Supporting Information

Ribonucleosides for an Artificially Expanded Genetic Information System

Hyo-Joong Kim, Nicole A. Leal, Shuichi Hoshika, and Steven A. Benner

[1]Foundation for Applied Molecular Evolution (FfAME), 720 SW 2nd Avenue, Suite 201, Gainesville, FL 32601

[2]The Westheimer Institute for Science and Technology (TWIST), 720 SW 2nd Avenue, Suite 208, Gainesville, FL 32601

[*] Corresponding author; sbenner@ffame.org

Table of Contents

Table S1. Sequences of standard and P-containing transcription templates with T7 RNA polymerase promoter segment (underlined)

		Sequence
Top Strand	NLT1	5'- GCGTAATACGACTCACTATAGG
Bottom Strand	NL Standard	3'- CGCATTATGCTGAGTGATAT CCG TCT CTC CTT CTT CAT GCT GTC CGT TCG-5'
	NLP1	3'- CGCATTATGCTGAGTGATAT CCG TCT CTC CTT CTT PAT GCT GTC CGT TCG-5'
	NLP2	3'- CGCATTATGCTGAGTGATAT CCG TCT CTC CTT CTT PAP GCT GTC CGT TCG-5'
	NLP3	3'- CGCATTATGCTGAGTGATAT CCG TCT CTC CTT CTT PAP GPT GTC CGT TCG-5'
	NLP4	3'- CGCATTATGCTGAGTGATAT CCG TCT CTC CTT CTT PPP GCT GTC CGT TCG-5'

