Pergamon

Bioorganic & Medicinal Chemistry, Vol. 5, No. 6, pp 1071–1079, 1997

PII: S0968-0896(97)00043-6
0968-0896/97 $17.00 + 0.00

RNA–RNA Interactions Between Oligonucleotide Substrates for Aminoacylation

Barry S. Henderson† and Paul Schimmel*
Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Abstract—RNA stem-loop microhelices with helix sequences based on tRNA acceptor stems can be charged with specific amino acids. Experiments were designed to test the possibility that microhelices could laterally associate through complementary loop sequences and thereby bring their attached aminoacyl groups close enough together to form a peptide bond. Computer simulations suggested that formation of such complexes would be sensitive to the number of loop nucleotides needed to span the grooves of the quasi-continuous helix of the intermolecular pseudoknot so formed. These predictions were confirmed experimentally by observation of complex formation sensitivity to loop size. Complexes with optimized loop sizes had apparent bimolecular dissociation constants of approximately 100 nM with only three complementary base pairs between the respective loops. Single nucleotide substitutions that disrupted the predicted intermolecular loop–loop base-pairing abolished detectable association. Similarly, placing a gap between the short helix formed by loop–loop pairing and the adjacent acceptor stems also diminished complex formation. These experiments establish an experimental basis for microhelix association for peptide synthesis. © 1997 Elsevier Science Ltd.

Introduction

Protein synthesis has probably not always involved the elaborate system found in nature today whereby polypeptides are made through the assembly of aminacyl-tRNAs[1] on a mRNA template embedded within a ribosome. Instead, a much simpler system probably first appeared that provided a transition from the RNA world to the theater of proteins. In part because tRNAs serve as the adapter molecules in this process, tRNA has been the topic of considerable speculation and experimentation concerning the evolution of peptide synthesis and coding.[2–7]

Transfer-RNAs are typically comprised of 73–95 nucleotides that can be represented by a cloverleaf secondary structure that in turn adopts a compact, L-shaped three-dimensional fold with two distinct domains (Fig. 1).[8] These domains are the acceptor-TΨC minihelix domain containing the amino acid attachment site at the 3′-end and the D-anticodon stem-biloop domain with the anticodon trinucleotide and its flanking sequences. Evidence that the minihelix domain represents an independent functional unit is growing. For example, RNase P catalyzes the maturation of the 5′-end of a preminihelix in vitro analogous to that for pre-tRNA.[9] In addition, at least 10 synthetases are known to specifically aminoacylate minihelices derived from their cognate tRNAs.[10–14] Once charged by their cognate synthetases, aminoacyl-tRNAs are carried to the ribosome by the G protein family member elongation factor Tu (Ef-Tu). As with the full tRNA molecule, the EF-Tu·GTP complex, but not the EF-TuGDP complex, specifically binds charged (and not uncharged) minihelices with high affinity.[15,16]

In addition to being an independent functional unit, the minihelix is thought to represent an ancient structural motif that appeared early in evolution and later combined with the anticodon-containing domain to form a single two-domain RNA. Weiner and Maizels argued that the acceptor-TΨC minihelix-like domain ending in the CCA trinucleotide is an ancient tag of RNA genomes[5] while, based on observation that the

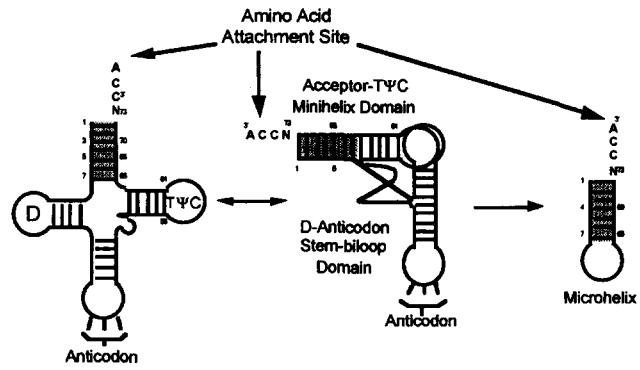

Figure 1. Schematic representation of the tRNA cloverleaf structure (left), how the segments of this structure are arranged into two domains within the three-dimensional structure (center) and the acceptor stem sequences that have given rise to the microhelix substrates for aminoacylation (right). The dihydrouridine (D) and pseudouridine loops (TΨC), the anticodon, and the amino acid attachment site are indicated in the cloverleaf structure. The base pairs of the acceptor stem that form the basis for the microhelix stem are shaded. The loop of the microhelix is typically based on the TΨC loop, although other sequences can be used because specificity for charging is determined by the stem itself.[10] The numbering used for the acceptor stem and single stranded NCCA at the 3′-end is that used for full-length tRNA. Adapted from ref. 6.

†Present address: Duke University Medical Center, Molecular Diversity Sciences Center, Box 3020, Durham, NC 27710, U.S.A.

two domains of tRNA interact with distinct rRNAs, Noller and coworkers suggested that two proto-RNAs (representing the precursors to the two contemporary tRNA domains) coevolved with distinct rRNAs.[4] The sequence-specific aminoacylation of small helical RNAs that lack the anticodon trinucleotides is also consistent with an independent origin for a charging system based on minihelix-like molecules.[17]

The plausible ancient nature of the minihelix domain has in turn led to speculation concerning the role of tRNA progenitors in the evolution of peptide synthesis and coded translation.[3-6] We suggested earlier that an operational RNA code for amino acids[17]—which is manifested by the sequence-specific charging of acceptor stem-like substrates—possibly started with ribozyme-catalyzed aminoacylation of RNA hairpins or hairpin-like structures at the 3′-ends of larger RNAs (Fig. 2). (Two possibilities for ribozyme-catalyzed aminoacylations have been demonstrated by the work of Illangasekare et al. and Picirilli et al.[18,19]) These hairpins, in turn, could have produced aminoacyl-RNAs for peptide bond formation.[6] Some of the peptides synthesized might then form RNA–peptide complexes for more efficient aminoacylations and for peptidyl transferase reactions. Evolution towards increased specificity and efficiency of these catalysts could lead to ribonucleoproteins (RNPs) for aminoacylations and eventually to the contemporary aminoacyl-tRNA synthetases.[20]

