Two distinct regions in *Staphylococcus aureus* GatCAB guarantee accurate tRNA recognition

Akiyoshi Nakamura1, Kelly Sheppard2, Junji Yamane1, Min Yao1,3, Dieter Söll2,4 and Isao Tanaka1,3,*

1Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan, 2Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8114, USA, 3Faculty of Advanced Life Science, Hokkaido University, kita-10, nishi-8, Sapporo, Hokkaido, 060-0810, Japan and 4Department of Chemistry, Yale University, New Haven, CT 06520-8114, USA

Received September 1, 2009; Revised October 6, 2009; Accepted October 10, 2009

ABSTRACT

In many prokaryotes the biosynthesis of the amide aminoacyl-tRNAs, Gln-tRNA\textsubscript{Gln} and Asn-tRNA\textsubscript{Asn}, proceeds by an indirect route in which mischarged Glu-tRNA\textsubscript{Gln} or Asp-tRNA\textsubscript{Asn} is amidated to the correct aminoacyl-tRNA catalyzed by a tRNA-dependent amidotransferase (AdT). Two types of AdTs exist: bacteria, archaea and organelles possess heterotrimeric GatCAB, while heterodimeric GatDE occurs exclusively in archaea. Bacterial GatCAB and GatDE recognize the first base pair of the acceptor stem and the D-loop of their tRNA substrates, while archaeal GatCAB recognizes the tertiary core of the tRNA, but not the first base pair. Here, we present the crystal structure of the full-length *Staphylococcus aureus* GatCAB. Its GatB tail domain possesses a conserved Lys rich motif that is situated close to the variable loop in a GatCAB:tRNA Gln docking model. This motif is also conserved in the tail domain of archaeal GatCAB, suggesting this basic region may recognize the tRNA variable loop to discriminate Asp-tRNA\textsubscript{Asn} from Asp-tRNA\textsubscript{Asp} in archaea. Furthermore, we identified a 3\textsubscript{10} turn in GatB that permits the bacterial GatCAB to distinguish a U1–A72 base pair from a G1–C72 pair; the absence of this element in archaeal GatCAB enables the latter enzyme to recognize aminoacyl-tRNAs with G1–C72 base pairs.

INTRODUCTION

Correct pairing of an amino acid with its cognate tRNA is an essential step to maintain the accuracy of translation. This is usually accomplished by aminoacyl-tRNA synthetases (aaRSs) that catalyze the direct attachment of an amino acid to its cognate tRNA (1). However, glutaminyl-tRNA synthetase (GlnRS) is absent in the majority of bacteria, in all known archaea, and in many organelles (2). These organisms utilize an indirect pathway for Gln-tRNA\textsubscript{Gln} formation where a non-discriminating glutamyl-tRNA synthetase (ND-GluRS) synthesizes mischarged Glu-tRNA\textsubscript{Gln} (3) that is then amidated to the cognate Gln-tRNA\textsubscript{Gln} by glutamyl-tRNA\textsubscript{Gln} amidotransferase (Glu-AdT) (2,4). Similarly, many prokaryotes lacking an asparaginyl-tRNA synthetase (AsnRS) generate Asn-tRNA\textsubscript{Asn} by the combined actions of a non-discriminating aspartyl-tRNA synthetase (ND-AspRS) and an aspartyl-tRNA\textsubscript{Asn} amidotransferase (Asp-AdT) (5,6).

Two types of AdTs exist: the heterotrimeric GatCAB (7) present in bacteria, archaea and organelles (2), and the heterodimeric GatDE found exclusively in archaea (8). GatDE specifically converts Glu-tRNA\textsubscript{Gln} (8), while bacterial GatCAB acts as both a Glu-AdT and an Asp-AdT in vitro (2). The *in vivo* role of bacterial GatCAB is defined by the nature of ND-aaRS (ND-GluRS and/or ND-AspRS) present in the cell (2). In archaea that lack AsnRS, GatCAB is encoded (9). This enzyme (e.g. from *Methanothermobacter thermautotrophicus*) in vitro strongly prefers Asp-tRNA\textsubscript{Asn} over the homologous Glu-tRNA\textsubscript{Gln} (10); thus archaeal GatCAB may act *in vivo* as an Asp-AdT.

AdTs accurately distinguish their mischarged aa-tRNA substrates (Glu-tRNA\textsubscript{Gln} and/or Asp-tRNA\textsubscript{Asn}) from the cognate Glu-tRNA\textsubscript{Glu} and Asp-tRNA\textsubscript{Asp} species. Bacterial GatCAB and GatDE achieve this by recognizing the first base pair of the acceptor stem and the D-loop of their tRNA substrates (11–13). In contrast, archaeal GatCAB does not recognize the first base pair of Asp-tRNA\textsubscript{Asn} (13,14). Instead the *M. thermautotrophicus* GatCAB makes use of the D-loop, the nucleotide in position 49, and to a lesser extent of the length of the variable loop to distinguish Asp-tRNA\textsubscript{Asn} from...
Asp-tRNA\textsuperscript{Asp} (14), while the \textit{Methanosarcina barkeri} GatCAB appears to use primarily the length of the variable loop for the same task (13). Previous work implicated the tail domain of the GatB and GatE in D-loop recognition of their respective aa-tRNA substrates (11,12). These tail domains share homology with the standalone YqeY proteins of unknown function (PFAM id: PF09424) that are present in a diverse array of organisms (11,12,15,16). Previous crystallographic studies did not resolve the YqeY-like domain of AdTs (11,12,17,18) or of \textit{Deinococcus radiodurans} GlnRS (15). Here, we present a full-length domain of AdTs (11,12,17,18) or of GatCAB appears to use primarily the length of the variable loop for the same task (13).