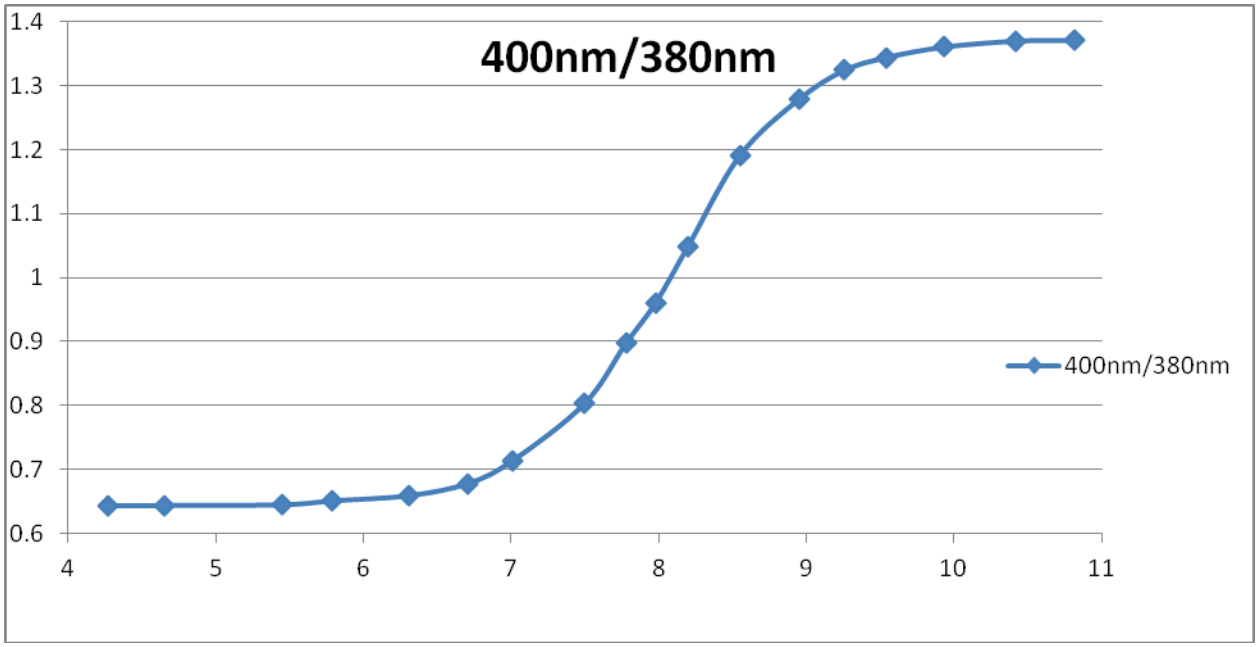
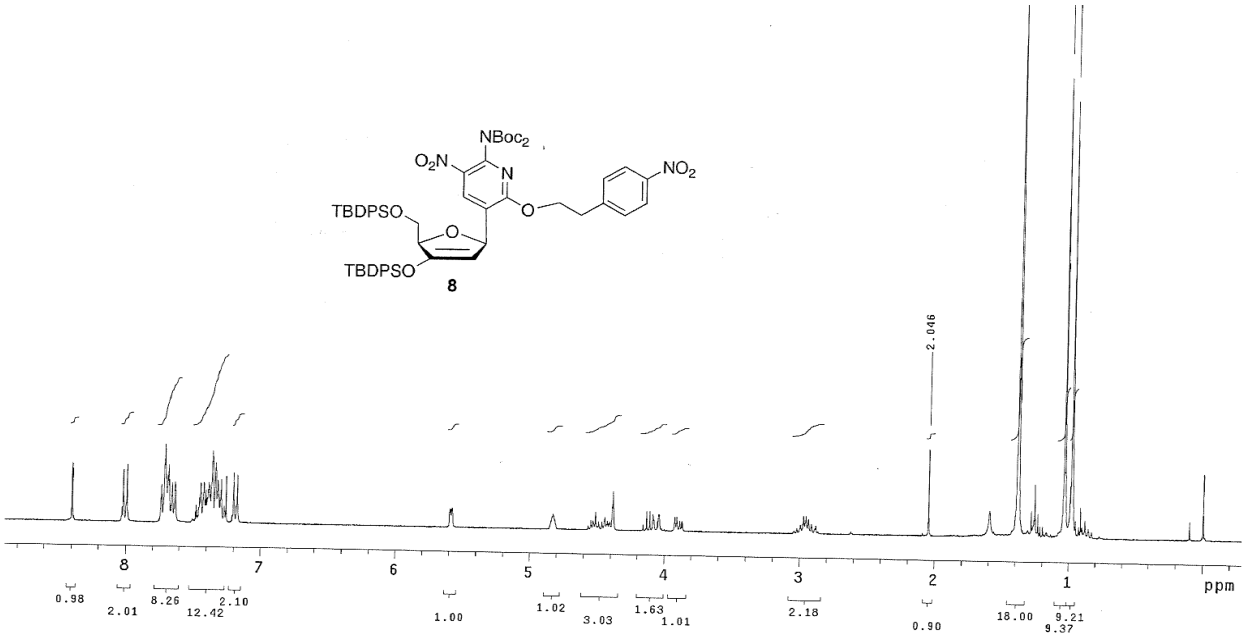
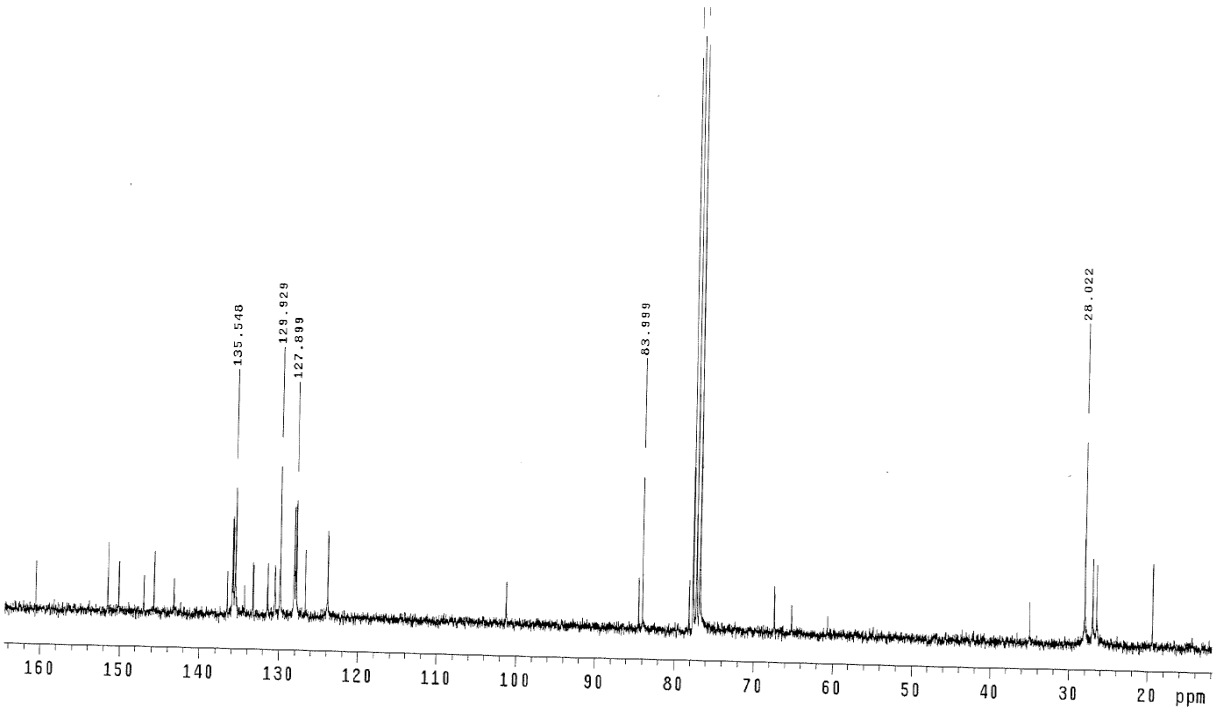
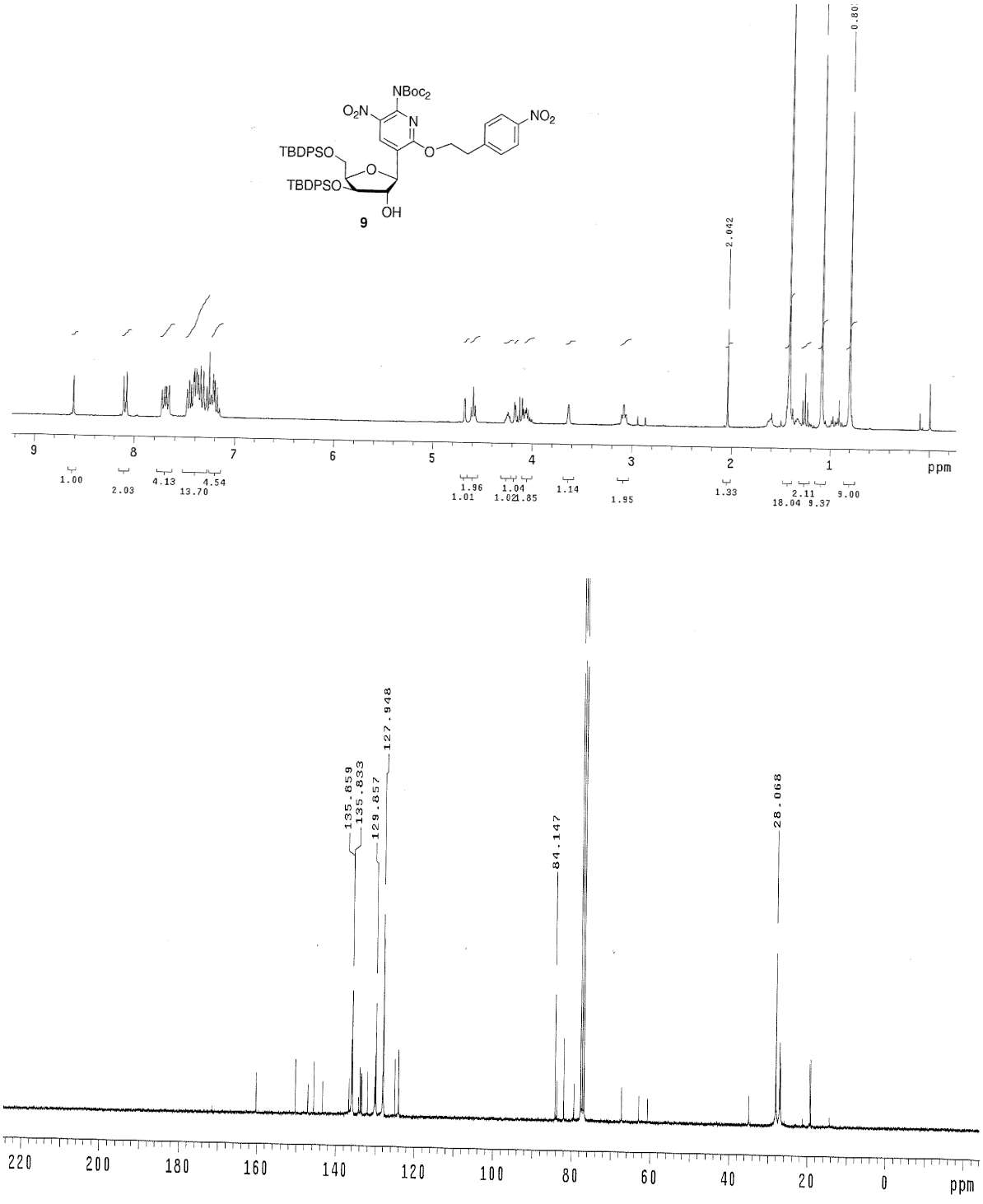


Figure S1. The ratio of absorbance at 400 nm and 380 nm measured as a function of pH gave an estimate of the pK_a of ribonucleoside **15** (r**Z**) to be 7.9.

