Electronic Supplementary Material (ESI) for Nanoscale.


**SUPPLEMENTARY INFORMATION**

# RNA intrusions change DNA elastic properties and structure

Hsiang-Chih Chiu[a,†], Kyung Duk Koh[b,†], Marina Evich[c], Annie L. Lesiak[d], Markus W. Germann[c,*], Angelo Bongiorno[d,*], Elisa Riedo[a,*] and Francesca Storici[b,*]

[a]School of Physics, Georgia Institute of Technology, Atlanta, Georgia, 30332, U.S.A.; [b]School of Biology, Georgia Institute of Technology, Atlanta, Georgia, 30332, U.S.A.; [c]Department of Chemistry, Georgia State University, Atlanta, Georgia, 30302, U.S.A; [d]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia, 30332, U.S.A.

[†]Equal contribution
[*]To whom correspondence should be addressed:
mwg@gsu.edu, angelo.bongiorno@chemistry.gatech.edu, elisa.riedo@physics.gatech.edu, storici@gatech.edu

## SUPPLEMENTARY MATERIALS AND METHODS

### 1. UV melting and circular dichroism for AFM substrates

Duplexes used in AFM measurements were analyzed by UV melting and circular dichroism (CD). The samples were prepared by annealing appropriate complementary oligos at concentrations of 1.38 μM in 100 mM NaCl, 10 mM phosphate, and 0.1 mM EDTA at pH 7.4. UV absorption changes at 260 nm were acquired by Cary 1E UV-Vis Spectrophotometer, between 25 °C to 91 °C. The temperature was raised at a rate of 0.5 °C/min. Absorbance values were normalized at the value at 25 °C. Only the $T_m$ was calculated from the UV melting curves since the substrates were all 30 bp long.

CD spectra were acquired on Jasco J-810 Circular Dichroism Spectrometer between 210 nm to 320 nm, with a scanning rate of 200 nm/min and a band width of 1 nm. The samples were placed in 1-cm path length cells at 25 °C.

### 2. Gold surface used in AFM experiments

The template stripped gold substrates were purchased from Platypus Technologies (http://www.platypustech.com/templatestrippedgold.html, Madison, WI). The thickness of gold coating is 100 nm. The roughness of gold surface is 0.4 nm within 1x1 μm area. The AFM image of a freshly stripped gold surface is shown in Supplementary Figure S3a.

### 3. Calibration of AFM cantilevers

The spring constants of all used cantilevers (Novascan, Ames, IA) are calibrated using the "reference beam method" with the same reference cantilever[1, 2]. This is to minimize the possible experimental error from the calibration procedure propagating into the final results. A reference cantilever with a known spring constant $k_{ref}$ = 0.08 N/m was used. As shown in Supplementary Figure S1, a cantilever with unknown spring constant is pressed onto the reference cantilever. Then, the unknown spring constant can be determined by

$$k = k_{ref}\left(\frac{S_{hard}}{S_{ref}} - 1\right)\left(\frac{L_{ref}}{L_{ref} - \Delta L}\right)^3$$

where $S_{ref}$ and $S_{hard}$ are the slopes of force curves when the tip is in contact with the reference cantilever and a hard surface such as silicon, respectively. $L_{ref}$ is the length of the reference cantilever, and $\Delta L$ is the offset between the AFM tips due to possible misalignment. For the triangular cantilevers used in this work, extra care was needed to position the AFM tip near the middle line of the reference cantilever to avoid errors owing to the torsional bending[3]. The $L_{ref}$ of our reference cantilever is purposely chosen to be 480 μm; thus the error coming from the tip alignment is negligible. Furthermore, to properly calibrate the unknown stiffness $k$ of a cantilever using this method, the following condition needs to be satisfied:

$$0.3k_{ref} < k < 3k_{ref}$$

According to the manufacturer's data, the spring constants of our cantilevers are approximately 0.06 N/m, which satisfies the condition above. The measured force can be calculated by $F = k \cdot$

$m \cdot D_{\mathrm{lever}}$, where $m$ is the optical sensitivity of the AFM cantilever and $D_{\mathrm{lever}}$ is the deflection signal of the cantilever recorded by the photodiodes of AFM. The optical lever sensitivities ($m$) are calibrated in solution on gold substrate during the force measurements. All calibrated $w$ and $k$ of the used cantilevers are listed in Supplementary Table S2.

**4. Selection and calibration of AFM force-distance curves**

We only consider force-distance curve showing only one DNA pick-up by the tip. We do not use data showing multiple DNA pick-ups. Typical force curves showing multiple DNA pick-ups are shown in Supplementary Figure S5.

During the data acquisition, the force-distance curve is recorded as cantilever deflection vs. distance moved by the piezo scanner, i.e. $\Delta z_{\mathrm{lever}}$ vs. $z_{\mathrm{piezo}}$, as shown in Supplementary Figure S6a. This is not the real separation distance, $d$, between the AFM tip apex and the gold surface. To accurately determine the tip-surface distance, $d$, the cantilever deflection $\Delta z_{\mathrm{lever}}$ has to be subtracted from $z_{\mathrm{piezo}}$, i.e., $d = z_{\mathrm{piezo}} - \Delta z_{\mathrm{lever}}$[4]. In addition, the cantilever deflection is assumed to be zero when the tip is far away from the surface and is used to offset the whole force curve. The calibrated force-distance curve is shown in Supplementary Figure S6b, which is identical to Figure 2d in the main text.

**5. Determination of $L_0$, $\delta$, and $F_{\mathrm{st}}$**

The calibrated force-distance curve during tip retracting shown in Supplementary Figure S6b is now presented in Supplementary Figure S7. As indicated by two red solid lines, we performed linear fitting to two linear sections of the retracting curve to determine the position when the DNA begins to be stretched. The intersected point of the fitted lines is determined to be the position where the stretching of DNA begins. The sudden jump in the force curve indicates when the DNA is suddenly detached from tip, i.e. when the bonding between streptavidin and biotin is broken. Then, the cantilever goes back to its initial position when there is no force, and the deflection of becomes zero again. The initial contour length $L_0$, extension $\delta$ of DNA, and the stretching force $F_{\mathrm{st}}$ can then be determined from the force curve and are used to calculate the stretch modulus by using $S = F_{\mathrm{st}} \cdot L_0 / \delta$.

**6. Removal of outliers**

We used Peirce's criterion to perform the outlier test[5-7]. Using Peirce's criterion, multiple outliers can be removed. Supplementary Table S7 summarizes the number of outliers that was excluded to obtain the final data sets for statistical consideration. Only as much as 6% of total data were removed.

**7. Gaussian fitting**

The best Gaussian fit to the data was obtained using the following equation:

$$y = y_0 + Ae^{-\frac{(x-x_c)^2}{2w^2}}$$

where the offset $y_0$ is set to 0 during the fitting procedure. The value of $x_c$ is the peak position of the Gaussian distribution.

