Double-stranded RNA bending by AU-tract sequences

Alberto Marin-Gonzalez[1], Clara Aicart-Ramos[1], Mikel Marin-Baquero[1], Alejandro Martín-González[1], Maarit Suomalainen[2], Abhilash Kannan[2], J. G. Vilhena[3, 4], Urs F. Greber[2], Fernando Moreno-Herrero[1*] and Rubén Pérez[4, 5*]

[1]Department of Macromolecular Structures, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas, 28049 Cantoblanco, Madrid, Spain

[2]Department of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

[3]Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland

[4]Departamento de Física Teórica de la Materia Condensada, Universidad Autónoma de Madrid, E-28049 Madrid, Spain

[5]IFIMAC - Condensed Matter Physics Center, Universidad Autónoma de Madrid, E-28049 Madrid, Spain

* To whom correspondence should be addressed. Email: fernando.moreno@cnb.csic.es or ruben.perez@uam.es

Supplementary Information

1. **Experimental details**
2. **Supporting figures**
3. **Supporting tables**

1. Experimental details

To produce the ExpAU-4 dsRNA molecule (**Table S1**) we fabricated a set of plasmids that contained repetitions of the ATAT sequence separated by seven random base pairs $(ATAT)N_7$, by several rounds of ligation of different pairs of hybridized oligonucleotides. In the first ligation, we employed the oligonucleotides *98.5P-F-T7* and *99.5P-R-T7* (fragment 98 and 99, **Table S3**) in order to locate the T7 promoter before the AT-periodic region. We then performed five additional rounds of ligations at different restriction sites of the hybridized oligonucleotides *100.5P-F-XhoI-blunt* and *101.5P-R-XhoI-blunt* (fragment 100 and 101, **Table S3**), *102.5P-F-blunt-blunt* and *103.5P-R-blunt-blunt* (fragment 102 and 103, **Table S3**). After six rounds of ligation, the plasmid (pBlueSK-T7-6oligos) contained a DNA segment with the final length. A similar strategy was followed to produce the ExpAU-5 dsRNA molecule. In this case, we fabricated a set of plasmids that contained a segment of repetitions of the TATAT sequence separated by six random base pairs $TATAT(N)_6$. The three pairs of hybridized oligonucleotides employed here were: *121.5P-F-T7* and *122.5P-R-T7* (fragment 121 and 122, **Table S3**), *123.5P-F-XhoI-blunt* and *124.5P-R-XhoI-blunt* (fragment 123 and 124, **Table S3**), *125.5P-F-blunt-blunt* and *126.5P-R-blunt-blunt* (fragment 125 and 126, **Table S3**). Finally, a control dsRNA with a GC content of ~50% and the same length of the ExpAU-4 molecule (612 bp) was produced. To fabricate this molecule, we PCR amplified a Lambda DNA fragment (NEB) of ~50% GC-content with the oligonucleotides *114.F RNA control 612* and *113.R RNA control 1316*. The PCR product was digested, purified and ligated with the long dephosphorylated fragment purified after digestion of pBlueSK-T7-6oligos plasmid with KpnI. All the plasmids were checked by DNA sequencing. In vitro transcription and annealing of the two complementary ssRNA was done as explained in the Experimental Section of the main text. The final dsRNA molecules were checked using gel electrophoresis. The three dsRNA sequences (Control, AU-4 and AU-5) showed similar migration in an agarose gel (see **Fig. S12**).

2. Supporting figures

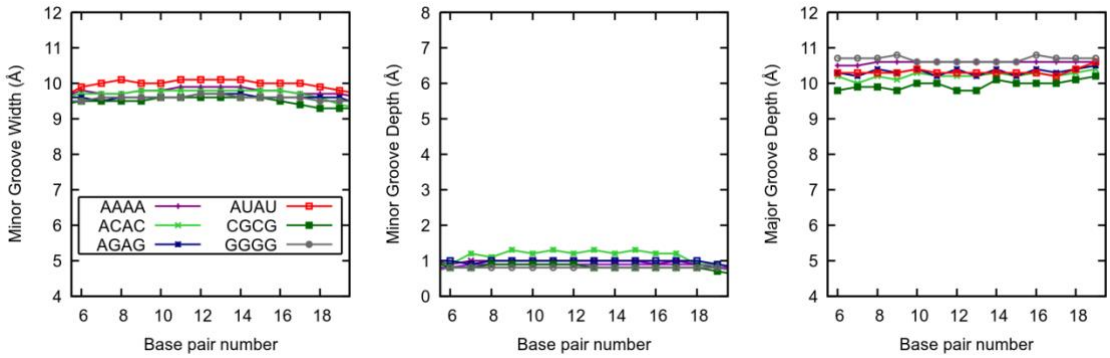

Figure S1. Groove dimensions of the benchmark sequences. Using the software Curves+ (Ref. 42, main text) the minor groove dimensions and the major groove depth were computed for the average structures of the simulated benchmark molecules. These quantities are represented as function of the dsRNA sequence. The flat lines obtained are consistent with the homogeneous sequences of the molecules. No significant sequence dependence was observed for these parameters.