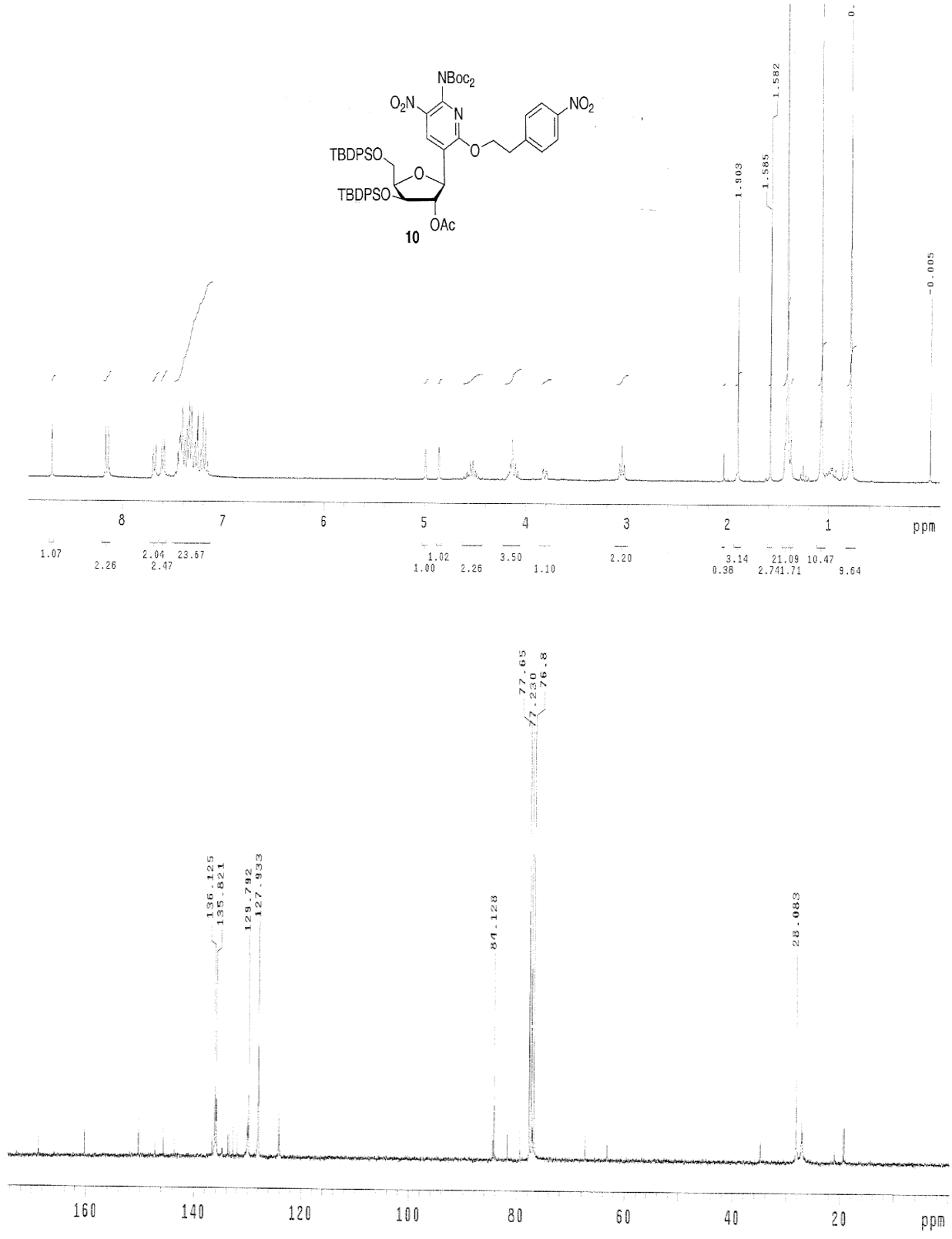




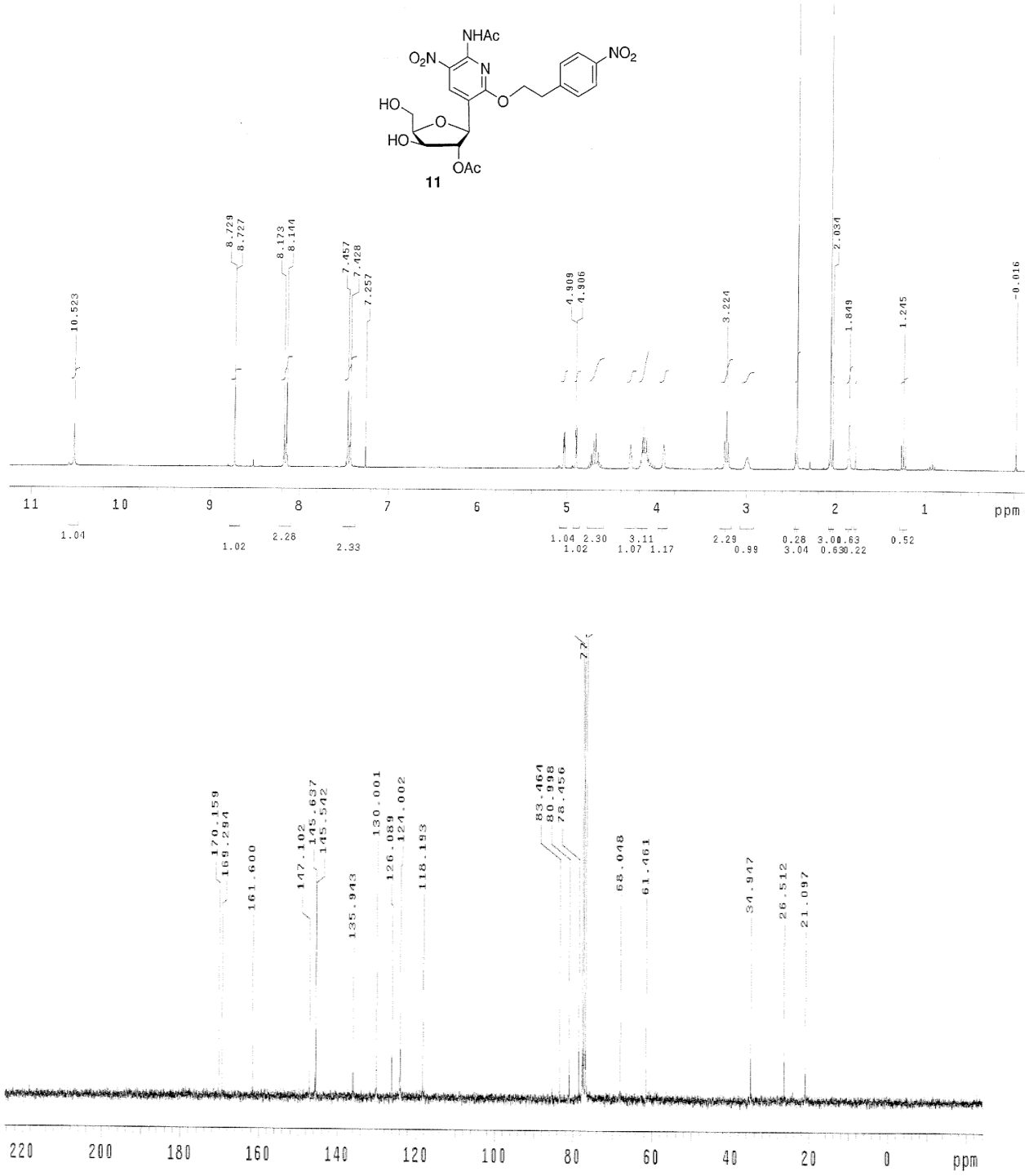
^{1}H and ^{13}C NMR spectra of compound **8**