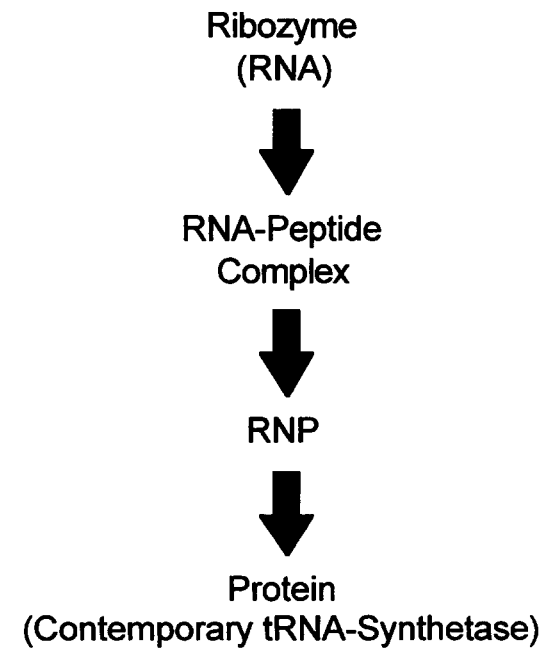

Figure 2. Schematic depiction of the stepwise evolution of proteins from an RNA world. Ribozymes capable of carrying out specific aminoacylations of small helical RNA substrates led to the first peptides. Those peptides that had selective advantage by virtue of enhancing the activity of ribozymes, and possibly by contributing to their own formation, led in turn to RNPs that were more efficient catalysts for carrying out aminoacylations. Eventually, the aminoacylations were taken over by contemporary aminoacyl-tRNA synthetases.

One difficulty in this progression is how aminoacyl moieties of charged minihelices could have been brought close enough to allow peptide bond formation in an early system of protein synthesis. One possibility is that aminoacyl-RNAs also contained sequences allowing for assembly of RNA–RNA complexes that positioned the aminoacyl moieties for reaction. These complexes could have served as the templates around which an RNA–peptidyl transferase might have evolved. Such complexes could in principle be formed through lateral hydrogen bonding of complementary loop nucleotides at opposite sides of aminoacyl-RNA hairpins (Fig. 3). Because the sequences for specific aminoacylation are contained in the first few base pairs of the helical stem plus the single stranded N^{73} 'discriminator base'[21] at the 3′-end[10-14] (Fig. 1), the loop–loop interactions are separated from those needed for aminoacylation specificity.

Loop–loop interactions have been described for several systems, including the association of tRNAs with complementary anticodons,[22,23] the regulation of HIV-1 mRNA transcriptional efficiency based on hairpin loop interactions within the transactivation response (TAR) element,[24] and the control of ColE1 plasmid replication by loop–loop structures.[25] These complexes have 'head-to-head' or 'kissing' loop–loop interactions that result in significant separation of the stems associated with the hairpin loops.[25] For example, the structure of one such complex, the ColE1 'kissing-hairpin' complex, reveals an angle between the two stems of ca. 45°.[25] On the other hand, 'lateral' or 'side-by-side' interactions of complementary loop nucleotides are intermolecular pseudoknots that locate the complementary sequences at the sides of the hairpin loops (Fig. 3), while maintaining a sufficient number of unpaired nucleotides to allow bridging of the grooves

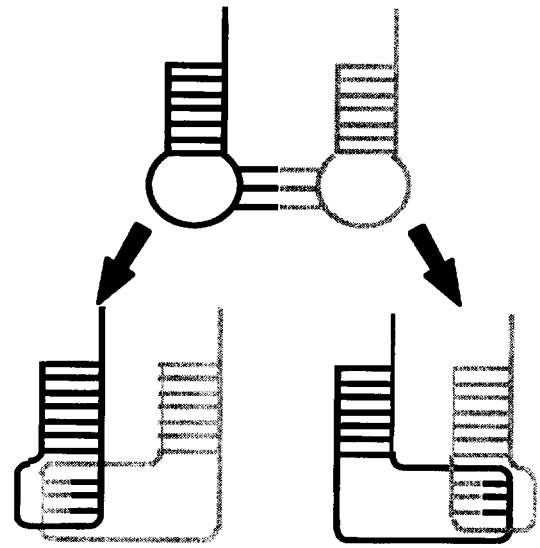

Figure 3. Illustration of lateral loop–loop base pairing between two RNA hairpin substrates (for aminoacylation). Two possible intermolecular pseudoknot secondary structures for the resulting complex are shown (bottom). The lengths of the segments for the single stranded loops are not drawn to scale.

of the quasi-continuous helix so formed. Sufficient flexibility could be achieved to align the acceptor stems so that the 3′-ends are proximal.

With the aminoacyl moieties positioned proximal to one another, peptide bond formation would involve either spontaneous or possibly ribozyme-catalyzed nucleophilic attack of the carbonyl of one aminoacyl moiety by the amino group of a neighbor. This reaction could be achieved with pairs of RNAs or with larger complexes having structural changes or thermal fluctuations affecting dissociation and binding of various aminoacyl-RNAs. If sufficient loop sequences were available to allow a set of different dipeptide combinations to form, and if a dipeptidyl-RNA could react with another aminoacyl-RNA and so on, then a system of noncoded peptide synthesis would be possible in principle.

As a first step in evaluating the feasibility of such assemblies for peptide synthesis, we set out to develop and study complexes of RNA microhelices capable of associating through lateral loop–loop interactions. Here we report the design and characterization of RNA microhelix complexes that form intermolecular pseudoknots through lateral base-pairing of loop nucleotides.

