

Novel RNA base pair with higher specificity using single selenium atom

Huiyan Sun, Jia Sheng, Abdalla E. A. Hassan, Sibio Jiang, Jianhua Gan and Zhen Huang*

Department of Chemistry, Georgia State University, Atlanta, Georgia, 30303, USA

Received October 23, 2011; Revised December 8, 2011; Accepted December 27, 2011

ABSTRACT

Specificity of nucleobase pairing provides essential foundation for genetic information storage, replication, transcription and translation in all living organisms. However, the wobble base pairs, where U in RNA (or T in DNA) pairs with G instead of A, might compromise the high specificity of the base pairing. The U/G wobble pairing is ubiquitous in RNA, especially in non-coding RNA. In order to increase U/A pairing specificity, we have hypothesized to discriminate against U/G wobble pair by tailoring the steric and electronic effects at the 2-exo position of uridine and replacing the 2-exo oxygen with a selenium atom. We report here the first synthesis of the 2-Se-U-RNAs as well as the 2-Se-uridine (^{Se}U) phosphoramidite. Our biophysical and structural studies of the ^{Se}U-RNAs indicate that this single atom replacement can indeed create a novel U/A base pair with higher specificity than the natural one. We reveal that the ^{Se}U/A pair maintains a structure virtually identical to the native U/A base pair, while discriminating against U/G wobble pair. This oxygen replacement with selenium offers a unique chemical strategy to enhance the base pairing specificity at the atomic level.

INTRODUCTION

DNA and RNA are crucial genetic information carriers (1,2). The base pairs of DNAs (T/A and C/G) and RNAs (U/A and C/G) need to be highly specific and accurate for the purpose of the precise genetic information storage, replication, transcription and translation. However, the wobble base pairs, where U in RNA (or T in DNA) pairs with G instead of A, may compromise the high specificity of the base pairing. In RNA, especially non-coding RNA, U/G wobble pair (Figure 1) is ubiquitous (3) and sometimes it has the similar stability as the Watson–Crick

U/A pair (4,5). U/G wobble pair offers unique structural and thermodynamic features (3–5). On the one hand, the U/G pairing increases structure and function diversities of RNA (6). But on the other hand, it may jeopardize the pairing specificity and can cause potential mutations in RNA transcription and protein translation. Codon–anticodon mismatch or misreading is observed with an error frequency at 10^{-5} or higher, which may affect the accuracy of synthesized proteins (7–9). For instance, the first position of the codon–anticodon interaction with wobble mismatch (U/G) was discovered in *Escherichia coli* (error frequency = 0.1%) with 100-fold higher than the normal error level (9). In this mis-incorporation of serine (codon: AGC) (9), glycine codon (GGC) in mRNA is recognized by Ser-charged tRNA (anticodon: GCU) instead of Gly-charged tRNA (anticodon: GCC). Similarly, the second position of the codon–anticodon interaction with wobble mismatch (U/G) was also observed, where Lys (codon: AAA) is mis-incorporated instead of normal incorporation of Arg (codon: AGA), with much higher error frequency (5–12%) (10). To avoid the negative impact of the wobble pairing on the level of protein synthesis, the genetic codes with degeneracy are used to deal with the consequence of the wobble pairing. Thus, wobble pairing is often observed at the third codon position through the codon degeneracy to limit errors. However, the codons forming the Watson–Crick pairs with tRNA anticodons are still preferred (11,12). Study shows that the third codon position with a Watson–Crick base pair can reduce the frequency of amino acid mis-incorporation by nearly 10-fold, and it is much more accurate than that with a wobble pair for the same amino acid (13). Nevertheless, the 3-nt genetic codes that accommodate the wobble pairing are used as the most ideal countermeasure at the level of protein synthesis in living organisms (14). Clearly, on the basis of the chemical principle, this degeneracy strategy properly guarantees the translation accuracy at the protein level by tolerating wobble pairs and silent mutations at the RNA and DNA levels.

Since the 2-exo-oxygen of uridine plays a significant role in U/G wobble pair, we hypothesized that tailoring the

*To whom correspondence should be addressed. Tel/Fax: +1 404 413 5535; Email: huang@gsu.edu

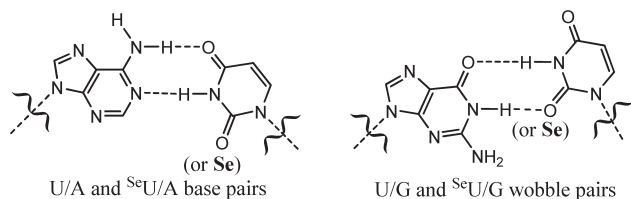


Figure 1. Native and Se-modified U/A pairs and U/G wobble pairs.

steric and electronic effects at this site may discriminate against the wobble pair, enabling the modified U/A base pair with higher specificity. Interestingly, selenium has been discovered in natural tRNAs in the 2-Se-uridine form, i.e. 5-methylaminomethyl-2-selenouridine ($\text{mm}^5\text{se}^2\text{U}$), in the wobble position on the anticodon loop (15,16). The function of such selenium modification is not completely clear yet, though it was proposed that such Se derivatization on tRNAs probably improves the accuracy and efficiency of protein translation (17). Similarly, the corresponding sulfur modification has been observed on natural tRNAs (18). Sulfur was chemically introduced to the 2-position of uridine (19,20). The S-modified U/G pair is slightly less stable than the native U/G pair (5), while the $^{\text{S}}\text{U/A}$ is more stable over the native U/A pair. Thus, we hypothesized that the 2-oxygen replacement with selenium ($^{\text{Se}}\text{U}$, Figure 1) can destabilize and discriminate against the U/G wobble pair, because the atomic size of selenium (1.16 Å) is larger than that of sulfur (1.02 Å) and oxygen (0.73 Å). Moreover, selenium has the least ability to form a hydrogen bond among O, S and Se, which weakens the hydrogen bond originally formed by the 2-oxygen of the wobble pair. Thus, it is expected that this 2-Se-replacement can largely destabilize U/G pair by generating a steric hindrance against the pair and significantly weakening the hydrogen bond. Furthermore, it is expected that the 2-Se-substitution does not significantly affect the hydrogen bonds within the U/A pair, since the 2-oxygen is not directly involved in the U/A base pairing. Therefore, we decided to incorporate selenium into the 2-position of uridine in RNA, in order to atom-specifically increase the U/A pair specificity and disrupt the U/G wobble.

MATERIALS AND METHODS

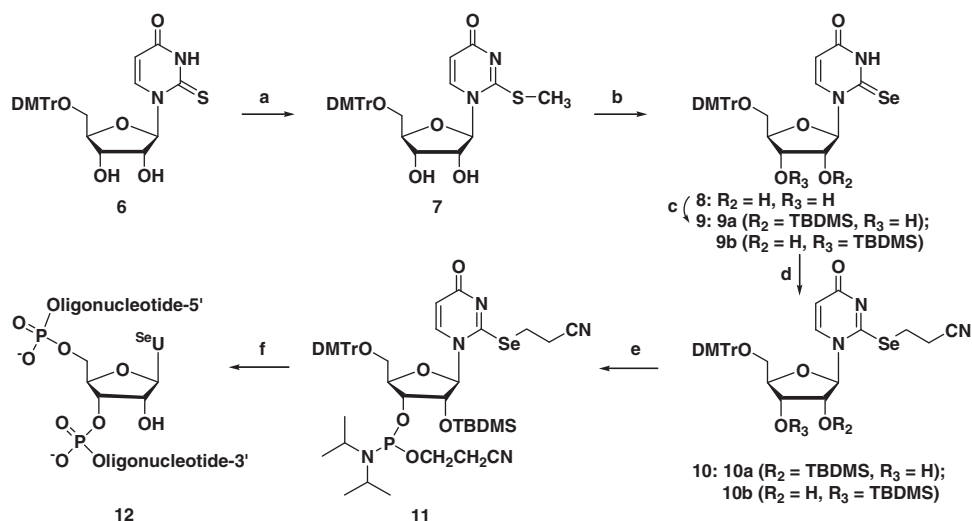
Synthesis of 2-Se-uridine phosphoramidite

