Nucleic Acids Research, 1993, Vol. 21, No. 19 4467–4475

Contributions of discrete tRNA^{Ser} domains to aminoacylation by *E.coli* seryl-tRNA synthetase: a kinetic analysis using model RNA substrates

Jeffrey R.Sampson and Margaret E.Saks
Division of Biology 147-75, California Institute of Technology, Pasadena, CA 91125, USA

Received June 28, 1993; Revised and Accepted August 13, 1993

ABSTRACT

The aminoacylation kinetics of T7 transcripts representing defined regions of *Escherichia coli* serine tRNAs were determined using purified *E.coli* seryl-tRNA synthetase (SerRS) and the kinetic values were used to estimate the relative contribution of various tRNA^{Ser} domains to recognition by SerRS. The analysis revealed that the extra stem/loop structure, characteristic of type II tRNAs such as tRNA^{Ser}, is the domain which makes the largest contribution to k_{cat}/K_m of aminoacylation. Moreover, K_m of aminoacylation was increased by a factor of about 1000 when the extra stem/loop was changed to the consensus sequence of type I tRNA extra loops indicating that the stem structure contributes significantly to the binding of tRNA^{Ser} to SerRS. A model RNA, which represents only the tRNA^{Ser} coaxial acceptor-TΨC stem/loop domain, was also specifically aminoacylated by SerRS having a k_{cat}/K_m about 1000-fold greater than background levels. A significant portion of the contribution of this domain to aminoacylation is attributable to the acceptor stem sequence making the acceptor stem the second most important domain for recognition by SerRS. Finally, k_{cat}/K_m was essentially unchanged when the entire anticodon stem/loop of tRNA^{Ser} was deleted indicating that neither the anticodon nucleotides nor the surrounding stem/loop structure are important for recognition by SerRS.

INTRODUCTION

Studies of the recognition of transfer RNA (tRNA) by aminoacyl-tRNA synthetases generally have two goals. One goal is to elucidate the nucleotide(s) that synthetases specifically recognize. Another goal is to determine whether a given recognition element contributes to the binding of the tRNA or to catalysis. If more than one recognition element is required for aminoacylation specificity, then determining the relative contribution of each element to the overall aminoacylation process becomes an important concern.

The crystal structures of yeast tRNA^{Phe} (1) and tRNA^{Asp} (2) reveal that tRNAs are essentially comprised of individual RNA structural domains which must contain the recognition elements. The two most extensive domains are the acceptor-TΨC stem/loop and the anticodon-D stem/loop which primarily consist of double-stranded helical RNA. The coaxial acceptor-TΨC stem/loop begins near the 3′-terminus and extends to form the top (horizontal) portion of the L-shaped tRNA structure (Figure 1). The anticodon and D stems are stacked on each other and are at nearly a 90° angle with respect to the coaxial acceptor-TΨC stem/loop. This L-shaped tertiary structure is maintained by base-pairing interactions between nucleotides in the D and TΨC loops. The structures of tRNA^{Phe} and tRNA^{Asp} also reveal that the 4 or 5 nucleotides in the extra loop join the two major tRNA domains and participate in determining tRNA global structure by making tertiary interactions with the D stem base pairs. In contrast, type II tRNAs, such as *Escherichia coli* tRNA^{Ser}, tRNA^{Leu} and tRNA^{Tyr} have 9 or more nucleotides in the extra loop that are predicted by their sequence to form a distinctive stem/loop structure.

Mutational studies of full-length tRNAs indicate that recognition elements for aminoacyl-tRNA synthetases can be located in virtually any region of the tRNA (3). Consequently, there is no a priori way to know the locations of all recognition elements in any given group of isoaccepting tRNAs. Moreover, a recognition element's location does not appear to correlate with the mode by which it affects the aminoacylation reaction. The discriminator nucleotide at position 73 (4) and anticodon nucleotides both contribute to catalysis in a number of tRNA-synthetase systems yet these nucleotides are located at opposite ends of the tRNA. Finally, although aminoacylation depends on both binding (K_m) and catalysis (k_{cat}), it is generally thought that specificity is primarily dictated at the catalytic step (5, 6).

Previous studies from this laboratory have shown that the acceptor stem, the extra stem/loop structure and the D stem of *E.coli* tRNA^{Ser} are important for serine identity *in vivo* (7, 8). Building on this knowledge, we investigated the contribution of discrete domains of tRNA^{Ser} to recognition by seryl-tRNA synthetase (SerRS) by synthesizing RNAs that represented defined domains of tRNA^{Ser} and measuring their *in vitro* aminoacylation kinetics with purified SerRS. The serine system is particularly attractive since serine tRNAs contain the extra stem/loop structural domain in addition to the domains normally present in type I tRNAs. Using this approach, we have identified a dispensable domain within tRNA^{Ser} and a domain that makes a large contribution to tRNA^{Ser}-SerRS binding. From the aminoacylation kinetic data for a series of related RNAs, the

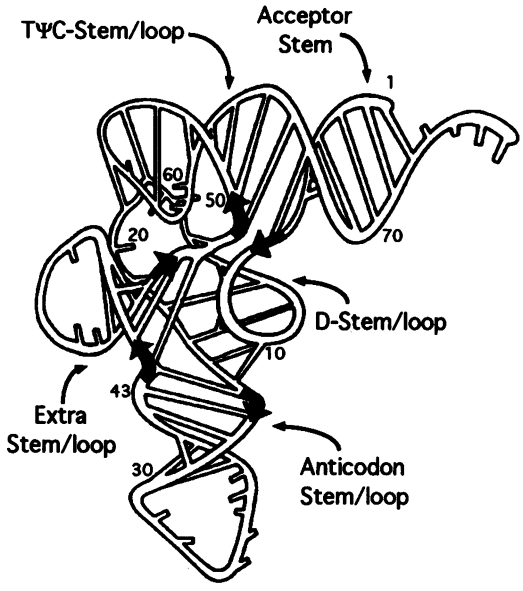

Figure 1. Ribbon diagram of the general tertiary structure for type II tRNAs. The orientation of the type II tRNA extra stem/loop is based on the model of yeast $tRNA^{Ser}$ (15). The dark arrows designate the 5'-to 3'-direction of the phosphate-backbone chain at positions where the chain was altered to create the model RNAs. The tRNA structural domains relevant to this study are indicated. Nucleotides are numbered using the conventional tRNA numbering system (37).

relative contribution of each domain to the overall aminoacylation reaction was estimated in terms of the difference in free energy of activation ($\Delta\Delta G^{\dagger}$).

