Tertiary core rearrangements in a tight binding transfer RNA aptamer

Timothy L. Bullock, Luke D. Sherlin and John J. Perona

Guided by an *in vitro* selection experiment designed to obtain tight binding aptamers of *Escherichia coli* glutamine specific tRNA (tRNA[Gln]) for glutaminyl-tRNA synthetase (GlnRS), we have engineered a tRNA mutant in which the five-nucleotide variable loop sequence 5′-[44]CAUUC[48]-3′ is replaced by 5′-[44]AGGU[48]-3′. This mutant tRNA binds to GlnRS with 30-fold improved affinity compared to the wild type. The 2.7 Å cocrystal structure of the RNA aptamer–GlnRS complex reveals major rearrangements in the central tertiary core of the tRNA, while maintaining an RNA–protein interface identical to the wild type. The repacked RNA core features a novel hydrogen bonding arrangement of the *trans* Levitt pair G15–U48, a new sulfate binding pocket in the major groove, and increased hydrophobic stacking interactions among the bases. These data suggest that enhanced protein binding to a mutant globular RNA can arise from stabilization of RNA tertiary interactions rather than optimization of RNA–protein contacts.

The specific recognition of RNA molecules by proteins is central to gene regulation in living cells. A key to the origins of binding affinity and specificity in these interactions lies in the ability of RNA to adopt a wide variety of irregular secondary and tertiary structures[1,2]. This provides a complex molecular surface for protein recognition, generating opportunities for sequence discrimination well beyond what is available in undistorted A-form duplexes[3]. From the perspective of high resolution structure, the most thoroughly characterized RNA–protein complexes are those involving the interaction of transfer RNA with various components of the cellular translational apparatus, including aminoacyl-tRNA synthetases[4–10], elongation factor Tu[11,12], and Met-tRNA$_f$[Met] transformylase[13]. The availability of these structures, together with the large sequence data base of molecules possessing the common L-shaped fold[14], makes tRNA an excellent model for elucidating general principles underlying the thermodynamics of protein binding to RNA.

The global L-shaped structure of tRNA is stabilized by numerous tertiary interactions, including a large cluster of noncanonical base pairs and base triples mediating the local architecture of the D-stem and adjacent region ('augmented D-stem') bridging to the T-stem and loop (Fig. 1). Many of the interactions include bases of the unpaired variable loop, which enters the D-stem major groove to form a local triplex structure. Crystal structures of the class I glutaminyl- and isoleucyl-tRNA synthetases bound to their respective cognate tRNAs show networks of intimate interactions along the minor groove and sugar-phosphate backbones in this region[4,9]. Thus, differences in local D-stem helical structure among at least a subset of tRNAs, with likely origins in distinct tertiary contacts, may well play an important role in modulating synthetase binding affinity and specificity.

To examine the tRNA tertiary interactions that help determine the binding affinity of a tRNA molecule to its cognate aminoacyl-tRNA synthetase, we have combined equilibrium binding and X-ray crystallographic studies of an *Escherichia coli* glutamine specific tRNA (tRNA[Gln]) variant with mutations in the central ter-

tiary core. Based on analysis of a set of tight binding tRNA aptamer sequences obtained through *in vitro* selection[15], we constructed a hybrid species in which the five-nucleotide CAUUC variable loop of native tRNA[Gln] is replaced with a four-nucleotide loop of sequence AGGU. The crystal structure of the resulting 'var-AGGU' mutant was determined at 2.7 Å resolution bound to GlnRS, and its improved binding affinity for the enzyme was characterized by gel retardation analysis[16]. The combined structural and thermodynamic data suggest a crucial role for RNA tertiary structure in modulating protein affinity, but by a mechanism not involving the generation of a rearranged macromolecular interface region.

Experimental design

A pool of tRNA[Gln] molecules randomized in the D, T and variable loops was selected on the basis of binding to GlnRS, using standard *in vitro* selection techniques[15]. The library comprised randomization of nucleotides at positions 15–21, 47–48 and 53–61, and RNAs capable of binding GlnRS were isolated by filtration through nitrocellulose. Optimization of conditions for the selection step resulted in binding selection under the following conditions: 40 mM sodium acetate (pH 5.8), 10 mM MgCl$_2$, 10 mM KCl and 5 mM dithiothreitol. The library was enriched by six rounds of selection for GlnRS binding; after the sixth round, the pooled RNAs were aminoacylated with glutamine and a final selection for elongation factor Tu (EF-Tu) binding was imposed[15]. Of 106 tRNAs sequenced from the final selected pool, 46 possessed a four-nucleotide replacement of the wild type five-nucleotide variable loop (5′-CAUUC-3′); 33 of these were of the sequence 5′-AGGU-3′. Filter binding assays performed for many selected tRNAs showed that the 5′-AGGU-3′ variable loop sequence consistently correlated with the tightest binding aptamers, with thermodynamic dissociation constants (K_d) for GlnRS reduced by up to 50 fold compared to wild type tRNA[Gln]. Although the selected tRNAs contained an average of 12–15 nucleotide alterations (Fig. 1*c*), sequence analysis of the tightest binding aptamers failed to reveal other

Department of Chemistry and Biochemistry, and Interdepartmental Program in Biochemistry and Molecular Biology, University of California at Santa Barbara, Santa Barbara, California 93106-9510, USA.