1-(5'-O-4,4'-dimethoxytrityl-β-D-ribofuranosyl)-2-methylthiouridine 7. Five grams of dry compound 6 (its synthesis: Scheme S1 in Supplementary Data) was dissolved in dry *N,N*-dimethylformamide (DMF), followed by addition of iodomethane (5.5 ml, 89 mmol), 1,8-Diazabicyclo[5.4.0]undec-7-ene (2 ml, 13.3 mmol) was then added to the reaction mixture at 0°C. The reaction was monitored by thin layer chromatography (TLC) plate (12% methanol in dichloromethane, $R_f = 0.4$) and completed in 4 h. Ethyl acetate (50 ml) was poured into the mixture and DMF was removed by washing the organic layer with saturated sodium chloride solution. The organic phase was dried over anhydrous magnesium sulfate and

evaporated under reduced pressure. The residue was purified by flash column chromatography (10% methanol in dichloromethane) and pure compound 7 was obtained in 95% yield. ^1H NMR (CDCl_3) δ : 7.87 (d, $J = 7.7$ Hz, 1H, H-6), 7.44–7.20 (m, 9H, Ar), 6.85 (m, 4H, Ar), 6.11 (br, 1H, OH), 5.88 (d, $J = 6.0$ Hz, 1H, H-1'), 5.54 (d, $J = 7.7$ Hz, 1H, H-5), 4.63 (m, 1H, H-4'), 4.44 (m, 1H, H-3'), 4.24 (d, $J = 2.3$ Hz, 1H, H-2'), 3.75 (d, $J = 3.1$ Hz, 6H, OCH_3), 3.42 (m, 2H, H-5'), 3.40–3.30 (br, 1H, OH), 2.55 (s, 3H, SCH_3); ^{13}C NMR (CDCl_3) δ : 169.19 (C-4), 164.36 (C-2), 158.88 (Ar), 144.49 (Ar), 140.13 (C-6), 135.37 (Ar), 135.22 (Ar), 130.41 (Ar), 130.28 (Ar), 128.32 (Ar), 128.28 (Ar), 127.29 (Ar), 113.54 (Ar), 108.92 (C-5), 91.95 (C-1'), 87.35 (C-Ar₃), 84.82 (C-4'), 75.24 (C-2'), 71.63 (C-3'), 63.40 (C-5'), 55.40 (OCH_3), 15.39 (SCH_3); High resolution mass spectra (HRMS) electrospray ionization-time of flight (ESI-TOF) $[\text{M} + \text{H}^+] = 577.2003$ (calc. 577.2008), Chemical formula: $\text{C}_{31}\text{H}_{33}\text{N}_2\text{O}_7\text{S}$.

1-(5'-O-4,4'-dimethoxytrityl-β-D-ribofuranosyl)-2-selenouridine 8. A solution of NaSeH was generated by addition of absolute ethanol (50 ml) to selenium (6.2 g, 78 mmol) and sodium borohydride (NaBH_4 , 4.43 g, 0.117 mol) at 0°C. The reaction was completed in 2 h and a clear solution was formed. The ethanolic solution was added to compound 7 (4.5 g, 7.80 mmol) and the mixture was stirred for 8 h under argon. The reaction mixture was then concentrated under reduced pressure and ethyl acetate (50 ml) was added to the residue. The organic layer was washed with water several times (5×30 ml), and then dried over anhydrous magnesium sulfate. Purification was performed by flash column chromatography (4% methanol in dichloromethane) and the light yellow compound (8) was obtained (85% yield). ^1H NMR (CDCl_3) δ : 10.95 (s, 1H, NH), 8.24 (d, $J = 8.2$ Hz, 1H, H-6), 7.44–7.19 (m, 9H, Ar), 6.84 (m, 4H, Ar), 6.48 (s, 1H, H-1'), 5.66 (d, $J = 8.1$ Hz, 1H, H-5), 4.48 (m, 2H, H-4', H-3'), 4.22 (m, 1H, H-2'), 3.89 (s, 1H, OH), 3.79 (s, 6H, OCH_3), 3.58 (dd, $J = 23.6, 9.2$ Hz, 2H, H-5'), 2.97 (br, 1H, OH); ^{13}C NMR (CDCl_3) δ : 175.74 (C-2), 159.21 (C-4), 158.98 (Ar), 158.94 (Ar), 144.45 (Ar), 140.82 (C-6), 135.38 (Ar), 135.18 (Ar), 130.35 (Ar), 130.27 (Ar), 128.30 (Ar), 128.28 (Ar), 127.45 (Ar), 113.58 (Ar), 108.37 (C-5), 96.86 (C-1'), 87.38 (C-Ar₃), 84.41 (C-4'), 76.33 (C-2'), 69.19 (C-3'), 61.20 (C-5'), 55.48 (OCH_3). HRMS (ESI-TOF) $[\text{M} - \text{H}^-] = 609.1136$ (calc. 609.1140), Chemical formula: $\text{C}_{30}\text{H}_{29}\text{N}_2\text{O}_7\text{Se}$; UV (MeOH): $\lambda_{\text{max}} = 311$ nm (in methanol).

1-(2'-O-tert-butyltrimethylsilyl-5'-O-4,4'-dimethoxytrityl-β-D-ribofuranosyl)-2-selenouridine 9a and *1-(3'-O-tert-butyltrimethylsilyl-5'-O-4,4'-dimethoxytrityl-β-D-ribofuranosyl)-2-selenouridine 9b*. 5'-DMTr-2-selenouridine **8** (0.5 g, 0.82 mmol) was dissolved in dry DMF, then *tert*-butyldimethylsilyl chloride (TBDMS-Cl, 0.15 g, 0.98 mmol) and imidazole (0.11 g, 1.64 mmol) were added into the solution under nitrogen gas. The reaction was monitored by TLC plate (15% ethyl acetate in dichloromethane, $R_f = 0.8$). The mixture was stirred overnight at room temperature and then directly poured into ethyl acetate



Scheme 1. Synthesis of ^{76}Se -phosphoramidite **11** and RNAs (**12**). Reagents and conditions: (a) CH_3I , DBU, DMF; (b) Se, NaBH_4 , EtOH; (c) TBDMS-Cl, imidazole, DMF; (d) $\text{ICH}_2\text{CH}_2\text{CN}$, $(i\text{-Pr})_2\text{NEt}$, CH_2Cl_2 ; (e) $(i\text{-Pr})_2\text{N}_2\text{P}(\text{Cl})\text{OCH}_2\text{CH}_2\text{CN}$, $(i\text{-Pr})_2\text{NEt}$, CH_2Cl_2 ; (f) solid-phase synthesis.

(20 ml) and washed with water (2×20 ml). The organic layer was dried by anhydrous magnesium sulfate and evaporated under reduced pressure. Two compounds, **9a** and **9b**, were obtained. The two regional isomers (ratio 1:1) were purified together by flash column chromatography (10% ethyl acetate in dichloromethane) and were not further separated. Since, it was both challenging and unnecessary to separate each isomer, we decided to move to the next step of synthesis without separation of these two isomers. HR-MS (ESI-TOF, **9a** and **9b**) $[\text{M}-\text{H}^+]^- = 723.1990$ (calc. 723.2005). Chemical formula: $\text{C}_{36}\text{H}_{43}\text{N}_2\text{O}_7\text{SeSi}$.

