Site-specific biotinylation of RNA molecules by transcription using unnatural base pairs

Kei Moriyama[1], Michiko Kimoto[2], Tsuneo Mitsui[1], Shigeyuki Yokoyama[2,3,4] and Ichiro Hirao[1,2,*]

[1]Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan, [2]Protein Research Group, RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan, [3]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan and [4]RIKEN Harima Institute at SPring-8, 1-1-1 Kohto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan

ABSTRACT

Direct site-specific biotinylation of RNA molecules was achieved by specific transcription mediated by unnatural base pairs. Unnatural base pairs between 2-amino-6-(2-thienyl)purine (denoted by s) and 2-oxo(1H)pyridine (denoted by y), or 2-amino-6-(2-thiazolyl)purine (denoted as v) and y specifically function in T7 transcription. Using these unnatural base pairs, the substrate of biotinylated-y (Bio-yTP) was selectively incorporated into RNA, opposite s or v in the DNA templates, by T7 RNA polymerase. This method was applied to the immobilization of an RNA aptamer on sensor chips, and the aptamer accurately recognized its target protein. This direct site-specific biotinylation will provide a tool for RNA-based biotechnologies.

INTRODUCTION

RNA molecules are frequently labeled with biotin for immobilization on avidin supports. In addition, the biotin in RNA molecules functions as a chemiluminescent marker using streptavidin coupled to alkaline phosphatase, to detect the RNA molecules and their interaction with target molecules. Several biotinylation methods for RNA have been reported. For example, biotin-modified substrates, such as biotinylated UTP or CTP, are available for the direct biotinylation of RNA during transcription (1). However, it is difficult to control the labeling position with this simple method, and this often alters the biological function of the labeled RNA. Chemical biotinylation of RNA at the 3′ or 5′end of the molecule with biotin-conjugated hydrazide (2–4) was developed to

circumvent this problem. RNA can also be biotin-labeled at the 5′ end by a one-step transcription procedure, using N6-biotin derivatives of AMP as transcription initiators under the T7 φ2.5 promoter (5) or biotinyl-guanosine 5′-monophosphate under the conventional T7 promoter (6). Another attractive method would be direct site-specific biotinylation at the desired positions by transcription using unnatural base pairs.
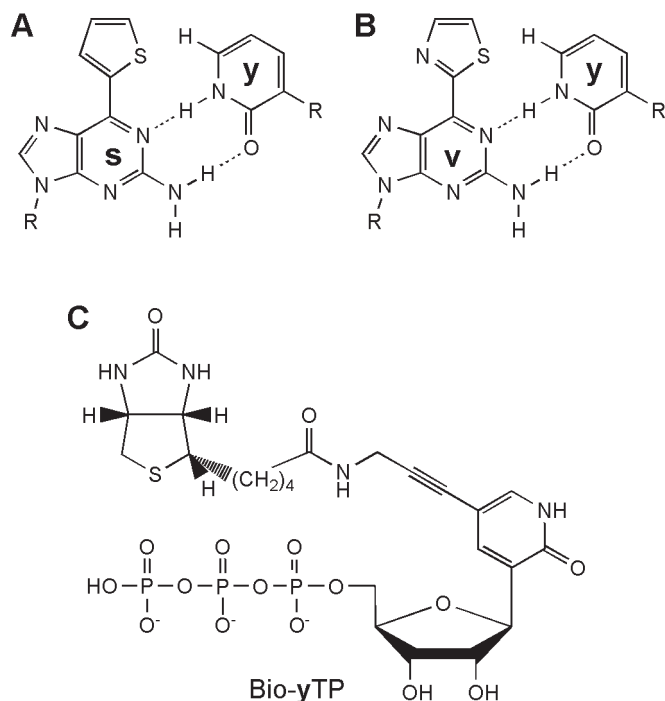
We recently developed unnatural base pairs between 2-amino-6-(2-thienyl)purine (denoted by **s**) and 2-oxo(1H)-pyridine (denoted by **y**) (7,8) or 2-amino-6-(2-thiazolyl)-purine (denoted as **v**) and **y** (Figure 1A and B) (9). These unnatural base pairs function in transcription, and the substrate of **y** (**y**TP) can be site-specifically incorporated into RNA, opposite **s** or **v** in templates, by T7 RNA polymerase. In addition, position 5 of **y** can be modified to introduce various functional groups via the nucleoside of 5-iodo-2-oxo-(1H)pyridine (5-iodo-**y**) (10,11). Therefore, these unnatural base pairs could be used in the direct site-specific biotinylation of RNA molecules. Here, we report the synthesis of the substrate of biotinylated-**y** (Bio-**y**TP, Figure 1C), the incorporation of Bio-**y** into RNA by T7 transcription, and the application of this method to the immobilization of an RNA aptamer.

MATERIALS AND METHODS

General

Reagents and solvents were purchased from standard suppliers and used without further purification. Reactions were monitored by thin-layer chromatography (TLC), using 0.25-mm silica gel 60 plates impregnated with 254 nm fluorescent indicator (Merck). ^{1}H-NMR (270 MHz), ^{13}C-NMR (68 MHz) and ^{31}P-NMR (109 MHz) spectra were recorded on a JEOL EX270 magnetic resonance spectrometer. Purification of

Figure 1. Chemical structures of the unnatural **s**–**y** (**A**) and **v**–**y** (**B**) pairs, and of biotinylated-**y**TP (Bio-**y**TP) (**C**).

nucleosides was performed on a Gilson HPLC system with a preparative C18 column (Waters μ-Bondsphere ϕ 19 × 150 mm). Triphosphates were purified on a DEAE-Sephadex A-25 column (300 × 15 mm), and the final purification was achieved using a Gilson HPLC system with an analytical column (Synchropak RPP, 250 × 4.6 mm, Eichrom Technologies). High resolution mass spectra (HRMS) and electrospray ionization mass spectra (ESI-MS) were recorded on a JEOL HX-110 mass spectrometer and a Waters micromass ZMD 4000 spectrometer equipped with a Waters 2690 LC system, respectively.