Results

General design considerations for complexes between RNA microhelices with complementary loops

Microhelix substrates for aminoacylation were selected as the basis for developing a system of RNA hairpins capable of association through loop–loop interactions. Microhelices are derived from tRNA acceptor-stem sequences (seven base pairs) with a common single stranded $NCCA_{OH}$ at the 3′-end (see Fig. 1) and are typically closed by a TΨC loop. These were chosen because they represent a class of truncated RNA substrates for aminoacylation that provide sufficient stem sequences for specific charging and provide loop sequences that are tolerant of variation. We refer to these systems as complementary microhelices. Larger RNAs, such as the minihelix substrates,[26] would have introduced additional degrees of freedom to the system and smaller substrates, such as the four-base-pair tetraloop substrates,[27] do not provide the flexibility in the loop sequences to allow for intermolecular interactions.

General tenants of nucleic acid structure[28] and results from structural studies involving small RNAs[25,29–31] were employed in developing a generic starting model for this system. NMR studies of RNA hairpins and microhelices have shown that the duplex regions of these molecules conform to general A-form characteristics. In the case of microhelix^{Ala}, the single-stranded ACCA at the 3′-end has been shown to stack with the double stranded stem, thereby maintaining general A-form structure.[30] Consequently, the seven base pair microhelix stems and single stranded ACCA at the 3′-ends were considered as rigid A-form structures. Constraining the acceptor stems and single-stranded ACCA sequences to A-form conformations required that the two microhelix loops provide for the intermolecular interactions and maintain sufficient flexibility to allow the acceptor stems to be oriented in parallel where the 3′-OHs would be proximal. Because maximization of stacking energy is a major factor in determining nucleic acid structure,[28] locating the loop–loop base pairing at the sides of the hairpin loops and directly adjacent to acceptor stem helices was expected to yield an intermolecular pseudoknot.

Computer simulation

Molecular simulations were performed using the constraint solution program MCSYM followed by energy minimization with XPLOR.[32–34] MCSYM produces stereochemically allowed combinations of known nucleotide conformations based on a set of input constraints. Energy minimization is used to provide a method of ranking the quality of the resulting solutions. This procedure was used to determine if positioning the 3′-OHs of the amino acid attachment sites close to one another (within the confines of an intermolecular pseudoknot) was stereochemically allowed. Moreover, the number of nonpaired nucleotides required to cross the different grooves of the quasi-continuous helix was also assessed.

The sequences used during the simulations are shown in Figure 4. The sequences of microhelix-1 and -2 were derived from the sequence of the *E. coli* alanine tRNA acceptor stem[35] with the following changes. The G3:U70 base pair was replaced with a G3:C70 base pair to allow for uniform Watson–Crick base pairing in the acceptor stem. The A7:U66 base pair of microhelix^{Ala} was changed to a G:C and C:G base pair for microhelix-1 and microhelix-2, respectively. Replacing the A7:U66 base pair with either C:G or G:C was anticipated to reduce fraying of the end of the helix and to increase the stability of the resulting complex by stronger stacking interactions. (The stacking interactions of G:C or C:G with the adjacent C:G pair of the intermolecular duplex were predicted to be 0.8 and 1.3 kcal/mol, respectively, stronger than interactions with an A:U pair.[28]) Finally, the loop sequence, which in earlier work on aminoacylation was chosen to be the TΨC loop,[10] was replaced with U_XGCG and $CGCU_Y$ for microhelix-1 and -2, respectively. The polyuridine tracks of the two loops were introduced to minimize potential steric interactions. The GCG and CGC sequences provided for intermolecular anti-parallel Watson–Crick base pairing (connect by dotted lines in Fig. 4).

Computer simulations were conducted in stages. Initially, the polyuridine tracks comprising the unpaired loop nucleotides were omitted, so that four strands were present. We refer to the resulting four-stranded complexes as quadraplexes, where two double stranded

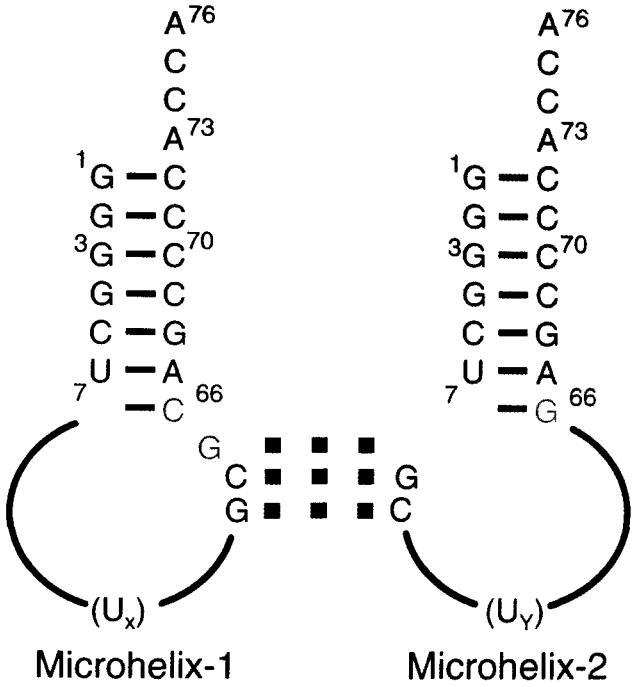

Figure 4. Nucleotide sequences used during computer simulation of intermolecular pseudoknots. The bases involved in loop–loop Watson–Crick base pairing (thus forming the interhairpin-loop duplex) are connected by dotted lines. The gold bases correspond to nucleotides that were allowed to deviate from canonical A-form nucleic acid conformation during simulation of a pseudoknot with the loop–loop duplex coaxially stacked with the acceptor stem of microhelix-1. The red bases correspond to nucleotides that were allowed to deviate from canonical A-form nucleic acid conformation during simulation of a pseudoknot with the interloop duplex coaxially stacked with the acceptor stem of microhelix-2. The single-stranded polyuridine loop sequences were also allowed to deviate from the A-form conformation.

helices are held together by a three-base-pair interduplex helix. In order for two acceptor stems to be positioned in parallel with the 3′-ends proximal to one another within the context of a quadraplex, conformational flexibility had to be introduced in a localized manner (e.g., at positions neighboring the intermolecular junction) and/or throughout the helical structure (e.g., variations in stacking conformations throughout the helical segments). Moreover, a quadraplex could, in principle, form through coaxial stacking of the interduplex three-base-pair helix with either of the acceptor stems (see Fig. 3). Consequently, we restricted ourselves to only localized variations in conformation and performed separate simulations with both possible stacking orientations.