MATERIALS AND METHODS

Materials

L-[3-^{3}H] serine (S.A. 29–33 Ci/mmole) was purchased from Amersham. Oligodeoxyribonucleotides were synthesized by the California Institute of Technology Biopolymer Synthesis and Analysis Resource Center on an Applied Biosystems DNA Synthesizer (Foster City, CA). Native serine tRNAs were purchased from Subriden RNA (Rollingbay, WA). BstN I restriction endonuclease was purchased from New England Biolabs. Purified *E.coli* inorganic pyrophosphatase (500 units/mg) was graciously provided by B. Cooperman. T7 RNA polymerase was purified to a specific activity of 650,000 units/mg, using the method of Grodberg and Dunn (9), from the overproducing strain *E.coli* BL21 harboring the plasmid pAR1219 (10) which was kindly provided by J. Dunn. *E.coli* SerRS was purified from *E.coli* JM101 harboring the plasmid pSerS2 which was kindly provided by M.Härtlein and R.Leberman. Purification was carried out essentially as described by Härtlein and co-workers (11) except that the SerRS active Sepharose CL6B fractions were applied to a hydroxylapatite column and the flow-through fractions containing SerRS were concentrated by chromatography on Sepharose CL6B. The specific activity of the purified SerRS was 1,200 units/mg protein. One unit is defined as the amount of SerRS required to aminoacylate 1.0 nmole of $tRNA^{Ser}$ per minute at 37°C at 5 μM $tRNA^{Ser}$ under the reaction conditions described below. Protein concentrations were determined using the Bradford Assay (BioRad) with bovine serum albumin (BSA) as the standard. SDS-polyacrylamide gel electrophoresis indicated that the SerRS was about 70% pure. All SerRS concentrations reported in this study are for the homodimer and are corrected for the 70% purity.

Gene construction

The genes for RNAs that were longer than 35 nucleotides (T7-CGA, T7-CUA, T7-ΔXL, T7-ΔAC; Figure 2) were cloned into the plasmid pUC18 and contained a T7 RNA polymerase promoter directly upstream of the tRNA gene and a downstream BstN I restriction endonuclease site which defines the 3'-terminal A76 in the RNA product. The genes for T7-CUA and T7-ΔXL were constructed by G.D. Tocchini-Valentini using PCR to introduce the T7-promoter and BstN I restriction site into genes that were previously constructed by J. Normanly (7, 8). The genes for T7-CGA and T7-ΔAC were constructed *de novo* using six overlapping DNA oligonucleotides. Plasmids were transformed into *E.coli* strain DH10B using standard methods (12) and sequenced using the chain-termination method (Sequenase, USB).

In vitro transcription and purification of RNAs

T7-CGA, T7-CUA, T7-ΔXL, and T7-ΔAC were transcribed from BstN I linearized plasmid DNA (13) in reaction mixtures containing: 40 mM Tris-Cl (pH 8.3 at 25°C), 20 mM $MgCl_2$, 5.0 mM DTT, 1.0 mM spermidine, 4 mM each NTP (pH 8.0), 10 mM GMP, 0.1 mg/ml BSA, 0.5 units/ml inorganic pyrophosphatase, 26,000 units/ml T7 RNA polymerase and 0.1 mg/ml template DNA. After incubation at 40°C for 4 hours, the reactions were stopped by adding EDTA to a final concentration of 25 mM, phenol/chloroform extracted, ethanol precipitated, and purified to single nucleotide resolution by electrophoresis on 8% 7 M urea polyacrylamide gels measuring 0.17×20×40 cm. Three native serine tRNAs were also gel purified. One isoacceptor was $tRNA^{Ser\text{-}1}$ which has the anticodon UGA. Two different serine isoacceptors were recovered from certain lots of $tRNA^{Ser\text{-}3}$. One corresponded to $tRNA^{Ser}$(GCU) and the other slower migrating species is presumed to be $tRNA^{Ser}$(UCA) (the selenocysteine tRNA) as deduced by partial T1 ladders of 3'-end labeled RNA (data not shown).

The mini and microhelix RNAs as well as 3'-Leumin (Figure 2) were transcribed from synthetic DNA templates (14) in reaction mixtures containing: 40 mM Tris-Cl (pH 8.3 at 25°C), 20 mM $MgCl_2$, 5.0 mM DTT, 1.0 mM spermidine, 7.5% PEG (mwt. 8,000), 0.01% Triton X-100, 4 mM each NTP (pH 8.0), 10 mM GMP, 0.3 to 0.4 μM DNA template, 0.05 mg/ml BSA, 0.15 units/ml inorganic pyrophosphatase, and 8,500 units/ml T7 RNA polymerase. After incubation at 40°C for 4 hours, the reactions were stopped as described above. The RNA products were purified to single nucleotide resolution on 15 or 20% 7 M urea polyacrylamide gels measuring 0.17×20×25 cm.

All RNAs were eluted from the gels at 4°C with 10 volumes of a buffer containing 200 mM KOAc (pH 5.2), 10 mM EDTA and were precipitated with 2.5 volumes of ethanol at $^{-}$20°C. RNAs were then reprecipitated 3 times with 2.5 volumes of ethanol in the presence of 2.0 M NH_4OAc. RNA concentrations were determined from measured A_{260} values of total T2 nuclease-digested RNA samples. The E^{M}_{260} for each RNA was calculated using the following E^{M}_{260} values for each Np: Gp (11.65×10^3), Ap (15.4×10^3), Cp (7.52×10^3) and Up (10.0×10^3).