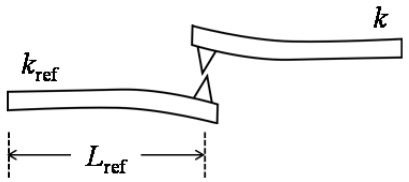
**8. UV melting for NMR substrates**

Thermal denaturation curves for an rGMP-containing 9-bp duplex, ATGGArGCTC (with rGMP III) and its DNA-control were obtained on a Cary UV-Vis Spectrophotometer. The duplexes were prepared in 100 mM NaCl, 10 mM phosphate, 0.5 mM EDTA at pH 6.6. Melting temperatures ($T_m$) were derived from a six-parameter fit of the melting curves for a series of duplex concentrations ranging from 2 to 20 μM[8]. Enthalpy and entropy values were then calculated from a linear fit of the van't Hoff plot.

## SUPPLEMENTARY TABLES AND FIGURES

| Name | Sequence |
|---|---|
| *Sequence 1* dG | 5'-Bi-CAGGTTCACGATGGAGCTCTCGATTCAGCT-SH-3' |
| *Sequence 1* rG | 5'-Bi-CAGrGTTCACrGATGGArGCTCTCrGATTCArGCT-SH-3' |
| *Sequence 1* compl_DNA | 5'-SH-AGCTGAATCGAGAGCTCCATCGTGAACCTG-Bi-3' |
| *Sequence 2* dG | 5'-Bi-ATCCGGTAGTGTTAGGCCTGAACAAGGTTT-SH-3' |
| *Sequence 2* rG | 5'-Bi-ATCCrGGTAGTrGTTAGrGCCTrGAACAArGGTTT-SH-3' |
| *Sequence 2* compl_DNA | 5'-SH-AAACCTTGTTCAGGCCTAACACTACCGGAT-Bi-3' |
| dG_III | 5'-ATGGAGCTC-3' |
| rG_III | 5'-ATGGArGCTC-3' |
| compl_DNA_III | 5'-GAGCTCCAT-3' |
| dG_VI | 5'-ATCCGGTAG-3' |
| rG_VI | 5'-ATCCrGGTAG-3' |
| compl_DNA_VI | 5'-CTACCGGAT-3' |
| dG_VIII | 5'-TTAGGCCTG-3' |
| rG_VIII | 5'-TTAGrGCCTG-3' |
| compl_DNA_VIII | 5'-CAGGCCTAA-3' |

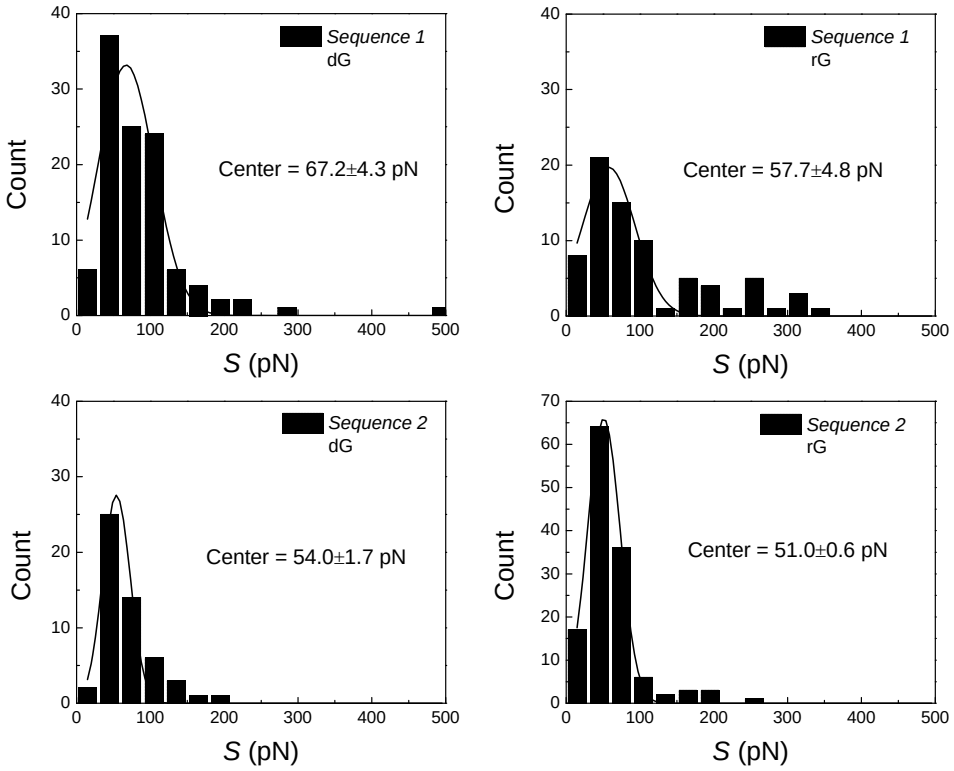
**Supplementary Table S1.** Sequences of synthetic oligonucleotides used in this study. Biotin and thiol group modifications are indicated by "Bi" and "SH," respectively. dNMPs are shown in blue while rNMPs are shown in red, preceded by letter "r." 30-mers were used in AFM experiments while 9-mers were used in NMR experiments. All oligonucleotides used in AFM experiments were PAGE-purified.




**Supplementary Figure S1.** Schematic of the reference beam method for the calibration of cantilever spring constant.

| Tips used for | Samples | *w* (nm/V) | *k* (N/m) |
|---|---|---|---|
| *Sequence 1* | dG-DNA Round 1 | 59.9±1.0 | 0.068±0.002 |
| | rG-DNA Round 1 | 62.3±1.0 | 0.057±0.002 |
| | dG-DNA Round 2 / rG-DNA Round 2 (Measured with the same tip) | 50.0±1.2 | 0.053±0.002 |
| | rG-DNA Round 3 | 53.5±1.7 | 0.415±0.002 |
| *Sequence 2* | dG-DNA Round 1 | 53.1±1.3 | 0.045±0.002 |
| | rG-DNA Round 1 | 48.0±1.0 | 0.051±0.004 |
| | dG-DNA Round 2 / rG-DNA Round 2 (Measured with the same tip) | 50.3±1.0 | 0.049±0.002 |

**Supplementary Table S2.** List of the optical lever sensitivity ($w$) and spring constant ($k_N$) of cantilevers used in the measurements.




