

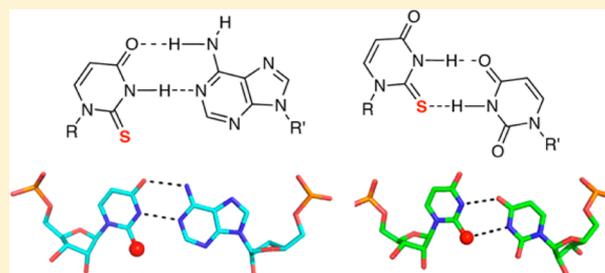
Crystal Structure Studies of RNA Duplexes Containing $s^2U:A$ and $s^2U:U$ Base Pairs

Jia Sheng,[†] Aaron Larsen, Benjamin D. Heuberger,[‡] J. Craig Blain,[§] and Jack W. Szostak*

Howard Hughes Medical Institute, Center for Computational and Integrative Biology, and Department of Molecular Biology, Simches Research Center, Massachusetts General Hospital, Boston, Massachusetts 02114, United States

Supporting Information

ABSTRACT: Structural studies of modified nucleobases in RNA duplexes are critical for developing a full understanding of the stability and specificity of RNA base pairing. 2-Thio-uridine (s^2U) is a modified nucleobase found in certain tRNAs. Thermodynamic studies have evaluated the effects of s^2U on base pairing in RNA, where it has been shown to stabilize U:A pairs and destabilize U:G wobble pairs. Surprisingly, no high-resolution crystal structures of s^2U -containing RNA duplexes have yet been reported. We present here two high-resolution crystal structures of heptamer RNA duplexes ($5'$ -uagc s^2U cc- $3'$ paired with $3'$ -aucgAgg- $5'$ and with $3'$ -aucgUgg- $5'$) containing $s^2U:A$ and $s^2U:U$ pairs, respectively. For comparison, we also present the structures of their native counterparts solved under identical conditions. We found that replacing O2 with S2 stabilizes the U:A base pair without any detectable structural perturbation. In contrast, an $s^2U:U$ base pair is strongly stabilized in one specific U:U pairing conformation out of four observed for the native U:U base pair. This $s^2U:U$ stabilization appears to be due at least in part to an unexpected sulfur-mediated hydrogen bond. This work provides additional insights into the effects of 2-thio-uridine on RNA base pairing.



INTRODUCTION

The genetic functions of RNA rely on accurate Watson–Crick base pairing, while the structural, regulatory, and catalytic functions of RNA are achieved by the formation of well-defined 3D structures resulting from a combination of normal Watson–Crick pairs and a wide variety of non-canonical base pairs as well as other tertiary interactions.^{1–5} An improved understanding of the structures and energetics of base–base interactions is important for the further elucidation of RNA functions, the development of new RNA-based therapeutics, and the study of the origins of life. Nature uses modified nucleobases to increase the specificity and diversity of RNA base–base interactions. Over 140 post-transcriptional modifications have been discovered so far in mRNA, rRNA, tRNA, and non-coding RNAs.⁶ It is possible that at least some of these modified nucleobases are relics of the RNA World, where they may have enhanced the chemical diversity of RNA prior to the emergence of coded proteins.

Among the nearly 60 known uridine modifications, 16 feature thiolation at the C2 position as in 2-thiouridine (s^2U) and its C5 modified derivatives.⁶ These modifications are observed at position 34 in certain tRNAs;^{7,8} this is the first position of the anticodon, which is base-paired with the nucleotide at the wobble position of the mRNA codon. The presence of s^2U and its S-modified derivatives have been demonstrated to increase codon-anticodon recognition efficiency and accuracy, enhance the aminoacylation kinetics of tRNA, and prevent frame-shifting during translation.⁹ These observations raise the question of whether this modification

might have played a role in non-enzymatic RNA replication at an early stage in the origin of life. The fidelity of non-enzymatic primer-extension on RNA templates has been reported to be quite poor, with U:G wobble pairing contributing a substantial fraction of the errors.¹⁰ We have recently observed that the non-enzymatic copying of $3'$ -phosphoramidate-DNA templates with activated $3'$ -amino nucleotides is highly error-prone, but that replacement of $3'$ -aminothymidine with $3'$ -amino, 2-thiothymidine greatly reduces wobble pairing and thereby strongly increases the fidelity of the copying process.¹¹ The question of whether replacing U with s^2U would enhance the fidelity of non-enzymatic RNA copying remains unresolved. As part of our efforts to address this question, we have become interested in the energetic and structural effects of this substitution on the RNA duplex.

It has been known for over two decades that in the context of an extended RNA:RNA duplex, s^2U at an internal U:A pair is strongly stabilizing, but the same modification at a wobble U:G pair is mildly destabilizing.^{12,13} Substitution at an internal location resulted in a significantly greater increase in thermal stability compared to terminal substitutions, most likely due to the inherently high flexibility of duplex ends. It remains unclear as to whether these stabilizing and destabilizing effects are primarily enthalpic or entropic in origin. Several hypotheses have been put forward to explain the greater stability of the $s^2U:A$ base pair relative to the standard U:A base pair. One

Received: August 5, 2014

Published: September 4, 2014

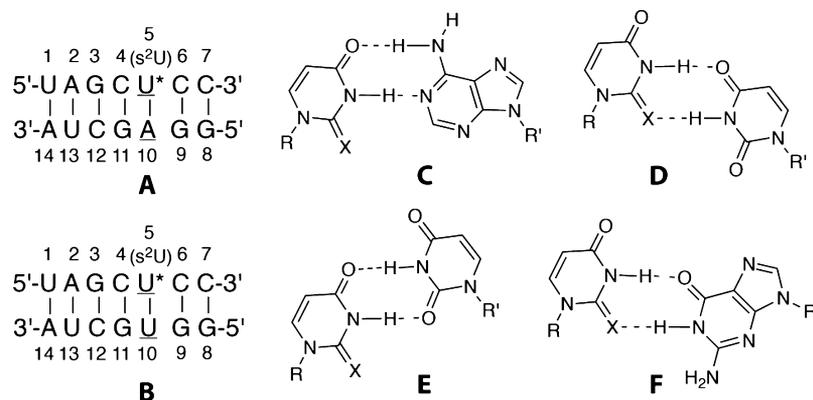


Figure 1. Sequences of heptamer RNA duplexes containing (A) $s^2U5:A10$ and $U5:A10$ pairs and (B) $s^2U5:U10$ and $U5:U10$ pairs. (C–F) Hydrogen bonding patterns of native and 2-thio-U modified U:A, U:U, and U:G base pairs. R, R' = ribose; X = O, uridine; X = S, 2-thio-uridine.

possibility is that s^2U enhances backbone preorganization and rigidity by locking the sugar pucker into the 3'-*endo* conformation and by extending this conformation to the 3'-adjacent nucleotides.^{14,15} Alternatively, the lower electronegativity of sulfur makes it a weaker hydrogen bond acceptor, which may reduce the desolvation penalty during duplex formation and thereby make the 2-thiolated duplex more stable.¹⁶ Other possibilities include increasing the strength of -N3-H as a hydrogen-bond donor^{12,17} and stronger stacking interactions due to the replacement of O2 with the more polarizable sulfur.¹⁸ A previously solved A-form DNA duplex structure containing 2'-O-[2-(methoxy)ethyl]-2-thiothymidine indicated that 2-thiolation causes only minor adjustments in the pattern of bound water molecules and a small overall structural perturbation, with the notable exception of altered base-pair "opening" and increased overlap of the s^2U nucleobase with adjacent residues.¹⁹ Surprisingly, no high resolution crystal structures of s^2U -containing RNA:RNA duplexes have yet been reported. Here, we present two crystal structures of heptamer RNA duplexes containing $s^2U:A$ and $s^2U:U$ base pairs (Figure 1A and B), along with the corresponding native structures. Our findings provide additional insight into the effect of 2-thiouridine on the specificity of RNA base pairing.