\textit{Escherichia coli} ND-GluRS was over-produced in an \textit{E. coli} B834 strain and purified over a HisTrap HP column (GE Healthcare) as described (11). The sample was then diluted 5-fold with Buffer A [50 mM Tris–HCl pH 7.5, 10 mM MgCl\textsubscript{2}, 1 mM DTT and 10% (v/v) glycerol], and applied to a HiTrap Heparin HP column (GE Healthcare) equilibrated with buffer A. The column was washed with buffer A containing 50 mM NaCl, and proteins were eluted with a linear gradient of 50–500 mM NaCl. The enzyme eluted at ~250 mM NaCl. The enzyme fractions were then loaded onto a HiLoad 26/60 Superdex 200 pg column (GE Healthcare) equilibrated with buffer B [50 mM Tris–HCl pH 7.5, 10 mM MgCl\textsubscript{2}, 1 mM DTT and 10% (v/v) glycerol]. Pooled fractions were concentrated by ultrafiltration using Vivaspin devices (VIVASCIENCE) to a final concentration of 18 mg/ml, and then diluted 2-fold with 100 mM KCl, 1 mM DTT and 10% (v/v) glycerol, and stored at ~30°C.

\textbf{Preparation of \textit{S. aureus} tRNA\textsuperscript{Gln}, and \textit{M. thermautotrophicus} and \textit{C. trachomatis} tRNA\textsuperscript{Asn}}

The tRNA isoacceptors were in vitro transcribed and purified as described (10,11).

\textbf{Preparation of aminoacyl-[\textsuperscript{32}P] labeled tRNA}

The tRNA isoacceptors were [\textsuperscript{32}P]-labeled and aminoacylated as described (10,20) with minor modifications. For glutamylation of \textit{S. aureus} tRNA\textsuperscript{Gln}, 5 \textmu M \textit{S. aureus} ND-GluRS and tRNA\textsuperscript{Gln} were added in the aminoacylation reaction.

\textbf{Crystallization and structure determination}

The high quality single crystals of \textit{S. aureus} GatCAB were obtained by using the micro-seeding technique (21) from hanging drops set up in a 1:1:0.1 ratio from protein, reservoir solution [25% (w/v) PEG 600, 5 mM MgCl\textsubscript{2}, 50 mM HEPES–NaOH pH 7.2 and 3% (v/v) 2-methyl-2,4-pentanediol (MPD)] and a micro-seeds stock solution (21). The crystal of \textit{S. aureus} GatCAB was rapidly soaked through the reservoir containing 50 mM MES-Na, pH 6.4, 25% (w/v) PEG 600, 5 mM MgCl\textsubscript{2}, 3% (v/v) MPD and 10% (v/v) glycerol as a cryoprotectant, and then a data set was collected to 1.9\AA{} resolution at SPring-8 beamline 41XU (Hyogo, Japan) under cryogenic condition (~173°C). The data set was processed and scaled using the HKL2000 package (22). The structure of \textit{S. aureus} GatCAB was solved by molecular replacement using \textit{AMoRe} in the CCP4 suite (23), with the refined model of the previous (11) \textit{S. aureus} GatCAB structure (PDB ID: 2G5H) as a search model. The model of \textit{S. aureus} GatCAB was rebuilt by automatic refinement program \textit{LAFIRE} (24,25) running with \textit{CNS} (26), and modified manually by using \textit{Coot} (27) followed yielded the final model with the crystallographic \textit{R}/\textit{R}\textsubscript{free} factor of 19.5/21.4%. The summary of data statistics is presented in Table 1. All figures were generated by \textit{PyMol} (28).

\textbf{Small-angle X-ray scattering}

For preparation of the tRNA\textsuperscript{Gln}-bound \textit{S. aureus} GatCAB, the \textit{S. aureus} GatCAB was mixed with tRNA\textsuperscript{Gln} in a molar ratio of 1:4, and then purified by a
gel filtration using a Superdex 200 10/300 GL column (GE Healthcare) in buffer B.

Small-angle X-ray scattering (SAXS) measurements were carried out at SPring-8 beamline 40B2 and 45XU (Hyogo, Japan) (29). The following SAXS analysis was carried out by using the data measured at BL45XU. A wavelength of 1.0 Å was used, and the specimen-to-detector distance was 246.5 cm. The scattering vector range was 0.008 Å⁻¹ < q < 0.334 Å⁻¹, with q being the scattering angle. Data were measured at 5.0 mg/ml of samples at 25°C. Both buffers and samples were exposed 100 s. The SAXS data were normalized to the intensity of the incident beam and processed to back-ground subtraction using the standard procedures with the program PRIMUS (30). Maximum particle dimensions Dmax were computed using the indirect transform package GNOM (31), which also gives the distance distribution functions p(r). The discrepancy between the calculated and the experimental scattering curves was minimized using the program CRYOSOL (32) as described previously (33,34). The 16 ab initio models of tRNA-free and -bound S. aureus GatCAB were generated and averaged with the program DAMMIN (35), DAMAVER (36) and SUPCOMB (37) as described previously (34).