**Supplementary Figure S2.** Histograms of stretch moduli of ss substrates with *Sequence 1* and *Sequence 2*. Peak position is presented as the fitted value ± standard error of the fit.

| Substrate | Mean | | | |
|---|---|---|---|---|
| | $L_0$ (nm) | $\delta$ (nm) | $F_{st}$ (pN) | $S$ (pN) |
| dG | 9.7 ± 3.6 | 3.0 ± 1.9 | 20.0 ± 11.1 | 84.9 ± 61.1 |
| rG | 13.7 ± 4.5 | 4.3 ± 2.0 | 26.2 ± 12.1 | 106.0 ± 84.7 |

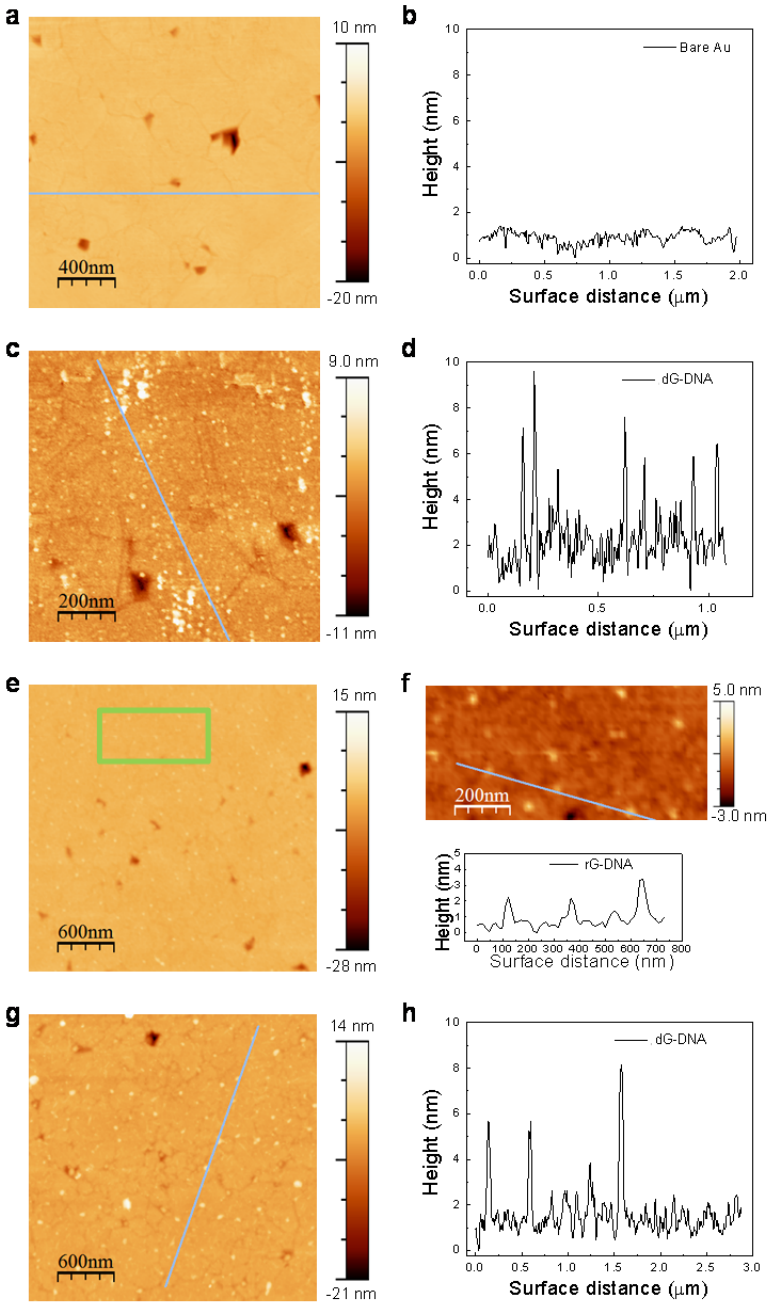
**Supplementary Table S3.** Mean values of all the parameters and stretch modulus of ss substrates with *Sequence 1*. Mean values are presented with standard deviation of the mean.

| Substrate | Mean | | | |
|---|---|---|---|---|
| | $L_0$ (nm) | $\delta$ (nm) | $F_{st}$ (pN) | $S$ (pN) |
| dG | 12.8 ± 3.4 | 5.2 ± 2.2 | 25.8 ± 9.3 | 69.7 ± 33.3 |
| rG | 12.0 ± 3.0 | 5.2 ± 1.9 | 22.7 ± 8.6 | 60.4 ± 38.8 |

**Supplementary Table S4.** Mean values of all the parameters and stretch modulus of ss substrates with *Sequence 2*. Mean values are presented with standard deviation of the mean.

| Substrate | Gaussian Peak | | | | Median |
|---|---|---|---|---|---|
| | $L_0$ (nm) | $\delta$ (nm) | $F_{st}$ (pN) | $S$ (pN) | $S$ (pN) |
| *Sequence 1* dG (n=108) | 9.1±0.2 | 2.2±0.1 | 20.2±0.1 | 67.2 ± 4.3 | 68.1 (59.4 – 86.9) |
| *Sequence 1* rG (n=75) | 13.5±0.7 | 3.7±0.2 | 23.4±0.1 | 57.7 ± 4.8 | 72.4 (55.9 – 102.3) |
| *Sequence 2* dG (n=52) | 12.0±0.2 | 5.0±0.2 | 24.6±0.1 | 54.0 ± 1.7 | 59.0 (51.6 – 73.5) |
| *Sequence 2* rG (n=132) | 12.1±0.2 | 5.1±0.2 | 23.0±0.2 | 51.0 ± 0.6 | 50.2 (44.4 – 60.1) |

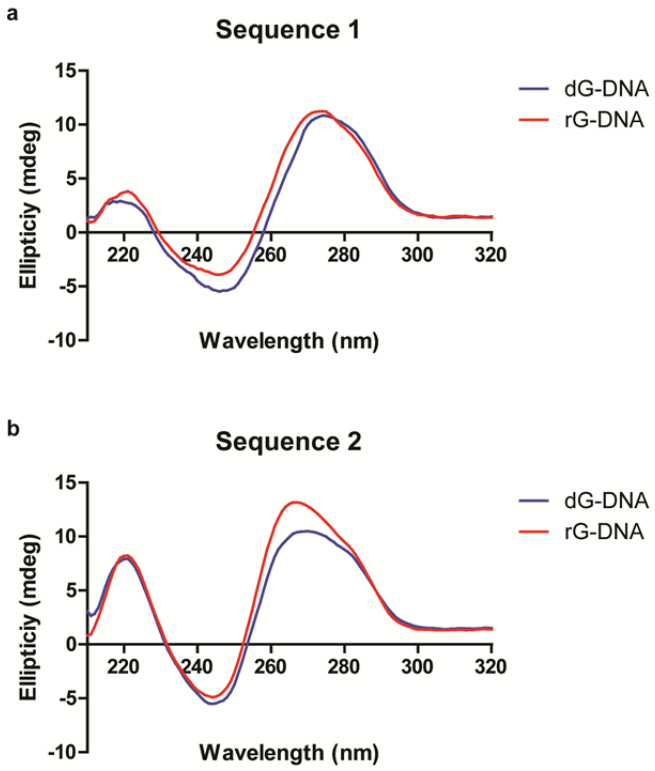
**Supplementary Table S5.** Comparison of Gaussian peak values and median values of stretch modulus of ss substrates. Gaussian peak values of all the parameters are also listed. Gaussian fitted values are presented with standard error of fit while median values are presented with 99% confidence interval of the median.