1-(2'-O-tert-butyl dimethylsilyl-5'-O-4,4'-dimethoxytrityl- β -D-ribofuranosyl)-2-cyanoethylselenyluridine 10a and *1-(3'-O-tert-butyl dimethylsilyl-5'-O-4,4'-dimethoxytrityl- β -D-ribofuranosyl)-2-cyanoethylselenyluridine 10b*. The mixture (0.52 g, 0.72 mmol) of **9a** and **9b** was dissolved in dried dichloromethane at 0°C . Iodopropionitrile (0.78 g, 4.31 mmol) was added to the solution, followed by addition of diisopropylethylamine (0.37 ml, 2.15 mmol). The reaction was monitored by TLC plates (30% ethyl acetate in dichloromethane). After 4-h reaction, the solvent was removed under reduced pressure and the residue was partitioned between ethyl acetate (20 ml) and water (20 ml). The organic phase was dried over anhydrous magnesium sulfate and evaporated into dryness. Two crude products were obtained: 1-(2'-O-tert-butyl dimethylsilyl-5'-O-4,4'-dimethoxytrityl- β -D-ribofuranosyl)-2-cyanoethylselenyluridine **10a** (Rf = 0.35) and 1-(3'-O-tert-butyl dimethylsilyl-5'-O-4,4'-dimethoxytrityl- β -D-ribofuranosyl)-2-cyanoethylselenyluridine **10b** (Rf = 0.30). These two compounds can be separated by flash column chromatography (15% ethyl acetate in dichloromethane). **10a** was obtained in 0.228 g (41% yield) and **10b** was obtained in 0.235 g (42% yield). **10a**: ^1H NMR (CDCl_3) δ : 7.96 (d, $J = 7.7$ Hz, 1H, H-6), 7.53–7.11 (m, 9H, Ar), 6.85 (m, 4H, Ar), 5.71

(d, $J = 7.7$ Hz, 1H, H-5), 5.60 (d, $J = 6.5$ Hz, 1H, H-1'), 4.61–4.49 (m, 1H, H-4'), 4.31 (m, 2H, H-3', H-2'), 3.80 (s, 6H, OCH₃), 3.54–3.34 (m, 4H, H-5', SeCH₂CH₂CN), 3.01 (m, 2H, SeCH₂CH₂CN), 2.91 (s, 1H, OH), 0.94 (s, 9H, SiCMe₃), 0.09 (d, 6H, SiMe₂); ^{13}C NMR (CDCl_3) δ : 167.70 (C-4), 159.04 (C-2), 158.51 (Ar), 144.17 (Ar), 139.19 (C-6), 134.91 (Ar), 134.74 (Ar), 130.26 (Ar), 130.17 (Ar), 128.30 (Ar), 128.11 (Ar), 127.58 (Ar), 118.78 (CN), 113.58 (Ar), 110.61 (C-5), 93.13 (C-1'), 87.82 (C-Ar₃), 85.39 (C-4'), 77.27 (C-2'), 72.40 (C-3'), 63.82 (C-5'), 55.44 (OCH₃), 25.84 (SiCMe₃), 24.06 (SeCH₂CH₂CN), 18.89 (SeCH₂CH₂CN), 18.14 (SiCMe₃), -4.56 (SiCH₃), -4.91 (SiCH₃). HRMS (ESI-TOF) $[\text{M}+\text{H}^+]^+ = 778.2464$ (calc. 778.2427). Chemical formula: $\text{C}_{39}\text{H}_{48}\text{N}_3\text{O}_7\text{SeSi}$. **10b**: ^1H NMR (CDCl_3) δ : 8.09 (d, $J = 7.7$ Hz, 1H, H-6), 7.32 (m, 9H, Ar), 6.88 (m, 4H, Ar), 5.75 (d, $J = 7.7$ Hz, 1H, H-5), 5.65 (d, $J = 3.7$ Hz, 1H, H-1'), 4.46 (m, 1H, H-2'), 4.21 (dd, $J = 9.3, 5.1$ Hz, 1H, H-4'), 4.18–4.08 (m, 1H, H-3'), 3.83 (s, 6H, OCH₃), 3.70 (m, 1H, H-5'), 3.55 (m, 1H, H-5'), 3.41 (m, 2H, SeCH₂CH₂CN), 3.22 (d, $J = 5.5$ Hz, 1H, OH), 3.08 (m, 2H, SeCH₂CH₂CN), 0.90 (s, 9H, SiCMe₃), 0.11 (d, 6H, SiMe₂). ^{13}C NMR (CDCl_3) δ : 167.90 (C-4), 159.05 (C-2), 157.82 (Ar), 143.99 (Ar), 138.96 (C-6), 135.02 (Ar), 134.89 (Ar), 130.35 (Ar), 130.34 (Ar), 128.37 (Ar), 128.29 (Ar), 127.58 (Ar), 118.95 (CN), 113.56 (Ar), 113.53 (Ar), 110.49 (C-5), 93.84 (C-1'), 87.57 (C-Ar₃), 84.72 (C-4'), 76.17 (C-2'), 71.05 (C-3'), 61.71 (C-5'), 55.48 (OCH₃), 25.82 (SiCMe₃), 24.12 (SeCH₂CH₂CN), 18.92 (SeCH₂CH₂CN), 18.17 (SiCMe₃), -4.59 (SiCH₃), -4.60 (SiCH₃). HRMS (ESI-TOF) $[\text{M}+\text{H}^+]^+ = 778.2401$ (calc. 778.2427). Chemical formula: $\text{C}_{39}\text{H}_{48}\text{N}_3\text{O}_7\text{SeSi}$.

1-[2'-O-tert-butyl dimethylsilyl-3'-O-(2-cyanoethyl-N,N-diisopropylamino) phosphoramidite-5'-O-(4,4'-dimethoxytrityl- β -D-ribofuranosyl)]-2-cyanoethylselenyluridine 11. Diisopropylethylamine (15.5 mg, 0.12 mmol) and 2-cyanoethyl N,N-diisopropylchlorophosphoramidite

