CHEMISTRY A European Journal

Supporting Information

CCR5 RNA Pseudoknots: Residue and Site-Specific Labeling correlate Internal Motions with microRNA Binding

[Bin Chen](http://orcid.org/0000-0001-7501-4229),^[a, b] Andrew P. Longhini,^[b] Felix Nußbaumer,^[c] Christoph Kreutz,^[c] [Jonathan D. Dinman,](http://orcid.org/0000-0002-2402-9698)*^[a] and T. Kwaku Dayie^{*[b]}

chem_201705948_sm_miscellaneous_information.pdf

Supplementary Materials

Residue/site-specific labeling of *CCR***5 RNA pseudoknot reveals internal motions that correlate with microRNA binding**

Bin Chen^[a,b], Andrew P. Longhini^[b], Felix Nußbaumer^[c], Christopher Kreutz^[c], Jonathan Dinman^{[a]*}, & T. Kwaku Dayie^{[b]*}

Contents

1 EXPERIMENTAL SECTION

1.1 Chemicals and materials

Chemical reagents and solvents were purchased from commercial vendors and used without further purification. The *Taq* DNA polymerase, T4 Polynucleotide Kinase, and T4 DNA ligase were purchased from New England Biolabs, Inc., and the RNA 5′ polyphosphatase from Epicentre. The T7 RNA polymerase used for *in vitro* transcription was prepared in house.^[1]

1.2 Plasmid and oligonucleotides

The plasmid containing *CCR*5 mRNA -1 PRF pseudoknot element was constructed. [2] Template DNA for *in vitro* transcription was generated by polymerase chain reactions (PCR) using primers (5′-CTA ATA CGA CTC ACT ATA GCC AGG ACG GTC ACC TTT GGG GTG-3′; 5′-TTG AGA TCT GGT AAA GAT GAT TCC TGG GAG AG-3′) directed against the 5′ and 3′ ends of 96nt *CCR*5 pseudoknot RNA. The 5′ primer and another 3′ primer (5′-TGG GAG AGA CGC AAA CAC AGC CAC CAC-3′) were used to amplify template DNA for *in vitro* transcription of 72nt-*CCR*5 acceptor fragment. A pair of complimentary DNA oligos (5′-mUmUG AGA TCT GGT AAA GAT GAT TCC TAT AGT GAG TCG TAT TAG-3′ and 5′-CTA ATA CGA CTC ACT ATA GGA ATC ATC TTT ACC AGA TCT CAA-3′) was used as a template for *in vitro* transcription of 24nt-*CCR*5 donor fragment. All of the primers and a DNA oligonucleotide used as a splint for RNA ligation (5′-GAT CTG GTA AAG ATG ATT CCT GGG AGA GAC GCA AAC ACA G-3′) were purchased from Integrated DNA Technologies, Inc. An RNA oligonucleotide (miR-1224: 5′-GUG AGG ACU CGG GAG GUG G-3′) was either purchased from Integrated DNA Technologies, Inc., or prepared in house by solid synthesis (**see below**).

1.3 Chemo-enzymatic synthesis of (8, 1′**- 13C) ATP and subsequent enzymatic dephosphorylation to (8, 1**′**- 13C)-adenosine** were described [3]

1.4 Synthesis of N⁶ -Benzoyl-(8, 1′**- 13C)-adenosine 3**′**-CEP 5**′**-DMT 2**′**-TBDMS protected building block (5)**

N6 -Benzoyl-(8, 1′**- 13C)-adenosine (2)**

(8- 13C)-adenosine **1** (1.0 eq, 1.00 g, 3.74 mmol) was coevaporated with anhydrous pyridine, dissolved in 15 mL of anhydrous pyridine and sealed with septum. After stirring for 5 minutes under argon, trimethylsilyl chloride (TMS-Cl, 5.0 eq, 2.03 g, 18.7 mmol) was added via syringe. Benzoyl chloride (1.2 eq, 631 mg, 4.49 mmol) was added dropwise via syringe over a period of 20 minutes after 30 minutes and stirring was continued for 2.5 h, while a clear yellow solution was formed. Then, 2 mL H_2O was added at once and after 5 minutes 4 mL of aqueous ammonia solution (28-30 %) was added at once, stirring was continued for 15 minutes. The mixture was evaporated to dryness and the oily residue coevaporated twice with toluene to give a yellow solid. This was dissolved in 100 mL of a mixture of CH_2Cl_2 / MeOH (1:1), insoluble solids were filterd off and the filtrate was evaporated to dryness to give a yellow solid. The crude product was applied to a silica gel column with methylene chloride and eluted using a gradient of 5 to 20 % MeOH in methylene chloride.