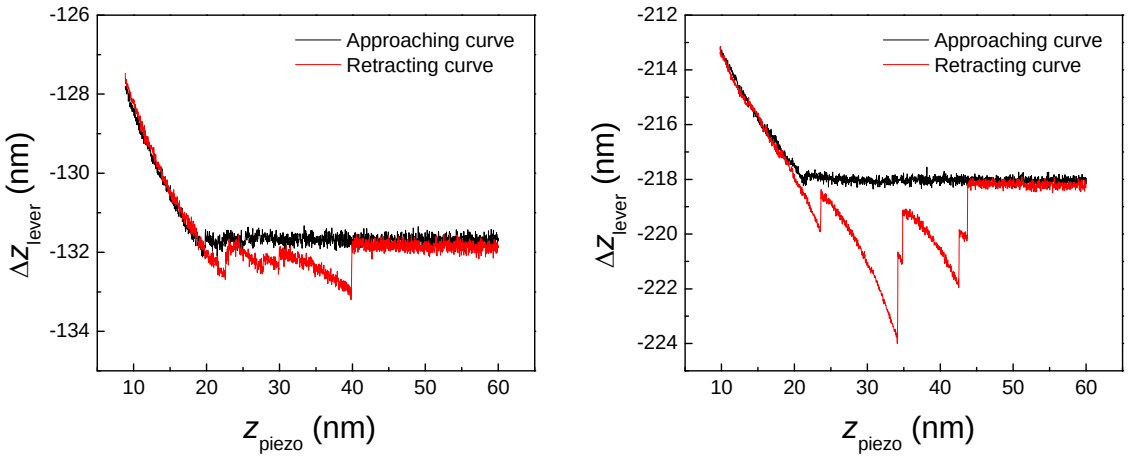

**Supplementary Figure S3. a**, AFM image of a fresh gold surface. **b**, Surface profile of the gold surface indicated by the blue line in **a**. **c**, A typical image of dG-DNA molecules of *Sequence 1* attached on gold surface in liquid. **d**, Height profile of the blue line indicated in **c** shows that DNAs are also standing up on the surface. **e**, A typical image of rG-DNA molecules of *Sequence 2* attached on gold surface in air. **f**, Top: Zoom-in image in **e** indicated by the green box; Bottom: Height profile of the blue line indicated in the top panel. **g**, A typical image of dG-DNA molecules of *Sequence 2* attached on gold surface in air. **h**, Height profile of the blue line indicated in **g**.

| | $T_m$ (°C) |
|---|---|
| *Sequence 1* dG-DNA | 73.6 |
| *Sequence 1* rG-DNA | 70.2 |
| *Sequence 2* dG-DNA | 69.6 |
| *Sequence 2* rG-DNA | 66.2 |

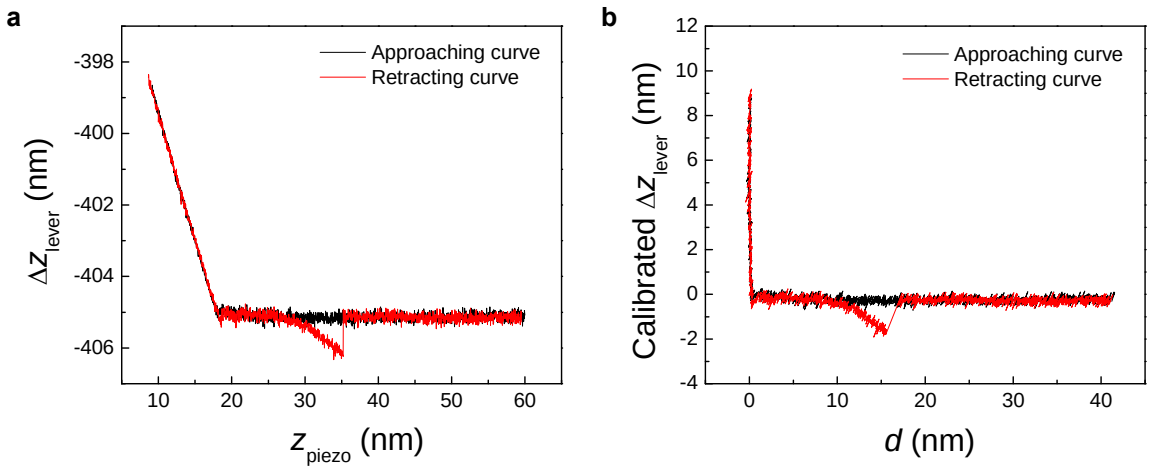
**Supplementary Table S6.** Thermal stability of 30-bp duplexes used in AFM experiments.



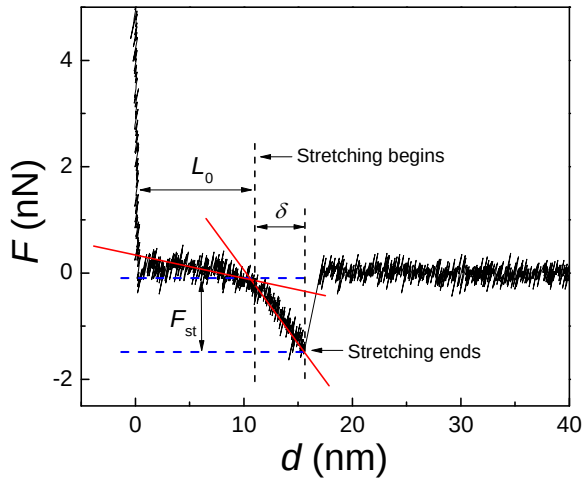

**Supplementary Figure S4.** CD spectra of dG-DNA and rG-DNA used in AFM experiments. **a**, CD spectra of duplexes with *Sequence 1*; **b**, CD spectra of duplexes with *Sequence 2*.




