The EMBO Journal vol.15 no.11 pp.2843–2849, 1996

Variant minihelix RNAs reveal sequence-specific recognition of the helical $tRNA^{Ser}$ acceptor stem by *E.coli* seryl-tRNA synthetase

Margaret E.Saks and Jeffrey R.Sampson

Division of Biology 147-75, California Institute of Technology, Pasadena, CA 91125, USA

Aminoacylation rate determinations for a series of variant RNA minihelix substrates revealed that *Escherichia coli* seryl-tRNA synthetase (SerRS) recognizes the 1–72 through 5–68 base pairs of the *E.coli* $tRNA^{Ser}$ acceptor stem with the major recognition elements clustered between positions 2–71 and 4–69. The rank order of effects of canonical base pair substitutions at each position on k_{cat}/K_m was used to assess the involvement of major groove functional groups in recognition. Conclusions based on the biochemical data are largely consistent with the interactions revealed by the refined structure of the homologous *Thermus thermophilus* $tRNA^{Ser}$–SerRS complex that Cusack and colleagues report in the accompanying paper. Disruption of an end-on hydrophobic interaction between the major groove *C*5(H) of pyrimidine 69 and an aromatic side chain of SerRS is shown to significantly decrease k_{cat}/K_m of a minihelix substrate. This type of interaction provides a means by which proteins can recognize the binary information of 'degenerate' sequences, such as the purine–pyrimidine base pairs of $tRNA^{Ser}$. The 3–70 base pair is shown to contribute to recognition by SerRS even though it is not contacted specifically by the protein. The latter effect derives from the organization of the specific contacts that SerRS makes with the neighboring 2–71 and 4–69 acceptor stem base pairs.

Keywords: RNA hairpin helix/RNA major groove/RNA–protein interaction/seryl-tRNA synthetase/tRNA acceptor stem recognition

Introduction

The specific aminoacylation of tRNA by the aminoacyl-tRNA synthetases provides a particularly attractive system for elucidating the molecular basis of RNA–protein recognition specificity. Although tRNAs share a common tertiary structure, there exists sufficient structural and sequence variation to provide the necessary information for each synthetase to recognize and aminoacylate only its set of cognate tRNAs. A variety of *in vitro* aminoacylation studies, *in vivo* amino acid identity determinations and X-ray diffraction analyses of tRNA–synthetase complexes have shown that synthetases recognize nucleotides that are most commonly located within the single-stranded anticodon loop, at the single-stranded position 73 and, with few exceptions, within the base-paired acceptor stem of their cognate tRNAs (Giegé *et al.*, 1993; McClain, 1993; Saks *et al.*, 1994; Pallanck *et al.*, 1995). Thus, synthetases can specifically recognize tRNA attributes that include single-stranded regions which present all functional groups, as well as base-paired regions, such as the acceptor stem, where there are constraints on the chemical information that is presented to the interacting synthetase.

The relationship between acceptor stem recognition and aminoacylation specificity depends on the overall binding topology of tRNA–synthetase complexes (Moras, 1992; Delarue and Moras, 1993) and on the content and accessibility of the information that is presented by the base pairs of A-form RNA helices (Seeman *et al.*, 1976; Steitz, 1990; Weeks and Crothers, 1993). In general, class I synthetases (Cusack *et al.*, 1990; Eriani *et al.*, 1990) such as glutaminyl-tRNA synthetase (GlnRS) approach the acceptor stem from the minor groove side, whereas class II synthetases such as AspRS and SerRS approach it from the major groove side. The major groove presents an array of functional groups that can be used to discriminate unambiguously among all canonical base pairs (Seeman *et al.*, 1976). However, because the major groove is deep and narrow, its information is only accessible at base pairs that are located near internal loops or helix termini (Weeks and Crothers, 1993). In fact, major groove functional groups near internal loops have been implicated in the recognition of the bovine and human immunodeficiency virus TAR RNA by the Tat protein (Weeks and Crothers, 1991; Puglisi *et al.*, 1992, 1993; Tao and Frankel, 1992; Hamy *et al.*, 1993; Chen and Frankel, 1995) and in the recognition of the human U1 spliceosomal RNA by the U1A protein (Howe *et al.*, 1994; Oubridge *et al.*, 1994). Although there are complementary biochemical and biophysical studies of the recognition of minor groove functional groups within the acceptor stem helix by a class I synthetase (Rould *et al.*, 1989; Jahn *et al.*, 1991; Hayase *et al.*, 1992), analogous studies of major groove recognition by a class II synthetase are lacking. Such studies are of particular interest given the anticipated binding topologies of class II synthetases (Ruff *et al.*, 1991; Biou *et al.*, 1994).

The use of hairpin helices and duplex RNAs that recapitulate portions of the tRNA acceptor–TΨC stem domain (including N73 and the 3′-terminal CCA) has facilitated studies of acceptor stem recognition. Because these RNAs lack recognition elements that are found in other regions of the tRNA (Giegé *et al.*, 1993; McClain, 1993; Saks *et al.*, 1994; Pallanck *et al.*, 1995), they are sometimes very poor substrates for the cognate synthetase (Martinis and Schimmel, 1995). Nevertheless, sequence-specific aminoacylation of these acceptor stem mimics is observed for alanyl- (Francklyn and Schimmel, 1989), aspartyl- (Frugier *et al.*, 1994), glutaminyl- (Wright *et al.*, 1993), glycyl- (Francklyn *et al.*, 1992), histidinyl- (Francklyn and Schimmel, 1990), isoleucyl- (Nureki *et al.*, 1993), methionyl- (Martinis and Schimmel, 1992), seryl-

