Supplementary Material

Movie S1: https://drive.google.com/open?id=0B-CNvwEZiM3rMl9WcnQ1dlZnbU0
Movie S2: https://drive.google.com/open?id=0B-CNvwEZiM3rOHZFbVJkNEhUeFU
Movie S3: https://drive.google.com/open?id=0B-CNvwEZiM3rXy16UWs5cVJTSEk
Movie S4: https://drive.google.com/open?id=0B-CNvwEZiM3rYlIzQVNSa2Z1dkk
Movie S5: https://drive.google.com/open?id=0B-CNvwEZiM3rS25XWTlPTTJNbm8
Movie S6: https://drive.google.com/open?id=0B-CNvwEZiM3rRFkyVThnb1hreXM
Movie S7: https://drive.google.com/open?id=0B-CNvwEZiM3rc3dEdjlkN1cwc1E
Movie S8: https://drive.google.com/open?id=0B-CNvwEZiM3rSGd2M2xOck1LaEU
Movie S9: https://drive.google.com/open?id=0B-CNvwEZiM3rWW1lQm5kOWMwOUU
Movie S10: https://drive.google.com/open?id=0B-CNvwEZiM3rMkxOZFlUc0JkZWc
Movie S11: https://drive.google.com/open?id=0B-CNvwEZiM3rYWJFMzBFc2EwRzA
Movie S12: https://drive.google.com/open?id=0B-CNvwEZiM3rSVhyVjJsdm5Vb3c

Supplementary note 1: In-vitro transcription
Supplementary note 2: RNA-seq
Supplementary note 3: AFM
Supplementary note 4: Cryo-EM
Supplementary note 5: Random RNA control
Supplementary note 6: AFM of HOTAIR functional modules
Supplementary note 7: DNA synthesis
Supplementary note 8: Molecular measurements and modeling
Supplementary note 9: Triplexator predictions and CD spectroscopy
Supplementary note 10: RNA joint analysis
Supplementary note 11: Analysis of computational predictions

Supplementary note 1: In-vitro transcription

>HOTAIR

GACUCGCCUGUGCUCUGGAGCUUGAUCCGAAAGCUUCCACAGUGAGGACUGCUCCGUGGGGGUAA
GAGAGCACCAGGCACUGAGGCCUGGGAGUUCCACAGACCAACACCCCUGCUCCUGGCGGCUCCCACCCGG
GGCUUAGACCCUCAGGUCCCUAAUAUCCCGGAGGUGCUCUCAAUCAGAAAGGUCCUGCUCCGCUUCGCAG
UGGAAUGGAACGGAUUUAGAAGCCUGCAGUAGGGGAGUGGGGAGUGGAGAGAGGGAGCCCAGAGUUACAG
ACGGCGGCGAGAGGAAGGAGGGGCGUCUUUAUUUUUUUAAGGCCCCAAAGAGUCUGAUGUUUACAAGACC
AGAAAUGCCACGGCCGCGUCCUGGCAGAGAAAAGGCUGAAAUGGAGGACCGGCGCCUUCCUUAUAAGUAU
GCACAUUGGCGAGAGAAUUAAGUGCUGCAACCUAAACCAGCAAUUACACCCAAGCUCGUUGGGGCCUAAG
CCAGUACCGACCUGGUAGAAAAAGCAACCACGAAGCUAGAGAGAGAGCCAGAGGAGGGAAGAGAGCGCCA
GACGAAGGUGAAAGCGAACCACGCAGAGAAAUGCAGGCAAGGGAGCAAGGCGGCAGUUCCCGGAACAAAC
GUGGCAGAGGGCAAGACGGGCACUCACAGACAGAGGUUUAUGUAUUUUUAUUUUUUAAAAUCUGAUUUGG
UGUUCCAUGAGGAAAAGGGAAAAUCUAGGGAACGGGAGUACAGAGAGAAUAAUCCGGGUCCUAGCUCGCC
ACAUGAACGCCCAGAGAACGCUGGAAAAACCUGAGCGGGUGCCGGGGCAGCACCCGGCUCGGGUCAGCCA
CUGCCCCACACCGGGCCCACCAAGCCCCGCCCCUCGCGGCCACCGGGGCUUCCUUGCUCUUCUUAUCAUC
UCCAUCUUUAUGAUGAGGCUUGUUAACAAGACCAGAGAGCUGGCCAAGCACCUCUAUCUCAGCCGCGCCC
GCUCAGCCGAGCAGCGGUCGGUGGGGGGACUGGGAGGCGCUAAUUAAUUGAUUCCUUUGGACUGUAAAAU
AUGGCGGCGUCUACACGGAACCCAUGGACUCAUAAACAAUAUAUCUGUUGGGCGUGAGUGCACUGUCUCU
CAAAUAAUUUUUCCAUAGGCAAAUGUCAGAGGGUUCUGGAUUUUUAGUUGCUAAGGAAAGAUCCAAAUGG
GACCAAUUUUAGGAGGCCCAAACAGAGUCCGUUCAGUGUCAGAAAAUGCUUCCCCAAAGGGUUGGCAGUG
UGUUUUGUUGGAAAAAAGCUUGGGUUAUAGGAAAGCCUUUCCCUGCUACUUGUGUAGACCCAGCCCAAUU
UAAGAAUUACAAGGAAGCGAAGGGGUUGUGUAGGCCGGAAGCCUCUCUGUCCCGGCUGGAUGCAGGGGAC
UUGAGCUGCUCCGGAAUUUGAGAGGAACAUAGAAGCAAAGGUCCAGCCUUUGCUUCGUGCUGAUUCCUAG
ACUUAAGAUUCAAAAACAAAUUUUUAAAAGUGAAACCAGCCCUAGCCUUUGGAAGCUCUUGAAGGUUCAG
CACCCACCCAGGAAUCCACCUGCCUGUUACACGCCUCUCCAAGACACAGUGGCACCGCUUUUCUAACUGG
CAGCACAGAGCAACUCUAUAAUAUGCUUAUAUUAGGUCUAGAAGAAUGCAUCUUGAGACACAUGGGUAAC
CUAAUUAUAUAAUGCUUGUUCCAUACAGGAGUGAUUAUGCAGUGGGACCCUGCUGCAAACGGGACUUUGC
ACUCUAAAUAUAGGCCCCAGCUUGGGACAAAAGUUGCAGUAGAAAAAUAGACAUAGGAGAACACUUAAAU
AAGUGAUGCAUGUAGACACAGAAGGGGUAUUUAAAAGACAGAAAUAAUAGAAGUACAGAAGAACAGAAAA
AAAAUCAGCAGAUGGAGAUUACCAUUCCCAAUGCCUGAACUUCCUCCUGCUAUUAAGAUUGCUAGAGAAU
UGUGUCUUAAACAGUUCAUGAACCCAGAAGAACGCAAUUUCAAUGUAUUUAGUACACACACAGUAUGUAU
AUAAACACAACUCACAGAAUAUAUUUUCCAUACAUUGGGUAGGUAUGCACUUUGUGUAUAUAUAAUAAUG
UAUUUUCCAUGCAGUUUUAAAAUGUAGAUAUAUUAAUAUCUGGAUGCAUUUUCAAAAAAAAAA

pLZRS-HOTAIR and pCDNA3-HOTAIR were a kind gift from Prof. Howard Chang. pJ-HOTAIR was purchased from DNA2.0. All plasmids were fully sequenced to ensure accuracy. To produce HOTAIR transcripts for in vitro transcription (IVT), DNA templates were either plasmids linearized with EcoRI restriction enzyme (NEB), or PCR amplicons. PCR primers for the amplification of DNA templates for IVT were synthesized by Integrated DNA Technologies and HPLC-purified.

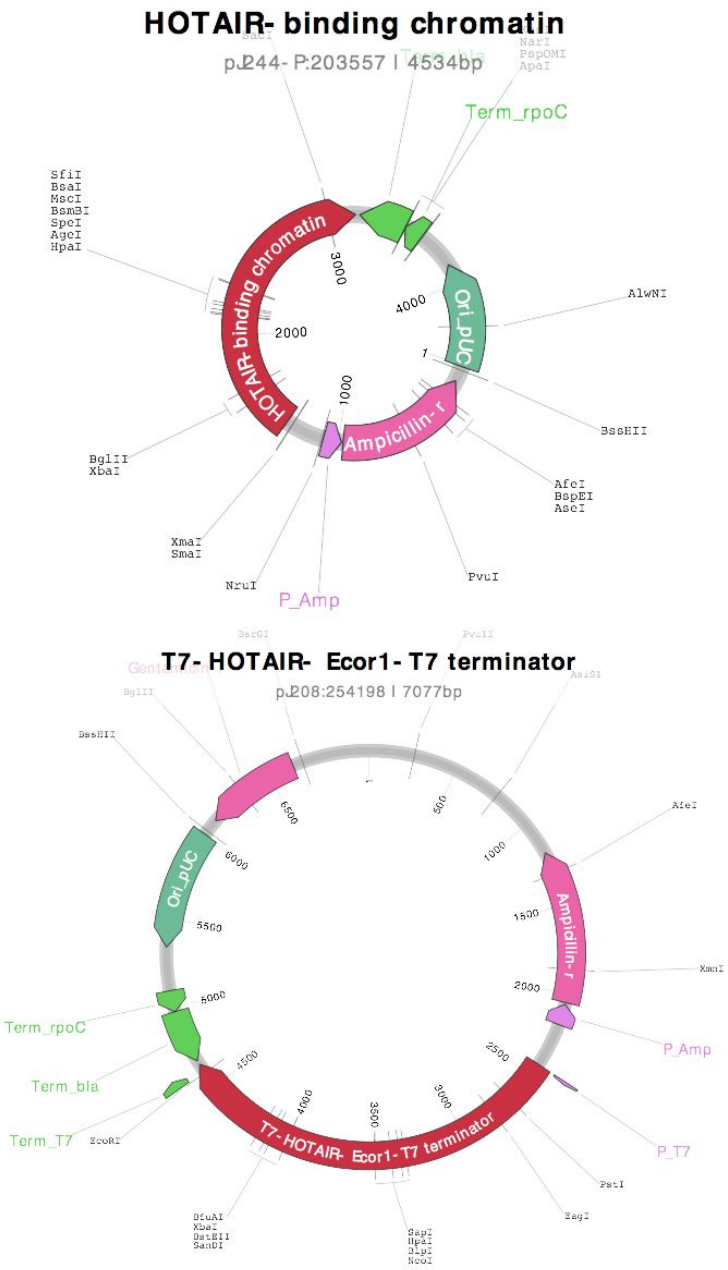


Figure S1. plasmid maps

DNA templates for IVT of whole HOTAIR, HOTAIR-300, HOTAIR-1200, and HOTAIR-658 RNA, were amplified using the following primers:

>HOTAIR_Fwd
TAATACGACTCACTATAGGG

HOTAIR_Rev
TTTTTTTTTTGAAAATGCATCCAGATATTAATATAT

>HOTAIR300_Fwd

TAATACGACTCACTATAGGG

>HOTAIR300_Rev

CGCCCCTCCTTCCTCTCGC

>HOTAIR1200_Fwd

GAGTACTAATACGACTCACTATAGGGATGGAGTACTAATACGACTCACTATAGGGATGTCTTTATTTTTT

TAAGGCCCCAAAGAG

>HOTAIR1200_Rev

TTTCACTTTTAAAAATTTGTTTTTGAATCTTAAGTCTAGGAATCAGCACGAAG

>HOTAIR658_Fwd

GAGTACTAATACGACTCACTATAGGGATGGCCTTTGCTTCGTGCTGATT

>HOTAIR658_Rev

TTTTTTTTTTGAAAATGCATCCAGATATTAATATATCTACATTTTAAAACTGCATGGAAAATAC

DNA templates for IVT were generated by PCR with high-fidelity DNA polymerase (KAPA HiFi HotStart PCR Kit, Kappa Biosystems). All products were confirmed by agarose gel electrophoresis before IVT. IVT was carried out using two separate kits (from New England Biolabs and Megascript). IVT was carried out for 3 h at 37 °C and followed by DNA template digestion (using DNase included in the kits). RNA was purified using MegaClear kit. Elution was carried out with ultrapure RNase free water in 37 °C. Samples were then diluted into nuclear-like buffer for imaging or further analysis (see **Supplementary note 3)**. All products were confirmed by agarose gel electrophoresis.