RESULTS AND DISCUSSION

Thermodynamic Studies of RNA Duplexes Containing $s^2U:A$ and $s^2U:U$ Pairs. We studied four 7-mer RNA:RNA duplexes, based on the canonical duplex formed by the pairing of 5'-uagcUcc-3' with 3'-aucgAgg-5'. This duplex is derived from the acceptor stem of *Escherichia coli* tRNA(Ala).²⁰ To investigate the effects of s^2U on A:U base pairing in duplex RNA, we synthesized a variant of the first oligo in which the highlighted U residue was replaced with s^2U . We also examined the consequences of U:U and $s^2U:U$ mispairing by annealing the two versions of this oligo with 3'-aucgUgg-5'. We then measured the T_M of all four duplexes by standard UV absorbance methods (Table 1). Consistent with previous measurements in different sequence contexts, the substitution of U with s^2U in a U:A base pair increases the T_M by 9.0 °C in buffer containing 100 mM NaCl and by 6.5 °C in buffer containing 100 mM MgCl₂ (Table 1, entries 1 and 2). Also consistent with previous results,^{21,22} the U:U mismatch strongly destabilized the RNA duplex, with the observed T_M decreasing by 16.8 °C in buffer containing 100 mM NaCl and by 16.8 °C as well in buffer containing 100 mM MgCl₂ (Table 1, entries 1 and 3). Changing the U:U mismatch to a $s^2U:U$ mismatch

Table 1. Melting Temperatures of the Four Heptamer RNA Duplexes

entry	duplex	base pair	T_M (°C) ^a	T_M (°C) ^b
1	5'-uagcUcc-3' 3'-aucgAgg-5'	U:A	51.0 ± 0.2	58.8 ± 0.8
2	5'-uagcs ² Ucc-3' 3'-aucgAgg-5'	$s^2U:A$	60.0 ± 0.4	65.3 ± 1.3
3	5'-uagcUcc-3' 3'-aucgUgg-5'	U:U	34.2 ± 0.04	42.0 ± 0.7
4	5'-uagcs ² Ucc-3' 3'-aucgUgg-5'	$s^2U:U$	42.5 ± 1.2	50.8 ± 1.8

^a200 mM HEPES, pH 7.5, 100 mM NaCl. ^b200 mM HEPES, pH 7.5, 100 mM MgCl₂. All samples contained 100 uM RNA duplex. T_M values are the average of duplicate measurements ± half of the difference between the two measured values.

resulted in a T_M increase of 8.3 °C in buffer containing 100 mM NaCl and 8.8 °C in buffer containing 100 mM MgCl₂ (Table 1, entries 3 and 4). These results are consistent with previous studies of T:T and 2-thio-T:T mismatch studies in a DNA duplex.²³

Crystallization and Structure Determination. To investigate the structural basis of the stability enhancement of duplex RNAs by 2-thiolation of U, we crystallized each duplex in Table 1. Several highly regular crystals of both native and 2-thio-modified RNA duplexes formed within 2–3 weeks at room temperature (20 °C) using the Hampton nucleic acid mini-screen kit and Natrix crystallization buffers. Most of these crystals diffracted at a resolution higher than 2.0 Å. To ensure a consistent comparison, the structures of two RNA duplexes containing U:A and $s^2U:A$ pairs were determined using crystals grown under identical conditions (10% MPD, 40 mM Na cacodylate pH 6.0, 12 mM spermine tetra-HCl and 80 mM NaCl). Similarly, the two structures containing U:U and $s^2U:U$ pairs were also determined from crystals grown under identical conditions (10% MPD, 40 mM Na cacodylate pH 7.0, 12 mM spermine tetra-HCl, 80 mM KCl and 20 mM BaCl₂). Data collection and structure refinement statistics are summarized in Table 2. All four structures were solved by molecular replacement using a model structure of an otherwise identical RNA duplex in which the U5:A10 base-pair is replaced with a U5:G10 wobble pair (PDB ID: 434D).²⁰

Crystal Packing and Overall Structures. To determine whether 2-thiolation affects molecular packing, we evaluated the helix–helix interactions in the unit cell. The two duplexes

Table 2. X-ray Data Collection and Structure Refinement Statistics^a

	UA	s ² UA	UU	s ² UU
		Scaling		
space group	C2	C2	P2 ₁	P2 ₁ 2 ₁ 2 ₁
unit cell parameters (Å, deg)	37.48, 38.49, 30.41, 90, 110.5, 90	37.39, 38.68, 30.07 90, 109.2, 90	29.03, 81.31, 36.67 90, 113.1, 90	21.68, 35.16, 47.30 90, 90, 90
resolution range, Å (last shell)	30–1.55 (1.61–1.55)	30–1.35 (1.40–1.35)	30–1.80 (1.86–1.80)	30–1.55 (1.61–1.55)
unique reflections	5625 (431)	8436 (579)	13573 (1114)	5355 (394)
completeness, %	95.1 (73.7)	93.5 (66.6)	93.1 (76.6)	95.1 (73.2)
R _{merge} ^b %	12.4 (17.3)	6.5 (12.9)	16.6 (35.0)	9 (25.8)
$\langle I/\sigma(I) \rangle$	12.3 (6.35)	24.2 (11.1)	10.2 (2.0)	23.3 (3.4)
redundancy	5.3 (2.6)	6.7 (4.0)	6.5 (3.5)	11.7 (6.3)
		Refinement		
molecules per asymmetric unit	1 duplex	1 duplex	4 duplex	1 duplex
resolution range, Å	28.49–1.55	22.06–1.35	30–1.80	28.22–1.55
no. of reflections	5375	8031	12816	5091
completeness, %	94.7	93.4	92.1	94.9
R _{work} %	20.3	18.8	21.1	18.3
R _{free} %	22.8	19.8	25.2	21.1
bond length rms Å	0.013	0.005	0.011	0.006
bond angle rms	1.997	1.412	1.842	1.620
overall B-factor with water, Å ²	20.33	11.11	37.70	18.64
av B-factor of RNA atoms, Å ²	18.7	7.36	37.30	15.48

^aData for the native 7mer duplex with a UA pair (UA): [5'-uagcUcc-3'/3'-aucgAgg-5'], the s²U:A-containing RNA 7mer duplex (s²UA): [5'-uagc(s²U)cc-3'/3'-aucgAgg-5'], the native 7mer duplex with a UU pair (UU): [5'-uagcUcc-3'/3'-aucgUgg-5'], and the s²U:U-containing 7mer duplex (s²UU): [5'-uagc(s²U)cc-3'/3'-aucgUgg-5']. ^bR_{merge} = $\sum |I - \langle I \rangle| / \sum I$.