Correspondence should be addressed to J.J.P. *email: perona@chem.ucsb.edu*

articles
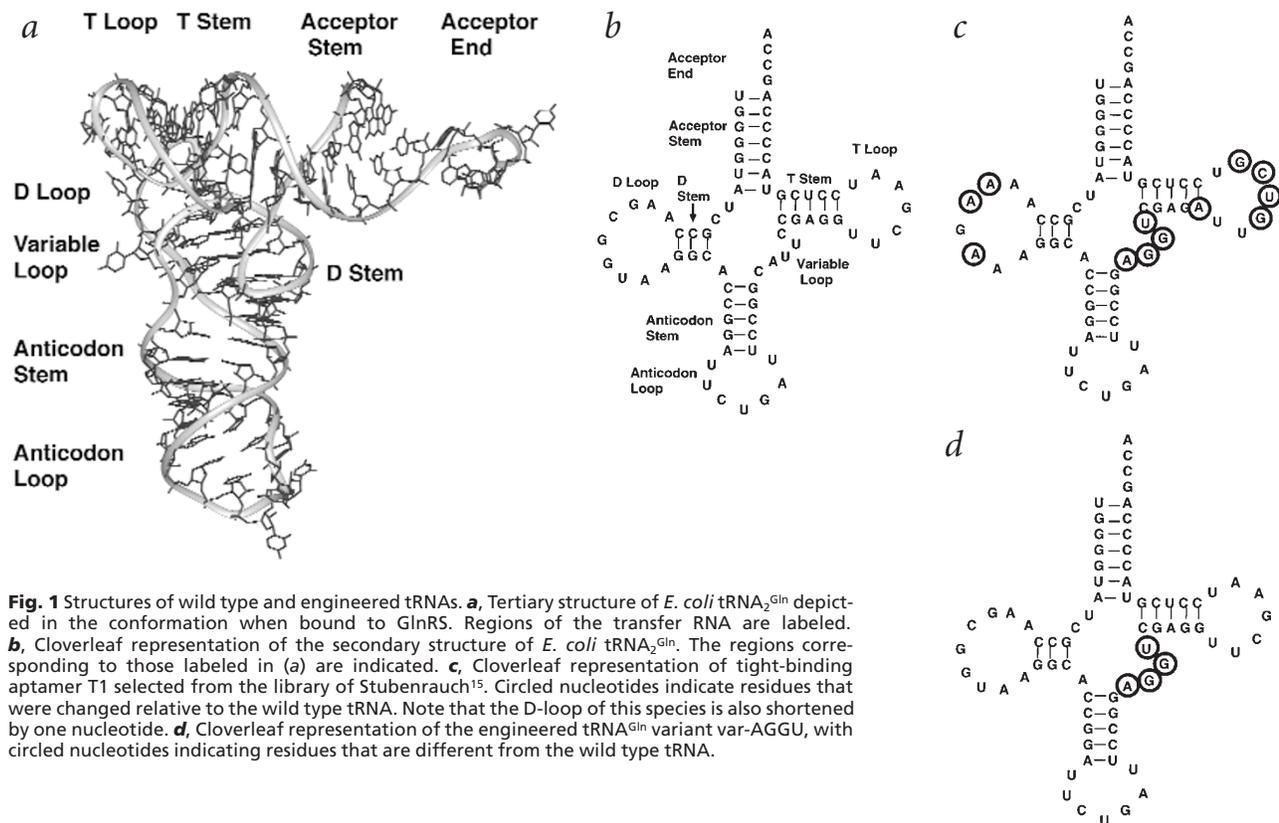
Fig. 1 Structures of wild type and engineered tRNAs. **a**, Tertiary structure of *E. coli* tRNA$_2^{Gln}$ depicted in the conformation when bound to GlnRS. Regions of the transfer RNA are labeled. **b**, Cloverleaf representation of the secondary structure of *E. coli* tRNA$_2^{Gln}$. The regions corresponding to those labeled in (*a*) are indicated. **c**, Cloverleaf representation of tight-binding aptamer T1 selected from the library of Stubenrauch[15]. Circled nucleotides indicate residues that were changed relative to the wild type tRNA. Note that the D-loop of this species is also shortened by one nucleotide. **d**, Cloverleaf representation of the engineered tRNA^{Gln} variant var-AGGU, with circled nucleotides indicating residues that are different from the wild type tRNA.

conserved nucleotides[15]. Therefore, we constructed a tRNA^{Gln} mutant in which the native variable loop was replaced with 5'-AGGU-3' (var-AGGU; Fig. 1*d*). This mutant was then analyzed for GlnRS binding using gel shift analysis and its crystal structure bound to GlnRS was determined by X-ray diffraction at 2.7 Å resolution.

Thermodynamic dissociation constants

The *in vitro* selection experiments showed that the tightest binding tRNA aptamer (aptamer T1; Fig. 1*c*) contains 13 nucleotide substitutions and two nucleotide deletions compared with native tRNA^{Gln} (ref. 15). The var-AGGU mutant possesses a subset of these: four nucleotide substitutions and one deletion, comprising replacement of the variable loop. To establish whether the variable loop replacement alone is sufficient to confer tight binding to GlnRS, we employed a gel shift binding assay to determine the dissociation constants of wild type tRNA^{Gln}, the var-AGGU mutant, and the T1 aptamer (Fig. 2; Table 1). The var-AGGU mutant binds with a K_d of 0.27 nM, ~30 fold tighter than wild type tRNA^{Gln}. This dissociation constant for the var-AGGU mutant is only two fold higher than that measured for the T1 aptamer, demonstrating that the 5'-AGGU-3' variable loop sequence indeed imparts most of the increased binding energy. This demonstrates that the additional mutations in the T1 aptamer contribute little to improved synthetase binding, and suggests that the AGGU variable loop operates largely independently of the surrounding structural context. The K_d values determined by gel shift for both native tRNA^{Gln} and the T1 aptamer match those measured by filter binding within experimental error[15]. Initial rates of glutaminylation for both the aptamer and the var-AGGU mutant are similar to that of the wild type species (data not shown).

Structure of the var-AGGU mutant

Crystals of the ternary complex of GlnRS bound to 5'-O-[N-(L-Gln)-sulfamoyl] adenosine (QSI) and the var-AGGU tRNA grew under conditions similar to those used for the wild type ternary complex with QSI; the crystals from both ternary complexes are in an isomorphous lattice. After cryocooling the cell constants of the crystals differ by 1.2% or less compared with native cryo-cooled crystals (Table 2)[17]. Initial refinement of the data against a model of the wild type structure lacking solvent molecules, QSI and the variable loop was used to generate calculated phases for electron density maps computed with coefficients (F_o^{AGGU} - F_o^{native}) and (F_o - F_c). The (F_o - F_c) map showed clearly connected density for each of the four nucleotides of the newly introduced AGGU loop (Fig. 3), and the (F_o - F_o) map showed positive and negative peaks consistent with the repositioning of nucleotides indicated by the (F_o - F_c) map. Cycles of rebuilding and addition of solvent molecules and iterative positional, B-factor and simulated annealing refinement produced a final model with an R_{free} of 29.3% and very low root mean square (r.m.s.) deviation in bond lengths and bond angles (Table 2).

A comparison of the wild type and mutant cocrystal structures shows that there are no significant differences in GlnRS conformation between the two complexes (r.m.s. deviation for all pro-

Table 1 Thermodynamic binding constants

tRNA[1]	K_d (nM)	Standard deviation (nM)
Wild type tRNA^{Gln}	7.1	1.1
44-AGGU-48 (var-AGGU)	0.27	0.1
T1	0.13	0.04

[1]All tRNAs contain the U1G mutation to increase transcriptional efficiency.
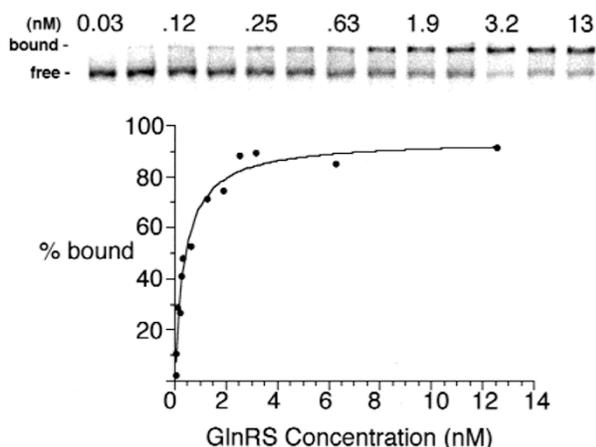
Fig. 2 Electrophoretic gel mobility shift analysis for determination of the dissociation constant (K_d) of the var-AGGU engineered tRNA$_2^{Gln}$ variant. GlnRS protein was titrated over the concentration range shown at the top. The native gel (top) and binding isotherm are shown. The ordinate plots the percentage of total label in the shifted band.

