**Insights on the behaviour of phosphorylated DNA breaks from molecular dynamic simulations**

**Li Zhang1, Outi Lampela1,2, Lari Lehtiö1,2, André H. Juffer1,2, \***

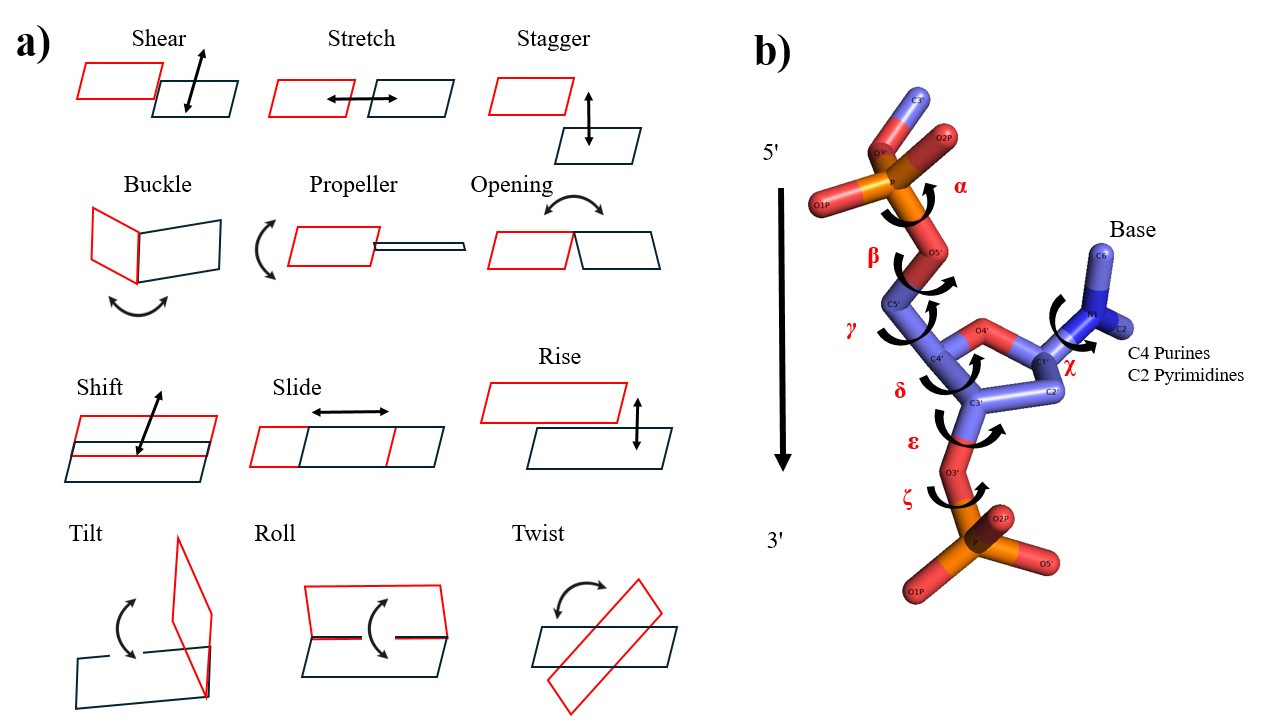
1Faculty of Biochemistry and Molecular Medicine, University of Oulu, PO Box 5400, 90014, Oulu, Finland.

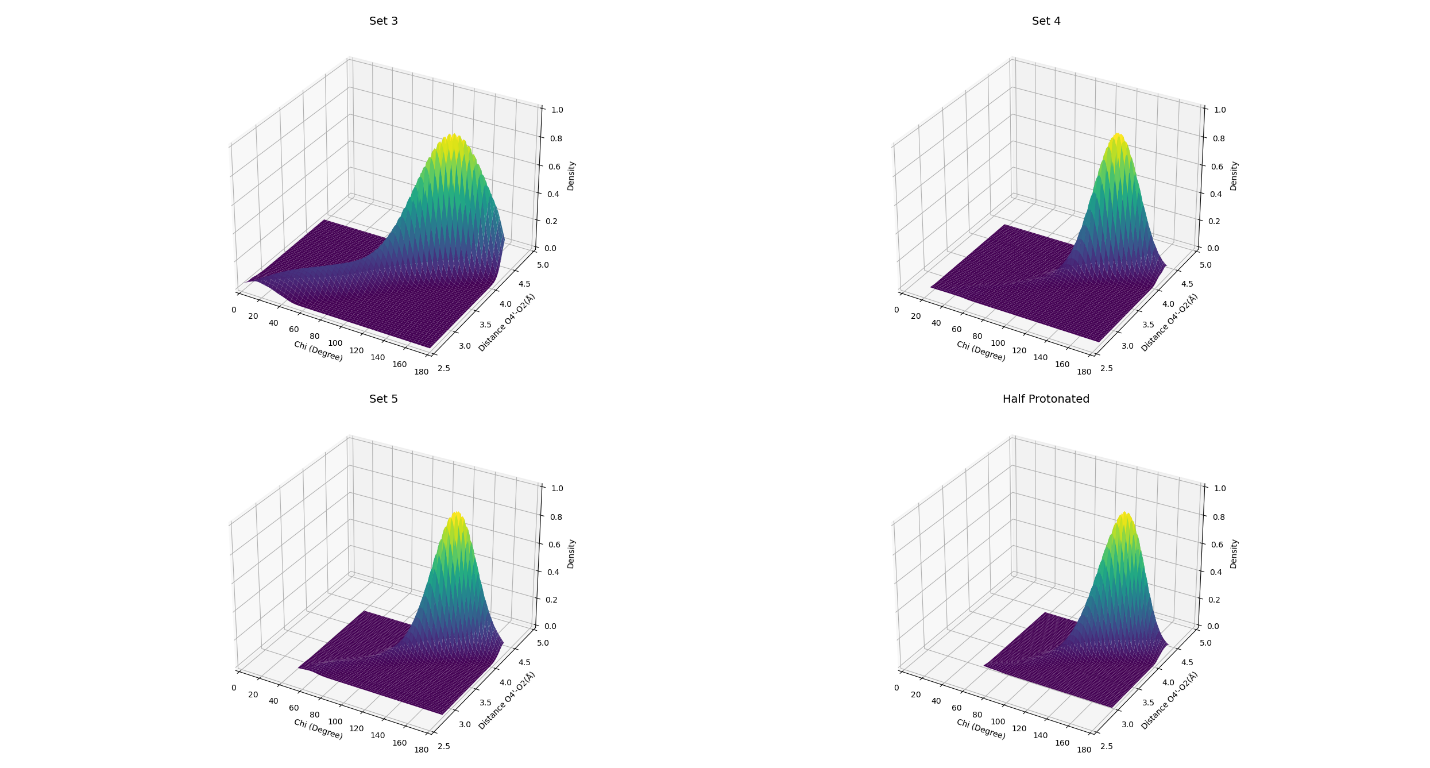
2Biocenter Oulu, University of Oulu, PO Box 5400, 90014, Oulu, Finland.

\*Author to whom all correspondence should be addressed. Email : [andre.juffer@oulu.fi](mailto:andre.juffer@oulu.fi)

**Table S1.** DNA sequences used in this study: the 5’ phosphorylated (5’P) break is in bold, the break is highlighted in yellow.

|  |  |
| --- | --- |
| C-break | 5’GGCATCATAG3’ **5’P** CAATACATGC3’ |
| T-break | 5’GGCATCATAG3’ **5’P** TAATACATGC3’ |
| G-break | 5’GGCATCATAG3’ **5’P** GAATACATGC3’ |
| A-break | 5’GGCATCATAG3’ **5’P** AAATACATGC3’ |