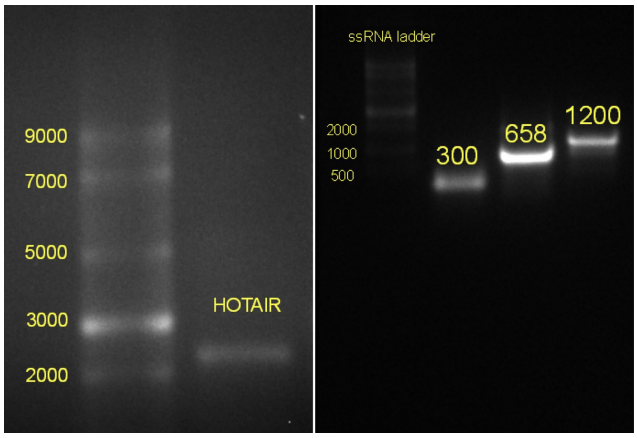
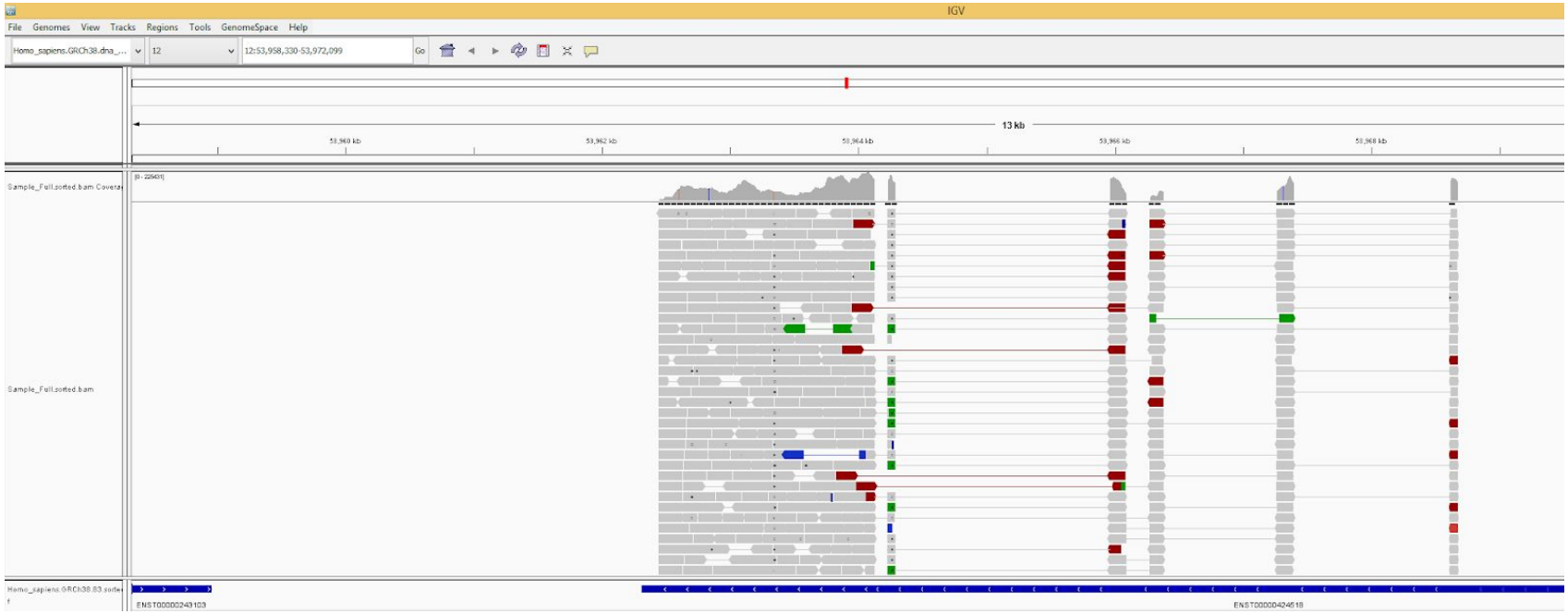


Figure S2. Gel electrophoresis of IVT products. Left, HOTAIR. Right, HOTAIR domains (300 nt, 658 nt, 1200 nt from left to right).

Supplementary note 2: RNA-seq

RNA-seq was performed at the Nancy and Stephen Grand Israel National Center for Personalized Medicine (G-INCPM) at the Weizmann Institute of Science. Library preparation was done using in-house protocols. Briefly, purified RNA samples from IVT of whole HOTAIR and HOTAIR-300 were fragmented and used as templates for cDNA generation. Then, end repair, a base addition, adapter ligation, and PCR amplification were performed. Quality of libraries was evaluated by Qubit and TapeStation. Sequencing library was constructed with barcodes to allow multiplexing of two samples from two bands in one run. 700,000-900,000 paired end 150x2 reads were sequenced on Illumina MiSeq. Next, reads were mapped to a human reference genome, GRCh38, using BWA-MEM v0.75 [1] with default parameters and coverage was calculated using Samtools v0.19 mpileup [2]. Analysis showed that both HOTAIR and HOTAIR-300 fully map to CRCh38.

Figure S3 (next page). RNA-seq results mapping HOTAIR to the HOTAIR gene. Upper track indicates the coverage, i.e. number of reads, that cover HOTAIR's genomic loci. Gray - a read pair has a reasonable insert size. Red - a read pair far from each other (ie on two different 'exons'). Blue- insert size too small. Green-inverted orientation of read pairs.


IGV
File Genomes View Tracks Regions Tools GenomeSpace Help
Homo_sapiens.GRCh38.dna_...
12
12:53,958,330-53,972,099
Go
13 kb
53,960 kb
53,962 kb
53,964 kb
53,966 kb
53,968 kb
Sample_Full.sorted.bam Coverage
Sample_Full.sorted.bam
Homo_sapiens.GRCh38.83.sorted
ENST00000243103
ENST00000424618

Supplementary note 3: AFM

Mica discs (Ø 8 mm, Ted Pela (1x3" sheets, Grade V1, Product Nr. 56)) were glued to a microscope slide. Immediately before use, the mica substrates were cleaved to produce atomically flat surfaces suitable for high-resolution imaging. 50 uL of 0.01%, sterile-filtered PLO (Poly-L-ornithine) solution (Sigma) was deposited on the mica for 10 minutes and then removed with a pipette and blow dried with N2 stream. Plastic rings were sonicated for 10 minutes ,cleaned with ethanol and glued to the slide. Prior to deposition on substrate, RNA and HBDs molecules were incubated in filtered NLB (Nuclear-Like Buffer: 5mM NaCl, 140mM K^+ , 0.5 mM Mg^{+2}, 10^{-4} mM Ca^{+2}, pH=7.2) for 30 min at 37 °C. 20uL of the investigate sample were deposited on the PLO- coated mica for 90-120 seconds and then removed with a pipette. 40uL of NLB were used to rinse the mica, and the chamber was gently filled with 1mL of NLB.

Samples were analysed using a NanoWizard® ULTRA Speed AFM (JPK Instruments, Germany) mounted on an inverted optical microscope (Nikon Eclipse TE2000-U or Zeiss AxioObserver.A1), or equipped with a JPK TopViewOptics™. Samples were imaged in buffer at ambient temperature in amplitude-modulation or phase-modulation AC mode. Fast-scanning high-resonant ultra-short cantilevers (USC-F0.3-k0.3, NanoWorld, Switzerland) with a nominal resonance frequency of 300 kHz in air, spring constant of 0.3 N/m, reflective chromium/gold-coated silicon chip, and high-density carbon tips with a radius of curvature of 10 nm were used.

For the DNA origami squares measurements, 0.5 nM samples in 1xTAE, 12.5 mM $MgCl_2$ were deposited on freshly cleaved mica (as described above) for 15 minutes, washed and scanned via JPK Nanowizard III Bio-AFM (HyperDrive and AC modes using ultra-short cantilevers with force constant of 0.3 N/m)). Images were analyzed using Nanoscope analysis v1.5 and JPK Data Processing v4.3.26..

Image processing was done by FIJI. Fig. 2D is a 3D surface plot of Fig. 2E (right panel), using the FIJI 3D surface plot function. No additional processing was made on the image. The following 3D surface plot parameters: Grid size 1024, Smoothing 12.0, Lighting 0.22, Fire LUT, Z-scale 0.10, Min 0%, Max 100%. Link to the original image file:
https://drive.google.com/open?id=0B1AS1gHBwUrRZGgtbUotTzRBR2M

Supplementary note 4: Cryo-EM

Cryo-EM was performed at the Weizmann Institute of Science by standard in-house protocols. Briefly, 3.5 µL of HOTAIR solution (1.5 mg/ml) was applied to glow-discharged Quantifoil holey carbon grids (R1.2/1.3, 300 mesh). Grids were plunge-frozen in liquid ethane cooled by liquid nitrogen, using a Leica EM-GP plunger (4 s blotting time, 80% humidity). Grids were imaged at liquid nitrogen temperature on an FEI Tecnai TF20 electron microscope operated at 200 kV with a Gatan side entry 626 cryo-holder. Images were recorded on a K2 Summit direct detector (Gatan) mounted at the end of a GIF Quantum energy filter (Gatan). Images were collected in counting mode, at a calibrated magnification of 16,218 yielding a pixel size of 3.083 Å. The dose rate was set to 8.5 electrons per physical pixel per second and a total exposure time of 8 s. Each image was fractionated into 40 subframes of 0.2 s. Defocus range was 1.5-3.5 µm. All dose-fractionated images were recorded using an automated low dose procedure implemented in SerialEM [3]. Recorded image frames were subjected to whole image beam-induced motion correction using MOTIONCORR [4]. We used e2boxer in EMAN2[5] for particle picking and RELION[5,6] for 2D classification.

Importantly, 2D classification and averaging did not produce meaningful results because the particles showed high conformational variability. This made it impossible to solve the structure entirely by cryo-EM.

Supplementary note 5: Random RNA control

RANDOM is a sequence formed by randomly scrambling HOTAIR sequence to yield a sequence with the same length and base composition. RANDOM was synthesized using the same IVT systems used for HOTAIR, and scanned in the AFM using the same protocols and settings. RANDOM formed no distinct shapes, and in about half of the cases did not even fold into a shape.

