. A 1-ml enzyme assay mixture contained 10 mM HEPES buffer (pH 7.6), 60 μg purified CntA and CntB, respectively, 0.25 mM carnitine and 0.25 mM NADH. To determine TMA formation from crude extract of recombinant *E. coli* harbouring *cntA, cntB* or *cntAB*, 500 μg of cell-free crude extracts were used. Coupling efficiency was determined as the ratio of the total amount of TMA formed to the amount of NADH consumed.

Substrate profile of CntAB

To test the substrate specificity of carnitine oxygenase (CntAB), the enzyme was purified from recombinant *E. coli*. Protein concentrations were quantified by the Bradford method using the Quick Start Bradford Protein Assay Kit (Bio-Rad). The following substrates were purchased from Sigma-Aldrich and tested, including D, L-carnitine, γ-butyrobetaine, glycine betaine, choline chloride, trimethylamine hydrochloride and trimethylamine *N*-oxide. Enzyme activity was assayed at room temperature (~ 22°C) by quantifying NADH oxidation. A 1-ml enzyme assay mixture contained 10 mM HEPES buffer (pH 7.6), 60 μg purified CntA and CntB, respectively, 0.25 mM substrate and 0.25 mM NADH.

Validation of TMA production from carnitine by gas chromatography- mass spectrometry (GC-MS)

To further confirm the identity of TMA production from carnitine oxidation by CntAB, the enzyme was purified and the activity was reconstituted *in vitro* by adding CntA and CntB to the enzyme assay as described above. Controls were set up by using CntA or CntB only. The reaction was initiated by adding NADH. The enzyme assay was then incubated at room temperature (~ 22°C) for 30 minutes. The reaction was terminated by adding 0.8 g KOH, and TMA produced from carnitine oxidation was subsequently extracted by toluene as described previously (daCosta *et al*., 1990). TMA identification was carried out using an Agilent 6890/5973 GC-MS platform equipped with an automatic liquid sampler. An aliquot (1 μl) of toluene layer was injected into GC-MS. GC conditions were as follows: column, Agilent HP-5ms capillary column (30 m × 0.25 mm i.d.; film thickness, 0.25 um); column temperature, 40°C for 2.5 min, then the temperature was increased at maximum rate up to 230°C, followed by 2 min at 230°C; carrier gas, helium; flow rate, 0.2 ml min^{-1}; split ratio, 10:1.

Characterization of carnitine oxygenase (CntA) and the CntA mutants (E205A; E205D) by circular dichroism (CD) and native polyacrylamide gel electrophoresis (PAGE)

Purified recombinant CntA and the mutants were dialyzed against 200 mM sodium phosphate (pH 7.0) containing 200 mM NaCl. Circular dichroism (CD) spectra were recorded in the range of 195–260 nm using a Jasco J-815 spectrometer (Jasco, UK) using a quartz cuvette of 1 mm path length at room temperature (~ 22°C). Spectra were collected 8 times per sample. Data were expressed as mean residue ellipticity (MRE) in degrees·cm^2·$dmol^{-1}$. Spectra were deconvoluted using the online programme DICHROWEB (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml) (Whitmore & Wallace, 2004; Whitmore & Wallace, 2008) and the CDSSTR algorithm was used to estimate the percentages of each

secondary structure in CntA and its mutants using the reference protein set 7 (Sreerama & Woody, 2000). Native-PAGE was performed using a NuSep® 12% (w/v) Tris-glycine precast polyacrylamide gel (NuSep Ltd.) at a constant voltage (200 V) at 4ºC for 1.5 hr with an Invitrogen electrophoresis system. The gels were stained with the Fast Blue reagent (Expedeon, UK). The NativeMark^{TM} unstained protein standard from Novex was used to estimate the native size of CntA and the mutants.

References

1. van Aartsen JJ, Rajakumar K. 2011. An optimized method for suicide vector-based allelic exchange in *Klebsiella pneumoniae. J Microbiol Methods* 86:313–319.
2. Simon R, Prifer U, Pühler A (1983). A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. *Bio/Technology* 1(9):784-791.
3. Choi KH, Schweizer HP (2005). An improved method for rapid generation of unmarked *Pseudomonas aeruginosa* deletion mutants. *BMC Microbiology* 5(30):30-41.
4. Choi KH, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR, Schweizer HP (2005). A Tn*7*-based broad-range bacterial cloning and expression system. *Nature Methods* 2(6): 443-448.
5. Kalivoda EJ, Horzempa J, Stella NA, Sadaf A, Kowalski RP, Nau GJ, Shanks RMQ (2011) New vector tools with a hygromycin resistance marker for use with opportunistic pathogens. *Mol Biotechnol*. 48:7-14.
6. daCosta KA, Vrbanac JJ, Zeisel SH (1990) The measurement of dimethylamine, trimethylamine, and trimethylamine N-oxide using capillary gas chromatography-mass spectrometry. *Anal Biochem* 187:234-239.

7. Whitmore, L. and Wallace, B.A. (2008) Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases. *Biopolymers* 89: 392-400.
8. Whitmore, L. and Wallace, B.A. (2004) DICHROWEB: an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research* 32: W668-673.
9. Sreerama, N, and Woody, RW (2000) Estimation of protein secondary structure from CD spectra: Comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set. *Anal. Biochem*. 287(2), 252-260.

Supplementary figure legends

Figure S1 A: Neighbour joining phylogenetic tree of CntB and other closely related reductase proteins (~ 320 amino acids). The bar represents 1 substitution per 10 amino acids. **B**: Growth of *Escherichia coli* SE11 and *Citrobacter freundii* M3 on carnitine as the sole carbon source and production of trimethylamine (TMA). Error bars represent standard deviation of measurements from three biological replicates.

Figure S2 A: Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of purified CntA and CntB from recombinant *Escherichia coli*. **B**: Substrate profile of the purified carnitine oxygenase.

Figure S3 Ion chromatography quantification (**A**) and gas chromatography- mass spectrometry (GC-MS) identification (**B**-**E**) of trimethylamine (TMA) production from carnitine using purified recombinant CntA alone, CntB alone or CntA+CntB. Mass spectrum of TMA from carnitine oxidation by CntAB were compared to that of authentic TMA standard (shown on the right panel of **D** and **E**, respectively).