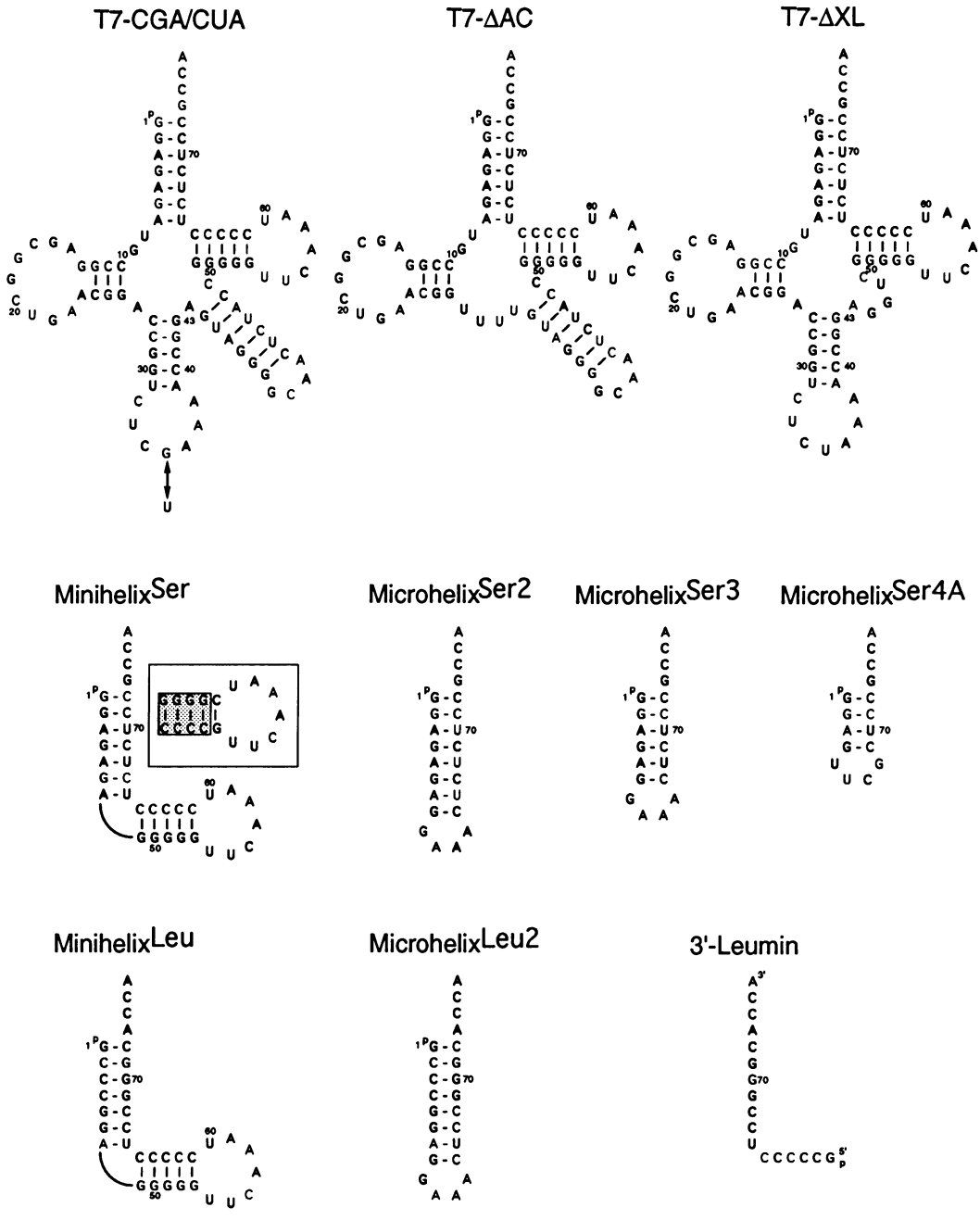

Figure 2. Model RNAs used in this study. The predicted secondary structure of each RNA is presented. The inset associated with minihelix^{Ser} indicates the nucleotide substitutions made to create minihelix$^{SerT\Psi C}$. 3′-Leumin is presented as a single stranded RNA, yet its true solution structure has not been determined. The acceptor stem sequences of minihelix^{Leu}, microhelix^{Leu2} and 3′-Leumin are based on tRNA^{Leu}(UAA) but have a U72C substitution. For the sequences of the native serine tRNAs used in this study see Sprinzl *et al.* (38).

Aminoacylation reactions

Unless otherwise indicated, aminoacylation reactions were performed in 50 or 100 μl reaction mixtures containing 30 mM HEPES-KOH (pH 7.4), 15 mM $MgCl_2$, 25 mM KCl, 2 mM DTT, 2 mM ATP, 40 μM ^{3}H-serine (S.A. 20 Ci/mmole) and 0.02 units/ml inorganic pyrophosphatase (PPiase). For the kinetic determinations, RNA concentrations ranged from 0.1 μM to 160 μM and SerRS concentrations ranged from 0.43 nM to 1.3 μM depending on the RNA substrate. All RNAs were annealed by heating for three minutes at 95°C in 10 mM HEPES-KOH (pH 7.4) at a concentration 2.5 times greater than their final concentration in the aminoacylation reaction. The mini and microhelix RNAs were snap cooled on ice and all other RNAs were cooled to 25°C over a 45 minute period. The total protein in each reaction was adjusted with BSA to a final concentration of 0.18 mg/ml. Reaction mixtures were incubated at 37°C and at appropriate time intervals, 8 μl aliquots were spotted on dried pieces of Whatman 3MM paper that had been pretreated with 10% TCA, 75 μM serine and 25 μM of each of the other 19 amino acids. The papers were washed two times with 10% TCA, three times with 5% TCA, and one time with 95% ethanol, for

approximately 15 minutes per wash at 4°C in a wash volume to paper area ratio of 1.8 ml:cm^2. Using this procedure, the residual background of ^{3}H-serine (cpm) for the reactions containing no RNA was no greater than 0.008% of the total radioactivity spotted on the paper and was independent of SerRS concentration. The precipitation and retention efficiency for the full-length tRNAs was essentially quantitative (>90%) at all RNA concentrations tested. For the RNAs which were 35 nucleotides or shorter, quantitative precipitation was obtained when greater than 20 pmoles of RNA were spotted onto the paper (data not shown). The counting efficiency of the ^{3}H-serine-tRNA^{Ser} on the Whatman filter paper was empirically determined to be 25% in a fluor consisting of 3.83 grams of 2,5-diphenyloxazole per liter of toluene.

Substrate design

Because this study involved the use of truncated tRNAs, it was critical that the elimination of discrete RNA domains caused neither distortion nor gross misfolding of the remaining RNA. The crystal structures of yeast tRNA^{Phe} (1), yeast tRNA^{Asp} (2) and the model of yeast tRNA^{Ser} (15) were used to achieve this goal. In creating the truncated RNAs, we considered the directionality of the phosphodiester backbone and the location of nucleotides which are involved in base-pairing and tertiary interactions. In principle, these factors should dictate where the tRNA can and cannot be altered without perturbing the remaining tRNA structure (Figure 1).

The model of yeast tRNA^{Ser} (15) suggests that the hairpin structure that comprises the extra stem/loop structure can be thought of as a discrete domain inserted within the variable loop of a type I tRNA. This suggests that the elimination of the extra stem/loop structure should have little or no effect on the overall tertiary structure of tRNA^{Ser} as long as nucleotide 43 of the anticodon stem and nucleotide 49 of the TΨC stem are correctly joined. Normanly *et al.* (8) had previously constructed a gene for *in vivo* studies in which the normal extra stem/loop structure of tRNA^{Ser} was eliminated and the phosphodiester chain was rejoined by constructing a small variable loop having the consensus sequence of type I *E.coli* tRNAs (A44, G45, G46, U47, C48). This gene was the basis for the RNA T7-ΔXL (Figure 2) which was used in our study.

The anticodon stem/loop can be similarly thought of as a separate domain which is stacked on the D stem. For T7-ΔAC, the anticodon stem/loop was deleted and the nucleotides C25 in the D stem and G45 in the extra stem/loop were joined with four uridine residues (Figure 2). Polyuridine was used to join the D stem and extra stem/loop since uridine is the least likely of the four ribonucleotides to induce gross refolding of the molecule or to stabilize unanticipated structures in this region of the tRNA.

The coaxial acceptor-TΨC stem/loop is essentially a discrete domain comprised of a hairpin loop that is interrupted between positions 7 and 49 by the remainder of the tRNA structure. McClain *et al.* (16) and Francklyn *et al.* (17) showed that this domain could be successfully represented by a 12 base-pair hairpin having a NCCA-3′ overhang. Thus minihelix^{Ser} was designed using the acceptor stem sequence of tRNA^{Ser}(CGA) and the *E.coli* tRNA^{Ser} consensus TΨC stem/loop sequence (Figure 2). Additional minihelices having the acceptor stem sequence of tRNA^{Leu}(UAA) (minihelix^{Leu}) or having an altered TΨC stem sequence (minihelix^{SerTΨC}) were synthesized to test the effect of sequence composition on aminoacylation (Figure 2).