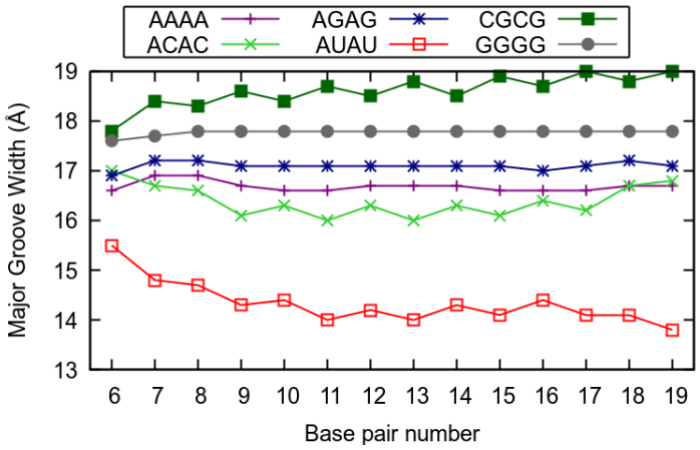

Figure S2. Major groove width profiles of the benchmark sequences computed using 3DNA (Ref. 43, main text). This figure was obtained in the same way as **Fig. 1a**, main text, the only difference being that 3DNA was used instead of Curves+.

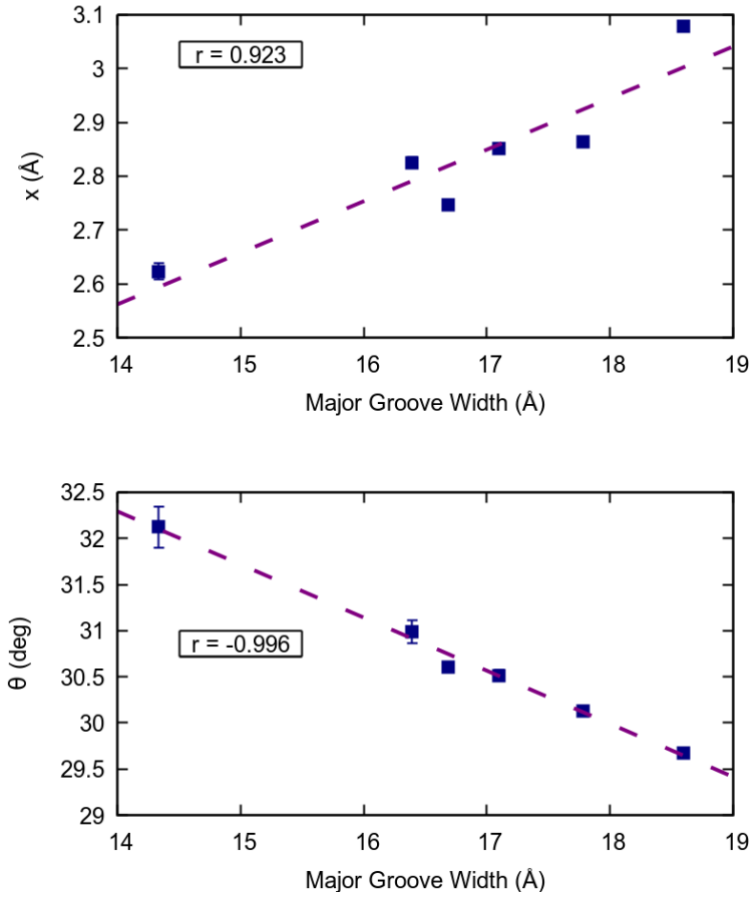

Figure S3. Major groove width correlations using 3DNA. The mean values of the major groove width, helical rise and helical twist of the 15 central base pair steps were obtained as in **Fig. 1b, c**, main text, using 3DNA instead of Curves+. Error bars are the standard error of the mean when averaging over the base pair steps. The dotted line represents a linear fit to the data and *r* is the value of the correlation coefficient.

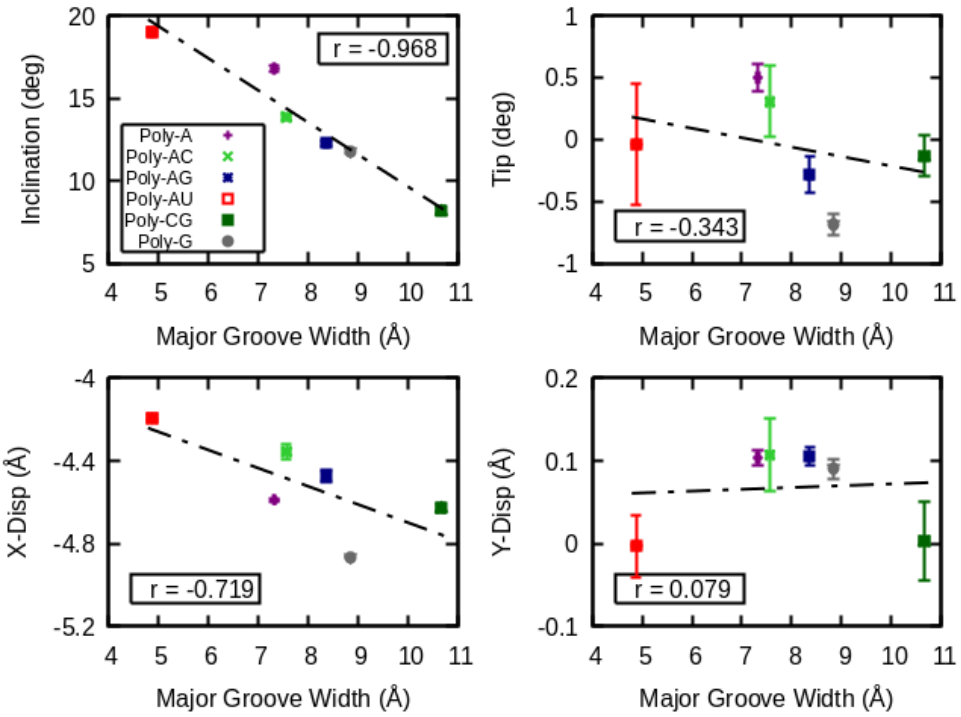

Figure S4. Helical parameters and correlations with the major groove width measured using Curves+. The values of the helical parameters and the major groove width were obtained by analyzing the average structures of the benchmark molecules using the software Curves+. These values were then averaged over the 15 central base pair steps as done in **Fig. 1b, c**, main text. Error bars are the s.e.m of this average. The dotted line is a linear fit to the data points and r is the correlation coefficient.