(26 mg, 0.11 mmol) were added to a solution of **10a** (100 mg, 0.10 mmol) in dry dichloromethane (5 ml) at room temperature under nitrogen gas. The mixture was monitored by TLC (15% ethyl acetate in dichloromethane). When the reaction was completed in 4 h, rapid Al₂O₃ column chromatography (dichloromethane as the eluent) was performed to remove the organic salts. The solvent was then evaporated under reduced pressure and the residue was dissolved in 0.5 ml dichloromethane and precipitated in dry hexane under vigorous stirring. The precipitate was collected by filtration, dried under reduced pressure and directly used for solid-phase synthesis. ¹H NMR (CDCl₃) δ: 7.95 (d, J = 7.7 Hz, 1H, H-6), 7.30 (m, 9H, Ar), 6.84 (m, 4H, Ar), 5.82 (d, J = 7.6 Hz, 1H, H-1'), 5.63 (d, J = 7.7 Hz, 1H, H-5), 4.68 – 4.43 (m, 1H, H-4'), 4.43 – 4.29 (m, 1H), 4.24 (s, 1H), 3.98 (dd, J = 15.6, 8.4 Hz, 2H), 3.80 (s, 6H, OCH₃), 3.63 (dd, J = 19.8, 7.3 Hz, 4H), 3.42 (dd, J = 19.3, 8.7 Hz, 5H), 3.01 (s, 2H), 2.70 (d, J = 5.8 Hz, 2H), 1.20 (d, J = 6.7 Hz, 18H), 1.08 (d, J = 6.5 Hz, 6H), 0.92 (s, 13H), 0.14 – 0.02 (m, 9H); ¹³C NMR (CDCl₃) δ: 167.83 (C-4), 159.04 (C-2), 158.71 (Ar), 144.16 (Ar), 139.02 (Ar), 134.96 (Ar), 134.74 (Ar), 130.21 (Ar), 130.14 (Ar), 128.38 (Ar), 128.04 (Ar), 127.58 (Ar), 118.86 (SeCH₂CH₂CN), 117.75 (OCH₂CH₂CN), 113.67 (Ar), 110.69 (C-5), 92.50 (C-1'), 87.92 (C-Ar3), 85.43 (C-4'), 77.15 (C-2'), 72.81 (C-3'), 63.72 (C-5'), 59.39 (OCH₂CH₂CN), 55.48 (OCH₃), 43.22-43.10 (NCMe₂), 29.90 (OCH₂CH₂CN), 26.12-25.97 (NCMe₂), 24.85 (SiCMe₃), 24.02 (SeCH₂CH₂CN), 18.94 (SeCH₂CH₂CN), 18.37 (SiCMe₃), -4.35 (SiCH₃), -4.56 (SiCH₃); ³¹P NMR (CDCl₃) δ: 148.81, 152.30. HRMS (ESI-TOF) [M + H]⁺ = 978.3528 (calc. 978.3505). Chemical formula: C₄₈H₆₅N₅O₈PSeSi.

Solid-phase synthesis of the 2-Se-functionalized RNAs

ABI3400 DNA/RNA Synthesizer was used for all the RNA oligonucleotides synthesis (1.0 μmol scale). All the non-modified nucleoside phosphoramidite reagents used were ultra-mild (Glen Research). RNA oligonucleotides were synthesized in DMTr-on form, cleaved from the beads and deprotected by the treatment of 0.05 M K₂CO₃ methanol solution for 10 h at room temperature. After evaporating the solution to dryness, the 2'-TBDMS deprotection was performed in TBAF (0.5 ml, 1 M) for 14 h at room temperature. Then the RNAs were treated with 1 M Tris-HCl buffer (0.5 ml, pH 7.5) for 5 min, followed by concentrating to 0.5 ml and desalting using G-25 Sephadex column. The 5'-DMTr deprotection was then performed using Glen-Pek RNA column, followed by desalting using Sep-Pak Vas column.

HPLC analysis and purification

The RNA oligonucleotides were analyzed and purified by reversed-phase high performance liquid chromatography (RP-HPLC), flow rate 6 ml/min [buffer A: 20 mM triethylammonium acetate (TEAAc, pH 7.1) in water; buffer B: 20 mM TEAAc (pH 7.1) in 50% acetonitrile]. The HPLC analysis was performed with a linear gradient from buffer A to 100% buffer B in 20 min.

Native RNAs were purchased from Integrated DNA Technologies. The concentrations of the native, Se-modified RNAs were adjusted to 1.0 mM in water. The Se-RNA samples were characterized by matrix assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS) (Table 1) and HPLC (Figure 2).

pH titration curve of 2-selenouridine

2-Selenouridine was prepared through detritylation of 1-(5'-O-4,4'-dimethoxytrityl-β-D-ribofuranosyl)-2-selenouridine (**8**) by acid treatment. The 2-selenouridine solutions

Table 1. MALDI-TOF MS of 2-Se-U RNAs

Entry	Oligonucleotide molecular formula	Measured (calc.) <i>m/z</i>
1	5'-rGUAUA ^{Se} UAC-3' C ₇₆ H ₉₄ N ₂₉ O ₅₃ P ₇ Se	[M + H] ⁺ = 2558.7 (2558.5)
2	5'-rAUCACC ^{Se} UCCUUA-3' C ₁₁₁ H ₁₄₁ N ₃₈ O ₈₂ P ₁₁ Se	[M + H] ⁺ = 3740.3 (3740.2)
3	5'-rAAUGC ^{Se} UGCACUG-3' C ₁₁₄ H ₁₄₂ N ₄₅ O ₈₁ P ₁₁ Se	[M + H] ⁺ = 3859.4 (3859.3)

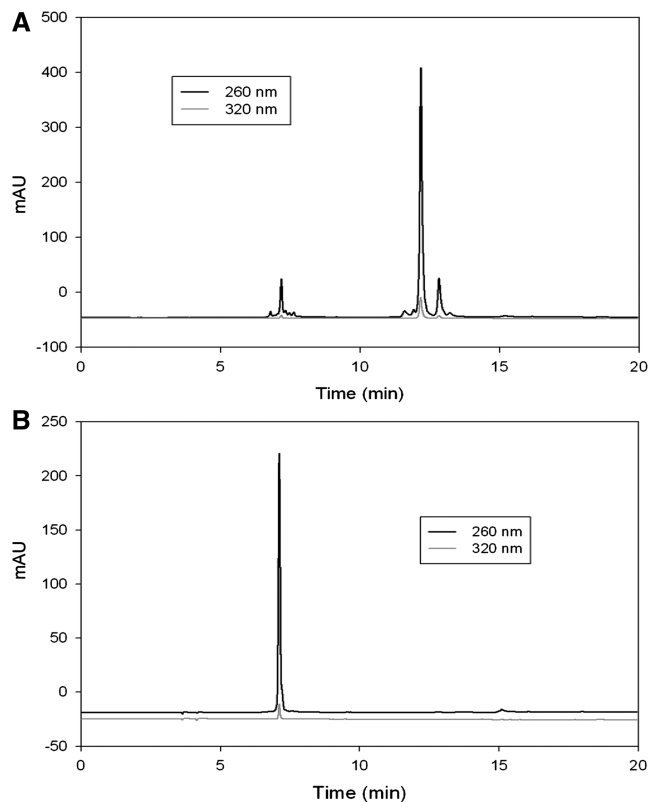


Figure 2. HPLC analysis of 2-Se-U modified RNA 12-mer (5'-rAUCACC^{Se}UCCUUA-3'). (A) The crude DMTr-on Se-RNA, retention time was 12.2 min. (B) The pure DMTr-off Se-RNA with same gradient and buffer, retention time was 7.1 min. Samples were eluted with a linear gradient from buffer A (20 mM triethylammonium acetate, pH 7.1) to 70% buffer B (50% acetonitrile, 20 mM triethylammonium acetate, pH 7.1) in 10 min, to 100% buffer B in 12 min and continuous 100% buffer B to 20 min.

were adjusted to desired pH values in the buffer of 50 mM Na_2HPO_4 at room temperature. The UV-Vis spectra were recorded every 0.1 pH unit between pH 6–8 and every 0.2–0.5 pH unit between pH 4–6 and pH 8–10. The pH of each solution was measured before and after its UV-Vis spectrum collection and the error was within ± 0.02 pH unit. The titration data was plotted and shown in Figure 3.