tein backbone atoms is 0.42 Å). Additionally, the entire tRNA–synthetase interface is identical within coordinate error (0.35 Å) to those of native complexes determined with tRNA^{Gln} derived either *in vivo* or *in vitro* (refs 17,18). Specifically, all hydrogen bonding and other electrostatic interactions are conserved (Fig. 4*a*). However, the new four-nucleotide 5′-AGGU-3′ variable loop adopts a distinct and more compact conformation compared with the wild type 5′-CAUUC-3′ loop (r.m.s. deviation for phosphorus atoms of nucleotides 44, 45, 46 and 48 is 3.1 Å, after superposition of protein structures). In the wild type tRNA, nucleotides C44, A45 and C48 stack into the tertiary core region and participate in the base triple A13-A22-A45 as well as the two noncanonical base pairs A26-C44 and G15-C48 (Fig. 4*b*). The bases of U46 and U47 are flipped out from the core and do not contribute to stacking interactions. In contrast, all four bases of the var-AGGU mutant are stacked into the core region (Fig. 4*c*). A44 and U48 occupy similar positions in the stack compared with C44 and C48 of wild type tRNA^{Gln}, but make different hydrogen bonding base pair interactions with A26 and G15, respectively. G46, replacing the two solvent exposed uridines in the wild type structure, inserts into the core by forming a base triple with A13 and A22. Nucleotide G45 is in turn displaced by ~7 Å from the position occupied by A45 in the wild type A13-A22-A45 stack layer, forming instead a new base triple with G10 and C25 (Fig. 4*c*). The base of G45 fills a pre-existing gap in the wild type core (Fig. 5), producing a significant increase in hydrophobic stacking interactions in the mutant structure.

Each of the bases in the new four-nucleotide variable loop makes hydrogen bonding and stacking interactions with surrounding bases in the core, with concomitant effects on the positions of both the base and sugar-phosphate moieties of these adjacent nucleotides. The altered backbone conformation of the shorter variable loop is propagated through the position of the 26-44 pair bridging the D and anticodon stems, to also include much of the anticodon stem. While the positions of bases and the base pairing within the anticodon

stem are preserved, the sugar-phosphate groups on both the 5′ and 3′ sides of this stem are displaced from their positions in the wild type structure, such that the major groove is widened. In the mutant structure, the N1 ring nitrogen of A44 accepts a hydrogen bond from the exocyclic amine of A26; this interaction replaces a similar hydrogen bond with the N3 ring nitrogen of C44 in the native tRNA (Fig. 6*a*). The new interaction is facilitated both by a significant displacement in the position of the A26 base, and by the introduction of significant propeller twist at the noncanonical A26-A44 pair.

The strongest feature in the initial difference electron density map computed with coefficients ($F_o^{AGGU} - F_o^{native}$) occurs in the widened major groove at this noncanonical A44-A26 base pair. The roughly spherical 10 σ difference peak appearing at this position is consistent with a small electron-dense molecule fitting tightly into the major groove, and interacting with the exocyclic amine nitrogens of both A26 and A44 (Fig. 3). The size, electron density, and hydrogen bonding potential of the group are consistent with binding of a sulfate anion, which is present in high concentration in the crystallization mother liquor. The sulfate ion appears to make a third hydrogen bond with the exocyclic nitrogen of C25 in the neighboring base triple (Fig. 5). Recently, a similar sulfate binding site was observed in the RNA major groove at a noncanonical C-C basepair[19].

The insertion of G45 into the D-stem major groove forms a base triple at the G10-C25 pair in the mutant (Fig. 6*b*). The exocyclic N2 amine group of G45 lies within close hydrogen bonding distance of both the N7 (2.6 Å) and O6 (2.9 Å) atoms of G10, but does not affect the G10-C25 Watson-Crick interactions. In the next layer, the base of C9 from the short acceptor/D-stem connector stacks nicely on G45 and is positioned evenly with the C11-G24 base pair. This is in contrast with the wild type structure, in which the C9 base is tilted to bridge the C11-G24 and C12-G23 base pairs, and indeed makes base interactions on the C12-G23 stack level (Fig. 4*b*).

G46 inserts into the major groove at the position of the noncanonical A13-A22 pairing, replacing A45 of the native tRNA in this position. While A45 makes no base interactions with A13-A22 in the wild type tRNA, G46 instead donates a new hydrogen bond to the A13 N1 ring nitrogen, to form a second base triple in
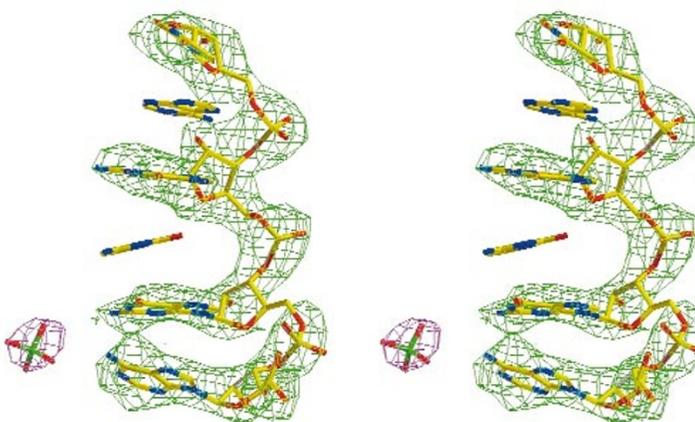
Fig. 3 Stereo view showing difference electron density maps computed in the region of the variable loop of the var-AGGU mutant. Density contours in green are computed from amplitudes ($F_o - F_c$), with phases from the wild type structure omitting the variable loop. The map was computed after scaling of the var-AGGU amplitudes to the native, and rigid body refinement of the model as a single rigid unit. The contour level is 2.5σ. Density contours in red are computed from amplitudes ($F_o^{AGGU} - F_o^{native}$) and are contoured at 6σ.