3-(Dichloroacetamido)-1-propyne (1)

Dichloroacetylchloride (2.12 ml, 22 mmol, 1.1 eq.) was added to a solution of propargylamine (1.37 ml, 20 mmol) and NaHCO$_3$ (2.02 g, 24 mmol, 1.2 eq.) in 40 ml of dry CH$_2$Cl$_2$, and the solution was stirred at room temperature for 3 h. The reaction mixture was washed once with 5% NaHCO$_3$ (aq) and twice with saturated NaCl (aq). The organic phase was dried over Na$_2$SO$_4$ and evaporated *in vacuo* to give a brown oil of **1** (2.67 g, 16.1 mmol, 81%), which was solid at room temperature. ^{1}H-NMR (270 MHz, CDCl$_3$) δ (p.p.m.) 6.62 (bs, 1H), 5.88 (s, 1H), 4.06 (dd, 2H, J = 2.6, 5.3 Hz), 2.26 (t, 1H, J = 2.6 Hz). ^{13}C-NMR (68 MHz, CDCl$_3$) δ (p.p.m.) 163.54, 77.80, 72.83, 65.92, 30.28. HRMS (FAB, 3-NBA matrix) for C$_5$H$_6$ONCl$_2$ (M + 1): calcd, 165.9826; found, 165.9852.

3-(β-D-Ribofuranosyl)-5-[3-(dichloroacetamido)-1-propynyl]-2-oxo(1H)pyridine (3)

Compound **2** (283 mg, 0.80 mmol) (10) was co-evaporated successively with dry acetonitrile and dry toluene (three times

each). Tetrakis(triphenylphosphine)palladium(0) (46.3 mg, 0.040 mmol, 0.05 eq.) and CuI (24.4 mg, 0.128 mmol, 0.16 eq.) were added and dissolved in 3.6 ml of dry DMF, and then triethylamine (167 μl, 1.20 mmol, 1.5 eq.) was added. To the solution was added a 400 μl aliquot of a 3 M DMF solution of **1** (1.20 mmol, 1.5 eq.). The reaction mixture was stirred at room temperature for 14 h. The product was extracted with H$_2$O/EtOAc, and the organic phase was washed three times with H$_2$O. The combined aqueous phase was evaporated *in vacuo*. After filtration, nucleoside **3** (226 mg, 0.578 mmol, 72% from **2**) was purified by reverse phase HPLC (0–10 min, 5–20% CH$_3$CN in H$_2$O, 10–18 min, 20% CH$_3$CN in H$_2$O. Flow rate, 10 ml/min, detected at 260 nm). ^{1}H-NMR (270 MHz, DMSO-d_6) δ (p.p.m.) 12.02 (bs, 1H), 9.09 (t, 1H, J = 5.1 Hz), 7.57–7.54 (m, 2H), 6.47 (s, 1H), 5.07 (bs, 1H), 4.90 (t, 1H, J = 5.4 Hz), 4.70 (d, 1H, J = 4.8 Hz), 4.64 (d, 1H, J = 4.3 Hz), 4.16 (d, 2H, J = 5.3 Hz), 3.87–3.78 (m, 3H), 3.75–3.45 (m, 2H). ^{13}C-NMR (68 MHz, DMSO-d_6) δ (p.p.m.) 163.06, 160.27, 138.44, 137.87, 131.08, 99.72, 84.84, 83.58, 80.29, 78.99, 74.56, 70.75, 66.45, 61.37, 29.73. HRMS (FAB, 3-NBA matrix) for C$_{15}$H$_{17}$O$_6$N$_2$Cl$_2$ (M+1): calcd., 391.0464; found, 391.0487.

3-(β-D-Ribofuranosyl)-5-[(3-biotinamido-1-propynyl)]-2-oxo(1H)pyridine (5)

Compound **3** (71 mg, 0.18 mmol) was dissolved in 5 ml of H$_2$O and 1 ml of concentrated ammonia solution in a screw-capped vial, and the solution was stirred for 24 h at room temperature to yield the deprotected nucleoside **4**. After evaporation *in vacuo*, the residue was dissolved in 1 ml of dry DMF together with biotin-*N*-hydroxysuccinimide (68.2 mg, 0.20 mmol, 1.1 eq.), and the solution was stirred at room temperature for 2 h. The reaction mixture was then diluted with H$_2$O, and compound **5** (72 mg, 0.14 mmol, 78% from **3**) was purified by reverse phase HPLC (0–15 min, 10–40% CH$_3$CN in H$_2$O, 15–18 min, 50% CH$_3$CN in H$_2$O. Flow rate, 10 ml/min, detected at 260 nm). ^{1}H-NMR (270 MHz, DMSO-d_6) δ (p.p.m.) 11.93 (bs, 1H), 8.27 (t, 1H, J = 5.4 Hz), 7.53 (s, 2H), 6.41 (s, 1H), 6.34 (s, 1H), 5.11 (bs, 1H), 4.92 (t, 1H, J = 5.3 Hz), 4.71 (d, 1H, J = 4.8 Hz), 4.63 (d, 1H, J = 4.6 Hz), 4.31–4.26 (m, 1H), 4.12–4.09 (m, 1H), 4.04 (d, 2H, J = 5.4 Hz), 3.87–3.78 (m, 3H), 3.62–3.47 (m, 2H), 3.11–3.04 (m, 1H), 2.80 (dd, 1H, J = 5.1 Hz, 12.5 Hz), 2.56 (d, 1H, J = 12.9 Hz), 2.09 (t, 2H, J = 7.4 Hz), 1.62–1.23 (m, 6H). ^{13}C-NMR (68 MHz, DMSO-d_6) δ (p.p.m.) 171.56, 162.48, 160.31, 138.57, 137.57, 130.97, 100.15, 86.46, 83.63, 80.34, 78.02, 74.57, 70.80, 61.41, 60.98, 59.13, 55.37, 34.88, 28.52, 28.19, 28.03, 25.15. HRMS (FAB, 3-NBA matrix) for C$_{23}$H$_{31}$O$_7$N$_4$S (M+1): calcd., 507.1913; found, 507.1937.

