Supporting Information

Efficient Synthesis of DNA Containing the Guanine Oxidation-Nitration Product 5-Guanidino-4-nitroimidazole: Generation by a Post-Synthetic Substitution Reaction

William L. Neeley, Paul T. Henderson,[†] and John M. Essigmann[*]

Department of Chemistry and Biological Engineering Division, Massachusetts Institute of Technology, Cambridge, MA 02139

[†] Present address: Lawrence Livermore National Laboratory, Biology and Biotechnology Research Program, L-411, Livermore, CA 94551

Email: jessig@mit.edu

Materials and Methods

Solvents were purchased anhydrous from Sigma-Aldrich, except for THF, which was distilled from CaH_2 prior to use. Reagents were purchased and used without further purification from Sigma-Aldrich, except for 3,5-di-*O*-toluoyl-α-1-chloro-2-deoxy-D-ribofuranose, which was purchased from Chemgenes Corporation. 5(4)-Bromo-4(5)-nitro-1*H*-imidazole (**2**) was prepared as previously described from 5(4)-bromo-1*H*-imidazole (**1**) (Sigma-Aldrich).[1] All reactions were conducted under argon atmosphere. DNA synthesis reagents and standard phosphoramidites were purchased from Glen Research. Unmodified ODNs were purchased from IDT, Inc. and HPLC purified prior to use. SVPD was purchased from ICN Biomedical. Nuclease P1 and alkaline phosphatase were purchased from Roche Applied Science.

^{1}H NMR (500 MHz) and ^{13}C NMR (125.8 MHz) spectra were obtained on a Varian Unity Inova spectrometer. ^{31}P NMR (121.5 MHz) spectra were obtained on a Varian Mercury spectrometer and referenced to 85% H_3PO_4. HRMS was performed by the MIT Department of Chemistry Instrumentation Facility. ESI-MS was performed by the MIT Center for Cancer Research and Howard Hughes Medical Institute Biopolymers Laboratory. MALDI-TOF MS were acquired on a PerSeptive Biosystems Voyager-DE STR spectrometer. Absorption spectra and T_m curves were obtained on a Varian Cary 100 or 300 Bio UV-vis spectrophotometer equipped with a temperature controller unit.

5-Bromo-4-nitroimidazole-3',5'-di-*O*-toluoyl-β-2'-deoxy-D-ribofuranose (3) and 4-Bromo-5-nitroimidazole-3',5'-di-*O*-toluoyl-β-2'-deoxy-D-ribofuranose (3A)

Compound **2** (0.500 g, 2.60 mmol) was treated with NaH (60% dispersion in mineral oil, 0.115 g, 2.87 mmol) in 15 mL CH_3CN for 25 min with stirring. 3,5-Di-*O*-toluoyl-α-1-chloro-2-deoxy-D-ribofuranose (1.11 g, 2.87 mmol) was added in portions and the reaction stirred for 3.5 h. The mixture was then diluted with 100 mL of CH_2Cl_2 and washed with 100 mL of H_2O. The aqueous layer was back extracted twice with 50 mL of CH_2Cl_2 and the combined organic layers were dried over Na_2SO_4, evaporated at room temperature on a rotary evaporator, suspended in a minimal amount of 30% hexanes in CH_2Cl_2, and loaded onto a silica column prepared in 30% hexanes in CH_2Cl_2. The first isomer was eluted with a step gradient of 17.5-0% hexanes:CH_2Cl_2 and the second isomer was eluted with a step gradient of 0-4% EtOAc:CH_2Cl_2. Fractions were pooled and concentrated to yield **3A** (0.499 g, 35%) as a white solid and **3** (0.716 g, 50%) as a white foam. Isomer **3A**: R_f = 0.48 (1:2 EtOAc:hexanes). ^{1}H NMR (CD_2Cl_2, 500 MHz): δ = 8.05 (1H, d, *J* = 0.8 Hz), 7.96 (2H, m), 7.81 (2H, m), 7.28 (4H, m), 6.68 (1H, t, *J* = 6.2 Hz), 5.60 (1H, dt, *J* = 3.1, 6.2 Hz), 4.76-4.66 (3H, m), 3.10 (1H, ddd, *J* = 3.1, 6.0, 14.6 Hz), 2.56 (1H, dt, *J* = 6.6, 14.6 Hz), 2.43 (3H , s), 2.41 (3H, s); ^{13}C NMR (CD_2Cl_2, 125.8 MHz): δ= 166.5, 166.3, 145.3, 145.2, 137.2, 130.25, 129.98, 129.91, 129.82, 126.93, 126.86, 122.3, 90.9, 84.7, 74.5, 64.2, 41.3, 22.02, 21.98; UV-vis (EtOH) λ_{max}: 240, 315 nm; ESI-HRMS calc'd for $C_{24}H_{22}N_3O_7Br$ [$(M+Na^+)$] 566.0533, found 566.0520. Isomer **3**: R_f = 0.30 (1:2 EtOAc:hexanes). ^{1}H NMR (CD_2Cl_2, 500 MHz): δ = 7.96 (2H, d, *J* = 8.2 Hz), 7.88 (1H, s), 7.85 (2H, d, *J* = 7.9 Hz), 7.28 (4H, dd, *J* = 7.9, 21.1 Hz), 6.25 (1H, dd, *J* = 6.1, 7.6 Hz), 5.68 (1H, m), 4.74-4.63 (3H, m), 2.96 (1H, ddd, *J* = 2.2, 5.9, 14.3 Hz), 2.61 (1H, m), 2.43 (3H, s), 2.41 (3H, s); ^{13}C NMR (CD_2Cl_2, 125.8 MHz): δ= 166.5, 166.3, 145.4, 145.2, 134.1, 130.2, 129.95, 129.92, 129.83, 127.0, 126.8, 103.6, 88.4, 84.3, 75.0, 64.2, 40.1, 22.02, 21.96; UV-vis (EtOH) λ_{max}: 241, 296 nm; ESI-HRMS calc'd for $C_{24}H_{22}N_3O_7Br$ [$(M+Na^+)$] 566.0533, found 566.0552.