Preliminary simulations were performed to define a minimal set of nucleotides whose conformations required deviation from A-form geometry to achieve an interduplex helix. For these simulations, a set of 70 MCSYM-derived nucleotide conformations, with variations in torsions ε, ζ, α, β, γ, δ and χ, were utilized for each of these positions. (The symbols ε, ζ, α, β, γ, δ and χ represent the torsional angles defined by the nucleotide $C_{3'}$–$O_{3'}$, $O_{3'}$–P, P–$O_{5'}$, $O_{5'}$–$C_{5'}$, $C_{5'}$–$C_{4'}$, $C_{4'}$–$C_{3'}$ and $C_{1'}$–base bonds, respectively.[33]) The 3′-ends of the acceptor stems were not constrained to be proximal during this stage. These simulations suggested that elements of base pair 7:66 in both stems and the first base pair of the interduplex helix had to be varied to achieve coaxial stacking of the interduplex helix with either acceptor stem (data not shown). Further refinement led to the identification of a different set of three nucleotides for each of the two stacking arrangements of the interduplex helix. The gold nucleotides in Figure 4 were identified as the minimal flexible set that allowed for intermolecular duplex stacking with acceptor stem-1. Likewise, the red nucleotides in Figure 4 were identified as the minimal flexible set that allowed for intermolecular duplex stacking with acceptor stem-2.

The two simulations described above employing the minimal flexible nucleotide sets were then repeated with the additional constraint that the 3′-OHs of the amino acid attachment sites be within 10 Å. This distance was considered a reasonable proximity and was expected to allow for some degree of rotational freedom of the two acceptor stems with respect to one another. With this constraint, stereochemically allowed solutions were only obtained when the interduplex helix was stacked with the acceptor stem of microhelix-2. A total of 64 solutions were obtained (from a search tree of 343,000 leaves).[36] The lowest energy quadraplex model from this set, after 10 cycles of energy minimization (conjugate gradient) using XPLOR 3.0, is shown in Figure 5(a). In this model, the GC dinucleotide of microhelix-1 that was not constrained to canonical A-form geometry (the red nucleotides of microhelix-1 in Figs 4 and 5a) forms a sharp turn allowing the acceptor stem of microhelix-1 to be positioned parallel to the acceptor stem of microhelix-2. The red G in microhelix-2 is located close in space to this turn and conformational variation at this position was required to accommodate the connection to the closed polyuridine loop in subsequent steps.

These findings suggested that positioning the 3′-OHs of two acceptor stems proximal to one another might be possible within the context of an intermolecular pseudoknot. A series of simulations were performed to assess the number of nonpaired nucleotides required to cross the different grooves of the quasi-continuous helix of the pseudoknotted dimer of hairpin loop microhelices. Using the quadraplex model shown in Figure 5(a) as a starting point and closing the two 'loops' with uridines, a series of simulations were performed where X and Y, representing the number of uridines in the two loops (see Fig. 4), were varied independently from 4 to 6. For the polyuridine tracks, a set of 30 MCSYM-derived nucleotide conformations representing variations in torsion angles ε, ζ, α, β, γ, δ and χ were employed at each position. Stereochemically allowed solutions were only found when both X and Y were greater than or equal to 5.[37] When X and Y were set equal to 5, a search tree of 5.9049×10^{14} leaves was produced that yielded 6631 stereochemically allowed solutions. Simulations with X/Y = 6 resulted in data sets too large ($\gg 10^4$) to analyze completely.

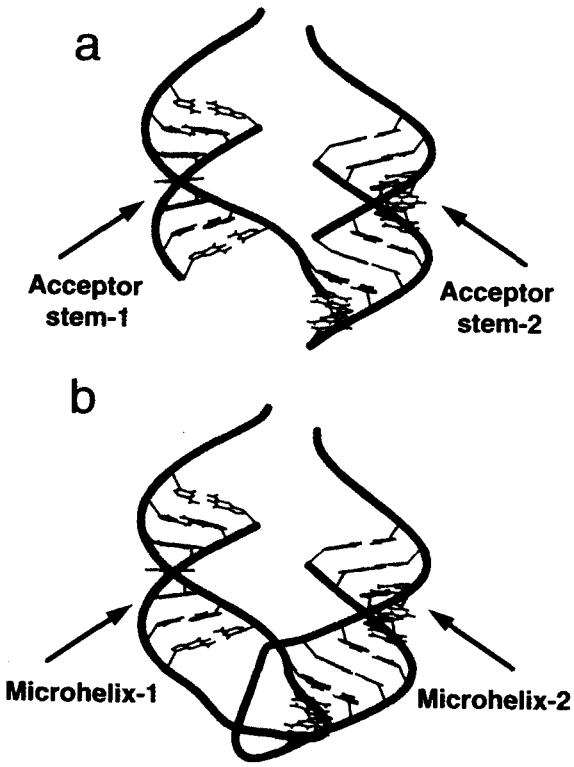

Figure 5. Models of intermolecular pseudoknots depicting specific solutions for a sterechemically allowed quadraplex (a) and pair of microhelices joined by loop–loop base pairing (b). The quadraplex model depicted represents the lowest energy solution obtained by analysis of 64 isomers. The quadraplex was constructed as the initial stage in developing a system of complementary microhelices using MCSYM v1.22 and energy minimization with XPLOR v3.0.[32–34] Canonical A-form parameters were used for the helices, including the three-base-pair interduplex helix that was coaxially stacked with acceptor stem-2. The conformation of the red nucleotides were allowed to vary in order to achieve proximal positioning of the 3′-ends. The 3′-OHs were restrained to within 10 Å. The model in panel a was then used as a starting point for incorporating the unpaired bases into microhelix loops. The model in panel b represents the lowest energy conformer obtained by analysis of 6631 isomers. The bases of the unpaired nucleotides are omitted for clarity. The orientations of the acceptor stems are the same in panels a and b. In these specific models the 3′-OHs are separated by 8.9 Å. Illustration generated using SETOR.[46]