**Supplementary Figure S5.** Typical force-distance curves when the AFM tip picks up multiple DNAs.



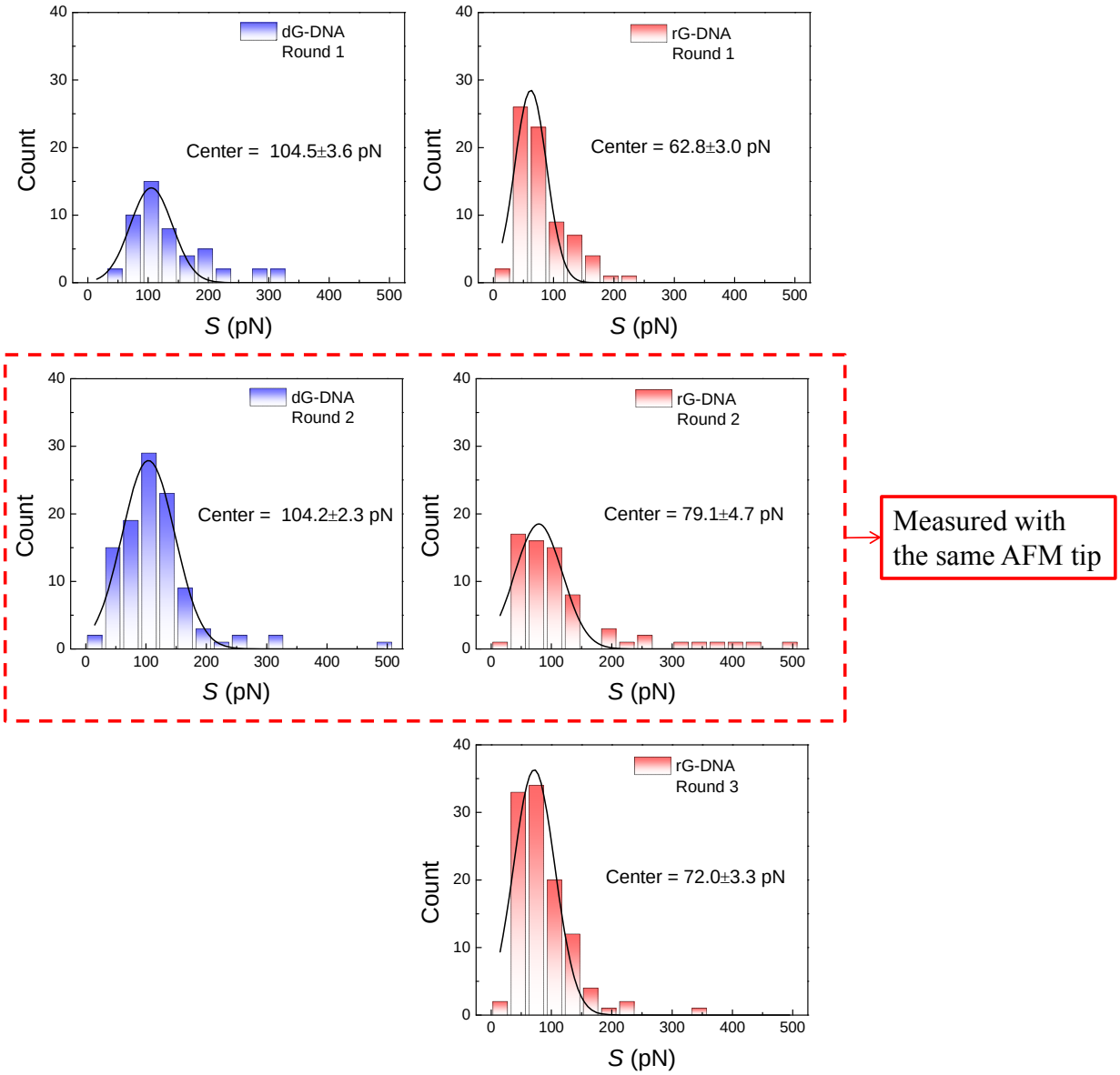

**Supplementary Figure S6.** Force-distance curves: **a**, before and **b**, after the calibration procedure.



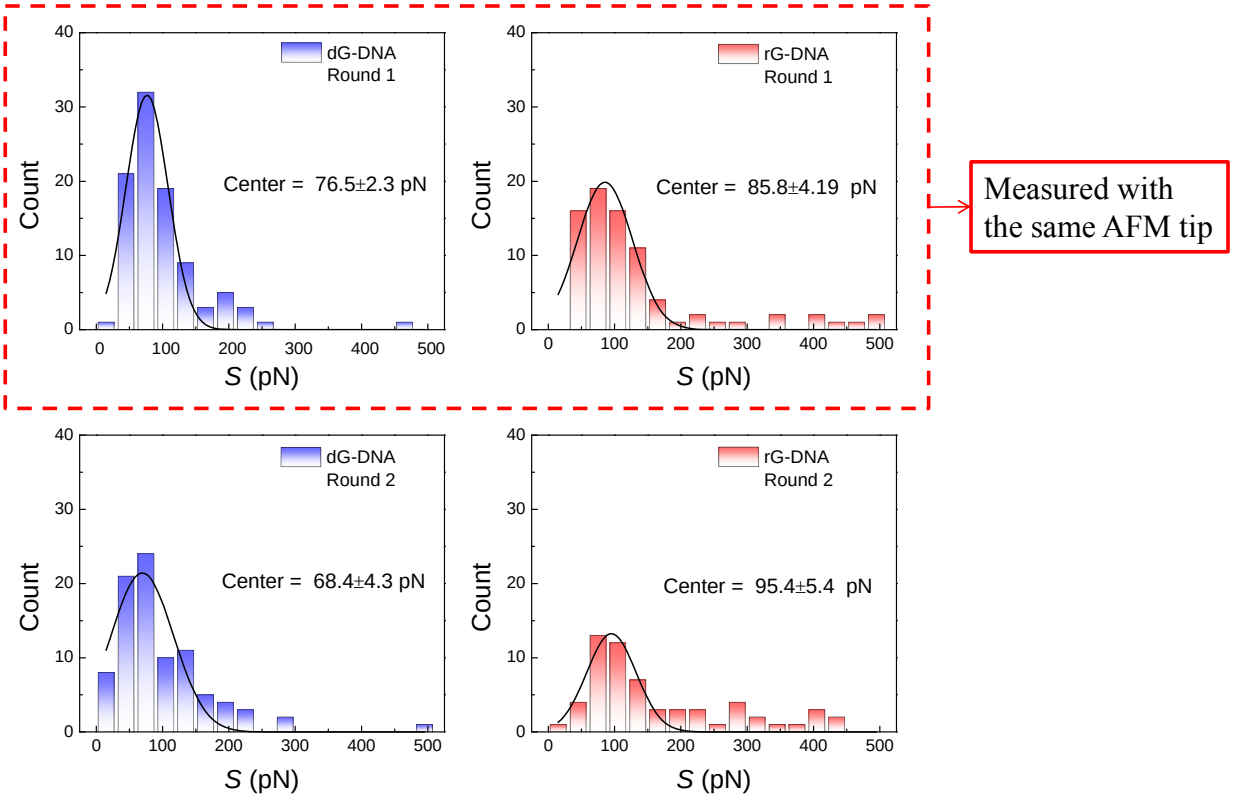

**Supplementary Figure S7.** The procedure to determine the contour length $L_0$, extension $\delta$, and $F_{st}$ during the stretching measurement.