Thermodenaturation of duplex RNAs

The UV-melting temperature studies were carried out by Cary 300 UV-Vis Spectrophotometer and a temperature control module. The samples (2 μM RNA duplexes) were dissolved in buffer of 150 mM NaCl, 2 mM MgCl_2 and 10 mM Na_2HPO_4 – NaH_2PO_4 (pH 6.8). The samples were heated to 80°C and cooled to room temperature slowly and then kept in 4°C overnight. The heating rate of

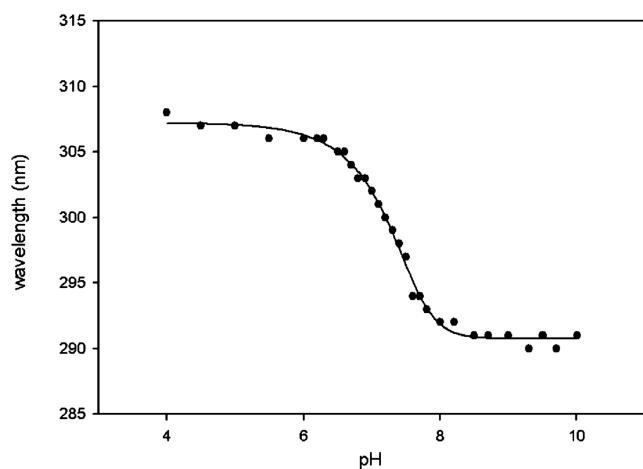


Figure 3. Plot of wavelength (nm) versus pH for 2-selenouridine nucleoside. The fitted titration curve yields the pK_a value (7.29 ± 0.02).

Table 2. Melting temperatures (T_m) of the native, S- and Se-modified RNA duplexes

Entry	Sequences	Base pairs	T_m (°C)
1	I: 5'-rAUCACCUCCUUA-3'		
2	I+3'-rUAGUGGAGGAAU-5'	U/A	62.8
3	I+3'-rUAGUGGGGAAU-5'	U/G	62.5
4	I+3'-rUAGUGGCGAAU-5'	U/C	50.6
5	I+3'-rUAGUGGUGAAU-5'	U/U	48.8
6	II: 5'-rAUCACCS^{Se}UCCUUA-3'		
7	II+3'-rUAGUGGAGGAAU-5'	SeU/A	65.8
8	II+3'-rUAGUGGGGAAU-5'	SeU/G	58.5
9	II+3'-rUAGUGGCGAAU-5'	SeU/C	50.3
10	II+3'-rUAGUGGUGAAU-5'	SeU/U	57.3
11	III: 5'-rAAUGCUGCACUG-3'		
12	III+3'-rUUACGACGUGAC-5'	U/A	64.1
13	III+3'-rUUACGCGUGAC-5'	U/G	59.4
14	III+3'-rUUACGCCGUGAC-5'	U/C	52.0
15	III+3'-rUUACGUGUGAC-5'	U/U	51.5
16	IV: 5'-rAAUGC^{Se}UGCACUG-3'		
17	IV+3'-rUUACGACGUGAC-5'	SeU/A	66.5
18	IV+3'-rUUACGCGUGAC-5'	SeU/G	55.5
19	IV+3'-rUUACGCCGUGAC-5'	SeU/C	51.9
20	IV+3'-rUUACGUGUGAC-5'	SeU/U	58.0

Bold and underlined sequences indicate the pairing and mis-pairing sites.

melting experiment was 0.5°C per min. The melting temperature data and curves of the matched and mismatched duplexes are presented in Table 2 and Figure 4.

Crystallization

The purified RNA oligonucleotide (5'-GUAUA-^{Se}U-AC-3', 1 mM) was heated to 80°C for 2 min, and cooled down slowly to room temperature. The Nucleic Acid Mini Screen Kit (Hampton Research) was applied to screen the crystallization conditions at different temperatures (10, 20 and 25°C) using the hanging-drop method by vapor diffusion.

Data collection

Perfluoropolyether was used as a cryoprotectant during the crystal mounting, and data collection was taken under the liquid nitrogen stream at -174°C . The Se-RNA crystal data were collected at beam line X12B and X12C in NSLS of Brookhaven National Laboratory. A number of crystals were screened to

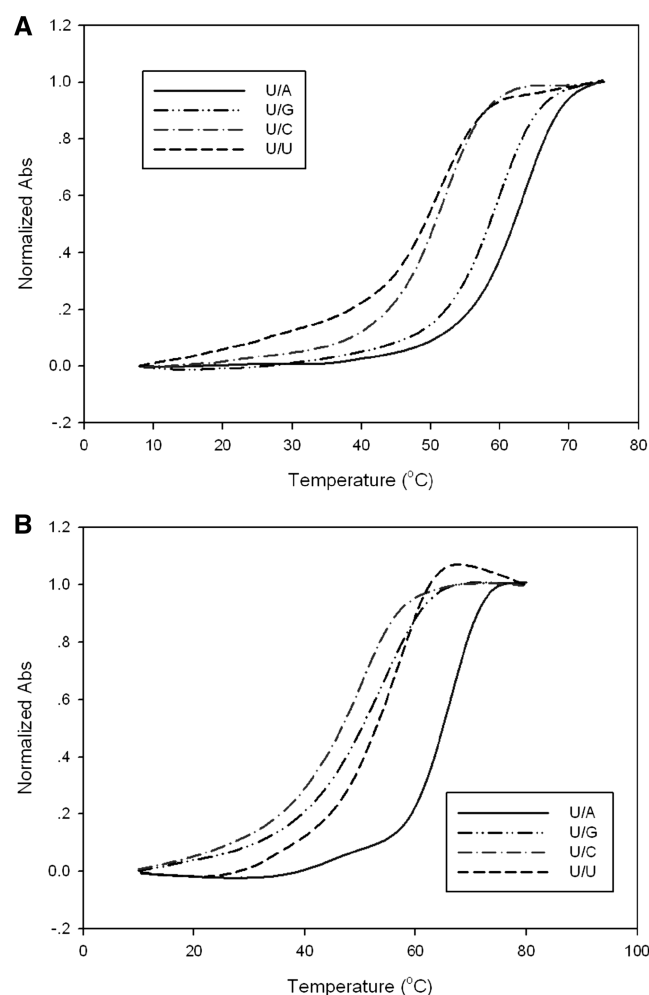


Figure 4. Normalized UV-melting curves of RNA duplexes. (A) Native RNA (5'-rAAUGCUGCACUG-3') paired with matched and mismatched strands. (B) Se-RNA (5'-rAAUGC^{Se}UGCACUG-3') with matched and mismatched strands.

Table 3. Data collection and refinement statistics of ⁷⁶Se-U-RNA

Structure (PDB ID)	GUAUA- ⁷⁶ Se-U-AC (3S49)
Data collection	
Space group	R32
Cell dimensions: <i>a</i> , <i>b</i> , <i>c</i> (Å)	47.095, 47.095, 424.655,
α , β , γ (°C)	90, 90, 120
90, 90, 120	
Resolution range (Å) (last shell)	50.0–2.28 (2.37–2.28)
Unique reflections	9870 (959)
Completeness (%)	99.8 (99.4)
<i>R</i> _{merge} (%)	7.7 (23.5)
<i>I</i> / σ (<i>I</i>)	21.0 (3.9)
Redundancy	18.8 (10.7)
Refinement	
Resolution range (Å)	30.0–2.3
<i>R</i> _{work} (%)	21.4
<i>R</i> _{free} (%)	26.9
Number of reflections	8206
Number of atoms	
Nucleic acid (double)	1162
Heavy atoms and ion	7 Se
Water	71
RMS deviations	
Bond length (Å)	0.007
Bond angle	1.169

identify the one with strong anomalous scattering at the K-edge absorption of selenium. The distance of the detector to the crystals was set to 150 mm. The wavelength of 0.9795 Å was chosen for selenium SAD phasing. The crystals were exposed for 10 or 15 s/image with 1° oscillation, and a total of 180 images were taken for each data set. All the data were processed using HKL2000 and DENZO/SCALEPACK (21).