Figure S4 Proposed electron transfer pathway in CntAB and the role of glutamate 205 in CntA. TMA: trimethylamine; NADH: nicotinamide adenine dinucleotide, reduced.

Figure S5 Characterization of CntA and the site-directed mutants by circular dichroism (CD) and native polyacrylamide gel electrophoresis (native-PAGE).

Table S1 CaiT homologues in sequenced Human Microbiome Project (HMP) reference genomes *

Locus Tag	Annotation	Organism
GCWU000182_02944	choline/carnitine/betaine transport	*Abiotrophia defectiva* ATCC 49176
HMPREF0004_3863	conserved hypothetical protein	*Achromobacter piechaudii* ATCC 43553
A60131_010100018802	putative transporter	*Acinetobacter baumannii* 6013113
A6013_010100011600	putative transporter	*Acinetobacter baumannii* 6013150
A6014_010100011000	putative transporter	*Acinetobacter baumannii* 6014059
HMPREF0010_01347	betaine/choline/glycine transporter	*Acinetobacter baumannii* ATCC 19606
HMPREF0014_03271	betaine/choline/glycine transporter	*Acinetobacter* sp. RUH2624
HMPREF0013_01187	conserved hypothetical protein	*Acinetobacter* sp. SH024
HMPREF1705_00723	osmoprotectant transporter, BCCT family	*Anaerobaculum hydrogeniformans* ATCC
HMPREF1013_04206	OpuD protein	*Bacillus* sp. 2_A_57_CT2
HMPREF1013_03192	hypothetical protein	*Bacillus* sp. 2_A_57_CT2
HMPREF1013_00505	glycine betaine transporter	*Bacillus* sp. 2_A_57_CT2
HMPREF1012_02673	OpuD protein	*Bacillus* sp. BT1B_CT2
HMPREF0178_03964	hypothetical protein	*Bilophila* sp. 4_1_30
HMPREF0179_03321	BCCT family transporter	*Bilophila wadsworthia* 3_1_6
HMPREF0179_00456	BCCT family transporter	*Bilophila wadsworthia* 3_1_6
CAMRE0001_1572	L-carnitine/gamma-butyrobetaine antiporter	*Campylobacter rectus* RM3267, CCUG 20446
HMPREF9428_01710	betaine/carnitine/choline transporter (BCCT) family	*Citrobacter freundii* 4_7_47CFAA
HMPREF9428_02735	L-carnitine/gamma-butyrobetaine antiporter	*Citrobacter freundii* 4_7_47CFAA
CSAG_01648	conserved hypothetical protein	*Citrobacter* sp. 30_2
CSAG_03337	L-carnitine/gamma-butyrobetaine antiporter	*Citrobacter* sp. 30_2
CIT292_00938	choline/carnitine/betaine transport	*Citrobacter youngae* ATCC 29220
CIT292_03024	choline/carnitine/betaine transport	*Citrobacter youngae* ATCC 29220
Cbac1_010100004187	putative transporter	*Clostridiales* sp. 1_7_47FAA

CLOSTASPAR_04953	Choline-glycine betaine transporter	*Clostridium asparagiforme* DSM 15981
HMPREF0240_03154	glycine betaine transporter	*Clostridium* sp. D5
HMPREF0240_03223	putative osmoprotectant transporter, BCCT family	*Clostridium* sp. D5
HMPREF0240_01437	osmoprotectant transporter, BCCT family	*Clostridium* sp. D5
HMPREF0322_04106	transporter, betaine/carnitine/choline family	*Desulfitobacterium hafniense* DP7 (draft 151
HMPREF0322_00857	transporter, betaine/carnitine/choline family	*Desulfitobacterium hafniense* DP7 (draft 151
HMPREF0322_00862	transporter, betaine/carnitine/choline family	*Desulfitobacterium hafniense* DP7 (draft 151
HMPREF9457_03501	hypothetical protein	*Dorea formicigenerans* 4_6_53AFAA
DORFOR_00040	choline/carnitine/betaine transport	*Dorea formicigenerans* ATCC 27755
EDWATA_03146	L-carnitine/gamma-butyrobetaine antiporter	*Edwardsiella tarda* ATCC 23685
EDWATA_03147	L-carnitine/gamma-butyrobetaine antiporter	*Edwardsiella tarda* ATCC 23685
HMPREF0864_03292	betaine/carnitine/choline transporter	*Enterobacteriaceae bacterium* 9_2_54FAA
HMPREF0358_1653	possible transporter	*Escherichia coli* 83972
HMPREF9345_01358	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 107-1
HMPREF9345_01115	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 107-1
HMPREF9540_03296	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 115-1
HMPREF9540_01454	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 115-1
HMPREF9541_04721	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 116-1
HMPREF9541_02156	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 116-1
HMPREF9346_01791	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 119-7
HMPREF9346_03001	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 119-7
HMPREF9347_04517	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 124-1
HMPREF9347_03685	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 124-1
HMPREF9348_01357	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 145-7
HMPREF9348_03244	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 145-7
HMPREF9543_02303	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 146-1

HMPREF9543_04874	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 146-1
HMPREF9547_02514	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 175-1
HMPREF9547_02765	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 175-1
HMPREF9548_02982	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 182-1
HMPREF9548_02182	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 182-1
HMPREF9549_02457	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 185-1
HMPREF9550_00256	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 187-1
HMPREF9550_01015	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 187-1
HMPREF9551_02675	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 196-1
HMPREF9551_02516	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 196-1
HMPREF9552_04960	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 198-1
HMPREF9553_02651	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 200-1
HMPREF9530_04260	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 21-1
HMPREF9530_03696	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 21-1
HMPREF9531_02296	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 45-1
HMPREF9534_01299	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 69-1
HMPREF9534_05004	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 69-1
HMPREF9535_02410	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 78-1
HMPREF9535_04534	transporter, betaine/carnitine/choline transporter family	*Escherichia* coli MS 78-1
HMPREF9536_03463	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 84-1
HMPREF9536_02737	transporter, betaine/carnitine/choline transporter family	*Escherichia coli* MS 84-1
ECSF_0045	putative carnitine transporter	*Escherichia coli* O150:H5 SE15
ECSE_0041	putative carnitine transporter	*Escherichia coli* SE11
ECSE_1975	putative transport protein	*Escherichia coli* SE11
ESCG_01223	choline/carnitine/betaine transporter	*Escherichia* sp. 1_1_43
ESAG_04067	L-carnitine/gamma-butyrobetaine antiporter	*Escherichia* sp. 3_2_53FAA