articles

a
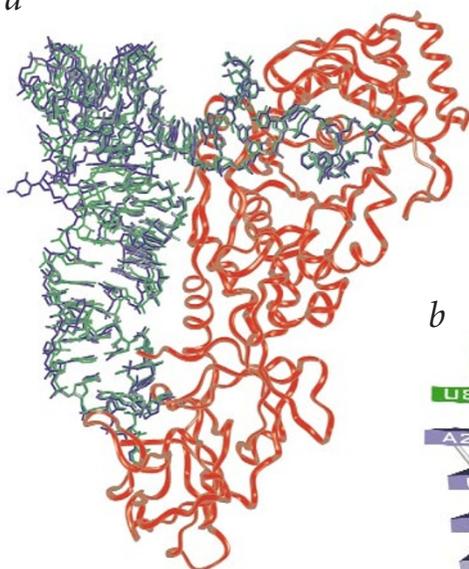
Fig. 4 Structures of wild type and var-AGGU complexes. *a*, Superposition of the structures of the wild type and var-AGGU GlnRS–tRNA[Gln] complexes. The superposition is based on all protein backbone atoms. The ribbon trace of the protein shown (red) is that of the wild type enzyme; no significant differences in protein conformation appear in the mutant complex. The wild type tRNA is shown in blue and the var-AGGU mutant in green. Note that the regions of the tRNA directly adjacent to the protein superimpose well and that distal portions of the molecule diverge. Schematic diagrams of *b*, wild type and *c*, var-AGGU tertiary cores. Nucleotides 8 and 9 (the connecting sequence between the acceptor and D-stems) are in green. D-stem nucleotides and the adjacent A26 and noncanonical A13-A22 pair are shown in magenta. D-loop residues that participate in the central tertiary core are red. Variable loop residues are blue. Dotted gray lines indicate positions where non-Watson-Crick base–base interactions occur.

b
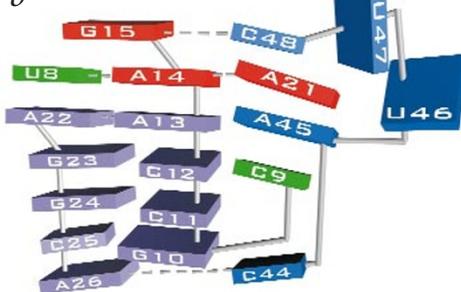
c
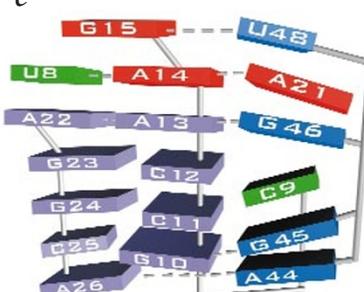
the mutant (Fig. 6*c*). The A13-A22 pair tolerates this substitution with little change in the position of the adenine bases. While there is also little change in the U8-A14-A21 triple, which lies above A13-A22-G46, the *trans* 'Levitt' base pair at positions 15-48 is markedly altered by the C48 to U48 substitution. In the mutant structure, the position of the U48 base deviates only slightly from that of C48, but the conserved G15 is significantly shifted (Fig. 6*d*). The shift creates a *trans* G-U pair featuring two hydrogen bonds between U48 (N3)–G15 (O6) and U48 (O4)–G15 (N1). This is in contrast to the wild type G15-C48 pairing, which instead utilizes the N1 and N2 groups of G15. The relative movement of bases in forming the G15-U48 pair from G15-C48 may be described as a 'trans wobble', and is analogous to the movement required for formation of a more conventional *cis* G-U wobble pair[20]. A *trans* wobble G-U pair was also observed in the structure of the hepatitis δ virus ribozyme[21], although in that case the same pairing was achieved with a *syn* guanosine and antiparallel chain orientations.

Formation of the new G15-U48 pair is accompanied by significant movements in adjacent structural regions. The *trans* wobble shift of the G15 base is accompanied by a backbone displacement, which propagates into the remainder of the D-loop until the loop rejoins the core at nucleotide A21. Additionally, the base of A59 from the T-loop shifts ~2 Å to improve stacking with the 15-48 pair, resulting in some displacement of the remainder of the T-loop and T-stem. However, both the pattern of base stacking interactions (including stacking of the D-loop G18 and G19 bases into the T-arm) and a crystal lattice contact between the G19-C56 pairs of adjacent symmetry-related molecules are conserved. Interestingly, analysis of B-factors between the two structures shows significant differences only in the variable and D-loops (Fig. 7). The var-AGGU mutant variable loop is much better ordered than the native five-nucleotide loop. Much of this difference is due to the replacement of wild type U46/U47 with G46 in the mutant, as these uracils have the highest average B-factors in the wild type tRNA[Gln] structure[4,17]. In contrast, the mobility of a portion of the D-loop

near G15 is higher in the mutant structure than in the native cocrystals. The altered mobilities of the variable and D-loops in the two structures likely reflect the direct coupling between these regions, *via* the 15-48 base pair.