Yield: 0.40 g (39 %)

¹H-NMR (300 MHz, DMSO-d₆, 25°C):

δ 11.20 (bs, 1H, N*H*); 8.76 (s, 1H, C2*H*); 8.72 (d, 1H, ¹J_{CH} = 214.6 Hz, ¹³C8*H*); 8.06 - 7.53 (m, 5H, C*H*ar); 6.05 (d, 1H, ¹J_{CH} = 165.8 Hz, C1'*H*); 5.55 (bs, 1H, C2'O*H*); 5.23 (bs, 1H, C3'O*H*); 5.12 (t, 1H, C5′O*H*); 4.66 (bs, 1H, C2′*H*); 4.20 (bs, 1H, C3′*H*); 3.99 (m, 1H, C4′*H*); 3.71 - 3.58 (m, 2H, C5′*H*) ppm.

13° C-NMR (75 MHz, DMSO-d₆, 25 $^{\circ}$ C):

 δ 151.54 (C2); 131.52 - 129.05 (C ar); 143.12 (¹³C8); 87.58 (¹³C1'); 85.28 (C4'); 73.38 (C2'); 69.07 (C3′); 60.05 (C5′) ppm.

N6 -Benzoyl-2′**-O-(tert-butyldimethylsilyl)-3**′**,5**′**-O-(di-tert-butylsilanediyl)-(8, 1´-** 1^{\prime} -¹³C)**adenosine (3)**

To a stirred suspension of **1** (1.0 eq, 0.40 g, 1.08 mmol) in 3 mL dry DMF was added dropwise under stirring and argon at 0 °C di-*tert*.-butylsilandiylditriflate (1.15 eq, 545 mg, 1.24 mmol). After stirring at the same temperature for 30 minutes, imidazole (5.0 eq, 366 mg, 5.39 mmol) was added at once. Then, the mixture was allowed to warm to room temperature, stirring was continued for further 15 minutes. *Tert.*-butyldimethylchlorosilane (TBDMS-Cl, 1.2 eq, 195 mg, 1.30 mmol) was added at once and stirring was continued at 60°C for 2 h. After TLC showed complete conversion, DMF was removed *in vacuo* to give viscose oil. This was dissolved in chloroform, extracted one time with saturated N aHCO₃ solution and two times with brine. The organic phase was dried over $Na₂SO₄$ and evaporated to dryness. The crude product was applied on a silica gel column with methylene chloride and eluted using a gradient of 0 to 5 % MeOH in methylene chloride.

Yield: 260 mg (40 %)

TLC: (CH₂Cl₂/MeOH = 9.5/0.5) R_f = 0.6

<u>¹H-NMR (300 MHz, CDCl₃, 25°C):</u>

δ 8.94 (bs, 1H, N*H*); 8.76 (s, 1H, C2*H*); 8.04 (d, 1H, ¹ JCH = 212.6 Hz, 13C8*H*); 8.01, 7.68 - 7.50 (m, 5H, CHar); 6.00 (d, 1H, ¹J_{CH} = 170.8 Hz, ¹³C1′*H*); 4.64 (d, 1H, C2′*H*); 4.53 - 4.46 (m, 2H, C3′*H*, C5′*H*); 4.25 (m, 1H, C4′*H*); 4.04 (t, 1H, C5′H); 1.08 - 0.95 (3 x s, 27H, 3 x tBu); 0.18 - 0.16 $(2 \times s, 6H, Si-Me₂)$ ppm.

$13C-NMR$ (75 MHz, CDCl₃, 25°C):

δ 153.30 (C2) ;143.11 (13C8); 88.23 (13C1′); 75.96 (C3′); 75.60 (C4′); 74.82 (C2′); 67.98 (C5′); 27.66, 27.20, 26.06 (3 x C(*C*H3)3); 22.88, 20.47, 18.44(3 x *C*(CH3)3); -4.18, -4.86 (Si(*C*H3)2) ppm.