5-Bromo-4-nitro-β-2'-deoxy-D-ribofuranose (4)

Guanidine hydrochloride (0.175 g, 1.83 mmol) was treated with 0.5 M NaOMe in MeOH (3.53 mL, 1.76 mmol) in 36 mL MeOH and the solution stirred for 15 min in an ice bath. **3** (0.400 g, 0.735 mmol) was added and the suspension stirred in an ice bath for 3.75 h. The reaction was

diluted with CH_2Cl_2 to form a 20% solution of MeOH in CH_2Cl_2 and then filtered through a bed of silica. The silica was washed with 100 mL of 20% MeOH in CH_2Cl_2, and the combined solutions evaporated. The mixture was adsorbed on to silica (1.5 g) by suspending in MeOH (5 mL) and then evaporating to dryness. The dry silica was suspended in CH_2Cl_2 and applied to a silica column prepared in CH_2Cl_2. Elution with a step gradient of 3-5% MeOH in CH_2Cl_2 afforded **4** as a white foam (0.204 g, 90%). R_f = 0.28 (10% MeOH:CH_2Cl_2). ^{1}H NMR (CD_3OD, 500 MHz): δ = 8.34 (1H, s, imidazolic H), 6.22 (1H, t, J = 6.1 Hz, 1'H), 4.50 (1H, m, 3'H), 4.03 (1H, ddd, J = 3.6, 3.6, 7.2 Hz, 4'H), 3.80 (1H, dd, J = 3.2, 12.1 Hz, 5'H) , 3.73 (1H, dd, J = 3.7, 12.2 Hz, 5'H), 2.54 (2H, m, 2'H); ^{13}C NMR (CD_3OD, 125.8 MHz): δ = 136.6, 105.4, 90.0, 89.2, 71.7, 62.4, 42.5; ESI-HRMS calc'd for $C_8H_{10}N_3O_5Br$ [(M+Na^+)] 329.9696, found 329.9698.

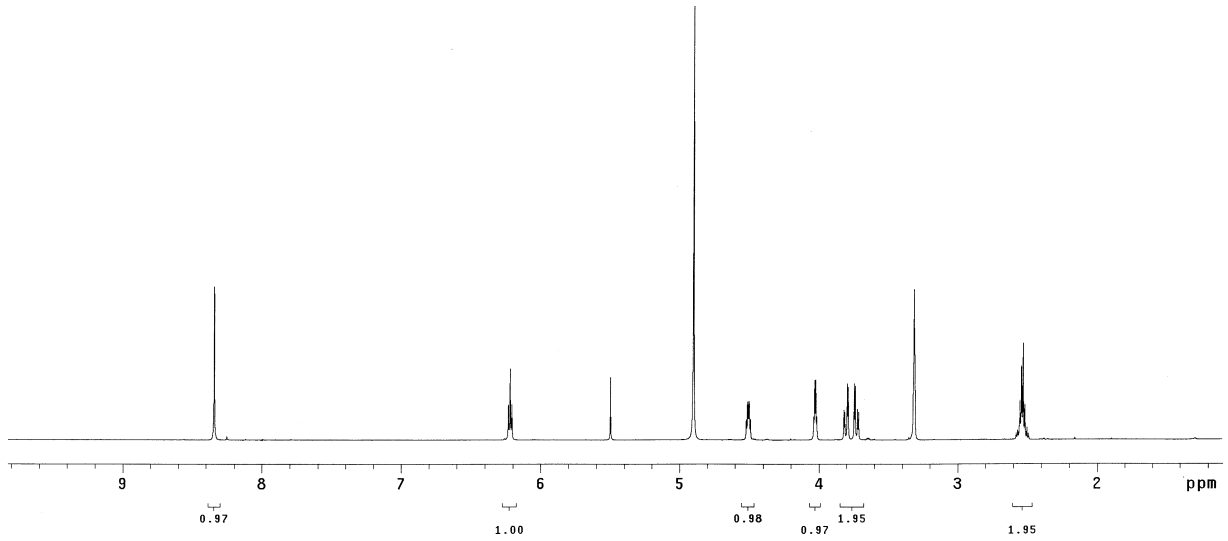

Figure S1. ^{1}H NMR spectrum of **4** in CD_3OD.

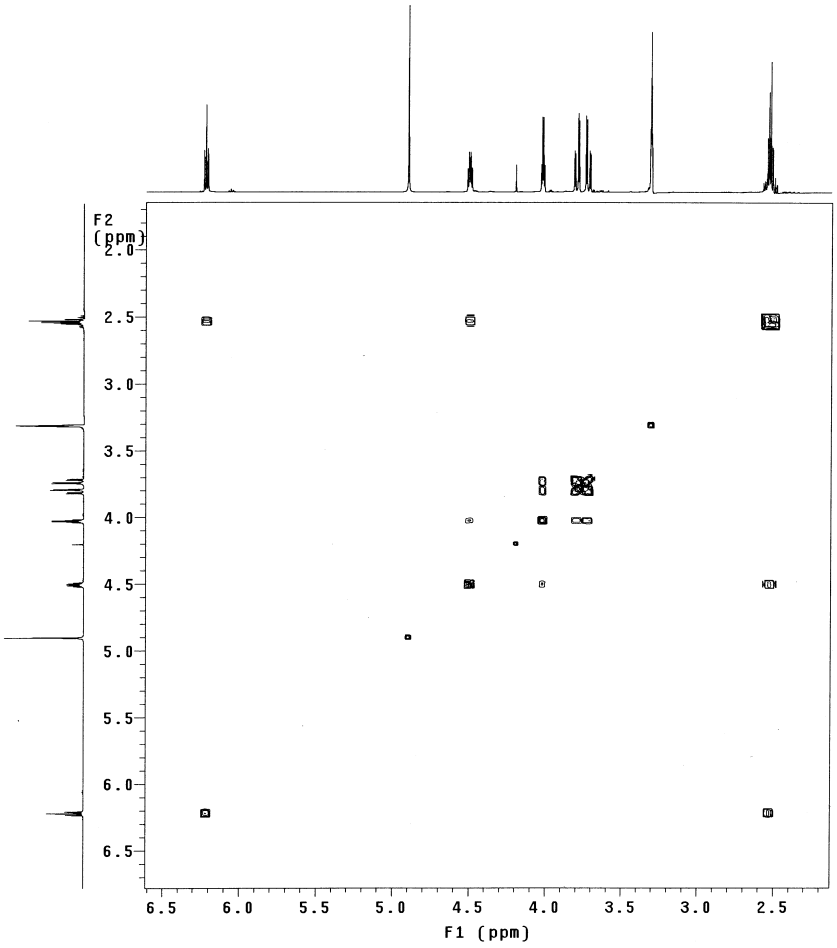

Figure S2. gCOSY spectrum of **4** in CD_3OD.