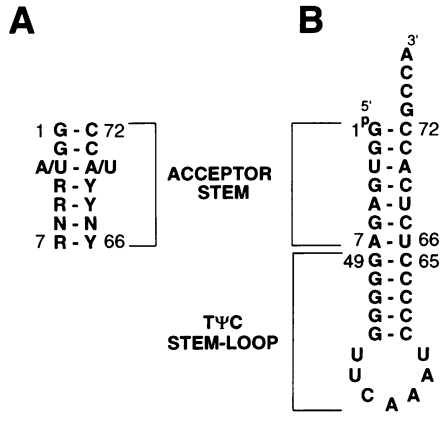

Fig. 1. (**A**) The conserved sequence of *E.coli* tRNA^{Ser} acceptor stems. R, purine (A or G); Y, pyrimidine (C or U); N, any nucleotide. The conserved acceptor stem sequence is based on the gene sequences reported for the five *E.coli* tRNA^{Ser} isoacceptors (Steinberg *et al.*, 1993). (**B**) Nucleotide sequence of the reference minihelix. The reference minihelix consists of the co-axial acceptor stem–TΨC stem-loop sequence and 3′-terminal GCCA of the *E.coli* tRNA^{Ser3} and tRNA^{Ser5} isoacceptors (anticodons GCU and GGA, respectively). Base pair substitutions in variant minihelices were restricted to the acceptor stem. Nucleotides are numbered according to the conventional tRNA numbering system (Gauss *et al.*, 1979).

(Sampson and Saks, 1993) and valyl- (Frugier *et al.*, 1992) tRNA synthetases, and indicates that recognition of the acceptor stem helix is a general property of tRNA–synthetase interactions (Schimmel *et al.*, 1993). Thus, studies of acceptor stem recognition provide a useful model with respect to the general problem of how proteins recognize A-form RNA helices.

Comparison of the acceptor stem sequences of the five *Escherichia coli* serine tRNAs reveals that they share elements that are conserved in three different ways: (i) the absolutely conserved G1–C72 and G2–C71 base pairs; (ii) the conserved chemistry of purine–pyrimidine base pairs at positions 4–69, 5–68 and 7–66; and (iii) the A–U and U–A alternatives at position 3–70 (Figure 1A). Although *in vivo* (Normanly *et al.*, 1986, 1992) and *in vitro* studies (Sampson and Saks, 1993; Asahara *et al.*, 1994) have implicated both the 2–71 and 3–70 base pairs in recognition by SerRS, the extent of acceptor stem recognition and the modes of interactions between the base pairs and SerRS were not addressed. To investigate further the details of acceptor stem recognition and to elucidate potential base functional groups that account for specificity, we synthesized a defined set of variant minihelix RNAs and determined their *in vitro* aminoacylation rates with purified SerRS. The aminoacylation data and their implications with respect to the molecular mechanisms that account for recognition specificity are discussed in the light of the co-crystal structure for the related *Thermus thermophilus* tRNA^{Ser}–SerRS complex that is presented by Cusack and colleagues in the accompanying paper (Cusack *et al.*, 1996).

Results

A wild-type serine minihelix RNA (Figure 1B) and a set of 21 RNA variants representing a canonical base pair substitution at each acceptor stem position were synthesized by *in vitro* transcription to examine SerRS recognition of the tRNA^{Ser} acceptor stem sequence. Two additional sets of minihelices having multiple base pair substitutions in the acceptor stem were also synthesized to investigate the relative contributions of nucleotide chemistry and helix geometry to recognition by SerRS. The initial rates of aminoacylation were determined under the buffer conditions described in Materials and methods, using 50 μM RNA and catalytic amounts of SerRS (1.3 μM). We have shown previously that the K_m for a minihelix having a wild-type serine acceptor stem sequence is ~240 μM (Sampson and Saks, 1993). Thus, although the individual kinetic parameters were not determined for each substrate, V_0/[S] should approximate k_{cat}/K_m under conditions where the RNA concentration is well below K_m, as it is in these experiments. For most of the substrates, SerRS had participated in less than one turnover event during the course of the measurements. However, we have shown previously for a wild-type serine minihelix that the rate determined from early time points accurately reflects the multiple turnover rate (Sampson and Saks, 1993).

Figure 2 shows that k_{cat}/K_m of aminoacylation is reduced for all base pair substitutions that deviate from the conserved tRNA^{Ser} acceptor stem sequence (Figure 1A). The largest decreases in k_{cat}/K_m were observed when the conserved G–C base pair at position 2–71 or the degenerate R–Y base pair at position 4–69 were changed. In addition, deviations from the wild-type A–U and U–A base pairs at position 3–70 substantially decreased k_{cat}/K_m, with the G–C substitution having the largest effect when compared with the wild-type A–U. Somewhat smaller but substantial effects were also revealed for substitutions at positions 1–72 and 5–68. Thus, correct recognition of the tRNA^{Ser} acceptor stem depends on its conformity to the sequence that is conserved and clustered at positions 2–71, 3–70 and 4–69, with smaller yet significant contributions from the base pairs at positions 1–72 and 5–68.

The different rank order of effects of base pair substitutions at each acceptor stem position indicates that SerRS uses different molecular mechanisms to recognize these base pairs (Table I). Base pair substitutions at positions 1–72 and 2–71 had similar effects in that any change away from the wild-type sequence caused a significant decrease in k_{cat}/K_m and each canonical base pair substitution had a discrete effect, giving an overall rank order for k_{cat}/K_m of G–C>>A–U>U–A>C–G. This type of complete discrimination among base pairs is expected when proteins recognize at least two major groove functional groups of a base pair (Seeman *et al.*, 1976). In contrast, for substitutions at positions 4–69 and 5–68, k_{cat}/K_m was reduced only when U–A and C–G base pairs were introduced, giving an overall rank order for k_{cat}/K_m of G–C≈A–U>U–A>C–G. The insensitivity of k_{cat}/K_m to the identity of the R–Y base pair is inconsistent with classical expectations for both major and minor groove recognition (Seeman *et al.*, 1976). Finally, k_{cat}/K_m reflected a gradient of effects for base pair substitutions at position 3–70, having a rank order of A–U>U–A>C–G>U·G>G·U>G–C. Major groove recognition can be ruled out because the wild-type A–U and U–A base pairs present completely different functional groups. In addition, minor groove recognition is unlikely because the C–G substitution, which introduced a minor groove exocyclic amino