N6 -Benzoyl-2′**-O-(tert-butyldimethylsilyl)-5**′**-O-(4,4**′**-dimethoxytrityl)-(8,1**′**- 13C)-adenosine (4)** To a stirred solution of Compound **2** (260 mg, 0,416 mmol) in 2 mL of methylene chloride was added at 0°C at once a pre-prepared mixture of 45 µL (70% hydrogen fluoride, 30% pyridine) diluted with 270 µL of pyridine. Completion of the reaction was monitored via thin layer chromatography. After 1 h of stirring at 0 $^{\circ}$ C, the mixture was diluted with methylene chloride and washed once with saturated sodium bicarbonate solution. The organic phase was dried over Na₂SO₄, evaporated to give a pale-yellow foam and dried in high vacuum for 30 minutes. The solid was co-evaporated twice with dry pyridine and dissolved in 5 mL of dry pyridine. To this solution was added one spatula 4-(dimethylamino)pyridine (DMAP, catalytic amount) and 4,4′-dimethoxytrityl chloride (DMT-Cl, 1.2 eq, 168 mg, 0.499 mmol) in three portions within 30 minutes. After stirring for 3 h at room temperature, the reaction was quenched with 1 mL of methanol, stirring was continued for 10 minutes and the mixture was evaporated to give an oily residue. This was dissolved in methylene chloride, washed once with saturated sodium bicarbonate solution and twice with 5 % citric acid solution. The organic phase was dried over $Na₂SO₄$, evaporated to dryness and dried in high vacuum. The yellow foam was applied to a silica gel column with methylene chloride and eluted using a gradient of 0 to 5 % methanol in methylene chloride.

Yield: 300 mg (91 %)

TLC: (CH₂Cl₂/MeOH = 9.5/0.5) R_f = 0.7

 1 H-NMR (300 MHz, DMSO-d $_6$, 25 $^{\circ}$ C):

δ 11.19 (bs, 1H, N*H*); 8.66 (s, 1H, C2*H*); 8.60 (d, 1H, ¹J_{CH} = 214.7 Hz, ¹³C8*H*); 8.05, 7.66 - 6.84 (C*H*ar); 6.09 (d, 1H, ¹J_{CH} = 167.8 Hz, C1′*H*); 5.18 (d, 1H, C3′O*H*); 4.90 (m, 1H, C2′*H*); 4.29 (m, 1H, C3′*H*); 4.14 (d, 1H, C4′*H*); 3.73 (s, 6H, 2 x OMe); 1.09, 1.03, 0.89 (3 x s, 27H, 3 x tBu); -0.01, -0.11 (2 x s, 6H, 2 x Si-Me) ppm.

$13C-NMR$ (75 MHz, DMSO-d₆, 25°C):

 δ 149.41 (C2); 141.26 (¹³C8); 131.61 - 124.49 (C ar); 110.24 (C ar); 92.69 (¹³C1'); 82.41 (C4'); 72.88 (C2′); 67.65 (C3′); 62.30 (C5′) 52.59 (OCH3) 22.13 (Si-C(*C*H3)3); -3.22, -3.49 (2 x s, 6H, $Si-Me₂$) ppm.

N6 -Benzoyl-2´-O-(tert-butyldimethylsilyl)-5´-O-(4,4′-dimethoxytrityl)-(8, 1′- 13C)-adenosine 3′-(cyanoethyl-*N***,***N***-diisopropylphosphoramidite) (5)**

To a stirred solution of **4** (1.0 eq, 300 mg, 0.38 mmol) in 4 mL of dry tetrahydrofuran was simultaneously added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (CEP-Cl, 1.8 eq, 160 mg, 0.680 mmol) and *N*,*N*-diisopropylethylamine (5.0 eq, 245 mg, 1.90 mmol). After 3 h, the mixture was diluted with methylene chloride and washed with saturated sodium bicarbonate solution. The organic phase was dried over $Na₂SO₄$, evaporated to dryness and purified by silica gel column chromatography using a gradient ethylacetate/hexane 2:8 to 1:1 (+ 3% triethylamine).