3-(β-D-Ribofuranosyl)-5-[(3-biotinamido-1-propynyl)]-2-oxo(1H)pyridine 5′-triphosphate (8, Bio-yTP)

Compound **5** (113 mg, 0.223 mmol) was co-evaporated with dry pyridine (three times) and dissolved in 2.23 ml of dry pyridine. To the solution was added 4,4′-dimethoxytritylchloride (79.2 mg, 0.234 mmol, 1.05 eq.), and the solution was stirred at room temperature for 18 h. The product was extracted with CH$_2$Cl$_2$/5% NaHCO$_3$, and the aqueous phase was washed with CH$_2$Cl$_2$. The combined organic phase was

dried over Na_2SO_4, and compound **6** (144 mg, 0.178 mmol, 80% from **5**) was purified by silica gel column chromatography (mobile phase, 0–10% MeOH in CH_2Cl_2). Compound **6** (144 mg, 0.178 mmol) was co-evaporated with dry pyridine (three times) and dissolved in 1.8 ml of dry pyridine. To the solution was added 50 µl of acetic anhydride (0.53 mmol, 3 eq.), and the solution was stirred at room temperature for 12 h. The 2′,3′-di-*O*-acetyl derivative of compound **6** was extracted with CH_2Cl_2/5% $NaHCO_3$ (aq). The organic phase was washed twice with 5% $NaHCO_3$ (aq) and once with saturated NaCl, and then was dried over Na_2SO_4. After the solvent was removed *in vacuo*, the residue was co-evaporated with dry toluene (three times) and dissolved in 890 µl of 80% acetic acid in H_2O. Deprotection of the dimethoxytrityl group was completed by stirring at room temperature for 2 h. The solvent was then removed *in vacuo*, and compound **7** (83 mg, 0.14 mmol, 79% from **6**) was purified by silica gel column chromatography (mobile phase, 0–10% MeOH in CH_2Cl_2). Compound **7** (83 mg, 0.14 mmol) was co-evaporated with dry pyridine (three times) and dissolved in 560 µl of dry pyridine/dry DMF (1:3, v/v). To the solution was added 154 µl of 1 M 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one (12) in dioxane (0.154 mmol, 1.1 eq.). The solution was stirred at room temperature for 10 min. After the addition of 140 µl of tri-*n*-butylamine and 420 µl of 0.5 M bis(tri-*n*-butylammonium)pyrophosphate in DMF (0.21 mmol, 1.5 eq.), the solution was stirred at room temperature for a further 10 min. To the reaction was added 2.8 ml of 1% I_2 in pyridine/H_2O (98:2, v/v), and the solution was stirred at room temperature for 15 min. A 210 µl aliquot of 5% $NaHSO_3$ (aq) was then added and the solvent was removed *in vacuo*. To remove the acetyl groups, the residue was dissolved in 14 ml of H_2O, and after the solution was stirred at room temperature for 30 min, a concentrated ammonia solution (28 ml) was added. After stirring at room temperature for 1 h, the solvent was removed *in vacuo*. Compound **8** was purified successively by anion exchange column chromatography (DEAE-Sephadex A-25, 0.05–1 M triethylammonium bicarbonate linear gradient) and reverse phase HPLC (Synchropak RPP, 0–30% CH_3CN in 100 mM triethylammonium acetate). ^{1}H-NMR (270 MHz, D_2O) δ (p.p.m.) 7.72 (d, 1H, $J = 2.3$ Hz), 7.47 (d, 1H, $J = 2.1$ Hz), 4.79 (d, 1H, $J = 3.8$ Hz), 4.32–4.27 (m, 1H), 4.16–4.12 (m, 1H), 4.06–3.98 (m, 5H), 3.95 (s, 2H), 3.01 (q, 22H, $J = 7.4$ Hz), 2.70 (dd, 1H, $J = 5.0$ Hz, 13.1 Hz), 2.48 (d, 1H, $J = 13.2$ Hz), 2.12 (t, 2H, $J = 6.8$ Hz), 1.54–1.18 (m, 6H), 1.09 (t, 33H, $J = 7.3$ Hz). ^{31}P-NMR (109 MHz, D_2O) δ (p.p.m.) -9.32 (d, 1P, $J = 13.4$ Hz), -10.48 (d, 1P, $J = 18.9$ Hz), -22.41 (t, 1P, $J = 19.5$ Hz). ESI-MS for $C_{23}H_{32}O_{16}N_4P_3S$ ([M-H]$^-$): calcd., 745.07; found, 744.89. UV-vis spectrum (in 10 mM sodium phosphate buffer, pH 7.0), $\lambda_{max} = 258\,nm\,(\varepsilon = 1.8 \times 10^4), 317\,nm\,(\varepsilon = 5.3 \times 10^3)$.

Preparation of templates for T7 transcription

Template DNAs for the anti-(Raf-1) RNA aptamer were amplified from an aptamer-encoding vector by PCR, using a 5′ end primer [39.45] and a 3′ end primer ([24.45] for the unmodified aptamer or [24.45**s**100] for the Bio-**y** aptamer) according to the literature (10). The sequences of the primers were as follows (the T7 promoter region is underlined, and **s** denotes 2-amino-6-(2-thienyl)purine): [39.45], 5′-GG<u>TA-

ATACGACTCACTATAGGG</u>AGTGGAGGAATTCATCG; [24.45], 5′-GCAGAAGCTTGCTGTCGCTAAGGC; [24.45**s**-100], 5′-**s**CAGAAGCTTGCTGTCGCTAAGGC.

T7 transcription for RNA molecules (17mer and 100mer) containing Bio-y

Transcription was carried out in buffer containing 40 mM Tris–HCl (pH 8.0), 24 mM $MgCl_2$, 5 mM DTT, 2 mM spermidine and 0.01% Triton X-100, in the presence of 1 mM of each NTP, 0.125 mM Bio-**y**TP, template DNA (2 µM for RNA 17mers and ~0.13 µM for Bio-**y** aptamer, 100mer), and 2.5U/µl T7 RNA polymerase (Takara), for 6 h at 37°C. To measure the transcription yields of the RNA 17mers, the reactions (20 µl) were performed in the presence of 2 µCi [γ-^{32}P]-GTP, and the products were analyzed on a 20% denaturing polyacrylamide gel containing 7 M urea. As for the Bio-**y** aptamer, the full-length transcripts were purified by elution from an 8% denaturing polyacrylamide gel containing 7 M urea.