5-Guanidino-4-nitroimidazole-β-2'-deoxy-D-ribofuranose (5)

Guanidine hydrochloride (77.4 mg, 0.810 mmol) was dissolved in 1.62 mL of 0.5 M NaOMe in MeOH. The solution was stirred, evaporated to dryness, and redissolved in 5 mL of absolute EtOH. To this solution was added 50.0 mg (0.162 mmol) of **4**, and the reaction refluxed for 2 h. The solvent was subsequently evaporated and the residue adsorbed onto silica (0.5 g) by suspending in MeOH and then evaporating to dryness. The dried silica was suspended in $CHCl_3$ and charged to a silica column prepared in $CHCl_3$. Elution with 15-30% MeOH in $CHCl_3$ gave **5** as a bright yellow, very hygroscopic solid (45 mg, 98%, by UV-vis spectrophotometry). R_f = 0.24 (30% MeOH:CH_2Cl_2). ^{1}H NMR (D_2O, 500 MHz): δ = 7.73 (1H, s), 6.08 (1H, t, J = 6.7 Hz), 4.52 (1H, dt, J = 4.0, 6.3 Hz), 4.03 (1H, m), 3.77 (1H, dd, J = 3.7, 12.5 Hz), 3.69 (1H, dd, J = 5.2, 12.5 Hz), 2.60 (1H, m), 2.46 (1H, ddd, J = 4.3, 6.6, 14.1 Hz); ^{13}C NMR (CD_3OD, 125.8

MHz): δ= 160.0, 143.4, 135.0, 132.2, 89.1, 85.5, 72.4, 63.1, 41.8; UV-vis (H_2O, pH 7) λ_{max} (log ε): 230 nm (4.13), 380 nm (3.82); ESI-HRMS calc'd for $C_9H_{14}N_6O_5$ [(M+Na^+)] 309.0918, found 309.0917.

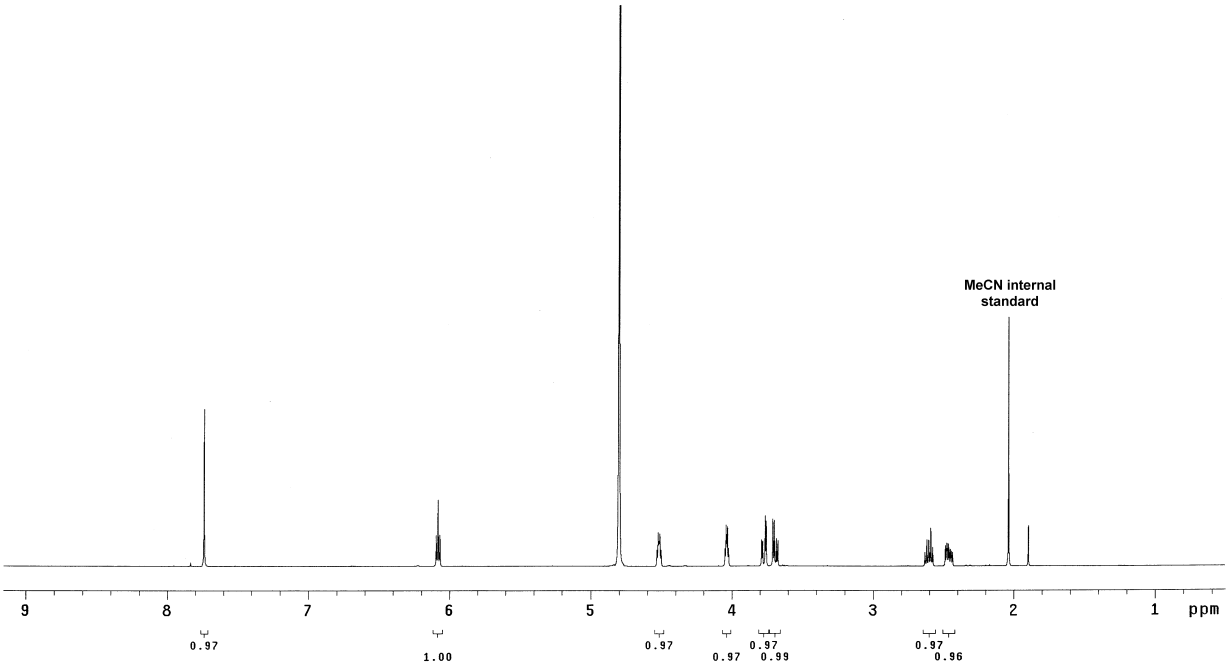

Figure S3. ^{1}H NMR spectrum of **5** in D_2O.

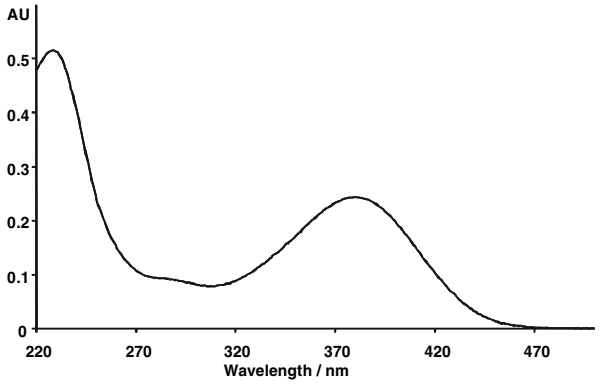

Figure S4. UV-vis spectrum of **5** in H_2O, pH 7.