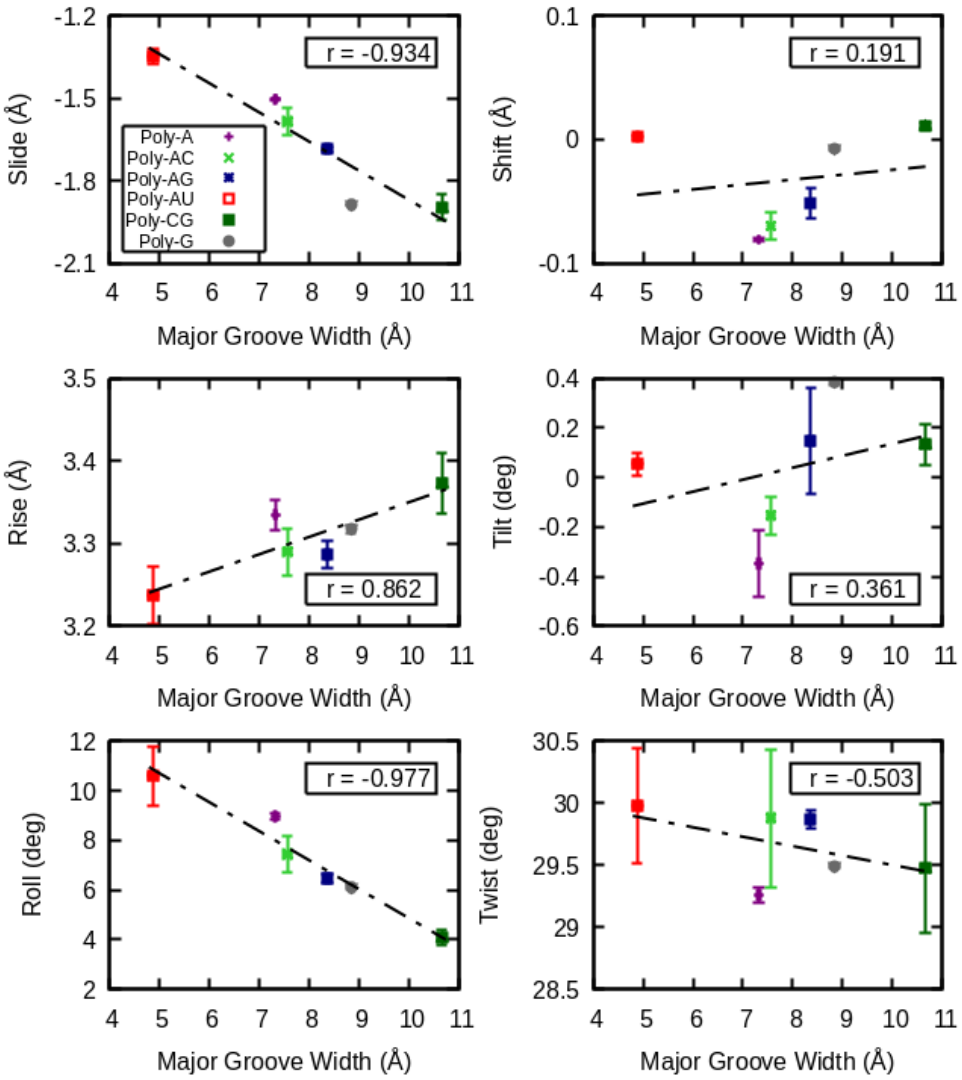

Figure S5. Base pair step parameters and correlations with the major groove width measured using Curves+. The mean values of the base pair step parameters were computed and represented as a function of the mean value of the major groove width in the same way as done with the helical parameters in **Fig. S4**.

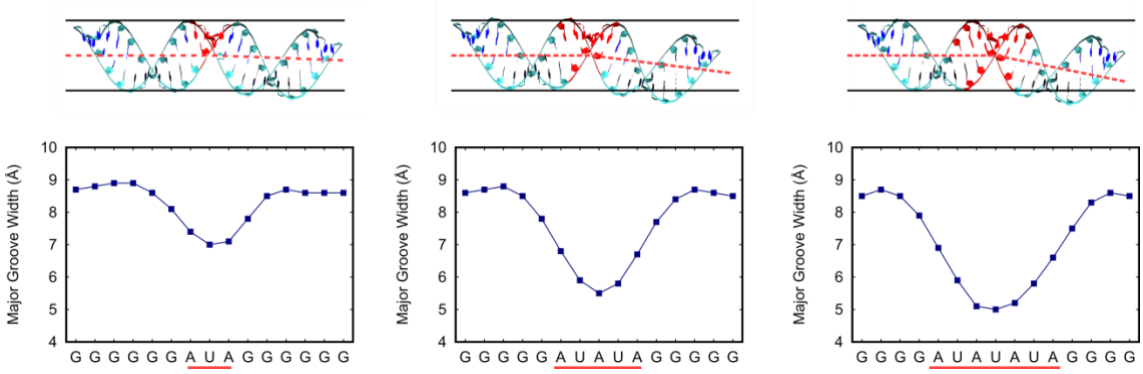

Figure S6. Major groove width and average structures of AU-3, AU-5 and AU-7 molecules, see **Table 1** main text. Data analysis and representation details are similar to **Fig. 2a-c**, main text.