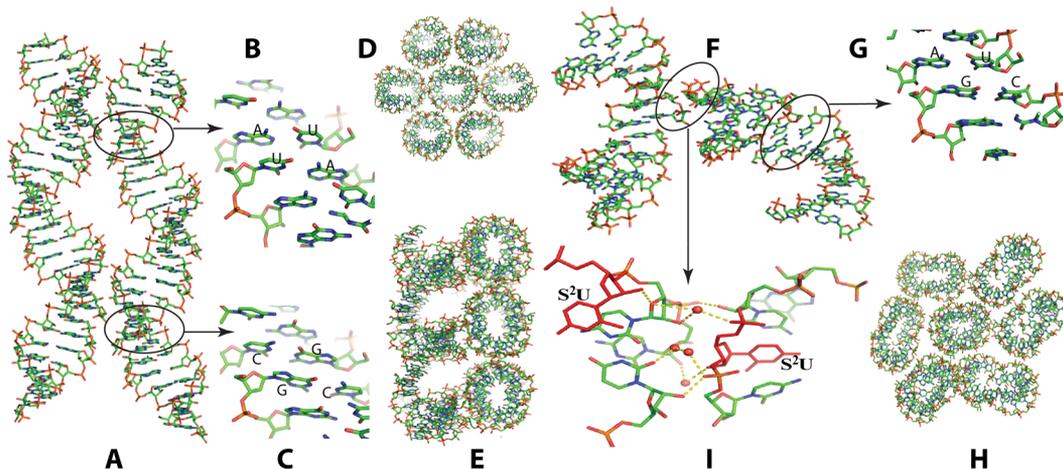


Figure 2. Molecular packing of native and s²U-modified RNA duplexes containing U:A and U:U pairs. (A) Side view of native heptamer RNA duplex with the U:A pair. (B) Zoom in view of the duplex terminal junction showing head-to-head stacking of two A:U pairs. (C) Zoom in view of duplex terminal junction showing tail-to-tail stacking of two G:C pairs. (D) Top view of native heptamer RNA duplex with U:A pair. (E) Packing overview of the RNA duplex with native U:U mismatch, showing two perpendicular axes of duplex packing. (F) Head-to-tail packing mode showing stacked U:A and C:G pairs in the heptamer RNA containing a s²U:U pair. (G) Side view of s²U:U-containing heptamer RNA duplex packing, with two axes in an angle of about 45°. (H) Top view of s²U:U-containing heptamer RNA duplex packing. (I) Zoom in view of duplex interaction by s²U and its two flanking nucleotides, mediated by water-bridging hydrogen bonds. (Red spheres represent water molecules; yellow dashed lines represent hydrogen bonds).

containing native U:A and s²U:A pack in the same manner and in the same space group (C2). As shown in Figure 2A, these duplexes stack head-to-head with two terminal U1:A14 pairs (Figure 2B) and tail-to-tail with two terminal C7:G8 pairs (Figure 2C), forming endless helices. Each long helix is surrounded by six other columns of stacked double helices with parallel axes (Figure 2D).

In the crystals of the U:U-containing duplex, there are four duplexes in the asymmetric unit. Although each duplex packs in the same head-to-head and tail-to-tail mode and forms long

helices as in the two U:A pair structures, these helices are aligned in two perpendicular axes, instead of being parallel (Figure 2E). In contrast, in the crystals of the s²U:U duplex, each duplex packs in a head-to-tail mode with terminal U1:A14 pairs stacking with C7:G8 pairs (Figure 2G). These long helices further pack along two axes at an angle of approximately 45° (Figure 2F). As a result, the top views of these molecules show a similar pattern as in the U:A pair, with six extended helical columns surrounding each vertically oriented helix (comparing Figure 2D and H). Further analysis indicated that the 2-thio-U

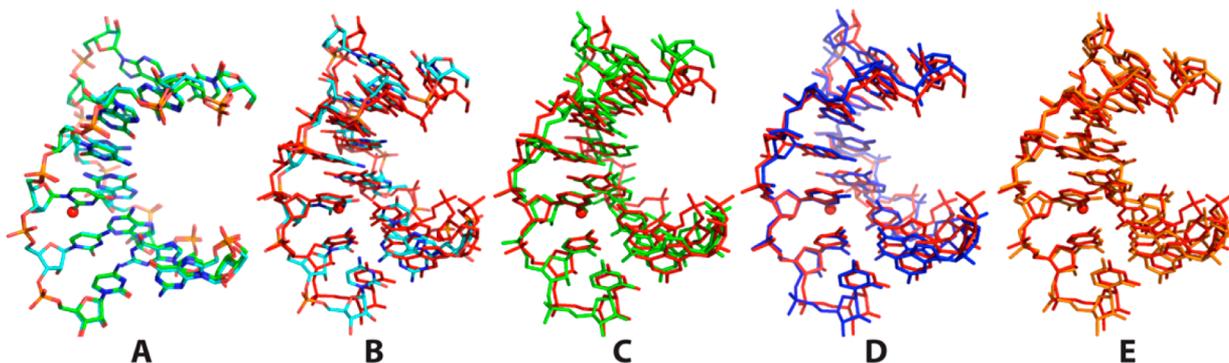


Figure 3. Comparison of the duplex structures of (A) native U:A and s^2 U:A-7mer duplexes, rmsd 0.19 Å; (B) s^2 U:U-7mer (red) and native UU-1 (cyan, chain AB), rmsd 1.26 Å; (C) s^2 U:U-7mer (red) and native UU-2 (green, chain CD), rmsd 1.27 Å; (D) s^2 U:U-7mer (red) and native UU-3 (blue, chain EF), rmsd 0.74 Å; (E) s^2 U:U-7mer (red) and native UU-4 (orange, chain GH), rmsd 0.653 Å. Sulfur atoms are shown as red spheres.

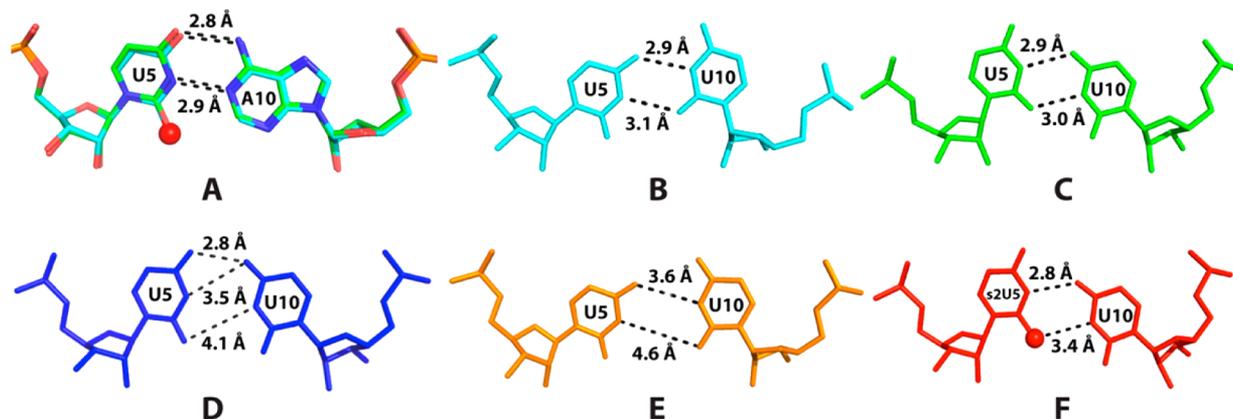


Figure 4. Base-pair structures. (A) Superposition of the U5:A10 and the s^2 U5:A10 pairs in RNA heptamer duplex structures. (B–E) The four native U5:U10 pairing patterns observed in the native heptamer RNA structure with the same color code as Figure 3: UU-1, chain AB (cyan); UU-2, chain CD (green); UU-3, chain EF (blue); and UU-4, chain GH (orange). (F) s^2 U5:U10 pair observed in the 2-thiolated heptamer RNA structure. Sulfur atoms are shown as red spheres.

residue and its two flanking bases are involved in duplex interactions through water-mediated hydrogen bonds connecting the 5'-phosphate oxygen and the 2'-hydroxyl group of the s^2 U (Figure 2I).