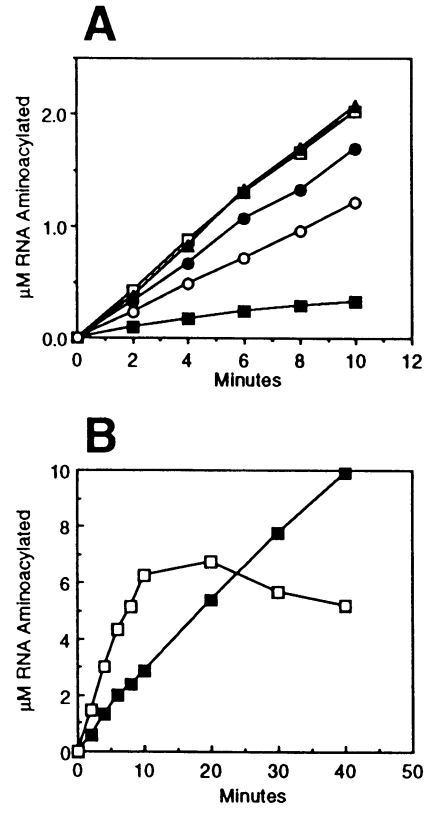

Figure 3. **(A)** Effect of inorganic pyrophosphatase on the aminoacylation rate of minihelix^{Ser}. Aminoacylation rates were determined at 50 μM RNA and 3.3 μM SerRS under the reaction conditions described in Materials and Methods. The final concentration of PPiase (units/ml) for each determination is as follows: (■) no PPiase; (○) 0.002; (●) 0.005; (▲) 0.02; (□) 0.2. **(B)** Effect of buffer on the aminoacylation rate of minihelix^{Ser}. The aminoacylation rates were determined, in the presence of PPiase, at 50 μM RNA and 3.3 μM SerRS under the reaction conditions described in Materials and Methods in either the (□) Tris-Cl or (■) HEPES-KOH buffer.

Table 1. Kinetic parameters for aminoacylation of native tRNA^{Ser} isoacceptors by *E.coli* seryl-tRNA synthetase

tRNA	K_m (μM)	k_{cat} (sec^{-1})	k_{cat}/K_m ($sec^{-1}M^{-1}$)
tRNA^{Ser} (GCU)	0.19 ± 0.022	3.5 ± 0.13	1.8×10^7
tRNA^{Ser} (UGA)	0.12 ± 0.011	3.1 ± 0.07	2.5×10^7
tRNA^{Ser} (UCA)	0.71 ± 0.079	5.4 ± 0.29	7.6×10^6

Kinetic parameters (± standard error) were determined for each substrate using six (GCU and UGA) or four (UCA) RNA concentrations which ranged between 0.1 and 1.8 μM at a SerRS concentration of 0.43 nM. Initial rates were calculated from a linear least squares analysis using five time points and all correlation coefficients were 0.99 or greater. Kinetic parameters and standard errors were calculated using the program ENZFITTER (Biosoft).

Microhelices comprised of eight base-pairs or less and closed by a thermodynamically stable tetra-loop (18, 19) were also synthesized (Figure 2). The tetra-loop sequences were chosen to minimize possible alternative hairpin structures and to drive the structures indicated in Figure 2. Microhelix^{Ser2} has the complete acceptor stem sequence of tRNA^{Ser}(CGA) whereas microhelix^{Ser3} and microhelix^{4A} have only its first six and first four base pairs, respectively. Microhelix^{Leu2} presents the same

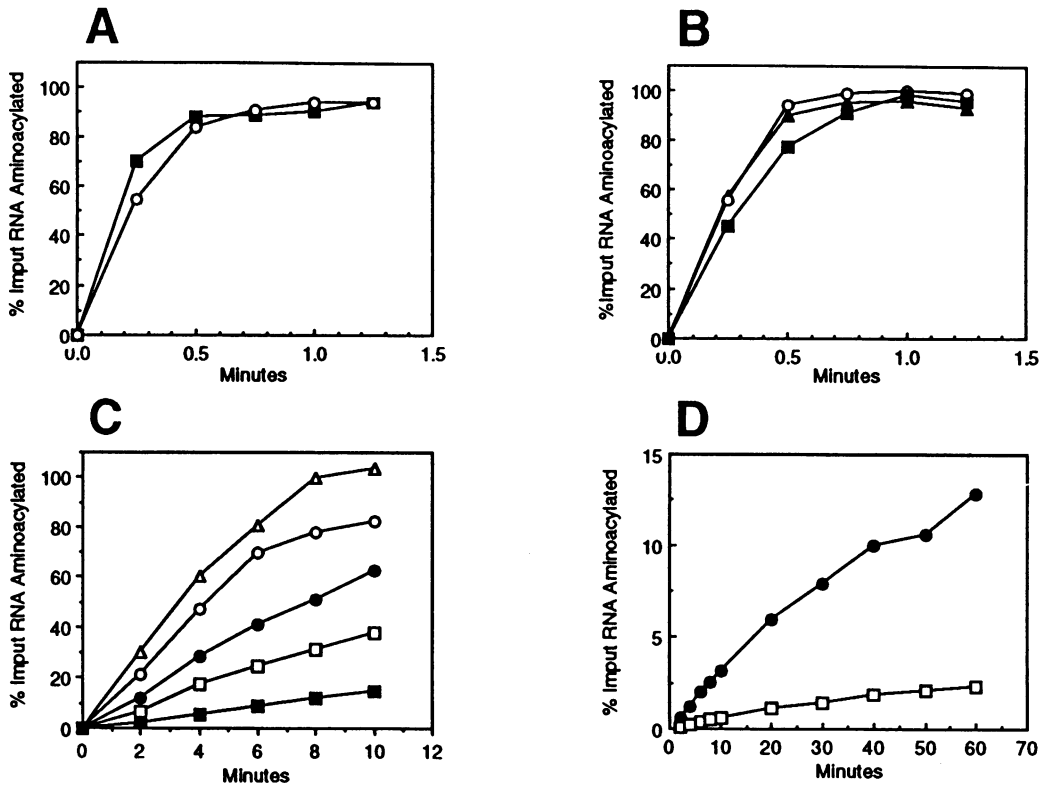

Figure 4. Aminoacylation plateau levels of RNAs. **(A)** Aminoacylation of native serine tRNA isoacceptors. Aminoacylation was measured at 5.0 μM RNA and 0.065 μM SerRS under the reaction conditions described in Materials and Methods. (■) $tRNA^{Ser}$(GCU); (○) $tRNA^{Ser}$(UGA). Data for $tRNA^{Ser}$(UCA) not shown. **(B)** Aminoacylation of T7-$tRNA^{Ser}$ transcripts. Aminoacylation was measured at 5.0 μM RNA and 0.065 μM SerRS under the reaction conditions described in Materials and Methods. (▲) T7-CGA; (○) T7-CUA; (■) T7-ΔAC **(C)** Aminoacylation of T7-ΔXL as a function of SerRS concentration. Aminoacylation was measured at 5.0 μM RNA using the indicated SerRS concentrations under the reaction conditions described in Materials and Methods. The final concentration SerRS (μM) for each determination is as follows: (■) 0.065; (□) 0.16; (●) 0.33; (○) 0.65; (△) 1.3. **(D)** Aminoacylation of $minihelix^{Ser}$ and $microhelix^{Ser2}$. Aminoacylation was measured at 100 μM RNA and 1.3 μM SerRS under the reaction conditions described in Materials and Methods. (●) $MinihelixSer$; (□) $MicrohelixSer2$.

structure as $microhelix^{Ser2}$ but has the acceptor stem sequence of $tRNA^{Leu}$(UAA). The microhelices do not necessarily represent a discrete structural domain of the full-length tRNA tertiary structure but rather they isolate the acceptor stem as a potential functional subdomain and provide a means for determining the contribution of the acceptor stem to recognition by SerRS in the absence of the remaining tRNA structure. Finally, the rate of non-specific aminoacylation by SerRS was determined using the RNA 3'-Leumin whose sequence corresponds to the 3'-half of $minihelix^{Leu}$.