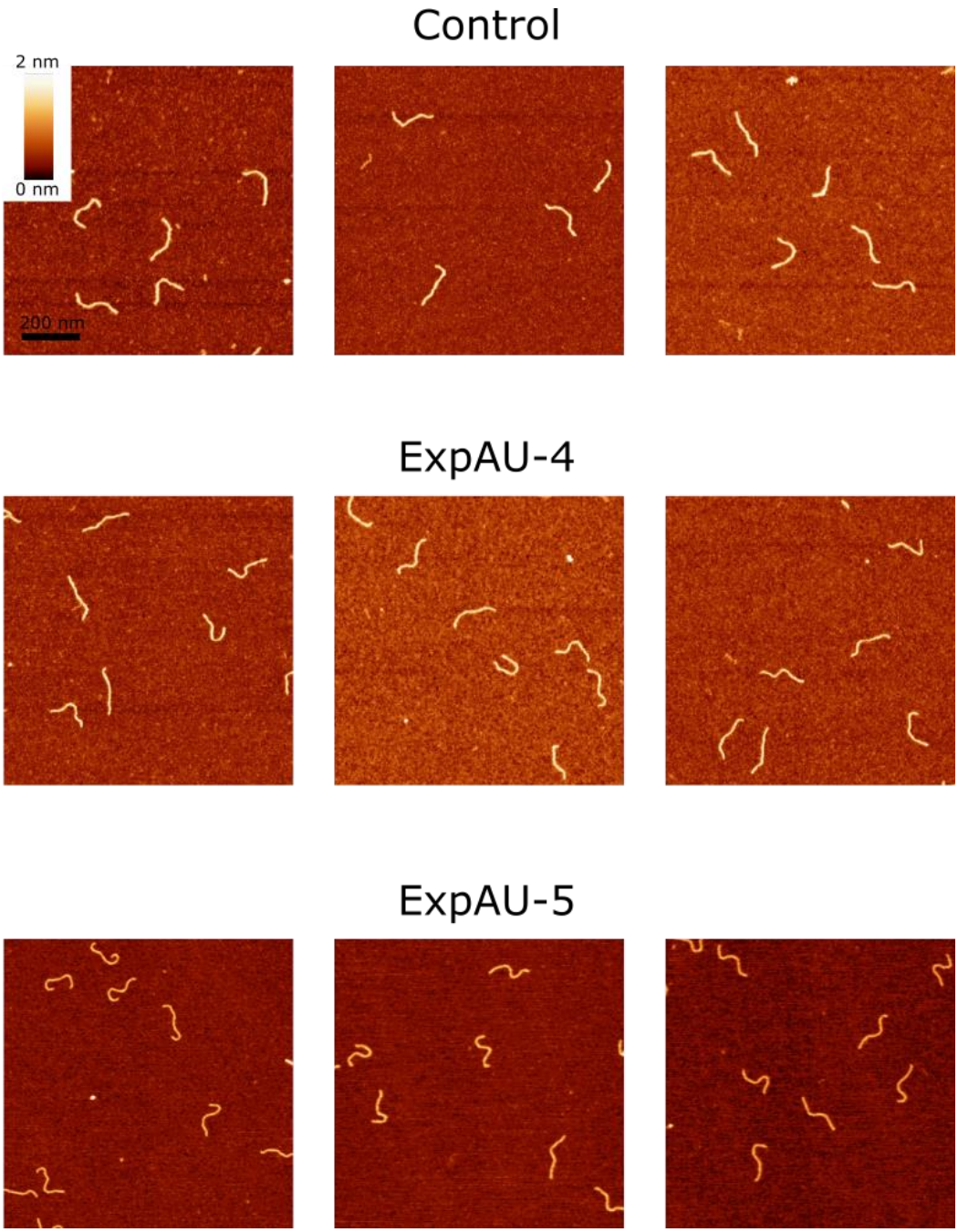

Figure S7. Representative AFM images of the Control and AU-tracts molecules. Z- and xy-scales are the same for all the images.