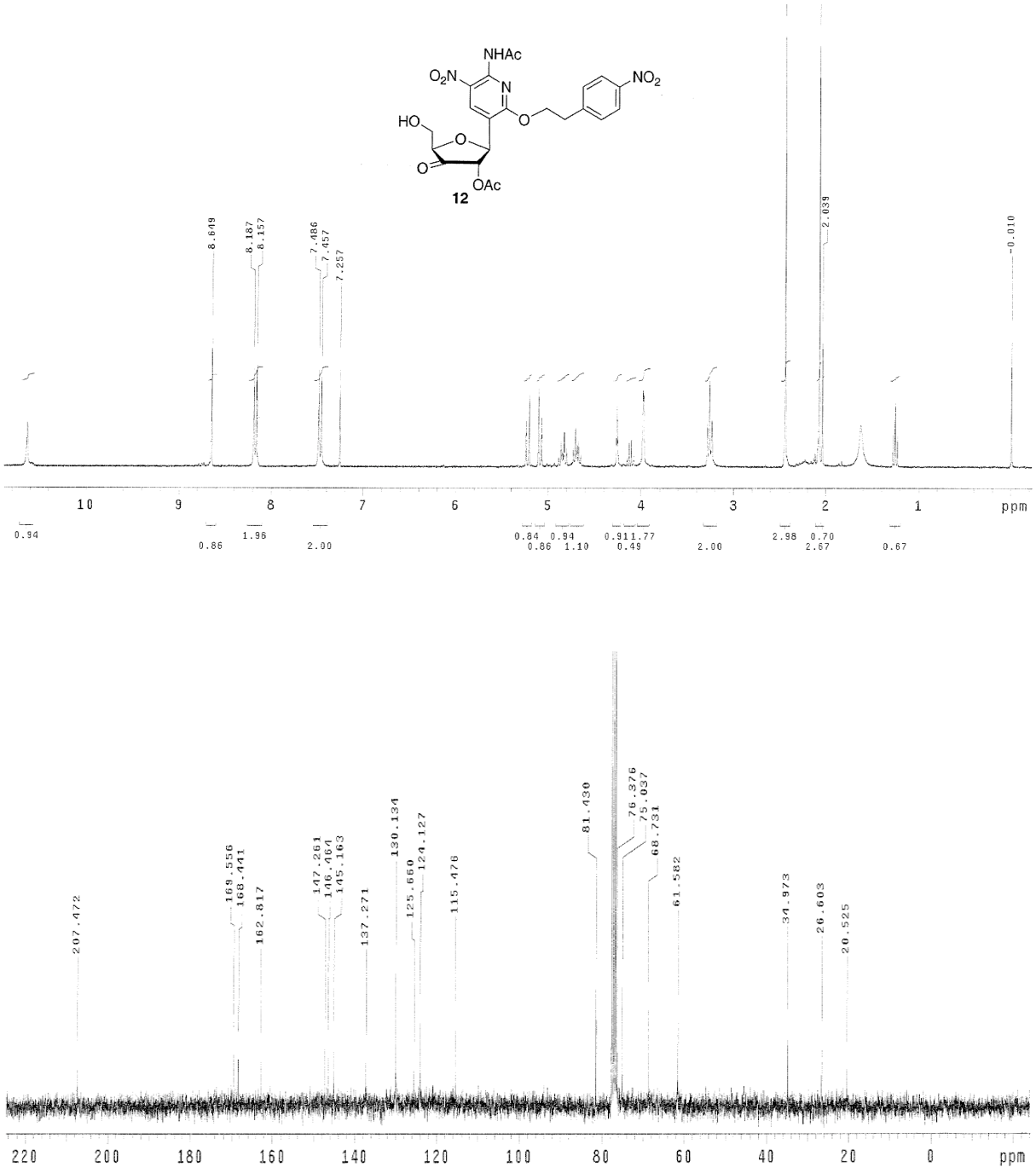
^{1}H and ^{13}C NMR spectra of compound **9**