E4_010100004119	putative transporter	*Escherichia* sp. 4_1_40B
E4_010100018447	L-carnitine/gamma-butyrobetaine antiporter	*Escherichia* sp. 4_1_40B
EUBVEN_02747	Choline-glycine betaine transporter	*Eubacterium ventriosum* ATCC 27560
HMPREF0402_02949	hypothetical protein	*Fusobacterium* sp. 12_1B
HMPREF0454_04935	L-carnitine/gamma-butyrobetaine antiporter	*Hafnia alvei* ATCC 51873
HOLDEFILI_01730	choline/carnitine/betaine transport	*Holdemania filiformis* VPI J1-31B-1, DSM
HMPREF9333_01454	hypothetical protein	*Johnsonella ignava* ATCC 51276
HMPREF9333_01061	hypothetical protein	*Johnsonella ignava* ATCC 51276
HMPREF0484_3753	conserved hypothetical protein	*Klebsiella pneumoniae rhinoscleromatis* ATCC
HMPREF0485_02497	BCCT family betaine/carnitine transporter	*Klebsiella* sp. 1_1_55
HMPREF1024_01931	betaine/carnitine/choline transporter (BCCT) family	*Klebsiella* sp. 4_1_44FAA
HMPREF9538_04403	transporter, betaine/carnitine/choline family	*Klebsiella* sp. MS 92-3
HMPREF0490_01850	hypothetical protein	*Lachnospiraceae bacterium* 6_1_37FAA
HMPREF0987_01668	hypothetical protein	*Lachnospiraceae bacterium* 9_1_43BFAA
HMPREF0531_2425	BCCT family betaine/carnitine/choline transporter	*Lactobacillus plantarum* ATCC 14917
HMPREF9371_2432	BCCT family osmoprotectant transporter	*Neisseria shayeganii* 871
HMPREF9370_2231	BCCT family osmoprotectant transporter	*Neisseria wadsworthii* 9715
POTG_03408	choline/carnitine/betaine transporter	*Paenibacillus* sp. D14
HMPREF9024_00300	BCCT family betaine/carnitine transporter	*Pediococcus acidilactici* 7_4
HMPREF0623_1089	glycine betaine/carnitine/choline transporter	*Pediococcus acidilactici* DSM 20284
HMPREF0693_2800	BCCT family betaine/carnitine/choline transporter	*Proteus mirabilis* ATCC 29906
PROPEN_01218	choline/carnitine/betaine transport	*Proteus penneri* ATCC 35198
PROVRETT_04514	choline/carnitine/betaine transport	*Providencia rettgeri* DSM 1131
HMPREF9373_0457	BCCT family betaine transporter	*Psychrobacter* sp. 1501
HMPREF9373_1383	BCCT family betaine/carnitine/choline transporter	*Psychrobacter* sp. 1501
HMPREF7215_2461	glycine betaine transporter OpuD	*Pyramidobacter piscolens* W5455

ShiD9_010100014456	L-carnitine/gamma-butyrobetaine antiporter	*Shigella* sp. D9
ShiD9_010100003772	putative transporter	*Shigella* sp. D9
HMPREF9372_1646	BCCT family osmoprotectant transporter	*Sporosarcina newyorkensis* 2681
HMPREF0782_2136	BCCT family osmoprotectant transporter	*Staphylococcus aureus aureus* ATCC 51811
HMPREF0783_0859	BCCT family osmoprotectant transporter	*Staphylococcus aureus aureus* ATCC BAA-39
HMPREF0769_10535	BCCT family osmoprotectant transporter	*Staphylococcus aureus aureus* MN8
HMPREF0774_1421	choline transporter	*Staphylococcus aureus aureus* TCH130
HMPREF0772_10576	BCCT family osmoprotectant transporter	*Staphylococcus aureus aureus* TCH60
HMPREF0773_0708	choline transporter	*Staphylococcus aureus aureus* TCH70
Sauraur_010100002150	BCCT family betaine/carnitine/choline transporter	*Staphylococcus aureus aureus*
HMPREF0786_01385	osmoprotectant transporter, BCCT family	*Staphylococcus caprae* C87
HMPREF0789_1995	BCCT family betaine/carnitine/choline transporter	*Staphylococcus epidermidis* BCM-HMP0060
HMPREF0793_0722	BCCT family betaine/carnitine/choline transporter	*Staphylococcus epidermidis* M23864:W1
HMPREF0794_1239	BCCT family osmoprotectant transporter	*Staphylococcus epidermidis* M23864:W2(grey)
HMPREF0797_0715	transporter, betaine/carnitine/choline transporter (BCCT)	*Staphylococcus epidermidis* SK135
HMPREF0791_2152	BCCT family betaine/carnitine/choline transporter	*Staphylococcus epidermidis* W23144
HMPREF0798_01170	osmoprotectant transporter, BCCT family	*Staphylococcus hominis hominis* C80
STAHO0001_0325	choline-glycine betaine transporter	*Staphylococcus hominis* SK119
HMPREF0790_1708	BCCT family osmoprotectant transporter	*Staphylococcus lugdunensis* M23590
STAWA0001_0151	osmoprotectant transporter, bcct family	*Staphylococcus warneri* L37603, SK66

*, 754 HMP reference genomes were available in the JGI/IMG-HMP database as of February 2013. 122 CaiT homologues were found (E≤-50) in 91 unique genomes.