5'-*O*-(4,4'-Dimethoxytrityl)- 5-bromo-4-nitroimidazole-β-2'-deoxy-D-ribofuranose (6)

Compound **4** (105 mg, 0.341 mmol) was dried by coevaporation twice with 2 mL of pyridine, once with 2 mL of toluene, and then by high-vacuum overnight. 4,4'-Dimethoxytritylchloride (DMT-Cl) (0.260 g, 0.767 mmol), 4-dimethylaminopyridine (DMAP) (4.2 mg, 10 mol %), and triethylamine (0.107 mL, 0.767 mmol) were combined and dissolved in pyridine (3.4 mL). The solution was stirred for 5 min and then added to a flask containing **4**. The reaction was stirred

for 3.5 h and then evaporated to dryness. The mixture was dissolved in CH_2Cl_2 and applied to a column of neutral alumina, Brockmann II, prepared in CH_2Cl_2 with several drops of pyridine. Elution with a step gradient of 0-0.5% MeOH in CH_2Cl_2 yielded 0.179 g of a yellow mixture. This mixture is unsuitable for the preparation of **7**; thus, 0.208 g of the mixture were flashed through a column of neutral alumina, Brockmann II, following the same procedure noted above. Compound **6** was isolated as an off-white foam (0.181 g, 76%). R_f = 0.46 (1:2 EtOAc:CH_2Cl_2). ^{1}H NMR (CD_2Cl_2, 500 MHz): δ = 7.79 (1H, s), 7.42-7.37 (2H, m), 7.33-7.21 (7H, m), 6.84 (4H, m), 6.13 (1H, t, *J* = 6.1 Hz), 4.56 (1H, br m), 4.15 (1H, ddd, *J* = 4.0, 4.0, 7.9 Hz), 3.78 (6H, s), 3.37 (2H, ddd, *J* = 4.5, 10.9, 15.4 Hz), 2.61 (1H, ddd, *J* = 4.4, 6.1, 13.6 Hz), 2.43 (2H, m); ^{13}C NMR (CD_2Cl_2, 125.8 MHz): δ= 159.2, 145.1, 136.0, 135.8, 134.6, 130.52, 130.50, 128.46, 128.45, 127.5, 113.69, 113.68, 103.7, 88.2, 87.6, 87.2, 71.6, 63.8, 55.7, 42.2; ESI-HRMS calc'd for $C_{29}H_{28}N_3O_7Br$ [(M+Na^+)] 632.1003, found 632.0995.

5'-*O*-(4,4'-Dimethoxytrityl)- 5-bromo-4-nitroimidazole-β-2'-deoxy-D-ribofuranose-3'-*O*-2-cyanoethyl-*N,N*-diisopropylphosphoramidite (7)

Compound **6** (50 mg, 0.082 mmol) was dried by coevaporation twice with 2 mL of pyridine, once with 2 mL of toluene, and then by high-vacuum overnight. To the flask containing **6** was added 0.8 mL of CH_2Cl_2 and diisopropylethylamine (0.043 mL, 0.25 mmol). 2-Cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.027 mL, 0.12 mmol) was added dropwise by syringe and the reaction was stirred for 45 min at room temperature. The reaction was poured into 20 mL of saturated aqueous $NaHCO_3$ and extracted 3 times with 20 mL of CH_2Cl_2. The organic layers were combined, dried over Na_2SO_4, and evaporated to an orange residue. The residue was applied to a column of basic alumina, Brockman II-III, prepared in CH_2Cl_2. Elution with CH_2Cl_2 gave **7** as a white solid (29 mg, 44%). R_f = 0.49 and 0.58 (two diastereomers) (10% EtOAc: CH_2Cl_2). ^{1}H NMR (CD_3CN, 500 MHz): δ = 7.88 (1H, s), 7.42-7.34 (2H, m), 7.32-7.19 (7H, m), 6.84 (4H, m), 6.12 (1H, dddd, *J* = 5.7, 5.7, 5.7, 11.2 Hz), 4.71 (1H, m), 4.18 (1H, d of br m, *J* = 21.2 Hz), 3.86-3.51 (10H, m including two s from –OCH_3 at 3.76 and 3.75 ppm), 3.37-3.20 (2H, m), 2.72-2.62 (3H, m), 2.53 (1H, t, *J* = 6.0 Hz), 1.19-1.04 (12H, m); ^{31}P NMR (CD_3CN, 121.5 MHz): δ= 150.2, 150.0 (two diastereomers); ESI-HRMS calc'd for $C_{38}H_{45}BrN_5O_8P$ [(M+Na^+)] 832.2081, found 832.2057.

NOE-Difference NMR Spectroscopy

The protons of **4** were assigned using a 2D gCOSY spectrum (Figure S2). Approximately 5 mg of **4** were dissolved in CD_3OD and transferred to an NMR tube fitted with a J. Young valve. The solution was degassed with four freeze-pump-thaw cycles, and then sealed. An NOE-difference experiment yielded the results in Table S1.

Table S1. NOE difference for **4** in CD_3OD (%)

Observed Proton	Irradiated Proton	
	1'H	Imidazolic H
Imidazolic H	1.41	—
1'H	—	1.73
2'H	6.43	3.27
3'H	0.28	2.65
4'H	2.20	0.49
5'H	0.10	1.47

Oligodeoxynucleotide Synthesis

ODNs were synthesized on an Applied Biosystems 392 DNA synthesizer using PAC-protected phosphoramidites and the reagents listed below (Table S2).

Table S2. DNA synthesis reagents

Instrument Function	Reagent
Activator	0.25 M 4,5-Dicyanoimidazole in CH_3CN
Cap A	5% Phenoxyacetic anhydride in THF/Pyridine
Cap B	10% 1-Methylimidazole in THF
Deblock	3% Dichloroacetic acid in CH_2Cl_2
Oxidizer	0.02 M I_2 in Pyridine/THF/H_2O

Syntheses were performed on a 1 μmol scale using 500-Å CPG supports and all coupling efficiencies, including that of the non-natural nucleoside, were greater than 95% as determined by trityl monitoring. The phosphoramidite **7** (16 mg) was dried by coevaporation twice with 2 mL of pyridine, once with 2 mL of toluene, and then by high-vacuum overnight. **7** was manually coupled by dissolving in 150 μL activator (0.25 M DCI in CH_3CN) and then pipetting directly on the CPG support. Coupling proceeded for 20 minutes.