| Substrate | | Data population before the removal of outliers | Number of outliers found | Stretch modulus removed as outliers (pN) |
|---|---|---|---|---|
| Double-stranded (ds) | *Sequence 1* dG-DNA Round 1 | 52 | 2 | 437.0, 455.6 |
| | *Sequence 1* dG-DNA Round 2 | 112 | 5 | 2086.7,1943.5, 903.2, 955.9, 1391.1 |
| | *Sequence 1* rG-DNA Round 1 | 76 | 3 | 416.7, 500.63, 407.9 |
| | *Sequence 1* rG-DNA Round 2 | 72 | 3 | 551.6, 830.2, 615.2 |
| | *Sequence 1* rG-DNA Round 3 | 112 | 3 | 1243.3, 709.4, 761.8 |
| | *Sequence 2* dG-DNA Round 1 | 99 | 4 | 590.6, 786.5, 1068.0, 1035.9 |
| | *Sequence 2* dG-DNA Round 2 | 94 | 3 | 1037.6, 1679.3, 2258.0 |
| | *Sequence 2* rG-DNA Round 1 | 84 | 5 | 662.1, 550.1, 989.5, 600.2, 697.8 |
| | *Sequence 2* rG-DNA Round 2 | 64 | 4 | 537.9, 520.1, 552.4, 647.8 |
| Single-stranded (ss) | *Sequence 1* dG | 113 | 5 | 662.1, 550.1, 898.5, 600.2, 697.8 |
| | *Sequence 1* rG | 78 | 3 | 399.8, 384.2, 408.4 |
| | *Sequence 2* dG | 55 | 3 | 352.2, 457.6, 338.2 |
| | *Sequence 2* rG | 140 | 8 | 444.7, 491.9, 432.1, 406.6, 297.4, 298.2, 295.5, 280.2 |

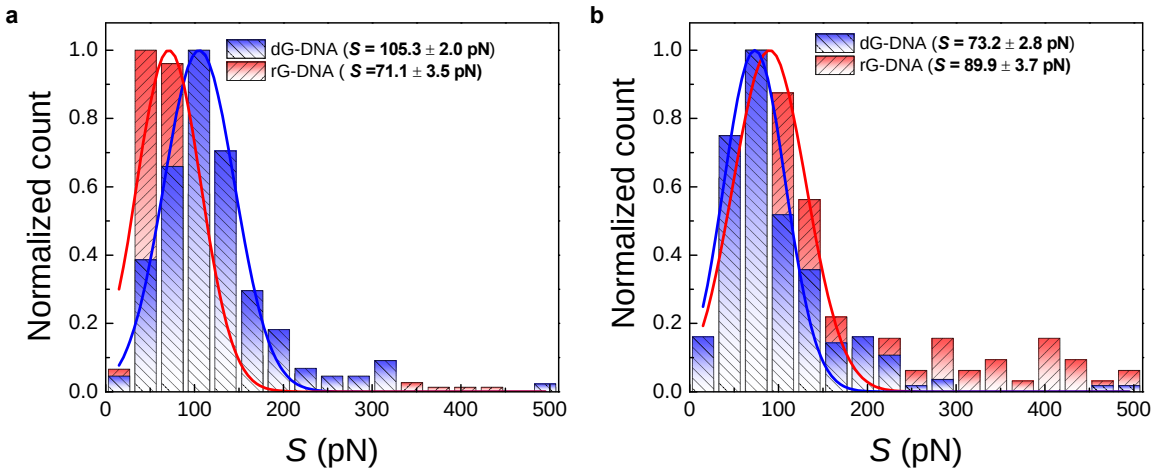
**Supplementary Table S7.** Summary of the number of data population before the removal of outliers and the values of stretch modulus that are considered as outliers and removed. See Section 6 of Supplementary Materials And Methods for details of the outlier test.




**Supplementary Figure S8.** Histograms of stretch moduli of ds substrates with *Sequence 1*. Peak position is presented as the fitted value ± standard error of the fit.




**Supplementary Figure S9.** Histograms of stretch moduli of ds substrates with *Sequence 2*. Peak position is presented as the fitted value ± standard error of the fit.




**Supplementary Figure S10.** Gaussian fitting for all combined data. **a**, *Sequence 1*. **b**, *Sequence 2*.

| Substrate | Mean | | | |
|---|---|---|---|---|
| | $L_0$ (nm) | $\delta$ (nm) | $F_{st}$ (pN) | $S$ (pN) |
| dG-DNA (n=157) | 10.0 ± 3.0 | 3.0 ± 1.8 | 34.2 ± 21.3 | 125.6 ± 83.3 |
| rG-DNA (n=251) | 11.4 ± 3.7 | 4.2 ± 2.1 | 30.9 ± 14.2 | 94.4 ± 67.7 |

**Supplementary Table S8.** Mean values of all the parameters and stretch modulus of ds substrates with *Sequence 1*. Mean values are presented with standard deviation of the mean.

| Substrate | Mean | | | |
|---|---|---|---|---|
| | $L_0$ (nm) | $\delta$ (nm) | $F_{st}$ (pN) | $S$ (pN) |
| dG-DNA (n=186) | 10.5 ± 4.0 | 3.3 ± 1.6 | 28.9 ± 20.1 | 106.4 ± 97.0 |
| rG-DNA (n=139) | 11.4 ± 3.9 | 3.2 ± 2.1 | 31.2 ± 14.5 | 147.5 ± 112.1 |

**Supplementary Table S9.** Mean values of all the parameters and stretch modulus of ds substrates with *Sequence 2*. Mean values are presented with standard deviation of the mean.

| Substrate | *S* (pN) | |
|---|---|---|
| | Gaussian Peak | Median |
| *Sequence 1* dG-DNA Round 1 | 104.5 ± 3.6 | 111.6 (99.3 – 147.4) |
| *Sequence 1* dG-DNA Round 2 | 104.2 ± 2.3 | 107.4 (91.3 – 121.3) |
| *Sequence 1* rG-DNA Round 1 | 62.8 ± 3.0 | 70.4 (56.3 – 84.6) |
| *Sequence 1* rG-DNA Round 2 | 79.1 ± 4.7 | 94.5 (67.2 – 108.7) |
| *Sequence 1* rG-DNA Round 3 | 72.0 ± 3.3 | 73.6 (65.7 – 88.4) |
| *Sequence 2* dG-DNA Round 1 | 76.5 ± 2.3 | 84.4 (65.2 – 98.8) |
| *Sequence 2* dG-DNA Round 2 | 68.4 ± 4.3 | 83.2 (71.8 – 99.1) |
| *Sequence 2* rG-DNA Round 1 | 85.8 ± 4.1 | 102.5 (73.4 – 120.1) |
| *Sequence 2* rG-DNA Round 2 | 95.4 ± 5.4 | 119.9 (93.5 – 181.6) |

**Supplementary Table S10.** Comparison of Gaussian peak values and median values of stretch modulus of ds substrates. Gaussian fitted values are presented with standard error of fit while median values are presented with 99% confidence interval of the median.