Yield: 300 mg (80 %)

TLC: (ethyl acetate/hexane = $6/4$) R_f = 0.7

<u>¹H-NMR (300 MHz, CDCl₃, 25°C):</u>

δ 8.97 (bs, 1H, N*H*); 8.72 (d, 1H, C2*H*); 8.24 (dd, 1H, ¹J_{CH} = 212.3 Hz, ¹³C8*H*); 8.03, 7.62 - 7.28, 6.82 (C*H*ar); 6.08 (d, 1H, ¹J_{CH} = 166.5 Hz); 5.09 (m, 1H, C2′*H*); 4.45 - 4.38 (m, 2H, C3′H, C4′*H*); 3.93 (m, 2H, 2 x NC*H*(CH3)2, 3.78 (s, 6H, 2 x OMe); 3.66 - 3.55 (m, 3H, C5′H, POC*H*2); 3.34 (m, 1H, POCH₂); 2.65, 2.32 (m, 2H, CH₂CN); 1.19 (m, 12H, 2 x NCH(CH₃)₂) 0.76 (s, 9H, Si-C(CH₃)₃); -0.02 , -0.20 (2 x d, 6H, Si-Me₂) ppm.

$31P$ -NMR (121 MHz, CDCl₃, 25°C):

δ 151.81; 149.76 ppm.

1.5 Solid phase synthesis of residue-specific labeled 24nt-*CCR***5 donor fragments**

The donor fragments were synthesized on a solid-phase DNA synthesizer (ABI 394) using phosphoramidites. Unlabeled 2′-*O*-TBDMS-nucleoside phosphoramidites, columns and reagents purchased from ChemGenes in Wilmington, Massachusetts, USA. Standard phosphoramidite protocols for the 1 µmol scale were used, except the coupling time per monomer was extended to 6 minutes. N^6 -benzoyl-2'-O-TBDMS-3'-CEP-5'-dimethoxytrityl-8,1'-¹³C-adenosine was incorporated into the donor fragments residue-specifically (**Supplementary Table 1**). The RNA samples were cleaved twice from the column using a mixture of 0.65ml aqueous ammonia and 0.65ml of 40% methylamine at 37˚C, first for 3h and second for overnight. The solution was combined and evaporated to dryness under reduced pressure. The dried white powder was dissolved in 0.5ml of anhydrous DMSO then mixed well with 0.5ml of triethylamine trihydrofluoride (TEA, 3HF) for overnight for removal of the protective group. The fullydeprotected RNA fragments were purified to homogeneity by 12% denaturing PAGE, recovered by electroelution, and washed extensively with H_2O for later on ligation.

1.6 RNA preparation by *in vitro* **transcription**

The full-length *CCR*5 RNA and the donor and acceptor fragments of the RNA were transcribed as described ^[4] with minor modifications. Transcriptions were performed in 40 mM Tris-HCl, pH 8.0, 1 mM spermidine, 0.01% Triton-10, 80 mg/mL PEG, 0.3 µM DNA template, 10 mM DTT, 2 U of inorganic pyrophosphatase per milliliter of reaction solution, 5-10 mM NTPs, 5-20 mM $MgCl₂$, and 1-2 mg/mL T7 RNA polymerase. Reactions proceeded for 3 hours at 37°C. For each RNA transcribed, the concentrations of MgCl₂, T7 RNA polymerase, and NTPs were optimized at small scales (50-100 μ L) before scaling up to large-scale transcriptions. Typically 10~20 mL of transcription were performed for an adequate NMR sample. DNA templates were either purchased from IDT or amplified from the plasmid.^[2] After transcription, samples were extracted and purified as described previously.^[5] Electroeluted samples were washed extensively with $ddH₂O$ and stored at -20 $^{\circ}$ C until use.