RESULTS

Aminoacylation reaction conditions

In the course of this study, we found that the RNAs described above differed in the aminoacylation specificity constant (k_{cat}/K_m) for SerRS by as much as nine orders of magnitude. It was possible to use a single detection method over such a broad window of measurability after defining reaction conditions which yielded reliable data when the SerRS concentration and/or the reaction incubation times were varied. The initial experiments which focused on defining reaction conditions that met these criteria are briefly outlined below.

The SerRS concentrations used to measure aminoacylation rates of both good and relatively poor substrates examined in this study ranged from as little as 0.43 nM to as much as 1.3 μM. The validity of comparing rates obtained using a wide range of SerRS concentrations was directly examined by determining the rates of aminoacylation at various SerRS concentration. The requisite proportional increases in the initial rates of aminoacylation were observed as SerRS concentration was increased as long as the concentration did not exceed 1.3 μM (data not shown). Thus, for all the kinetic studies or when direct comparisons were made between two different RNAs, the SerRS concentration never exceeded 1.3 μM.

Inorganic pyrophosphate (PPi) is a product of the aminoacylation reaction and has been shown to inhibit the reaction (20). At saturating $tRNA^{Ser}$ concentrations, the aminoacylation activity of SerRS is inhibited 50% by about 25 μM exogenously added PPi (data not shown) which is consistent with a study published earlier (21). Although at low SerRS concentrations the amount of PPi released as a result of aminoacyl-adenylate formation would be negligible, at high SerRS concentrations and/or during long incubation times, the concentration of PPi could approach inhibitory levels. This was of concern since high SerRS concentrations and relatively long incubation times were required for determining the aminoacylation rates of poor substrates such as the mini- and microhelices. The addition of inorganic pyrophosphatase (PPiase) to the aminoacylation reactions could potentially solve the problem since this enzyme catalyzes the hydrolysis of PPi to inorganic phosphate (22). To test whether the addition of PPiase affected aminoacylation by

SerRS, the aminoacylation rate of a poor substrate (minihelix^{Ser}) was determined using a high SerRS concentration (3.3 μM) and various PPiase concentrations. Figure 3A shows that the aminoacylation rate increased with increasing PPiase concentration until, at an optimum of about 0.02 units PPiase/ml, the reaction rate was improved 4-fold compared to the control which contained no PPiase. More importantly, the addition of PPiase maintained the linearity of the rate for at least one hour (see Figure 4D). The possibility that the addition of PPiase might alter the effects of mutations on aminoacylation rates was also examined. These experiments indicated that the presence or absence of PPiase in the reaction mixture did not affect aminoacylation rates of mutant minihelices relative to minihelix^{Ser} (data not shown). Thus 0.02 units/ml PPiase was included in the aminoacylation reaction mixture that was used for all RNAs.

The effect of buffer composition on the linearity of the aminoacylation rates over extended time periods was also examined. Figure 3B shows the aminoacylation rates for minihelix^{Ser} using the standard reaction mixture but in the presence of either 30 mM HEPES-KOH (pH 7.4) or 30 mM Tris-Cl (pH 7.4 at 37°C). During the first 15 minutes, the rate obtained in the Tris-Cl buffer was 2.5-fold greater than that obtained in the HEPES-KOH buffer. However, as the reaction proceeded, the rate in the Tris-Cl buffer declined markedly suggesting that the RNA was deacylated. The ability of nucleophilic buffers such as Tris-Cl to enhance both the forward and reverse rates of aminoacylation has been previously documented (23). In contrast, the aminoacylation rate was relatively linear for one hour in the presence of HEPES-KOH. Consequently, the HEPES-KOH buffer was used for all experiments.

Initial characterization of RNAs

Given the unusual nature of the RNAs which were used, and the necessity to maintain tRNA-like structures, it was important to distinguish changes in aminoacylation kinetics which are attributable to the loss of specific phosphate or nucleoside contacts with SerRS from those due to gross misfolding of the RNA. Independent structural assays involving specific Pb^{+2} cleavage and UV-induced intramolecular cross-linking have been successfully used to differentiate between these two types of effects for mutants of both yeast and *E.coli* tRNA^{Phe} (24, 25, 26). However, there is no simple method for determining the effect of a mutation on tRNA structure for tRNAs such as tRNA^{Ser} which have neither a specific Pb^{+2} cleavage site nor the nucleotides required for intramolecular cross-linking. We therefore relied on measurements of the aminoacylation plateau of each model RNA to assess its overall aminoacylation activity and to reveal whether or not it was grossly misfolded. However, it is important to note that subtle effects on tRNA structure could not be assessed by this assay.

At 5 μM RNA and 0.065 μM SerRS, native tRNA^{Ser}(GCU), tRNA^{Ser}(UGA), tRNA^{Ser}(UCA) and the full-length tRNA^{Ser} transcripts (T7-CGA, and T7-CUA) showed similar aminoacylation rates and were aminoacylated to a level of nearly 100% after 40 seconds (Figure 4A,B). Under the same reaction conditions, T7-ΔAC was also aminoacylated with a rate and final level similar to that of the full-length tRNAs (Figure 4B). Using the above RNA and SerRS concentrations, only about 15% of T7-ΔXL could be aminoacylated after 10 minutes (Figure 4C). However, as the SerRS concentration was increased, a greater percentage of T7-ΔXL was aminoacylated and reached nearly 100% at 1.3 μM SerRS. This indicates that T7-ΔXL, like T7-ΔAC, is in a form that can be aminoacylated and is thus most likely correctly folded. That T7-ΔXL is an efficient glutamine-inserting amber suppressor when expressed *in vivo* (8) also suggests that it adopts a normal tRNA structure.

Minihelix^{Ser} is also aminoacylated by SerRS, but to a lower extent than any of the above RNAs (Figure 4D). At 100 μM minihelix^{Ser} and 1.3 μM SerRS, only about 13% of the RNA was aminoacylated after 60 minutes. Microhelix^{Ser} was an even less efficient substrate and was aminoacylated to only about 2% under these conditions (Figure 4D). Although higher SerRS concentrations could not be used to test whether these RNAs could be fully aminoacylated, native gel electrophoresis indicated that they are structurally homogeneous and form hairpin structures of their predicted molecular weights (data not shown).

Approximately 13 μM minihelix^{Ser} had been aminoacylated in the presence of 1.3 μM SerRS after 60 minutes (Figure 4D). This indicates that each molecule of SerRS has participated in multiple cycles of the aminoacylation reaction. Because the aminoacylation rate was essentially linear, it is likely that rates measured when less than one mole of RNA was aminoacylated per mole of SerRS reflect the rates which would be obtained during multiple turnover events.