These particular solutions represent stereochemically allowed variations in the conformations of the single-stranded segments of the two loops obtained by using extended backbone conformations. In each case, the interhairpin-loop duplex and acceptor stem conformations are those depicted in Figure 5(a). The resulting solutions were subject to ten cycles of energy minimization (conjugate gradient) using XPLOR 3.0 with the provided nucleic acid parameter sets[34] and ranked according to the total energy following the minimization process. A specific model representing the lowest energy solution produced from this set is shown in Figure 5(b).[37]

The bases of the polyuridine segments in the two microhelix loops for the low energy model depicted in Figure 5(b) show no particular regularity. This finding contrasts with the ordered nature of nonpaired loop bases found in the three dimensional structures for RNA hairpin loops whose bases stack with adjacent helical segments.[29,38,39] The disordered structure of the polyuridine tracks in the complementary microhelix model is likely to be a consequence of the conformational sets used in the simulations which were selected to search extended backbone conformational space instead of to impose additional structural constraints. Even with the crude nature of the loop simulations, inspection of the complementary microhelix model shown in Figure 5(b) demonstrates that the structural requirements for the two polyuridine tracks to be used for the hairpin loops of microhelix-1 and microhelix-2 were different, however.

The unpaired polyuridine segment of the loop for microhelix-2 crosses the minor groove in this model. The distance spanned from the end of the intermolecular duplex to the 7-66 base pair of microhelix-2 is about 14 Å. The crowding together of the two acceptor stems that results must also be accommodated and is most acute at the 7-66 base pairs of the two microhelices. Although five nucleotides were sufficient to span this juncture in the simulations, complex formation was predicted to be sensitive to the total number of nucleotides in the loop of microhelix-2. In contrast, the unpaired nucleotides of microhelix-1 cross the major groove face of the pseudoknot and no unusual structural features needed to be considered. Complex formation was predicted not to be strongly dependent on the number of nucleotides within this loop.

Characterization of complexes by native gel electrophoresis

The predictions of the computer simulations were tested experimentally with microhelices that incorporate complementary nucleotides for lateral base pairing within the context of variable loop sizes. Two microhelix^{Ala} variants were prepared (containing a G3:U70 base pair) with a loop sequence of $CGCU_5$ or $CGCU_6$ (Fig. 6). (Thus, in the simulations of Figure 5, microhelix^{Ala} corresponds to microhelix-2.) These are referred to as Micro^{Ala}-L8 and Micro^{Ala}-L9, respectively, and each can be aminoacylated with alanine. Two microhelix^{Gly} variants (Micro^{Gly}-L8 and Micro^{Gly}-L9) were prepared with loop sequences of AU_4GCG or AU_5GCG (Fig. 6). [The replacement of the 5′-U of the loop (for these variants of microhelix-1) with A was to avoid the potential formation of a G:U pair at the beginning of the loop.] Additionally, the use of two different stem sequences was to avoid intermolecular duplex formation between the stems themselves. These microhelices can be charged with glycine.

The association of these RNA hairpins through loop–loop interactions was assessed by native gel electrophoresis. When [^{32}P]Micro^{Ala}-L8 was titrated with increasing amounts of either Micro^{Gly}-L8 or Micro^{Gly}-L9, a complex is observed (Fig. 7; top and bottom left). The apparent dissociation constant (K_d) for this

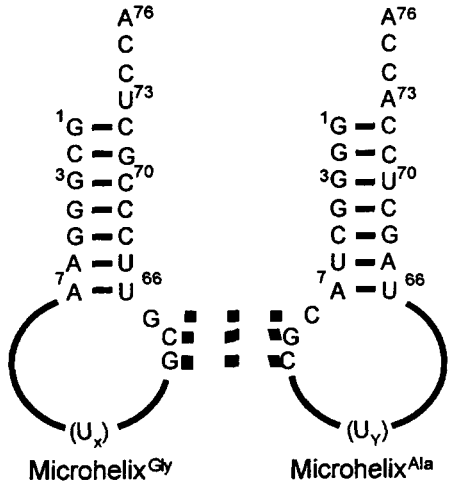

Figure 6. Sequences of complementary microhelices for alanine and glycine. The loops for microhelix^{Gly} (Micro^{Gly}) and microhelix^{Ala} (Micro^{Ala}) are identical to microhelices-1 and -2, respectively, shown in Figure 4. In order to facilitate studies involving charged complementary microhelices, the sequences for the stems of these hairpins were based on the wild-type microhelices.[35,47] The bases involved in loop–loop Watson–Crick base pairing, thus forming the interloop duplex, are connected by dotted lines.

complex was approximately 1 μM, regardless of the loop size of Micro^{Gly} (microhelix-1).[40] However, when [^{32}P]Micro^{Ala}-L9 is titrated with increasing amounts of either Micro^{Gly}-L8 or Micro^{Gly}-L9, the K_d decreases by approximately an order of magnitude to 100 nM (Fig. 7; top and bottom right). These data demonstrate that the association of complementary RNA hairpins is differentially sensitive to the number of unpaired nucleotides within the hairpin loops of Micro^{Ala} but not of Micro^{Gly}. This pattern was predicted based on the computer simulations which showed that complex formation would be sensitive to the number of nucleotides in the loop of the alanine microhelix (microhelix-2 in the simulations) but not sensitive to the number nucleotides in the loop of the glycine microhelix (microhelix-1 in the simulations).