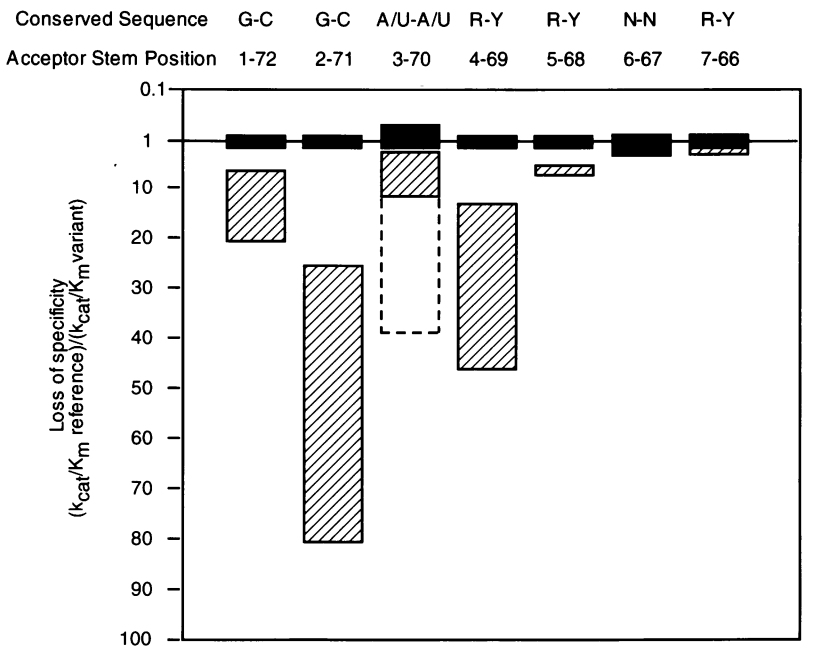

Fig. 2. Relationship between the conserved sequence of *E.coli* $tRNA^{Ser}$ acceptor stems and the effects of acceptor stem base pair substitutions in minihelix RNAs on k_{cat}/K_m of aminoacylation by SerRS. The black boxes correspond to the range of effects on k_{cat}/K_m for substitutions which conform to the conserved sequence at each acceptor stem position. The hatched boxes correspond to the range of effects on k_{cat}/K_m for substitutions which differ from the conserved sequence. The dashed box indicates the effect on k_{cat}/K_m for the G3–C70 substitution relative to the A3–U70 base pair.

group, reduced k_{cat}/K_m by only 1.9-fold compared with the wild-type U–A. These inconsistencies with respect to classical models of sequence-specific recognition, in conjunction with the fairly undifferentiated effects for each base pair substitution on k_{cat}/K_m, suggest that the 3–70 base pair contributes indirectly to SerRS recognition.

The acceptor stem sequences of the five *E.coli* tRNAs and the results discussed above for the single base pair substitutions at positions 4–69 and 5–68 indicate that some aspect of the R–Y base pairs at these positions is important for recognition by SerRS. However, since not all R4–Y69/R5–Y68 combinations are represented in nature, the possibility exists that only a subset promote aminoacylation by SerRS. To test this possibility, we synthesized additional minihelicies having all possible R4–Y69/R5–Y68 combinations embedded in both the A3–U70 and U3–A70 background. As shown in Figure 3A, the k_{cat}/K_m for each of these minihelices was within a factor of 2 with respect to its reference minihelix. Thus, SerRS is insensitive to both the identity and combination of R–Y base pairs at these two positions. Moreover, because the predicted base stacking free energies (Freier *et al.*, 1986) for this region of the helix differ significantly among these substrates, it seems unlikely that SerRS requires a defined helix stability at these positions.

An additional set of minihelices was synthesized (Figure 3B) to examine whether subtle changes in helix geometry or changes in base functional groups account for the observed effects of Y4–R69 base pairs on aminoacylation by SerRS. The set had base pair substitutions at position 4–69 embedded within a minihelix having an A3–U70 and a G5–C68 base pair. The G–C, A–U and G·U substitutions at position 4–69 conform to the conserved R–Y chemistry whereas the C–G does not. The effect of the G·U substitution was of particular interest since this wobble base pair should introduce a slight perturbation into the normal A-form helical geometry while maintaining the R–Y nucleotide chemistry. Whereas the G4·U69 substitution decreased k_{cat}/K_m by 9.3-fold, the C4–G69 substitution resulted in a much more dramatic decrease of 177-fold. Thus, these results support the idea that some aspect of R–Y nucleotide chemistry is more important to recognition by SerRS than is helix geometry.

Discussion

The biochemical results presented in this paper indicate that the acceptor stem base pairs that are recognized by SerRS begin at position 1–72 and extend to position 5–68, with the major recognition cluster located between positions 2–71 and 4–69. Several factors were considered in using the biochemical data to draw conclusions about the mechanisms by which SerRS recognizes these base pairs. Based on previous $tRNA^{Ser}$ footprinting (Schatz *et al.*, 1991) and X-ray crystallographic studies of the $tRNA^{Ser}$–SerRS complex (Biou *et al.*, 1994), it was clear that SerRS primarily binds from the variable loop side of the tRNA. This orients SerRS toward the major rather than the minor groove of the acceptor stem helix. Although each base pair presents a distinctive array of major groove hydrogen bond donors, acceptors and hydrophobic groups (Seeman *et al.*, 1976), their accessibility to amino acid side chains is predicted to be greater at positions 1–72 and 2–71 than at positions 4–69 and 5–68 due to differences in the proximity of each base pair to the helix terminus (Weeks and Crothers, 1993). Thus, based on the topology of the $tRNA^{Ser}$–SerRS complex and aspects of the geometry of A-form RNA helices, different molecular mechanisms were expected to dictate the recognition of

Table I. Magnitude of the effect of base pair substitutions in minihelix RNAs on the k_{cat}/K_m of aminoacylation by SerRS