[^32]P|tRNA/nuclease P1 amidotransferase assay

AdT activities of wild-type and mutant GatCAB enzymes were monitored using a[^32]P|tRNA/nuclease P1 amidotransferase assay as previously described (38). Determination of kinetic parameters with the S. aureus GatCAB enzymes was carried out as described for the Helicobacter pylori GatCAB (20). Determination of the kinetic parameters with the M. thermotrotophicus GatCAB enzymes was carried out as described (10).

RESULTS

Overall structure of a full-length GatCAB

We crystallized S. aureus GatCAB by adding 3–10% (v/v) MPD to the previous crystallization condition (11), which dramatically improved the crystal quality (maximum resolution from 2.3 to 1.9 Å). Surprisingly, the high-resolution structure clearly shows the GatB C-terminal region encompassing amino acids 412–475 and a histidine-tag (Figure 1A and Supplementary Figure S1), whose electron density map was not visible in the previous crystal structures. Although the unit cell dimension and crystal packing of the full-length and the previous apo-form S. aureus GatCAB are very similar, in the current structure the C-terminal GatB region is sandwiched between two GatA molecules related by 2-fold screw axes. Successful resolution of the tail-domain may be due to MPD addition in the crystallization mixture making the protein domain less flexible.

The GatA apo-form is essentially the same as the previously deposited S. aureus GatA structures (11). Interestingly, the high-resolution structure shows an alternative conformation for R358, which recognizes the carboxyl group of the substrate Gln (Supplementary Figure S2). Furthermore, Y310, which made a hydrogen bond with D425 in the Gln-bound S. aureus GatCAB (PDB ID: 2F2A), is flipped out from the active site of GatA, indicating that Gln binding induces only minor conformational changes in the GatA active site. GatC is also nearly identical to the deposited S. aureus GatC structures except the six C-terminal residues (95–100) were disordered and could not be modeled.

GatB is comprised of three consecutive domains: a cradle domain (1–294), a helical domain (295–407) and a tail-domain (295–475), belongs to the same protein family.
as standalone YqeY polypeptides of unknown function present in many bacteria and in yeast (PFAM ID: PF09424) (15,16). The YqeY-like tail domain appended to the *D. radiodurans* GlnRS enables the enzyme to productively bind to tRNA\textsubscript{Gln} (15). A similar role is proposed for this structure in GatB and GatE via recognition of the D-loop (11,12). The only YqeY structure so far resolved is from *Bacillus subtilis* (PDB ID: 1NG6); it can now be compared with the YqeY region in the present full-length of *S. aureus* GatCAB structure. The anti-parallel helix bundle of the tail domain superimposes well with the C-terminal domain of *B. subtilis* YqeY protein (YqeY-C: 92-146) with an r.m.s. deviation of 1.97 Å for the 51 C\textsubscript{a} pairs compared (Figure 1C). Interestingly, the GatB tail domain has an additional \(\alpha\)-helix (\(\alpha\)14) and a short loop between \(\alpha\)14 and \(\alpha\)15 (439–442) not found in the YqeY-C. Based on the multiple sequence alignment with the tail domains and YqeY-C, this extended region contains a Lys rich motif (KXXGXX) that is highly conserved in bacterial and archaeal GatB enzymes (Figure 1D and Supplementary Figure S3). The C-terminal extension of *D. radiodurans* GlnRS also contains this extended region with a similar motif (RGGKTA). In contrast to GatB and *D. radiodurans* GlnRS, the tail domain of GatE lacks \(\alpha\)14 and the Lys rich motif. Interestingly, the Lys rich motif of GatB is instead replaced in GatE with a GXXAXGX motif that has been implicated in GatDE distinguishing Glu-tRNA\textsuperscript{Gln} from Asp-tRNA\textsuperscript{Asn} (16).

**Docking tRNA\textsuperscript{Gln} into the *S. aureus* GatCAB structure**

The GatB C-terminal tail is essential for GatCAB binding tRNA\textsuperscript{Gln} (11). In the co-crystal structure of the GatDE enzyme (PDB ID: 2D6F), the GatE tail domain was in the vicinity of the tRNA D-loop; however, its detailed structure could not be solved and the D-loop was fitted with the help of the *B. subtilis* YqeY-C structure (12). We decided to create a docking model of full-length *S. aureus* GatCAB with *E. coli* tRNA\textsuperscript{Gln} to understand how the enzyme uses the tail domain to bind the tRNA. The use of *E. coli* tRNA\textsuperscript{Gln} is reasonable as the sequences of the variable and D-loops of the *S. aureus* and *E. coli* tRNA\textsuperscript{Gln} are identical (Figure 2A), and *E. coli* Glu-tRNA\textsuperscript{Gln} serves as substrate for GatCAB \textit{in vivo} (39).