Analysis of the insertion position of Bio-y in the 17mer transcript

RNA transcripts (17mer) containing Bio-**y** or unmodified **y** at position 13 were prepared by T7 transcription using temp35**v**-1 as the template, in the presence of 0.125 mM Bio-**y**TP or 1 mM unmodified **y**TP, and were purified by gel electrophoresis. After dephosphorylation by calf intestinal alkaline phosphatase (Takara), the RNAs were labeled with ^{32}P at the 5′ end by [γ-^{32}P]ATP (Perkin Elmer) and T4 polynucleotide kinase (Takara), and were purified by gel electrophoresis. Alkaline digestion was performed in 40 mM sodium carbonate (pH 9.0), 0.8 mM EDTA and 0.2 mg/ml *E.coli* tRNA for 17 min at 90°C. RNase T1 digestion (0.09 U/µl) was performed in 16 mM sodium citrate (pH 5.0), 5.5 M urea, 0.8 mM EDTA and 0.2 mg/ml *E.coli* tRNA for 5 min at 55°C. The digested samples were analyzed on a 20% polyacrylamide gel containing 7 M urea.

Quartz crystal microbalance (QCM) analysis

QCM experiments were performed on an Affinix-Q apparatus (13,14) (Initium, Japan). The clean, bare Au (4.9 mm^2) electrode side of a QCM sensor chip was incubated with an aqueous solution of 3,3′-dithiodipropionic acid (1 mM) at room temperature for 30 min. Streptavidin was immobilized on the sensor chip by amine-coupling chemistry. The surfaces were activated with a mixture of 1-ethyl-3-(3-dimethyl-aminopropyl)carbodiimide hydrochloride (EDC, 0.2 M) and *N*-hydroxysuccinimide (NHS, 0.05 M) for 10 min. A streptavidin solution (100 µl, 200 µg/ml) in 10 mM sodium acetate (pH 5.2) was applied to the surface for 1 h, and the surface was treated with a monoethanolamine hydrochloride/NaOH solution (100 µl, 1 M, pH 8.5) for 10 min to block the remaining activated groups. After the immobilization of streptavidin, the frequency decreased by an average of 610 ± 140 Hz. The chip was then allowed to equilibrate at 22°C, with stirring at 600 r.p.m., in 2 ml phosphate buffer (10.1 mM Na_2HPO_4 and 1.8 mM KH_2PO_4) supplemented with 587 mM NaCl, 2.7 mM KCl, 5 mM $MgCl_2$, 2 mM DTT and 0.01% Tween-20 (Buffer A).

For the Bio-**y** aptamer immobilization, the aptamer (60 pmol; 6.7 µl) was injected (time = 0 min) and the

frequency changes were monitored. After 30 min, the frequency decreased by an average of 400 ± 100 Hz.

Since a frequency decrease (ΔF) of 1 Hz corresponded to a mass increase of 30 pg on the electrode (0.049 cm^2), the concentration (C pmol/cm^2) of immobilized streptavidin (approx. MW = 60 000) or Bio-**y** aptamer (approx. MW = 32 000) was calculated using the following formula: $C = (\Delta F \times 30)/(MW \times 0.049)$.

To detect the interaction between the aptamer and its target protein, the RNA-containing buffer was exchanged for 2 ml of Buffer A, and a BSA solution (25 µl; 20 mg/ml) was added to prevent non-specific interactions between RNA and proteins. After the chip attained equilibrium, a protein solution (20 µl) containing 64 µM GST-RBD of Raf-1 was injected (time = 0 min), and the frequency changes were monitored. For comparison, the frequency changes were adjusted to 0 Hz, 15 s after the protein injection.

Surface plasmon resonance (SPR) analysis

SPR experiments were performed on a BIACORE 3000 apparatus (Biacore AB, Sweden) equipped with the BIAcore 3.2 software. Streptavidin was immobilized on the CM5 sensor chip (Research grade), by amine-coupling chemistry. The immobilization steps were carried out at a flow rate of 5 µl/min in 10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA and 0.005% P20 surfactant. The flow cells were activated for 7 min with a mixture of NHS (0.05 M) and EDC (0.2 M), and the streptavidin solution was then injected during a period of 4 min. The monoethanolamine hydrochloride/NaOH solution was injected during a period of 7 min, to block the remaining activated groups. Streptavidin on each flow cell was 2000 RU on average. The chip was allowed

to equilibrate at 22°C in phosphate buffer (10.1 mM Na$_2$HPO$_4$ and 1.8 mM KH$_2$PO$_4$) supplemented with 137 mM NaCl, 2.7 mM KCl, 5 mM MgCl$_2$, 2 mM DTT and 5% glycerol (Buffer B). For immobilization of the Bio-**y** aptamer, a 30 nM solution of the aptamer in Buffer B was injected (time = 0 min) at a flow rate of 10 µl/min for 15 min. About 300 RU of the aptamer were immobilized on the flow cell.

After the chip attained equilibrium in Buffer B, the protein solutions (640 nM) were injected (time = 0 min) at a flow rate of 20 µl/min for 2 min. One non-coated or one streptavidin-coated flow cell was used to monitor the non-specific binding of proteins. The RU changes were corrected by using the signals from the non-coated flow cells. The sensor chip surface was regenerated with two 10-µl pulses of 6 M urea, followed by four 10-µl pulses of distilled water, and equilibration in Buffer B.