Following DNA synthesis, ODNs containing 5-bromo-4-nitroimidazole were converted to ODNs containing NI by treating with 1 mL of a 0.5 M solution of guanidine in THF for 30

min at room temperature (**8NI**) or with 1 mL of a 0.5 M solution of guanidine in *tert*-butanol for 15 h at 55 °C (**9NI**). The guanidine solutions were prepared by treating 48 mg (0.5 mmol) of guanidine hydrochloride with 1 mL (0.5 mmol) of 0.5 M NaOMe in MeOH for 5 min, evaporating the MeOH, and redissolving in 1 mL of solvent. **8NI** and **9NI** were deprotected and cleaved from the solid support by treating with concentrated NH_4OH at room temperature for 4 hours and 2 hours, respectively. In the case of **8NI**, the guanidine was first washed from the solid support with 10 mL of THF prior to treating with NH_4OH. After treating **9NI** with NH_4OH, the solution was neutralized with triethylammonium acetate and acetic acid and then desalted using a NAP-10 column (Sephadex G-25, Amersham Biosciences).

ODNs containing NI were purified by reversed-phase HPLC on a Varian Microsorb-MV analytical C-18 column (4.6 mm × 25 cm, 5 μm), using 0.1 M triethylammonium acetate and CH_3CN as solvents. The concentration of CH_3CN was increased from 0-7% in 0.7 min, then 7-15% in 30 min with a 1.0 mL/min flow rate.

ESI Mass Spectra

Mass spectra (Figures S5 and S6) were acquired by direct infusion on a Sciex Model API 365 triple stage mass spectrometer.

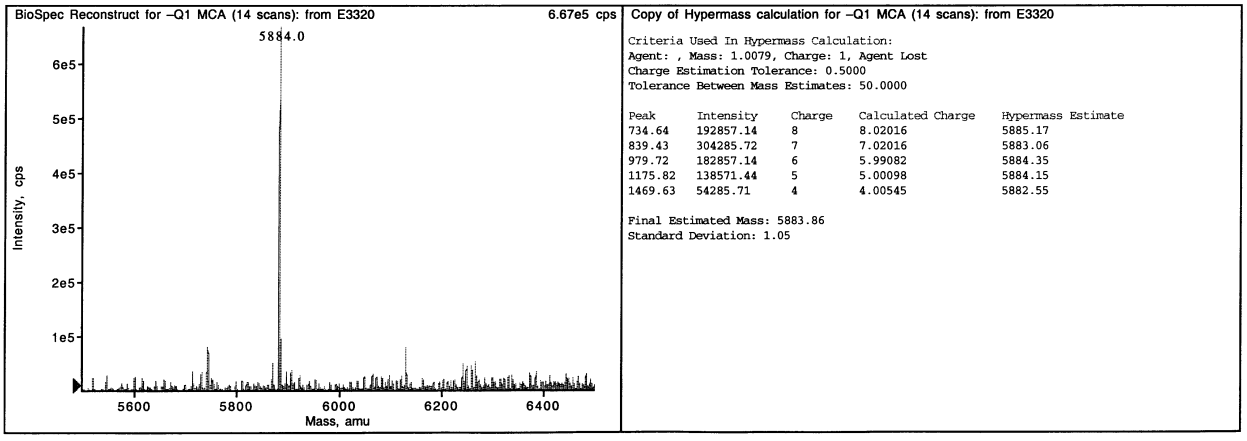

Figure S5. ESI mass spectrum for **8NI**: calc'd 5882.9, found 5884.0.

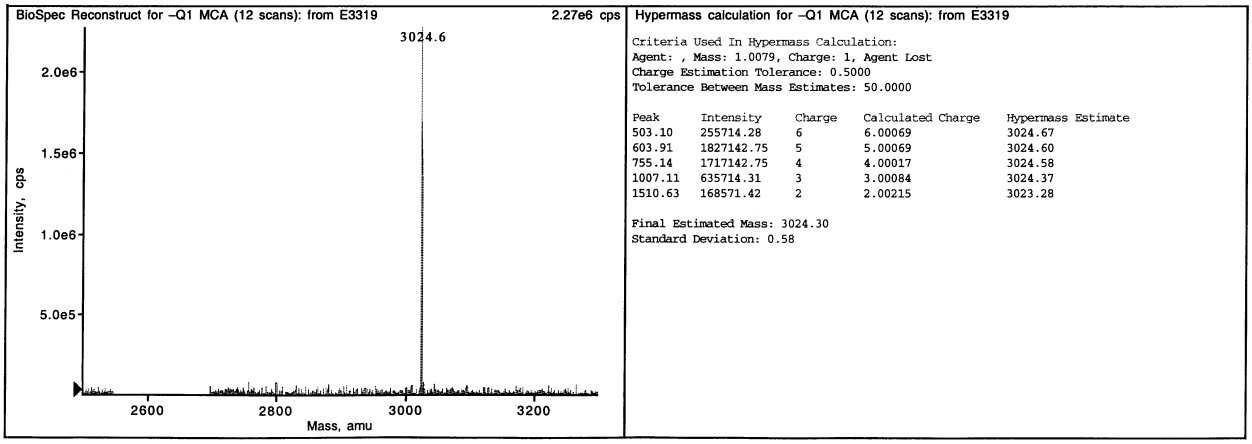

Figure S6. ESI mass spectrum for **9NI**: calc'd 3024.0, found 3024.6.

MALDI-TOF Mass Spectra

Mass spectra (Figures S7 and S8) were acquired using a matrix consisting of anthranilic acid, nicotinic acid, and diammonium citrate (2:1:0.003) in H_2O and CH_3CN as described by Zhang and Gross.[2] The following instrument parameters were used:

Table S3. MALDI-TOF Instrument Settings

Instrument Parameter	Setting
Mode of operation	linear
Polarity	negative
Accelerating voltage	25,000 volts
Grid voltage	95% of accelerating voltage
Extraction delay time	150 nsec
Laser wavelength	337 nm