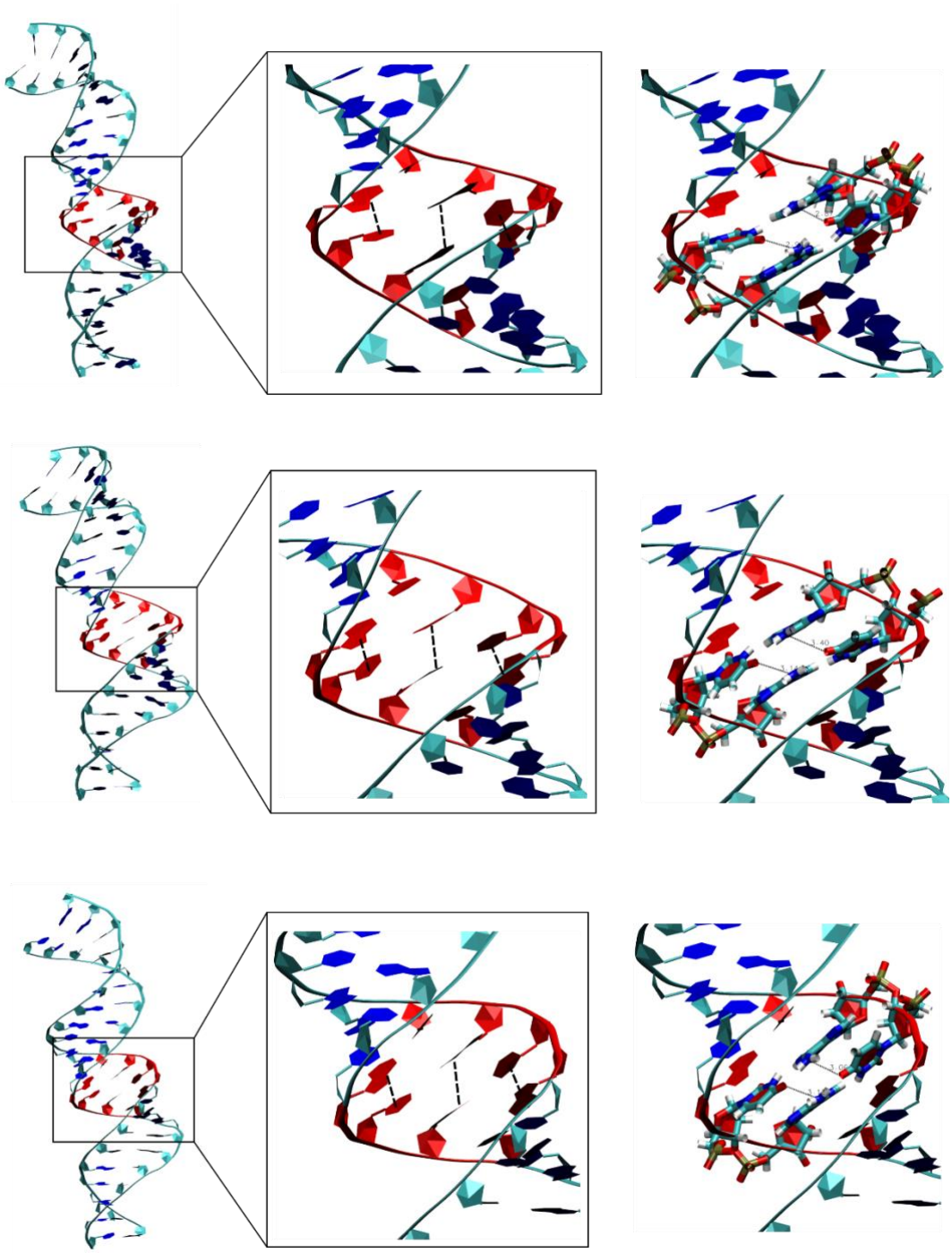

Figure S8. Snapshots of the trajectory of Seq. 3, which contains a central AU-tract (marked in red). Left, snapshot of the entire dsRNA molecule showing bending events. Center, zoom into the indicated region showing two adenines in a UpA step in close proximity, potentially engaging in intrastrand stacking interactions (dotted lines); and the highly propeller twisted configuration of the base pairs of that UpA step. Propeller twisting could further enhance interstrand interactions of the uracils with their adjacent adenines (dotted lines). Right, atomistic view on UpA step, showing that propeller twisting could also promote the formation of non-canonical hydrogen bonds, e.g. between the O4 atom of uracil and the N6 atom of a neighboring adenine.