RESULTS AND DISCUSSION

The specificity and efficiency of Bio-**y** incorporation into RNA were tested by T7 transcription of short DNA templates (35mer). Bio-**y**TP, in which a biotin residue is linked to position 5 of **y** through a propynylamine linker, was synthesized from the ribonucleoside of 5-iodo-**y** by a six-step reaction (Figure 2). Transcription of the templates containing **v** or **s** in the presence of 0.125 mM Bio-**y**TP (Figure 3A) yielded full-length (17mer) transcripts with the site-specific incorporation of Bio-**y**, as indicated by their mobility shift on the gel (Figure 3B). The transcripts into which Bio-**y** was incorporated (lanes 1 and 3) migrated more slowly than the 17mer transcript consisting of only the natural bases (lane 6). Although transcription of the **v**- or **s**-containing templates
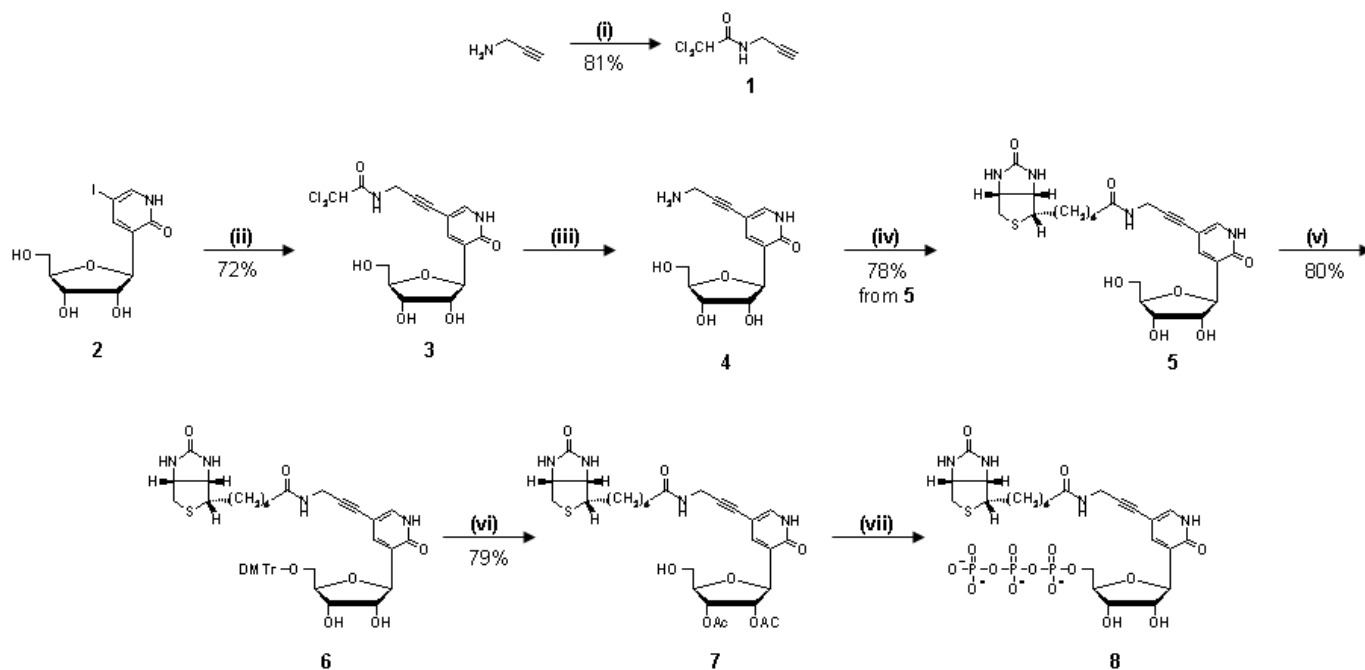
Figure 2. Chemical synthesis of Bio-**y**TP. (i) Cl$_2$CHCOCl, NaHCO$_3$, CH$_2$Cl$_2$, r.t., 3 h, 81%; (ii) **1**, Pd(PPh$_3$)$_4$, CuI, Et$_3$N, DMF, r.t., 14 h, 72%; (iii) NH$_4$OH, r.t., 24 h; (iv) biotin-*N*-hydroxysuccinimide, DMF, r.t., 2 h, 78% from **3**; (v) DMTr-Cl, pyridine, r.t., 18 h, 80%; (vi) (1) (CH$_3$CO)$_2$O, pyridine, r.t., 12 h and (2) 80% CH$_3$COOH, r.t., 2 h, 79%; and (vii) (1) 2-Chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one, pyridine, DMF, dioxane, r.t., 10 min; (2) tri-*n*-butylamine, bis(tri-*n*-butylammonium)pyrophosphate, DMF, r.t., 10 min; (3) NH$_4$OH, r.t., 1 h. Abbreviations: DMTr, 4,4'-dimethoxytrityl; Ac, acetyl.

Figure 3. Specific T7 transcription using the unnatural base pairs. (**A**) Experimental scheme. (**B**) Gel electrophoresis of transcripts obtained from the templates (N = **v**, **s** or A) in the presence or absence of Bio-**y**TP. The relative yields of the full-length transcripts (17mer) were derived by comparison with the yield of transcripts obtained from transcription of the natural template (N = A; lane 6), and each yield was averaged from 4 to 5 datasets.

Figure 4. Sequence analysis of the 17mer transcripts containing Bio-**y** or **y** at position 13. The 5′-labeled transcripts (N′ = Bio-**y** or **y**) were partially digested with alkali (OH⁻) or with RNase T1 (RNase T1), and the digested fragments and the intact transcripts (Intact) were analyzed on a 20% polyacrylamide gel containing 7 M urea.

in the absence of Bio-**y**TP also yielded 17mer transcripts, due to the misincorporation of natural NTPs opposite **v** or **s** in the template (lanes 2 and 4), the bands corresponding to the misincorporated 17mer were barely detectable in lanes 1 and 3, and thus, this misincorporation was prevented by the addition of Bio-**y**TP. In addition, shifted bands corresponding to the Bio-**y**-containing transcripts were not observed upon transcription of the natural template in the presence of Bio-**y**TP (lane 5), indicating that Bio-**y**TP was not incorporated into RNA opposite the natural bases in the templates.

