Supporting Information

Xu et al. 10.1073/pnas.1612745113

SI Materials and Methods

RNA Pol II Purification. Wild-type RNA pol II used for transcription was purified from *Saccharomyces cerevisiae*, as previously described unless otherwise stated (51, 52). The AA-mutant was introduced into a protease-deficient yeast strain GRY3175 by plasmid shuffle (56). Hexahistidine-tagged Rpb3 was used to pull down pol II from the whole cell lysate essentially as previously described (56, 57). In parallel, the wild-type RNA pol II used as comparison in the mutation study was also generated and purified via the same method as that for the AA-mutant. The DNA template and nontemplate oligonucleotides were purchased from IDT. RNA primers were purchased from TriLink Biotechnologies and radiolabeled using [γ-^{32}P] ATP and T4 Polynucleotide Kinase (New England Biolabs).

DNA Template Design. DNA sequences of template and nontemplate strand used for TS binding orientation were 5′-GACTCTTCTGACTTGGTCATACACACACTTCTCTCTCTCGTTGTTCCTCTCGATTGTTAAGTGATGTGTCGTTGTAAT-3′ and 5′-ATTACAACGACACATCACTTATGTCAGATCTACGCTCACGAGAGAGAGAAGTGTGTGTATGACCAAGTCAGAAGAGTC-3′, and the binding site for compounds **1**–**4** is underlined. Sequences of template and nontemplate strand used for NTS binding orientation are 5′-GACTCTTCTGACTTGACCATACACACACTTCTCTCTCTCGTTGTTCCTCTCGATTGTTAAGTGATGTGTCGTTGTAAT-3′ and 5′-ATTACAACGACACATCACTTATGTCAGATCTACGCTC ACGAGAGAGAGAAGTGTGTGTATGGTCAAGTCAGAAGAGTC-3′, and the binding site for compounds **1**–**4** is underlined. The single mutation was introduced in both TS binding orientation and NTS binding orientation. The single mutated template strand for the TS binding orientation is 5′-GACTCTTCTGACTTGTTCATACACACACTTCTCTCTCTCGTTGTTCCTCTCGATTGTTAAGTGATGTGTCGTTGTAAT-3′ and the corresponding nontemplate strand is 5′-ATTACAACGACACATCACTTATGTCAGATCTACGCTCACGAGAGAGAGAAGTGTGTGTA TGAACAAGTCAGAAGAGTC-3′, and the mutated binding site is underlined. The single mutated template and nontemplate strand for the template strand for NTS binding orientation are 5′-GACTCTTCTGACTTGAACATACACACACTTCTCTCTCTCGTTGTTCCTCTCGATTGTTAAGTGATGTGTCGTTGTAAT-3′ and 5′-ATTACAACGACACATCACTTATGTCAGATCTACGCTCACGAGAGAGAGAAGTGTGTGT ATGTTCAAGTCAGAAGAGTC-3′, and the mutated binding site is underlined. The RNA primer used in all these scaffolds was 5′-AUCGAGAGGA-3′.

RNA Pol II Elongation Complex Assembly with a Full Transcription Bubble. The pol II elongation complexes in the full transcription bubble were assembled based on previous reported methods with slight modifications (7, 58). Briefly, an aliquot of 5′-^{32}P-labeled RNA was annealed with a 1.2-fold amount of template DNA and 1.5-fold amount of nontemplate DNA to form the RNA/DNA bubble scaffold in elongation buffer [20 mM Tris·HCl (pH 7.5), 40 mM KCl, 5 mM $MgCl_2$]. An aliquot of annealed scaffold of RNA/DNA was then incubated with a fourfold excess amount of pol II at room temperature for 10 min to ensure the formation of a pol II elongation complex.

In Vitro RNA Pol II Transcription Assay. The assay was carried out as previously described (7, 58–60). Briefly, the preincubated polymerase/scaffold complex was mixed with an equal volume of solution containing 40 mM KCl, 20 mM Tris·HCl (pH 7.5), 10 mM DTT, 10 mM $MgCl_2$, and 2 mM nucleotide triphosphate (NTP). Final reaction concentrations after mixing were 25 nM scaffold, 100 nM pol II, 5 mM $MgCl_2$, and 1 mM NTP in elongation buffer unless stated differently somewhere else. Reactions were quenched at various times by addition of one volume of 0.5 M EDTA (pH 8.0). For the TFIIS treatment, the elongation complex is mixed with equal volume of solution containing 40 mM KCl, 20 mM Tris·HCl (pH 7.5), 10 mM DTT, 10 mM $MgCl_2$, 2 mM NTP, and 2 μM TFIIS. Final reaction concentrations after mixing were 25 nM scaffold, 100 nM pol II, 5 mM $MgCl_2$, 1 mM NTP, and 1 μM TFIIS in the elongation buffer. The transcript was analyzed by 12% (wt/vol) denatured PAGE.