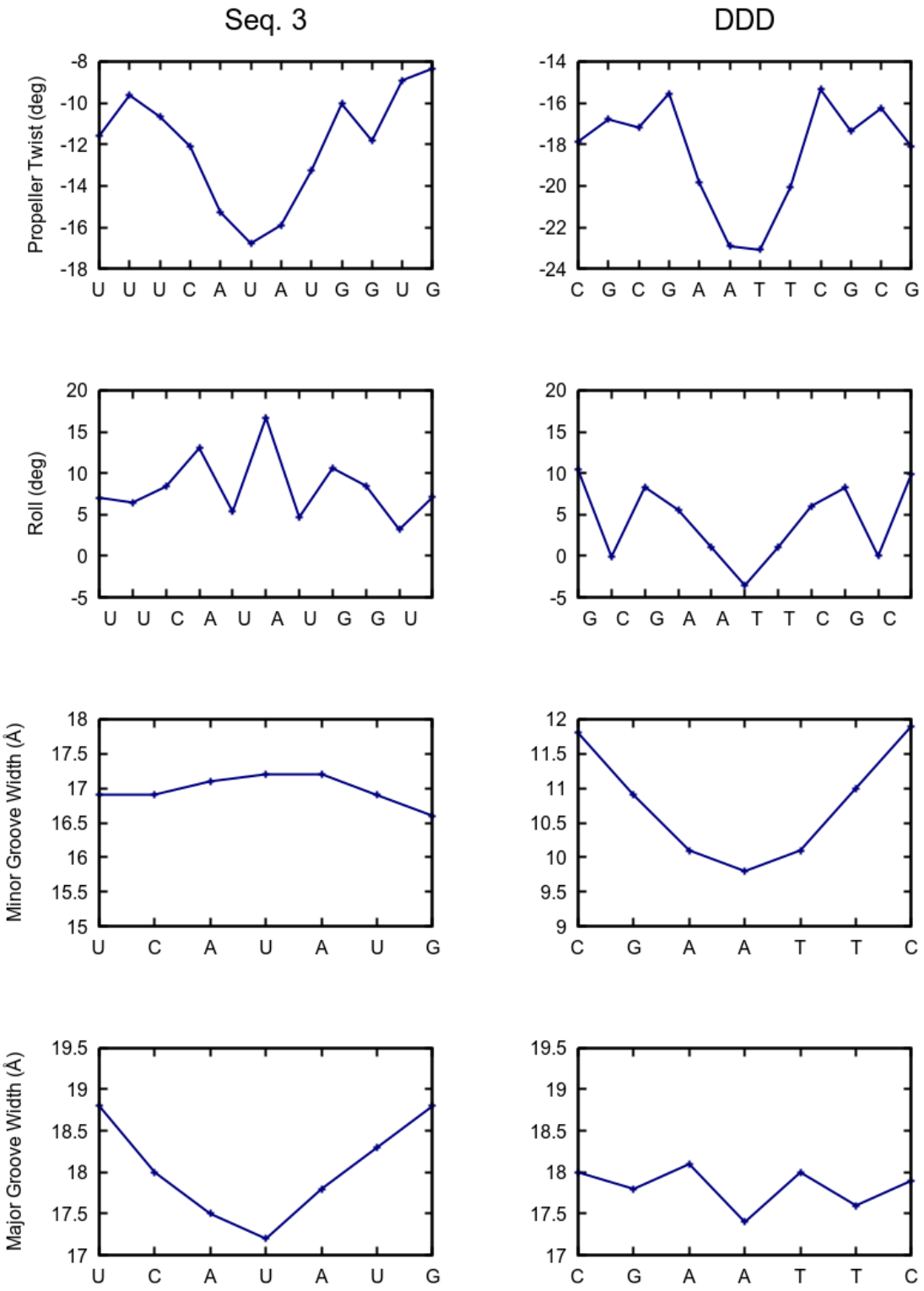

Figure S9. Comparison of structural parameters of a DNA A-tract and an RNA AU-tracts using 3DNA. The structural parameters shown in **Fig. 6**, main text, were computed using the alternative software 3DNA. The results obtained are very similar to the ones from Curves+ (**Fig. 6**).

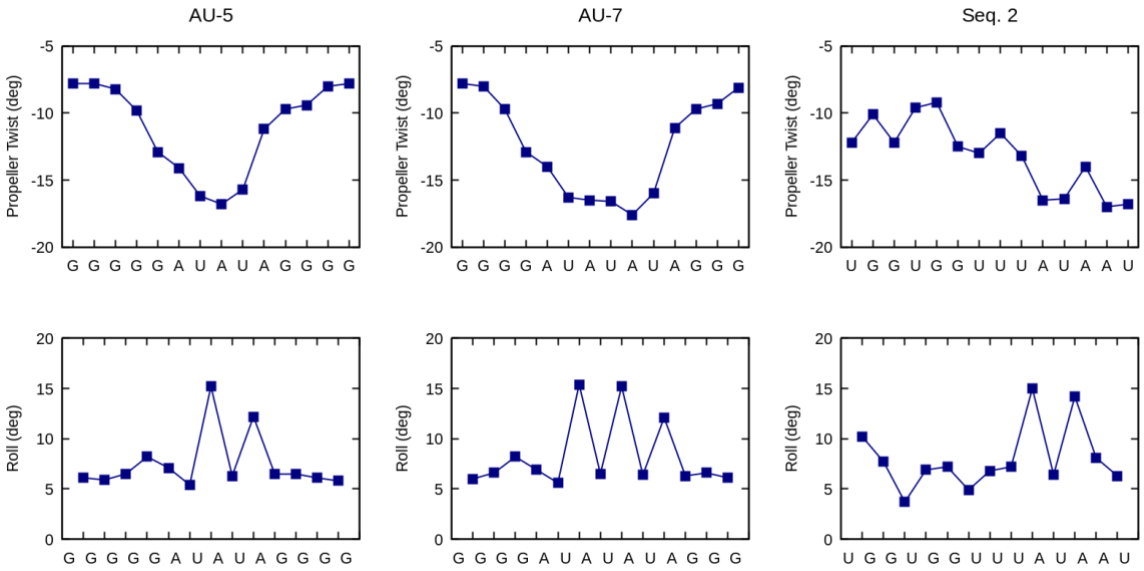

Figure S10. Propeller twist and roll profiles for three different sequences containing an AU-tract: AU-5, AU-7 and Seq. 2. Propeller twist and roll values were obtained by analyzing the average structure of the molecules over the 1 μs simulation time using the software Curves+.

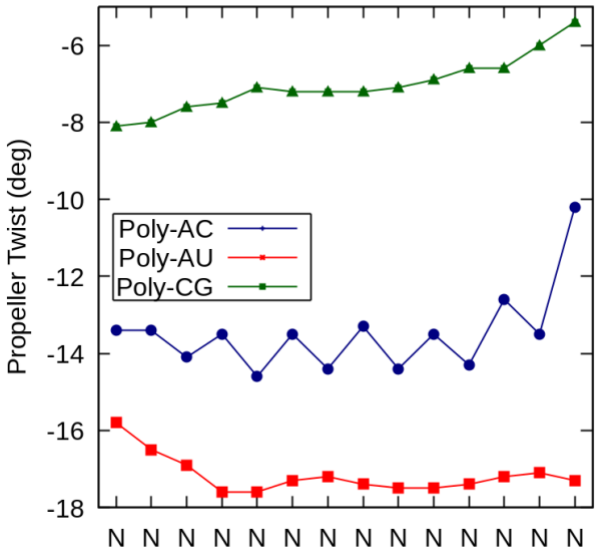

Figure S11. Propeller twist angles for benchmark sequences consisting on alternating purine-pyrimidine. Data were computed and represented as in **Fig. S10.**