Furthermore, to confirm the insertion position of Bio-**y** in the transcript, we analyzed the sequences of the 17mer transcripts containing Bio-**y** or **y** at position 13. The 5′-labeled transcripts containing Bio-**y**TP or **y**TP were partially digested by a treatment with alkali or RNase T1. On the sequencing ladder of the transcript containing Bio-**y** (Figure 4), the bands corresponding to the 13mer to 17mer fragments were shifted, as compared with the bands resulting from the digestion of the transcript containing **y**. These mobility shifts were caused by the incorporation of Bio-**y** into the fragments, and clearly indicate that the Bio-**y** was site-specifically incorporated into position 13 of the transcript, opposite **v** in the template.

The efficiency of transcription mediated by the **v**–**y** pair (Figure 3, lane 1) was slightly higher than that achieved by the **s**–**y** pair (lane 3), and the relative yield (47%) of the Bio-**y**-containing transcript from the **v**-containing template was higher than that (38%) from the **s**-containing template. In contrast, the relative yield (70%) of the truncated product

(12mer) obtained by pausing before **v** in the template (lane 1) was slightly lower than that (86%) generated by pausing before **s** in the template (lane 3). In addition, this relative yield (70%) of the 12mer also decreased, in comparison with the truncated 12mer (103%) obtained from the transcription in the absence of Bio-**y**TP, using the **v**-containing template (lane 2). Thus, decreasing the amounts of truncated products increases the yield of full-length transcripts containing Bio-**y** in the unnatural base pair transcription system.

This site-specific Bio-**y** incorporation can be utilized with RNA molecules with different sequence contexts, although the transcription efficiency depends on the neighboring-base composition of **v** in the templates. Transcription reactions using the templates containing an A**v**A or T**v**A sequence generated higher relative yields (59–84%) of the 17mer transcripts (Figure 5B, lanes 3 and 7). However, the transcription using the template containing a G**v**C sequence produced a slightly lower relative yield (36%) of the transcript (Figure 5B, lane 5). Furthermore, transcripts containing two Bio-**y** components could also be synthesized from the templates containing two **v** bases (Figure 5C, lanes 3 and 5). Although the relative yield (24%) of the transcript containing two adjacent Bio-**y** components was relatively low, an adequate amount of the transcript could be obtained for subsequent biological experiments.

We used this procedure to transcribe a biotinylated anti-(Raf-1) RNA aptamer (15) for immobilization on an avidin support; the immobilized aptamer retained its ability to bind human Raf-1 specifically. The RNA aptamer (100mer), in which Bio-**y** was incorporated at the 3′-terminus, was derived from its **s**-containing DNA template prepared by PCR of the original, unmodified DNA template using a 3′-primer containing **s** (Figure 6). The biotinylated RNA aptamer was efficiently immobilized on both sensor chips (Figure 7A and C). The unmodified RNA aptamer, in which Bio-**y** was not incorporated, did not bind to either sensor chip. The frequency of QCM decreased upon binding of the modified RNA aptamer containing Bio-**y** (Bio-**y** aptamer) to the streptavidin on the QCM plate; 8 ± 2 pmol/cm^2 of Bio-**y** aptamer were immobilized on the sensor chip containing 6 ± 1 pmol/cm^2 of streptavidin. For the Bio-**y** aptamer, 300 ± 50 RU were immobilized on the SPR sensor chip. The interaction of the immobilized aptamer with its target protein, the Ras binding domain (RBD) of Raf-1 fused with glutathione *S*-transferase (GST-RBD of Raf-1; 640 nM), was detected by the sensors (Figure 7B and D). The protein bound on the SPR sensor chip could be removed by a 6 M urea treatment, to regenerate the sensor chip surface (data not shown). The binding of another RBD, GST-RBD of RGL (640 nM), to the sensor chip-bound aptamer was not detected [Figure 7D, (b)], indicating that the immobilized RNA aptamer retains its specificity for Raf-1.

This study demonstrates the site-specific biotinylation of RNA molecules using a transcription system that included unnatural base pairs. In this system, the unnatural base pair between **v** and Bio-**y** showed relatively high efficiency, as well as high selectivity, in transcription. Thus, the **v**–**y** pair is capable of efficiently incorporating **y**, modified with a large functional group, into RNA. DNA templates containing **v** can be easily prepared by direct DNA chemical synthesis using the amidite of **v**, or by PCR with a **v**-containing 3′-primer and the original templates. This is very useful, especially for RNA
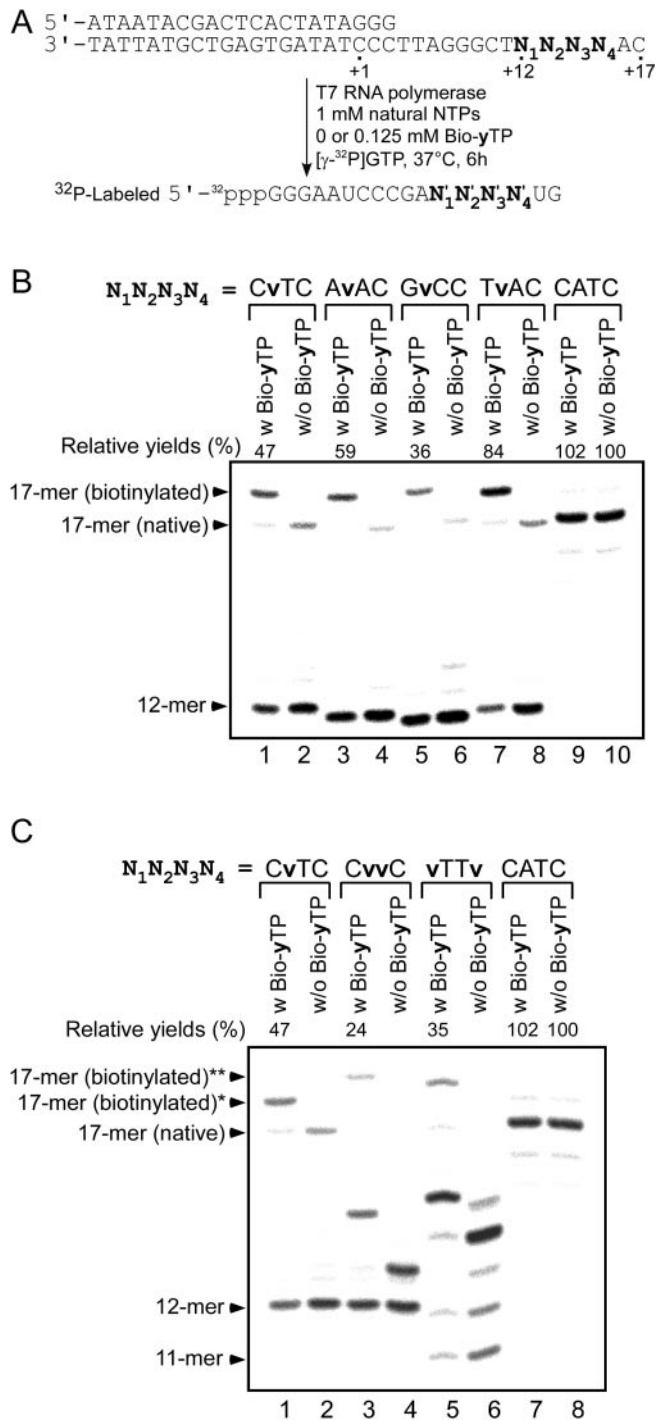
Figure 5. Specific T7 transcription using templates with different sequence contexts containing one **v** or two **v** bases. (**A**) Experimental scheme. (**B**) Gel electrophoresis of transcripts obtained from the templates containing **v** within different sequence contexts. (**C**) Gel electrophoresis of transcripts obtained from templates containing one **v** or two **v** bases. The relative yields of the full-length transcripts (17mer) were derived by comparison with the yield of transcripts obtained from transcription of the natural template [lane 10 in (B), lane 8 in (C)], and each yield was averaged from 3 to 5 datasets.