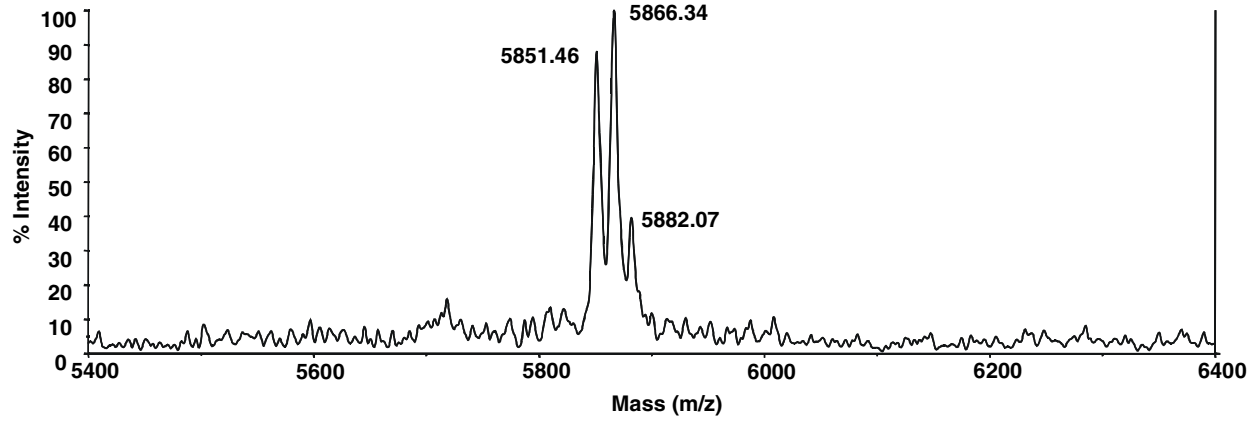

Figure S7. MALDI-TOF mass spectrum for **8NI**: calc'd 5882.9, found 5882.1.

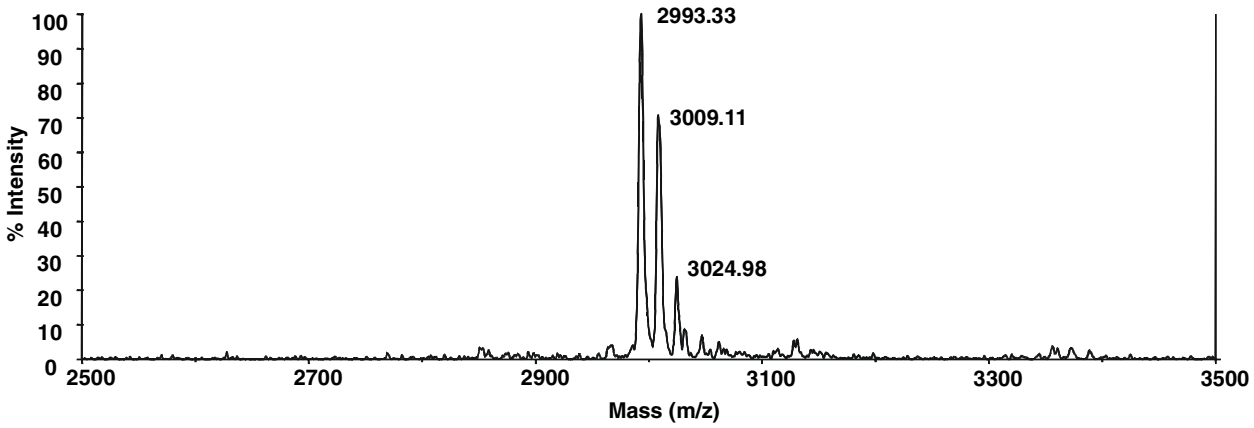

Figure S8. MALDI-TOF mass spectrum for **9NI**: calc'd 3024.0, found 3025.0.

Enzymatic Digestion Analyses

The purity and content of the ODNs **8NI** and **9NI** were established by HPLC analysis of the digested ODNs (Figure S9). The following conditions were used: 100 mM Tris-Cl, pH 8.8, 15 mM $MgCl_2$, 500 mU SVPD, and 2 nmol ODN in 40 μL were incubated at 37 °C for 4.5 h, then 20 U of alkaline phosphatase were added and the mixture incubated for an additional 2 h.

ODN **9NI** was digested with nuclease P1 using these conditions: 30 mM NaOAc, pH 5.3, 0.1 mM $ZnSO_4$, 17 U nuclease P1, and 2 nmol **9NI** in 50 μL were incubated at 37 °C for 2 h. Alkaline phosphatase (10 U, 10 μL) and 6.6 μL 10X buffer (500 mM Tris-Cl, pH 8.5, 1 mM EDTA) were added and the solution incubated for an additional 2 h (Figure S10).

Digestion mixtures were analyzed at 254 nm by reversed-phase HPLC with a Supelco Supercosil C-18 analytical column (2.1 mm × 25 cm, 5 μm), using 150 mM NH_4OAc and CH_3CN as eluting solvents. The concentration of CH_3CN was increased from 0-15% over 40 min at a flow rate of 0.25 mL/min.

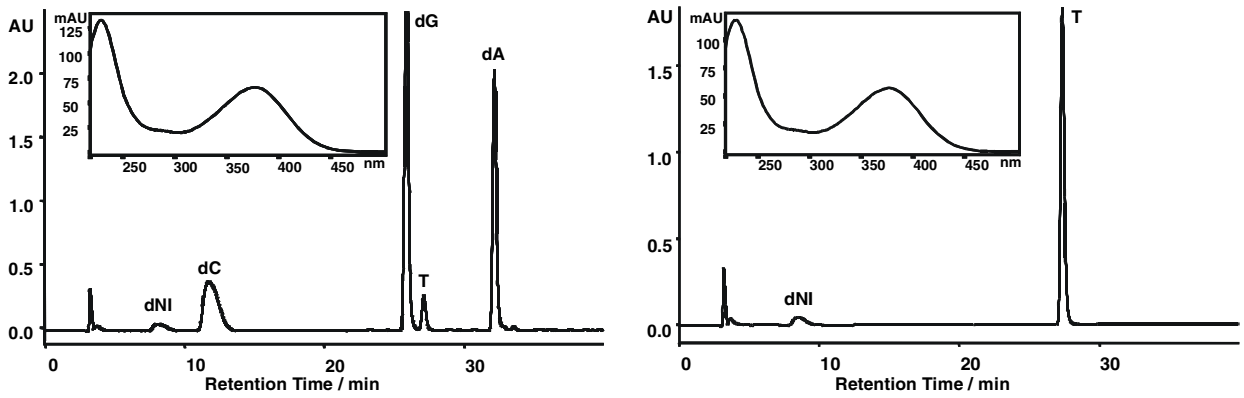

Figure S9. Digestion analyses of **8NI** (Left) and **9NI** (Right) using SVPD and alkaline phosphatase. The insets show the UV-vis spectra of the dNI peaks.