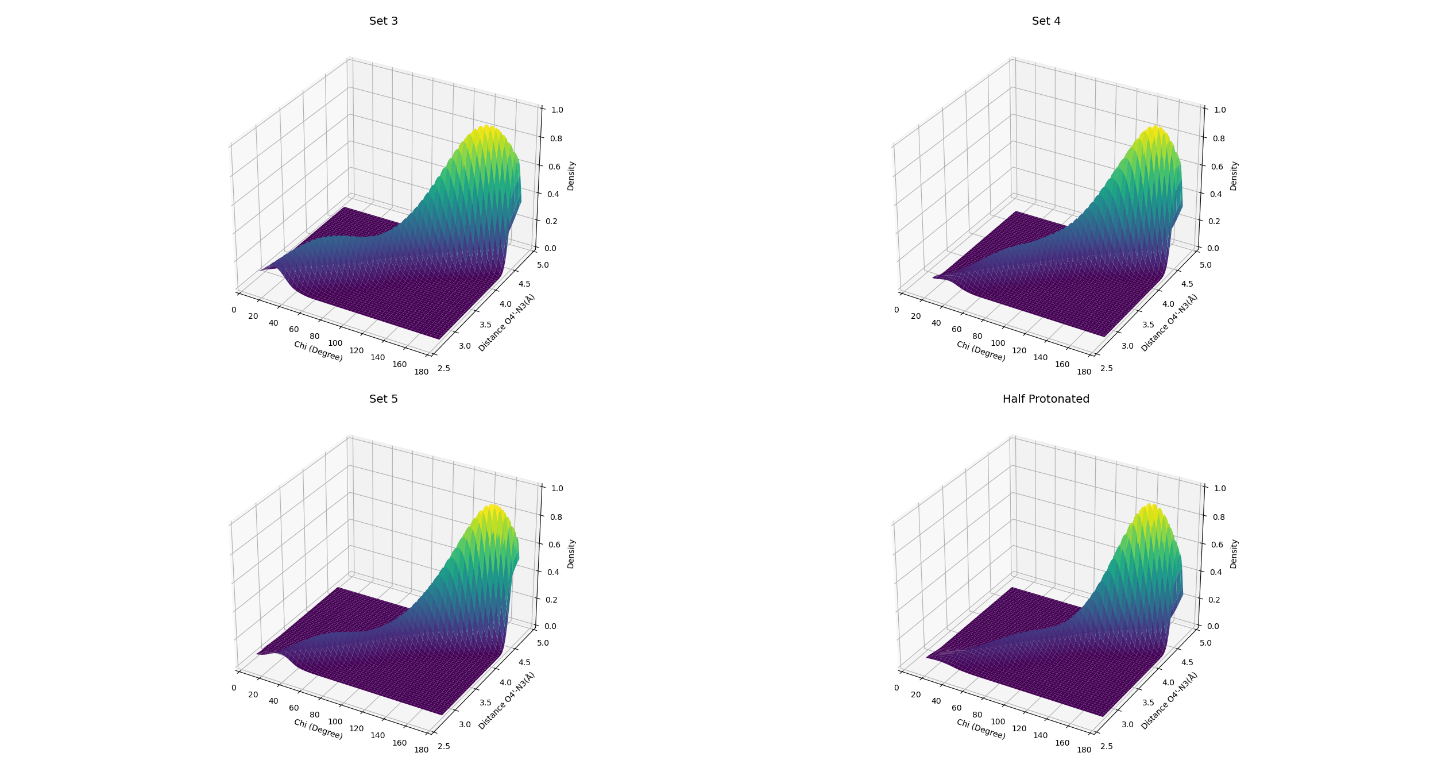
**Figure S1** a). Graphic illustration helical properties for intra-base pair (*shear*, *stretch, stagger, buckle, propeller, opening)*, inter-base pair steps (*shift, slide, rise, tilt, roll, twist*); b). DNA backbone torsion angles (α, β, γ, δ, ε, ζ and χ): α: O3'(i-1)-P-O5'-C5'; β: P-O5'-C5'-C4'; γ: O5'-C5'-C4'-C3'; δ: C5'-C4'-C3'-O3'; ε: C4'-C3'-O3'-P(i+1); ζ: C3'-O3'-P(i+1)-O5'(i+1); χ: purines:O4′-C1′-N9-C4, pyrimidines:O4′-C1′-N1-C2.



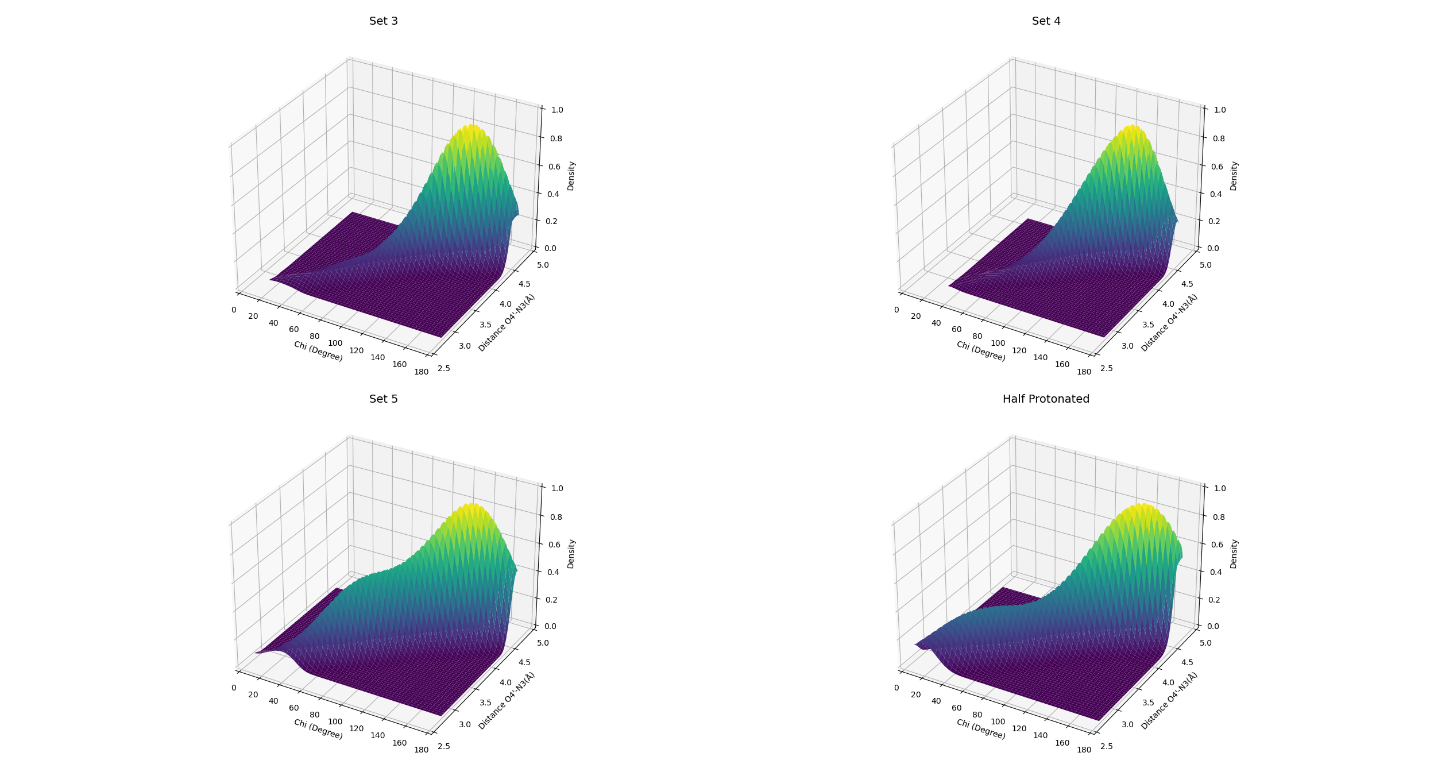
**Figure S2.** Distribution of the dihedral angle χ (in degrees) and distance *d*OO (in Å) of O4’ to O2. Top row: Charge Set3 (left) and 4, bottom row: Charge Set5 (left) and half-protonated (“Half”) for the C-break.



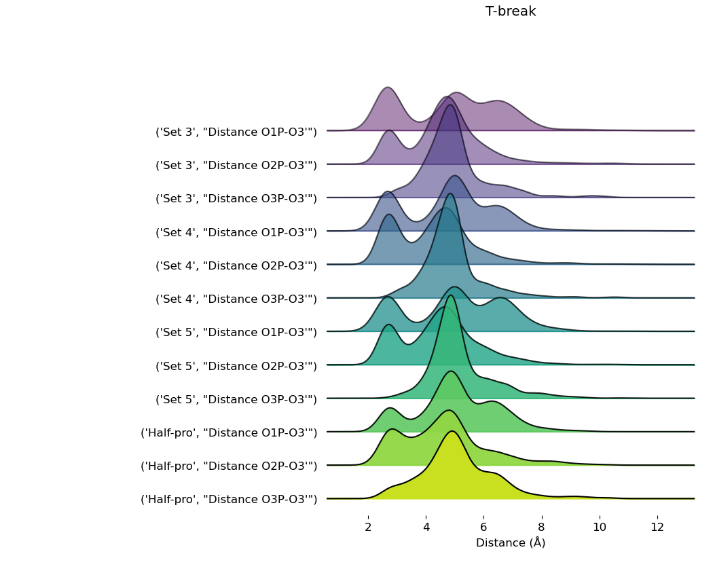
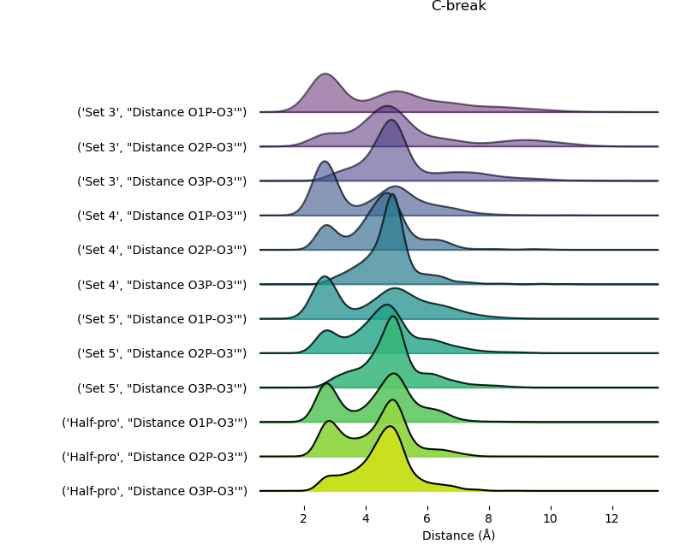
**Figure S3.** As **Figure S2** but now for the T-break.