aptamers, because RNA aptamers contain 5′- and 3′-constant regions. In addition, since two Bio-**y** components can be incorporated into one RNA molecule, the immobilization efficiency and the detection sensitivity of the RNA molecule can be
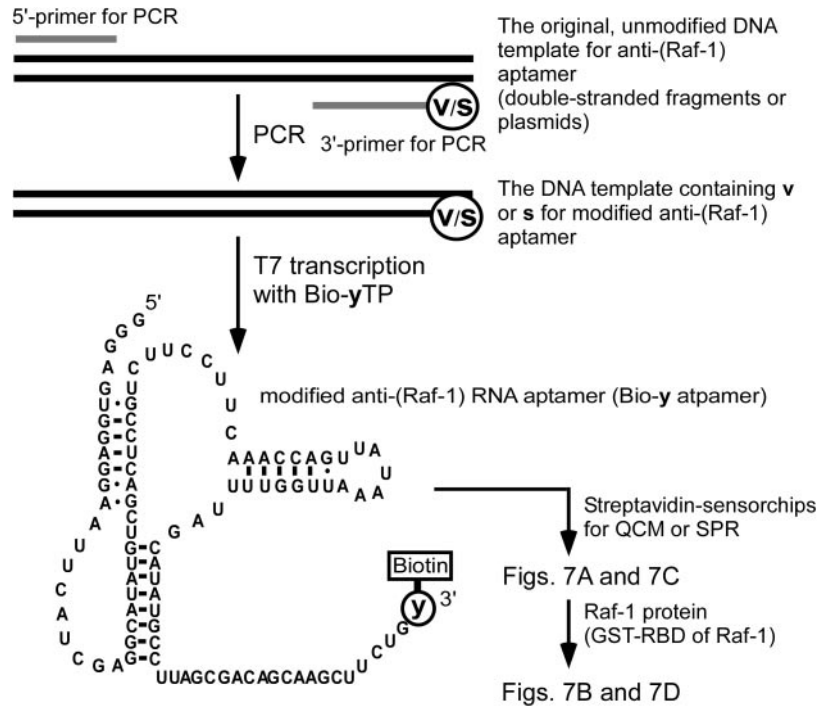
Figure 6. Scheme for Bio-**y** incorporation into the 3′-terminus of the anti-(Raf-1) RNA aptamer.
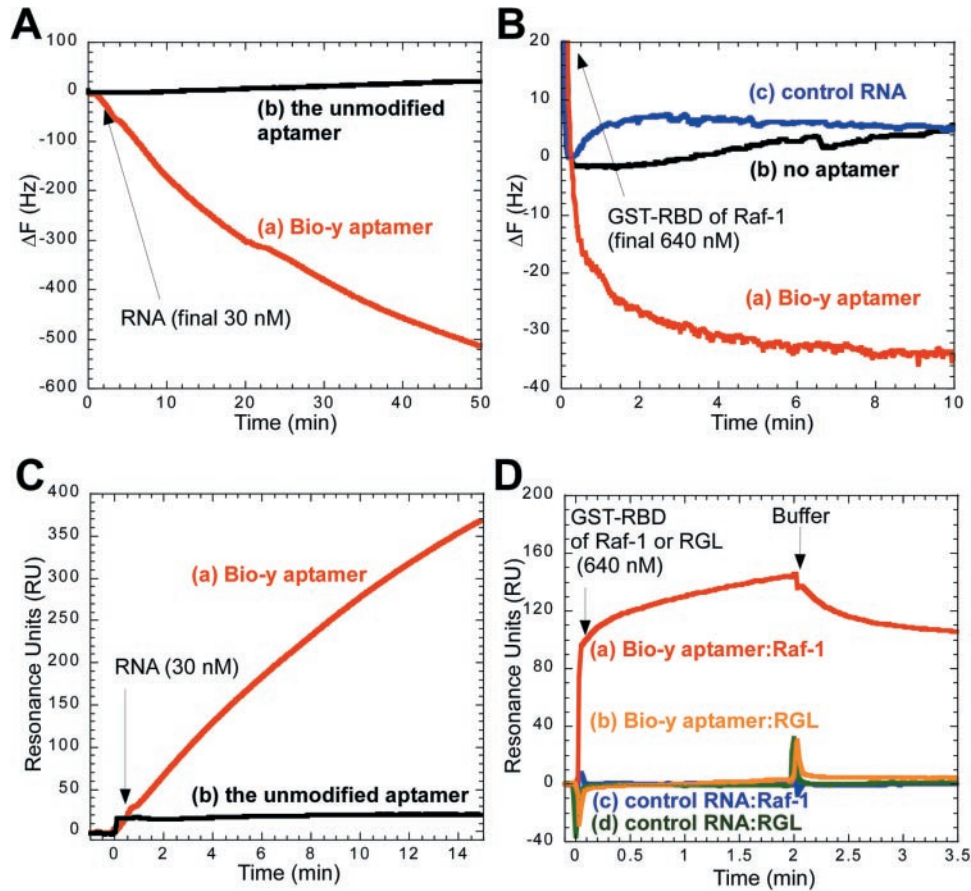