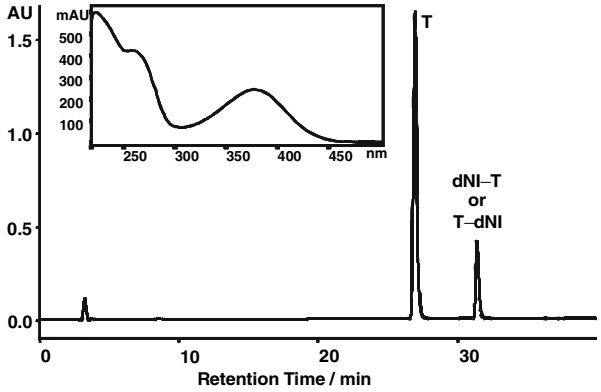

Figure S10. Digestion analysis of **9NI** using nuclease P1 and alkaline phosphatase. The inset shows the UV-vis spectrum of the dinucleotide peak.

Thermal Denaturation Studies

Melting temperatures were determined in teflon-stoppered 1 cm pathlength quartz cells. Absorbances were monitored at 260 nm and the temperature was cycled between 20 and 85 °C at a rate of 0.5 °C per min with a 5 min delay at each temperature extreme. T_m values and thermodynamic parameters were extracted from the melting curves (Figures S11 and S12) using the program MeltWin 3.5. Solutions consisted of 0.75 μM ODN in phosphate buffer (5 mM sodium phosphate, 50 mM NaCl, pH 7.0) in a total volume of 1.32 mL. All duplexes obeyed an apparent two state transition, and annealing curves matched melting curves.

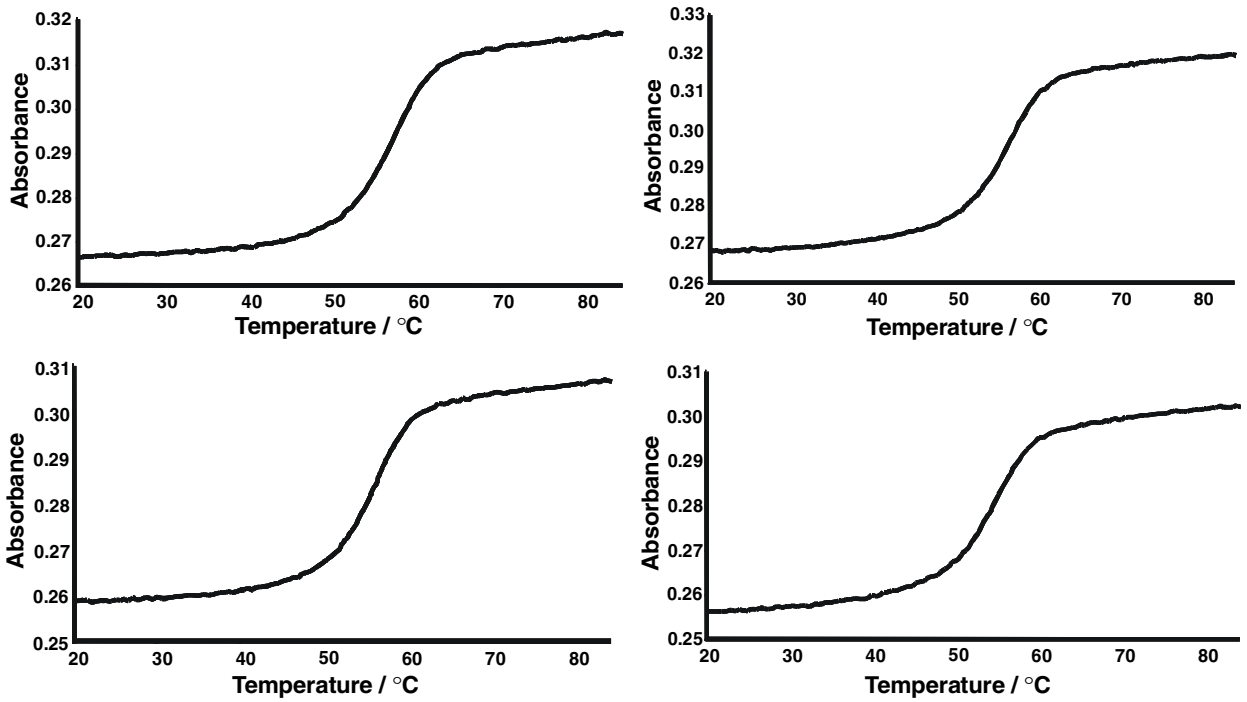

Figure S11. Representative melting curves for **8NI** annealed with **10G** (top left), **10A** (top right), **10C** (bottom left), and **10T** (bottom right).