Acceptor stem base pair	k_{cat}/K_m (/s/M)	Loss of specificity
G1–C72	12	1.0
A1–U72	1.9	6.2
U1–A72	0.97	12
C1–G72	0.60	20
G2–C71	12	1.0
A2–U71	0.45	26
U2–A71	0.20	58
C2–G71	0.15	80
U3–A70	12	1.0
A3–U70	42	(3.6)
C3–G70	6.1	1.9
U3–G70	4.2	2.8
G3–U70	1.7	6.7
G3–C70	1.1	11
G4–C69	12	1.0
A4–U69	13	(1.1)
U4–A69	0.98	12
C4–G69	0.26	46
A5–U68	12	1.0
G5–C68	25	(2.1)
U5–A68	2.2	5.4
C5–G68	1.9	6.3
G6–C67	12	1.0
A6–U67	6.8	1.7
U6–A67	3.8	2.6
C6–G67	4.1	3.1
A7–U66	12	1.0
G7–C66	9.3	1.3
U7–A66	11	1.0
C7–G66	11	1.1

The sequence of the reference minihelix at each acceptor stem position is given in bold with the base pair substitutions indented below. The loss of specificity is calculated as $(k_{cat}/K_{m\ reference})/(k_{cat}/K_{m\ variant})$ for each base pair substitution with the value for the reference minihelix set at 1.0. Parentheses indicate increases in k_{cat}/K_m relative to the reference minihelix.

the 1–72 and 2–71 as compared with the 4–69 and 5–68 base pairs.

The effects of base pair substitutions at positions 1–72 and 2–71 were similar in that canonical base pair substitutions had discrete effects on k_{cat}/K_m that followed the same rank order. This suggests that SerRS interacts with similar base functional groups at these two acceptor stem positions. However, at position 2–71, changing the wild-type G–C base pair to A–U reduced k_{cat}/K_m by 26-fold, whereas the same change at position 1–72 decreased k_{cat}/K_m by only 6-fold. At both positions there was an additional 2- to 4-fold decrease in k_{cat}/K_m for Y–R substitutions. The G–C to A–U substitution reverses the polarity of the hydrogen bond acceptors and donors on the inside of the major groove without altering the *N*7 and *C*5(H) on its outside. Thus, the observed decrease in k_{cat}/K_m for the A–U substitution could reflect a disruption of contacts between SerRS and either or both the carbonyl oxygen and exocyclic amino groups of the wild-type G–C base pair. The *T.thermophilus* $tRNA^{Ser}$–SerRS co-crystal structure (Cusack *et al.*, 1996) shows that S261 and F262 make specific contacts with these two functional groups of the G2–C71 base pair. In addition, S261 is within hydrogen bonding distance of G1; but no other interactions are apparent with respect to the 1–72 base pair. It is likely that the similar rank order but different magnitudes of effects for base pair substitutions at these two positions are due to differences in the extent to which S261 interacts with the carbonyl oxygens of G1 and G2 and to the fact that SerRS contacts both pairing partners at 2–71 but only G1 at position 1–72.

Although helix geometry may make a small contribution to aminoacylation by SerRS, some aspect of the chemistry of R–Y base pairs is primarily responsible for the observed effects of base pair substitutions at position 4–69. The purine *N*7 and pyrimidine *C*5(H) functional groups on the outside of the major groove provide a basis for SerRS to distinguish between R–Y and Y–R base pairs. Although recognition of either functional group could account for our aminoacylation data, earlier footprinting studies of the *E.coli* $tRNA^{Ser}$–SerRS complex showed protection on the 3′ strand of the helix. Moreover, chemical probing experiments of model RNA helices indicate that, at the fourth position from the helix termini, the major groove base functional groups associated with the 3′ but not the 5′ strand of the helix should be accessible to interacting proteins (Weeks and Crothers, 1993). Together, these studies argue in favor of a direct interaction between SerRS and the *C*5(H) at positions 68 and 69. In fact, the *T.thermophilus* $tRNA^{Ser}$–SerRS complex shows that the aromatic ring of F262 is positioned perpendicular to the plane of the pyrimidine base at position 69 and is directly over its *C*5(H) functional group, creating an end-on hydrophobic interaction (Cusack *et al.*, 1996). A similar interaction is expected in the *E.coli* cognate complex between Y69 and the homologous Y274 (Cusack *et al.*, 1996).

Although hydrophobic ring–ring interactions between nucleotides and amino acid functional groups have been described in other systems (Betts *et al.*, 1994), to our knowledge this is the first study of the biochemical contribution of the *C*5(H) to such an interaction when presented in a structured RNA helix. Our aminoacylation data indicate that the hydrophobic contact between SerRS and Y69 contributes to the overall aminoacylation reaction to an extent similar to that derived from the network of hydrogen bonding interactions with the 2–71 base pair. Given that the analogous *C*5(CH_3) hydrophobic group in thymidine residues of DNA is known to contribute to DNA–protein recognition (Luisi *et al.*, 1991; Mazzarelli *et al.*, 1992), and that the average length of an A-form RNA helix is relatively short and is thus amenable to major groove recognition, it is likely that a large number of proteins will utilize the *C*5(H) and/or the *N*7 functional groups as a source of binary information to discriminate between R–Y and Y–R base pairs.