Mutational analysis

The specificity of the RNA–RNA dimer formation between microhelices with complementary loop sequences was assessed by introducing mutations designed to disrupt either the base pairing or the stacking of the loop–loop helix onto an adjacent acceptor stem. The results are shown in Figure 8. Substitution of the Micro^{Gly}-L9 3′-G loop nucleotide with C, creates a CC mismatch, disrupting one of the intermolecular base pairs. This substitution abolished detectable dimer formation (Fig. 8; top right). Introduction of a single nucleotide at the 3′-end of the loop of Micro^{Gly} or at the 5′-end of Micro^{Ala} (both insertions at the position between the junction of sequences for intra- and intermolecular base pairing) reduced the extent of dimer formation (Fig. 8; bottom left and bottom right). No reduction in the extent of dimer formation was observed when as many as four nucleotides were introduced in either or both loops at points distal to the bases for intermolecular base pairing (e.g., increasing X/Y to 9 and 10, respectively) (data not shown).

Discussion

Our initial attempts to simply incorporate complementary nucleotides at the opposite sides of two microhelices did not result in detectable complex formation (data not shown). The loops of these constructs contained a total of seven nucleotides each and provided in principle for the same intermolecular base pairing shown in Figure 4. Because 'head-to-head' loop–loop interactions had been reported with as few as three base pairs,[23] it is likely that the four unpaired nucleotides within the loops of these initial constructs were not sufficient to allow for coaxial stacking of the putative inter-loop duplex with an adjacent acceptor stem duplex. This hypothesis was supported by the detection of a four-stranded quadraplex complex representing two acceptor stems joined by an interduplex helix of six base pairs (data not shown). These

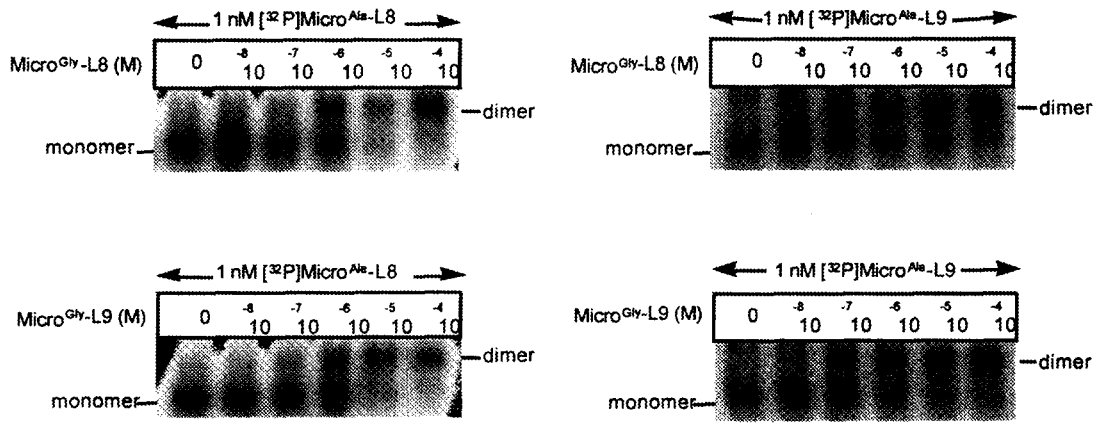

Figure 7. Sensitivity of complementary microhelix dimerization to variation in loop size. Radiolabeled Micro^{Ala}-L8 (left) or -L9 loop (right) variants (~1 nM) were annealed in reactions with varying concentrations (0–100 μM) of Micro^{Gly}-L8 or -L9 loop variants and were then electrophoresed through nondenaturing polyacrylamide gels at 4 °C. After drying the gels were quantified with a Molecular Dynamics Phosphorimager.

observations suggested that a more rigorous design for the loop–loop interactions was warranted.

Computer simulations indicated that positioning the amino acid attachment sites proximal to one another within the context of an intermolecular pseudoknot (formed through loop–loop base pairing) was stereochemically feasible. Although stereochemically allowed conformations were obtained with the intermolecular duplex coaxially stacked with either acceptor stem, only those solutions with coaxial stacking of the acceptor stem for microhelix-2 allowed the amino acid attachment sites to be positioned within 10 Å of one another. (The 3′-OHs of the specific models shown in Figure 5a and b are separated by 8.9 Å.) Moreover, coaxial stacking of the intermolecular duplex with the acceptor stem of microhelix-1 requires the minor grooves of the two acceptor stems to be oriented to the interior of the complex. This arrangement would represent a considerably higher energy conformation due to the repulsive electrostatic interactions between the phosphate backbones. Ulyanov et al. ruled out a similar arrangement of two DNA duplexes in their NMR and model building studies of slipped loop-DNA.[41]

The apparent dissociation constants for Micro^{Ala}-L9 with Micro^{Gly}-L9 or -L8 are approximately several orders of magnitude lower than expected for a complex between a trinucleotide and an RNA loop[23,28] and between one and two orders of magnitude lower than observed for tRNAs with complementary anticodons.[22,23] The extra G–C stacking interaction generated by formation of a quasi-continuous helix could account for roughly two to three orders of magnitude of extra stability (~34 kcal mol^{-1}).[28,42]

The mutational analysis demonstrated that complex formation was sequence-specific and that disruption of just one of the intermolecular base pairs was sufficient to abolish complex formation. The constructs that introduced a base insertion between the intermolecular duplex and either of the two acceptor stems (see Fig. 8) were designed to determine if one insertion would destabilize the complex to a greater extent than the others, as might be expected for a model where the quasi-continuous helix of the pseudoknot aligns with Micro^{Ala}. As can be seen, however, both destabilized the complex beyond the point where differences could be measured (Fig. 8). Thus, while some features of our model are supported by these experiments, the fine details need investigation by rigorous structural methods.