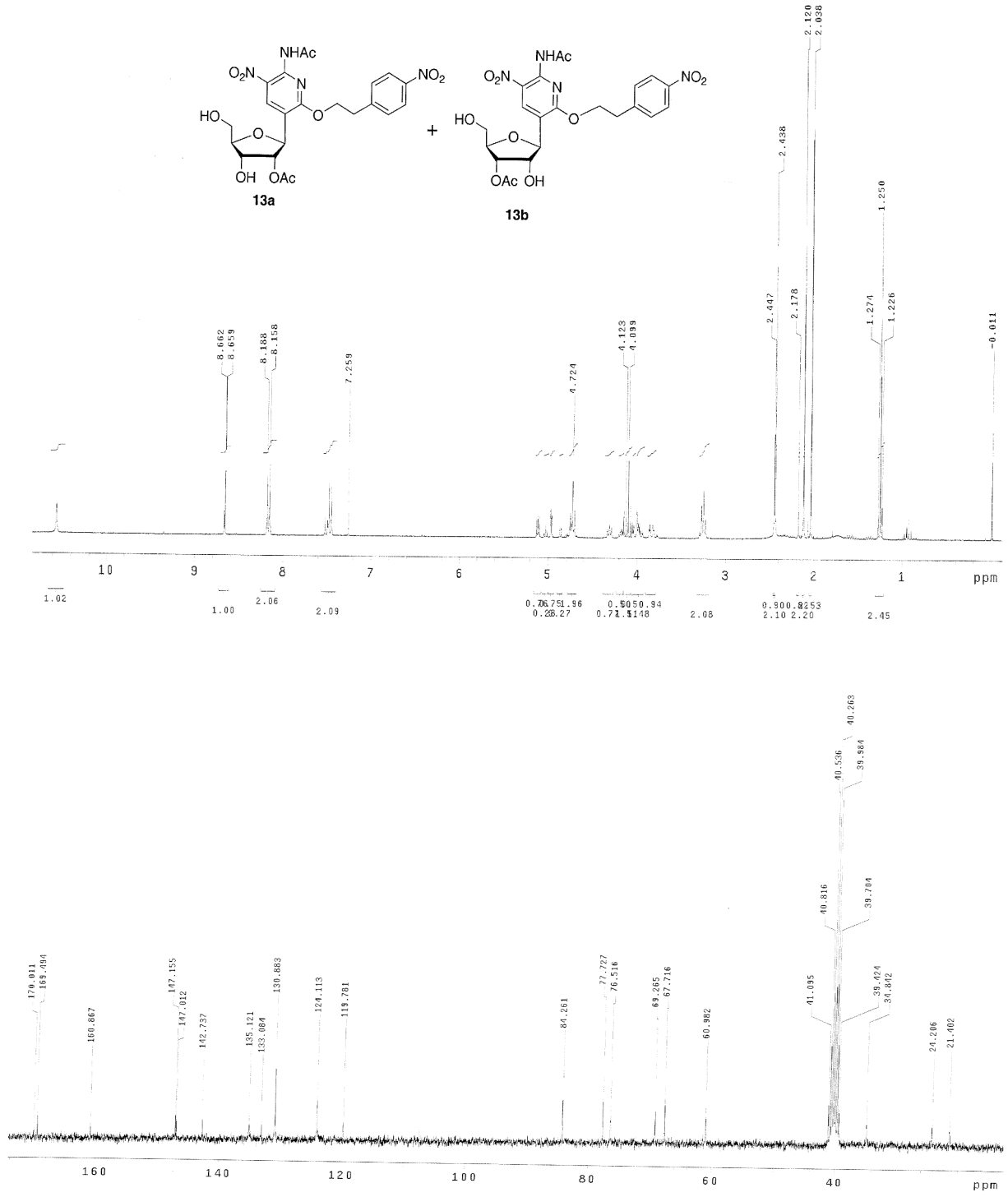
^{1}H and ^{13}C NMR spectra of compound **10**



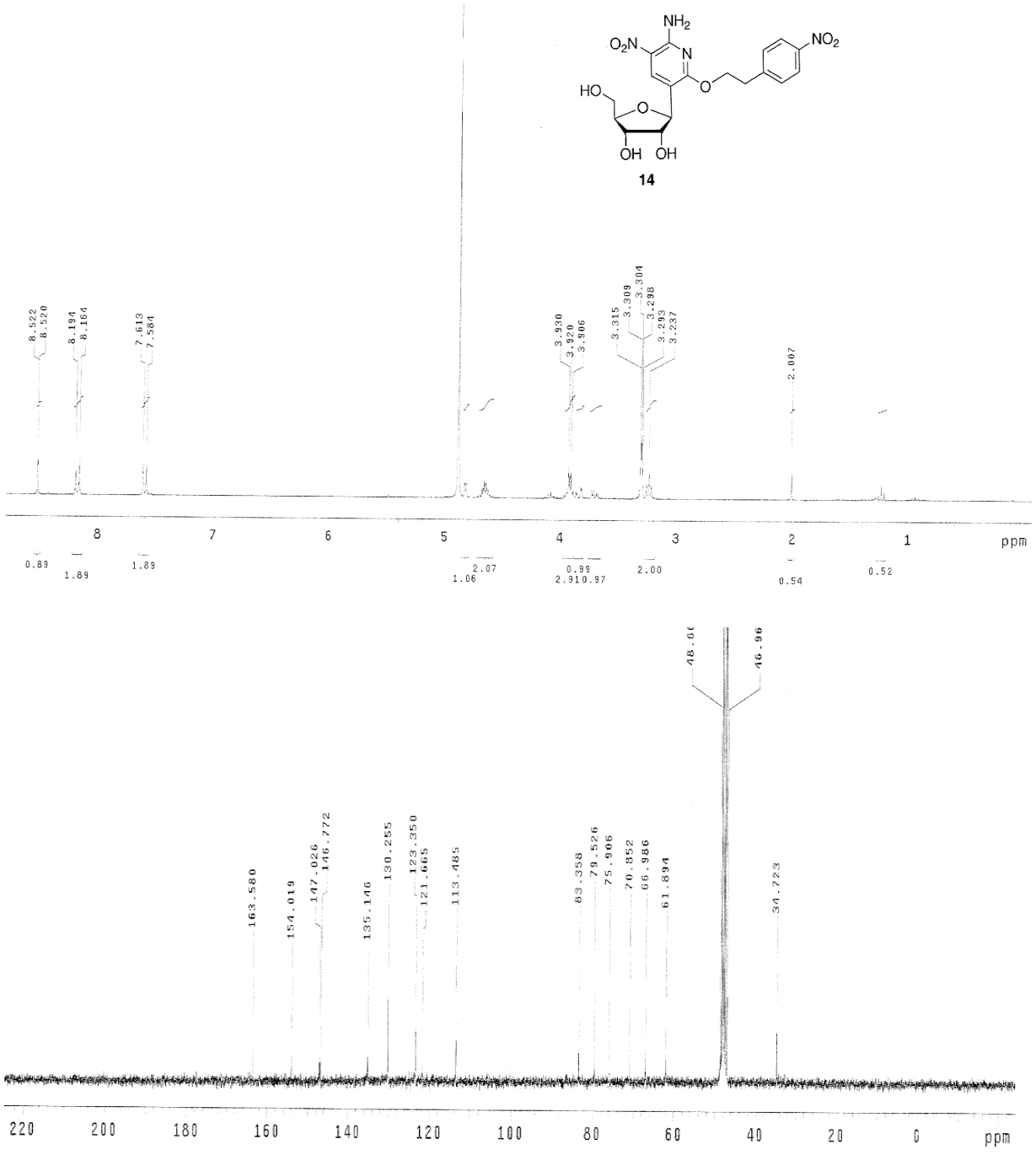
^{1}H and ^{13}C NMR spectra of compound **11**