Consistent with the unchanged molecular packing, s^2 U has a minimal effect on the overall U:A-duplex structure compared to the native structure. As shown in Figure 3A, both duplexes form ideal A-form helices and are aligned very well with a root-mean-square deviation (rmsd) of the two duplexes of only 0.19 Å. A slight terminal backbone rotation is responsible for this minor deviation. This perturbation is smaller than that observed in the 2'-O-[2-(methoxy)ethyl]- s^2 T-containing A-form DNA duplex,¹⁹ suggesting that the RNA duplex might be more flexible and thus more able to accommodate the 2-thio substitution than the highly modified A-form DNA duplex.

In the native U:U-duplex structure, each of the four duplexes in the asymmetric unit exhibits significant structural differences from the others. Such structural variability suggests that the U:U mismatch-containing duplex structure is quite flexible and that the energy differences between these conformations are very small. In contrast, when one of the U residues is replaced with s^2 U, only one conformation is seen and there is only one duplex in the asymmetric unit of the s^2 U:U-duplex crystal. Considering that these crystals were grown under identical conditions, the thiolation of U is likely to be responsible for the reduction in conformational variability, although it is difficult to rule out crystal packing effects on the overall duplex

conformation since the s^2 U residue is directly involved in lattice interactions. It is also noteworthy that the space groups of native U:U and s^2 U:U-duplex structures are $P2_1$ and $P2_12_12_1$ respectively. When we attempted to index the diffraction data of the native U:U duplex crystals to the higher symmetry space group $P2_12_12_1$, the R -factors calculated during refinement failed to drop below 35%. In contrast, refinement in $P2_1$ progressed smoothly to a final R_{free} of 25.2%. As a result, four slightly different duplexes were captured in one asymmetric unit with different modes of U:U pairing. We then aligned the single s^2 U:U-duplex with each of native U:U-duplexes (labeled as UU-1, UU-2, UU-3, and UU-4), as shown in Figure 3B–E. The rms deviations of these alignments range from 0.65 to 1.27 Å. The s^2 U:U-duplex (red in Figure 3B–E) is more strongly bent than each of the two native duplex structures, such that the distance between the two terminal phosphate atoms in the s^2 U:U-duplex decreased by 1.3, 1.9, 0.8, and 0.8 Å, respectively, compared to the four native duplexes (21.4 Å in s^2 U:U vs 22.7 Å in UU-1; 23.3 Å in UU-2; 22.2 Å in UU-3; 22.2 Å in UU-4).

Base Pairing Studies. We studied the effect of 2-thiolation on U:A and U:U base pairing in more detail by examining the conformations of the nucleotides involved and the nature of their interactions. In each structure, the 2-thiouridine adopts the 3'-endo sugar pucker conformation. Consistent with the overall duplex comparisons, the s^2 U:A and native U:A base-pairs aligned precisely, with the same Watson–Crick hydrogen bonding distances (Figure 4A). Examination of the four distinct

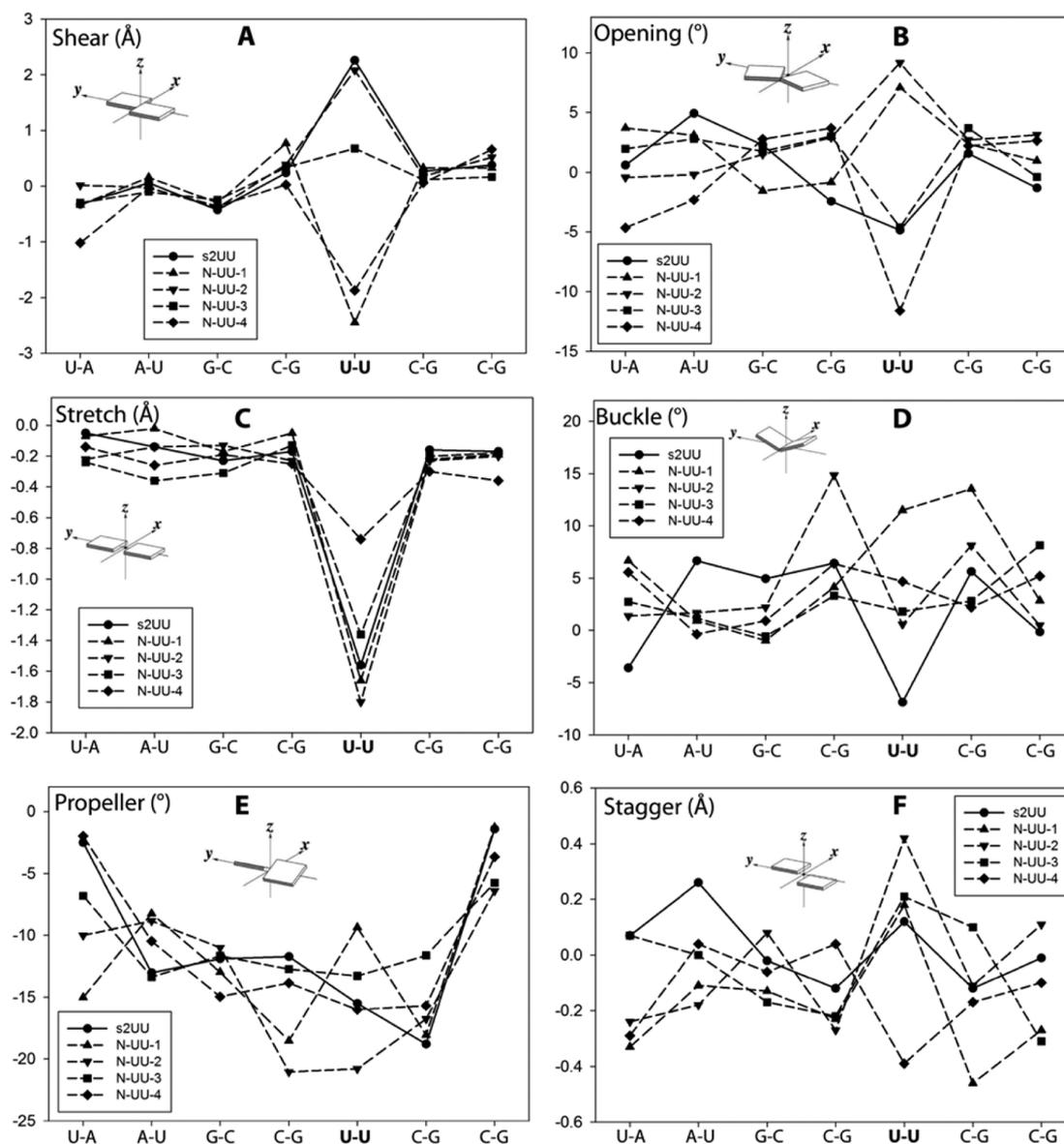


Figure 5. Local base-pair parameters for duplexes containing $s^2U:U$ or $U:U$ mismatches: (A) shear, (B) opening, (C) stretch, (D) buckle, (E) propeller, and (F) stagger. Duplex with $s^2U:U$ pair, ● and solid lines; duplexes with a $U:U$ pair, dashed lines (▲ for UU-1, ▼ triangle for UU-2, ■ for UU-3, and ◆ for UU-4). All schematics are from 3DNA.²⁴