Previous studies have shown that the identity of the 4–69 base pair can affect recognition by IleRS (Nureki *et al.*, 1993, 1994) and ThrRS (Hasegawa *et al.*, 1992). In addition, MetRS may inspect the minor groove of the 4–69 base pair such that the exocyclic amino group of C–G in non-cognate tRNAs functions as an anti-determinant that discourages misacylation (Martinis and Schimmel, 1993). Although the molecular mechanisms that account for the observed effects in these three systems have not yet been elucidated by a structural analysis, both MetRS

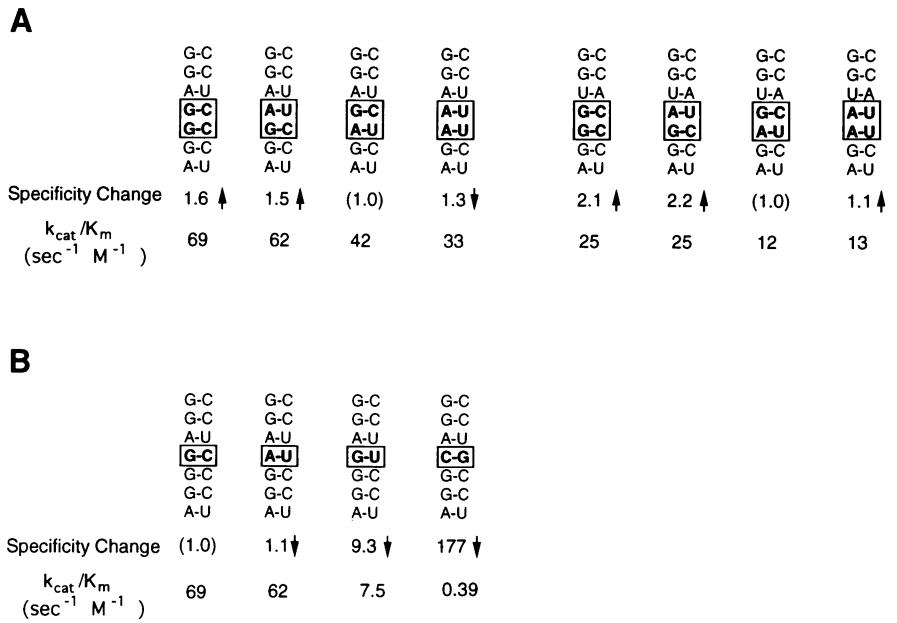

Fig. 3. Contributions of helix geometry and nucleotide chemistry at positions 4–69 and 5–68 of the minihelix acceptor stem to recognition by SerRS. (**A**) Effects of purine–pyrimidine (R–Y) base pair combinations at positions 4–69 and 5–68 on the k_{cat}/K_m of aminoacylation by SerRS in minihelices having either an A3–U70 or U3–A70. (**B**) The effects of altering helix geometry versus nucleotide chemistry on the k_{cat}/K_m of aminoacylation by SerRS. The magnitude of the effect on k_{cat}/K_m of boxed base pairs was calculated relative to the reference minihelix as $(k_{cat}/K_{m\ reference})/(k_{cat}/K_{m\ variant})$. The value for the reference minihelix was set at 1.0. Increases and decreases in k_{cat}/K_m relative to the reference minihelix are respectively indicated by upward- and downward-pointing arrows.

and IleRS are class I synthetases and are expected to approach the acceptor stem from the minor groove side. Thus, access to the 4–69 base pair should be fairly unrestricted. In contrast, ThrRS falls into the same subcategory of class II synthetases as that of SerRS (Cusack *et al.*, 1990; Eriani *et al.*, 1990). Thus, ThrRS is expected to approach the acceptor stem from the major groove side where helix geometry limits the accessibility of functional groups. Although the molecular mechanism that accounts for the contribution of the 4–69 base pair to recognition by ThrRS is not yet known, based on the available $tRNA^{Thr}$ and ThrRS sequences, it is possible that the recognition of the 4–69 base pair of $tRNA^{Thr}$ will involve hydrogen bonding rather than hydrophobic interactions.

The biochemical results indicate that the 3–70 base pair contributes to recognition by SerRS. However, this contribution appears to be indirect rather than direct since the effects of base pair substitutions on k_{cat}/K_m are not consistent with a simple model of either major or minor groove recognition. The network of interactions that are revealed in the structure of the *T.thermophilus* $tRNA^{Ser}$–SerRS complex (Cusack *et al.*, 1996) provides an explanation for how indirect effects might arise from substitutions at position 3–70 (Figure 4). The structure shows that F262 provides a hydrogen bond acceptor for interaction with the exocyclic amino group of C71 and that its aromatic ring participates in a hydrophobic interaction with the *C5*(H) of Y69. To accomplish this, F262 must span a distance of ~6 Å in the narrow major groove. Although the 3–70 base pair is not contacted directly, it is 'sandwiched' between the C71 and Y69 contact points. Consequently, any perturbation at position 3–70 could affect the direct contacts between SerRS and the neighboring base pairs. Interestingly, with the exception of minihelices having either a G3·U70 or U3·G70 base pair which introduce

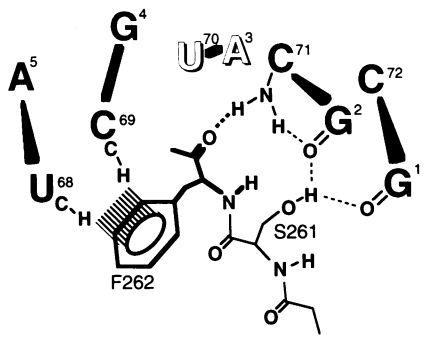

Fig. 4. Diagrammatic representation of the interactions between SerRS and the $tRNA^{Ser}$ acceptor stem base pairs as elucidated by the crystal structure of the *T.thermophilus* $tRNA^{Ser}$–SerRS complex that is discussed in detail in the accompanying paper by Cusack *et al.* The critical F262 residue that specifically interacts with two non-contiguous base pairs is shown in bold. The intervening A3–U70 base pair that indirectly affects recognition of the acceptor stem is shown in outline.

helical distortions, there is a tendency for the better substrates to have a less favorable predicted base stacking free energy ($-\Delta G$) when the value is calculated within the context of the wild-type G2–C71 and G4–C69 base pairs (Freier *et al.*, 1986). This would argue that some degree of flexibility is required to facilitate the correct orientation of F262 with respect to the 2–71 and 4–69 base pairs.