>RANDOM

UAAUACGACUCACUAUAGGUGUUUUGCACUACCGUCAGGUACGUUAGUAUGCGUUCUUCCUUCCAGAGGU
AUGUGGCUGCGUGGUCAAAAGUGCGGCAUUCGUAUUUGCUCCUCGUGUUUACUCUCACAAACUUGACCUG
GAGAUCAAAGAGAUGCUUCUUGUGGAACUGGACAACGCAUCAACGCAACGGAUCUACGUUACAGCGUGCA
UAGUGAAAACGAAGUUGCUGACGACGAAAGCGACAUUGGGAUCUGUCUGUUGUCAUUCGCGAAAAACAUC
CGUUCACGAGGCGGACACUGAUUGACACGGUUUUGUAGAAGGUUAGGGGAAUAGGUUAAAUUGAGUGGCU
UAAAAAUGUUAUAUCUGGGAUUAAAGUGUAGUAAACUGUAAUUAACGGAGACGGUUUUAAGACAGGAGUU
CGCAAAAUCAAGCGGGGUCAUUACAACAGUUAUUCCUGAUGGUUUAGGCGUACAAUGUCCUGAAGAAUAU
UUAAGAAAAAAGCACCCCUCGUCGCCUAGAAUUACCUACCGCGGUCGACCAUACCUUCGAUUAUCGCGCC
CACUCUCCCAUUAGUCGGCAGAGGUGGUUGUGUUGCGAUAGCCCAGUAUGAUAUUCUAAGGCGUUACGCU
GAUGAAUAUUCUACAGAAUUGCCAUAGGCGUUGAACGCUACACGGACGAUACGAAUUUAUGUAUAGAGCG
GGUCAUCGAAAGGUUAUACUCUUGUAGUUAACAUGUAGCCCGGCCCUAUUAGUACAGCAGUGCCUUGAAU
GACAUUCUCAUUAUUAAAUUUUCUCUACAGCCAAACGACCAAGUGCAUUUCCACGGAGCGCGAUGGAGAU
UCAUUCACUCGGCAGCUCUGUAAUAGGGACUAAAAGAGUGAUGAUAAUCAUGAGUGCCGUGUUAUGAUGG
UGUCGGAACAAAGCGGUCUUACGGUCAGUCGUAUUCCUUCUCGAGUUCCGUCCAGUUAAGCGUGACACUC
CCAGUGUACCUGCAAACCGUGAUGGCUGUGCUUGGAGUCAAUCGCAUGUAGGAUGGUCUCCAGACACCGG
GGCACCAGUUUUCACGCCUAAAGCAUAAACGACGAGCAGUCAUGAAAGUCUUAGAACUGGACGUGCCGUU
UCUCUGCGAAUAAUACCUCAAGCUGUACCGUUAUUGCGCUGCUUAGAUGCAGUGCUGCUCUUAUCACAUU
UGUUUCGACGACUGCCGCCUUCGCUGUUUCCUUAGACACUUAACAAUAAGCGCUUUUUGUAGGCAGAGGC
ACCCCCUAUUAGUGGCUGCGCCAAAAUAUCUUCGGAUCCCCUUGUCCAAUCAAAUUGAUCGAAUUCUUUC
AUUUAAGACCCUAAUAUGACAUCAUUAGUGAUUAAAUGCCACUCCCAAAAUUCUGCCUAGAAAUGUUUAA
GUUCGCUCCACUAAAGUUGUUUAAAACGACUACUAAAUCCGCGUGAUAGGGGAUUUCAUAUUUAAUCUUU
UAUCGUAAGGAACAGCCGAUCUUAAUGGAUGGCCGCAGGUGGUAUGGAAGCUAUAAGCGCGGGUGAGAGG
GUAAUUAGGCGUGUUCACCUACACUACGCUAACGGGCGAUUCUAUAAGAUUGCACAUUGCGUCUACUUAU
AAGAUGUCUCAACGGCAUGCGCAACUUGUGAAGUGUCUACUAUCCUUAAACGCAUAUCUCGCACAGUAAC
UCCCCAAUAUGUGAGCAUCUGAUGUUGCCCGGGCCGAGUUAGUCUUGUGCUCACGGAACUUAUUGUAUGA
GUAGUGAUUUGAAAGAGUUGUCAGUUAGCUCGUUCAGGUAAUAGUUCCUCACACUACGUCAAAAUAAGAG
AACGGUCGUAACAUUAUCCGUGAUUUUCUCACUACUAUCAGUACUCACGACUCGAUUCUGCCGCAGCCAC
GUAUCGCCAGAAAGCCAGUCAGCAUUAAGGAGUGCUCUGAGCAGGACAACUCGCAUAGUGAGAGUUACAU
GUUCGUUGGGCUCUUCCGACACGAACCUCAGUUGGCCUACAUCCUACCUGAGGUCUGUGCCCCGGUGAUG
AGAAGUAUGCAUUUCGUUCUUGCAGCUUGUCAGUACUUUCAGAAUCAUGGCCUGCAUGGUAGAAUGACGC
UUAUAAUGGACUUCGACAUGGCAAUAACCCCCCGUUUCUACUUCAAGAGGAGAAAAGU

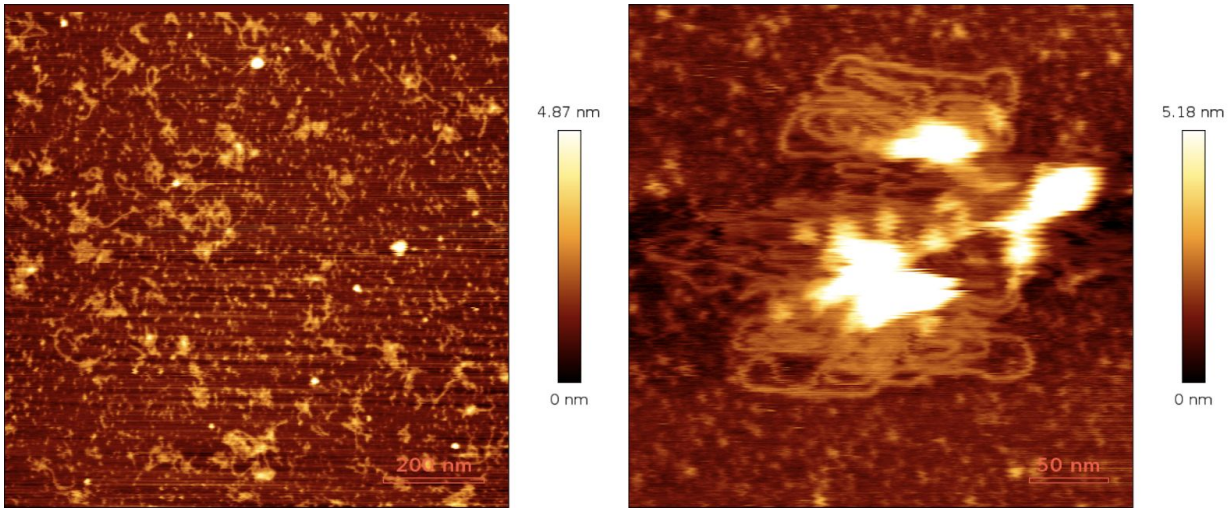


Figure S4. RANDOM RNA AFM scans.

Supplementary note 6: AFM of HOTAIR modules

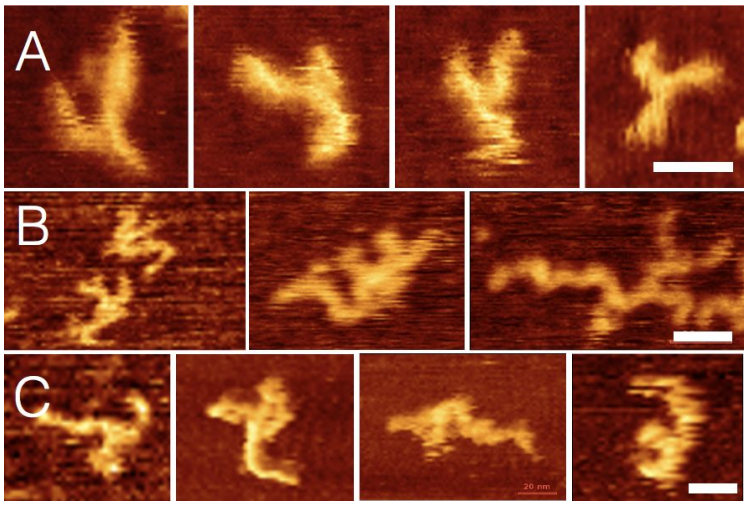


Figure S5. AFM scanning of the proposed functional modules of HOTAIR: 300 nt (A), 1200 nt (B), 658 nt (C) (bars = 15 nm).

Supplementary note 7: DNA synthesis

p-HBD was purchased from DNA2.0. Plasmids was fully sequenced to ensure accuracy. PCR primers for the amplification of DNA templates for IVT were synthesized by Integrated DNA Technologies. HBD1 was the most enriched sequence in a ChIRP[7] study, where HOTAIR was precipitated. HBD4 was not precipitated with HOTAIR.

>HBD1_chromosome13:56624796-56625228
TGTAGACTGAAAGAAAAGAGATTAAAAAAAGATACTCCATGTAAATGGAAACCATAAATCACCAAGAGTA
GCCACATGTCTACCAGATAAAACATATTTGACTCAAAAAACATAACAGAGACAAAGAAGTTTATTATATA
ATGATAAAGAAATCACTTTAGCAAGAGGACATAATAATTCTAAGCACACATGCAAGCAACACTAGAGCAC
TTGGATATATAAGGCAAATATTATTAGATCTAATGGGAGAGAAAACTCCAATACAAAAATATTTGGGGAC
TTTGACACACCCTTCTCAGCATTATAAAGTTCATCTAGAAAGATAATTAACAGAGCAATTTTGGCTTTAA
ACTACACTTTGGAACAAATAGAACTAACAGATATTTAAAGAACATTTTATCCAAAAACTACAGAATACAC
AGAATACACATTC

>HBD4_chromosome13:57000000-57000432
TTGCAAGTTCAATGTTAGATATTTGAAATGTTCTGATCTTTTTAAACAATTTGAATAGAAAGCAGACAAT
GTAGGTTAAAATTTTAATATAAATTGAAGGATCTAATGTAAGAAATTAAAATACTTTAAGCACCTAACTC
CTCCATGCCTGATATATAGAAGATGATAATTAAATGTTTGCTAAATGAATGCATTGACAACACAGTTTCT
CTTTTATTCCTAAAAGCATTTCTTTGGATAAGTACTACTACCATCAGAGCTCAGGTCATAAAATTACATC
CAGGAAGAAATTCTCAGAATGTTAAACTTGTCAGAATATACTGTATACTGCTGTCAAAGTTGATTCTCAG
GAAAATGCTGTGCCATTCCTTCTTTAACATCATGTTTCTTATGGCATATTACACCCAGAACTAGGCAGTC
ATATCAAATTAGA

HBD1 and HBD4 were amplified using the following primers:
>HBD1_Fwd
TGTAGACTGAAAGAAAAGAGATTAAAAA
>HBD1_Rev
GAATGTGTATTCTGTGTATTCTGTAG
>HBD4_Fwd
TTGCAAGTTCAATGTTAGATATTTGAAA
>HBD4_Rev
TCTAATTTGATATGACTGCCTAGTTC

HBD1 and HBD4 were generated by PCR with high-fidelity DNA polymerase HyFy High Fidelity mix (HyLabs) according to manufacturer's instructions. All products were confirmed by agarose gel electrophoresis. PCR products were then purified using Qiagen PCR purification kit.

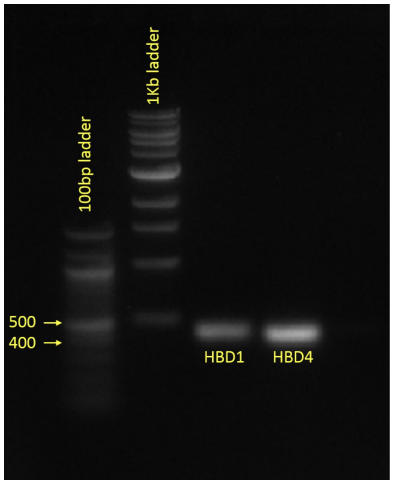


Figure S6. Gel image showing products

Supplementary note 8: Molecular measurements and modeling

In order to measure observed HOTAIR molecules, AFM images were first scaled to have similar resolution of 1px/nm, and then processed to improve clarity and traced using splines to build a tree structure that follows the observed geometry. Processing included conversion to grayscale, histogram auto-leveling and strong contrast filter, performed in ADOBE Photoshop. Tracing was done in SOLIDWORKS using a sketch with the true resolution image in the background. Spline length was later measured with the Measurement tool.