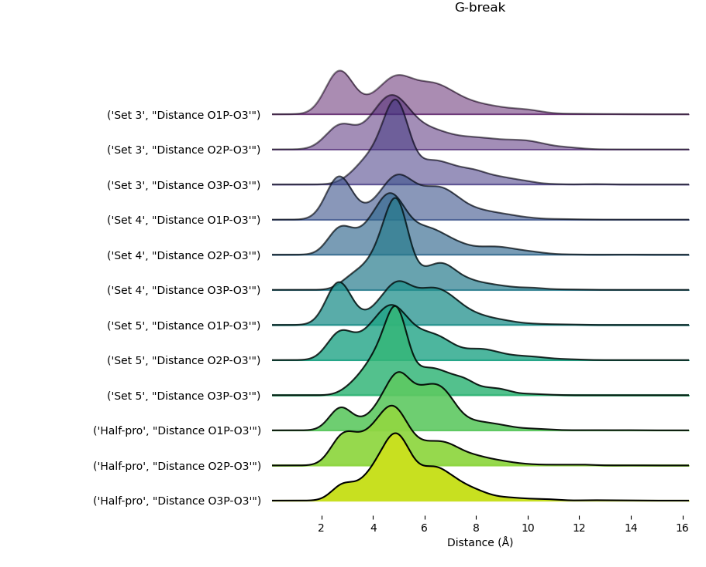


**Figure S4.** As **Figure S2** but now for the G-break, but the distance *d*OO is now the distance between O4’ and N3.

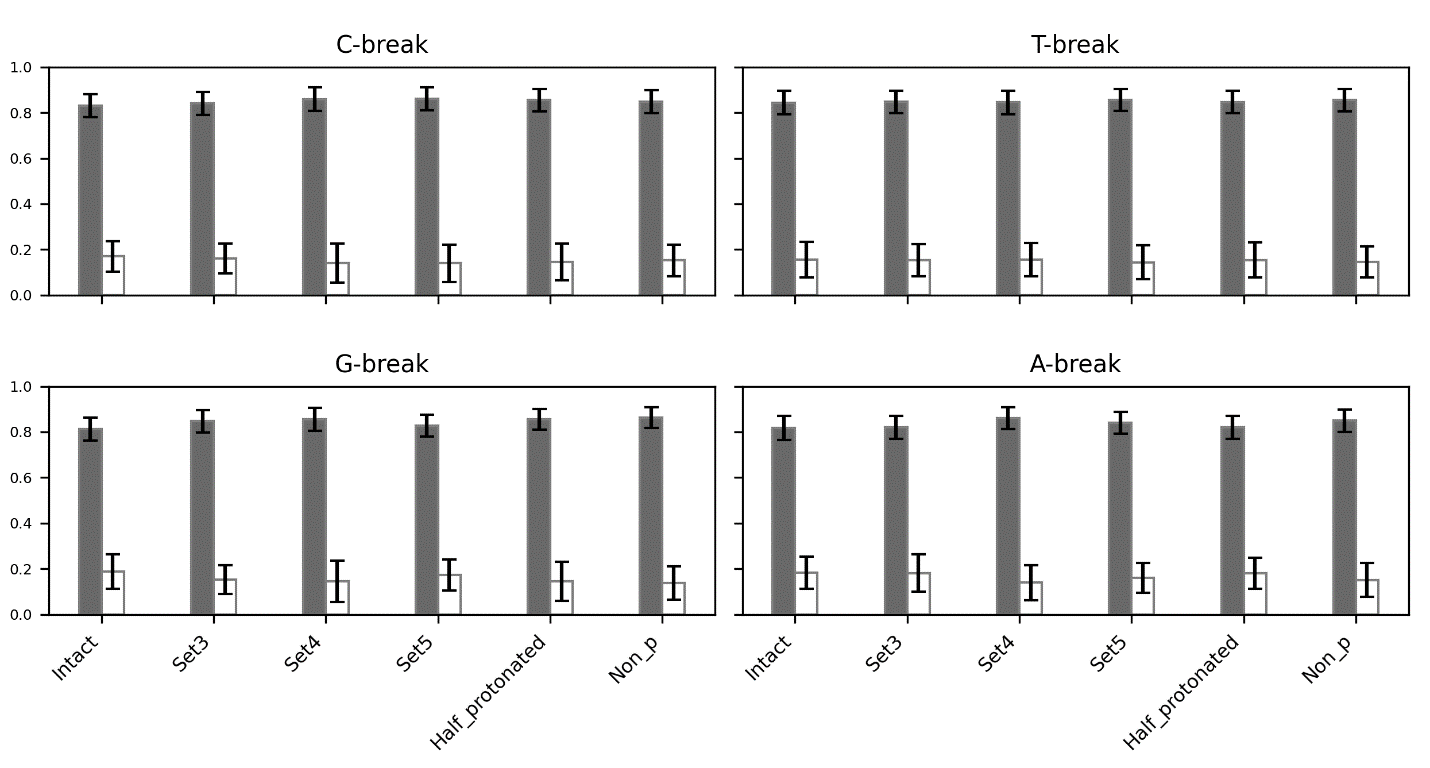


**Figure S5.** As **Figure S4** but now for the A-break.

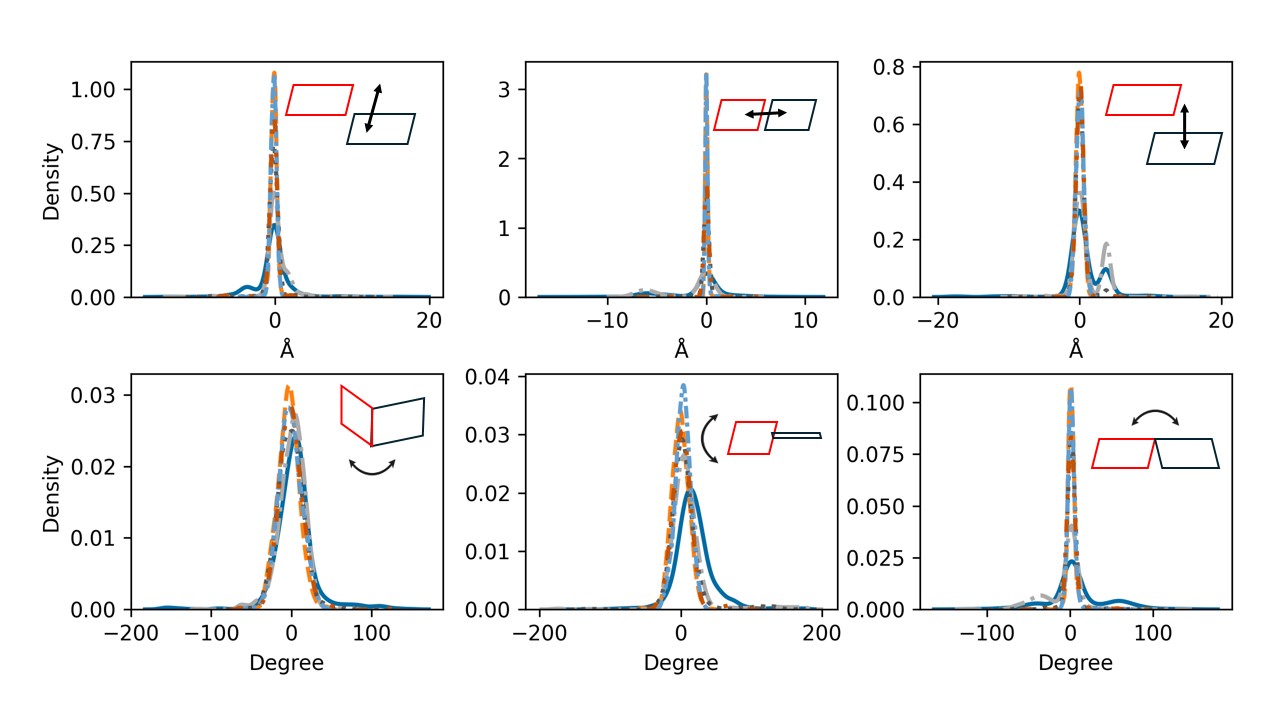


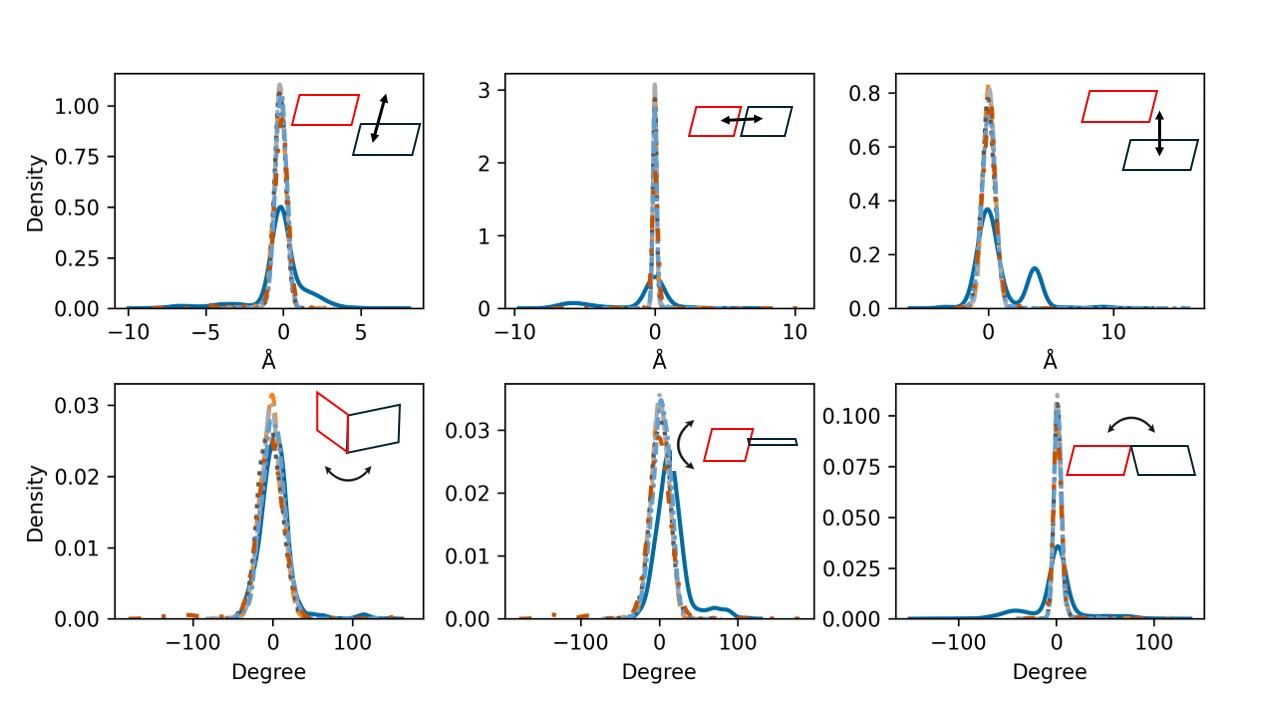


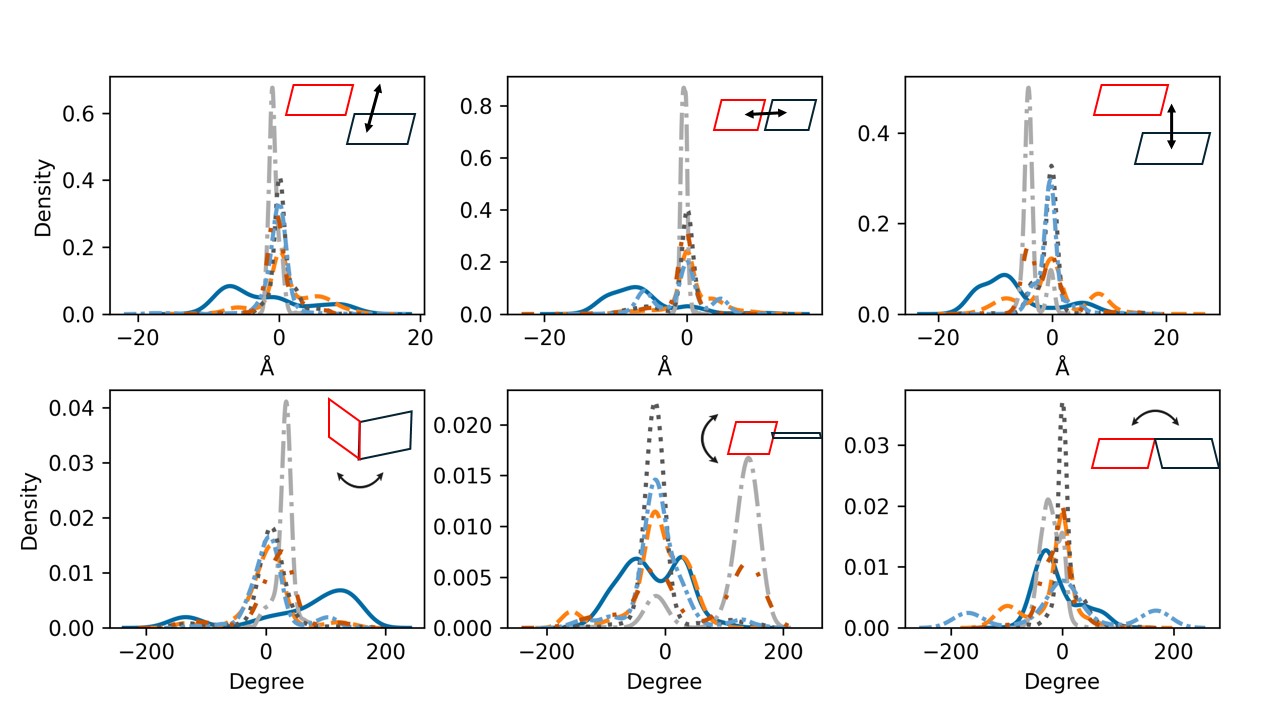
**Figure S6:** Distance distributions for all break types and pairs of oxygen atoms (the 3’ end oxygen O3’ and the 5’ end OPs). Top row: C-break (left) and