The results that are obtained using model RNA substrates such as minihelices sometimes differ from those obtained using full-length tRNA variants. For example, large contributions of G73 and the first three acceptor stem base pairs to recognition by GlnRS were revealed by *in vitro* aminoacylation determinations using variants of a full-length $tRNA^{Gln}$ (Jahn *et al.*, 1991; Hayase *et al.*, 1992; Sherman *et al.*, 1995). However, when RNA hairpin

helices were used, G73 retained its large contribution to aminoacylation but the relative contributions of the 2–71 and 3–70 base pairs were diminished (Wright *et al.*, 1993). In contrast, the contribution of G73 to recognition by yeast AspRS was essentially identical when tested in both RNA hairpin helix and full-length $tRNA^{Asp}$ substrates (Frugier *et al.*, 1994). Similarly, the G2–C71 of $tRNA^{Ser}$ was identified as a major recognition element for SerRS using both the minihelix (this work) and $tRNA^{Ser}$ substrates (Asahara *et al.*, 1994; J.R.Sampson and M.E.Saks, unpublished observation). However, although the 3–70 and 4–69 base pairs of $tRNA^{Ser}$ make substantial contributions to aminoacylation specificity in the minihelix substrate (this work), they make negligible contributions to aminoacylation when presented in a full-length $tRNA^{Ser}$ (J.R.Sampson and M.E.Saks, unpublished observation).

The biochemical data that were obtained for base pair substitutions at positions 3–70 and 4–69 in the minihelix and full-length $tRNA^{Ser}$ present a paradox with respect to the definition of recognition elements and the ways in which they contribute to the overall aminoacylation reaction. The 3–70 and 4–69 base pairs would be excluded from the SerRS recognition set if recognition elements (Normanly and Abelson, 1989; Schulman, 1991) are strictly limited to those nucleotides which have large effects on aminoacylation in the full-length tRNA. However, they deserve serious consideration based on the co-crystal structure for the full-length $tRNA^{Ser}$–SerRS complex and the biochemical data for the minihelix substrates. Based on these observations, the 3–70 and 4–69 base pairs are 'cryptic' recognition elements with respect to a steady-state kinetic analysis of aminoacylation of the full-length tRNA by SerRS. Clearly, a more detailed analysis of RNA body-dependent acceptor stem recognition is necessary in order to understand fully the mechanism and biological significance of these types of 'cryptic' recognition elements.

Materials and methods

Materials

All oligodeoxyribonucleotides were synthesized by the California Institute of Technology Biopolymer Synthesis and Analysis Resource Center on an Applied Biosystems DNA synthesizer (Foster City, CA). L-[3-^{3}H] serine (sp. act. 29–33 Ci/mmol) was purchased from Amersham. Seryl-tRNA synthetase was purified to a specific activity of 1200 U/mg protein as previously described (Sampson and Saks, 1993) from the *E.coli* strain JM101 harboring the plasmid pSer2 which was kindly provided by M.Härtlein and R.Leberman (Härtlein *et al.*, 1987). T7 RNA polymerase was purified to a specific activity of 650 000 U/mg protein from the overproducing *E.coli* strain BL21 harboring the plasmid pAR1219 (Davanloo *et al.*, 1984). Inorganic pyrophosphatase was kindly provided by B.Cooperman.

RNA synthesis and purification

The minihelices were synthesized by *in vitro* transcription using T7 RNA polymerase and synthetic DNA templates (Milligan *et al.*, 1987) and purified to single nucleotide resolution by electrophoresis on 15% denaturing polyacrylamide gels. The conditions for the transcription, purification and quantification of the RNA products are all described in detail elsewhere (Sampson and Saks, 1993).

Aminoacylation reactions

All RNAs were reannealed prior to aminoacylation by heating to 90°C for 3 min and snap cooling on ice. The initial rate for each minihelix was determined in 50 µl reaction mixtures containing 30 mM HEPES–KOH (pH 7.4), 15 mM $MgCl_2$, 25 mM KCl, 2 mM dithiothreitol, 2 mM ATP, 40 µM [^{3}H]serine (sp. act. 20 Ci/mmol), 0.02 U/ml inorganic pyrophosphatase, 50 µM RNA and 1.3 µM SerRS. The reactions were incubated at 37°C and, at 2 min time intervals, 8 µl aliquots were spotted on Whatmann 3MM paper and washed twice with 10% trichloroacetic acid (TCA), three times with 5% TCA and once with 95% ethanol (Sampson and Saks, 1993). The initial rates were calculated from a linear least squares analysis of the data and the correlation coefficients for all rates were 0.99 or greater. The aminoacylation rate for each minihelix was determined at least twice, and the standard error of each mean rate was determined. The standard error ranged from 0.09 to 6.72% of the mean rate for each minihelix, giving an overall average standard error for all minihelices in the study of 2.5%. The k_{cat}/K_m that was estimated at 42/s/M based on initial velocity measurements for the A3–U70 minihelix (Table I) is consistent with the k_{cat}/K_m for this substrate which was calculated from measurements of the individual kinetic parameters (Sampson and Saks, 1993).

Acknowledgements

We are very grateful to John Abelson for providing the laboratory and intellectual support that was critical to the development of this project. We also wish to thank Stephen Cusack for engaging in discussions of our biochemical data throughout the past few years, for keeping us apprised of his structure of the $tRNA^{Ser}$–SerRS complex and for helping us to incorporate his structural data into Figure 4 of our paper. Reuben Leberman and Michael Härtlein kindly provided us with their seryl-tRNA synthetase overproducing strain and Barry Cooperman kindly provided purified *E.coli* inorganic pyrophosphatase. This work was supported by NIH grant GM48560 (J.A.) and by NIH (GM13776; M.E.S.) and American Cancer Society (PF-3404; J.R.S.) postdoctoral fellowships.

References

Asahara,H., Himeno,H., Tamura,K., Nameki,N., Hasegawa,T. and Shimizu,M. (1994) *Escherichia coli* seryl-tRNA synthetase recognizes $tRNA^{Ser}$ by its characteristic tertiary structure. *J. Mol. Biol.*, **236**, 738–748.

Betts,L., Xiang,S.B., Short,S.A., Wolfenden,R. and Carter,C.W. (1994) Cytidine deaminase. The 2.3 Å crystal structure of an enzyme:transition state analog complex. *J. Mol. Biol.*, **235**, 635–656.