Response of globular RNA to mutation

The crystal structure of the var-AGGU mutant provides the first opportunity to examine in detail the structural effects of mutation in the core region of a large globular RNA. Relative to wild type, the *in vitro* selected variable loop sequence 5′-AGGU-3′ appears to improve intramolecular interactions within the tertiary core according to several criteria: (i) a net increase in base stacking interactions (Fig. 5), (ii) elimination of poorly ordered solvent exposed nucleotides, (iii) optimization of base–base tertiary hydrogen bonding geometry, and (iv) formation of a new anion binding site linking two base pairs in the major groove. The structural compensations within the mutated variable loop are facilitated by the ability of surrounding nucleotides to adapt to the new core. Rearrangements to the sugar-phosphate backbone on the
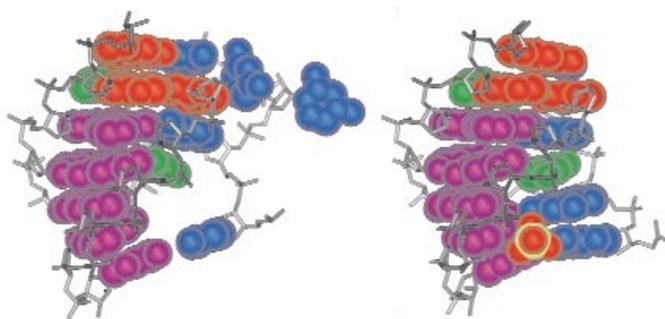
Fig. 5 van der Waals representation of bases in the tertiary core of the wild type (left) and var-AGGU (right) structures. The bound sulfate molecule (oxygen atoms in red; sulfur in yellow) is shown in the major groove of the var-AGGU structure. The region of the tRNA shown spans base pair 26-44 (bottom) to base pair 15-48 (top), as also depicted in Fig. 4*b*,*c*. Color coding of nucleotides is the same as for Fig. 4*b*,*c*.
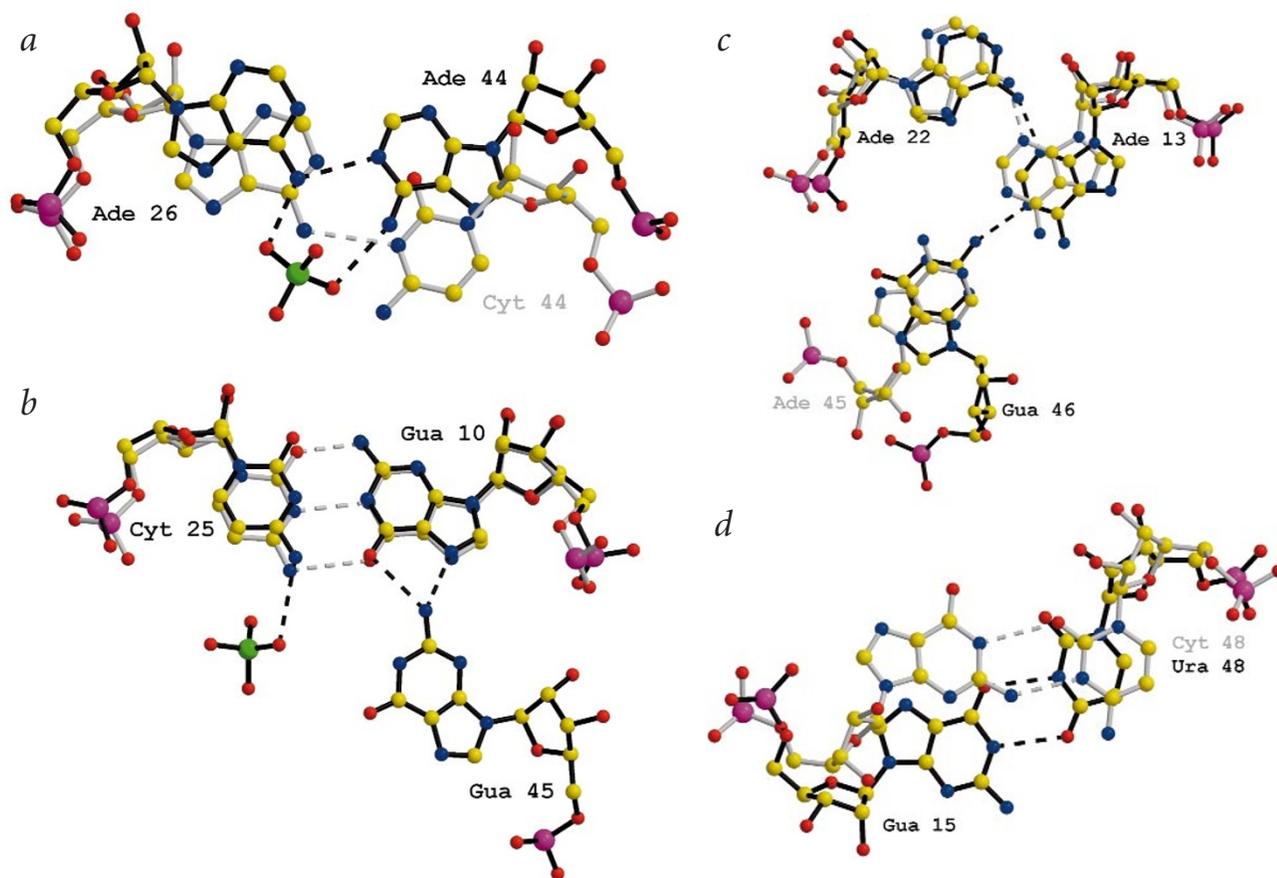
Fig. 6 Overlay of mutant (black bonds) and wild type (gray bonds) tRNA structures in the region of the variable loop. Atom coloring is as follows: carbon, yellow; sulfur, green; oxygen, red; nitrogen, blue. All superpositions are based on protein backbone atoms (not shown). Black dashed lines indicate hydrogen bonds in the mutant; gray dashed lines indicate wild type hydrogen bonds. *a*, 26-44 base pair and the associated sulfate molecule (found only in the var-AGGU structure). *b*, 10-25-45 base triple, showing the same sulfate molecule as in (*a*). There is no G45 equivalent in the wild type tRNA. *c*, 13-22-45/46 base triple. Note that the A45 base makes no hydrogen bonds to A13-A22 in the wild type. *d*, 15-48 Levitt base pair exhibiting a *trans* wobble to accommodate the U48 substitution.

order of 1–3 Å occur on both sides of the core in the D-loop, T-stem and T-loop, as well as anticodon stem regions, encompassing over half of the nucleotides in the tRNA (Fig. 4*a*). It appears that the intrinsic flexibility of the sugar-phosphate backbone permits such large scale propagation of structural effects, and that a low free-energy state of the RNA is thus attained by multiple small readjustments of main chain torsion angles. The ability of surrounding RNA nucleotides to adapt to altered core interactions may provide a general explanation for why some covariations in the core regions of globular RNAs are allowed even when they are not precisely isosteric[20]. Another feature of interest is that these rearrangements occur only in those regions not interacting with synthetase, so that all intermolecular protein–RNA contacts are preserved. A similar conservation of protein–RNA contacts, but in the context of altered RNA secondary rather than tertiary structure, was observed in the crystal structures of bulged-hairpin RNA aptamers bound to bacteriophage MS2 coat protein[22,23].

Several core repacking studies of proteins have emphasized the importance of maintaining van der Waals contacts and tight packing among amino acid side chains in the hydrophobic interiors[24–26]. In these studies of single and multiple amino acid substitutions, the structural changes are modest (on the order of 1 Å or less), with only local repositioning of interior hydrophobic amino acids and little movement of the polypeptide backbone. By

contrast, the structure of the var-AGGU mutant shows both substantial divergence of the sugar-phosphate backbone and repacking of the hydrophobic base side chains, such that the stacking arrangement is fundamentally altered. The structural response in this RNA mutant resembles more closely the effects of insertion or deletion of amino acids from core structural elements of a protein[27]. In both cases long range effects propagated from the site of the mutation occurs, and there are differences in the three-dimensional arrangement of side chains forming the interiors. Additional structural studies of globular RNA mutants will be very informative in providing a more thorough comparison of how RNA and protein structures respond to mutational alteration.

Origins of RNA–protein binding affinity
By providing both binding thermodynamics and high resolution structural information for wild type and mutant protein–RNA complexes, the data presented here offer an opportunity to isolate the contribution to binding energy of a noninterfacial portion of the RNA. Based on the relative dissociation constants, the var-AGGU mutant exhibits an enhanced binding free energy of 1.9 kcal mol[-1] compared with native tRNA. Because the crystal structure shows no differences in direct protein–RNA contacts (although differences in solvent structure undetectable at 2.7 Å resolution may be present), the

articles
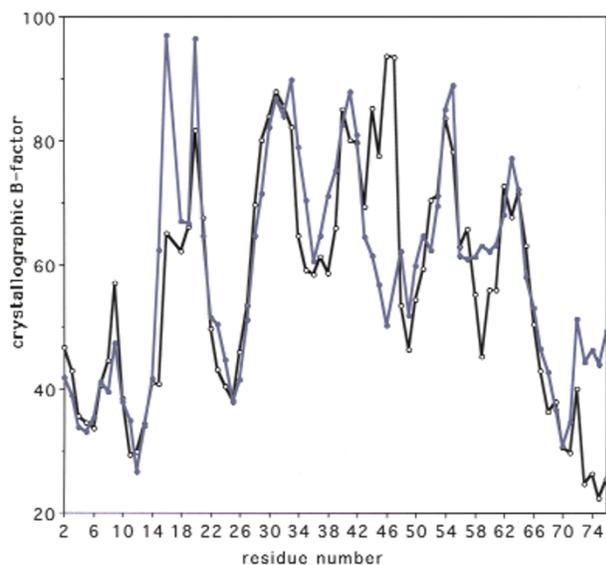
Fig. 7 Plot of the average crystallographic thermal factor (B-factor) per nucleotide, as a function of nucleotide residue along the tRNA sequence. B-factors for all nonhydrogen atoms of each nucleotide are averaged for this plot. The var-AGGU mutant is shown in blue (filled circles) and the wild type in black (open circles). The higher B-factors of the mutant at the 3' acceptor terminus may be due to incomplete inhibitor binding in this structure (see Methods).

tighter binding cannot be attributed in a straightforward way to improved static interactions generating a favorable enthalpic contribution. Instead, these data highlight the role of factors not easily assessed by structure–function approaches. Such factors can include entropic contributions arising from altered flexibility upon binding, release of bound water and ions upon complex formation, and altered conformational equilibria in the unliganded state[28].