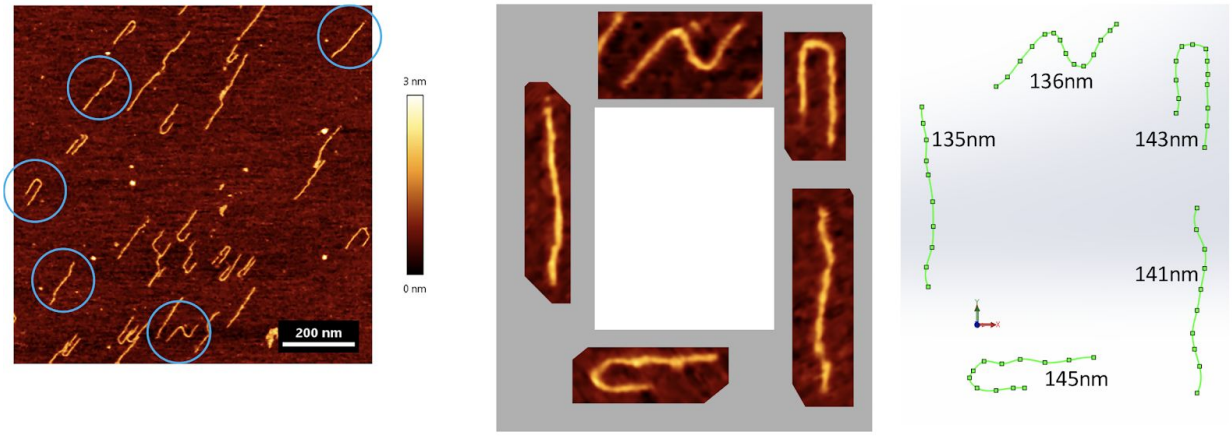


Figure S7. Measurements of size reference DNA (HBD1). On the left an imaged field of HBD1 dsDNA fragments. Right, selected objects and their measured traces.

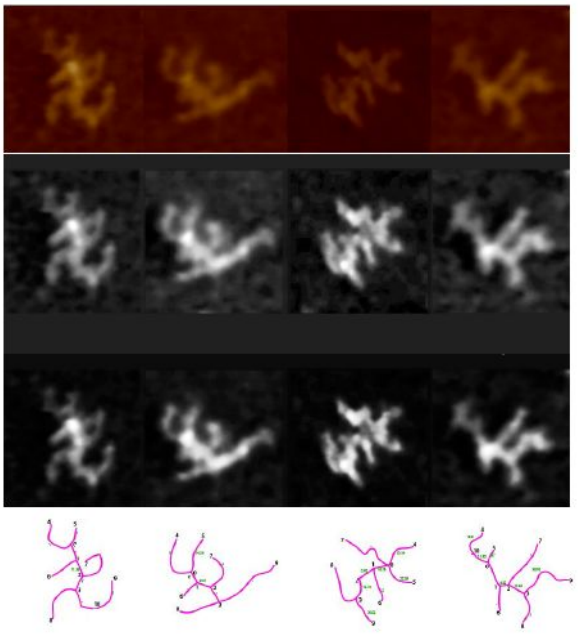


Figure S8. Image processing for dimension measurements of 4 selected HOTAIR molecules. Entire images were converted to grayscale and each specimen was autolevelled locally (row 2).

Next strong contrast curves were added to entire images (row 3). Row 4 shows the specimen traces for size measurements.

Visual structural model of HOTAIR

The aim was to create a three-dimensional computer generated model of the HOTAIR structure, proportional to scale and representative of the observed structure. The model was created and rendered using Autodesk Maya 2016 and mental ray renderer.

An A-DNA molecule (PDB ID: 1TTD) was downloaded and imported into Autodesk Maya through the molecular maya Protein Data Bank (PDB) import kit. Chain B of the dodecamer duplex was selected and scaled to 2.3 scale units (SU) wide and 2.86 SU long, and edited to smoothly connect to another subunit (each scale unit in maya was considered to be 1nm). This subunit was duplicated 140 times to generate a single strand DNA (404.4 SU long). The ssDNA was rigged using a lattice and joints. To create the dsDNA mesh according to the observed structure, a path was drawn, which can be adjusted in the three directions of space using control points, and assigned as as the IK spline handle of the mesh. Finally the structure was rendered in high quality using the mental ray renderer.

Supplementary note 9: Triplexator

Triplexator is a prediction algorithm that detects potential nucleic acid triple helices in genomic and transcriptomic data[8]. With the kind assistance of Dr. Buske we ran HBD1 and HOTAIR sequences in Triplexator to highlight potential triple helical structures, in an attempt to define the mechanism of observed physical interaction between these two molecules. One assumption that was made in this analysis is that the interaction is between genomic DNA in its dsDNA state, and HOTAIR through an ssRNA segment (i.e. most likely a loop).

Below is the output. The algorithm produced three scenarios, each beginning with a line describing the parameters on the list appearing at the beginning. At the end is the log file output including all the parameters used.

```
# Sequence-ID
TFO start
TFO end Duplex-ID
TTS start
TTS end Score
Error-rate
Errors
Motif
Strand
Orientation
Guanine-rate

seq2   538    548    seq1   287    297    9      0.1    d1     R      -      A      0.6
     5'- CACCCTTCTC -3'
TTS: 3'- GtGGGAAGAG -5'
         |*|||||||
TFO: 5'- GAGGGAAGAG -3'

seq2   1252   1262   seq1   8      18     9      0.1    t2     M      +      P      0.3
TFO: 5'- GTGTGTTTTG -3'
         ||*|||||||
TTS: 5'- GAAAGAAAAG -3'
     3'- CTTTCTTTTC -5'

seq2   1253   1266   seq1   8      21     12     0.077  t1     M      +      A      0.31
TFO: 3'- GGTTGTTTTGTGT -5'
         |*||||||||||
TTS: 5'- GAAAGAAAAGAGA -3'
     3'- CTTTCTTTTCTCT -5'


************************************************************
*** Triplexator - Finding nucleic acid triple helices     ***
***         (c) Copyright 2011 by Fabian Buske            ***
***     Comments, Bugs, Feedback: f.buske@uq.edu.au       ***
************************************************************
*** Version 1.3.2 (30/03/2012) SeqAn Revision: 1225
*** COMMAND:
>triplexator -ss RNA.fasta -ds DNA.fasta -po -l 10 -g 30 -of 1 -e 10 -o hali.tpx
************************************************************
*** INPUT:
- single-stranded file supplied : Yes
```

```
- duplex file supplied : Yes
-> search putative triplexes (matching triplex-forming oligonucleotides and target sites)
************************************************************
*** Output Options:
- output directory : ./
- output file : hali.tpx
- output format : 1 = extended Triplex (+Alignment)
- merge features : Yes
- report duplicate locations : No
- error reference : 0 = Watson strand (TTS)
************************************************************
*** Main Options:
- maximum error-rate : 10%
- maximum total error : 3
- minimum guanine content with respect to the target : 30%
- maximum guanine content with respect to the target : 100%
- minimum length : 10 nucleotides
- maximum length : 30 nucleotides
- maximum number of tolerated consecutive pyrimidine interruptions in a target: 1
- include GT-motif : Yes
- include GA-motif : Yes
- include TC-motif : Yes
- minimum guanine-percentage in anti-parallel mixed motif TFOs : 0%
- maximum guanine-percentage in parallel mixed motif TFOs : 100%
- number of consecutive matches required in a feature : 1
- longest match only : yes
- detect duplicates : 0 = off
- same sequence duplicates : on
************************************************************
*** Filtration Options :
- filter repeats : Yes
- minimum repeat length : 10
- maximum repeat period : 4
- duplicate cutoff : -1
- filtering : none - brute force
************************************************************
*** Runtime mode:
- OpenMP support : Yes
- runtime mode : 0 = serial
************************************************************
*** Log messages:
[02/10/16 21:32:49] * Started checking duplex file
[02/10/16 21:32:49] * Finished checking duplex file
[02/10/16 21:32:49] * Started reading single-stranded file:RNA.fasta
[02/10/16 21:32:49] * Finished reading single-stranded file (1 sequences read)
[02/10/16 21:32:49] * Started detecting triplex-forming oligonucleotides in single-stranded
sequences
[02/10/16 21:32:49] * Finished detecting TFOs within 0.014 seconds (46 TFOs detected)
[02/10/16 21:32:49] * Started searching for triplexes
[02/10/16 21:32:49] * Processing DNA.fasta
[02/10/16 21:32:49] * Finished processing DNA.fasta
[02/10/16 21:32:49] * Finished searching for triplexes  within 0.00775 seconds (summed over
all cpus)
[02/10/16 21:32:49] * Exit without errors
[02/10/16 21:32:49] * Finished program within 0.0456 seconds
```

Supplementary note 10: RNA joint analysis

DNA origami rectangles were designed using caDNAno v2.0 (http://cadnano.org/), using M13mp18 (p7249) bacteriophage genome (Tilibit Nanosystems) as scaffold strand. The rectangle is identical to the basic origami object described by Rothemund [9], with additional side-facing handles used as RNA joint connectors. DNA staple strands were from Integrated DNA Technologies (IDT).

The caDNAno file can be downloaded from the following link:
https://drive.google.com/open?id=0B-CNvwEZiM3rcWpnUDNnVFViRnM

DNA staple strands were reconstituted at 100 µM with DNase/RNase free DDW. For folding reactions of DNA origami shapes, scaffold and staple strands were mixed at a 1:10 molar ratio in 1X TAE (40 mM Tris-acetate, 1 mM EDTA, pH 8.3) or folding buffer (5 mM Tris, 1 mM EDTA, 5 mM sodium chloride at pH 8), either supplemented with 10 mM magnesium chloride. Folding was performed in a thermal cycler (BioRad C1000 Touch Thermal Cycler) by a thermal annealing ramp as follows: 3 min at 90 °C, followed by stepwise cooling to 4 °C at -0.1 °C steps for a span of 12 hours. Folded objects were purified by 3X sequential centrifugal gel filtration using Amicon Ultra-0.5 mL 100K centrifugal filters (Millipore). DNA concentration was measured by spectrophotometer (Thermo Sci. NanoDrop 2000c). Objects were then electrophoresed on 1% agarose gels (ice cold 0.5X TBE, 10 mM magnesium chloride) stained with ethidium bromide.

Folded squares were incubated with x24molar ratio of RNA overnight in 1xTAE, 12.5 mM $MgCl_2$ at 37°C.

```
>RNAjoint_Sense
CCGACCCGAGGCGUGUUUAGCAAUUUACACCUCUGUUGGCAUUAAUCAGCGC

>RNAjoint_Antisense+loadingseq
GCGCUGAUUAAUGCCAACAGUACUACCCCCCCCUAAACACGCCUCGGGUCGGGAACUGGAGUAGCACAA
```

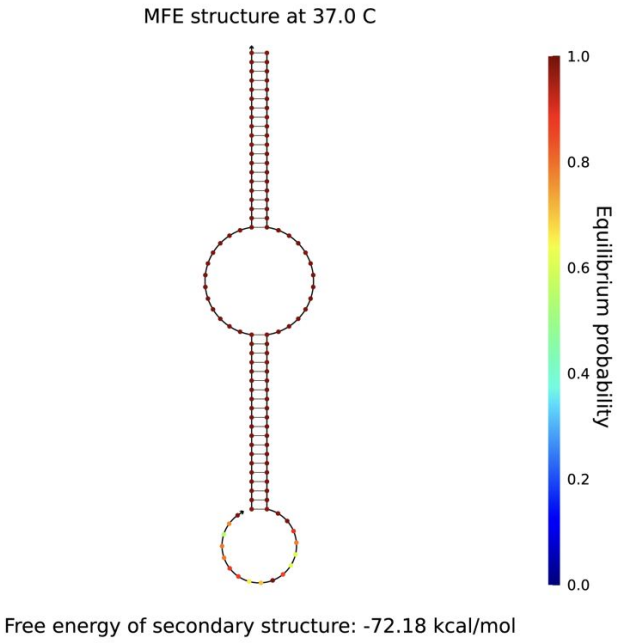


Figure S9. NUPACK prediction showing RNA joint structure. Small loop is in fact not a loop but a tail used for docking onto the origami rectangle.