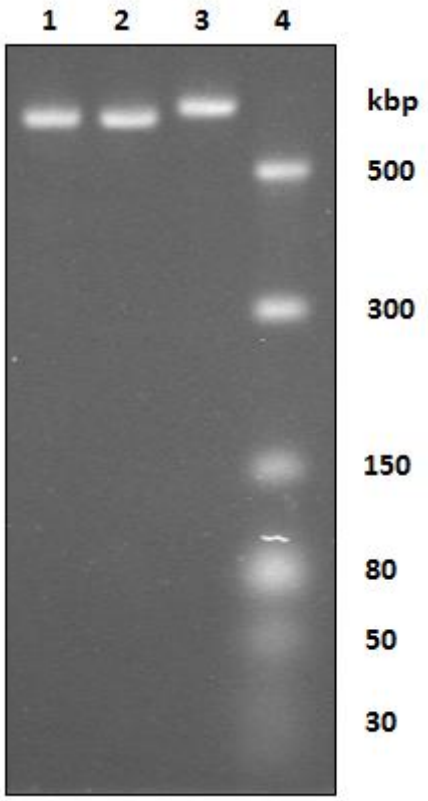

Figure S12. Migration of control, ExpAU-4 and ExpAU-5 dsRNA constructs in agarose gels. 2% agarose gel electrophoresis of 612 bp-dsRNA AFM constructs (control DNA, lane 1, and ExpAU-4, lane 2) and 624 bp-dsRNA ExpAU-5 (lane 3) were run (40 ng/lane) in 1xTBE buffer. Lane 4 corresponds to the dsRNA ladder (NEB). The electrophoresis was run at room temperature for 100 min at 60 V (4V/cm) in the absence of intercalator and the gel was later stained with SYBR®Safe (Invitrogen). AU-tracts showed no anomalous migration compared to the control.

3. Supporting tables

Table S1. dsRNA fragments used in this work. Sequences are written from the 5' to the 3' end omitting the complementary strand.

Fragment	Size (bp)	Sequence	GC-Content
Control dsRNA	612	GGGUACCUAGUAACCAUUCAGGAACGCAACCGCAGCUUAGACCAAAACAGGAAGC UAUGGGCCUGCUUAGGUGACGUCUCUCGUCAGGUUGAAUGGCAUGGUCGCUGGCU GGAUGCAGAAAGCUGGAAGUGUGUGUUUACCGCAGCAUUAAAGCAGCAGGAUGUU GUUCCUAACCUUGCCGGGAAUGGCUUUGUGGUAAUAGGCCAGUCAACCAGCAGGA UGCGUGUAGGCGAAUUUGCGGAGCUAUUAGAGCUUAUACAGGCAUUCGGUACAGA GCGUGGCGUUAAGUGGUCAGACGAAGCGAGACUGGCUCUGGAGUGGAAAGCGAGA UGGGGAGACAGGGCUGCAUGAUAAAUGUCGUUAGUUUCUCCGGUGGCAGGACGUC AGCAUAUUUGCUCUGGCUAAUGGAGCAAAAGCGACGGGCAGGUAAAGACGUGCAU UACGUUUUCAUGGAUACAGGUUGUGAACAUCCAAUGACAUAUCGGUUUGUCAGGG AAGUUGUGAAGUUCUGGGAUAUACCGCUCACCGUAUUGCAGGUUGAUAUCAACCC GGAGCUUGGACAGCCAAAUGGUUAUACGGUAUGGGAACCAAAGGAUAUUCAGACG GGUACCC	50.3 %
ExpAU-4 dsRNA	612	GGGUACCauauGGUCGAGauauGGUUGGGauauGAAGCUUauauGUAACCCauau CCUGUGGauauGGACUCCauauGGUUCCCauauGGGUGUGauauCAAGGGCauau GUUGGACauauGUGGACCauauGGGGAACauauCCGAACCauauGGGUUCUauau GGUCGAGauauGGUUGGGauauGAAGCUUauauGUAACCCauauCCUGUGGauau GGACUCCauauGGUUCCCauauGGGUGUGauauCAAGGGCauauGUUGGACauau GUGGACCauauGGGGAACauauCCGAACCauauGGGUUCUauauGGUCGAGauau GGUUAACauauCCGAACCauauGGGUUCUauauGGUCGACauauCUACGUAauau CCAGGAGauauGGUUAGGauauCCUACCGauauGGAGUCUauauGGGGGUAauau CCAGGAGauauGGUUAGGauauCCUACCGauauGGAGUCUauauGGGGGUAauau CCAGGAGauauGGUUAGGauauCCUACCGauauGGAGUCUauauGGGGGUAauau CCUCGAGauauGAGGCCUauauGGCCUACauauGGCUAGCauauGUCGCGAauau GGUACCC	42.3 %
ExpAU-5 dsRNA	624	GGGUACCuauauGUCGAGuauauGUUGGGuauauGAAGCUuauauGUACCCuaua uCCGUGGuauauGGACCCuauauGGUUCCuauauGCGUGGuauauCAGGGCuaua uGUGGACuauauGUGGCCuauauGGGAACuauauCCGACCuauauGGGUUCuaua uGUCGAGuauauGUUGGGuauauGAAGCUuauauGUACCCuauauCCGUGGuaua uGGACCCuauauGGUUCCuauauGCGUGGuauauCAGGGCuauauGUAGACuaua uGUGGCCuauauGGGGGGuauauGAAGCUuauauGUACCCuauauCCGUGGuaua uGGACCCuauauGGUUCCuauauGCGUGGuauauCAGGGCuauauGUGGACuaua uGUGGCCuauauGGGAACuauauCCGACCuauauGGGUUCuauauGUCGACuaua uGCCGGCuauauCCAGGGuauauGGUUGGuauauCCUGCGuauauGGAGUCuaua uGCGGGCuauauCCAGGGuauauGGUUGGuauauCCUGCGuauauGGAGUCuaua uGCGGGCuauauCUCGAGuauauGGCGCCuauauGGCCUCuauauGCUAGCuaua uGUCGCGuauauGGUACCC	41.5 %