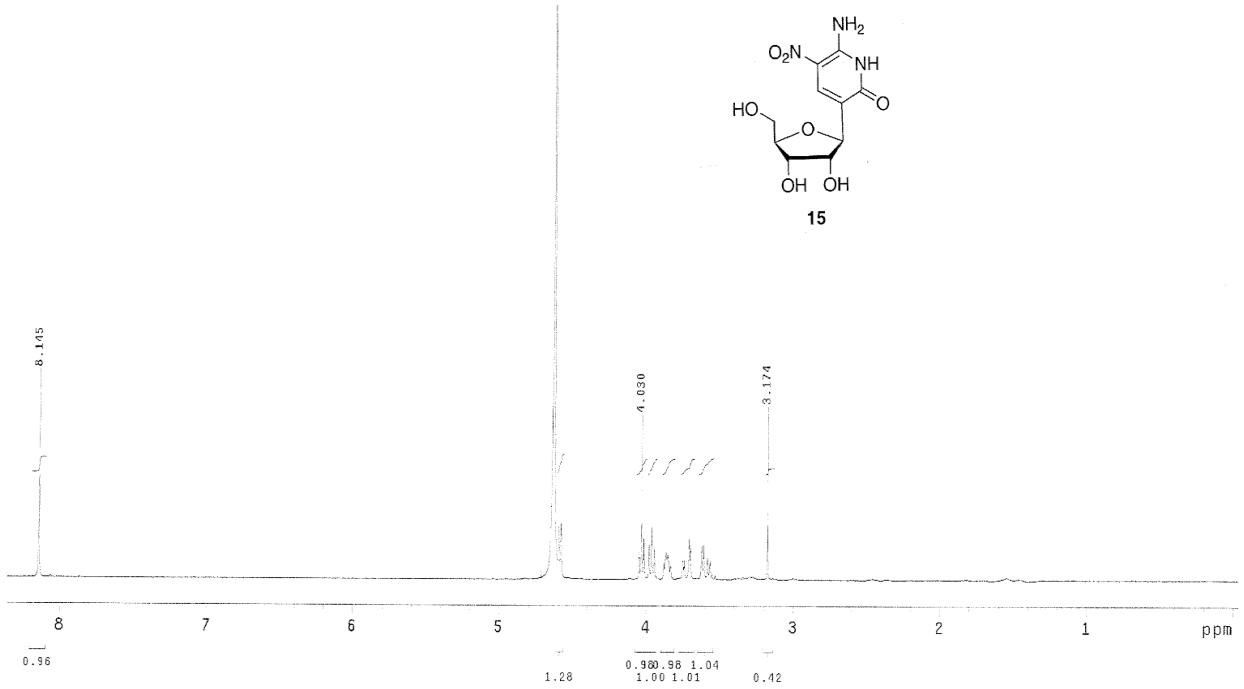
^{1}H and ^{13}C NMR spectra of compound **12**