1.7 RNA ligation

Based on previous studies that mutation at the Stem 2 of the RNA caused a decreased binding capacity to miR-1224, the RNA was split into two fragments such that the 24nt donor fragment spanned the proposed interaction site.^[2] The unlabeled 72nt acceptor fragment and the synthesized residue/site-specific labeled 24-nt donor fragment of CCR5 RNA were ligated using a method described.^[6] The unlabeled acceptor fragment with a protective 5'-triphosphate group and a reactive 3'-OH group was synthesized by *in vitro* transcription using T7 RNA polymerase without any further treatment. The donor was prepared by either *in vitro* transcription followed by enzymatic dephosphorylation using RNA 5′ polyphosphatase or solid phase synthesis followed by phosphorylation with T4 polynucleotide kinase to obtain the desired monophosphorylated group at the 5'terminus. A 40mer DNA splint was used to facilitate RNA ligation. Excess amounts of the acceptor fragment and the DNA splint were used in the ligation reaction to achieve maximum conversion of the donor fragment into the ligated residue/site-specific isotope labeled *CCR*5 product with a typical ligation yield (based on the donor fragment) of >60%.

1.8 Sample preparation

The purified ligated RNA was lyophilized and dissolved in a buffer containing 10 mM sodium phosphate, pH 6.2, and 50 mM NaCl, 4 mM Mg^{2+} , 0.2 mM EDTA, 10 µM DSS, 0.1% sodium azide in 100% D_2O without further denaturing/refolding process.^[2] The RNA concentration was 60~200 µM for NMR experiments and dynamic measurements, respectively, as determined by the UV absorbance at 260 nm and the calculated extinction coefficient (1,030,900 M^{-1} cm⁻¹) of the RNA's base composition.

1.9 NMR experiments

All NMR spectra were processed and analyzed using the Bruker TOPSPIN 3.5 and NMRPipe/ NMRDraw.^[7,8] All NMR experiments were performed at a temperature of 298 K either on a Bruker Avance 600 MHz (14.1 T) spectrometer equipped with actively shielded z-axis gradient triple resonance probes or on a Bruker Advance III HD 800 MHz (18.8T) instrument equipped with a TCI cryo probe. The CPMG frequencies (v_{CPMG}) of 0, 0, 50, 100, 100, 150, 200, 300, 400, 500, 600, 600, 700, 800, 900, and 1000 Hz, were used at 600 and 800 MHz. The constant time relaxation delay *T*relax was 20 ms for all experiments. A total of 2048 x 48 complex data points were recorded at 600 MHz and 800 MHz with a spectral width of 3.2 ppm in the carbon dimensions. The number of transients was 48 and the recovery delay was set to 2.0 s at 600 MHz and 800 MHz, yielding total experimental times of ~70 hours.

1.10 Data analysis

CPMG RD data were analyzed as described $[3,9]$. Peak intensities were extracted and used for calculate effective transverse relaxation rates $R_{2,eff}$ according to:

$$
R_{2, \text{ eff}} = -\frac{1}{T_{\text{relax}}} \ln \frac{I \left(v_{\text{CPMG}} \right)}{I_0}
$$

with T_{relax} being the constant time relaxation delay, $I(v_{CPMG})$ the peak intensity at the CPMG frequency v_{CPMG} , and I_0 the peak intensity from the averaged reference experiments. Dispersion curves were obtained for seven residues except that the signal intensity of A79-C1' rapidly decreases during the CPMG relaxation delay which prevents further characterization and analysis. Nessy software was used to fit the data to the Carver–Richards equation describing two-state exchange.^[10] The kinetic and thermodynamic parameters were derived from 1000 Monte Carlo runs.

1.11 Supplementary references

- [1] B. Chen, X. Zuo, Y.-X. Wang, T. K. Dayie, *Nucleic Acids Res.* **2012**, *40*, 3117–30.
- [2] A. T. Belew, A. Meskauskas, S. Musalgaonkar, *Nature* **2014**, *512*, 265–269.
- [3] A. P. Longhini, R. M. Leblanc, O. Becette, C. Salguero, C. H. Wunderlich, B. A. Johnson, V. M. D'Souza, C. Kreutz, T. K. Dayie, *Nucleic Acids Res.* **2015**, *44*, e52.
- [4] J. F. Milligan, D. R. Groebe, G. W. Witherell, O. C. Uhlenbeck, *Nucleic Acids Res.* **1987**, *15*, 8783–8798.
- [5] O. H. Gumbs, R. A. Padgett, K. T. Dayie, *RNA* **2006**, *12*, 1693–1707.
- [6] B. Chen, R. LeBlanc, T. K. Dayie, *Angew. Chemie - Int. Ed.* **2016**, *55*, 2724–2727.
- [7] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J.Biomol.NMR* **1995**, *6*, 277–293.
- [8] B. A. Johnson, R. A. Blevins, *J.Biomol.NMR* **1994**, *4*, 603–614.
- [9] C. H. Wunderlich, R. Spitzer, T. Santner, K. Fauster, M. Tollinger, C. Kreutz, *J. Am. Chem. Soc.* **2012**, *134*, 7558–7569.
- [10] M. Bieri, P. R. Gooley. *BMC Bioinformatics*, **2011**,12:421.