$U:U$ duplexes revealed four different $U:U$ pairing patterns (Figure 4B–E). In duplex UU-1, U5 and U10 interact by hydrogen bonds between N3 of U5 and O2 of U10 and between O4 of U5 and N3 of U10 (Figure 4B). In UU-2, the $U:U$ interaction shifts such that the two hydrogen bonds are between N3 of U5 and O4 of U10 and between O2 of U5 and N3 of U10 (Figure 4C). In the UU-3 and UU-4 structures, the two uridines are further apart, resulting in much weaker base pairing (Figure 4D and E). We note that U10 in Figure 4D may be present in the enol tautomeric form as indicated by its density map (Supplementary Figure S1A) and the short distance between the two O4 atoms (2.8 Å), while the UU pair in Figure 4E may represent a disordered mixture of the first two pairing patterns, as indicated by the partial positive density map observed for the U10 residue (Supplementary Figure S1B). The four pairing patterns captured in this structure indicate that $U:U$ pairing is quite flexible. The energetic differences between these structures must be small, otherwise

only the most stable structure would be seen. In contrast, when one U is replaced with s^2U , only one pairing conformation is seen (Figure 4F). Considering that the sulfur atom is both larger and less electronegative than oxygen, one would expect the preferred $s^2U5:U10$ pairing pattern to be the UU-1 state (Figure 4B) where the sulfur would not be involved in a hydrogen bond. Surprisingly, the observed $s^2U5:U10$ pairing is similar to the UU-2 pattern (Figure 4C), in which the sulfur atom forms a hydrogen bond with N3 of U10 (Figure 4F). The $S2-N3$ distance of 3.4 Å is reasonable for a hydrogen bond considering that the atomic radius of sulfur is ~ 0.3 Å greater than that of oxygen. Thiones have previously been observed to act as hydrogen bond acceptors in a few thoroughly studied small molecule systems;^{24–28} our structure provides strong evidence that thiones such as that in s^2U can also act as hydrogen bond acceptors in RNA structures. Given the large increase in T_M of the $s^2U:U$ duplex compared to the $U:U$ duplex and the fact that only one conformation is observed in

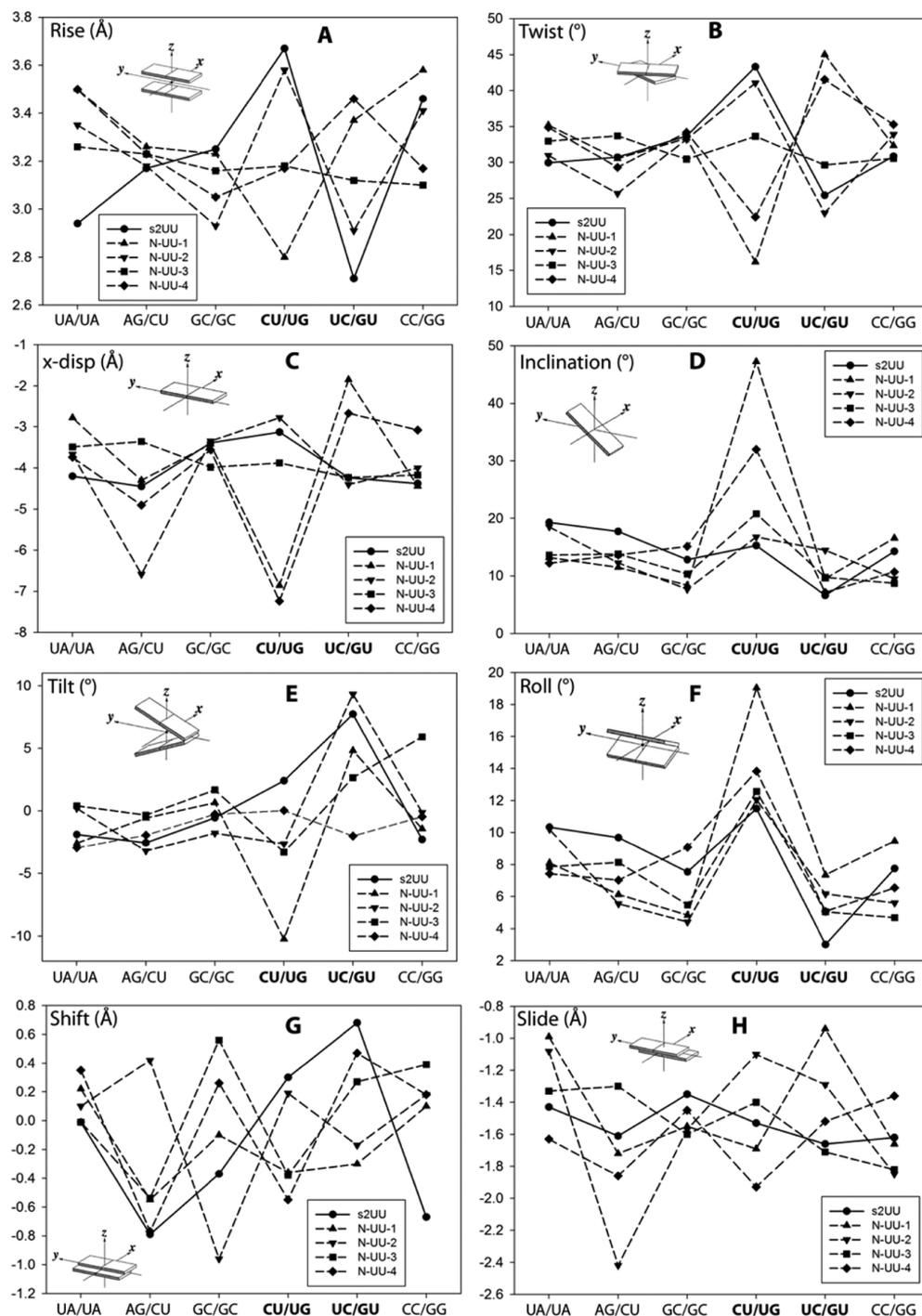


Figure 6. Local base-pair step parameters for duplexes containing $s^2U:U$ or $U:U$ mismatches: (A) rise, (B) twist, (C) x-displacement, (D) inclination, (E) tilt, (F) roll, (G) shift, and (H) slide. Schematics from the program 3DNA.²⁴

the $s^2U:U$ interaction, it appears that replacing the oxygen with sulfur has greatly stabilized the $UU-2$ pairing pattern (compare Figure 4C and F). Additional structural studies along with detailed computational simulations will be required to confirm the generality of this type of sulfur-mediated hydrogen bonding interaction, which might be also affected by sequence-specific, position-specific, or crystal packing effects.