The system of complementary microhelices described represents a starting point for developing systems of hairpins for peptide synthesis. The structural details for complexes of microhelices that associate through lateral or side-by-side loop–loop interactions remain to be elucidated. However, the specificity and strength of the complexes described here illustrate the potential utility of such interactions and approaches to the design of stable RNA complexes. In terms of a system for peptide synthesis, this represents a starting point for the exploration of additional loop–loop interactions and the development of an RNA peptidyl transferase through the application of in vitro evolution techniques that have already shown the possibilities for amino acid transfer reactions catalyzed by a ribozyme.[43]

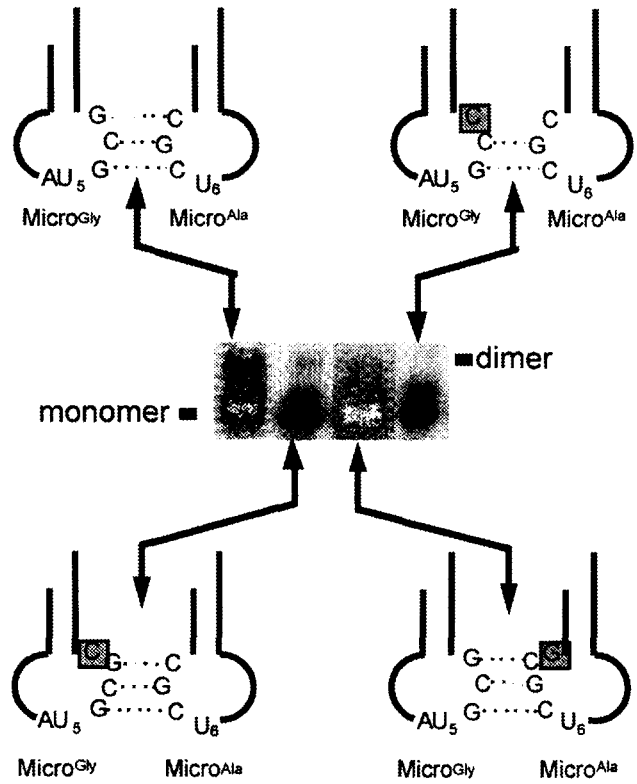

Figure 8. Effect of changes in complementary loop–loop interactions on complex formation. Radiolabeled Micro^{Ala}-L9 loop variants (~1 nM) were annealed with Micro^{Gly}-L9 loop variants (1 μM) and electrophoresed through nondenaturing polyacrylamide gels at 4 °C. After drying the gels were quantitated with a Molecular Dynamics Phosphorimager. In the series of experiments shown, dimerization of the control microhelices, Micro^{Ala} and Micro^{Gly}-L9, was incomplete. The shaded boxes indicate changes in the hairpins from the starting pairs shown in the upper left.

Materials and methods

Materials

Oligonucleotides were synthesized on a Pharmacia Gene Assembler Plus and deprotected using standard procedures.[44] Phosphoramidites were routinely obtained from ChemGenes, Glen Research or CPG, Inc. Deprotected oligonucleotides were purified by electrophoresis on a 16% denaturing polyacrylamide gel.[26] RNA concentrations were determined by using microhelix extinction coefficients determined previously.[35]

Computer simulations

Computer models of intermolecular pseudoknots formed by Watson–Crick base pairing of nucleotides

at the sides of two microhelix substrates for aminoacylation were generated using the constraint solution program MCSYM v1.22[32] in combination with energy minimization using XPLOR v3.0[34] using a Silicon Graphics Indigo^2 workstation. Stereochemically allowed solutions were subject to 10 cycles of energy minimization (conjugate gradient) using XPLOR 3.0 with the provided nucleic acid parameter sets and ranked according to the total energy following the minimization process.[34] The models were then evaluated visually using Quanta v4.0.

Native gel electrophoresis

Native gel electrophoresis was used to characterize the association of RNA.[24,45] Briefly, [^{32}P]-RNA hairpins (~1 nM) were incubated in separate reactions with increasing concentrations of complementary RNAs (0–100 μM) at 4 °C in 89 mM Tris borate, 0.1 mM EDTA, and 20 mM Mg^{2+}. Following annealing, aliquots were electrophoresed at 4 °C overnight through a non-denaturing polyacrylamide gel prepared with the same buffer. Gels were dried and quantitated by exposure to a Molecular Dynamics Phosphorimager. Apparent dissociation constants were taken as the point of 50% dimerization.

Acknowledgements

We thank Dr Lluís Ribas de Pouplana and Dr Douglas D. Buechter for critical discussions throughout the course of this work. Support was provided by grant GM 15539 from the National Institutes of Health and by a grant from the National Foundation for Cancer Research. B.H. was an NIH postdoctoral fellow.

References and Notes

1. Abbreviations: D, dihydrouridine; EDTA, ethylenediamine tetraacetic acid; HIV, human immunodeficiency virus; mRNA, messenger ribonucleic acid; TΨC, ribothymidine-pseudouridine-cytidine loop; RNP, ribonucleoprotein; TAR, transactivation response element.

2. Crick, F. H. C.; Brenner, S.; Klug, A.; Pieczenik, G. *Origins of Life* **1976**, *7*, 389.

3. Orgel, L. E. *J. Mol. Evol.* **1989**, *29*, 465.

4. Noller, H. F. In *The RNA World*; Gesteland, R. F.; Atkins, J. F., Eds.; Cold Spring Harbor Laboratory Press: Plainview, 1993; pp 137–156.

5. Maizels, N.; Weiner, A. M. In *The RNA World*; Gesteland, R. F.; Atkins, J. F., Eds.; Cold Spring Harbor Laboratory Press: Plainview, 1993, pp 577–602.

6. Schimmel, P.; Henderson, B. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11283.

7. Rodin, S.; Rodin, A.; Ohno, S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4537.

8. Sprinzl, M.; Steegborn, C.; Hübel, F.; Steinberg, S. *Nucl. Acids Res.* **1996**, *24*, 68.

9. McClain, W. H.; Guerrier-Takada, C.; Altman, S. *Science* **1987**, *238*, 527.

10. Martinis, S. A.; Schimmel, P. In *tRNA: Structure, Biosynthesis, and Function*; 1st ed.; Söll, D.; RajBhandary, U., Eds.; American Society for Microbiology: Washington, DC, 1995, pp 349–369.