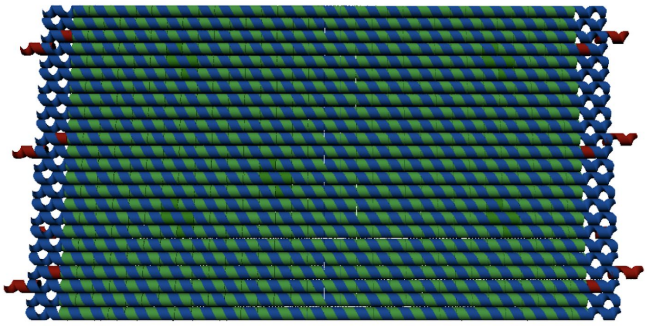


Figure S10. DNA origami rectangle with side-facing loading sites for RNA joints, used as AFM imaging guides.

Scanned images were analyzed using the angle tool in FIJI. Only joints that could be reliably assigned were used in the analysis.

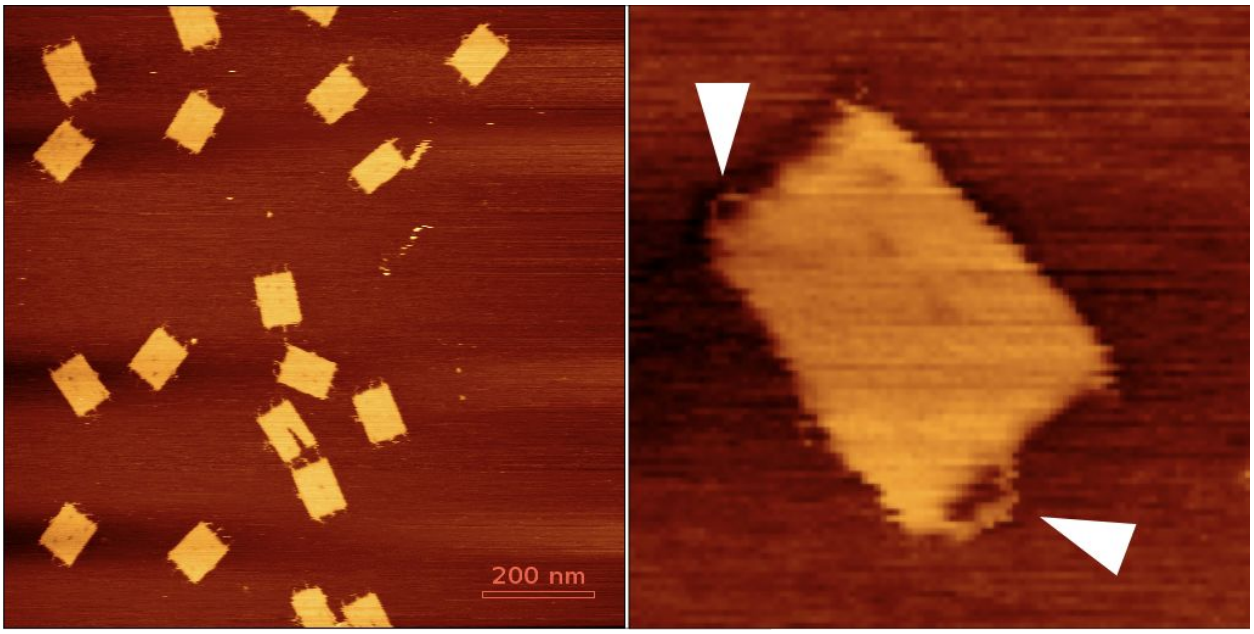


Figure S11. Joint

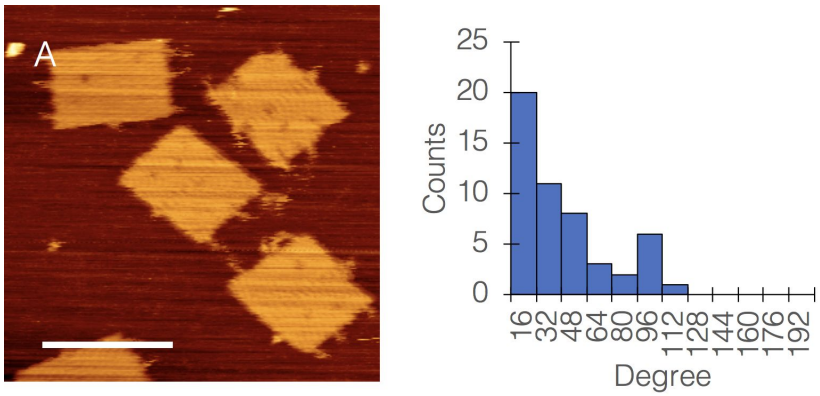


Figure S12. Joint analysis. Note the peak at ~ 90 degrees, which derives from the interaction between two adjacent joints. This phenomenon results from the joint ending in a blunt end which is hydrophobic, driving adjacent joints to interact in order to minimize contact with the aqueous buffer.

DNA origami scaffold and staple sequences are as follows:

>M13mp18_p7249

AATGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAGCTCGCGCCCCAAATGAAAATATAGCTAAACAGGTTATTGACCATTTGCGAAATG
TATCTAATGGTCAAACTAAATCTACTCGTTCGCAGAATTGGGAATCAACTGTTATATGGAATGAAACTTCCAGACACCGTACTTTAGTTGCATA
TTTAAAACATGTTGAGCTACAGCATTATATTCAGCAATTAAGCTCTAAGCCATCCGCAAAAATGACCTCTTATCAAAAGGAGCAATTAAAGGTA
CTCTCTAATCCTGACCTGTTGGAGTTTGCTTCCGGTCTGGTTCGCTTTGAAGCTCGAATTAAAACGCGATATTTGAAGTCTTTCGGGCTTCCTC
TTAATCTTTTTGATGCAATCCGCTTTGCTTCTGACTATAATAGTCAGGGTAAAGACCTGATTTTTGATTTATGGTCATTCTCGTTTTCTGAACT
GTTTAAAGCATTTGAGGGGGATTCAATGAATATTTATGACGATTCCGCAGTATTGGACGCTATCCAGTCTAAACATTTTACTATTACCCCCTCT
GGCAAAACTTCTTTTGCAAAAGCCTCTCGCTATTTTGGTTTTTATCGTCGTCTGGTAAACGAGGGTTATGATAGTGTTGCTCTTACTATGCCTC
GTAATTCCTTTTGGCGTTATGTATCTGCATTAGTTGAATGTGGTATTCCTAAATCTCAACTGATGAATCTTTCTACCTGTAATAATGTTGTTCC
GTTAGTTCGTTTTATTAACGTAGATTTTTCTTCCCAACGTCCTGACTGGTATAATGAGCCAGTTCTTAAAATCGCATAAGGTAATTCACAATGA

TTAAAGTTGAAATTAAACCATCTCAAGCCCAATTTACTACTCGTTCTGGTGTTTCTCGTCAGGGCAAGCCTTATTCACTGAATGAGCAGCTTTG
TTACGTTGATTTGGGTAATGAATATCCGGTTCTTGTCTTTCAAAGTTGGTCAGTTCGGTTCCCTTATGATTGACCGTCTGCGCCTCGTTCCGGC
TAAGTAACATGGAGCAGGTCGCGGATTTCGACACAATTTATCAGGCGATGATACAAATCTCCGTTGTACTTTGTTTCGCGCTTGGTATAATCGC
TGGGGGTCAAAGATGAGTGTTTTAGTGTATTCTTTTGCCTCTTTCGTTTTAGGTTGGTGCCTTCGTAGTGGCATTACGTATTTTACCCGTTTAA
TGGAAACTTCCTCATGAAAAAGTCTTTAGTCCTCAAAGCCTCTGTAGCCGTTGCTACCCTCGTTCCGATGCTGTCTTTCGCTGCTGAGGGTGAC
GATCCCGCAAAAGCGGCCTTTAACTCCCTGCAAGCCTCAGCGACCGAATATATCGGTTATGCGTGGGCGATGGTTGTTGTCATTGTCGGCGCAA
CTATCGGTATCAAGCTGTTTAAGAAATTCACCTCGAAAGCAAGCTGATAAACCGATACAATTAAAGGCTCCTTTTGGAGCCTTTTTTTTGGAGA
TTTTCAACGTGAAAAAATTATTATTCGCAATTCCTTTAGTTGTTCCTTTCTATTCTCACTCCGCTGAAACTGTTGAAAGTTGTTTAGCAAAATC
CCATACAGAAAATTCATTTACTAACGTCTGGAAAGACGACAAAACTTTAGATCGTTACGCTAACTATGAGGGCTGTCTGTGGAATGCTACAGGC
GTTGTAGTTTGTACTGGTGACGAAACTCAGTGTTACGGTACATGGGTTCCTATTGGGCTTGCTATCCCTGAAAATGAGGGTGGTGGCTCTGAGG
GTGGCGGTTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTACTAAACCTCCTGAGTACGGTGATACACCTATTCCGGGCTATACTTATATCAACCC
TCTCGACGGCACTTATCCGCCTGGTACTGAGCAAAACCCCGCTAATCCTAATCCTTCTCTTGAGGAGTCTCAGCCTCTTAATACTTTCATGTTT
CAGAATAATAGGTTCCGAAATAGGCAGGGGGCATTAACTGTTTATACGGGCACTGTTACTCAAGGCACTGACCCCGTTAAAACTTATTACCAGT
ACACTCCTGTATCATCAAAAGCCATGTATGACGCTTACTGGAACGGTAAATTCAGAGACTGCGCTTTCCATTCTGGCTTTAATGAGGATTTATT
TGTTTGTGAATATCAAGGCCAATCGTCTGACCTGCCTCAACCTCCTGTCAATGCTGGCGGCGGCTCTGGTGGTGGTTCTGGTGGCGGCTCTGAG
GGTGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGCTCTGAGGGAGGCGGTTCCGGTGGTGGCTCTGGTTCCGGTGATTTTGATTATGAAA
AGATGGCAAACGCTAATAAGGGGGCTATGACCGAAAATGCCGATGAAAACGCGCTACAGTCTGACGCTAAAGGCAAACTTGATTCTGTCGCTAC
TGATTACGGTGCTGCTATCGATGGTTTCATTGGTGACGTTTCCGGCCTTGCTAATGGTAATGGTGCTACTGGTGATTTTGCTGGCTCTAATTCC
CAAATGGCTCAAGTCGGTGACGGTGATAATTCACCTTTAATGAATAATTTCCGTCAATATTTACCTTCCCTCCCTCAATCGGTTGAATGTCGCC
CTTTTGTCTTTGGCGCTGGTAAACCATATGAATTTTCTATTGATTGTGACAAAATAAACTTATTCCGTGGTGTCTTTGCGTTTCTTTTATATGT
TGCCACCTTTATGTATGTATTTTCTACGTTTGCTAACATACTGCGTAATAAGGAGTCTTAATCATGCCAGTTCTTTTGGGTATTCCGTTATTAT
TGCGTTTCCTCGGTTTCCTTCTGGTAACTTTGTTCGGCTATCTGCTTACTTTTCTTAAAAAGGGCTTCGGTAAGATAGCTATTGCTATTTCATT
GTTTCTTGCTCTTATTATTGGGCTTAACTCAATTCTTGTGGGTTATCTCTCTGATATTAGCGCTCAATTACCCTCTGACTTTGTTCAGGGTGTT
CAGTTAATTCTCCCGTCTAATGCGCTTCCCTGTTTTTATGTTATTCTCTCTGTAAAGGCTGCTATTTTCATTTTTGACGTTAAACAAAAAATCG
TTTCTTATTTGGATTGGGATAAATAATATGGCTGTTTATTTTGTAACTGGCAAATTAGGCTCTGGAAAGACGCTCGTTAGCGTTGGTAAGATTC
AGGATAAAATTGTAGCTGGGTGCAAAATAGCAACTAATCTTGATTTAAGGCTTCAAAACCTCCCGCAAGTCGGGAGGTTCGCTAAAACGCCTCG
CGTTCTTAGAATACCGGATAAGCCTTCTATATCTGATTTGCTTGCTATTGGGCGCGGTAATGATTCCTACGATGAAAATAAAAACGGCTTGCTT
GTTCTCGATGAGTGCGGTACTTGGTTTAATACCCGTTCTTGGAATGATAAGGAAAGACAGCCGATTATTGATTGGTTTCTACATGCTCGTAAAT
TAGGATGGGATATTATTTTTCTTGTTCAGGACTTATCTATTGTTGATAAACAGGCGCGTTCTGCATTAGCTGAACATGTTGTTTATTGTCGTCG
TCTGGACAGAATTACTTTACCTTTTGTCGGTACTTTATATTCTCTTATTACTGGCTCGAAAATGCCTCTGCCTAAATTACATGTTGGCGTTGTT
AAATATGGCGATTCTCAATTAAGCCCTACTGTTGAGCGTTGGCTTTATACTGGTAAGAATTTGTATAACGCATATGATACTAAACAGGCTTTTT
CTAGTAATTATGATTCCGGTGTTTATTCTTATTTAACGCCTTATTTATCACACGGTCGGTATTTCAAACCATTAAATTTAGGTCAGAAGATGAA
ATTAACTAAAATATATTTGAAAAAGTTTTCTCGCGTTCTTTGTCTTGCGATTGGATTTGCATCAGCATTTACATATAGTTATATAACCCAACCT
AAGCCGGAGGTTAAAAAGGTAGTCTCTCAGACCTATGATTTTGATAAATTCACTATTGACTCTTCTCAGCGTCTTAATCTAAGCTATCGCTATG
TTTTCAAGGATTCTAAGGGAAAATTAATTAATAGCGACGATTTACAGAAGCAAGGTTATTCACTCACATATATTGATTTATGTACTGTTTCCAT
TAAAAAAGGTAATTCAAATGAAATTGTTAAATGTAATTAATTTTGTTTTCTTGATGTTTGTTTCATCATCTTCTTTTGCTCAGGTAATTGAAAT
GAATAATTCGCCTCTGCGCGATTTTGTAACTTGGTATTCAAAGCAATCAGGCGAATCCGTTATTGTTTCTCCCGATGTAAAAGGTACTGTTACT
GTATATTCATCTGACGTTAAACCTGAAAATCTACGCAATTTCTTTATTTCTGTTTTACGTGCAAATAATTTTGATATGGTAGGTTCTAACCCTT
CCATTATTCAGAAGTATAATCCAAACAATCAGGATTATATTGATGAATTGCCATCATCTGATAATCAGGAATATGATGATAATTCCGCTCCTTC
TGGTGGTTTCTTTGTTCCGCAAAATGATAATGTTACTCAAACTTTTAAAATTAATAACGTTCGGGCAAAGGATTTAATACGAGTTGTCGAATTG
TTTGTAAAGTCTAATACTTCTAAATCCTCAAATGTATTATCTATTGACGGCTCTAATCTATTAGTTGTTAGTGCTCCTAAAGATATTTTAGATA
ACCTTCCTCAATTCCTTTCAACTGTTGATTTGCCAACTGACCAGATATTGATTGAGGGTTTGATATTTGAGGTTCAGCAAGGTGATGCTTTAGA
TTTTTCATTTGCTGCTGGCTCTCAGCGTGGCACTGTTGCAGGCGGTGTTAATACTGACCGCCTCACCTCTGTTTTATCTTCTGCTGGTGGTTCG
TTCGGTATTTTTAATGGCGATGTTTTAGGGCTATCAGTTCGCGCATTAAAGACTAATAGCCATTCAAAAATATTGTCTGTGCCACGTATTCTTA
CGCTTTCAGGTCAGAAGGGTTCTATCTCTGTTGGCCAGAATGTCCCTTTTATTACTGGTCGTGTGACTGGTGAATCTGCCAATGTAAATAATCC
ATTTCAGACGATTGAGCGTCAAAATGTAGGTATTTCCATGAGCGTTTTTCCTGTTGCAATGGCTGGCGGTAATATTGTTCTGGATATTACCAGC
AAGGCCGATAGTTTGAGTTCTTCTACTCAGGCAAGTGATGTTATTACTAATCAAAGAAGTATTGCTACAACGGTTAATTTGCGTGATGGACAGA
CTCTTTTACTCGGTGGCCTCACTGATTATAAAAACACTTCTCAGGATTCTGGCGTACCGTTCCTGTCTAAAATCCCTTTAATCGGCCTCCTGTT
TAGCTCCCGCTCTGATTCTAACGAGGAAAGCACGTTATACGTGCTCGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAAGCGCGGC
GGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTC
GCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTTGG
GTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCA
AACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGAACCACCATCAAACAGGATTTTCGCC
TGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAG
AAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAA
AGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGT
GGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACC
TGCAGGCATGCAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATC
CCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTG
GTTTCCGGCACCAGAAGCGGTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGCAC
GGTTACGATGCGCCCATCTACACCAACGTGACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGGTTGTTACTCGC

TCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTCCTATTGGTTAAAAAATGAGCTGATTTAA
CAAAAATTTAATGCGAATTTTAACAAAATATTAACGTTTACAATTTAAATATTTGCTTATACAATCTTCCTGTTTTTGGGGCTTTTCTGATTAT
CAACCGGGGTACATATGATTGACATGCTAGTTTTACGATTACCGTTCATCGATTCTCTTGTTTGCTCCAGACTCTCAGGCAATGACCTGATAGC
CTTTGTAGATCTCTCAAAAATAGCTACCCTCTCCGGCATTAATTTATCAGCTAGAACGGTTGAATATCATATTGATGGTGATTTGACTGTCTCC
GGCCTTTCTCACCCTTTTGAATCTTTACCTACACATTACTCAGGCATTGCATTTAAAATATATGAGGGTTCTAAAAATTTTTATCCTTGCGTTG
AAATAAAGGCTTCTCCCGCAAAAGTATTACAGGGTCATAATGTTTTTGGTACAACCGATTTAGCTTTATGCTCTGAGGCTTTATTGCTTAATTT
TGCTAATTCTTTGCCTTGCCTGTATGATTTATTGGATGTT

START	END	SEQUENCE
20[207]	18[208]	CCAGCAGCAGGAAGGTTATCTAAAGCGGAACA
16[79]	14[80]	AAACAGGAATGCCGGAGAGGGTAGCAAGGATA
13[256]	15[255]	GAATAAACATAGTGAATTTATCAAAATCGTCG
11[224]	13[223]	CCTGTTTATTGAGAATCGCCATATTTTGAAAT
2[47]	0[48]	AAAGGAGCACTACAACGCCTGTAGAGCCACCA
11[192]	13[191]	ATAAACAAGTAATTTAGGCAGAGGGTTAATTT
5[160]	6[144]	GTGGCAACAACGTAGAAAATACATCAATCATA
2[79]	0[80]	TCACGTTGCTCATAGTTAGCGTAAAGAACCGC
3[128]	5[127]	GCGGGATCCGAGGGTAGCAACGGCATGTTACT
14[79]	12[80]	AAAATTTTGCATTAACATCCAATAATAACAGT
22[271]	20[272]	AATATCCAGCGCGAACTGATAGCCCCACCAGC
14[111]	12[112]	GCGGGAGACAAAGAATTAGCAAAACTAAAGTA
8[175]	6[176]	ATAACATAACTGAACACCCTGAACACGCAGTA
7[160]	8[144]	GAGAATTAAAAACAGGGAAGCGCATTGGGAAG
1[96]	3[95]	AGTTTTGTAAGGAATTGCGAATAATTCGGTCG
14[207]	12[208]	AGGTTGGGGACCTAAATTTAATGGTTAACAAC
8[143]	7[159]	AAAAATCTTATACCAGTCAGGACGTTAGACGG
22[143]	21[159]	TGCGTATTGCCAACGCGCGGGGAGACGACCAG
1[192]	3[191]	TCTGAATTTGACAGGAGGTTGAGGTTAGCGTC
13[128]	15[127]	TAAAGCCTACATTATGACCCTGTATCATTGCC
23[96]	22[112]	AGTCCACTATTAAAGAACGTGGACTTTCTTTT
21[192]	23[191]	GAACCCTTCTCAATCGTCTGAAATTGTTTTTA
10[159]	9[143]	AGATATAGATTCATTGAATCCCCCCGGAATCG
12[271]	10[272]	TCTTACCAAATATCCCATCCTAATCTGTCTTT
15[96]	17[95]	GAGAGATCGGTTGATAATCAGAAAATCAACAT
20[47]	18[48]	CTGTGTGACTGTTGGGAAGGGCGATCGCACTC
18[143]	17[159]	GGGATAGGAACAAACGGCGGATTGCAATTCAT
22[207]	20[208]	TTTTGACGCTGACCTGAAAGCGTAGCTGAGAG
0[143]	1[127]	GTATAGCCCGGAATAGGTGTATCAAGTAAATG
13[224]	15[223]	ACCGACCGGACTACCTTTTTAACCAATATATG
15[128]	17[127]	TGAGAGTCCGTAAAACTAGCATGTTCGGATTC
15[32]	17[31]	ACCATCAAGTTAATATTTTGTTAATTTTTTAA
2[207]	0[208]	GCCAGCATTACCGTTCCAGTAAGCAGAGAAGG
6[175]	4[176]	TGTTAGCAATATAAAAGAAACGCAAACGTCAC
20[271]	18[272]	AGAAGATAAATACATTTGAGGATTACTCGTAT
16[175]	14[176]	TGATGAAAAACAATTTCATTTGAAAAATGCTG
15[160]	16[144]	TTACATTTCAAACATCAAGAAAACATCGATGA
15[192]	17[191]	TTTAATGGTTCAATTACCTGAGCATATACTTC
22[111]	20[112]	CACCAGTGAAACCTGTCGTGCCAGATGCCTGC
9[224]	11[223]	TTAAATCACTCATCGAGAACAAGCAGAACGCG
10[47]	8[48]	GGATTGCAGACGATAAAAACCAAAAACTAATG
5[192]	7[191]	ACGGAATAATTAAGACTCCTTATTAAAGTCAG
4[143]	3[159]	CATCGGAAGTCACCCTCAGCAGCGCGTAATCA
7[224]	9[223]	CCCACAAGAAGAAACGATTTTTTGTTGAAGCC
20[79]	18[80]	CTCGAATTCGCTATTACGCCAGCTAGGGGACG
11[32]	13[31]	TTCAAAGCTTAGATACATTTCGCATTTGGGGC
21[256]	23[255]	TCTTTAATGAACAATATTACCGCCAGCAATAC
4[47]	2[48]	CTACGAAGAGTTGCGCCGACAATGAGGCTCCA