Aminoacylation kinetic determinations

The aminoacylation kinetics for both the native serine tRNAs and the full-length T7 transcripts were compared. The native tRNA^{Ser}(GCU) and tRNA^{Ser}(UGA) isoacceptors showed similar

Table 2. Kinetic parameters for aminoacylation of model RNAs by *E.coli* seryl-tRNA synthetase

RNA	K_m (μM)	k_{cat} (sec^{-1})	k_{cat}/K_m ($sec^{-1} \cdot M^{-1}$)
T7-CGA	0.12 ± 0.023	4.2 ± 0.20	3.4×10^7
T7-CUA	0.079 ± 0.018	3.9 ± 0.17	5.0×10^7
T7-ΔAC	0.26 ± 0.035	4.4 ± 0.21	1.7×10^7
T7-ΔXL	107 ± 34	0.38 ± 0.081	3.5×10^3
Minihelix^{Ser}	236 ± 12	0.0082 ± 0.00028	3.5×10^1
Minihelix^{SerTΨC}	164 ± 19	0.0059 ± 0.00041	3.6×10^1
Microhelix^{Ser2}	274 ± 73	0.0019 ± 0.00035	6.8
Microhelix^{Ser3}	nd	nd	6.3 ± 0.24
Microhelix^{Ser4A}	nd	nd	0.18 ± 0.051
Minihelix^{Leu}	nd	nd	0.097 ± 0.016
Microhelix^{Leu2}	nd	nd	0.062 ± 0.0028
3′-Leumin	nd	nd	0.036 ± 0.011

Kinetic parameters (± standard error) for T7-CGA, T7-CUA, and T7-ΔAC were determined using six RNA concentrations which ranged between 0.1 and 1.6 μM at a SerRS concentration of 0.43 nM and those for T7-ΔXL were determined using six RNA concentrations which ranged between 2 and 70 μM at a SerRS concentration of 43 nM. Kinetic parameters for minihelix^{Ser}, minihelix^{SerTΨC}, and microhelix^{Ser2} were determined using five RNA concentrations which ranged between 10 and 160 μM at a SerRS concentration of 1.3 μM. The individual kinetic parameters and standard errors for these substrates were calculated using the program ENZFITTER (Biosoft). Individual kinetic parameters were not obtained for the remaining substrates. Instead, $V_o/[S]$ was assumed to reflect k_{cat}/K_m since the measured initial rates of aminoacylation were essentially proportional to RNA concentration and thus indicated that the RNA concentrations used were well below K_m and that the reaction was first order. The reported k_{cat}/K_m for microhelix^{Ser3} and microhelix^{Ser4A} represents the mean $V_o/[S]$ (± standard error) obtained at five RNA concentrations ranging between 10 and 160 μM at a SerRS concentration of 1.3 μM. The reported k_{cat}/K_m for minihelix^{Leu}, microhelix^{Leu2}, and 3′-Leumin represents the mean $V_o/[S]$ (± standard error) obtained at 50 and 150 μM RNA and at 1.3 μM SerRS. Initial rates for all substrates were calculated from a linear least squares analysis using five time points and the correlation coefficients were all 0.98 or greater.

k_{cat} and K_m values and differed less than 2-fold in k_{cat}/K_m (Table 1). The k_{cat}/K_m of $tRNA^{Ser}$(UCA) was lower than that of the other native serine tRNAs by a factor of about three primarily due to an increased K_m. Although the specificity constants for these three native serine tRNAs fall within a relatively narrow range, it is noteworthy that there is some variation among them. T7-CGA has the sequence of yet another serine isoacceptor but lacks the modified nucleotides of native tRNAs. Its k_{cat} and K_m were within the range found for native serine tRNAs (Table 2) indicating that, under our buffer conditions, the modified nucleotides make little or no contribution to recognition by SerRS. This is consistent with the results of a previous study by Himeno *et al.* (27).

The contributions of the anticodon stem and loop nucleotides to aminoacylation by SerRS were examined using two different substrates. The k_{cat} and K_m for the amber suppressor T7-CUA were essentially the same as those for T7-CGA (Table 2). Since no anticodon nucleotide is held in common among the serine isoacceptors it is not altogether surprising that anticodon nucleotides have no effect on aminoacylation. Although T7-ΔAC lacks the entire anticodon stem/loop domain, its k_{cat}/K_m was within the range observed for serine isoacceptors. These results clearly demonstrate that neither the anticodon nucleotides nor the surrounding stem/loop structure are involved in recognition by SerRS.

T7-ΔXL was used to examine the contribution of the extra stem/loop to aminoacylation by SerRS. This substrate is a T7-CUA in which the extra stem/loop was changed to a variable loop of five nucleotides. T7-ΔXL had a k_{cat}/K_m of aminoacylation that was 1.4×10^4-fold lower than that of T7-CUA (Table 2 and 3). Although K_m was extrapolated, the value of approximately 100 μM is reproducible. Thus the low specificity constant for T7-ΔXL is primarily due to an increased K_m which is about 1000-fold greater than that measured for any full-length $tRNA^{Ser}$. A large difference in the binding of T7-CGA and T7-ΔXL by SerRS was also indicated by the results of a nitrocellulose filter-binding assay (data not shown). Whereas the dissociation constant (K_d) for T7-CGA was 0.2 μM, there was no binding of T7-ΔXL at 4 μM SerRS. Thus these data indicate that the extra stem/loop makes a significant contribution to the binding of $tRNA^{Ser}$ to SerRS.

The mini and microhelices were used to examine the structural contributions of the acceptor stem and TΨC stem/loop domains to aminoacylation by SerRS (Table 2 and 3). The 1000-fold greater k_{cat}/K_m of $minihelix^{Ser}$ compared to the single-stranded RNA, 3′-Leumin, indicates that the acceptor-TΨC stem/loop structure makes an important contribution to aminoacylation. The nearly 200-fold difference between $microhelix^{Ser2}$ and 3′-Leumin should primarily reflect the contribution of the acceptor stem domain to aminoacylation. The approximately 5-fold difference in k_{cat}/K_m between $minihelix^{Ser}$ and $microhelix^{Ser2}$ should primarily reflect the contribution of the TΨC stem/loop structure to aminoacylation. Since k_{cat}/K_m was essentially the same for $microhelix^{Ser2}$ and $microhelix^{Ser3}$, neither the first base pair of the TΨC stem nor the last acceptor stem base pair seem to contribute to aminoacylation by SerRS. In contrast, the approximately 35-fold lower k_{cat}/K_m for $microhelix^{Ser4A}$ compared to $microhelix^{Ser2}$ indicates that the acceptor stem region encompassing base pairs 5-68 and/or 6-67 contributes to recognition by SerRS.

The potential contributions of the acceptor stem and TΨC stem base-pair sequence to aminoacylation were also examined (Table 2). There was no difference in k_{cat}/K_m between $minihelix^{Ser}$ and $minihelix^{SerTΨC}$, which has an altered TΨC stem sequence, indicating that the contribution of the TΨC stem/loop domain indicated above is base-pair sequence independent. This conclusion is consistent with the fact that, with the exception of the G53-C61 base pair found in most tRNAs, no TΨC stem base pair is held in common among serine isoacceptors. In contrast, the acceptor stem contributes to aminoacylation in a sequence-specific manner. The k_{cat}/K_m was about 360-fold greater for the minihelix having a $tRNA^{Ser}$ ($minihelix^{Ser}$) versus a $tRNA^{Leu}$ ($minihelix^{Leu}$) acceptor stem sequence and it was approximately 100-fold greater for $microhelix^{Ser2}$ compared to $microhelix^{Leu2}$. The somewhat smaller effect of changing the acceptor stem sequence in the microhelix background could be due to the low activity of $microhelix^{Leu2}$ whose k_{cat}/K_m was only 2-fold greater than that obtained for the non-specific RNA 3′-Leumin.