Structure determination and refinement

The structure of the Se-RNA [5'-GUAUA-⁷⁶Se-U-AC-3']₂ was solved by molecular replacement with both CNS (22) and Phaser (23). The refinement protocol includes simulated annealing, positional refinement, restrained B-factor refinement, and bulk solvent correction. The stereo-chemical topology and geometrical restrain parameters of DNA/RNA (24) have been applied. The topologies and parameters for modified uridine with 2-selenium (US) were constructed and applied. After several cycles of refinement, a number of highly ordered waters were added. Finally, the occupancies of selenium were adjusted. Cross-validation (25) with a 5–10% test set was monitored during the refinement. The σ A-weighted maps (26) of the ($2_m|F_o| - D|F_c|$) and the difference ($m|F_o| - D|F_c|$) density maps were computed and used throughout the model building (Table 3).

RESULTS AND DISCUSSION

Synthesis of 2-Se-uridine phosphoramidite

Though selenium was incorporated into uridine four decades ago (27,28), RNA containing 2-Se-uridine (⁷⁶SeU) has not been synthesized because of the synthetic challenge. Recently, our laboratory has successfully developed a novel strategy to incorporate the selenium functionality

to the 2-position of thymidine in DNA (29). This successful strategy has encouraged us to introduce the selenium functionality to the 2-position of uridine in RNA. Herein, we report the first synthesis of the 2-selenouridine derivatives and RNAs. The synthesis (Scheme S1 in Supplementary Data) started from the glycosidation (30) of the acylated ribofuranose (1) with silylated 2-thiouracil (3), followed by benzoyl deprotection and trityl protection of the 5'-hydroxyl group to offer 6 (31). After methylation of 6 to activate the 2-thio-functionality (29), NaSeH was used to displace the 2-S-functionality and offer the 2-Se-uridine 8 in 85% yield. Following the protections of the 2'-hydroxyl group and the 2-Se-functionality with ICH₂CH₂CN, the Se-phosphoramidite 11 was synthesized by phosphorylation of 10a (29,32,33). The ⁷⁶Se-U-phosphoramidite was finally incorporated into RNAs by solid-phase synthesis. The synthesized ⁷⁶SeU-RNAs (12) were deprotected, purified and confirmed by HPLC and MS (Table 1 and Figure 2). The characterization of the Se-nucleosides and Se-nucleotides is presented in Supplementary Figures S1–S23.

Characterization of the 2-Se-functionalized RNA

After cleavage from solid support and deprotection, the crude DMTr-on RNAs were purified by RP-HPLC and lyophilized to dryness. As shown in Figure 2A, the coupling yield is ~90%. The 5'-DMTr deprotection of the oligonucleotides was then performed using Glen-Pek RNA column. The HPLC analysis of the DMTr-off RNA was shown in Figure 2B. All the pure seleno-RNA oligonucleotides were characterized by MALDI-TOF MS (Table 1).

Thermodenaturation study of 2-Se-uridine RNAs containing match and mismatch base pairs

UV-melting temperatures (*T*_m) of the native and Se-modified duplexes with match and mismatch sequences are shown in Table 2, Figures 4 and 5. *T*_m of the Se-RNA duplex containing the ⁷⁶SeU/A Watson–Crick pair was 3.0°C higher for one duplex (or 2.4°C higher for the other duplex) than those of the corresponding duplexes containing native U/A pair (Table 2). Comparing with native U/G, the ⁷⁶SeU/G pair is ~4°C less stable than the native U/G pair, suggesting that ⁷⁶SeU discourages the ⁷⁶SeU/G pair formation. While the ⁷⁶SeU/C mis-pair is slightly less stable than the native U/C mis-pair, the ⁷⁶SeU/U mis-pair is more stable than the native U/U mis-pair. The higher stability may be attributed to the higher acidity of the imino group (3-NH) of ⁷⁶SeU [*pK*_a = 7.29 ± 0.02, Figure 3, compared to that of the native uridine (*pK*_a = 9.18 ± 0.02) (34)], which may promote U/U interaction via hydrogen bond. In addition, considering a selenium atom is 0.43 Å larger in atomic radius than an oxygen atom, the 2-Se atom may strengthen the stacking interaction between ⁷⁶SeU and its 3'-nucleobase (Supplementary Figure S24). When directly comparing the Watson–Crick base pairs (U/A and ⁷⁶SeU/A) with their own corresponding mis-pairs, it is clear that ⁷⁶SeU/A pair has the balanced discrimination against all mis-pairs,

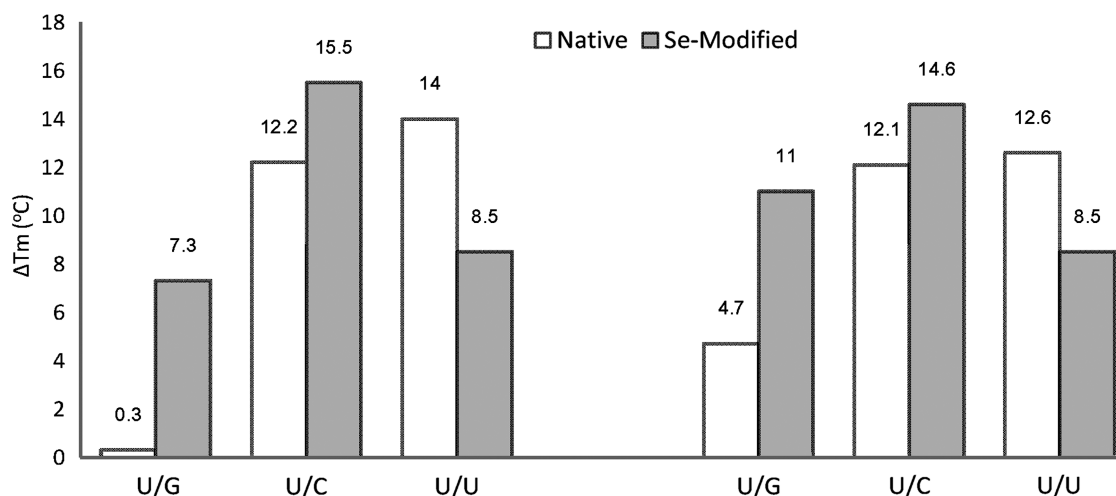


Figure 5. Differences^a of melting temperatures (T_m) of the native and Se-modified U/A pairs and their corresponding mis-pairs. Native (white bar) refers to the T_m difference between the native U/A pair and the other mis-pairs (U/G, U/C and U/U); Se-Modified (gray bar) refers to the T_m difference between the ^{Se}U/A pair and the other modified mis-pairs (^{Se}U/G, ^{Se}U/C and ^{Se}U/U).

with the T_m differences of ^{Se}U/G (7.3°C for one duplex or 11°C for the other in Table 2, and Figure 5), ^{Se}U/C (15.5°C or 14.6°C), and ^{Se}U/U (8.5°C for both sequences). While maintaining fine discrimination against U/C pair (T_m difference: 12.2 or 12.1°C) and U/U pair (T_m difference: 14 or 12.6°C), the native U/A pair poorly discriminates against U/G wobble pair (the T_m differences: 0.3 or 4.7°C in Figure 5 and Table 2). Therefore, in general, ^{Se}U/A has higher base pair fidelity than the native U/A pair.