To construct the docking model we first superposed *E. coli* tRNA\textsuperscript{Gln} (PDB ID: 1QTQ) into the GatDE:tRNA\textsuperscript{Gln} co-crystal structure by aligning all atoms of nucleotides U8, A14–G19, A21, C48 and U54–G57 of the tRNAs. The U8–A14–A21 base triple in the augmented D-stem is commonly found in tRNA\textsuperscript{Gln} while nucleotide 15 in the D-loop and nucleotide 48 in the variable loop form the conserved Levitt pair (40,41).
The tertiary interaction between the D- and TΨC-loops of tRNA is also well conserved.

Next, we docked \textit{S. aureus} GatCAB to tRNA\textsuperscript{Gln} by taking advantage of the homology between GatB and GatE \cite{8,16}, and the co-crystal structure of GatDE:tRNAGln \cite{12}, superposing GatB with \textit{M. thermautotrophicus} GatE. The cradle domain of \textit{S. aureus} GatB and \textit{M. thermautotrophicus} GatE superpose well with an r.m.s. deviation of 1.7 Å for 244 pairs of C\textsubscript{a} atoms compared. However, initially there were severe clashes between the helical domain and tRNA\textsuperscript{Gln}. Therefore, we separated the helical and tail domains into three parts (293–363, 364–381 and 382–475) and then superposed into GatE independently. Such movements are predicted based on previous AdT structures \cite{11,12,17}. Finally, the cradle and helical domains of \textit{S. aureus} GatB could be superposed into that of \textit{M. thermautotrophicus} GatE with an r.m.s. deviation of 1.9 Å for 321 pairs of C\textsubscript{a} atoms compared (Supplementary Figure S4A).

The GatB tail domain possesses a highly conserved hydrophobic pocket comprised by residues (V449, M452, G457, A459 and P461) from the α15, α16 and the loop between them (Figure 3A), surrounded by positively charged residues (Figure 3B). These observations suggest that this conserved hydrophobic pocket may recognize the shape of the tRNA\textsuperscript{Gln} D-loop with the surrounding positive residues interacting with the tRNA phosphate backbone (Figure 3C).

Interaction of the tail domain with tRNA is expected, as deletion of the C-terminal portion of GatB gives rise to a GatCAB mutant enzyme unable to bind tRNA \cite{11} and as mentioned the electron density map of the GatDE:tRNAGln structure places the YeqY-like tail domain in proximity of the tRNA D-loop \cite{12}. However, the initial \textit{S. aureus} GatCAB:tRNAGln docking model placed the conserved hydrophobic pocket of the tail domain ~12 Å away from the D-loop of tRNAGln (Supplementary Figure S4B). It is likely that the tail domain can move to interact with tRNAGln due to the flexible loop connecting the tail and helical domains. This domain flexibility may explain why it has

---

**Figure 2.** Comparison with tRNA\textsuperscript{Gln} and tRNA\textsuperscript{Glu}. (A) Secondary structures of tRNAGln tRNAGlu from \textit{S. aureus}, \textit{E. coli} and \textit{T. thermophilus}. The G18:C56 base pair are colored green. The two guanines conserved in the tRNA D-loop are enclosed with black boxes. U8, A14–G19, A21, C48 and U54–G57, which employed on superposing of tRNA\textsuperscript{Gin} and tRNAGlu, are shown as black large characters. U47 of tRNA\textsuperscript{Gln} and U20\textsubscript{b} of tRNAGlu are colored red. (B) Superposition of the \textit{E. coli} tRNAGln and \textit{T. thermophilus} tRNAGlu. Orange and cyan ribbon diagram indicate \textit{E. coli} tRNAGln and \textit{T. thermophilus} tRNAGlu, respectively. Supplemental nucleotides of the variable loop of \textit{E. coli} tRNAGln and the D-loop of \textit{T. thermophilus} tRNAGlu are shown as thick stick models. The U8–A14–A21 base-triple, the conserved ‘Levitt’ pair and tRNA D-loop and TΨC-loop associating nucleotides are indicated as thin stick models.
been difficult to resolve the YqeY-like tail of previous AdT structures (11,12,17,18). Given the above, we manually fitted the pocket of the tail domain into the D-loop of the E. coli tRNAGln (Figure 3C). The distance between the N-terminal end of the tail domain (Q411) and the C-terminal end of the helical domain (N407) is 10 Å, a distance the inter-domain loop connecting the two domains can bridge (Figure 3C). Recognition of the D-loop is consistent with the fact bacterial GatCAB uses that tRNA element to distinguish transamidation substrates (Glu-tRNA Gln and Asp-tRNA Asn) from Glu-tRNAGlu and Asp-tRNA Asp (11,13).