Table S2 Putative *cntA* gene in sequenced Human Microbiome Project (HMP) reference genomes*

Locus Tag	Annotation	Organism	Sources of isolation
HMPREF0004_3864	conserved hypothetical protein	*Achromobacter piechaudii* ATCC 43553	Nose wound, *Homo sapiens*, France
A60131_010100018812	Rieske [2Fe-2S] domain protein	*Acinetobacter baumannii* 6013113	Human skin
A6013_010100011610	Rieske [2Fe-2S] domain protein	*Acinetobacter baumannii* 6013150	Human skin
A6014_010100011010	Rieske [2Fe-2S] domain protein	*Acinetobacter baumannii* 6014059	Human skin
HMPREF0010_01349	dioxygenase alpha subunit	*Acinetobacter baumannii* ATCC 19606	Human urine
HMPREF0014_03273	dioxygenase alpha subunit	*Acinetobacter* sp. RUH2624	Human skin
HMPREF0013_01185	conserved hypothetical protein	*Acinetobacter* sp. SH024	Human skin
HMPREF9428_01711	putative dioxygenase subunit alpha *yeaW*	*Citrobacter freundii* 4_7_47CFAA	Human sigmoid colon
CSAG_01649	Rieske domain-containing protein	*Citrobacter* sp. 30_2	Human faeces
CIT292_03025	ring-hydroxylating dioxygenases, large terminal subunit	*Citrobacter youngae* ATCC 29220	Human faeces
HMPREF9345_01357	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 107-1	Human gastrointestinal tract
HMPREF9540_01453	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 115-1	Human intestinal tract
HMPREF9541_04722	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 116-1	Human gastrointestinal tract
HMPREF9346_03000	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 119-7	Human gastrointestinal tract
HMPREF9347_04518	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 124-1	Human gastrointestinal tract
HMPREF9348_01358	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 145-7	Human gastrointestinal tract

HMPREF9543_02302	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 146-1	Human gastrointestinal tract
HMPREF9547_02513	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 175-1	Human gastrointestinal tract
HMPREF9548_02983	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 182-1	Human gastrointestinal tract
HMPREF9550_00257	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 187-1	Human gastrointestinal tract
HMPREF9551_02674	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 196-1	Human gastrointestinal tract
HMPREF9552_03168	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 198-1	Human gastrointestinal tract
HMPREF9530_03695	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 21-1	Human gastrointestinal tract
HMPREF9534_05005	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 69-1	Human gastrointestinal tract
HMPREF9535_02411	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 78-1	Human gastrointestinal tract
HMPREF9536_02738	rieske [2Fe-2S] domain protein	*Escherichia coli* MS 84-1	Human gastrointestinal tract
ECSE_1976	putative dioxygenase alpha subunit	*Escherichia coli* SE11	Human gastrointestinal tract
E4_010100004114	predicted 2Fe-2S containing protein	*Escherichia* sp. 4_1_40B	Human gastrointestinal tract
HMPREF0484_3754	conserved hypothetical protein	*Klebsiella pneumoniae rhinoscleromatis* ATCC 13884	Human airways sample
HMPREF0485_02496	dioxygenase subunit alpha *yeaW*	*Klebsiella* sp. 1_1_55	Human gastrointestinal tract
HMPREF1024_01930	hypothetical protein	*Klebsiella* sp. 4_1_44FAA	Human gastrointestinal tract
HMPREF9538_04404	rieske [2Fe-2S] domain protein	*Klebsiella* sp. MS 92-3	Human gastrointestinal tract
PROVRETT_04639	ring-hydroxylating dioxygenases, large terminal subunit	*Providencia rettgeri* DSM 1131	Human faeces
PstuA_020100020983	dioxygenase, alpha subunit	*Providencia stuartii* ATCC 25827	Human gastrointestinal tract
SBO_1286	hypothetical protein	*Shigella boydii* Sb227	Epidemic in China in 1950s
SDY_1706	hypothetical protein	*Shigella dysenteriae* Sd197	Epidemic in China in 1950s

SSON_1359	hypothetical protein	*Shigella sonnei* Ss046	Epidemic in China in 1950s
ShiD9_010100003767	predicted 2Fe-2S containing protein	*Shigella* sp. D9	Human gastrointestinal tract
HMPREF9372_1747	Rieske (2Fe-2S) domain protein	*Sporosarcina newyorkensis* 2681	Human blood

*, 754 HMP reference genomes were available in the JGI/IMG-HMP database as of February 2013. 39 CntA homologues (E≤-50, query sequence: CntA of *A. baumannii* ATCC19606) were found in 39 unique genomes.

Table S3 Rieske-type proteins used in CntA phylogenetic analyses

GenBank or PDB accession number	Gene	Enzyme	Organism
M64747	*xylX*	Toluate 1,2-dioxygenase	*Pseudomonas putida*
AF071556	*antA*	Anthranilate dioxygenase	*Acinetobacter* sp. ADP1
AF119621	*ditA*	Abietane diterpenoids oxygenase	*Pseudomonas abietaniphila*
AJ223220	*dxnA*	Dioxin dioxygenase	*Sphingomonas* sp. RW1
2B1X	*nahAc*	Naphthalene 1,2-dioxygenase	*Rhodococcus* sp. NCIMB12038
2XR8	*bhpAE*	Biphenyl dioxygenase	*Burkholderia xenovorans* Lb400
1WQL	*cumA*	Cumene dioxygenase	*Pseudomonas fluorescens* Ip01
3EN1	*todC1*	Toluene 2,3-dioxygenase	*Pseudomonas putida*
2GBW	*bphA*	Biphenyl 2,3-dioxygenase	*Sphingomonas yanoikuyae* B1
2CKF	*pahAc*	PAH dioxygenase	*Sphingomonas* sp. Chy-1
2BMO	*nbzAC*	Nitrobenzene dioxygenase	*Comamonas* sp. JS765
1O7W	*nahAc*	Naphthalene 1,2-dioxygenase	*Pseudomonas putida*
2ZYL	*kshA*	3-ketosteroid-9-alpha-hydroxylase	*Mycobacterium tuberculosis*
1Z02	*oxoO*	2-oxoquinoline 8-monooxygenase	*Pseudomonas putida*
1WW9	*carA*	Carbazole 1, 9 a-dioxygenase	*Janthinobacterium* sp. J3
3GKE	*dmo*	Dicamba monooxygenase	*Stenotrophomonas maltophilia*
Q44256	*cbaA*	3-chlorobenzoate-3,4-dioxygenase	*Comamonas testosteroni*
3VCP	*stc2*	Dimethylproline demethylase	*Sinorhizobium meliloti* RM2011
AAG08795	*gbcA*	Glycine betaine demethylase	*Pseudomonas aeruginosa*
U85780	*cmo*	Choline monooxygenase	*Spinacia oleracea*
AB303389	*cmo*	Choline monooxygenase	*Amaranthus tricolor*
CAE17671	*cmo*	Choline monooxygenase	*Oryza sativa*