19[96]	21[95]	GGGGATGTTCTAGAGGATCCCCGGCCGCTTTC
6[271]	4[272]	TAGCCGAAACATTCAACCGATTGACATTAAAG
20[175]	18[176]	TTGCTGAATCAGTTGGCAAATCAATTATCATC
16[111]	14[112]	TGTACCCCTACAAAGGCTATCAGGATACTTTT
20[143]	19[159]	AACGACGGGTTTTCCCAGTCACGACTCAATCA
1[32]	3[31]	CAGTACAACTTTAATTGTATCGGTACAGCTTG
23[64]	22[80]	TAGGGTTGAGTGTTGTTCCAGTTTATTGCCCT
9[128]	11[127]	CAATACTGTCAAATGCTTTAAACAGCGGATGG
22[175]	20[176]	TACATTGGGGGACATTCTGGCCAAGCATCACC
1[256]	3[255]	TAAGTTTTCCACCCTCAGAGCCACCCATCTTT
14[47]	12[48]	CAATGCCTAAAAGGTGGCATCAATAGATTTAG
6[143]	5[159]	AGGGAACCACGAGGCGCAGACGGTACATAAAG
8[79]	6[80]	TTGAGATTAGATGGTTTAATTTCACGCATAGG
23[256]	22[272]	TTCTTTGATTAGTAATAACATCACTTGCTGGT
17[64]	19[63]	CTGGCCTTATCGGCCTCAGGAAGATCGGTGCG
2[239]	0[240]	GAGCCGCCGATGATACAGGAGTGTACATGAAA
5[32]	7[31]	CCCAGCGAATATTCATTACCCAAAAGGCTTGC
15[256]	17[255]	CTATTAATTAACGGATTCGCCTGAAATAAAGA
1[224]	3[223]	TGGCTTTTACCAGAACCACCACCAATTTTCGG
6[47]	4[48]	AGAACCGGTTATACCAAGCGCGAATAATGCCA
1[128]	3[127]	AATTTTCTTTCAGCGGAGTGAGAACCGCTTTT
3[160]	4[144]	GTAGCGACCCATCGATAGCAGCACAAAGACAG
11[256]	13[255]	AGAAAAATGTATAAAGCCAACGCTTTAAATAA
12[175]	10[176]	AGCCAGTAAGTAATTCTGTCCAGAACCGCGCC
10[111]	8[112]	AACGAGAAAAAATGTTTAGACTGGCTAACGGA
19[128]	21[127]	ACGCCAGGCCAGTGCCAAGCTTGCCTGCATTA
3[32]	5[31]	ATACCGATGCACCAACCTAAAACGCTTTGACC
12[79]	10[80]	TGATTCCCGGATTAGAGAGTACCTATCAGGTC
17[192]	19[191]	TGAATAATACCAGAAGGAGCGGAACAGTTGAA
20[111]	18[112]	AGGTCGACGCTGCAAGGCGATTAAGGGCGCAT
5[128]	7[127]	TAGCCGGAGAACTGACCAACTTTGTTAAGAAC
18[47]	16[48]	CAGCCAGCAACGCCATCAAAAATATATTTAAA
23[160]	22[176]	ACGGTACGCCAGAATCCTGAGAAGGGATTATT
10[79]	8[80]	TTTACCCTGCAAAAGAAGTTTTGCTTCATCAG
11[64]	13[63]	ACAGGTCAAATTCTGCGAACGAGTTCTACTAA
18[111]	16[112]	CGTAACCGAGCGAGTAACAACCCGCAATCATA
8[111]	6[112]	ACAACATTATTACCTTATGCGATTAAAGAGGA
18[79]	16[80]	ACGACAGTCCTGTAGCCAGCTTTCAGCCCCAA
9[32]	11[31]	TACCAGACTCAAAAAGATTAAGAGATTCGAGC
13[160]	14[144]	GAAAACTTCCAATCGCAAGACAAAATCGGTTG
5[224]	7[223]	AAAATTCAATAATAACGGAATACCAGAGATAA
2[271]	0[272]	CAGAACCGAACGGGGTCAGTGCCTCCCCTGCC
16[207]	14[208]	TTATTCATAAACAGTACATAAATCTCCGGCTT
13[96]	15[95]	AGGCAAGGAGCCTTTATTTCAACGCTATTTTT
6[111]	4[112]	CAGATGAAATCCGCGACCTGCTCCTACAGAGG
11[128]	13[127]	CTTAGAGCTGTTTTAAATATGCAATTAAGCAA
23[192]	22[208]	TAATCAGTGAGGCCACCGAGTAAATACCTACA
2[143]	1[159]	TCAACAGTGTATGGGATTTTGCTACAAATAAA
18[239]	16[240]	AAGTTTGATTGCACGTAAAACAGATTGCTTTG
10[239]	8[240]	GTACCGCAAGATTAGTTGCTATTTTTATCCCA
4[207]	2[208]	ACCATTACGCGCGTTTTCATCGGCGAGCCGCC
2[175]	0[176]	ACGATTGGAAAGCCAGAATGGAAAAGTACCAG
0[207]	1[191]	ATTAGGATTAGCGGGGTTTTGCTCGCGCAGTC
17[96]	19[95]	TAAATGTGTGCATCTGCCAGTTTGGGCGAAAG
19[64]	21[63]	GGCCTCTTCGTAATCATGGTCATACTAACTCA
4[239]	2[240]	TGGGAATTCCCTTATTAGCGTTTGCACCCTCA
7[64]	9[63]	TGGGCTTGTAGGAATACCACATTCATAGCGAG

19[256]	21[255]	CAATAGATAAACAGAGGTGAGGCGGCTATTAG
18[175]	16[176]	ATATTCCTTCCTGATTGTTTGGATAAAGAAGA
14[271]	12[272]	AAGAGTCAACCGGAATCATAATTAATACAAAT
12[207]	10[208]	GCCAACATCATGTTCAGCTAATGCAAGCCGTT
21[224]	23[223]	TGGCACAGAAAACGCTCATGGAAAAGAGTCTG
0[271]	1[255]	TATTTCGGAACCTATTATTCTGAAACTGGTAA
3[96]	5[95]	CTGAGGCTACTAAAGACTTTTTCAGATAAATT
4[271]	2[272]	GTGAATTAAAAATCACCGGAACCAGCCACCCT
10[271]	8[272]	CCTTATCATTTATCCTGAATCTTATTTGCCAG
23[128]	23[159]	CAAAGGGCGAAAAACCGTCTATCAAGACAGGA
20[239]	18[240]	TAACACCGTAACAACTAATAGATTAATTTTAA
0[239]	1[223]	GTATTAAGAGGCTGAGACTCCTCAGTCATACA
10[207]	8[208]	TTTATTTTCGACTTGCGGGAGGTTTTTAACGT
3[192]	5[191]	AGACTGTACATTAGCAAGGCCGGAAAGACACC
22[79]	20[80]	TCACCGCCGCGTTGCGCTCACTGCGTACCGAG
1[160]	2[144]	TCCTCATTCCTTGATATTCACAAAAACAACTT
5[96]	8[95]	GTGTCGAACGGTGTACAGACCAGGACTTTAATCATTGTGAATTACAGGT
17[128]	19[127]	TCCGTGGGTCACGTTGGTGTAGATGTTGGGTA
8[207]	6[208]	CAAAAATGTGAGCGCTAATATCAGCAAAAGAA
7[32]	9[31]	CCTGACGATAACGCCAAAAGGAATCCCTCGTT
22[239]	20[240]	CAACAGGAACAATATTTTTGAATGGTCAGTAT
22[47]	20[48]	CGGTCCACGTGCCTAATGAGTGAGGCTGTTTC
10[175]	8[176]	CAATAGCATCTAAGAACGCGAGGCCAGAGAGA
16[271]	14[272]	AGAAACAATAATTTTCCCTTAGAAACGCTGAG
13[32]	15[31]	GCGAGCTGGAGTAATGTGTAGGTAGTCAAATC
7[192]	9[191]	AGGGTAATAAAATAGCAGCCTTTAGTTTTAGC
5[64]	7[63]	CAACGGAGCCTTCATCAAGAGTAATAGTAAAT
16[143]	15[159]	ACGGTAATTGGAGCAAACAAGAGAAAAATTAA
0[175]	0[144]	GCGGATAAGTGCCGTCGAGAGGGTTGATATAA
13[192]	15[191]	CATCTTCTTTATATAACTATATGTTTACCTTT
17[256]	19[255]	AATTGCGTTTGCCCGAACGTTATTAGAGCCGT
19[160]	20[144]	ATATCTGGCCTCAAATATCAAACCCGTTGTAA
2[111]	0[112]	AACAACTACGTCTTTCCAGACGTTCCGTACTC
18[271]	16[272]	TAAATCCTAGATTTTCAGGTTTAAACATCGGG
14[143]	13[159]	TACCAAAACAGAGCATAAAGCTAAGAACGCGA
12[47]	10[48]	TTTGACCAGAACCAGACCGGAAGCAGCAAAGC
19[32]	21[31]	CTGCGCAAAATTGTTATCCGCTCAAAGTGTAA
9[64]	11[63]	AGGCTTTTGACTATTATAGTCAGAAAACTCCA
0[47]	1[31]	CCCTCATTTTCAGGGATAGCAAGCTTCGTCAC
0[111]	1[95]	AGGAGGTTTAGTACCGCCACCCTCCGATCTAA
9[256]	11[255]	GCTACAATTTCCAAGAACGGGTATCCTGAACA
16[239]	14[240]	AATACCAATAACCTTGCTTCTGTAAATCATAG
8[47]	6[48]	CAGATACAGAAACACCAGAACGAGTCTTGACA
18[207]	16[208]	AAGAAACCGGAAGGGTTAGAACCTGAGGCGAA
23[224]	22[240]	TCCATCACGCAAATTAACCGTTGTAGCCATTG
21[32]	23[31]	AGCCTGGGGCTGGTTTGCCCCAGCATCGGCAA
13[64]	15[63]	TAGTAGTATAGAACCCTCATATATTCTAGCTG
17[160]	18[144]	CAATATAAGATTATCAGATGATGGACCGTAAT
9[192]	11[191]	GAACCTCCCATCGTAGGAATCATTCGACGACA
8[271]	6[272]	TTACAAAAACAATGAAATAGCAATTAAGCAGA
21[64]	23[63]	CATTAATTTGGCCCTGAGAGAGTTGCCCGAGA
23[32]	22[48]	AATCCCTTATAAATCAAAAGAATAGCAGCAAG
17[32]	19[31]	CCAATAGGTTTCCGGCACCGCTTCCATTCAGG
1[64]	3[63]	AGACAGCCAAAATCTCCAAAAAAAACAACAAC
11[160]	12[144]	AAAGGTAAATAAGAGAATATAAAGATGCTGTA
19[224]	21[223]	AGGAGCACCCTGCAACAGTGCCACAGAATACG
4[111]	2[112]	CTTTGAGGTGCAGGGAGTTAAAGGTAGAAAGG