In order to verify our final S. aureus GatB:tRNAGln docking model, we recorded small-angle X-ray diffusion spectra on the tRNA-free and -bound S. aureus GatCAB purified by a size exclusion chromatography (Supplementary Figure S5). The predicted curve calculated from the crystal structure of the full-length S. aureus GatCAB with the program CRYSOL (32) fits closely to the experimental curve of tRNA-free S. aureus GatCAB, as characterized by a discrepancy value $\chi$ of 0.050 (Figure 4A). Furthermore, the crystal structure fits well to the ab initio envelope calculated with the program DAMAVER (36), of 16 dummy atom models calculated with DAMMIN (35), as reflected by a normalized spatial discrepancy (NSD) value of 1.21 (4A). The predicted scattering curve of the tRNA$^{Gln}$ docking model fits well the experimental curve of tRNA-bound S. aureus GatCAB ($\chi$ of 0.098) and the ab initio envelope (NSD of 0.97) (Figure 4A). However, the tRNA$^{Gln}$ docking model before fitting the tail domain shows a significantly higher discrepancy value with the experimental curve of tRNA-bound S. aureus GatCAB and the ab initio envelope ($\chi$ of 0.186 and NSD of 1.08, respectively) than using the tail domain fitting model. These results suggest that our docking model with the tail domain interacting with the tRNA is consistent with the solution structure of the S. aureus GatCAB:tRNA$^{Gln}$ complex.

In our docking model, the hydrophobic pocket of the GatB tail domain nicely accommodates the curve of the tRNA$^{Gln}$ D-loop (Figure 3C). The S. aureus enzyme distinguishes tRNA$^{Gln}$ from tRNA$^{Glu}$ based on the presence of an extra base (U20B) in the D-loop of tRNAGlu (11). To gain a better understanding of how this is accomplished we superposed the structures of E. coli tRNA$^{Gln}$ and T. thermophilus tRNAGlu (PDB ID: 1G59) in a similar manner as described above. The D-loop of T. thermophilus tRNAGlu like that of S. aureus tRNAGln possesses a U20B supernumerary base (Figure 2A). That extra base flips out from the tRNA D-loop and TP'C-loop associating region (Figure 2B). In the docking model of the tail domain with tRNA, the flipped out U20B in the D-loop of tRNAGlu could not be accommodated; the extra base sterically clashes with the surface of the protein in the model (Figure 3C), suggesting that is the mechanism by which S. aureus GatCAB rejects tRNA$^{Glu}$.

In addition to recognition of the D-loop, the GatB tail domain may also bind to the variable loop of the tRNA. In our docking model, the GatB-specific Lys rich motif of
the tail domain is situated in proximity to the variable loop of tRNA\textsubscript{Gln}, in particular U46 which is pushed out from the tertiary core of the tRNA (Figures 2B and 3C).

Consistent with this prediction, replacement of this Lys rich motif with the GXXAXGX motif from GatE results in a mutant \textit{S. aureus} GatCAB enzyme with reduced affinity for tRNA (Supplementary Figure S6).

The GatB cradle domain recognizes the first base pair of the acceptor stem

Bacterial GatCAB enzymes recognize the U1–A72 base pair of the acceptor stem of Glu-tRNA\textsubscript{Gln} and Asp-tRNA\textsubscript{Asn} to discriminate them from Glu-tRNA\textsubscript{Glu} and Asp-tRNA\textsubscript{Asp} (11,13). In contrast, archaeal GatCAB does not use the first base pair of the tRNA acceptor stem to distinguish Asp-tRNA\textsubscript{Asn} from Asp-tRNA\textsubscript{Asp}, recognizing aa-tRNA species with either a U1–A72 or a G1–C72 base pair (10,13,14). However, the process of distinguishing a U1–A72 base pair from a G1–C72 base pair is not known.

In our GatCAB:tRNA\textsubscript{Gln} docking model the tRNA\textsubscript{Gln} acceptor stem U1–A72 base pair is located in the center of the cradle domain in a space constructed by a 3\_10 turn between \(\beta 2–\beta 11\). The 3\_10 turn is adjacent to the U1–A72 base pair (Figure 4B). This 3\_10 turn is constructed by a hydrogen bond between the main-chain carbonyl oxygen of K183 and amide of E186. Furthermore, M184 in the 3\_10 turn forms a hydrophobic core with L167, L189, C191 and F217, suggesting the 3\_10 turn is fixed with a 3\_10 and \(\beta 11\) by a hydrophobic interaction. In the GatDE enzyme, GatE has a short loop in place of a 3\_10 turn at the corresponding region (Figure 4B and C).

In an alignment of GatB and GatE sequences (Figure 4C), the region including the 3\_10 turn (183–186 in \textit{S. aureus} GatB) is conserved in bacterial GatB sequences, and the corresponding region in archaeal GatB is 1–3 residues shorter, suggesting archaeal GatB has a loop in place of the 3\_10 turn like GatE. The loop region of GatE is well conserved and is rich in positive residues, however the corresponding region in archaeal GatB is not well conserved. The docking model and alignment suggest these structural differences between bacterial GatB, archaeal GatB and GatE may enable the differences in first base pair recognition between the enzymes. In particular, they imply that the 3\_10 turn enables bacterial GatB to distinguish a U1–A72 from a G1–C72 base pair.
Table 2. Transamidase activity of the *S. aureus* (Sa) GatCAB and *M. thermautotrophicus* (Mt) GatCAB with different mischarged tRNA substrates