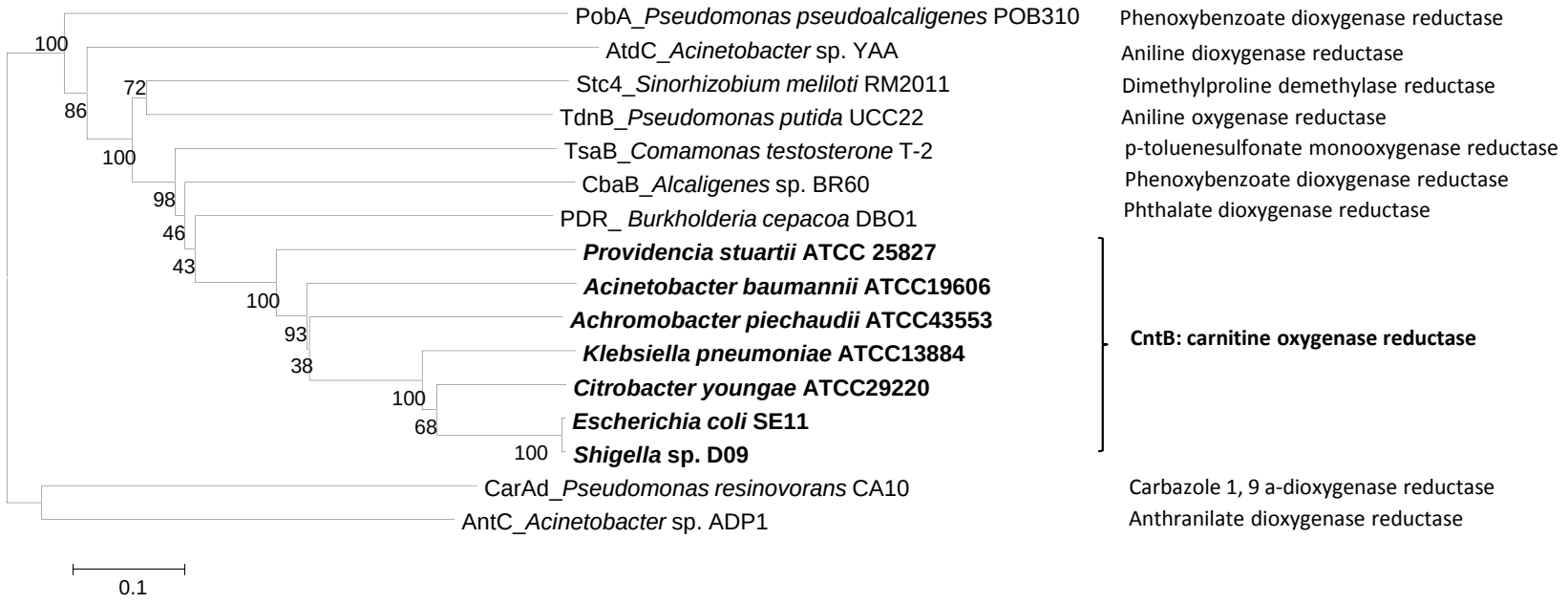
Table S4 Bacterial strains, plasmids and primers used in this study

Strains and plasmids	Description, genotype or relevant characteristics	Reference of source
Acinetobacter baumannii		
ATCC 19606		University of Leicester lab collection
Δ*cntA*::*aacC1*	ATCC 19606 derivative with *cntA* deletion	This study
Δ*cntB*::*aacC1*	ATCC 19606 derivative with *cntB* deletion	This study
Escherichia coli		
JM109	General cloning	Promega
BLR(DE3)pLysS	Heterologous expression of *cntA/cntB* under the T7 promoter	Novagen
CC118*λpir*	Δ(*ara-leu*) *araD* _*lacX74 galE galK phoA thi-1 rpsE rpoB argE*(Am) *recA1*; lysogenized with λ*pir* phage	Simon *et al.*, 1983
S17-1*λpir*	*hsdR recA pro* RP4-2 (Tc::Mu; Km::Tn*7*)(*λpir*)	Simon *et al.*, 1983
Plasmids		
pET28a	For overexpression of *cntA* and its mutants (E205D, E205A)	Novagen
pCOLADuet-1	For co-expression of *cntA* and *cntB*	Novagen
pJTOOL-3	Suicide vector for *Acinetobacter baumannii*, R6K *ori mob*RP4 *sacB*; Cm^R	van Aartsen and Rajakumar, 2011
pUC18R6K-mini-Tn*7*T-Gm	Source of gentamicin cassette	Choi *et al.*, 2005
pKR609	pJTOOL-3::Δ*cntA*-GM 5	This study
pKR612	pJTOOL-3::Δ*cntB*-GM 22	This study
pMQ300	For complementation of Δ*cntA* and Δ*cntB* mutants; Hph^R; pBBR1 *ori*, *oriT*, URA3 marker	Kalivoda *et al*., 2011
pKR706	The 7.8 kb carnitine oxygenase gene cluster cloned into the *Bam*HI/*Kpn*I sites of pMQ300	This study
Primer	**Sequence (5′-3′)**	**Note**
cntAF	GGATCCGATGAGTGCAGTTGAAAAATTACCTGAAGA	Overexpression of CntA
cntAR	ATGATGAAAGCTTTTATTGATGGTACTGCGCCACAAGAT	
cntBF	GGATCCGATGGCGAGTCATTATGAAATGTT	Overexpression of CntB
cntBR	AAGCTTTTAAAGATCTAATATCAATTTTTTGCC	
cntBF_Co	CATATGGCGAGTCATTATGAAATGTT	Co-expression of CntA and CntB in pCOLADuet-1
cntBR_Co	GGTACCTTAAAGATCTAATATCAATTTTTTGCC	
PR351	CGAATTAGCTTCAAAAGCGCTCTGA	Amplification of the gentamicin cassette (*aacC1*)
PR2773	AATTGGGGATCTTGAAGTTCCT	