T-break. Bottom row: G-break (left) and A-break. The label on the left of each graph identifies the oxygen pair in question. The distance is in Å.

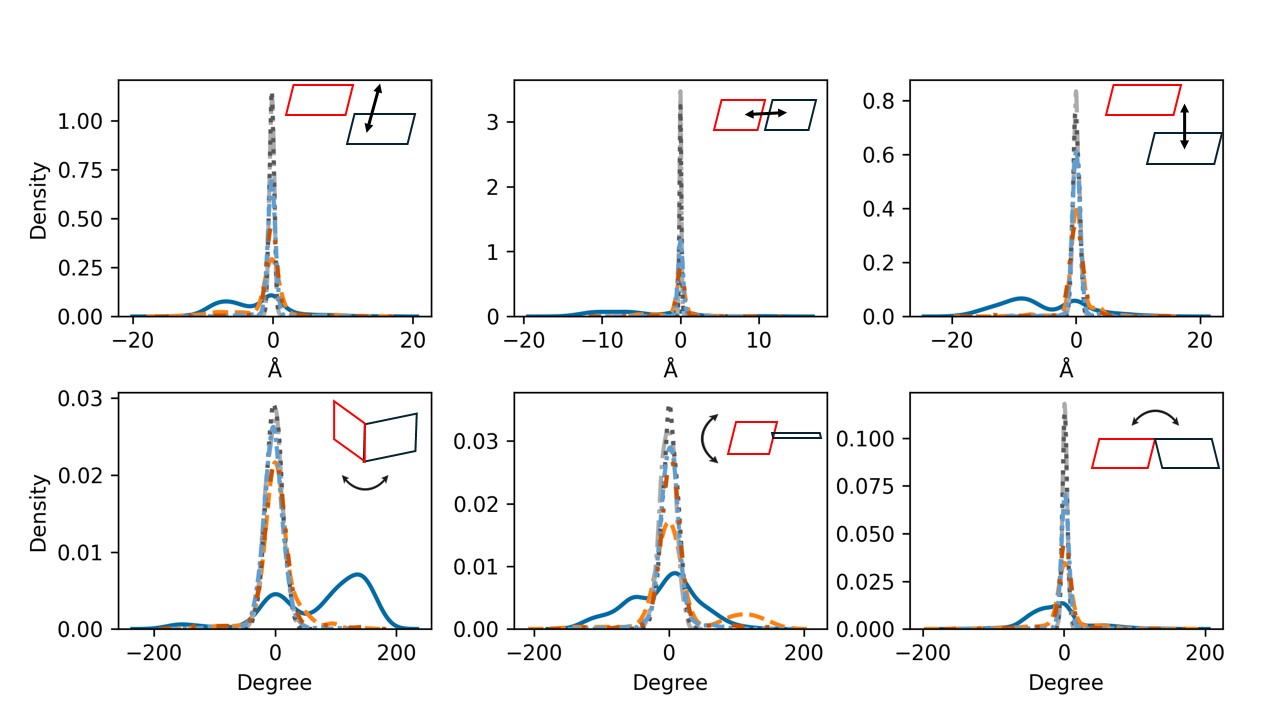


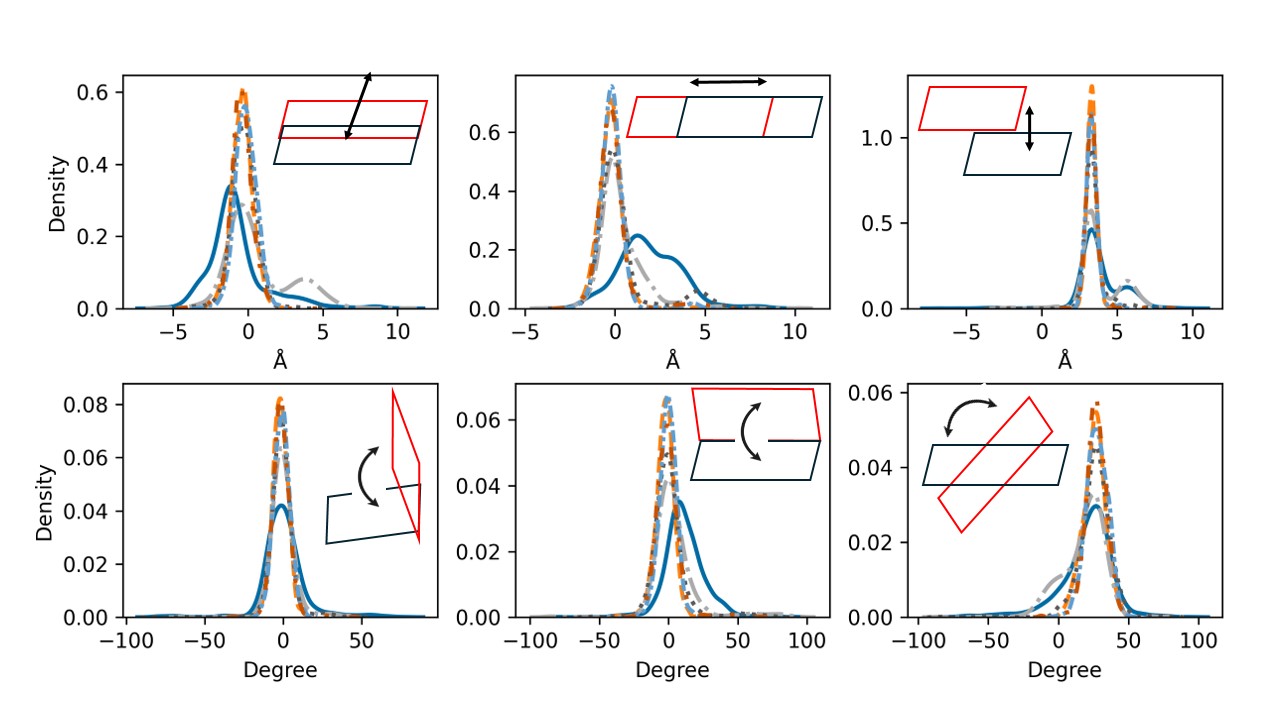
**Figure S7.** The DNA B form analysis for intact DNA, and charge sets 3 to 5, half-protonated, and dephosphorylated (“Non\_P”) break for all X-breaks. Grey is BI, white represents BII. Shown is the overall probability of occurrence (in %) for BI and BII, respectively, including a measure for the standard deviation.