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Sa Glu-tRNA&lt;sub&gt;Gln&lt;/sub&gt;&lt;sup&gt;Glu&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</th>
<th>Sa Glu-tRNA&lt;sub&gt;Gln&lt;/sub&gt;&lt;sup&gt;GC&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</th>
<th>Mt Asp-tRNA&lt;sub&gt;Asn&lt;/sub&gt;&lt;sup&gt;Asn&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</th>
<th>Mt Asp-tRNA&lt;sub&gt;Asn&lt;/sub&gt;&lt;sup&gt;GC&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Transamidase activity (s&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>wt</td>
<td>3&lt;sub&gt;10&lt;/sub&gt;A</td>
<td>wt</td>
<td>3&lt;sub&gt;10&lt;/sub&gt;ins</td>
</tr>
<tr>
<td>Sa Glu-tRNA&lt;sub&gt;Gln&lt;/sub&gt;&lt;sup&gt;Glu&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</td>
<td>0.71±0.03</td>
<td>0.05±0.01</td>
<td>0.0035±0.0003</td>
<td>0.0047±0.0002</td>
</tr>
<tr>
<td>Sa Glu-tRNA&lt;sub&gt;Gln&lt;/sub&gt;&lt;sup&gt;GC&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</td>
<td>0.05±0.01</td>
<td>0.03±0.01</td>
<td>0.0014±0.0002</td>
<td>0.0004±0.0002</td>
</tr>
<tr>
<td>Mt Asp-tRNA&lt;sub&gt;Asn&lt;/sub&gt;&lt;sup&gt;Asn&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</td>
<td>0.05±0.03</td>
<td>0.04±0.01</td>
<td>0.05±0.01</td>
<td>0.01±0.01</td>
</tr>
<tr>
<td>Mt Asp-tRNA&lt;sub&gt;Asn&lt;/sub&gt;&lt;sup&gt;GC&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</td>
<td>0.24±0.01</td>
<td>0.02±0.01</td>
<td>0.04±0.01</td>
<td>0.11±0.01</td>
</tr>
</tbody>
</table>

Measurements were from three separate experiments. Standard deviations are reported. Reactions were carried out at 37°C in the presence of ATP (4 mM), amide donor (4 mM) and aa-tRNA (1.25μM) indicated.

<sup>a</sup>The aa-tRNA substrates tested were the *S. aureus* (Sa) wild-type Glu-tRNA<sub>Gln</sub><sup>Glu</sup><sub>UA</sub> and mutant Glu-tRNA<sub>Gln</sub><sup>GC</sup><sub>UA</sub>, the wild type of the *M. thermautotrophicus* (Mt) Asp-tRNA<sub>Asn</sub><sup>Asn</sup><sub>UA</sub> and the *C. trachomatis* (Ct) Asp-tRNA<sub>Asn</sub><sup>Asn</sup><sub>UA</sub>.

<sup>b</sup>In the reactions, concentrations of the GatCAB indicated [Su wild-type (wt) GatCAB, Su 3<sub>10</sub> turn deleted mutant GatCAB (3<sub>10</sub>Δ), Mt wt GatCAB, or Mt 3<sub>10</sub> turn insertion mutant GatCAB (3<sub>10</sub> ins)] ranged from 20 nM to 1 μM.

Table 3. Kinetic data for the transamidase activity of the *S. aureus* and *M. thermautotrophicus* GatCAB mutants with different mischarged tRNA substrates

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Sa Glu-tRNA&lt;sub&gt;Glun&lt;/sub&gt;&lt;sup&gt;Glu&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</th>
<th>Sa Glu-tRNA&lt;sub&gt;Glun&lt;/sub&gt;&lt;sup&gt;GC&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</th>
<th>Mt Asp-tRNA&lt;sub&gt;Asn&lt;/sub&gt;&lt;sup&gt;Asn&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</th>
<th>Mt Asp-tRNA&lt;sub&gt;Asn&lt;/sub&gt;&lt;sup&gt;GC&lt;/sup&gt;&lt;sub&gt;UA&lt;/sub&gt;</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Km (mM)</td>
<td>k&lt;sub&gt;cat&lt;/sub&gt; (s&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>k&lt;sub&gt;cat&lt;/sub&gt;/K&lt;sub&gt;m&lt;/sub&gt; (s&lt;sup&gt;-1&lt;/sup&gt;/mM)</td>
<td>k&lt;sub&gt;cat&lt;/sub&gt;/K&lt;sub&gt;m&lt;/sub&gt; (s&lt;sup&gt;-1&lt;/sup&gt;/mM)</td>
<td></td>
</tr>
<tr>
<td>Sa wt</td>
<td>1.80±0.22</td>
<td>1.7±0.1</td>
<td>930±120</td>
<td>2.05±0.48</td>
<td>0.1±0.01</td>
</tr>
<tr>
<td>Sa 3&lt;sub&gt;10&lt;/sub&gt;Δ</td>
<td>1.85±0.59</td>
<td>0.1±0.1</td>
<td>54±18</td>
<td>0.61±0.40</td>
<td>0.04±0.01</td>
</tr>
<tr>
<td>Mt wt</td>
<td>0.70±0.26</td>
<td>0.006±0.001</td>
<td>8±3</td>
<td>0.78±0.30</td>
<td>0.10±0.01</td>
</tr>
<tr>
<td>Mt 3&lt;sub&gt;10&lt;/sub&gt;ins</td>
<td>0.97±0.22</td>
<td>0.009±0.001</td>
<td>9±2</td>
<td>2.51±0.51</td>
<td>0.03±0.01</td>
</tr>
</tbody>
</table>

Measurements were from three separate experiments. Standard deviations are reported. Reactions were carried out at 37°C in the presence of excess ATP (4 mM) and Gln (4 mM). L is the relative catalytic efficiency ([k<sub>cat</sub>/K<sub>m</sub> of homologous substrate]/[k<sub>cat</sub>/K<sub>m</sub> of non-homologous substrate]).