Effects of $U:U$ Pairing and 2-Thiolation on Base-Pair Conformations. In order to further explore the conformational variability of $U:U$ pairing and the effects of 2-thio- U on $U:U$ pairing, we calculated the geometric parameters of all of

the base-pairs and base-pair steps in the $s^2U:U$ -duplex and each of the four $U:U$ -duplexes using the 3DNA software tools (Supplementary Tables S1 and S2).²⁹ Figure 5 summarizes the parameters describing each base-pair. Most parameters at the $U:U$ site are clearly different from the normal base-pair parameters within each duplex. The most obvious base pairing perturbations introduced by the $U:U$ pairing are in shear, opening, and stretch. As expected, the four different $U:U$ pairing patterns in the native duplex cause wide variations, especially in shear (-2 to 2 Å), opening (-12° to 7°), stretch (-1.9 to -0.8 Å), propeller (-22° to -10°), and stagger (-0.4

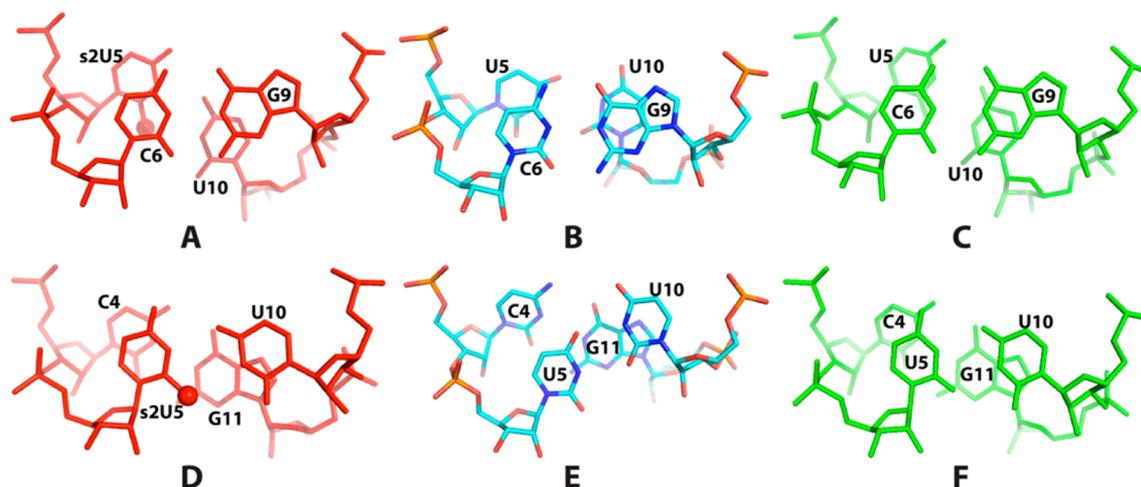


Figure 7. Stacking interactions of the two base-pair steps in the $s^2U:U$ -containing duplex and native UU-1 and UU-2 duplexes: (A) $s^2U5-C6/G9-A10$ step in the s^2UU -duplex; (B) U5-C6/G9-A10 step in the native duplex UU-1; (C) U5-C6/G9-A10 step in the native duplex UU-2; (D) C4- $s^2U5/A10-G11$ step in the s^2UU -duplex; (E) C4-U5/A10-G11 step in the native duplex UU-1; and (F) C4-U5/A10-G11 step in the native duplex UU-2. The color code is same as in Figures 3 and 4. Sulfur atoms are labeled as red spheres.

to 0.4 Å), confirming the multiple distinct geometries and the highly flexible pairing properties of U:U mismatches. Figure 6 depicts the structural parameters for all base-pair steps in each duplex. As expected, the four different U:U pairing patterns in the native duplex cause a wide range of structural variation. Although the U:U pair displays large perturbations for the two flanking steps, these effects compensate for each other in rise, twist, x-displacement and tilt (Figure 6A, B, C, and E). In addition, the step differences in shift and slide (Figure 6G, H) caused by the U:U pair are minimal relative to the normal base-pair steps. Therefore, the perturbing effects of the U:U pair on the overall duplex structure appear to result mainly from increased inclination and roll (Figure 6D and F). The integration of these two effects could significantly decrease the stacking interactions of the U:U pair with the neighboring base-pairs, which may be the main reason for the reduced thermal stability of U:U-containing duplexes.

The $s^2U:U$ local structural parameters are generally similar to one of the native U:U patterns. Of the base-pair parameters, the most significant change associated with 2-thiolation is buckle (Figure 5D). The $s^2U:U$ pair shows an $\sim 7^\circ$ negative buckle, while the U:U pairs show positive values from 0.6° to 12° (Supplementary Table S1). Consistent with the previous base-pairing studies, the $s^2U:U$ pair shows geometries very similar to those of the native UU-2 conformation in shear (Figure 5A), stretch (Figure 5C), and propeller twist (Figure 5E). Similar results are observed in most of the other base-pair step parameters including twist, x-displacement, inclination, tilt, roll, and shift (Figure 6B–G). Surprisingly, in terms of opening (Figure 5B), stagger (Figure 5F), and slide (Figure 6H), the $s^2U:U$ pair is most similar to the geometry seen in the UU-3 conformation (Figure 4D). Similarly to the native duplexes, neighboring base-pair steps exhibit compensatory geometric changes in rise, twist, x-displacement, and tilt. As a result the overall average values of these parameters are quite close to each other (Supplementary Table S2) in native and $s^2U:U$ duplexes, despite the individual base-pair steps in the two structures showing significant differences. This provides further structural evidence that the RNA duplex is flexible enough to accommodate base modifications while minimizing the overall structural perturbation.

Base Stacking and Enhancement of Duplex Stability by 2-Thiolation.

Previous biophysical studies and simulations suggest that the more polarizable sulfur atom could make the stacking interactions of s^2U with its neighboring bases more favorable^{13,18} and thereby increase the overall duplex stability. In order to explore in more detail the effects of uridine 2-thiolation on stacking interactions, we calculated the overlap areas of the two base-pair steps containing s^2U in each duplex using the program 3DNA. Comparing the U:A and $s^2U:A$ duplexes, 2-thiolation only causes a very small increase in the overlap areas of the U5C6/G9A10 and C4U5/A10G11 steps (~ 0.5 and ~ 0.2 Å² respectively) and only a 1 Å² increase in total overlap area in the whole duplex. A direct comparison of these base-pair steps by superimposition is shown in Supplementary Figure S2-A and B. In addition, the distance between S2 of s^2U5 and N1 of C6 is 3.7 Å, the same as in its native counterpart. Thus, any changes in the energetics of the stacking interactions due to s^2U have occurred in the absence of significant geometrical changes and must be due to the enhanced polarizability of sulfur relative to oxygen.

We have compared the UU-1 and UU-2 conformations of the U:U duplex with the $s^2U:U$ duplex. Views of the overlap of the U5C6/G9A10 and C4U5/A10G11 steps are shown in Figure 7. Although the overlap area of the U5-C6 step is similar in all three cases (2.88, 3.31, 2.76 Å²) (Figure 7A–C), the G9-U10 overlap areas for UU-1 and UU-2 (7.22 and 3.56 Å²) are greater than in the $s^2U:U$ structure (1.97 Å²). Similarly, the total overlap areas of C4- $s^2U5/A10G11$ (Figure 7D) are less than those of the UU-1 and UU-2 duplexes (Figure 7E and F) (2.04, 2.26, and 3.12 Å², respectively). As a result, the total overlap area in the whole $s^2U:U$ duplex is smaller than that in the native duplexes UU-1 and UU-2 (28.29, 33.77, and 30.24 Å², respectively). Similar overlap patterns are also observed in UU-3 and UU-4 duplexes (Supplementary Figure S3). Again, it appears that any favorable changes in the energetics of the stacking interactions due to s^2U have occurred in the absence of significantly increased base overlap. It is also possible that the enhanced polarizability of sulfur increases the ability of the sulfur to form a hydrogen bond with N3-H of U10.