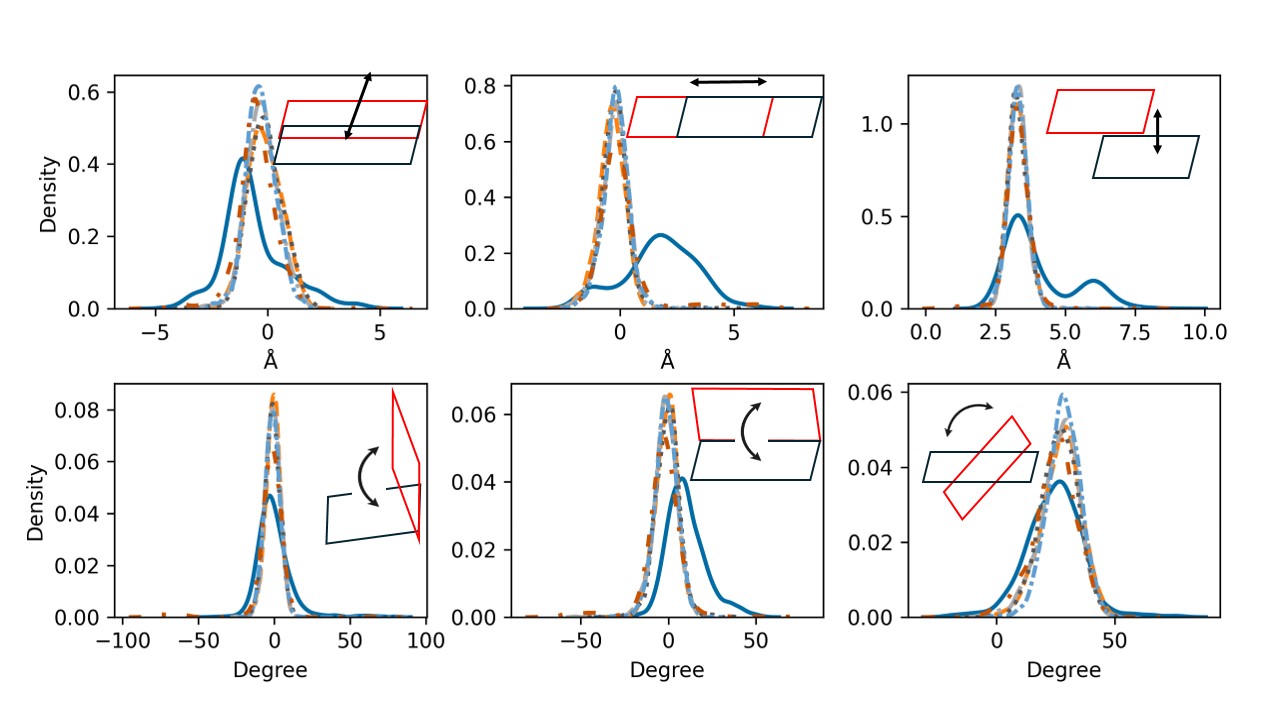
**Figure S8.** The intra-base pair parameters calculated for C-break compared to intact DNA. Top row: the translational parameters from left to right: *shear*, *stretch,* and *stagger*; bottom row: therotational parameter from left to right: *buckle, propeller,* and *opening*. Shown are the parameters’ probability density functions. Intact DNA: solid cerulean blue line; damaged DNA with charge set3: dashed pumkin orange; charge set4: dashdot dark gray; charge set5: dotted mortar grey; half-protonated: densely dashdotted picton blue; non-phosphorylated: dashdotted tenne orange.

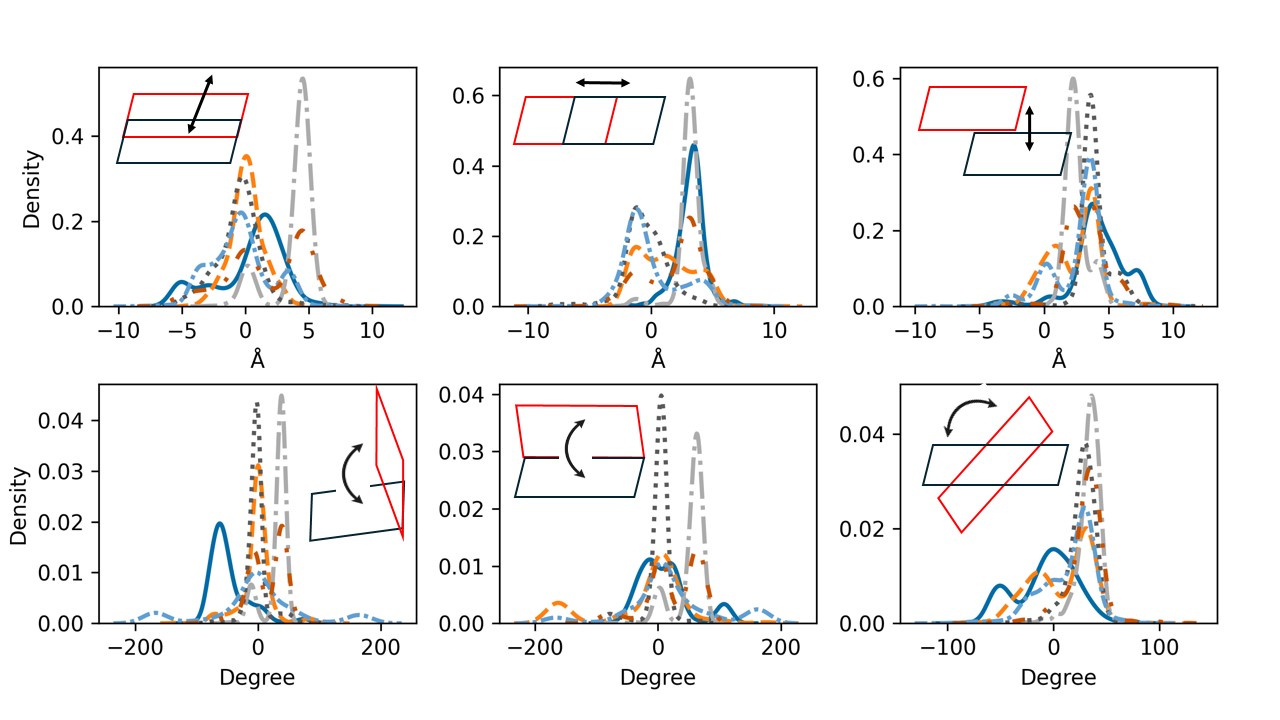
**Figure S9.** As **Figure S8**, helical parameters for T-break.

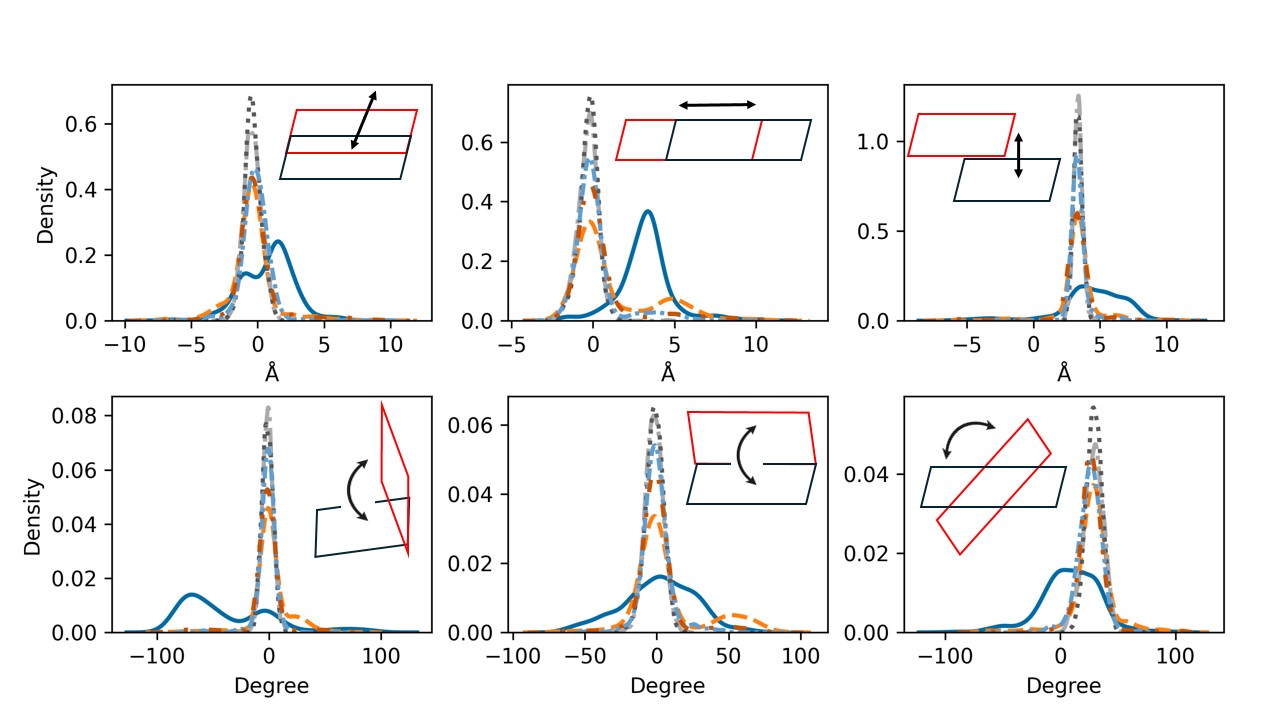
**Figure S10.** As **Figure S8**, helical parameters for G-break.