2 SUPPLEMENTARY RESULTS

Supplementary Table 1 Chemical synthesis of the donor fragments bearing 8,1′- 13Cmodified nucleotide

*: Lowercase letter 'a' colored in red in the sequences indicates the position at which a selective (8, 1′)- 13C-labeled adenosine residue is inserted *via* solid phase syntheses.

+ : based on the scale of 1 µmol synthetic column.

Supplementary Table 2 Longitudinal relaxation rates of *CCR*5 RNA in the absence of miR-1224

Supplementary Figure 1. (a) Enzymatic Synthesis of adenosine from labeled precursors. (8,1'-¹³C)-ATP was synthesized in a single pot reaction. (1-¹³C)-ribose was first converted to $(1 - {^{13}C})$ -ribose-t-phosphate by ribokinase (RK), which was then transformed to (1-¹³C)-(5-phospho-D-ribosyl-α-1-pyrophosphate) by phosphoribosyl pyrophosphate synthetase (PRPPS). Next, PRPP was coupled to (8-¹³C)-adenine by adenine phosphoribosyl transferase (APRT) to produce (8,1'-¹³C)-AMP, which was phosphorylated in two steps by myokinase (MK) and creatine kinase (CK) to ATP. After purification ATP was dephosphorylated for 4 hours to adenosine by the action of recombinant shrimp alkaline phosphatase (rSAP). **(b)** The site-specific labeled (8,1′- 13° C)-ATP (cyan) was completely converted to the product (magenta) demonstrated by the FPLC profile using a C18 reverse-phase Vydac analytical column.

Supplementary Figure 2 Scheme illustrating a cycle of solid-phase RNA synthesis *via* the phosphoramidite approach. The filled-in black circle implies the CPG support. Base 1 & base 2 represent uracil-1-yl, or any of adenin-9-yl, cytosine-1-yl, and guanin-9-yl that are protected by acetyl groups.

Supplementary Figure 3 NMR data of the ligated *CCR*5 RNA labeled only the eight site-specific (8,1'-¹³C)-adenosine residues in the donor fragment in the absence of miR-1224, 298 K, 600 MHz. 2D HMQC spectra of the RNA at base C-8 (a) and ribose C-1′ (b) regions. (c) 2D edited-edited NOESY spectrum obtained for the RNA sample in the absence of miR-1224, 298 K, 800 MHz, showed inter-residue cross-peaks between A75H1′ and A76H8, and between A95H1′ and A96H8, allowing straightforward and unambiguous assignments of residues A75 and A96.

Supplementary Figure 4 ¹³C CPMG RD profiles of ribose anomeric carbons from residues A75 (a), A76 (b), A88 (c), and A95 (d) of the ligated *CCR*5 RNA recorded at 150 (red dots) and 200 MHz (blue dots) carbon Larmor frequency, 298 K. Experimental data are shown as circles and the fits are depicted as dotted lines. Black circles represent repeat experiments.

Supplementary Figure 5 Longitudinal relaxation rates of the ligated CCR5 containing eight ¹³C probes across the donor fragment recorded on a 600 MHz NMR spectrometer at 298 K. Different residues are color-coded.

Supplementary Figure 6 Electrophoresis mobility shift assay of *CCR*5 RNA pseudoknot. 12% native PAGE of the pseudoknot in 50 mM Tris-Hepes, 0.1 M NaCl, 2 mM Mg^{2+} , pH 8.0, in the absence or presence of miR-1224. L1 & L6: no miR-1224; L2~L5, different molar ratios between *CCR*5 and miR-1224, L2: 1/0.5; L3: 1/1; L4: 1/2; L5: 1/4.