The 3<sub>10</sub> turn of bacterial GatB is crucial for tRNA<sub>Gln</sub> acceptor helix U1–A72 base pair recognition

To evaluate the relation between the presence of the 3<sub>10</sub> turn in GatB and the specificity of first base pair recognition, we constructed a mutant *S. aureus* GatCAB enzyme in which the 3<sub>10</sub> turn (K183, M184 and E185) is replaced by the short putative loop from *M. thermautotrophicus* GatB. Based on the multiple sequence alignment (Figure 4C) we also constructed a mutant *M. thermautotrophicus* GatCAB where the 3<sub>10</sub> turn (residues Lys, Met and Gln) is inserted into GatB between residues Gly169 and Glu170. We then tested the transamidase activities (Tables 2 and 3) by the [32P]tRNA/nuclease P1 assay (38) of these mutant GatCAB enzymes with a variety of tRNA substrates.

As expected (10,13,14), wild-type *S. aureus* GatCAB prefers tRNAs with a U1–A72 base pair over those with a G1–C72 base pair, while wild-type *M. thermautotrophicus* GatCAB has no strong preference (Table 2). *S. aureus* Glu-tRNA<sub>Gln</sub> was a poor substrate for wild-type *M. thermautotrophicus* GatCAB (Tables 2 and 3), like *B. subtilis* Glu-tRNA<sub>Gln</sub> for unknown reasons (10). However, the transamidase activity of wild-type *M. thermautotrophicus* GatCAB with wild-type Glu-tRNA<sub>Gln</sub> containing a U1–A72 base pair was approximately the same as with the G1–C72 mutant Glu-tRNA<sub>Gln</sub> (0.0035 and 0.0014 s<sup>-1</sup>, respectively; Table 2); this is in line with previous results archaeal GatCAB (10,13,14).

Deletion of the 3<sub>10</sub> turn from *S. aureus* GatCAB resulted in a mutant enzyme that was 17-fold less efficient than wild-type mostly due to a difference in k<sub>cat</sub> (Table 3), possibly due to the enzyme no longer recognizing the first base pair of the substrate aa-tRNA. However, consistent with our model, the removal of the 3<sub>10</sub> turn from the *S. aureus* GatB resulted in a mutant *S. aureus* GatCAB that no longer strongly preferred tRNA substrates with a U1–A72 base pair to those with a G1–C72 base pair (Tables 2 and 3). For example the mutant *S. aureus* GatCAB could use the mutant *S. aureus* Glu-tRNA<sub>Gln</sub><sup>Glu</sup><sub>UA</sub> as a substrate about as well as wild-type *S. aureus* Glu-tRNA<sub>Gln</sub><sup>Glu</sup><sub>UA</sub> (0.03 and 0.05 s<sup>-1</sup>, respectively, Table 2). In addition, the mutant *S. aureus* GatCAB could use *M. thermautotrophicus* Asp-tRNA<sub>Asn</sub><sup>Asn</sup><sub>UA</sub>, with its G1–C72 base pair, about as well as *Chlamydia trachomatis* Asp-tRNA<sub>Asn</sub><sup>Asn</sup><sub>UA</sub>, which has a U1–A72 base pair (Table 2) and was about as efficient in amidating *M. thermautotrophicus* Asp-tRNA<sub>Asn</sub><sup>Asn</sup><sub>UA</sub> as *S. aureus* Glu-tRNA<sub>Gln</sub><sup>Glu</sup><sub>UA</sub> (Table 3).

Consistent with our model, the insertion of the 3<sub>10</sub> turn into *M. thermautotrophicus* GatCAB gave a mutant enzyme with preference for tRNA substrates with a U1–A72 base pair (Tables 2 and 3). The mutant *M. thermautotrophicus* GatCAB was about as efficient as wild-type *M. thermautotrophicus* GatCAB using *S. aureus* GatCAB.
Glu-tRNA^{Gln}_{UA} as a substrate (Table 3). However, compared to wild-type *M. thermautotrophicus* GatCAB, the mutant enzyme was 10-fold less efficient with the *M. thermautotrophicus* Asp-tRNA^{Asn}_{GC} substrate (Table 3). In addition to the mutant *M. thermautotrophicus* GatCAB, the insertion mutant of *C. trachomatis* Asp-tRNA^{Asn}_{UA} as with *M. thermautotrophicus* Asp-tRNA^{Asn}_{GC}, the insertion GatCAB mutant preferred *C. trachomatis* Asp-tRNA^{Asn}_{UA}, which has a U1–A72 base pair (Table 2). These results indicate that insertion of the 3_10 turn into GatB enables *M. thermautotrophicus* GatCAB to distinguish aa-tRNA with a U1–A72 base pair from ones with a G1–C72 base pair. Taken with the mutant *S. aureus* GatCAB data, these results strongly suggest that the 3_10 turn in *S. aureus* GatCAB is crucial for distinguishing aa-tRNA with a U1–A72 from those with a G1–C72 base pair.