Since the T_m differences between the native U/A pair and U/G wobble pair were relatively small (0.3 and 4.7°C in Figure 5). The small T_m differences indicate possible changes between U/A and U/G pairs without a significant decrease in duplex stability. This is consistent with the ubiquitous presence of U/G wobble pair in RNAs, which diversifies the structure and function of RNAs, especially non-coding RNAs. Such small thermostability difference between native U/A pair and U/G wobble pair has been previously observed in the literature (4,5). On the contrary, the T_m differences between the ^{Se}U/A and ^{Se}U/G pairs were significant, such as 7.3°C (versus 0.3°C in the native) and 11°C (versus 4.7°C in the native) in Figure 5. The single selenium atom replacement directly decreases the thermal stability of the U/G wobble pair by 4.0 and 3.9°C in RNA duplexes (Table 2). This experimental observation reveals that the U/G wobble pair can be greatly discriminated by incorporating a selenium atom to the 2-position of uridine. The strong discrimination against U/G pair is mainly attributed to the selenium disruption of the hydrogen bond formed by the 2-oxygen (Figure 1) and to the steric effect of the bulky selenium atom at the 2-position. Our results indicate that the 2-Se-modification on uridine significantly increases the high specificity of the U/A base pair.

Crystallization and data collection of Se-RNA

Consistently, our crystal structure study of the ^{Se}U-RNA [5'-rGUAUA-^{Se}U-AC-3'] supports the biophysical results

of ^{Se}U/A pairing. Similar to the native, the Se-RNA crystal is also in rhombohedral space group R32. The Se-RNA structure, determined at 2.3 Å resolution, is virtually identical to the native one (35) (at 2.2 Å resolution, Figure 6). Interestingly, the Se-RNA crystals grew much faster than the native ones. In six days, the Se-RNA formed diffraction-quality crystals in decent sizes (approximately 0.05 × 0.05 mm), while the corresponding native did not crystallize in 3–4 weeks under the same conditions. Moreover, the Se-RNA crystals could form in broader buffer conditions (12 out of 24 conditions in Hampton buffers) than the corresponding native (2 out of 24 conditions). This observation of faster crystal growth of the Se-RNA is consistent with the Se-facilitated duplex stability. As shown in Figure 6A, there are seven self-complementary RNA molecules in a unit cell, and the overall shape of the duplexes is almost linear (~8° inclination to the screw axes). Although this assembling pattern results in the discontinued backbones and grooves, the duplexes stack on top of each other in a head-to-tail fashion, and a pseudo-fiber is formed. The data collection and structure refinement statistics are summarized in Table 3.

Since 2-exo-oxygen of uridine is not involved in the hydrogen bond interactions of U/A pairing, it's expected that the U/A pair will accommodate the larger selenium atom at this position (Figures 1 and 6C). The Se-modification also leads to the acidity increase of the 3-imino group (NH) in the 2-Se-uridine, which strengthens the hydrogen bond between N3 of U6 and N1 of A11. Indeed, after the selenium modification, the U/A hydrogen bond length between N3 of U6 and N1 of A11 is shortened from the native distance (3 Å) to the Se-modified distance (2.81 Å). Moreover, after the Se-modification (Figure 6D), the U/A hydrogen bond length between O4 of U6 and N6 of A11 decreases by 0.47 Å from the native distance (3.39 Å) to the Se-modified distance (2.92 Å). The shortened H-bond lengths indicate stronger H-bonds, which may explain

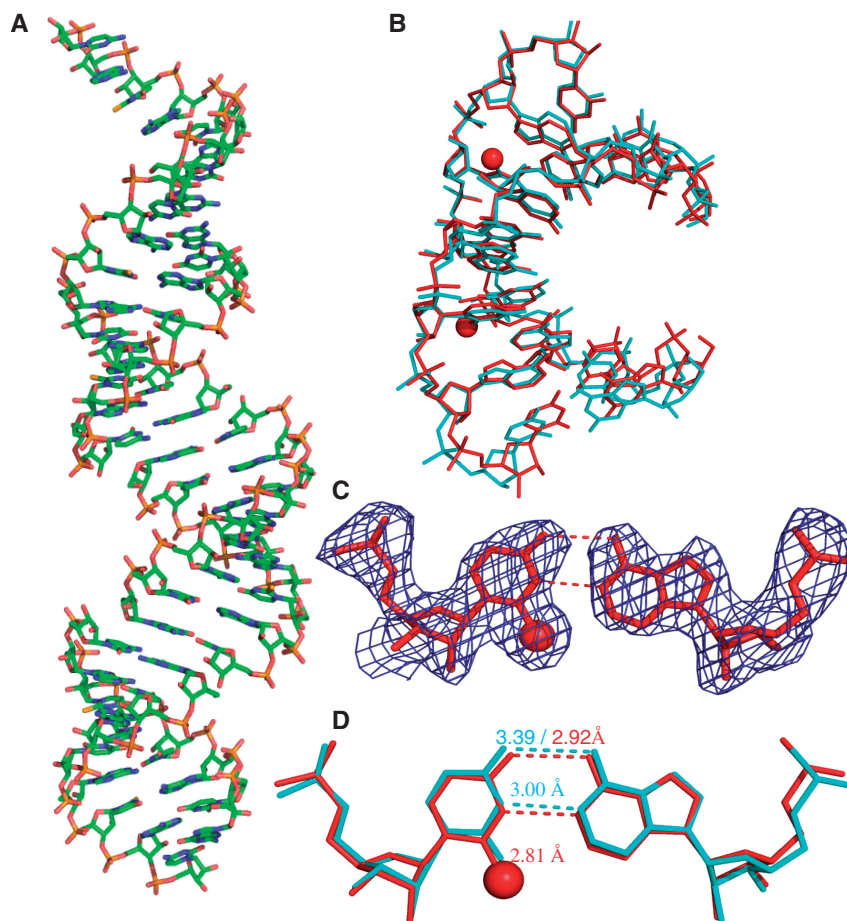


Figure 6. Global and local structures of the ^{76}SeU -containing RNA $r[5'\text{-GUAUA}({}^{76}\text{SeU})\text{AC-3}']_2$ with a resolution of 2.3 Å. (A) The overall structure of duplex. (B) The superimpose comparison of one ^{76}SeU -RNA duplex (red; PDB ID: 3S49) with its native counterpart $[5'\text{-r(GUAUAUA)-dC-3}']_2$ (cyan; PDB ID: 246D) with a RMSD value 0.55. The two red balls represent the selenium atoms. (C) The experimental electron density of $^{76}\text{SeU6/A11}$ base pair with $\sigma = 1.0$. (D) The superimpose comparison of the local base pair $^{76}\text{SeU6/A11}$ (red) and the native U6/A11 (cyan). The numbers indicate the distance between the corresponding atoms.

the increase of duplex stability after the Se-modification. On the contrary, the distance between Se2 of U6 and C2 of A11 in the Se-modified duplex is slightly increased. This distance increase is likely due to a steric effect. This steric clash at the position 2 of the Se-uridine can be a driving force to increase $^{76}\text{SeU/A}$ pair specificity. Consistent with our biophysical study, our structure study has indicated that the selenium bulkiness at the uridine 2-position discourages the U/G wobble pairing. Moreover, due to the electronic effect of a selenium atom, the inability of a Se atom to form a stable hydrogen bond is another main factor responsible for the discrimination against U/G wobble pair.

CONCLUSIONS

In summary, we have first synthesized the ^{76}SeU -RNAs as well as the ^{76}SeU -phosphoramidite. Our biophysical and structural studies on the ^{76}SeU -RNAs indicate that the native and Se-modified structures are virtually identical. Moreover, the 2-Se-modification can largely discriminate against the U/G wobble pair without significant impact on

U/A pair, thereby providing a unique chemical strategy to further enhance base pair fidelity. Furthermore, the 2-Se-modification will provide a useful tool in X-ray crystal structure studies of RNAs and their protein complexes. The atom-specific mutagenesis with selenium opens a new research avenue for investigating base-pair recognition, fidelity and RNA modification. This novel base pair ($^{76}\text{SeU/A}$) with higher specificity likely enables better preservation of genetic information at the RNA level.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–24 and Supplementary Methods.