11. Quinn, C. L.; Tao, N.; Schimmel, P. *Biochemistry* **1995**, *34*, 12489.

12. Hou, Y.-M.; Sterner, T.; Bhalla, R. *RNA* **1995**, *1*, 707.

13. Frugier, M.; Florentz, C.; Giegé, R. *EMBO J.* **1994**, *13*, 2219.

14. Nureki, O.; Niimi, T.; Muto, Y.; Kanno, H.; Kohno, T.; Muramatsu, T.; Kawai, G.; Miyazawa, T.; Giegé, R.; Florentz, C.; Yokoyama, S. In *The Translation Apparatus*; Nierhaus, K. H.; Franceschi, F.; Subramanian, A. R.; Erdmann, V. A.; Wittmann-Liebold, B., Eds.; Plenum: New York, 1993, pp 59–66.

15. Nazarenko, I. A.; Uhlenbeck, O. C. *Biochemistry* **1995**, *34*, 2545.

16. Rudinger, J.; Blechschmidt, B.; Ribeiro, S.; Sprinzl, M. *Biochemistry* **1994**, *33*, 5682.

17. Schimmel, P.; Giegé, R.; Moras, D.; Yokoyama, S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8763.

18. Illangasekare, M.; Sanchez, G.; Nickles, T.; Yarus, M. *Science* **1995**, *267*, 643.

19. Piccirilli, J. A.; McConnell, T. S.; Zaug, A. J.; Noller, H. F.; Cech, T. R. *Science* **1992**, *256*, 1420.

20. Discussions relating to possible mechanisms for the transition from the RNA world to the theater of proteins have been published (see, for example refs 2–7).

21. Crothers, D. M.; Seno, T.; Söll, D. G. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 3063.

22. Eisinger, J. *Biochem. Biophys. Res. Commun.* **1971**, *43*, 854.

23. Grosjean, H.; Söll, D. G.; Crothers, D. M. *J. Mol. Biol.* **1976**, *103*, 499.

24. Chang, K.-Y.; Tinoco, I., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8705.

25. Marino, J. P.; Gregorian, R. S., Jr.; Csankovszki, G.; Crothers, D. M. *Science* **1995**, *268*, 1448.

26. Francklyn, C.; Schimmel, P. *Nature (London)* **1989**, *337*, 478.

27. Shi, J.-P.; Martinis, S. A.; Schimmel, P. *Biochemistry* **1992**, *31*, 4931.

28. Turner, D. H.; Sugimoto, N. *Ann. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 167.

29. Davis, P. W.; Thurmes, W.; Tinoco, I., Jr. *Nucl. Acids Res.* **1993**, *21*, 537.

30. Limmer, S.; Hofmann, H.-P.; Ott, G.; Sprinzl, M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6199.

31. Ott, G.; Arnold, L.; Limmer, S. *Nucl. Acids Res.* **1993**, *21*, 5859.

32. Major, F.; Turcotte, M.; Gautheret, D.; Lapalme, G.; Fillion, E.; Cedergren, R. *Science* **1991**, *253*, 1255.

33. Gautheret, D.; Major, F.; Cedergen, R. *J. Mol. Biol.* **1993**, *229*, 1049.

34. Brünger, A.; Kuriyan, J.; Karplus, M. *Science* **1987**, *235*, 458.

35. Francklyn, C.; Shi, J.-P.; Schimmel, P. *Science* **1992**, *255*, 1121.

36. A total of 43 solutions were obtained with coaxial stacking of the interduplex helix with the acceptor stem of microhelix-1.

These solutions had the 3′-ends of the acceptor stems constrained to within 15 Å, however. Distances less than 15 Å failed to yield solutions. For the 43 solutions obtained with a distance of 15 Å, the minor grooves of the acceptor stems were positioned to the interior of the complex where the phosphate backbones were brought into close proximity. This orientation was therefore considered unlikely to be productive for the purposes at hand.

37. A preliminary description of these simulations has been reported previously. See Schimmel, P. In *Tracing Biological Evolution in Protein and Gene Structures*; Go, M.; Schimmel, P., Eds.; Elsevier: Nagoya, 1995, pp 1–10.

38. Woo, N. H.; Roe, B. A.; Rich, A. *Nature (London)* **1980**, *286*, 346.

39. Moras, D.; Comarmond, M. B.; Fischer, J.; Weiss, R.; Thierry, J. C.; Ebel, J. P.; Giegé, R. *Nature (London)* **1980**, *288*, 669.

40. For the experiments reported in Figure 7 (top and bottom), the amount of dimer detected was dependent on the length of time the samples were electrophoresed. The values reported here represent values obtained for ~12 h electrophoresis runs. Electrophoresis times from 3–6 h yielded K_ds that were roughly an order of magnitude lower for all pairs of hairpins, but the decreased resolution of monomeric from dimeric species made quantitation more difficult. For times longer than 12 h (ca. 16–20 h), the K_ds increase by about an order of magnutude.

41. Ulyanov, N. B.; Bishop, K. D.; Ivanov, V. I.; James, T. L. *Nucl. Acids Res.* **1994**, *22*, 4242.

42. Walter, A. E.; Turner, D. E.; Kim, J.; Lyttle, M. H.; Müller, P.; Mathews, D. H.; Zuker, M. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 9218.

43. Lohse, P. A.; Szostak, J. W. *Nature (London)* **1996**, *381*, 442.

44. Scaringe, S. A.; Francklyn, C.; Usman, N. *Nucleic Acids Res.* **1990**, *18*, 5433.

45. Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456.

46. Evans, S. V. *J. Mol. Graphics* **1993**, *11*, 134.

47. Hipps, D.; Shiba, H.; Henderson, B.; Schimmel, P. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5550.

(*Received in U.S.A.* 2 *September* 1996; *accepted* 18 *February* 1997)