Figure 7. Typical responses of the streptavidin-coated sensor chip to the immobilization of the biotinylated anti-(Raf-1) aptamer (**A** and **C**) and its interaction with its target Raf-1 (**B** and **D**), detected by 27 MHz QCM (A and B) or SPR (C and D). To check for the non-specific binding of proteins, a control RNA (100mer) with negligible affinity for Raf-1 was immobilized on streptavidin-coated sensor chips [(c) in panel B; (c) and (d) in panel D].

increased (16). Thus, this is a practical method for the immobilization and detection of RNA molecules.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

Conflict of interest statement. None declared.

REFERENCES

1. Langer,P.R., Waldrop,A.A. and Ward,D.C. (1981) Enzymatic synthesis of biotin-labeled polynucleotides: novel nucleic acid affinity probes. *Proc. Natl Acad. Sci. USA*, **78**, 6633–6637.
2. Bayer,E.A., Ben-Hur,H. and Wilchek,M. (1988) Biocytin hydrazide—a selective label for sialic acids, galactose, and other sugars in glycoconjugates using avidin-biotin technology. *Anal. Biochem.*, **170**, 271–281.
3. Wu,T., Ruan,K. and Liu,W. (1996) A fluorescence-labeling method for sequencing small RNA on polyacrylamide gel. *Nucleic Acids Res.*, **24**, 3472–3473.
4. Seki,M., Carninci,P., Nishiyama,Y., Hayashizaki,Y. and Shinozaki,K. (1998) High-efficiency cloning of Arabidopsis full-length cDNA by biotinylated CAP trapper. *Plant J.*, **15**, 707–720.
5. Huang,F., Wang,G., Coleman,T. and Li,N. (2003) Synthesis of adenosine derivatives as transcription initiators and preparation of 5′ fluorescein- and biotin-labeled RNA through one-step *in vitro* transcription. *RNA*, **9**, 1562–1570.
6. Collett,J.R., Cho,E.J., Lee,J.F., Levy,M., Hood,A.J., Wan,C. and Ellington,A.D. (2005) Functional RNA microarrays for high-throughput screening of antiprotein aptamers. *Anal. Biochem.*, **338**, 113–123.
7. Fujiwara,T., Kimoto,M., Sugiyama,H., Hirao,I. and Yokoyama,S. (2001) Synthesis of 6-(2-thienyl)purine nucleoside derivatives that form unnatural base pairs with pyridine-2-one nucleosides. *Bioorg. Med. Chem. Lett.*, **11**, 2221–2223.
8. Hirao,I., Ohtsuki,T., Fujiwara,T., Mitsui,T., Yokogawa,T., Okuni,T., Nakayama,H., Takio,K., Yabuki,T., Kigawa,T., Kodama,K., Yokogawa,T., Nishikawa,K. and Yokoyama,S. (2002) An unnatural base pair for incorporating amino acid analogs into proteins. *Nat. Biotechnol.*, **20**, 177–182.
9. Mitsui,T., Kimoto,M., Harada,Y., Yokoyama,S. and Hirao,I. (2005) An efficient unnatural base pair for a base-pair-expanded transcription system. *J. Am. Chem. Soc.*, **127**, 8652–8658.
10. Kimoto,M., Endo,M., Mitsui,T., Okuni,T., Hirao,I. and Yokoyama,S. (2004) Site-specific incorporation of a photo-crosslinking component into RNA by T7 transcription mediated by unnatural base pairs. *Chem. Biol.*, **11**, 47–55.
11. Endo,M., Mitsui,T., Okuni,T., Kimoto,M., Hirao,I. and Yokoyama,S. (2004) Unnatural base pairs mediate the site-specific incorporation of unnatural hydrophobic component into RNA transcripts. *Bioorg. Med. Chem. Lett.*, **14**, 2593–2596.
12. Ludwig,J. and Eckstein,F. (1989) Rapid and efficient synthesis of nucleoside 5′-*O*-(1-thiotriphosphates), 5′-triphosphates and 2′,3′-cyclophosphorothioates using 2-chloro-4H-1,3,2-benzodioxaphosphorin-4-one. *J. Org. Chem.*, **54**, 631–635.
13. Okahata,Y., Niikura,K., Sugiura,Y., Sawada,M. and Morii,T. (1998) Kinetic studies of sequence-specific binding of GCN4-bZIP peptides to DNA strands immobilized on a 27-MHz quartz-crystal microbalance. *Biochemistry*, **37**, 5666–5672.
14. Matsuno,H., Furusawa,H. and Okahata,Y. (2002) Direct monitoring of DNA cleavages catalyzed by an ATP-dependent deoxyribonuclease on a 27 MHz quartz-crystal microbalance. *Chem. Commun.*, 470–471.
15. Kimoto,M., Shirouzu,M., Mizutani,S., Koide,H., Kaziro,Y., Hirao,I. and Yokoyama,S. (2002) Anti-(Raf-1) RNA aptamers that inhibit Ras-induced Raf-1 activation. *Eur. J. Biochem.*, **269**, 697–704.
16. Rodgers,J.T., Patel,P., Hennes,J.L., Bolognia,S.L. and Mascotti,D.P. (2000) Use of biotin-labeled nucleic acids for protein purification and agarose-based chemiluminescent electromobility shift assays. *Anal. Biochem.*, **277**, 254–259.