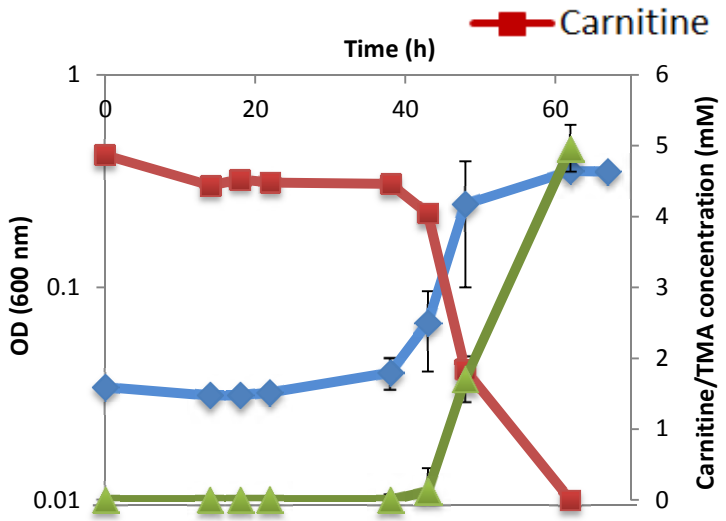
PR2768	GCGCGGCCGCGCAGCATTTATTCACCAGCA	Amplification of *cntA* upstream flanking region
PR2769	TCAGAGCGCTTTTGAAGCTAATTCGTTCAACTGCACTCATGTTTTTACTCC	
PR2770	AGGAACTTCAAGATCCCCAATTGCGCAGTACCATCAATAAAAAC	Amplification of *cntA* downstream flanking region
PR2771	GCGCGGCCGCACAGGTCTGGCCTGCATTAC	
PR2774	GCGCGGCCGCGCTTGCAGAAAAGGCAGGTAA	Amplification of *cntB* upstream flanking region
PR2775	TCAGAGCGCTTTTGAAGCTAATTCGATGACTCGCCATATTAGAATC	
PR2776	AGGAACTTCAAGATCCCCAATTTTGGTATTAGATCTTTAATATA	Amplification of *cntB* downstream flanking region
PR2777	GCGCGGCCGCCGAAATGTGGCTTTCCAAGT	
PR2767	TGCAAGCTCAATCAAACTGG	Detection of cross-over at *cntA* upstream flanking region
PR319	CTCCGAACTCACGACCGA	
PR2772	GCCAATTTGTACGCCATCTT	Detection of cross-over at *cntA* downstream flanking region
PR318	GACATAAGCCTGTTCGGTT	
PR2778	CGAAGAAGGCAAGCGTATTT	Detection of cross-over at *cntB* upstream flanking region
PR319	CTCCGAACTCACGACCGA	
PR2764	TTTCCGGCAATTTATTCAGC	Detection of cross-over at *cntB* downstream flanking region
PR318	GACATAAGCCTGTTCGGTT	
PR3103	TGCCATCTTCTGTCAGGCTA	Amplification of the 7.8 kb carnitine oxygene gene cluster from ATCC 19606
PR3104	TATGCCCACCAAATGAACAA	
PR3120	CGGCGGTACCTGTAAAACGACGGCCAGT	Amplification of the carnitine oxygenase gene cluster from pGEM-Teasy
PR3121	CGGCGGATCCATTTAGGTGACACTATAGAAT	

A)

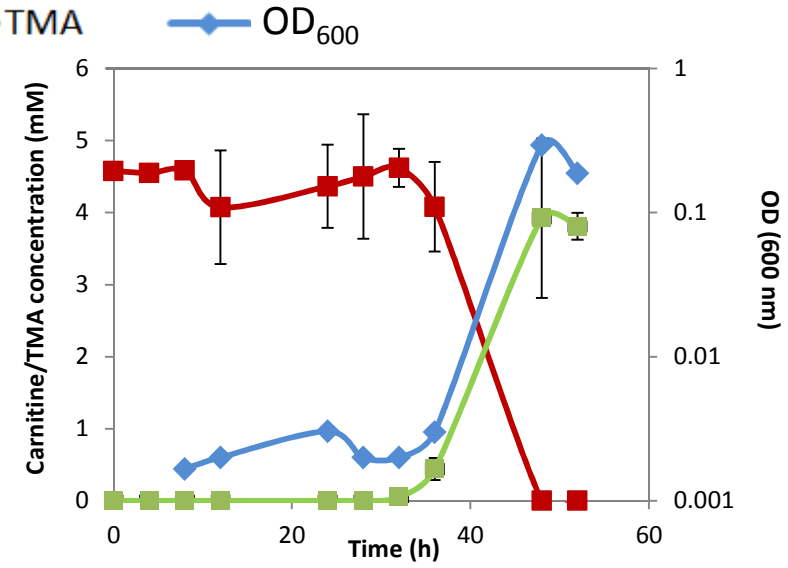
Figure S1



B)