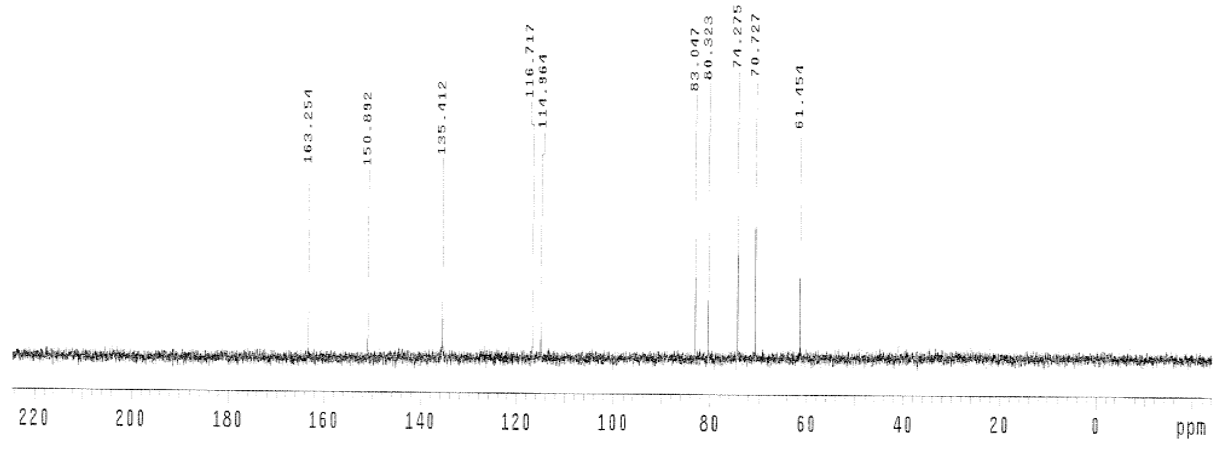


^{1}H and ^{13}C NMR spectra of compound **13**

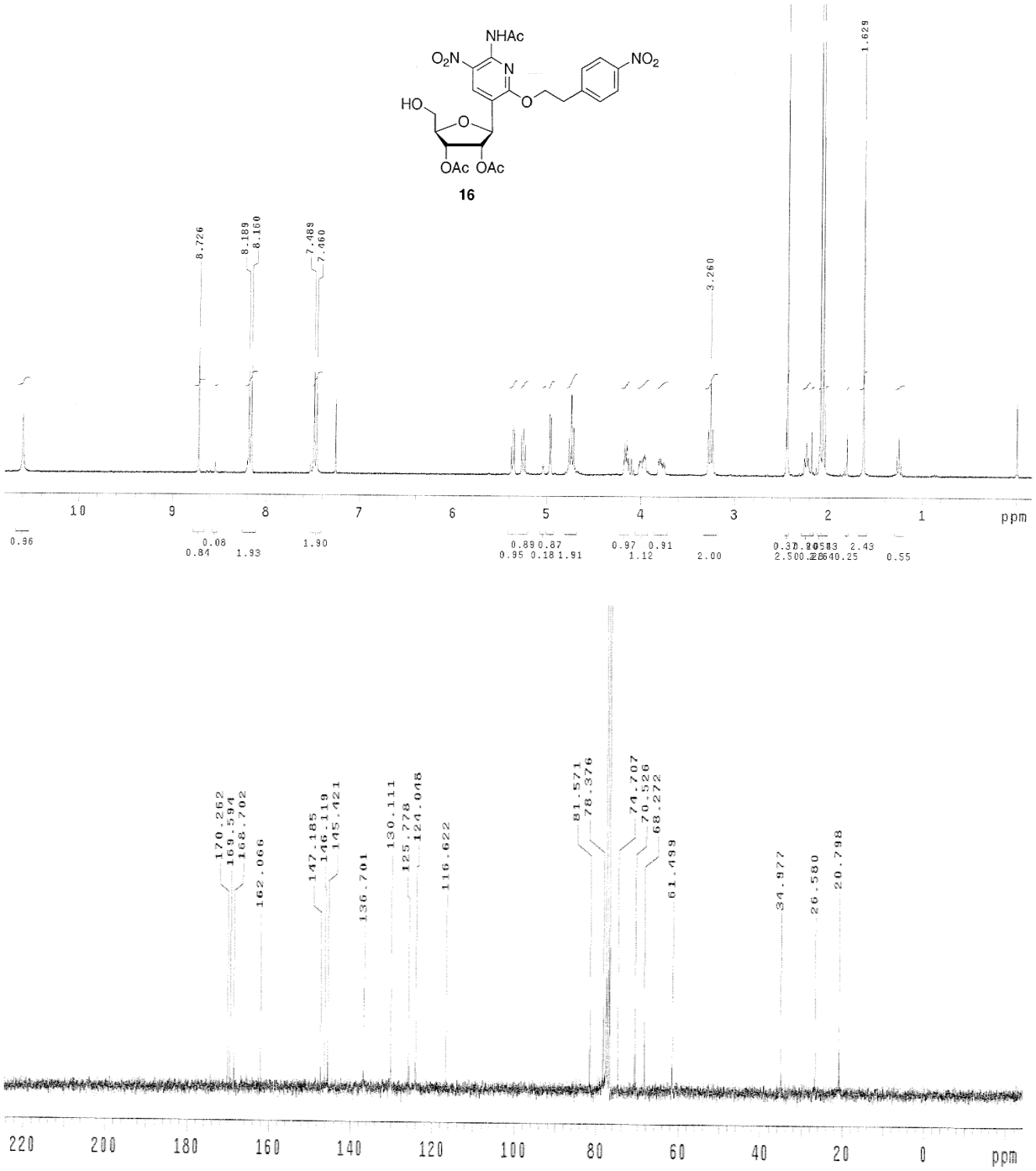


^{1}H and ^{13}C NMR spectra of compound **14**

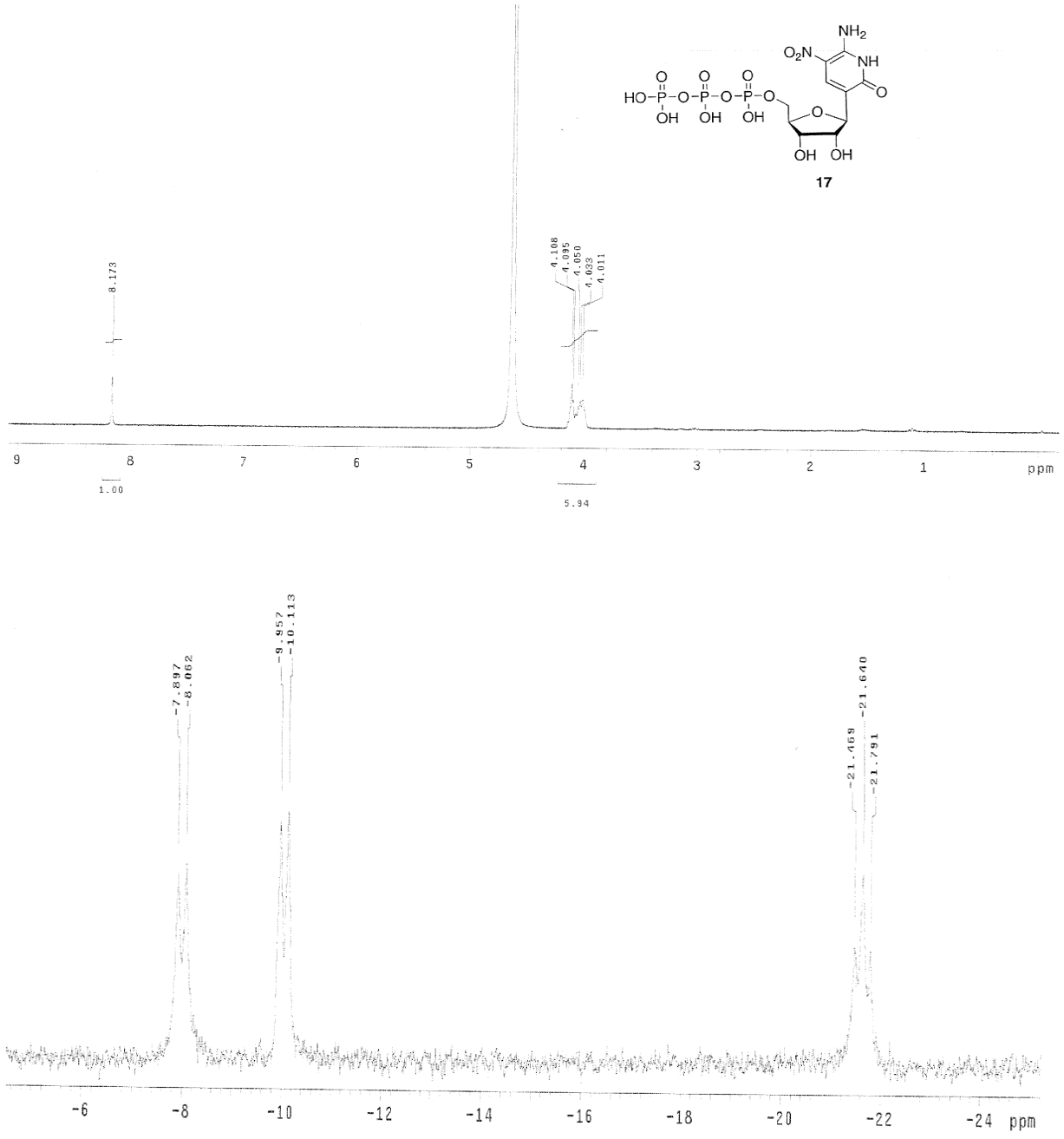




^{1}H and ^{13}C NMR spectra of compound **15**



^{1}H and ^{13}C NMR spectra of compound **16**



^{1}H and ^{31}P NMR spectra of compound **17**

Table S2. Sequences of standard and Z-containing transcription templates with T7 RNA polymerase promoter segment (underlined)