Despite the formidable difficulties in partitioning binding free energies among molecular events, experimental structure–function correlations nonetheless provide a valuable counterpoint to theoretical free energy analyses. Although such calculations have not been widely applied to protein–RNA complexes, systematic efforts to analyze the contribution from free energy components are underway in analogous protein–DNA systems[29]. Indeed, it is possible that many of the energy terms considered in predicting the binding free energy for a single native protein–nucleic acid complex can be disregarded in comparative mutant studies (particularly when structures of mutants are available), because these terms may not be altered by the perturbation.

In the case of wild type *versus* var-AGGU tRNA binding to GlnRS, the structural adaptation of the (unaltered) protein, the desolvation and solvation terms associated with formation of the macromolecular interface, and the favorable enthalpic components to complex formation may contribute little to the value of $\Delta\Delta G°_{bind}$[29]. Perhaps the most likely origin of the increased binding affinity resides in the ability of the mutant tRNA to adopt the required binding conformation at a lower free energy cost than the native species. If the tertiary core of the unliganded mutant is better ordered than that of the unliganded wild type species (as observed in the bound complexes), then tighter binding could arise from a decreased entropic penalty for the tRNA upon complex formation. Ordering of RNA upon protein binding has been documented in nuclear magnetic resonance experiments[30]. It is of considerable interest to note that this mechanism may also explain some structure–function data in the MS2 translational operator complex. In this case, higher protein affinity for a mutant containing a uracil to cytidine alteration was explained on the basis of an additional intramolecular RNA hydrogen bond in the mutant complex[31]. A different possibility is that dynamic motions after complex formation might increase for the AGGU aptamer, thereby increasing the conformational entropy[32]. Combined thermodynamic and structural data for additional GlnRS-tRNA mutant complexes with a range of binding affinities may be helpful in evaluating these alternative hypotheses.

Note added in proof: Recently, a sulfate anion has been identified that bridges two noncanonical base pairs in the crystal structure of 4.5S RNA (ref. 42).

Methods

tRNA gene synthesis, transcription and purification. DNA oligonucleotides were purchased from Integrated DNA Technologies. Complementary strands of DNA were purchased; one oligonucleotide contains the T7 RNA polymerase promoter (5'-AATTCCTGCAGTAAT-ACGACTCACTATA-3') while the other encodes the complement of the tRNA gene and a 10-base overlap (underlined) with the T7 promoter. The sequence of the wild type G1-tRNA[Gln] is: 5'-TGGCTGGGGTACGAG-GATTCGAACCTCGGAATGCCGGAATCAGAATCCGGTGCCTTACCGCTTG-GCGATACCCCCTATAGTGAGT-3'. Templates for the engineered tRNAs were made by incorporating the appropriate changes in the second oligonucleotide (Fig. 1). A double stranded DNA template was then produced by annealing the two strands at 10 °C followed by cycling the temperatures between 10 °C and 37 °C four times in the presence of DNA polymerase I Klenow fragment (Promega, Madison, Wisconsin). The 500 µL polymerization reactions contained 4 µM of each oligonucleotide, 800 µM of each dNTP, 50 mM Tris-HCl (pH 7.5), 10 mM magnesium sulfate, 0.1 mM DTT and 25 U Klenow fragment. Double stranded DNAs were then precipitated in 0.3 M sodium acetate (pH 5.2), 65% (v/v) ethanol, and resuspended in 100 µl distilled water.

T7 RNA polymerase was purified and stored as described[33,34]. Run-off transcription reactions (5 ml) contained 250 mM Hepes-KOH (pH 7.5), 30 mM magnesium chloride, 2 mM spermidine, 0.1 mg ml⁻¹ bovine serum albumin (BSA), 5 mM each of ATP, CTP, GTP, and UTP (Pharmacia Biotech), 40 mM dithiothreitol (DTT), 5 µg inorganic pyrophosphatase (Boehringer-Mannheim), 50 U RNasin (Pharmacia Biotech) and 40 µg ml⁻¹ T7 RNA polymerase[35]. All of the recovered duplex DNA from the Klenow extensions was used as template. After incubation at 37 °C for 10 h, reactions were diluted 1:1 with distilled water and passed over a 5 ml DE-52 (Whatman) column, preequilibrated with 100 mM HEPES-KOH (pH 7.5), 12 mM magnesium chloride, 200 mM sodium chloride. After loading, the column was washed with 30 ml of the preequilibration buffer. tRNA was eluted with 100 mM HEPES-KOH (pH 7.5), 12 mM magnesium chloride, 600 mM sodium chloride. Fractions containing RNA were dialyzed into 10 mM piperazine-N,N'–bis (2-ethanesulfonic acid) (PIPES, pH 7.5), 10 mM magnesium chloride, heated to 75 °C for 10 min and slowly cooled to 20 °C at a concentration of ~0.5 mg ml⁻¹. The tRNA was then precipitated in 0.3 M sodium acetate (pH 5.2), 65% ethanol (v/v), dried and resuspended in 200 µl 10 mM PIPES (pH 7.5), 10 mM magnesium chloride.

Gel shift assays. To obtain RNAs labeled at a high specific activity for binding studies, a separate set of wild type and mutant tRNAs lacking the 3'-terminal adenosine nucleotide (tRNA-CC) were synthesized and purified as described above. Radioactive AMP was then transferred enzymatically to the 3' end of each tRNA-CC using tRNA nucleotidyltransferase[36] and [α-³²P]ATP (purchased from Amersham**)**. The purified tRNA-CC species were incubated with enzyme in the presence of [α-³²P]ATP for 15 min at 20 °C. The conditions for this reaction were: 0.3 µM [α-³²P]ATP, 0.8 µM tRNA-CC, 3.6 µg tRNA nucleotidyltransferase, 50 mM Tris-HCl (pH 8.0), 12 mM magnesium chloride and 5 mM DTT.

Labeled tRNA was repurified as described above, using a 1 ml DEAE column. The tRNA was then exchanged into 25 mM sodium acetate, 10 mM potassium chloride (pH adjusted to 5.8 with acetic acid) using a Bio-spin chromatography column containing 1 ml of P6 Bio-Gel (Bio-Rad). Polyacrylamide gel electrophoresis (PAGE) was employed to determine the amount of incorporation.