Biou,V., Yaremchuk,A., Tukalo,M. and Cusack,S. (1994) The 2.9 Å crystal structure of *T.thermophilus* seryl-tRNA synthetase complexed with $tRNA^{Ser}$. *Science*, **263**, 1404–1410.

Chen,L. and Frankel,A.D. (1995) A peptide interaction in the major groove of RNA resembles protein interactions in the minor groove of DNA. *Proc. Natl Acad. Sci. USA*, **92**, 5077–5081.

Cusack,S., Berthet-Colominas,C., Härtlein,M. and Nassar,N. (1990) A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å. *Nature*, **347**, 249–255.

Cusack,S., Yaremchuk,A. and Tukalo,M. (1996) The crystal structure of the ternary complex of *T.thermophilus* seryl-tRNA synthetase with $tRNA^{Ser}$ and a seryl-adenylate analogue reveals a conformational switch in the active site. *EMBO J.*, **15**, 2834–2842.

Davanloo,P., Rosenberg,A.H., Dunn,J.J. and Studier,F.W. (1984) Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proc. Natl Acad. Sci. USA*, **81**, 2035–2039.

Delarue,M. and Moras,D. (1993) The aminoacyl-transfer RNA synthetase family: modules at work. *BioEssays*, **15**, 675–687.

Eriani,G., Delarue,M., Poch,O., Gangloff,J. and Moras,D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature*, **347**, 203–206.

Francklyn,C. and Schimmel,P. (1989) Aminoacylation of RNA minihelices with alanine. *Nature*, **337**, 478–481.

Francklyn,C. and Schimmel,P. (1990) Enzymatic aminoacylation of an eight-base pair microhelix with histidine. *Proc. Natl Acad. Sci. USA*, **87**, 8655–8659.

Francklyn,C., Shi,J.-P. and Schimmel,P. (1992) Overlapping nucleotide determinants for specific aminoacylation of RNA microhelices. *Science*, **255**, 1121–1125.

Freier,S.M., Kierzek,R., Jaeger,J.A., Sugimoto,N., Caruthers,M.H., Neilson,T. and Turner,D.H. (1986) Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl Acad. Sci. USA*, **83**, 9373–9377.

Frugier,M., Florentz,C. and Giegé,R. (1992) Anticodon-independent aminoacylation of an RNA minihelix with valine. *Proc. Natl Acad. Sci. USA*, **89**, 3990–3994.

Frugier,M., Florentz,C. and Giegé,R. (1994) Efficient aminoacylation of resected RNA helices by class II aspartyl-tRNA synthetase dependent on a single nucleotide. *EMBO J.*, **13**, 2218–2226.

Gauss,D.H., Grüter,F. and Sprinzl,M. (1979) Proposed numbering system of nucleotides in tRNAs based on yeast $tRNA^{Phe}$. In Schimmel,P.R., Söll,D. and Abelson,J.N. (eds), *Transfer RNA: Structure, Properties and Recognition*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 518–519.

Giegé,R., Puglisi,J.D. and Florentz,C. (1993) tRNA structure and aminoacylation efficiency. *Prog. Nucleic Acid Res. Mol. Biol.*, **45**, 129–206.

Hamy,F., Asseline,U., Grasby,J., Iwai,S., Pritchard,C., Slim,G., Butler,P.J.G., Karn,J. and Gait,M.J. (1993) Hydrogen-bonding contacts in the major groove are required for human immunodeficiency virus type-1 tat protein recognition of TAR RNA. *J. Mol. Biol.*, **230**, 111–123.

Härtlein,M., Madern,D. and Leberman,R. (1987) Cloning and characterization of the gene for *Escherichia coli* seryl-tRNA synthetase. *Nucleic Acids Res.*, **15**, 1005–1017.

Hasegawa,T., Miyano,M., Himeno,H., Sano,Y., Kimura,K. and Shimizu,M. (1992) Identity determinants of *E.coli* threonine tRNA. *Biochem. Biophys. Res. Commun.*, **184**, 478–484.

Hayase,Y., Jahn,M., Rogers,M.J., Sylvers,L.A., Koizumi,M., Inoue,H., Ohtsuka,E. and Söll,D. (1992) Recognition of bases in *Escherichia coli* $tRNA^{Gln}$ by glutaminyl-tRNA synthetase: a complete identity set. *EMBO J.*, **11**, 4159–4165.

Howe,P.W.A., Nagai,K., Neuhaus,D. and Varani,G. (1994) NMR studies of U1 snRNA recognition by the N-terminal RNP domain of the human U1A protein. *EMBO J.*, **13**, 3873–3881.

Jahn,M., Rogers,M.J. and Söll,D. (1991) Anticodon and acceptor stem nucleotides in $tRNA^{Gln}$ are major recognition elements for *E.coli* glutaminyl-tRNA synthetase. *Nature*, **352**, 258–260.

Luisi,B.F., Xu,W.X., Otwinowski,Z., Freedman,L.P., Yamamoto,K.R. and Sigler,P.B. (1991) Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature*, **352**, 497–505.

Martinis,S.A. and Schimmel,P. (1992) Enzymatic aminoacylation of sequence-specific RNA minihelices and hybrid duplexes with methionine. *Proc. Natl Acad. Sci. USA*, **89**, 65–69.

Martinis,S.A. and Schimmel,P. (1993) Microhelix aminoacylation by a class I tRNA synthetase. *J. Biol. Chem.*, **268**, 6069–6072.

Martinis,S.A. and Schimmel,P. (1995) Small RNA oligonucleotide substrates for specific aminoacylations. In Söll,D. and RajBhandary,U.L. (eds), *tRNA: Structure, Biosynthesis and Function*. ASM Press, Washington, DC, pp. 349–370.

Mazzarelli,J.M., Rajur,S.B., Iadarola,P.L. and McLaughlin,L.W. (1992) Interactions between the *trp* repressor and its operator sequence as studied by base analogue substitution. *Biochemistry*, **31**, 5925–5936.