Replacing uridine with 2-thiouridine has been proposed to stabilize the 3'-endo sugar pucker conformation,³⁰ which would

preorganize the single-stranded RNA structure, thereby decreasing the entropic cost of duplex formation and increasing duplex stability. It is difficult to test this hypothesis directly on the basis of the duplex crystal structures, since all nucleotides in the duplex have a 3'-*endo* conformation. However, the experimental B-factors can be used to evaluate the relative disorder (dynamic and static) of each atom. In the duplexes containing the s²U:A and U:A base-pairs, the average B-factors for the U5 sugar are 5.66 and 17.37, respectively. Thus, the ribose moiety of s²U5 is much more highly ordered than the sugar of the native U5. However, this increased order also extends to the entire duplex structure, with the s²U:A duplex having a much smaller average B factor (7.36) than the native duplex (18.7).

In the U:U-containing duplexes, the presence of s²U locks the U:U mismatch into one specific pattern, instead of the four different conformations observed in the native U:U duplexes. The increased order conferred by s²U is also seen in a B-factor comparison: the average B-factor for the s²U5 sugar in the s²U:U-duplex is 13.0, much smaller than observed in the four native duplexes (32.0, 27.8, 34.4, and 36.7, respectively), and the average B-factors of atoms in the RNA for the s²U:U and native U:U duplexes are 15.48 and 37.30. Determining the extent to which these ordering effects derive from a constrained sugar conformation in s²U, or from other less direct effects, will require detailed modeling studies.

Since a decreased energetic cost of desolvation has been cited as a possible reason for the duplex stabilization conferred by s²U,¹⁶ we examined the pattern of bound water molecules in the vicinity of s²U (and the corresponding normal U) in both the matched and mismatched duplex structures (Supplementary Figure S4). Interestingly there are two well-defined waters that are hydrogen-bonded to the sulfur in the s²U:A structure, and only one water that is hydrogen-bonded to O2 in the U:A structure. In the U:U structures, there are no well-defined waters close to either S2 or O2. The larger sulfur of s²U may interact with more water molecules than the O2 of U in the unpaired state; however since the strength and the number of these interactions is unknown, and an unknown number of waters are lost during duplex formation in each case, it is not possible at this point to determine whether or not desolvation energetics contributes to the duplex stabilization conferred by s²U.

The overall duplex stabilization caused by replacing U with s²U could potentially result from any one of, or some combination of, the distinct factors explored in this study, including enhanced stacking interactions (and in the case of the s²U:U mismatch, enhanced hydrogen bonding) due to the greater size and polarizability of the sulfur, the more constrained sugar conformation, and possible effects on interactions with bound water molecules. The two new high resolution s²U-containing RNA duplex structures presented here should provide useful starting points for the detailed modeling studies that will be required to help to disentangle these effects. We are attempting to extend our studies of s²U with structural and computational studies of s²U in different sequence and positional contexts, including the other two possible mismatches, s²U:C and s²U:G, with the goal of providing additional insight into the effects of 2-thiolation on RNA base-pair specificity and diversity.

■ EXPERIMENTAL SECTION

Synthesis and Deprotection of s²U and Native RNA Oligonucleotides. The s²U phosphoramidite was synthesized according to literature protocols.¹² All RNA oligonucleotides were chemically synthesized at the 1.0- μ mol scale by solid phase synthesis. 2'-TBDMS-protected RNA phosphoramidites were obtained from Chemgenes (Wilmington, MA). The s²U-phosphoramidite was dissolved in acetonitrile to a concentration of 0.1 M. Coupling was performed using 5-(benzylmercapto)-1H-tetrazole (0.25 M) in acetonitrile for 10 min; 0.02 M I₂ in THF/pyridine/H₂O solution was used as a mild oxidizing reagent to prevent oxidation of the s²U during oligonucleotide synthesis. All other reagents were obtained from Glen Research (Sterling, VA). Synthesis was performed on the appropriate nucleoside immobilized via a succinate linker to controlled-pore glass (CPG-500). All oligonucleotides were prepared in DMTr-off form. After synthesis, RNAs were cleaved from the solid support and fully deprotected with concentrated NH₃ (aq)/EtOH (3:1 v/v) at 55 °C overnight. Solvent was completely removed by Speed-Vac concentration and the dried material was treated with 1 mL of Et₃N·3HF at room temperature for 8 h. The reaction was quenched with 1 mL of water, and the RNA was precipitated by adding 0.2 mL of 3 M sodium acetate and 6 mL of *n*-butanol. The solution was cooled to -30 °C for 1 h before the RNA was recovered by centrifugation and finally dried under vacuum.

HPLC Purification and Analysis. RNA oligonucleotides were purified by ion-exchange HPLC using a PA-100 column from Dionex at a flow rate of 1 mL/min. Buffer A was pure water, and buffer B contained 2 M ammonium acetate (pH 7.1). The RNA oligonucleotides were eluted with a linear gradient of 0% to 35% buffer B over 20 min. Collected fractions were lyophilized, desalted with Waters Sep-Pac C18 columns and reconcentrated. All samples were verified by LC-MS. LC-MS analysis was performed using an Agilent 6520 Q-TOF mass analyzer and 1200 series HPLC with a Waters XBridge C18 column (3.5 μ m, 1 mm \times 100 mm). Mobile phase A was aqueous 200 mM HFIP and 3 mM TEA at pH 7.0, and mobile phase B was methanol. The HPLC method for 35 μ L of a 2.5 μ M solution was a linear increase of 5% to 50% B over 20 min at 0.1 mL/min, with the column heated to 60 °C. Sample elution was monitored by absorbance at 260 nm, and the eluate was passed directly into an ESI source with 325 °C drying nitrogen gas flowing at 8.0 L/min, a nebulizer pressure of 30 psig, and a capillary voltage of 3500 V. Agilent MassHunter Qualitative Analysis software was used to analyze the MS data.

Thermal Denaturation Studies. Stock solutions of duplex RNAs (1 mM) were prepared by dissolving the purified RNAs in HEPES buffer (200 mM, pH 7.5) containing either 100 mM NaCl or 100 mM MgCl₂. The solutions were heated to 85 °C for 3 min, then cooled down slowly to room temperature, and stored at 4 °C for 2 h before *T_M* measurements. Thermal denaturation was performed at a duplex RNA concentration of 100 μ M in an Agilent Cary 60 ultraviolet spectrophotometer with a Quantum Northwest LC 600 temperature controller. The temperature reported is the block temperature. Melting curves were acquired at 260 nm by heating and cooling from 4 to 89 °C twice at a rate of 1 °C/min.

Crystallization and Diffraction Data Collection. RNA samples (1 mM duplex) were heated to 80 °C for 3 min, cooled slowly to room temperature, and placed at 4 °C overnight before crystallization. Nucleic Acid Mini Screen Kits (Hampton Research), Natrix (Hampton Research), and Nuc-Pro-HTS (Jena Bioscience) were used to screen crystallization conditions at different temperatures using the hanging drop method. Perfluoropolyether was used as cryoprotectant for crystal mounting. Data was collected under a liquid nitrogen stream at -174 °C. All diffraction data was collected at beamlines ALS 8.2.2 and 8.2.1 at Lawrence Berkeley National Laboratory. A number of crystals were scanned to find one that diffracted with the highest resolution. Data was collected at a wavelength of 1.0 Å. Crystals were exposed for 1 s per image with a 1° oscillation angle. All data were processed using HKL2000 and DENZO/SCALEPACK.³¹

Structure Determination and Refinement. All four RNA structures presented here were solved by molecular replacement with PHASER using PDB structure 434D (a similar RNA 7-mer duplex with a U5:G10 wobble pair) as the search model, followed by refinement using Refmac. The usual refinement protocol included 10 cycles of simulated annealing, positional refinement, restrained B-factor refinement, and bulk solvent correction. The stereochemical topology and geometrical restraint parameters of DNA/RNA were applied.³² The topologies and parameters for 2-thio-uridine were constructed using Jligand.³³ After several cycles of refinement, a number of highly ordered waters were added. Cross-validation³⁴ with a 10% test set was monitored during the refinement. The σ_A -weighted maps³⁵ of the $(2mF_o - DF_c)$ and the difference $(mF_o - DF_c)$ density maps were computed and used throughout the model building.