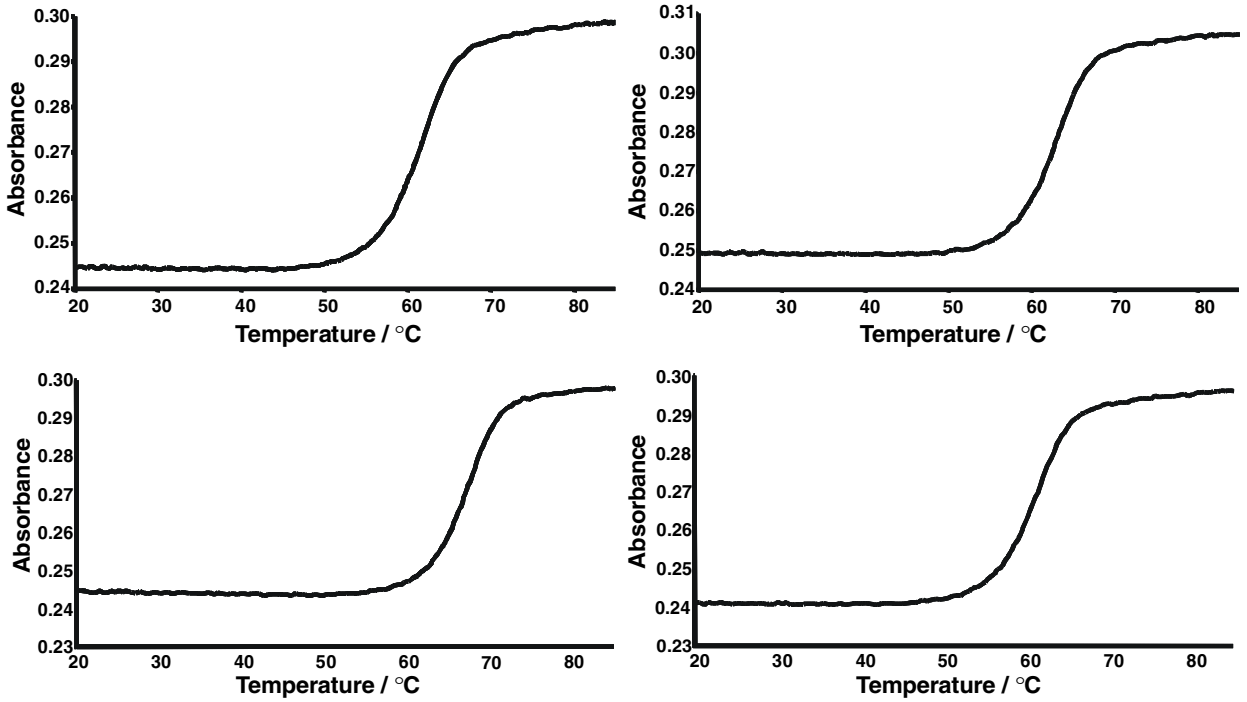

Figure S12. Representative melting curves for **8G** annealed with **10G** (top left), **10A** (top right), **10C** (bottom left), and **10T** (bottom right).

Semi-Empirical MO Calculations

Calculations were performed using Hyperchem 6.03 on a Pentium IV 1.3 GHz personal computer. The geometry of the molecule was first optimized using the MM+ molecular mechanics method (Hyperchem's version of MM2), and then optimized using the AM1 semi-empirical method without constraints. Several minima of similar energy in addition to that shown in Figure S13 were found (including syn conformations), and all depicted the guanidino group out of plane with respect to the imidazole ring. Geometries were optimized until a root-mean-square gradient of less than or equal to 0.01 kcal/(Å mol) was achieved.

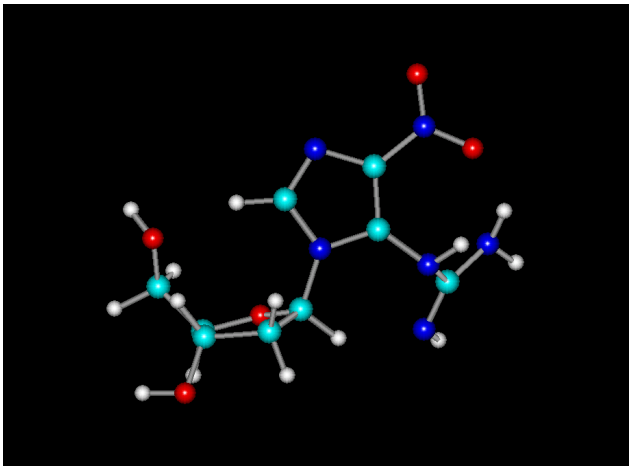

Figure S13. Energy-minimized (AM1) structure of dNI (**5**). The imidazole ring and nitro group are in the plane of the page, while the carbon and two of the nitrogen atoms of the guanidino group are projecting through the back of the page.

M13 Genome Constructions

ODNs **8NI** and **8G** were inserted into an M13mp7(L2) bacteriophage genome using a previously reported procedure (Figure S14).[3] The desalted genomes (0.3 μg) were run on a 0.9% agarose gel in 0.5X TBE buffer. The gel was run for 5 h at 100 V, and then stained in a 1X solution of SYBR Gold (Molecular Probes, Inc.) in 1X TBE buffer (200 mL). The ligation efficiency of ODN insertion into the genome was approximately 38% as determined by quantification of the stained DNA. Circular wild-type M13 DNA and linear M13 DNA (linearized by digestion with *Eco*RI restriction endonuclease) were used as markers.

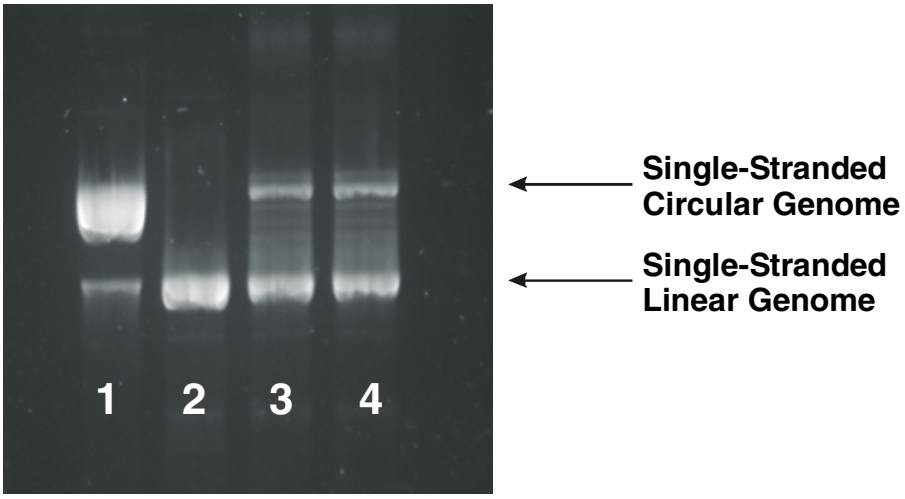

Figure S14. Agarose gel analysis of the M13 genome constructions. (1) Circular M13 DNA. (2) Linear M13 DNA. (3) Insertion of **8G** into M13 DNA. (4) Insertion of **8NI** into M13 DNA.

Reference List

1. Del Carmen, M.; Barrio, G.; Scopes, D. I. C.; Holtwick, J. B.; Leonard, N. J. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 3986-3988.

2. Zhang, L. K.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 854-865.

3. Henderson, P. T.; Delaney, J. C.; Muller, J. G.; Neeley, W. L.; Tannenbaum, S. R.; Burrows, C. J.; Essigmann, J. M. *Biochemistry* **2003**, *42*, 9257-9262.