**DISCUSSION**

The bacterial GatCAB enzyme uses the first base pair of the acceptor stem and the tRNA D-loop for precise tRNA discrimination (11,13). Our results suggest that the U1–A72 base pair of tRNA^{Asn} and tRNA^{Gln} is recognized by a 3_10 turn in the cradle domain of GatB, while the tail domain of GatB binds the D-loop.

The transamidosome is a ternary complex of *T. thermophilus* GatCAB, ND-AspRS and tRNA^{Asn} stable during the overall catalytic process (42). Such a structure would enable Asn-tRNA^{Asn} formation without the risk of free mischarged Asp-tRNA^{Asn} being used in protein synthesis. The complex would also protect the Asn-tRNA^{Asn} product from deacylation until it will be bound by EF-Tu and transported to the ribosome (42,43). Complexes between ND-GluRS, tRNA^{Gln}, and either Glu-AdT (GatCAB or GatDE) are also predicted to exist (12,42). Formation of these complexes between ND-aaRS, AdT and tRNA may explain why the AdTs recognize the specific identity elements in their tRNA substrates.

The structural models of the transamidosomes predict the ND-aaRS enzyme binds to the acceptor stem and anticodon loop of the tRNA (tRNA^{Gln} or tRNA^{Asn}) (12,42) (Supplementary Figure S7) in the same fashion as they do in the absence of AdT (44,45). This places the tRNA 3’-end into the synthetase active site to be aminoacylated; initially the first base pair of the tRNA’s acceptor helix is not accessible to the 3_10 turn of bacterial GatB in its role to discriminate tRNA isoacceptors. However, the tRNA’s tertiary core (including the D-loop) is recognized (Supplementary Figure S7). Recognition of the D-loop by the AdT tail domain (11,12) may permit the amidotransferases to distinguish the ND-aaRS complexed with their tRNA transamidation substrates (tRNA^{Gln} or tRNA^{Asn}) from other tRNA isoacceptors (tRNA^{Glu} or tRNA^{Asp}) (12–14). For example, GatCAB recognition of the D-loop would enable the AdT to discriminate ND-AspRS bound to tRNA^{Asn} from the aaRS bound to tRNA^{Asp} despite the U1–A72 base pair being initially inaccessible to GatCAB. The variable loop of the tRNA in the transamidosome models is also accessible to the AdT, in particular the Lys rich motif in the tail domain of GatB, which may explain why archaeal GatCAB uses this element to discriminate tRNA^{Asn} from tRNA^{Asp} (13,14).

The transamidosome models (12,42) predict that after aminoacylation the tRNA’s 3’ CCA terminus flips from active site in the ND-aaRS to the kinase active site of the AdT (42), similar to the tRNA movements seen in certain aaRSs with editing domains (46,47). In the case of the transamidosome, this movement enables the aminoacyl moiety of the mischarged tRNA to be amidated by the AdT in the complex. Once the 3’ end of the acceptor stem flips from the aaRS active site into the transamidase active site of the AdT, the U1–A72 base pair may become accessible to the 3_10 turn in bacterial GatB.

Recognition of the first base pair by bacterial GatCAB and the archaeal GatDE may be a proofreading step to ensure amidation of the mischarged tRNA substrate (Glu-tRNA^{Gln} and/or Asp-tRNA^{Asn}) and not the properly aminoacylated product of the ND-aaRS (Glu-tRNA^{Gln} or Asp-tRNA^{Asp}). Why this proofreading step is not required by the archaeal GatCAB is unclear, but presumably recognition of the tRNA tertiary core may be enough to ensure that GatCAB amidates Asp-tRNA^{Asn} and not Asp-tRNA^{Asp} (13,14).

It was speculated that the ancestor of GatB and GatE recognized the first base pair of its tRNA substrates (16). However, as the 3_10 turn in the cradle domain of the *S. aureus* GatB for recognition of the U1–A72 base pair in tRNA^{Gln} is conserved only in bacterial GatB and not in archaeal GatB and GatE, this may not be the case. Instead, the common ancestor of GatB and GatE may not have recognized the first base pair of its tRNA substrate, with bacterial GatCAB and GatDE independently evolving to recognize the first base pair of tRNA. Thus, bacterial GatCAB and GatDE both recognizing the first base pair may have been a case of convergent and not divergent evolution.

**ACCESSION NUMBERS**

Coordinate and structure factor have been deposited in PDB with accession number 3IP4.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

The authors are grateful to S. Chimnaronk, I. Heinemann, J. Ling and Y. Tanaka for many stimulating discussions, and to T. Tamura (National Institute of Advanced Industrial Science and Technology) for his gift of the *R. erythropolis* expression system. They thank the staff of SPring-8 BL41XU, 40B2 (2009A1119) and 45XU.
beam lines for their help during collection of the X-ray data.

**FUNDING**

A.N. held a Research Fellowship of the Japan Society for the Promotion of Science for Young Scientists. Grants-in-Aid for Scientific Research (19370037 and 21370041 to M.Y.); Targeted Proteins Research Program (to I.T.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan; Institute of General Medical Sciences and the US Department of Energy (to D.S.).

**Conflict of interest statement.** None declared.

**REFERENCES**