Table S2. Pucker angles of the benchmark sequences. Values of the sugar pucker angles were computed for the nucleotides of the benchmark sequences and were averaged over the nucleotides of the same strand (Columns 2 & 3) and over all the nucleotides (Column 4). Errors are the standard deviation. For comparison, we include the sugar pucker values reported in Ref. [30] main text, where similar dsRNA sequences were simulated; and the canonical value of the pucker angle for RNA, which is ~18 deg. Values of the sugar pucker angle are given in degrees.

Sequence	Strand 1	Strand 2	Mean	Values from Ref. [30]	RNA canonical (C3'-endo)
Poly-A	13.5 (0.3)	21.2 (1.1)	17.4 (4)	18.3	18
Poly-AC	16.6 (2.6)	15.5 (2.9)	16.1 (2.8)	16.2	18
Poly-AG	12.20 (0.27)	19.0 (1.7)	15 (4)	16.0	18
Poly-AU	16.9 (1.6)	16.8 (1.5)	16.9 (1.6)	17.8	18
Poly-CG	14.2 (2.6)	14 (3)	14.4 (2.8)	14.6	18
Poly-G	11.81 (0.10)	17.63 (0.11)	15 (3)	15.2	18

Table S3. DNA oligonucleotides used in this work. Sequences are written from the 5' to the 3' end. The T7 RNA polymerase promoter sequence was underlined.

fragment	oligonucleotide	Sequence
control dsRNA	114.F RNA control 612	GCGTAAGTGGTACCTAGTAACCATTCAGGAACGC
	113.R RNA control 1316	GCGTAAGTGGTACCCGTCTGAATATCCTTTGGTTC
98 and 99	98.5P-F-T7	(Pho)TAATACGACTCACTATAGGGTACCatatGGTCGACatatGTACGTAatatCCTCGAGatatGAGGCCTatatGGCCTACatatGGCTAGCatatGTCGCGAatatGGTACCC
	99.5P-R-T7	(Pho)GGGTACCatatTCGCGACatatGCTAGCCatatGTAGGCCatatAGGCCTCatatCTCGAGGatatTACGTACatatGTCGACCatatGGTACCCTATAGTGAGTCGTATTACTAG
100 and 101	100.5P-F-XhoI-blunt	(Pho)TCGAGatatGGTTAACatatCCGAACCatatGGGTTCTatatGGTCGACatatCTACGTAatatCCAGGAGatatGGTTAGGatatCCTACCGatatGGAGTCTatatGGGG
	101.5P-R-XhoI-blunt	(Pho)CCCCatatAGACTCCatatCGGTAGGatatCCTAACCatatCTCCTGGatatTACGTAGatatGTCGACCatatAGAACCCatatGGTTCGGatatGTTAACCatatC
102 and 103	102.5P-F-blunt-blunt	(Pho)GGGatatGAAGCTTatatGTAACCCatatCCTGTGGatatGGACTCCatatGGTTCCCatatGGGTGTGatatCAAGGGCatatGTTGGACatatGTGGACCatatGGGG
	103.5P-R-blunt-blunt	(Pho)CCCCatatGGTCCACatatGTCCAACatatGCCCTTGatatCACACCCatatGGGAACCatatGGAGTCCatatCCACAGGatatGGGTTACatatAAGCTTCatatCCC
121 and 122	121.5P-F-T7	(Pho)TAATACGACTCACTATAGGGTACCtatatGTCGACtatatGCCGGCtatatCTCGAGtatatGGCGCCtatatGGCCTCtatatGCTAGCtatatGTCGCGtatatGGTACCC
	122.5P-R-T7	(Pho)GGGTACCatataCGCGACatataGCTAGCatataGAGGCCatataGGCGCCatataCTCGAGatataGCCGGCatataGTCGACatataGGTACCCTATAGTGAGTCGTATTACTAG
123 and 124	123.5P-F-XhoI-blunt	(Pho)TCGAGtatatGTTAACtatatCCGACCtatatGGGTTCtatatGTCGACtatatGCCGGCtatatCCAGGGtatatGGTTGGtatatCCTGCGtatatGGAGTCtatatGCG
	124.5P-R-XhoI-blunt	(Pho)CGCatataGACTCCatataCGCAGGatataCCAACCatataCCCTGGatataGCCGGCatataGTCGACatataGAACCCatataGGTCGGatataGTTAACatataC
125 and 126	125.5P-F-blunt-blunt	(Pho)GGGtatatGAAGCTtatatGTACCCtatatCCGTGGtatatGGACCCtatatGGTTCCtatatGCGTGGtatatCAGGGCtatatGTGGACtatatGTGGCCtatatGGG
	126.5P-R-blunt-blunt	(Pho)CCCatataGGCCACatataGTCCACatataGCCCTGatataCCACGCatataGGAACCatataGGGTCCatataCCACGGatataGGGTACatataAGCTTCatataCCC