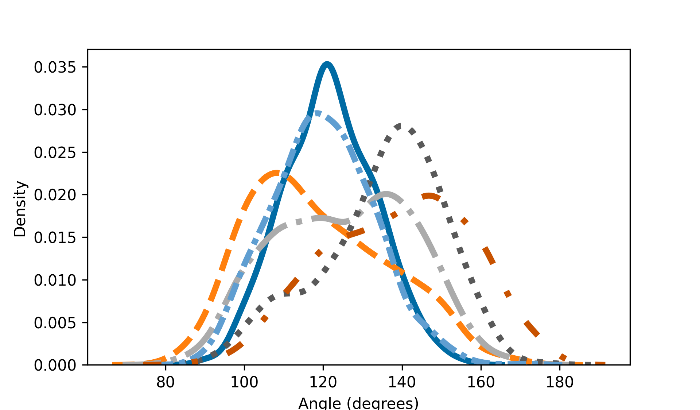
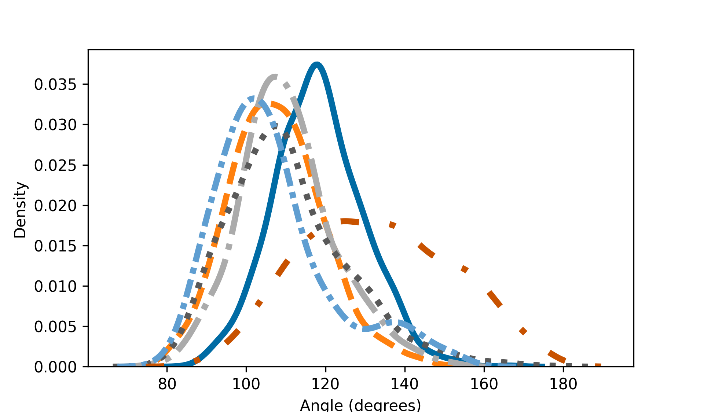
**Figure S11.** As **Figure S8**, helical parameters for A-break.

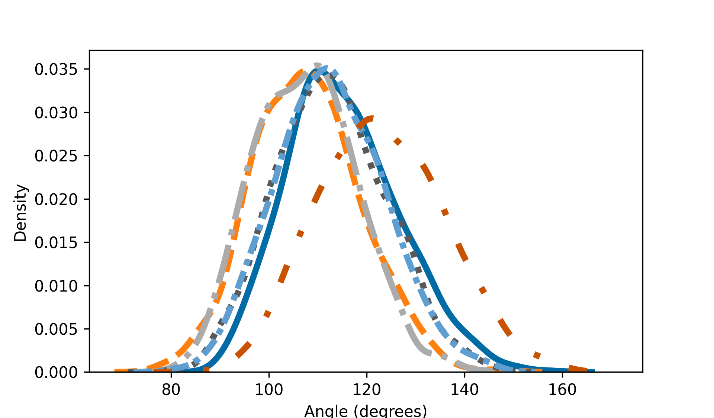
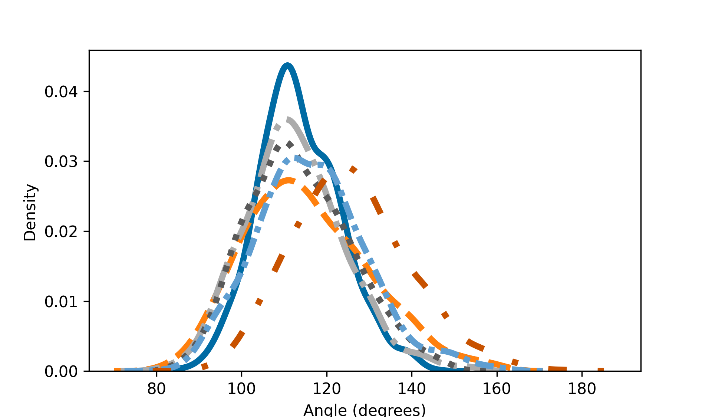
**Figure S12.** The inter-base pair parameters calculated for C-break compared to intact DNA. Top row: the translational parameters from left to right: *shift, slide,* and *rise*; bottom row: therotational parameter from left to right: *tilt, roll,* and *twist*. Shown are the parameters’ probability density functions. Intact DNA: solid cerulean blue line; damaged DNA with charge set3: dashed pumkin orange; charge set4: dashdot dark gray; charge set5: dotted mortar grey; half-protonated: densely dashdotted picton blue; non-phosphorylated: dashdotted tenne orange.

**Figure S13.** As **Figure S12**, inter-base pairs parameters for each T-break.

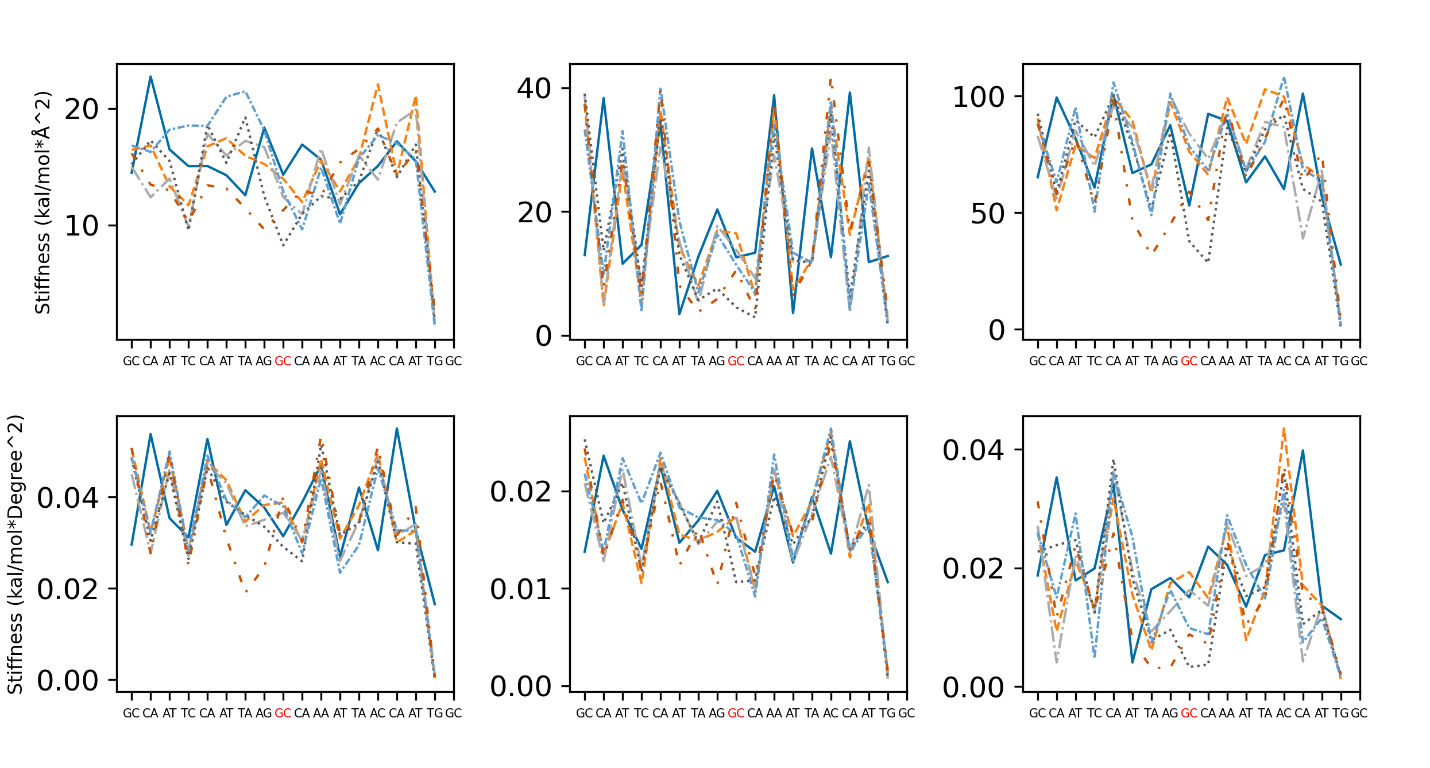
**Figure S14.** As **Figure S12**, inter-base pairs parameters for each G-break.