McClain,W.H. (1993) Rules that govern tRNA identity in protein synthesis. *J. Mol. Biol.*, **234**, 257–280.

Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, **15**, 8783–8798.

Moras,D. (1992) Structural and functional relationships between aminoacyl-tRNA synthetases. *Trends Biochem. Sci.*, **17**, 159–164.

Normanly,J. and Abelson,J. (1989) tRNA identity. *Annu. Rev. Biochem.*, **58**, 1029–1049.

Normanly,J., Ogden,R.C., Horvath,S.J. and Abelson,J. (1986) Changing the identity of a transfer RNA. *Nature*, **321**, 213–219.

Normanly,J., Ollick,T. and Abelson,J. (1992) Eight base changes are sufficient to convert a leucine-inserting tRNA into a serine-inserting tRNA. *Proc. Natl Acad. Sci. USA*, **89**, 5680–5684.

Nureki,O., Niimi,T., Muto,Y., Kanno,H., Kohno,T., Muramatsu,T., Kawai,G., Miyazawa,T., Giegé,R., Florentz,C. and Yokoyama,S. (1993) Conformational change of tRNA upon interaction of the identity-determinant set with aminoacyl-tRNA synthetase. In Nierhaus,K.H., Franceschi,F., Subramanian,A.R., Erdmann,V.A. and Wittmann-Liebold,B. (eds), *The Translational Apparatus*. Plenum Press, New York, pp. 59–66.

Nureki,O., Niimi,T., Muramatsu,T., Kanno,H., Kohno,T., Florentz,C., Giegé,R. and Yokoyama,S. (1994) Molecular recognition of the identity-determinant set of isoleucine transfer RNA from *Escherichia coli*. *J. Mol. Biol.*, **236**, 710–724.

Oubridge,C., Ito,N., Evans,P.R., Teo,C.-H. and Nagai,K. (1994) Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432–438.

Pallanck,L., Pak,M. and Schulman,L.H. (1995) tRNA discrimination in aminoacylation. In Söll,D. and RajBhandary,U.L. (eds), *tRNA: Structure, Biosynthesis and Function*. ASM Press, Washington, DC, pp. 371–394.

Puglisi,J.D., Tan,R., Calnan,B.J., Frankel,A.D. and Williamson,J.R. (1992) Conformation of the TAR RNA–arginine complex by NMR spectroscopy. *Science*, **257**, 76–80.

Puglisi,J.D., Chen,L., Frankel,A.D. and Williamson,J.R. (1993) Role of RNA structure in arginine recognition of TAR RNA. *Proc. Natl Acad. Sci. USA*, **90**, 3680–3684.

Rould,M.A., Perona,J.J., Söll,D. and Steitz,T.A. (1989) Structure of *E.coli* glutaminyl-tRNA synthetase complexed with $tRNA^{Gln}$ and ATP at 2.8 Å resolution. *Science*, **246**, 1135–1142.

Ruff,M., Krishnaswamy,S., Boeglin,M., Poterszman,A., Mitschler,A., Podjarny,A., Rees, B., Thierry,J.C. and Moras,D. (1991) Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with $tRNA^{Asp}$. *Science*, **252**, 1682–1689.

Saks,M.E., Sampson,J.R. and Abelson,J.N. (1994) The transfer RNA identity problem: a search for rules. *Science*, **263**, 191–197.

Sampson,J.R. and Saks,M.E. (1993) Contributions of discrete $tRNA^{Ser}$ domains to aminoacylation by *E.coli* seryl-tRNA synthetase: a kinetic analysis using model RNA substrates. *Nucleic Acids Res.*, **21**, 4467–4475.

Schatz,D., Leberman,R. and Eckstein,F. (1991) Interaction of *Escherichia coli* $tRNA^{Ser}$ with its cognate aminoacyl-tRNA synthetase as determined by footprinting with phosphorothioate-containing tRNA transcripts. *Proc. Natl Acad. Sci. USA*, **88**, 6132–6136.

Schimmel,P., Giegé,R., Moras,D. and Yokoyama,S. (1993) An operational RNA code for amino acids and possible relationship to genetic code. *Proc. Natl Acad. Sci. USA*, **90**, 8763–8768.

Schulman,L.H. (1991) Recognition of tRNAs by aminoacyl-tRNA synthetases. *Prog. Nucleic Acid Res. Mol. Biol.*, **41**, 23–87.

Seeman,N.C., Rosenberg,J.M. and Rich,A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl Acad. Sci. USA*, **73**, 804–808.

Sherman,J.M., Rogers,M.J. and Söll,D. (1995) Recognition in the glutamine tRNA system: from structure to function. In Söll,D. and RajBhandary,U.L. (eds), *tRNA: Structure, Biosynthesis and Function*. ASM Press, Washington, DC, pp. 395–409.

Steinberg,S., Misch,A. and Sprinzl,M. (1993) Compilation of transfer RNA sequences and sequences of transfer RNA genes. *Nucleic Acids Res.*, **21**, 3011–3015.

Steitz,T.A. (1990) Structural studies of protein–nucleic acid interaction: the sources of sequence-specific binding. *Q. Rev. Biophys.*, **23**, 205–280.

Tao,J. and Frankel,A.D. (1992) Specific binding of arginine to TAR RNA. *Proc. Natl Acad. Sci. USA*, **89**, 2723–2726.

Weeks,K.M. and Crothers,D.M. (1991) RNA recognition by Tat-derived peptides: interaction with the major groove? *Cell*, **66**, 577–588.

Weeks,K.M. and Crothers,D.M. (1993) Major groove accessibility of RNA. *Science*, **261**, 1574–1577.

Wright,D.J., Martinis,S.A., Jahn,M., Söll,D. and Schimmel,P. (1993) Acceptor stem and anticodon RNA hairpin helix interactions with glutamine tRNA synthetase. *Biochimie*, **75**, 1041–1049.

Received on January 10, 1996; revised on Febuary 5, 1996