GlnRS was purified and stored as described[34,37]. Thermodynamic dissociation constants (K_d) were determined by gel shift binding assay[16] using the 3′ end-labeled tRNAs. A concentrated stock solution of GlnRS was diluted in binding buffer (25 mM sodium acetate, 10 mM potassium chloride, 10 mM $MgCl_2$, 5 mM DTT; pH adjusted to 5.8 with acetic acid). Prior to performing the binding reactions G1-tRNA^{Gln} was diluted to 800 pM in binding buffer containing 12% glycerol (v/v), while AGGU and T1 tRNAs were diluted to 200 pM. For G1-tRNA^{Gln}, GlnRS was titrated from 660 nM to 66 pM and for AGGU and T1 tRNAs the concentration range was 66 pM to 6.6 nM. Twenty reactions evenly sampled across these concentration ranges were set up for each species.

Equal volumes of tRNA and GlnRS solutions were mixed and incubated for 15 min at ambient temperature. The mixtures were then loaded onto a 20% polyacrylamide gel at 4 °C and run at 200 V for 5 h. The gels were exposed for 72 h. Autoradiography and densitometry of the gels was performed with the Storm 840 phosphorimager (Molecular Dynamics). For AGGU and T1 tRNAs, the intensity of the shifted band was quantified to determine the relative amount of complex formed at each enzyme concentration. For G1-tRNA^{Gln}, the shifted band was diffuse due to dissociation of tRNA and protein during gel electrophoresis; in this case the K_d was determined by quantifying the free RNA band. Equilibrium binding constants were determined by fitting this data to a standard hyperbolic curve using Kaleidagraph (Synergy Software). All measurements were made in triplicate. Each reported K_d is the average of three determinations and the error represents the standard deviation from the mean (Table 2).

Crystallization and X-ray structure determination. Crystallization of the mutant complex was carried out under conditions similar to those used for the native complex[4,38]. Prior to crystallization, the tRNA was dialyzed into 10 mM PIPES (pH 7.5), 10 mM magnesium chloride, and concentrated to 20 mg ml^{-1} in a 3000 MW cutoff microconcentrator (Amicon). QSI (2 mM), an analog of glutaminyl adenylate[17], was added to the tRNA solution. The tRNA solution was mixed with equal volumes of 9 mg ml^{-1} wild type GlnRS (giving a tRNA:protein mole ratio of ~6) in 5 mM PIPES (pH 7.0), 5 mM β-mercaptoethanol. Crystals were grown by hanging drop vapor diffusion using 4 μl drops of the tRNA–protein solution over a reservoir containing 2.0 M ammonium sulfate, 70 mM PIPES (pH 7.0), 20 mM $MgCl_2$ and 20 mM β-mercaptoethanol. Crystals typically grew to full size (~0.1 mm × 0.15 mm × 0.5 mm) within two weeks.

Cryoprotection of tRNA crystals was carried out as described[17], except that the QSI inhibitor was not included in the stabilization and cryostabilization solutions. X-ray diffraction data were collected at the Advanced Light Source at the Lawrence Berkeley Laboratory on beamline 5.0.2 equipped with a 2 × 2 CCD detector and an Oxford cryosytems cryostream cooler. Data were collected in 1° oscillations at a wavelength of 0.97 Å. Data were reduced with MOSFLM and the CCP4 suite of programs[39].

Initial (F_o^{AGGU} - F_o^{native}) difference Fourier maps were generated using TNT[40]. Refinement was performed with TNT and X-PLOR[41] using

Table 2 Data collection and refinement statistics for var-AGGU

Space group	C222$_1$
Cell constants	
a, b, c (Å)	236.9, 94.0, 113.9
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	30.0–2.7
Total number of observations	383,591
Unique reflections	44,056
R_{merge} (%)[1,2]	6.0 (33.2)
Completeness (%)[2]	97.7 (94.8)
R_{cryst} (30–2.7 Å; %)[3]	22.1
R_{free} (30–2.7 Å; %)[4]	29.3
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.60

[1]$R_{merge} = (\Sigma_h \Sigma_i |<F_h> - F_{hi}|) / (\Sigma_h F_h)$ where $<F_h>$ is the mean structure factor magnitude of i observations of symmetry-related reflections with Bragg index h.
[2]Numbers in parentheses refer to statistics for the highest 0.2 Å resolution shell.
[3]$R_{cryst} = (\Sigma_h \Sigma_i ||F_o| - |F_c||) / (\Sigma |F_o|)$ where F_o and F_c are the observed and calculated structure factor magnitudes.
[4]R_{free} was calculated with a test set containing 5% of the data removed at the start of refinement.

the 100 K structure of the wild type complex as a starting model[17]. All solvent molecules, all nucleotides of the variable loop (44–48) and the QSI inhibitor were removed. After rigid body and positional refinement using TNT, the four nucleotides of the variable loop were replaced using (F_o - F_c) difference density as a guide. Water molecules were placed where there was spherical difference density with proper hydrogen bonding geometry to at least one electronegative atom in the tRNA or protein model. In the place of the QSI inhibitor, an adenosyl-monophosphate molecule was fit into the inhibitor density, due to the lack of complete density for the inhibitor. Further positional refinement as well as B-factor refinement and bulk solvent correction was carried out using XPLOR (Table 1). The final model consists of 529 amino acids, 74 nucleotides, one adenosyl-5′-monophosphate, three sulfate ions and 102 water molecules.

Coordinates. Coordinates have been submitted to the Protein Data Bank (accession code 1exd) and the Nucleic Acid Database (code PR0027).