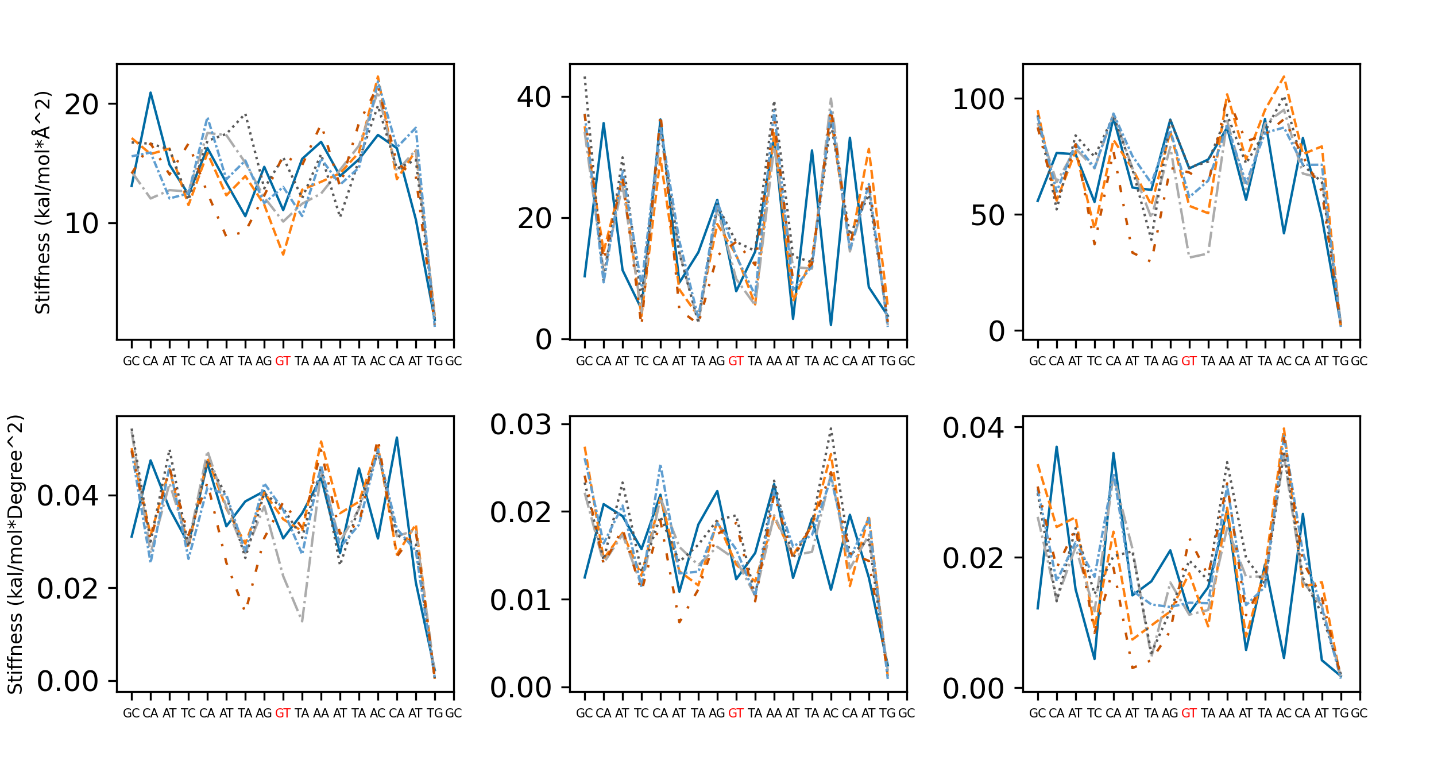
**Figure S15.** As **Figure S12**, inter-base pairs parameters for each A-break.