| Substrate | dG-DNA | rG-DNA | dG | rG |
|---|---|---|---|---|
| **dG-DNA** | – | < 0.0001 | < 0.0001 | 0.0008 |
| **rG-DNA** | < 0.0001 | – | 0.1357 | 0.9233 |
| **dG** | < 0.0001 | 0.1357 | – | 0.4098 |
| **rG** | 0.0008 | 0.9233 | 0.4098 | – |

**Supplementary Table S11.** Summary of p values of all combined data for *Sequence 1*. Mann-Whitney U test was performed to obtain the p values.

| Substrate | dG-DNA | rG-DNA | dG | rG |
|---|---|---|---|---|
| **dG-DNA** | – | < 0.0001 | 0.0008 | < 0.0001 |
| **rG-DNA** | < 0.0001 | – | < 0.0001 | < 0.0001 |
| **dG** | 0.0008 | < 0.0001 | – | 0.0138 |
| **rG** | < 0.0001 | < 0.0001 | 0.0138 | – |

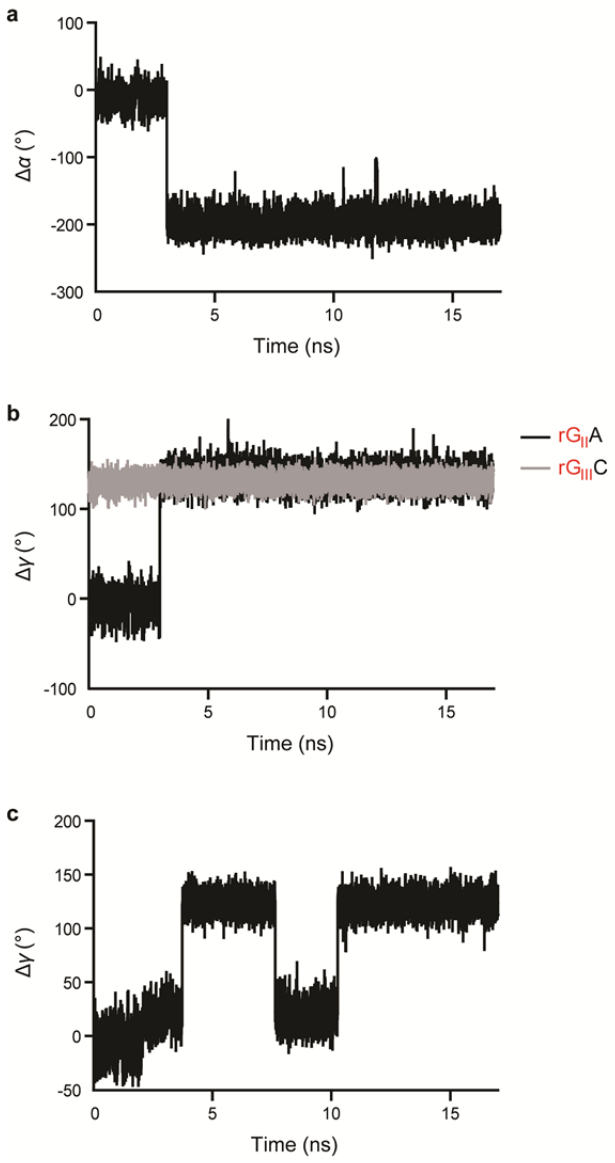
**Supplementary Table S12.** Summary of p values of all combined data for *Sequence 2*. Mann-Whitney U test was performed to obtain the p values.

| Substrate | dG-DNA Round 1 | dG-DNA Round 2 | rG-DNA Round 1 | rG-DNA Round 2 | rG-DNA Round 3 |
|---|---|---|---|---|---|
| **dG-DNA Round 1** | – | 0.0425 | < 0.0001 | 0.0041 | < 0.0001 |
| **dG-DNA Round 2** | 0.0425 | – | < 0.0001 | 0.1352 | < 0.0001 |
| **rG-DNA Round 1** | < 0.0001 | < 0.0001 | – | 0.0143 | 0.2921 |
| **rG-DNA Round 2** | 0.0041 | 0.1352 | 0.0143 | – | 0.0855 |
| **rG-DNA Round 3** | < 0.0001 | < 0.0001 | 0.2921 | 0.0855 | – |

**Supplementary Table S13.** Summary of p values of each individual round of measurements of ds substrates with *Sequence 1*. Mann-Whitney U test was performed to obtain the p values.

| Substrate | dG-DNA Round 1 | dG-DNA Round 2 | rG-DNA Round 1 | rG-DNA Round 2 |
|---|---|---|---|---|
| **dG-DNA Round 1** | – | 0.5045 | 0.0326 | 0.0001 |
| **dG-DNA Round 2** | 0.5045 | – | 0.0726 | < 0.0001 |
| **rG-DNA Round 1** | 0.0326 | 0.0726 | – | 0.0370 |
| **rG-DNA Round 2** | 0.0001 | < 0.0001 | 0.0370 | – |

**Supplementary Table S14.** Summary of p values of each individual round of measurements of ds substrates with *Sequence 2*. Mann-Whitney U test was performed to obtain the p values.




**Supplementary Figure S11. a**, Instantaneous deviation of alpha (α) torsional angle of the dAMP following rGMP in the 5' to 3' direction in CrGATGGArGCT for rGMP II. **b**, Instantaneous deviations of gamma (γ) torsional angles of the dAMP following rGMP and dCMP following rGMP in the 5' to 3' direction in CrGATGGArGCT for rGMPs II and III. **c**, instantaneous deviation of γ torsional angle of the rGMP I in GrGTTCArGGTT.

| | Enthalpy (kJ/mol) | Entropy (kJ/mol) | $T_m$ (K) |
|---|---|---|---|
| **ATGGArGCTC** | 241 ± 6 | 0.661 | 317.6 |
| **DNA-control** | 261 ± 9 | 0.733 | 314.9 |

**Supplementary Table S15.** Thermal stability of an rGMP-containing 9-bp duplex and its DNA-control in 100 mM NaCl, 10 mM phosphate, 0.5 mM EDTA at pH 6.6. The duplex concentration was 30 μM.