Table 3. Estimated relative contribution of each $tRNA^{Ser}$ domain to aminoacylation by *E.coli* seryl-tRNA synthetase

tRNA domain	Defined operational rate constant	Relative k_1/k_2	$-\Delta\Delta G^{\dagger}$ (kcal/mole)
Full length $tRNA^{Ser}$	$k_{(T7\text{-}CUA)}/k_{(3'\text{-}Leumin)}$	1.4×10^9	(13.0)
Extra Stem/loop	$k_{(T7\text{-}CUA)}/k_{(T7\text{-}\Delta XL)}$	1.4×10^4	5.9
Acceptor Stem	$k_{(Microhelix^{Ser2})}/k_{(3'\text{-}Leumin)}$	1.9×10^2	3.2
TΨC Stem/loop	$k_{(Minihelix^{Ser})}/k_{(Microhelix^{Ser2})}$	5.1	1.0
Anticodon Stem/loop	$k_{(T7\text{-}CUA)}/k_{(T7\text{-}\Delta AC)}$	2.9	0.65
D Stem/loop	nd	nd	2.2

The $-\Delta\Delta G^{\dagger}$ values were calculated using the equation $\Delta\Delta G^{\dagger} = -RT\ln(k_1/k_2)$ where k_1 and k_2 correspond to the k_{cat}/K_m values for the indicated RNAs (see Table 2), R=1.987 cal·°K^{-1}·mole^{-1}, and T=310°K (37°C).

Relative contribution of RNA domains to recognition

The contribution of discrete tRNA domains to aminoacylation can be estimated, under the simple assumption that each domain contributes independently to aminoacylation, by calculating the difference in free energy of activation ($\Delta\Delta G^{\dagger}$) using the specificity constants (k_{cat}/K_m) of appropriate pairs of RNAs as the operational rate constant (k) (28, 29). Since $\Delta\Delta G^{\dagger}$ is equal to $-RT\ln(k_1/k_2)$, the difference in free energy of activation for any pair of RNA substrates can be expressed as $-RT\ln[(k_{cat}/K_m)_{RNA1}/(k_{cat}/K_m)_{RNA2}]$. Table 3 presents the $-\Delta\Delta G^{\dagger}$ values for the contribution of each domain to aminoacylation using the specificity constants of the pair of RNAs that define each domain. The overall window of $-\Delta\Delta G^{\dagger}$ defined in this study was calculated using k_{cat}/K_m for T7-CUA, a full-length tRNA, and 3′-Leumin which represented a non-specific RNA with an aminoacylation rate which was reproducibly above background levels. The analysis indicates that the extra stem/loop structure makes the largest overall contribution to aminoacylation whereas the anticodon stem/loop makes essentially none. The acceptor stem makes the second largest contribution to aminoacylation and smaller contributions are attributable to the D stem/loop and TΨC stem/loop domains.

Since no pair of RNAs in the study represented a specific deletion of the D stem/loop, it was necessary to indirectly estimate the contribution of this domain. This was accomplished by calculating the difference between the $-\Delta\Delta G^{\dagger}$ (13.0 kcal/mole) for the full-length $tRNA^{Ser}$-SerRS interaction and the sum of the $-\Delta\Delta G^{\dagger}$ values that were determined for the anticodon stem/loop, extra stem/loop and acceptor-TΨC stem/loop. This

calculation indicates that the contribution of the D stem/loop is about twice that of the TΨC stem/loop. However, since G18 and G19 in the D loop and Y55 and C56 in the TΨC loop form tertiary base pairs which stabilize the folded tRNA structure, it is possible that the contribution of the D stem/loop depends on its interaction with the TΨC stem/loop. In this regard, it is interesting to note that k_{cat}/K_m for T7-ΔXL, which presents both the D and TΨC stem/loops, is about 100-fold greater than that of minihelix^{Ser} which presents only the TΨC/loop (Table 2). These RNAs seem to have similar K_ms but differ in k_{cat} suggesting that catalysis may depend on whether or not the D and TΨC stem/loops are presented together.

DISCUSSION

The ability to estimate the relative contributions of discrete tRNA domains to recognition and subsequent aminoacylation by SerRS is due, in large part, to the window of measurability (nine orders of magnitude) that we have obtained using a single assay system. This was achieved using reaction conditions which were not optimized for initial rates of aminoacylation. In fact, the initial rate of aminoacylation for minihelix^{Ser} was greater in a Tris-Cl buffer than it was in the HEPES-KOH buffer used throughout the study. However, the relatively rapid deacylation that was observed in the Tris-Cl buffer clearly illustrates that optimizing for initial rate is not necessarily desirable and could profoundly affect the outcome of this type of analysis. The addition of inorganic pyrophosphatase to the aminoacylation reaction resulted in the maintenance of a linear initial rate and increased the initial rate by a factor of about four compared to reactions without PPiase. Although it has been shown for some synthetases that PPiase can alter the kinetic pathway (30) and can affect the rates of misacylation of non-cognate tRNAs (31), its presence had a relatively small effect on the magnitude of the initial rates of aminoacylation for minihelices and did not affect the aminoacylation specificity of SerRS. Thus it is unlikely that the results presented here were biased by the use of PPiase.

The tRNA^{Ser} domains important for recognition by SerRS that have been revealed in this study are consistent with those revealed using two other methodologies. Footprinting studies of the tRNA^{Ser}-SerRS complex indicated that the tRNA phosphate backbone is protected by SerRS on the 3'-side of the acceptor stem, in the TΨC stem, and in the extra stem/loop (32). *In vivo* studies indicated that acceptor stem, D stem and extra stem/loop nucleotides contribute to serine identity (7, 8). Since the footprinting studies reveal no protection of D stem/loop nucleotides whereas the *in vivo* studies and our *in vitro* aminoacylation kinetic analysis indicate that the D stem/loop contributes to aminoacylation, it is possible that nucleotides in this region of the tRNA indirectly contribute to aminoacylation.

Our analysis clearly shows that the tRNA^{Ser} extra stem/loop domain makes the greatest contribution of all the domains to recognition by SerRS. Earlier studies of tRNA^{Ser} mutants showed that the extra stem/loop is important for the *in vivo* identity of serine tRNAs (8) and for the *in vitro* aminoacylation of tRNA^{Ser} by SerRS (27, 33). The work presented here extends these earlier studies and clearly shows that the extra stem/loop structure contributes significantly to the binding of tRNA^{Ser} to SerRS. Although k_{cat} is presumed to be the kinetic parameter primarily responsible for maintaining *in vivo* aminoacylation specificity (5, 6), our *in vitro* aminoacylation data, in conjunction with previous *in vivo* identity studies from this laboratory (8), suggest that tRNA binding can contribute to specificity.

The *in vivo* studies of Normanly *et al.* (7, 8) showed that it was essential to change the first three base pairs of the acceptor stem of tRNA^{Leu} to those conserved in tRNA^{Ser} in order to convert the *in vivo* identity of the tRNA^{Leu} amber suppressor from leucine to serine. However, it is difficult from identity swap studies alone to discern whether the changes were required to improve productive interactions with SerRS or to prevent productive interactions from occurring with a competing synthetase such as LeuRS. The *in vitro* aminoacylation kinetic results for the mini and microhelices now clearly establish that SerRS specifically recognizes the tRNA^{Ser} acceptor stem nucleotides. This result has also been confirmed by a mutational analysis in both the tRNA^{Ser} and T7-ΔXL RNA backgrounds (unpublished data). Since the acceptor stem must necessarily be in close contact with the enzyme during catalysis, recognition in this region is not surprising. Studies using mini and microhelix substrates in Schimmel's laboratory have shown that acceptor stem nucleotides of tRNA^{Ala}, tRNA^{His}, tRNA^{Gly} and tRNA^{Met} are specifically recognized by the cognate synthetase (29, 34).