ACKNOWLEDGEMENTS

We thank Dr Alex Soares, the mail-in program and PXrr in Brookhaven National Laboratory (BNL). The National Synchrotron Light Source (NSLS) at BNL is funded by NIH's National Center for Research Resources and the

Department of Energy's (DOE's) Office of Biological and Environmental Research.

FUNDING

USA National Foundation (NSF CHE-0750235 and MCB-0824837), USA National Institute of Health (NIH GM095086) and Georgia Cancer Coalition Distinguished Cancer Clinicians and Scientists award. Funding for open access charge: National Science Foundation (CHE-0750235).

Conflict of interest statement. None declared.

REFERENCES

- Watson, J.D. and Crick, F.H.C. (1953) Genetical implications of the structure of deoxyribonucleic acid. *Nature*, **171**, 964–967.
- Watson, J.D. (1963) The involvement of RNA in the synthesis of proteins. *Science*, **140**, 17–26.
- Juhling, F., Morl, M., Hartmann, R.K., Sprinzl, M., Stadler, P.F. and Putz, J. (2009) tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.*, **37**, D159–D162.
- Freier, S.M., Kierzek, R., Caruthers, M.H., Neilson, T. and Turner, D.H. (1986) Free energy contributions of G.U and other terminal mismatches to helix stability. *Biochemistry*, **25**, 3209–3213.
- Testa, S.M., Disney, M.D., Turner, D.H. and Kierzek, R. (1999) Thermodynamics of RNA-RNA duplexes with 2- or 4-thiouridines: implications for antisense design and targeting a group I intron. *Biochemistry*, **38**, 16655–16662.
- Varani, G. and McClain, W.H. (2000) The GU wobble base pair - a fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Reports*, **1**, 18–23.
- Bouadloun, F., Donner, D. and Kurland, C.G. (1983) Codon-specific missense errors in vivo. *EMBO J.*, **2**, 1351–1356.
- Stansfield, I., Jones, K.M., Herbert, P., Lewendon, A., Shaw, W.V. and Tuite, M.F. (1998) Missense translation errors in *Saccharomyces cerevisiae*. *J. Mol. Biol.*, **282**, 13–24.
- Toth, M.J., Murgola, E.J. and Schimmel, P. (1988) Evidence for a unique first position codon-anticodon mismatch in vivo. *J. Mol. Biol.*, **201**, 451–454.
- Seetharam, R., Heeren, R.A., Wong, E.Y., Braford, S.R., Klein, B.K., Aykent, S., Kotts, C.E., Mathis, K.J., Bishop, B.F., Jennings, M.J. et al. (1988) Mistranslation in IGF-1 during over-expression of the protein in *Escherichia coli* using a synthetic gene containing low frequency codons. *Biochem. Biophys. Res. Commun.*, **155**, 518–523.
- Ikemura, T. (1982) Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes. *J. Mol. Biol.*, **158**, 573–597.
- Bernardi, G. (1985) Codon usage and genome composition. *J. Mol. Evol.*, **22**, 363–365.
- Parker, J. (1989) Errors and alternatives in reading the universal genetic code. *Microbiol. Mol. Biol. Rev.*, **53**, 273–298.
- Crick, F.H.C. (1966) Codon-anticodon pairing: the wobble hypothesis. *J. Mol. Biol.*, **19**, 548–555.
- Wittwer, A.J., Tsai, L., Ching, W.M. and Stadtman, T.C. (1984) Identification and synthesis of a naturally occurring selenonucleoside in bacterial tRNAs: 5-[(methylamino)methyl]-2-selenouridine. *Biochemistry*, **23**, 4650–4655.
- Ching, W.M. and Alznerdeweerd, B. (1985) A selenium-containing nucleoside at the 1st position of the anticodon in seleno-transfer RNA Glu from *Clostridium sticklandii*. *Proc. Natl Acad. Sci. USA*, **82**, 347–350.
- Lim, V.I. (1994) Analysis of action of wobble nucleoside modifications on codon-anticodon pairing within the ribosome. *J. Mol. Biol.*, **240**, 8–19.
- Agris, P.F. (1991) Wobble position modified nucleosides evolved to select transfer RNA codon recognition: a modified-wobble hypothesis. *Biochimie*, **73**, 1345–1349.
- Dunkel, M., Cook, P.D. and Acevedo, O.L. (1993) Synthesis of novel C-2 substituted pyrimidine nucleoside analogs. *J. Heterocyclic Chem.*, **30**, 1421–1430.
- Seio, K., Sasami, T., Tawarada, R. and Sekine, M. (2006) Synthesis of 2-O-methyl-RNAs incorporating a 3-deazaguanine, and UV melting and computational studies on its hybridization properties. *Nucleic Acids Res.*, **34**, 4324.
- Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Method. Enzym.*, **276**, 307–326.
- 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. and Pannu, N.S. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D: Biol. Crystallogr.*, **54**, 905–921.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
- Parkinson, G., Vojtechovsky, J., Clowney, L., Brunger, A. and Berman, H. (1996) New parameters for the refinement of nucleic acid-containing structures. *Acta Crystallogr. D: Biol. Crystallogr.*, **52**, 57–64.
- Brunger, A.T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
- Read, R.J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst. A*, **42**, 140–149.
- Wise, D. and Townsend, L. (1972) Synthesis of the selenopyrimidine nucleosides 2-seleno and 4-selenouridine. *J. Heterocyclic Chem.*, **9**, 1461–1462.
- Shiue, C.Y. and Chu, S.H. (1975) A facile synthesis of 1-beta-D-arabinofuranosyl-2-seleno- and -4-selenouracil and related compounds. *J. Org. Chem.*, **40**, 2971–2974.
- Hassan, A., Sheng, J., Zhang, W. and Huang, Z. (2010) High fidelity of base pairing by 2-selenothymidine in DNA. *J. Am. Chem. Soc.*, **132**, 2120–2121.
- Vorbrüggen, H. and Strehlke, P. (1973) Nucleosidsynthesen, VII. Eine einfache Synthese von 2-Thiopyrimidin nucleosiden. *Chem. Ber.*, **106**, 3039–3061.
- Kumar, R.K. and Davis, D.R. (1997) Synthesis and studies on the effect of 2-thiouridine and 4-thiouridine on sugar conformation and RNA duplex stability. *Nucleic Acids Res.*, **25**, 1272–1280.
- Salon, J., Jiang, J.S., Sheng, J., Gerlits, O.O. and Huang, Z. (2008) Derivatization of DNAs with selenium at 6-position of guanine for function and crystal structure studies. *Nucleic Acids Res.*, **36**, 7009–7018.
- Salon, J., Sheng, J., Jiang, J.S., Chen, G.X., Caton-Williams, J. and Huang, Z. (2007) Oxygen replacement with selenium at the thymidine 4-position for the Se base pairing and crystal structure studies. *J. Am. Chem. Soc.*, **129**, 4862–4863.
- Knobloch, B., Da Costa, C.P., Linert, W. and Sigel, H. (2003) Stability constants of metal ion complexes formed with N3-deprotonated uridine in aqueous solution. *Inorg. Chem. Commun.*, **6**, 90–93.
- Wahl, M., Ban, C., Sekharudu, C., Ramakrishnan, B. and Sundaralingam, M. (1996) Structure of the purine-pyrimidine alternating RNA double helix, r (GUAUAUA) d (C), with a 3'-terminal deoxy residue. *Acta Crystallogr. D: Biol. Crystallogr.*, **52**, 655–667.