■ ASSOCIATED CONTENT

🔗 Supporting Information

Detailed base-pair, base-pair step, and base-pair helical parameters of the native and s²U:U mismatch containing duplexes; base stacking comparison of the native and the s²U:A-containing duplex. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Accession Codes

The four RNA structures have been deposited in Protein Data Bank (www.rcsb.org) with the PDB IDs 4U34 (s²U:A duplex), 4U35 (s²U:U duplex), 4U37 (native U:A duplex), and 4U38 (native U:U duplex).

■ AUTHOR INFORMATION

Corresponding Author

szostak@molbio.mgh.harvard.edu

Present Addresses

[†]Department of Chemistry, the RNA Institute, University at Albany, State University of New York, Albany, NY 12222.

[‡]Lathrop & Gage, LLP, Boston, MA 02109.

[§]Ra Pharmaceuticals, Inc., Cambridge, MA 02139.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Jianhua Gan for help with structure determination; Dr. Garib Murshudov for help with Jligand; and Dr. Aaron E. Engelhart, Dr. Shenglong Zhang, and Dr. Li Li for helpful discussions. All diffraction data were collected at the Advanced Light Source (ALS) beamlines 8.2.1 and 8.2.2. The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract DE-AC02-05CH11231. This work was supported in part by a grant from the Simons Foundation. J.W.S. is an Investigator of the Howard Hughes Medical Institute.

■ REFERENCES

- (1) Fedor, M. J. *Annu. Rev. Biophys.* **2009**, *38*, 271–299.
- (2) Doudna, J. A.; Lorsch, J. R. *Nat. Struct. Mol. Biol.* **2005**, *12*, 395–402.
- (3) Henkin, T. M. *Genes Dev.* **2008**, *22*, 3383–3390.
- (4) Roth, A.; Breaker, R. R. *Annu. Rev. Biochem.* **2009**, *78*, 305–334.
- (5) Montange, R. K.; Mondragon, E.; van Tyne, D.; Garst, A. D.; Ceres, P.; Batey, R. T. *J. Mol. Biol.* **2010**, *396*, 761–772.

- (6) Machnicka, M. A.; Milanowska, K.; Osman Oglou, O.; Purta, E.; Kurkowska, M.; Olchowik, A.; Januszewski, W.; Kalinowski, S.; Dunin-Horkawicz, S.; Rother, K. M.; Helm, M.; Bujnicki, J. M.; Grosjean, H. *Nucleic Acids Res.* **2013**, *41*, D262–267.

- (7) Agris, P. F. *Nucleic Acids Res.* **2004**, *32*, 223–238.

- (8) Rozenski, J.; Crain, P. F.; McCloskey, J. A. *Nucleic Acids Res.* **1999**, *27*, 196–197.

- (9) Rodriguez-Hernandez, A.; Spears, J. L.; Gaston, K. W.; Limbach, P. A.; Gampfer, H.; Hou, Y. M.; Kaiser, R.; Agris, P. F.; Perona, J. J. *J. Mol. Biol.* **2013**, *425*, 3888–3906.

- (10) Rajamani, S.; Ichida, J. K.; Antal, T.; Treco, D. A.; Leu, K.; Nowak, M. A.; Szostak, J. W.; Chen, I. A. *J. Am. Chem. Soc.* **2010**, *132*, 5880–5885.

- (11) Zhang, S.; Blain, J. C.; Zielinska, D.; Gryaznov, S. M.; Szostak, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 17732–17737.

- (12) Kumar, R. K.; Davis, D. R. *Nucleic Acids Res.* **1997**, *25*, 1272–1280.

- (13) Testa, S. M.; Disney, M. D.; Turner, D. H.; Kierzek, R. *Biochemistry* **1999**, *38*, 16655–16662.

- (14) Smith, W. S.; Sierzputowska-Gracz, H.; Sochacka, E.; Malkiewicz, A.; Agris, P. F. *J. Am. Chem. Soc.* **1992**, *114*, 7989–7997.

- (15) Agris, P. F.; Sierzputowska-Gracz, H.; Smith, W.; Malkiewicz, A.; Sochacka, E.; Nawrot, B. *J. Am. Chem. Soc.* **1992**, *114*, 2652–2656.

- (16) Siegfried, N. A.; Kierzek, R.; Bevilacqua, P. C. *J. Am. Chem. Soc.* **2010**, *132*, 5342–5344.

- (17) Vormbrock, R.; Morawietz, R.; Gassen, H. G. *Biochim. Biophys. Acta* **1974**, *340*, 348–358.

- (18) Mazumdar, S. K.; Saenger, W. *J. Mol. Biol.* **1974**, *85*, 213–219.

- (19) Diop-Frimpong, B.; Prakash, T. P.; Rajeev, K. G.; Manoharan, M.; Egli, M. *Nucleic Acids Res.* **2005**, *33*, 5297–5307.

- (20) Mueller, U.; Schubel, H.; Sprinzl, M.; Heinemann, U. *RNA* **1999**, *5*, 670–677.

- (21) Davies, R. W.; Waring, R. B.; Ray, J. A.; Brown, T. A.; Scazzocchio, C. *Nature* **1982**, *300*, 719–724.

- (22) Michel, F.; Netter, P.; Xu, M. Q.; Shub, D. A. *Genes Dev.* **1990**, *4*, 777–788.

- (23) Sintim, H. O.; Kool, E. T. *J. Am. Chem. Soc.* **2006**, *128*, 396–397.

- (24) Dabrowska, U.; Dabrowski, J. *Adv. Mol. Relax. Interact. Processes* **1973**, *5*, 81–87.

- (25) Francuski, B. M.; Novakovic, S. B.; Bogdanovic, G. A. *CrystEngComm* **2011**, *13*, 3580–3591.

- (26) Krystkowiak, E.; Koput, J.; Maciejewski, A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 8842–8851.

- (27) Simanek, E. E.; Tsoi, A.; Wang, C. C. C.; Whitesides, G. M.; McBride, M. T.; Palmore, G. T. R. *Chem. Mater.* **1997**, *9*, 1954–1961.

- (28) Valdés-Martínez, J.; Hernández-Ortega, S.; Rubio, M.; Li, D.; Swearingen, J.; Kaminsky, W.; Kelman, D.; West, D. *J. Chem. Crystallogr.* **2004**, *34*, 533–540.

- (29) Lu, X. J.; Olson, W. K. *Nat. Protoc.* **2008**, *3*, 1213–1227.

- (30) Yokoyama, S.; Watanabe, T.; Murao, K.; Ishikura, H.; Yamaizumi, Z.; Nishimura, S.; Miyazawa, T. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 4905–4909.

- (31) Otwinowski, Z.; Minor, W. *Method Enzymol.* **1997**, *276*, 307–326.

- (32) Parkinson, G.; Vojtechovsky, J.; Clowney, L.; Brunger, A. T.; Berman, H. M. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1996**, *52*, 57–64.

- (33) Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 431–440.

- (34) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *54*, 905–921.

- (35) Read, R. J. *Acta Crystallogr., Sect. A* **1986**, *42*, 140–149.