Anticodon recognition is important in the majority of isoaccepting groups in *E.coli* (3). That no anticodon nucleotide is held in common among all the tRNA^{Ser} isoacceptors and that all are aminoacylated by a single SerRS led to an early proposal that the anticodon is not important for SerRS recognition (35). The results presented here are consistent with these earlier studies and clearly show that neither the three anticodon nucleotides nor the entire anticodon stem/loop domain are involved in recognition. The data also indicate that the anticodon stem/loop is an independent structural domain in tRNA^{Ser} even though it is normally stacked on the D stem/loop and is directly adjacent to the functionally important extra-stem/loop structure. Thus it seems that the nucleotides important for dictating the spatial orientation of the extra-stem/loop structure are confined to the core region of the tRNA tertiary structure and perhaps include the D loop nucleotides as well (27).

Knowing the regions of tRNA^{Ser} that are important for aminoacylation and having a high resolution crystal structure of SerRS (36) makes it possible to differentiate among potential tRNA^{Ser}-SerRS binding topologies. Initial model-building suggested that the 3' terminus of tRNA^{Ser} could bind the SerRS active site with the anticodon stem interacting with the long coiled-coil of SerRS (36). However, the results presented here rule out the possibility that the anticodon stem and loop significantly contribute to recognition by SerRS. A subsequent model proposed by Asahara *et al.* (33) places the tRNA^{Ser} extra stem/loop structure in contact with the coiled-coil of SerRS. This model is consistent with the biochemical data presented here. However it does not completely explain our observation that the D stem/loop and TΨC stem/loop also contribute to aminoacylation. The co-crystal structure of the tRNA^{Ser}-SerRS complex currently being solved by Cusack and co-workers should help to resolve this problem.

ACKNOWLEDGEMENTS

We are very grateful to John N.Abelson, in whose laboratory this work was performed, for his encouragement and support of this project. We wish to thank Giuseppe Tocchini Valentini for constructing the T7-CUA and T7-ΔXL genes and for

providing purified SerRS for our initial experiments. We also thank Barry Cooperman for supplying purified *E.coli* inorganic pyrophosphatase. This work was supported by NIH grant GM48560 (J.N.A.) and by NIH (GM13776; M.E.S.) and American Cancer Society (PF-3404; J.R.S.) postdoctoral fellowships.

REFERENCES

1. Kim, S.H., Suddath, F.L., Quigley, G.J., McPherson, A., Sussman, J.L., Wang, A.H.J., Seeman, N.C., Rich, A. (1974) *Science* **185**, 435–440.
2. Westhof, E., Dumas, P., Moras, D. (1985) *J. Mol. Biol.* **184**, 119–145.
3. Schulman, L.H. (1991) *Prog. Nucleic Acid Res. Mol. Biol.* **41**, 23–87.
4. Crothers, D.M., Seno, T., Söll, D.G. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3063–3067.
5. Schimmel, P.R., Söll, D. (1979) *Annu Rev. Biochem.* **48**, 601–648.
6. Hou, Y.-M., Schimmel, P. (1989) *Biochemistry* **28**, 4942–4947.
7. Normanly, J., Ogden, R.C., Horvath, S.J., Abelson, J. (1986) *Nature* **321**, 213–219.
8. Normanly, J., Ollick, T., Abelson, J. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5680–5684.
9. Grodberg, J., Dunn, J.J. (1988) *J. Bact.* **170**, 1245–1253.
10. Davanloo, P., Rosenberg, A.H., Dunn, J.J., Studier, F.W. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2035–2039.
11. Härtlein, M., Madern, D., Leberman, R. (1987) *Nucleic Acids Res.* **15**, 1005–1017.
12. Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
13. Sampson, J.R., Uhlenbeck, O.C. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1033–1037.
14. Milligan, J.F., Groebe, D.R., Witherell, G.W., Uhlenbeck, O.C. (1987) *Nucleic Acids Res.* **15**, 8783–8798.
15. Dock-Bregeon, A.C., Westhof, E., Giegé, R., Moras, D. (1989) *J. Mol. Biol.* **206**, 707–722.
16. McClain, W.H., Guerrier-Takada, C., Altman, S. (1987) *Science* **238**, 527–530.
17. Francklyn, C., Schimmel, P. (1989) *Nature* **337**, 478–481.
18. Tuerk, C., Gauss, P., Thermes, C., Groebe, D.R., Gayle, M., Guild, N., Stormo, G., D'Aubenton-Carafa, Y., Uhlenbeck, O.C., Tinoco, I. Jr., Brody, E.N., Gold, L. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1364–1368.
19. SantaLucia, J. Jr., Kierzek, R., Turner, D.H. (1992) *Science* **256**, 217–219.
20. Dignam, J.D., Deutscher, M.P. (1979) *Biochemistry* **18**, 3165–3170.
21. Katze, J.R., Konigsberg, W. (1970) *J. Biol. Chem.* **245**, 923–930.
22. Cooperman, B.S. (1982) *Methods Enzymol.* **87**, 526–548.
23. Eigner, E.A., Loftfield, R.B. (1974) *Methods Enzymol.* **29**, 601–619.
24. Behlen, L.S., Sampson, J.R., DiRenzo, A.B., Uhlenbeck, O.C. (1990) *Biochemistry* **29**, 2515–2523.
25. Behlen, L.S., Sampson, J.R., Uhlenbeck, O.C. (1992) *Nucleic Acids Res.* **20**, 4055–4059.
26. Peterson, E.T., Uhlenbeck, O.C. (1992) *Biochemistry* **31**, 10380–10389.
27. Himeno, H., Hasegawa, T., Ueda, T., Watanabe, K., Shimizu, M. (1990) *Nucleic Acids Res.* **18**, 6815–6819.
28. Fersht, A. (1977) Enzyme Structure and Mechanism. W. H. Freeman and Company, San Francisco, CA.
29. Francklyn, C., Shi, J.-P., Schimmel, P. (1992) *Science* **255**, 1121–1125.
30. Freist, W., Sternbach, H., Cramer, F. (1982) *Eur. J. Biochem.* **128**, 315–329.
31. Khvorova, A.M., Motorin, Y.A., Wolfson, A.D. (1992) *FEBS Lett.* **311**, 139–142.
32. Schatz, D., Leberman, R., Eckstein, F. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 6132–6136.
33. Asahara, H., Himeno, H., Shimizu, M. (1991) *Chem. Lett.* **3**, 363–366.
34. Martinis, S.A., Schimmel, P. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 65–69.
35. Sundharadas, G., Katze, J., Söll, D., Konigsberg, W., Lengyel, P. (1968) *Proc. Natl. Acad. Sci. U.S.A.* **61**, 693–700.
36. Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., Leberman, R. (1990) *Nature* **347**, 249–255.
37. Gauss, D.H., Grüter, F., Sprinzl, M. (1979) In P. R. Schimmel, D. Söll, J. N. Abelson (ed.) Transfer RNA: Structure, Properties, and Recognition. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY pp. 518–519.
38. Sprinzl, M., Dank, N., Nock, S., Schön, A. (1991) *Nucleic Acids Res.* **19 (supplement)**, 2127–2171.