articles

1. Conn, G.L. & Draper, D.E. RNA structure. *Curr. Opin. Struct. Biol.* **8**, 278–285 (1998).
2. Ferre-D'Amare, A.R. & Doudna, J.A. RNA folds: insights from recent crystal structures. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 57–73 (1999).
3. Cusack, S. RNA–protein complexes. *Curr. Opin. Struct. Biol.* **9**, 66–73 (1999).
4. Rould, M.A., Perona, J.J., Soll, D. & Steitz, T.A. Structure of *E. coli* glutaminyl-tRNA synthetase complexed with tRNA^Gln and ATP at 2.8 Å resolution. *Science* **246**, 1135–1142 (1989).
5. Ruff, M. *et al*. Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA^Asp. *Science* **252**, 1682–1689 (1991).
6. Biou, V., Yaremchuk, A., Tukalo, M. & Cusack, S. The 2.9 Å crystal structure of *T. thermophilus* seryl-tRNA synthetase complexed with tRNA^Ser. *Science* **263**, 1404–1410 (1994).
7. Cusack, S., Yaremchuk, A. & Tukalo, M. The crystal structures of *T. thermophilus* lysyl-tRNA synthetase complexed with *E. coli* tRNA^Lys and a *T. thermophilus* tRNA^Lys transcript: anticodon recognition and conformational changes upon binding of a lysyl-adenylate analogue. *EMBO J.* **15**, 6321–6334 (1996).
8. Cusack, S., Yaremchuk, A., Krikliviy, I & Tukalo, M. tRNA^Pro anticodon recognition by *Thermus thermophilus* prolyl-tRNA synthetase. *Structure* **6**, 101–108 (1998).
9. Silvian, L.F., Wang, J. & Steitz, T.A. Insights into editing from an Ile-tRNA synthetase structure with tRNA^Ile and mupirocin. *Science* **285**, 1074–1077 (1999).
10. Sankaranarayanan, R. *et al*. The structure of threonyl-tRNA synthetase-tRNA^Thr complex enlightens its repressor activity and reveals an essential zinc ion in the active site. *Cell* **97**, 371–381 (1999).
11. Nissen, P. *et al*. Crystal structure of the ternary complex of Phe-tRNA^Phe, EF-Tu, and a GTP analog. *Science* **270**, 1464–1472 (1995).
12. Nissen, P., Thirup, S., Kjeldgaard, M. & Nyborg, J. The crystal structure of Cys-tRNA^Cys-EFTu-GDPNP reveals general and specific features in the ternary complex and in tRNA. *Structure* **7**, 143–156 (1999).
13. Schmitt, E., Panvert, M., Blanquet, S. & Mechulam, Y. Crystal structure of methionyl-tRNA^fMet transformylase complexed with the initiator formyl-methionyl-tRNA^fMet. *EMBO J.* **17**, 6819–6826 (1998).
14. Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A. & Steinberg, S. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **26**, 148–153 (1998).
15. Stubenrauch, M. *In vitro* selection and characterization of tRNA-like substrates of *E. coli* glutaminyl-tRNA synthetase. Doctoral thesis, Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado; 1996.
16. Carey, J. Gel retardation. *Methods Enzymol.* **208**, 103–117 (1991).
17. Rath,V.L., Silvian, L.F., Beijer, B., Sproat, B.S. & Steitz, T.A. How glutaminyl-tRNA synthetase selects glutamine. *Structure*, **6**, 439–449 (1998).
18. Arnez, J.G. & Steitz, T.A. Crystal structure of unmodified tRNA^Gln complexed with glutaminyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry* **33**, 7560–7567 (1994).
19. Masquida, B., Sauter, C. & Westhof, E. A sulfate pocket formed by three G-U pairs in the 0.97 Å resolution X-ray structure of a nonameric RNA. *RNA* **5**, 1384–1395 (1999).
20. Leontis, N.B. & Westhof, E. Conserved geometrical base-pairing patterns in RNA. *Quat. Rev. Biophys.* **31**, 399–455 (1998).
21. Ferre-D'Amare, A.R., Zhou, K. & Doudna, J.A. Crystal structure of a hepatitis delta virus ribozyme. *Nature* **395**, 567–574 (1998).
22. Convery, M.A. *et al*. Crystal structure of an RNA aptamer–protein complex at 2.8 Å resolution. *Nature Struct. Biol.* **5**, 133–139 (1998).
23. Rowsell, S. *et al*. Crystal structures of a series of RNA aptamers complexed to the same protein target. *Nature Struct. Biol.* **5**, 970–975 (1998).
24. Lim W.A. & Sauer, R.T. Alternative packing arrangements in the hydrophobic core of λ repressor. *Nature* **339**, 31–36 (1989).
25. Baldwin, E.P., Hajiseyedjavadi, O., Baase, W.A. & Matthews, B.W. The role of backbone flexibility in the accommodation of variants that repack the core of T4 lysozyme. *Science* **262**, 1715–1718 (1993).
26. Lim, W.A., Hodel, A., Sauer, R.T. & Richards, F.M. The crystal structure of a mutant protein with altered but improved hydrophobic core packing. *Proc. Natl. Acad. Sci. USA* **91**, 423–427 (1994).
27. Vetter, I.R. *et al*. Protein structural plasticity exemplified by insertion and deletion mutants in T4 lysozyme. *Protein Sci.* **5**, 2399–2415 (1996).
28. Jen-Jacobson, L. Structural-perturbation approaches to thermodynamics of site-specific protein–DNA interactions. *Methods Enzymol.* **259**, 305–344 (1995).
29. Jayaram, B., McConnell, K.J., Dixit, S.B. & Beveridge, D.L. Free energy analysis of protein–DNA binding: the *Eco*RI endonuclease–DNA complex. *J. Comp. Phys.* **151**, 333–357 (1999).
30. Allain, F. H.-T., Gubser, C.C., Howe, P.W.A., Nagai, K., Neuhaus, D. & Varani, G. Specificity of ribonucleoprotein interaction determined by RNA folding during complex formation. *Nature* **380**, 646–650 (1996).
31. Valegard, K., Murray, J.B., Stonehouse, N.J., van den Worm, S., Stockley, P.G. & Liljas, L. The three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein–RNA interactions. *J. Mol. Biol.* **270**, 724–738 (1997).
32. Zidek, L., Novotny, M.V. & Stone, M.J. Increased protein backbone conformational entropy upon hydrophobic ligand binding. *Nature Struct. Biol.* **6**, 1118–1121 (1999).
33. Grodberg, J. & Dunn, J.J. *OmpT* encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *J. Bacteriol.* **170**, 1245–1258 (1988).
34. Nissan, T.A., Oliphant, B. & Perona, J.J. (1999). An engineered class I transfer RNA with a class II tertiary fold. *RNA* **5**, 434–445.
35. Milligan, J.F., Groebe, D.R., Witherell, G.W. & Uhlenbeck, O.C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **21**, 8783–8798 (1987).
36. McGann, R.G. & Deutcher, M.P. Purification and characterization of a mutant tRNA nucleotidyl transferase. *Eur. J. Biochem.* **106**, 321–328 (1980).
37. Hoben,P. *et al*. *Escherichia coli* glutaminyl-tRNA synthetase. *J. Biol. Chem.* **257**, 11644–11650 (1982).
38. Perona, J.J., Swanson, R., Steitz, T.A. & Soll, D. Overproduction and purification of *Escherichia coli* tRNA₂^Gln and its use in crystallization of the glutaminyl-tRNA synthetase-tRNA^Gln complex. *J. Mol. Biol.* **202**, 121–126 (1988).
39. Collaborative Computational Project Number 4, *Acta Crystallogr. D* **50**, 760–776 (1994).
40. Tronrud, D.E., Ten Eyck, L.F. & Matthews, B.W. An efficient general-purpose least-squares refinement program for macromolecular structures *Acta Crystallogr. A* **43**, 489–501 (1987).
41. Brunger, A.T., Kuriyan, J. & Karplus, M. Crystallographic R-factor refinement by molecular dynamics. *Science* **235**, 458–460 (1987).
42. Jovine, L. *et al*. Crystal structure of the Ffh and EF-G binding sites in the conserved domain IV of *E. coli* 4.55 RNA. Structure 8, 527–540 (2000)