12[239]	10[240]	GGGCTTAATCAACAATAGATAAGTTAAACCAA
17[224]	19[223]	AAAATTATGTAACATTATCATTTTATATCTTT
21[96]	23[95]	CAGTCGGGAGACGGGCAACAGCTGGGAACAAG
5[256]	7[255]	AAAGGGCGCAAAGTTACCAGAAGGATAATAAG
19[192]	21[191]	AGGAATTGAAATGAAAAATCTAAACAGAGATA
0[79]	1[63]	CACCCTCAGAACCGCCACCCTCAGCATTCCAC
4[79]	2[80]	TTTCCATTACGCATAACCGATATATAATTTTT
16[47]	14[48]	TTGTAAACTATGATATTCAACCGTTTTAAATG
4[175]	2[176]	CAATGAAAAGAATCAAGTTTGCCTCAGGTCAG
6[207]	4[208]	CTGGCATGAGTTTATTTTGTCACACCAGTAGC
14[239]	12[240]	GTCTGAGATGTGATAAATAAGGCGCAACAGTA
7[128]	9[127]	TGGCTCATACGTTAATAAAACGAAATAGCGTC
21[160]	22[144]	TAATAAAACAGATTCACCAGTCACAGGCGGTT
8[239]	6[240]	ATCCAAATAATTGAGTTAAGCCCAAAACCGAG
7[256]	9[255]	AGCAAGAATAAACAGCCATATTATTGCACCCA
6[239]	4[240]	GAAACGCATATGGTTTACCAGCGCGAGCCATT
6[79]	4[80]	CTGGCTGAATTTGTATCATCGCCTTGAGGAAG
14[175]	12[176]	ATGCAAATTTTCAAATATATTTTACATTTTCG
3[256]	5[255]	TCATAATCTCACCGTCACCGACTTCAAAGACA
12[111]	10[112]	CGGTGTCTATAAGAGGTCATTTTTGTTCAGAA
12[143]	11[159]	GCTCAACATTAATTGCTGAATATATACCGACA
21[128]	23[127]	ATGAATCGGGGCGCCAGGGTGGTTTCCAACGT
10[95]	13[95]	AATCAAAATTAATTGCTCCTTTTGGGAAGTTTCATTCCATAATCATAC
8[94]	24[81]	AGAAAGACAGAGGGGGTAATAGTTGACCATATTGTGCTACTCCAGTTC
15[224]	24[207]	TGAGTGAAGTTACAAAATCGCGCAACCATATCTTGTGCTACTCCAGTTC
3[224]	25[239]	TCATAGCCAGAGCCAGCAAAATCAATCAATAGTTGTGCTACTCCAGTTC
3[64]	25[79]	CATCGCCCAAACGGGTAAAATACGACAAAGTATTGTGCTACTCCAGTTC
9[144]	24[144]	TCATAAATAAGGCTTATCCGGTATAGCAAATCTTGTGCTACTCCAGTTC
15[64]	24[47]	ATAAATTAAGATTGTATAAGCAAAATTCGCGTTTGTGCTACTCCAGTTC
16[287]	5[15]	TACCTTTTCGTCAGATGAATATACACACTAAA
20[287]	9[15]	CGAACGAACTAAAACATCGCCATTAGCAACAC
22[287]	11[15]	TATCGGCCTTGCCTGAGTAGAAGACAAATATC
12[287]	1[15]	TATGCGTTCTAGAAAAAGCCTGTTTACCGTAA
14[287]	3[15]	AGATTAAGTCCTTGAAAACATAGCGAGGTGAA
18[287]	7[15]	ATTCGACATAGAAGTATTAGACTTGCTCATTC
18[287]	19[15]	ATTCGACATAGAAGTATTAGACTTGCAAAGCG
21[16]	20[288]	GAAGCATACAATTCCACACAACATAAAAATAC
22[287]	23[15]	TATCGGCCTTGCCTGAGTAGAAGATGATGGTG
20[287]	21[15]	CGAACGAACTAAAACATCGCCATTACGAGCCG
23[16]	22[288]	GTTCCGAAAGGCGAAAATCCTGTTCTCAAAC
19[16]	18[288]	CCATTCGCTGGTGCCGGAAACCAGTACAAACA
0[287]	1[15]	GTTAATGCTGAGTAACAGTGCCCGTACCGTAA
2[287]	3[15]	TCAGAGCCGAGCCACCACCGGAACGAGGTGAA
5[16]	4[288]	ACACTCATAAAGAGGCAAAAGAATATTGACGG
3[16]	2[288]	TTTCTTAATTATCAGCTTGCTTTCCGCCTCCC
1[16]	0[288]	CACTGAGTCCAATAGGAACCCATGTATAAACA
4[287]	5[15]	AAATTATTGGGAGGGAAGGTAAATACACTAAA
11[16]	22[288]	GCGTTTTAGAAGCCCGAAAGACTTACTCAAAC
1[16]	12[288]	CACTGAGTCCAATAGGAACCCATGTAGTATCA
9[16]	20[288]	TATCATAATACGAGGCATAGTAAGAAAAATAC
5[16]	16[288]	ACACTCATAAAGAGGCAAAAGAATAGTAACAG
7[16]	18[288]	AGTGAATATCAACGTAACAAAGCTTACAAACA
3[16]	14[288]	TTTCTTAATTATCAGCTTGCTTTCGATAGCTT

18[287]	19[15]	ATTCGACATAGAAGTATTAGACTTGCAAAGCG
19[16]	18[288]	CCATTCGCTGGTGCCGGAAACCAGTACAAACA
20[287]	21[15]	CGAACGAACTAAAACATCGCCATTACGAGCCG
21[16]	20[288]	GAAGCATACAATTCCACACAACATAAAAATAC
22[287]	23[15]	TATCGGCCTTGCCTGAGTAGAAGATGATGGTG
23[16]	22[288]	GTTCCGAAAGGCGAAAATCCTGTTACTCAAAC
0[287]	1[15]	GTTAATGCTGAGTAACAGTGCCCGTACCGTAA
1[16]	0[288]	CACTGAGTCCAATAGGAACCCATGTATAAACA
2[287]	3[15]	TCAGAGCCGAGCCACCACCGGAACGAGGTGAA
3[16]	2[288]	TTTCTTAATTATCAGCTTGCTTTCCGCCTCCC
4[287]	5[15]	AAATTATTGGGAGGGAAGGTAAATACACTAAA
5[16]	4[288]	ACACTCATAAAGAGGCAAAAGAATATTGACGG
2[287]	3[15]	TCAGAGCCGAGCCACCACCGGAACGAGGTGAA
3[16]	2[288]	TTTCTTAATTATCAGCTTGCTTTCCGCCTCCC
20[287]	21[15]	CGAACGAACTAAAACATCGCCATTACGAGCCG
21[16]	20[288]	GAAGCATACAATTCCACACAACATAAAAATAC

Supplementary note 11: Analysis of computational predictions

We compared our microscopic observations with structural predictions using two available tools. We fed HOTAIR and RANDOM **(see Supplementary note 5)** into Mfold[10] and RNAfold[11], using their default prediction parameters. Mfold yielded 28 and 30 predicted secondary structures for HOTAIR and RANDOM, respectively, with average minimum free energy (MFE) of -697.93±5.99 kcal/mol for HOTAIR. The shape range for both sequences included shapes resembling the observed anatomy as well as more remote ones. RNAfold retrieved a single structure for each sequence (MFE -668 kcal/mol for HOTAIR and -632 kcal/mol for RANDOM). However, difference between predictions, and similarity between observed and predicted HOTAIR, were not conclusive. We therefore chose not to rely on these predictions for our understanding of the global structure of HOTAIR. Nevertheless, some strictly-local predictions (e.g. lengths of stems and loops) were used as assumptions in certain aspects of our modeling approach.

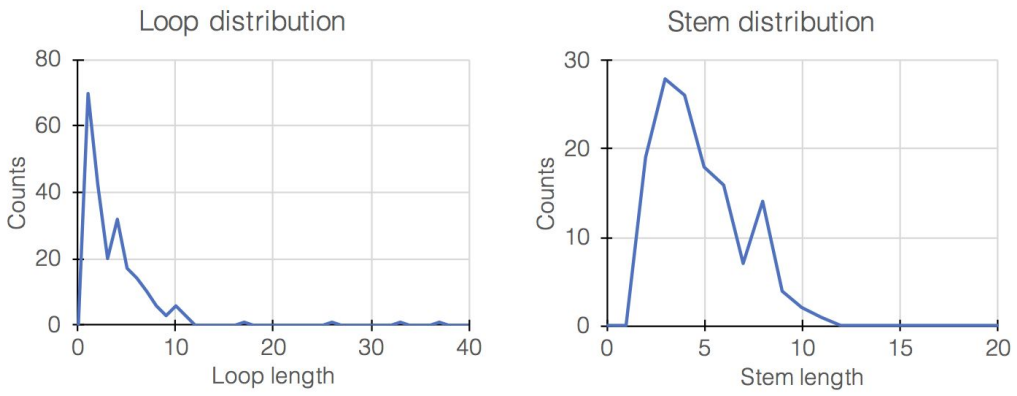
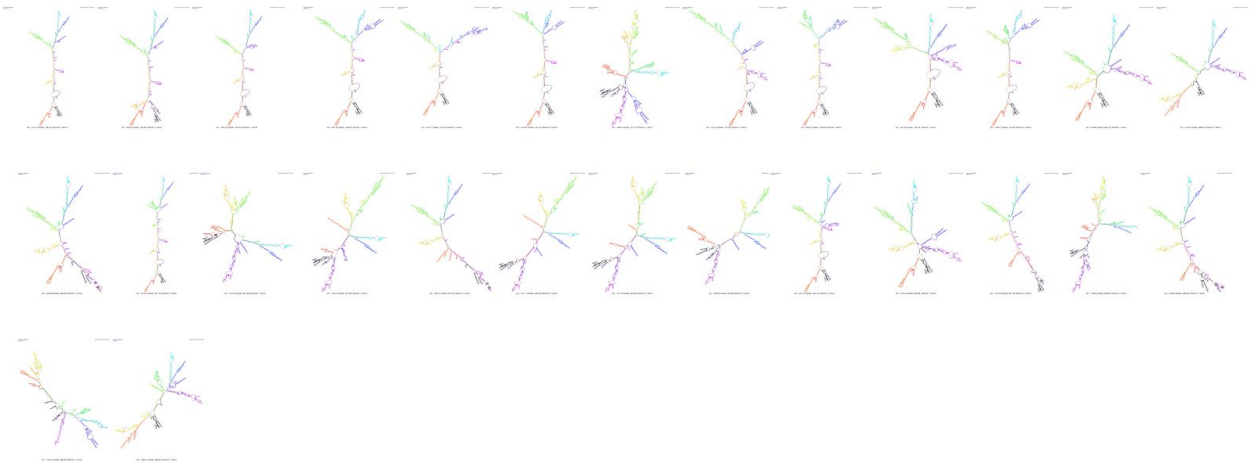


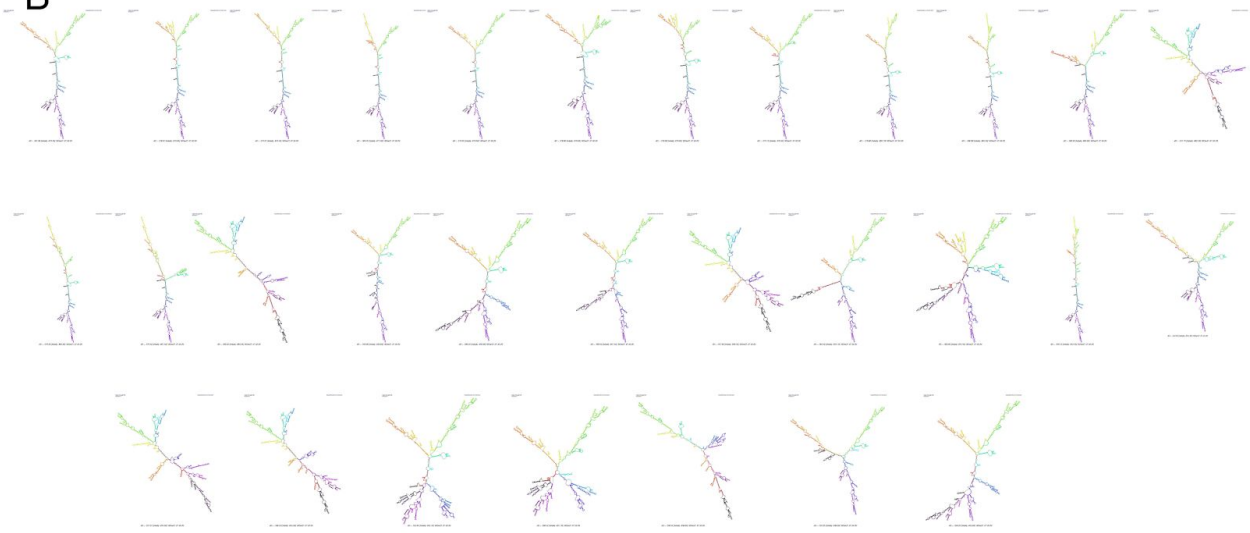
Figure S13. Distribution of stem and loop sizes in HOTAIR predicted by Mfold.

(On next page) Figure S14. Predicted minimum free energy configurations of HOTAIR and a random sequence generated from HOTAIR (same base composition, same length, scrambled sequence). A, HOTAIR. B, Random.

A



B



References

1. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
2. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
3. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
4. Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
5. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
6. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
7. Chu, C., Qu, K., Zhong, F. L., Artandi, S. E. & Chang, H. Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **44**, 667–678 (2011).
8. Buske, F. A., Bauer, D. C., Mattick, J. S. & Bailey, T. L. Triplexator: detecting nucleic acid triple helices in genomic and transcriptomic data. *Genome Res.* **22**, 1372–1381 (2012).
9. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
10. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
11. Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011).