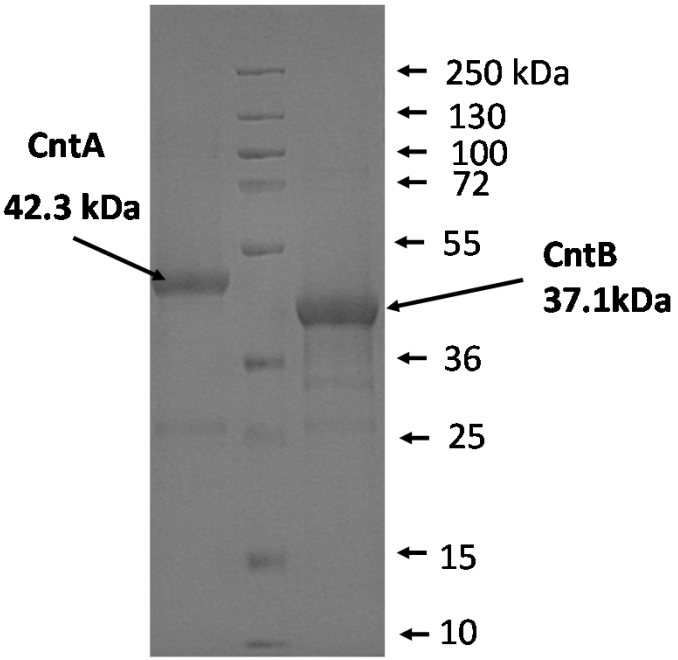


Escherichia coli SE11

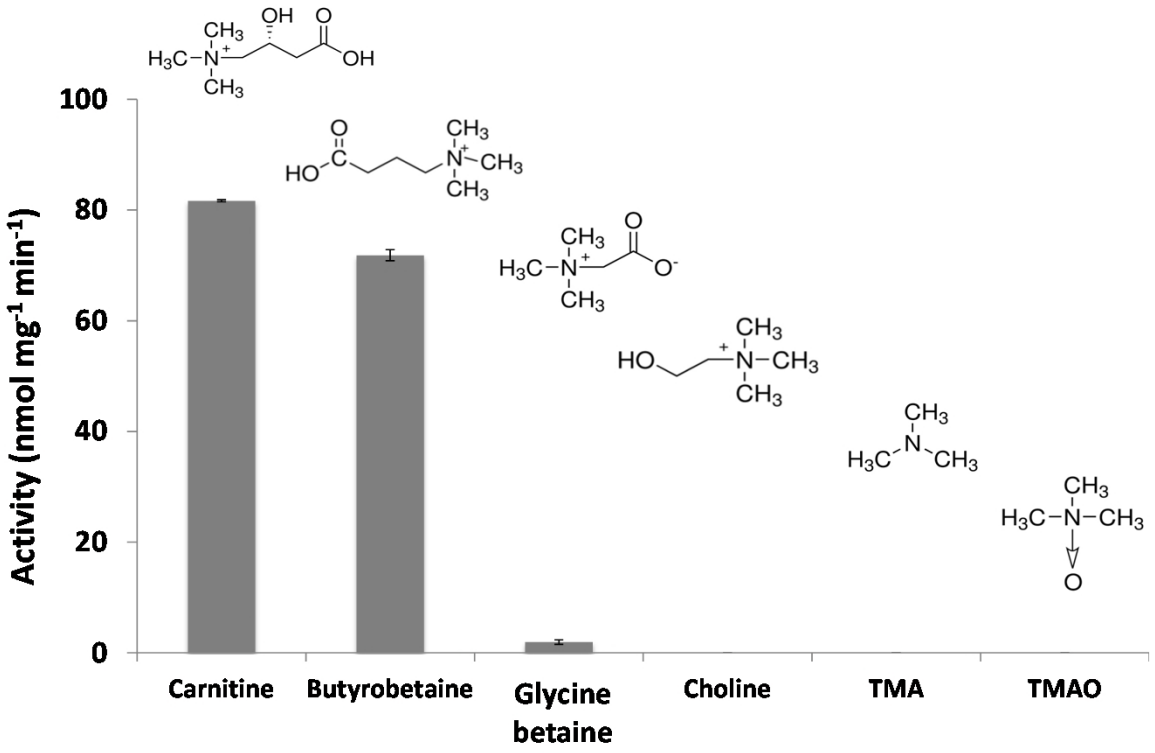


Citrobacter freundii M3

Figure S2


CntA
42.3 kDa
CntB
37.1kDa
250 kDa
130
100
72
55
36
25
15
10

(A)


Activity (nmol mg-1 min-1)
100
80
60
40
20
0
Carnitine
Butyrobetaine
Glycine betaine
Choline
TMA
TMAO

(B)