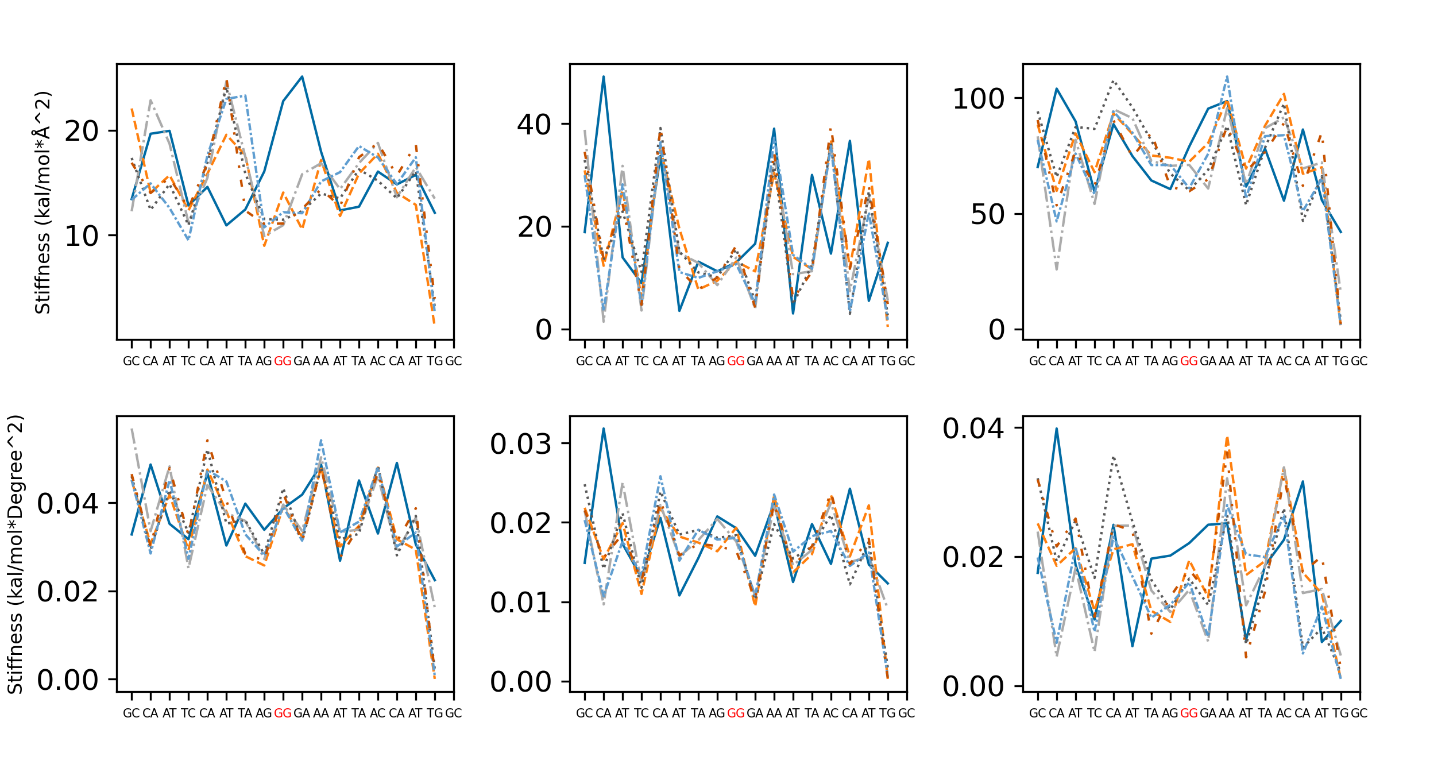


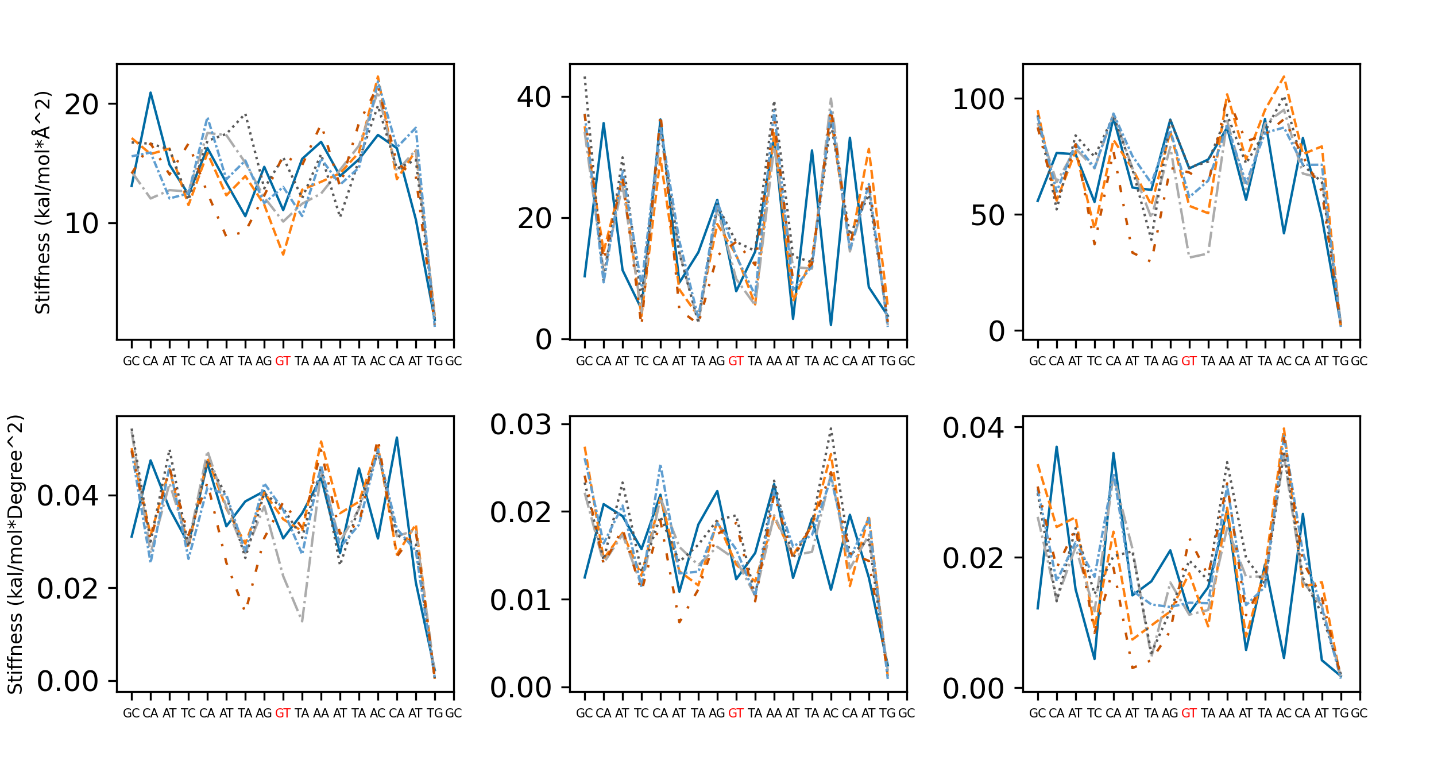


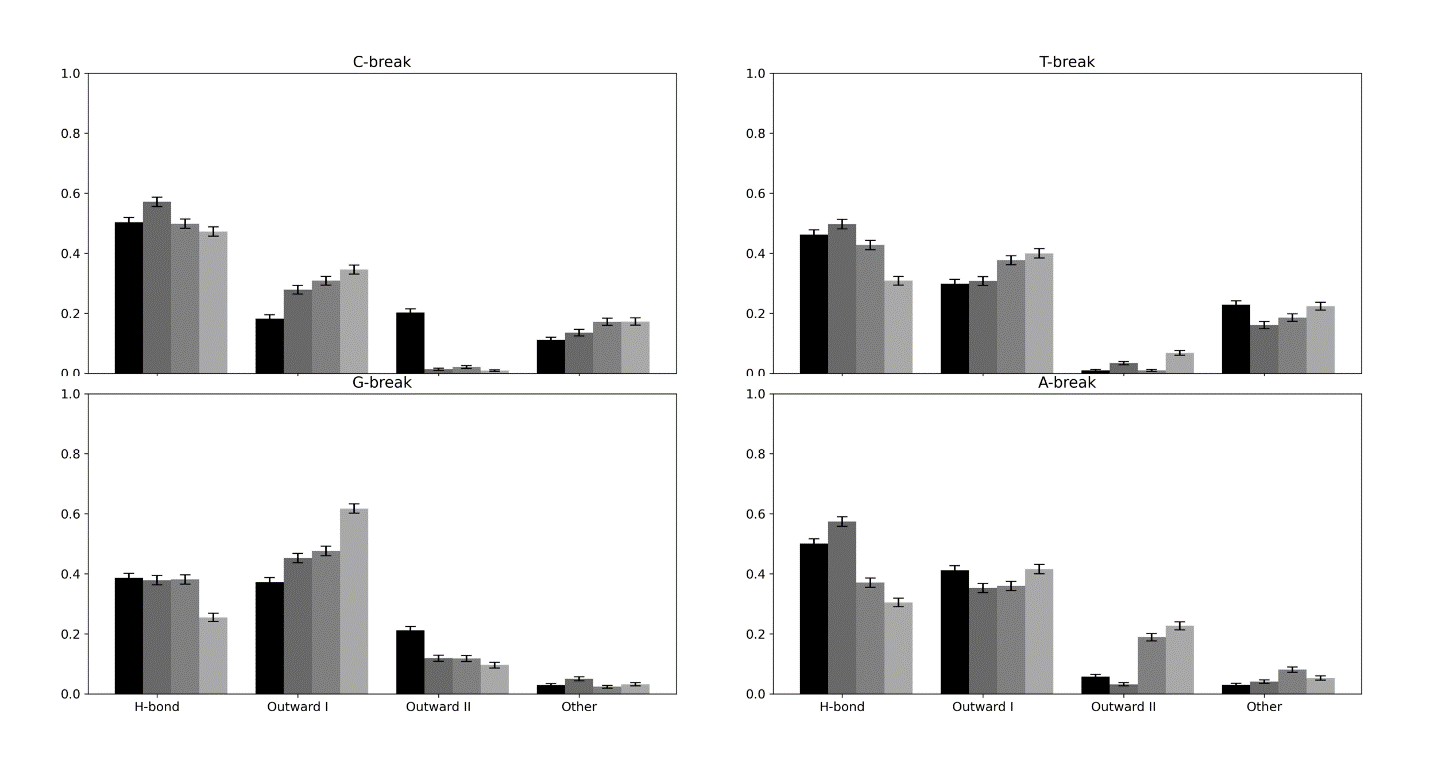
**Figure S16.** The local bending angle distribution of X-breaks, Top row: C-break (left), T-break (right) (pyrimidines); bottom row: G-break (left), A-break (right) (purines). Shown are the probability density functions of these angles. Intact DNA: solid cerulean blue line; damaged DNA with charge set3: dashed pumkin orange; charge set4: dashdot dark gray; charge set5: dotted mortar grey; half-protonated: densely dashdotted picton blue; non-phosphorylated: dashdotted tenne orange.

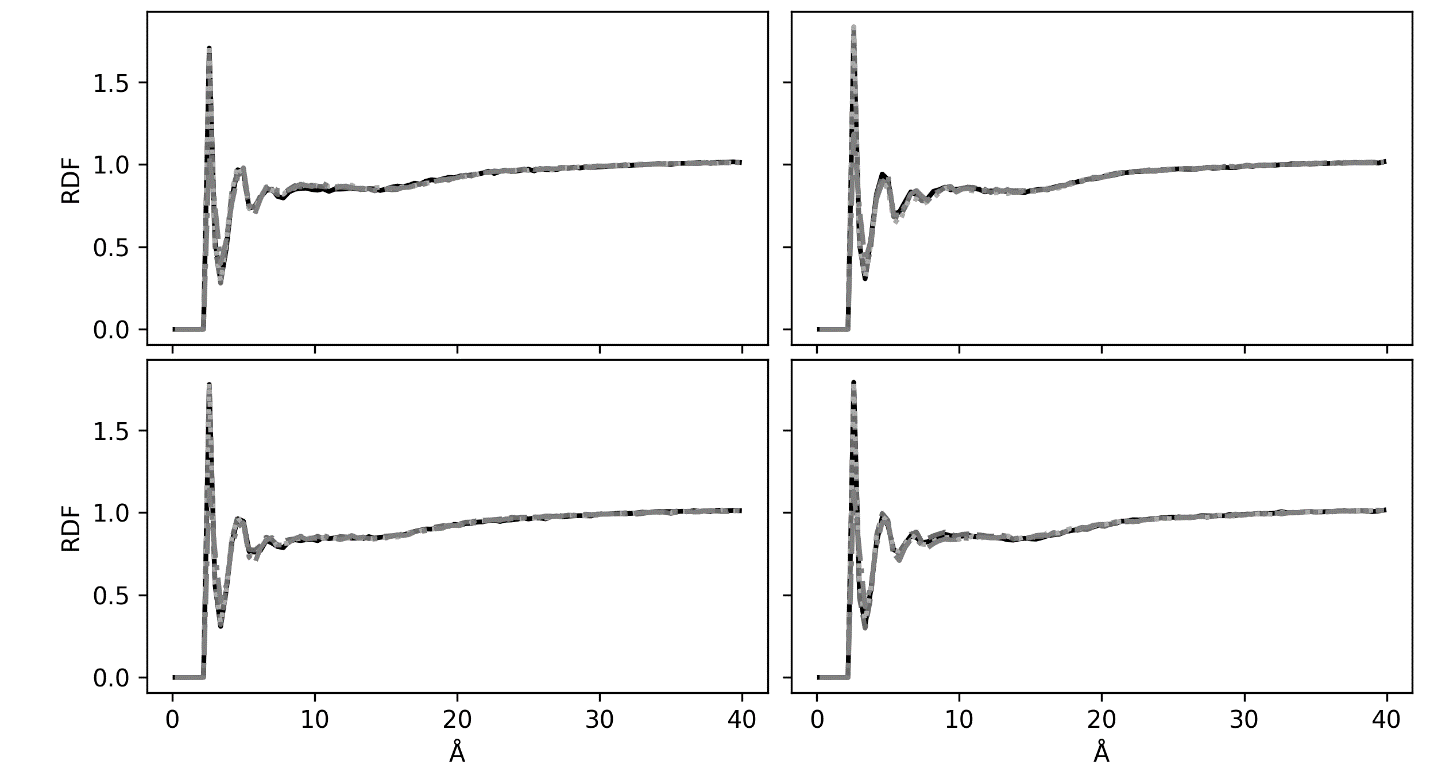
**Figure S17.** The helical stiffness of each base pair step for intact DNA and damaged DNA of C-break. Top row: from left to right: *shift, slide,* and *rise;* bottom row: from left to right: *tilt, roll* and *twist*. Intact DNA: solid cerulean blue line; damaged DNA with charge set3: dashed pumkin orange; charge set4: dashdot dark gray; charge set5: dotted mortar grey; half-protonated: densely dashdotted picton blue; non-phosphorylated: dashdotted tenne orange. The break site is highlighted in red.

**Figure S18.** As **Figure S17**, inter-base pairs parameters for each T-break.

**Figure S19.** As **Figure S17**, inter-base pairs parameters for each G-break.

**Figure S20.** As **Figure S17**, inter-base pairs parameters for each A-break.

**Figure S21.** The probability of occurrence of 5'P orientational states for X-breaks calculated for charge set3-5 and half-protonated 5'P: Top row: C-break (left) and T-break; bottom row: G-break (left) and A-break. Each probability includes a measure for its error. Top row: C-break (left), T-break (right) (pyrimidines); bottom row: G-break (left), A-break (right) (purines). Damaged DNA with charge set3 in black; charge set4: dimgray; charge set5: darkgray; half-protonated: gray. The actual numbers are listed in **Table 4**.

**Figure S22.** The radial distribution function (RDF) between P of 5’P and the O of water for each charge set and for of each break. Top row: C-break (left), T-break (right) (pyrimidines); bottom row: G-break (left), A-break (right) (purines). Damaged DNA with charge set3: solid black line; charge set4: dot dimgray; charge set5: dotted darkgray; half-protonated: dashdotted gray.