		Sequence
Top Strand	NLT1	5'- GCGTAATACGACTCACTATAGG
Bottom Strand	NL Standard	3'- CGCATTATGCTGAGTGATAT CCG TCT CTC CTT CTT CAT GCT GTC CGT TCG-5'
	NLZ1	3'- CGCATTATGCTGAGTGATAT CCG TCT CTC CTT CTT ZAT GCT GTC CGT TCG-5'
	NLZ2	3'- CGCATTATGCTGAGTGATAT CCG TCT CTC CTT CTT ZAZ GCT GTC CGT TCG-5'
	NLZ3	3'- CGCATTATGCTGAGTGATAT CCG TCT CTC CTT CTT ZAZ GZT GTC CGT TCG-5'
	NLZ4	3'- CGCATTATGCTGAGTGATAT CCG TCT CTC CTT CTT ZZZ GCT GTC CGT TCG-5'

Incorporation of rPTP opposite template dZ

Transcription using standard and dZ templates

Transcription templates were prepared by independently combining equimolar ratios of top strand (NLT1) and bottom strand (NL standard, NLZ1, NLZ2, NLZ3 or NLZ4) in 1X transcription buffer (20 mM NaCl, 40 mM Tris pH 7.8, 6 mM $MgCl_2$, 2 mM spermidine, and 10 mM DTT), heating to 95°C and then cooling to room temperature. Transcription reactions contained a final concentration of 0.2 μM template DNA, 1X transcription buffer (20 mM NaCl, 40 mM Tris pH 7.8, 6 mM MgCl2, 2 mM Spermidine, and 10 mM DTT), 1 μCi/μL α^{32}P-GTP, MC T7 RNA Pol RNA polymerase (0.05 μg/μL final) and 0.5 mM each rNTP in the minus experiment or 0.5 mM each rNTP and rZTP in the plus experiment. Two different sources of rPTP were tested were tested and similar results were found; only the HJK rPTP preparation is shown. Reactions were incubated at 37°C for 40 min and 2 hours. Reactions were quenched with 3-fold formamide quench buffer and samples were resolved on a 20% PAGE.

Varying pH in transcription reactions using standard and NLZ2 template

Transcription reactions were done as outlined above except at pH 7.3, 7.8 and 8.3 at 25°C. The amount of 30mer transcription product was determined using the adjusted volume of radiation on the phosphorimager.

Figure S2: PAGE (20%) of transcription products by internal labeling. These reactions show minus (rNTPs) and plus (rNTPs + rPTP) experiments. Transcription templates (right hand panel) were radiolabeled with γ^{32}P-ATP to show purity. In the minus reactions there was no significant pausing seen in the regions where dZ is located. With increasing Z's there is an increase in the amount of abortive transcripts and a decrease in the amount of full length product. The increase in the amount of abortive transcripts may indicate that Z is not far enough away from the promoter region and the polymerase falls off before the transition into the elongation phase.

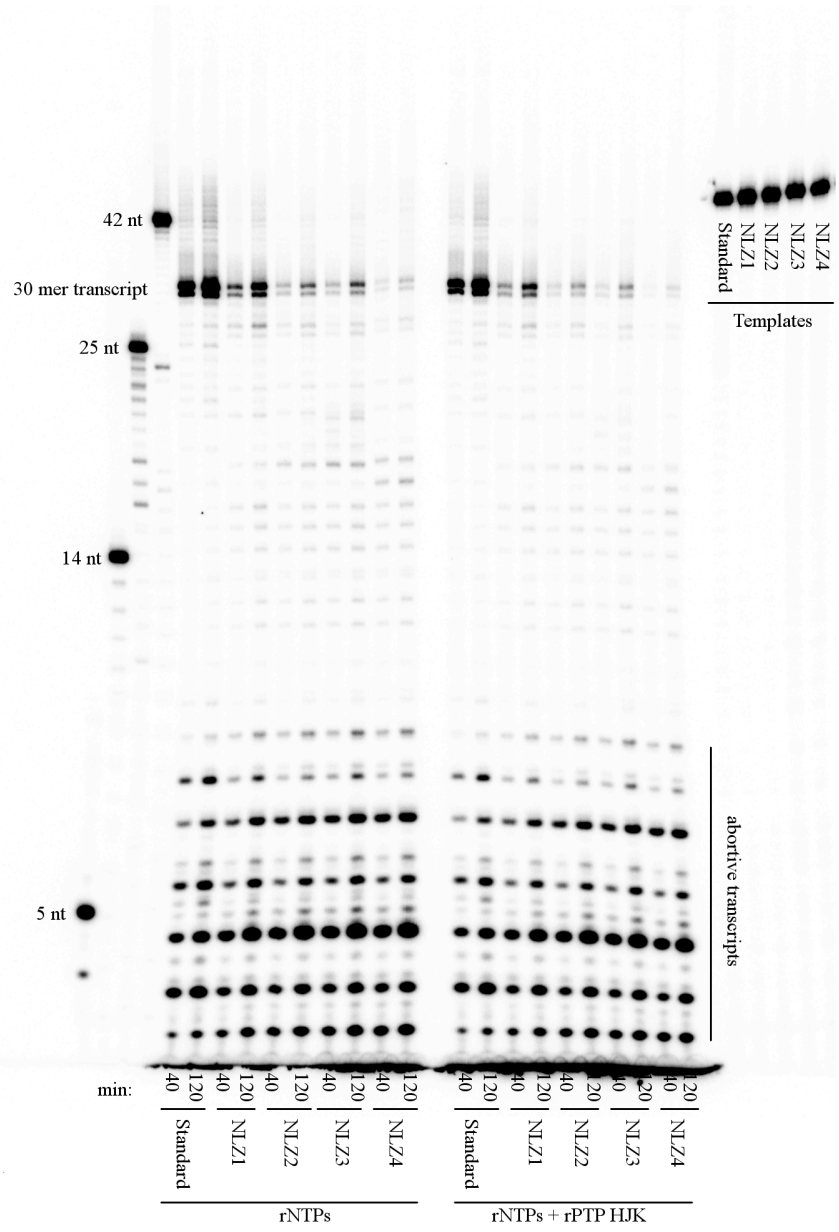


Figure S3: PAGE (20%) of transcription products by internal labeling. These reactions show minus (rNTPs) and plus (rNTPs + rPTP) experiments at varying pH's. At higher pH, less overall product in the + assays is seen.