Figure S3

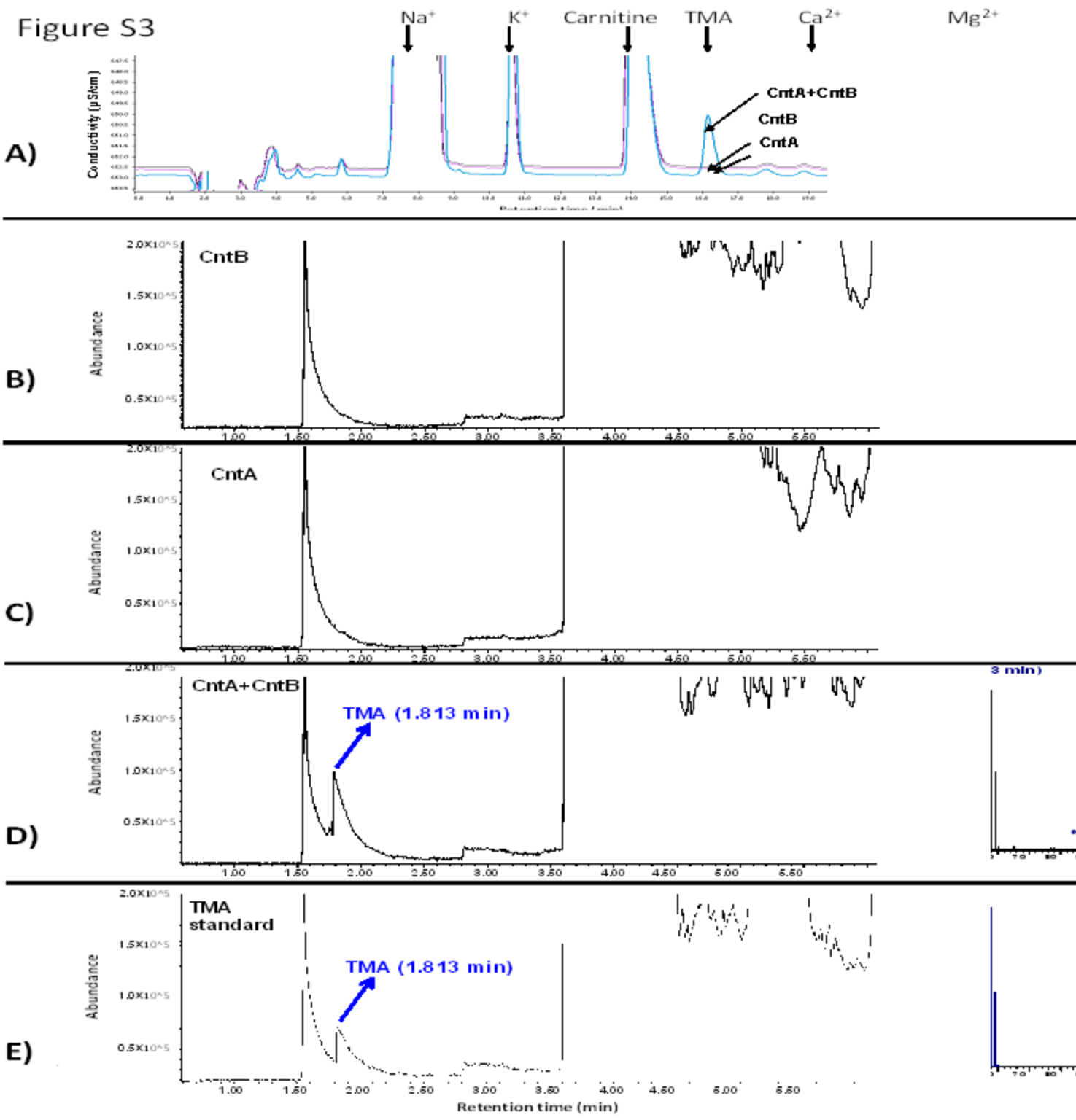

Na+
K+
Carnitine
TMA
Ca2+
Mg2+
A)
Conductivity (µS/cm)
CntA+CntB
CntB
CntA
B)
CntB
Abundance
C)
CntA
D)
CntA+CntB
TMA (1.813 min)
3 min)
E)
TMA standard
TMA (1.813 min)
Retention time (min)

Figure S4

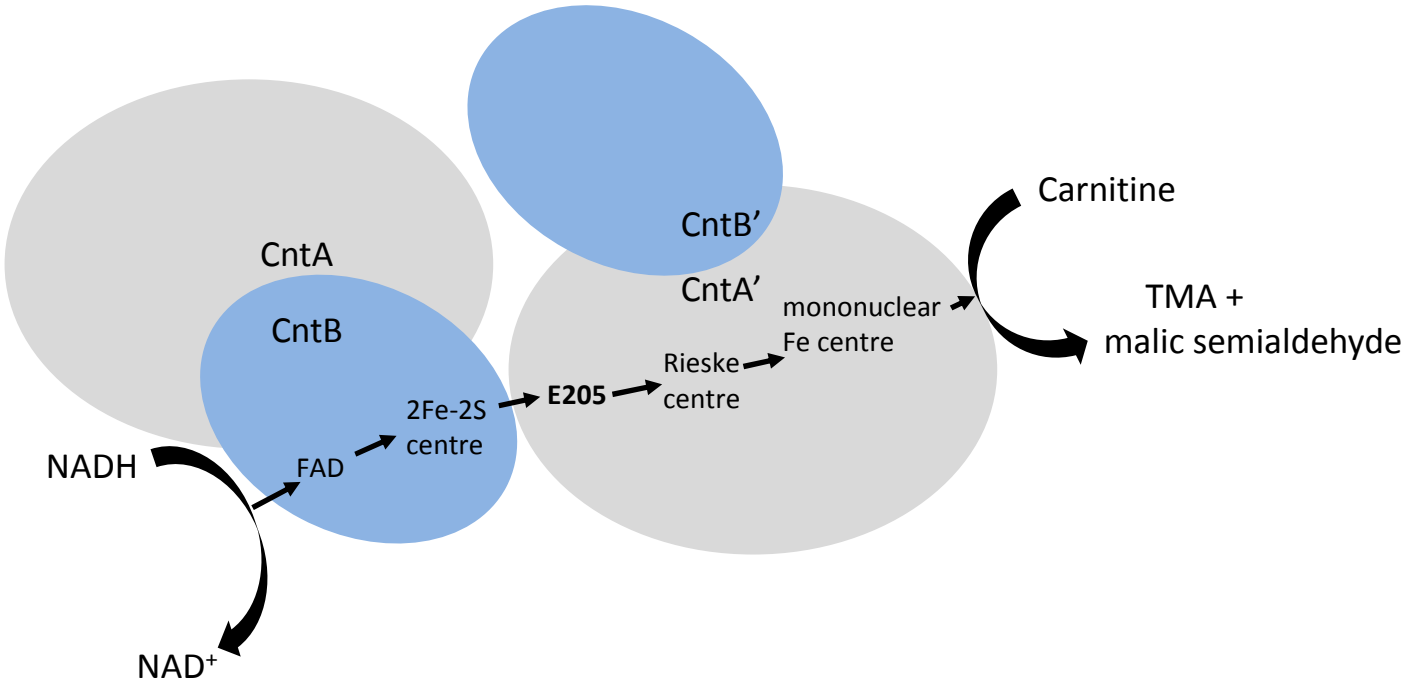
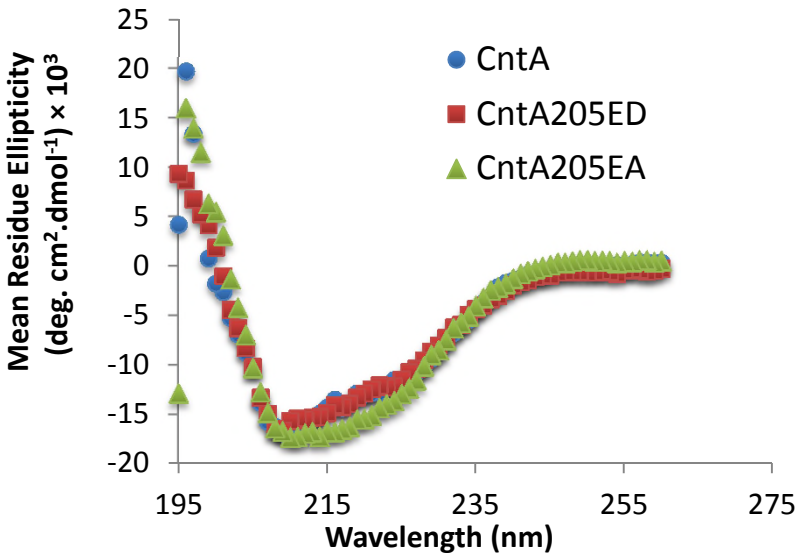

CntB'
CntA
CntA'
CntB
mononuclear
Fe centre
Carnitine
TMA +
malic semialdehyde
Rieske
centre
E205
2Fe-2S
centre
FAD
NADH
NAD+

Figure S5

A)



Result	Helix1	Helix2	Strand1	Strand2	Turns	Unordered	Total
CntA	0.36	0.06	0.18	0.12	0.07	0.21	1
CntA E205A	0.36	0.06	0.17	0.12	0.07	0.21	0.99
CntA E205D	0.38	0.07	0.17	0.11	0.06	0.2	0.99

B)