| Base | δ (ppm) | | Δδ (ppm) |
|---|---|---|---|
| | DNA-control | ATGGArGCTC | |
| T18 | 13.03 | 12.98 | 0.05 |
| T2 | 13.87 | 13.82 | 0.05 |
| G3 | 12.90 | 12.84 | 0.06 |
| G4 | 12.92 | 12.66 | 0.26 |
| T14 | 13.92 | 13.81 | 0.11 |
| G6/rG6 | 12.76 | 12.90 | -0.14 |
| G12 | 12.90 | 13.05 | -0.15 |
| T8 | 14.20 | 14.14 | 0.06 |
| G10 | 12.81 | 12.80 | 0.01 |
| **Base** | **δ (ppm)** | | **Δδ (ppm)** |
| | **DNA-control** | **ATCCrGGTAG** | |
| T18 | 13.36 | 13.36 | 0.00 |
| T2 | 13.87 | 13.88 | -0.01 |
| G16 | 12.79 | 12.79 | 0.00 |
| G15 | 13.06 | 12.95 | 0.11 |
| G5/rG5 | 13.06 | 13.00 | 0.06 |
| G6 | 12.82 | 12.84 | -0.02 |
| T7 | 13.72 | 13.64 | 0.08 |
| T11 | 13.85 | 13.88 | -0.03 |
| G9 | 13.16 | 13.19 | -0.03 |
| **Base** | **δ (ppm)** | | **Δδ (ppm)** |
| | **DNA-control** | **TTAGrGCCTG** | |
| T1 | - | - | - |
| T2 | 13.59 | 13.69 | -0.10 |
| T16 | 13.84 | 13.82 | 0.02 |
| G4 | 12.93 | 12.86 | 0.07 |
| G5/rG5 | 12.96 | 13.06 | -0.10 |
| G13 | 12.99 | 13.11 | -0.12 |
| G12 | 12.95 | 13.00 | -0.05 |
| T8 | 14.19 | 14.12 | 0.07 |
| G9 | 12.98 | 12.92 | 0.05 |

**Supplementary Table S16.** Imino proton NMR chemical shift data for three rGMP-containing 9-bp duplexes, ATGGArGCTC (with rGMP III), ATCCrGGTAG (with rGMP VI), and TTAGrGCCTG (with rGMP VIII), and their DNA-controls. Spectra were recorded of 1.1 mM duplexes in 100 mM NaCl, 10 mM phosphate, 10% $D_2O$ buffer (pH 6.4) at 280K using the solvent suppression jump and return pulse program.

| Nucleotide | $^{31}$P δ (ppm) | | Δδ (ppm) |
|---|---|---|---|
| | DNA-control | ATGGArGCTC | |
| A1 | – | – | – |
| T2 | -0.68 | -0.68 | 0.00 |
| G3 | -0.24 | -0.19 | 0.05 |
| G4 | -0.32 | -0.31 | 0.01 |
| A5 | -0.49 | -0.39 | 0.10 |
| G6/rG6 | -0.64 | -0.40 | 0.24 |
| C7 | -0.30 | 0.50 | 0.80 |
| T8 | -0.83 | -1.12 | -0.29 |
| C9 | -0.42 | -0.47 | -0.05 |
| G10 | – | – | – |
| A11 | -0.48 | -0.53 | -0.05 |
| G12 | -0.61 | -0.57 | 0.04 |
| C13 | -0.34 | -0.45 | -0.11 |
| T14 | -0.86 | -0.92 | -0.05 |
| C15 | -0.68 | -0.70 | -0.02 |
| C16 | -0.38 | -0.36 | 0.02 |
| A17 | -0.28 | -0.23 | 0.05 |
| T18 | -0.61 | -0.60 | 0.01 |

| Nucleotide | $^{31}$P δ (ppm) | | Δδ (ppm) |
|---|---|---|---|
| | DNA-control | ATCCrGGTAG | |
| A1 | - | - | - |
| T2 | -0.77 | -0.77 | 0.00 |
| C3 | -0.65 | -0.63 | 0.02 |
| C4 | -0.51 | -0.42 | 0.09 |
| G5/rG5 | -0.36 | -0.50 | -0.14 |
| G6 | -0.43 | 0.85 | 1.28 |
| T7 | -0.80 | -1.07 | -0.27 |
| A8 | -0.59 | -0.69 | -0.10 |
| G9 | -0.53 | -0.54 | -0.01 |
| C10 | - | - | - |
| T11 | -0.73 | -0.73 | 0.00 |
| A12 | -0.51 | -0.52 | -0.01 |
| C13 | -0.66 | -0.66 | 0.00 |
| C14 | -0.54 | -0.50 | 0.04 |
| G15 | -0.43 | -0.40 | 0.03 |
| G16 | -0.43 | -0.48 | -0.05 |
| A17 | -0.43 | -0.48 | -0.05 |
| T18 | -0.67 | -0.67 | 0.00 |

| Nucleotide | $^{31}$P δ (ppm) | | Δδ (ppm) |
|---|---|---|---|
| | DNA-control | TTAGrGCCTG | |
| T1 | - | - | - |
| T2 | -0.72 | -0.79 | -0.07 |
| A3 | -0.47 | -0.54 | -0.07 |
| G4 | -0.50 | -0.50 | 0.00 |
| G5/rG5 | -0.38 | 0.02 | 0.40 |
| C6 | * | -0.44 | * |
| C7 | * | -0.87 | * |
| T8 | -0.70 | -1.02 | -0.32 |
| G9 | -0.43 | -0.54 | -0.11 |

| | | | |
|---|---|---|---|
| C10 | - | - | - |
| A11 | -0.43 | -0.53 | -0.10 |
| G12 | -0.50 | -0.61 | -0.11 |
| G13 | -0.40 | -0.36 | 0.04 |
| C14 | -0.57 | -0.67 | -0.10 |
| C15 | -0.55 | -0.69 | -0.14 |
| T16 | -0.79 | -0.94 | -0.15 |
| A17 | -0.61 | -0.62 | -0.01 |
| A18 | -0.57 | -0.60 | -0.03 |

**Supplementary Table S17.** $^{31}$P NMR chemical shift data for three rGMP-containing 9-bp duplexes, ATGGArGCTC (with rGMP III), ATCCrGGTAG (with rGMP VI), and TTAGrGCCTG (with rGMP VIII), and their DNA-controls at 294K. The phosphorous resonances are a good indicator on the status of the nucleic acid backbone. Shown here, large deviations in chemical shift between rGMP-containing duplexes and their DNA-contorls suggest localized perturbations in the backbone 3' of the damage site on both the top and bottom strand. Note: 5' nucleotides A1 and G10 do not have phosphate groups.
* Denotes uncertainty; $^{31}$P resonances for nucleotides C6 and C7 were in the range of -0.47 to -0.74 ppm.

## REFERENCES


1. T. G. Christopher, S. W. Gregory and M. Sverre, *Nanotechnology*, 1996, **7**, 259.
2. D. S. Ashley, J. B. Adam, S. Q. Jamie and T. G. Christopher, *Nanotechnology*, 2013, **24**, 15710.
3. J. C. Peter, A. C. Charles and H. John, *Measurement Science and Technology*, 2004, **15**, 1337.
4. H.-J. Butt, B. Cappella and M. Kappl, *Surface Science Reports*, 2005, **59**, 1-152.
5. B. Peirce, *The Astronomical Journal*, 1852, **2**, 161-163.
6. B. A. Gould, *The Astronomical Journal*, 1855, **4**, 81-87.
7. H. Li and R. Durbin, *Bioinformatics*, 2009, **25**, 1754-1760.
8. J. M. Aramini and M. W. Germann, *Biochemistry*, 1999, **38**, 15448-15458.