Transcription Inhibition Assay by Py-Im Polyamide. For the Py-Im polyamide transcription inhibition assay, the RNA/DNA scaffold was preformed and incubated with various Py-Im polyamides for 3 h before the addition of pol II. The in vitro transcription was performed as described above. For the concentration titration assays, the final concentration of elongation complex is 2 nM.

K_i values for these compounds were determined by fitting the curves of the percentage of pol II transcriptional full-length bypass in the presence of the Py-Im polyamide with concentrations of the Py-Im polyamide in a log scale. The percentage of bypass was calculated by comparing the full-length transcript with the actively elongated transcript. Relative transcriptional bypass percentage was normalized against the maximal percentage of full-length transcript in the absence of ligand. Inhibition curves were determined using GraphPad Prism 6 by variable-slope, nonlinear regression fit to a dose–response model with a bottom constraint of 0 and a top constraint of 100.

Py-Im Polyamide Synthesis. Py-Im polyamides **1**, **2**, and **3** were synthesized by microwave-assisted, solid-phase synthesis according to previously described protocols (54). Boc protected *N*-methylpyrrole and *N*-methylimidazole monomers were attached on oxime resin (Peptides International) using published PyBOP coupling conditions. All hairpin polyamides were cleaved from resin with 3,3′-diamino-*N*-methyldipropylamine, precipitated with Et_2O, and coupled to isophthalic acid before purification by reverse phase HPLC. Polyamide **4** was synthesized on 2-chlorotrityl chloride resin (Bachem) using FmocPyOH and FmocImOH monomers (Wako), as previously reported (55). Purity and identity of compounds were verified by analytical HPLC and MALDI-TOF mass spectrometry.

MALDI-TOF characterization for polyamides **1**–**4** are summarized as follows:

i) MALDI-TOF $[M + H]^+$ calculated for $C_{64}H_{76}N_{23}O_{12}^+$ = 1358.6, observed = 1358.9

ii) MALDI-TOF $[M + H]^+$ calculated for $C_{66}H_{77}N_{23}O_{13}^+$ = 1399.6, observed = 1400.7

iii) MALDI-TOF $[M + H]^+$ calculated for $C_{64}H_{76}N_{23}O_{12}^+$ = 1358.6, observed = 1358.9

iv) MALDI-TOF $[M + H]^+$ calculated for $C_{53}H_{61}N_{22}O_{10}^+$ = 1165.5, observed = 1165.5.

Thermal Denaturation Analysis. DNA thermal stabilization analysis of polyamides **1**–**4** was performed on a thermos-controlled cell holder equipped Varian Cary 100 spectrophotometer. The match oligo (5′-GACT TGGTCA TACA-3′) and mismatch oligo (5′-GACT TGTTCA TACA -3′) were purchased from Valuegene. DNA duplexes and polyamides were added to a final concentration of 2 μM and 3 μM, respectively, to a degassed solution

of 10 mM sodium cacodylate, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0, as previously described (53). Denaturation profiles were recorded at $\lambda = 260$ nm from 25 °C to 95 °C with a heating rate of 0.5 °C/min. Thermal stabilization (ΔTm) by polyamides was calculated relative to naked control oligonucleotides. Results are reported as an average of three experiments.

Molecular Modeling. The 3D model of Py-Im polyamide **1** was oriented from the crystal structure of cyclic Py-Im polyamide inhibitor with DNA (PDB ID code 3I5L) (31), and modified by Chemdraw3D software. The DNA sequence was mutated according to the Py-Im polyamide **1**-binding sequence with COOT (61). The docking of ligand with the DNA-binding site was optimized to get the Py-Im polyamide **1**-DNA complex using the software ICM-browser-pro (23, 62). To get the register of the first pol II pausing state (state *i*), part of Py-Im polyamide **1**-DNA complex (from the 1st base to the 10th base of Chain A) was first aligned with the downstream NTS DNA region of the pol II complex (from *i*+5 site to *i*+14 site) (PDB ID code 3M3Y). Then, the portion of downstream DNA duplex of pol II elongation complex (from *i*+5 to *i*+14) was substituted by the superimposed 10 bp Py-Im polyamide 1-binding DNA duplex structure and extended with the ideal B-form DNA duplex downstream (beyond *i*+15 site) added with COOT. Finally, the upstream DNA and RNA sequences were modified accordingly to generate the model of the arrested state *i*. To mimic the stepwise pol II progression toward Py-Im polyamide binding site, we aligned and substituted a portion of downstream DNA duplex of pol II elongation complex with the superimposed 10 bp Py-Im polyamide 1-binding DNA duplex structure to generate the other three arrested states (from *i*+4 site to *i*+13 site for state *ii*, from *i*+3 site to *i*+12 site for state *iii*, from *i*+2 site to *i*+11 site for state *iv*, respectively) in a similar manner. All structure figures were rendered in PyMOL (www.pymol.org).

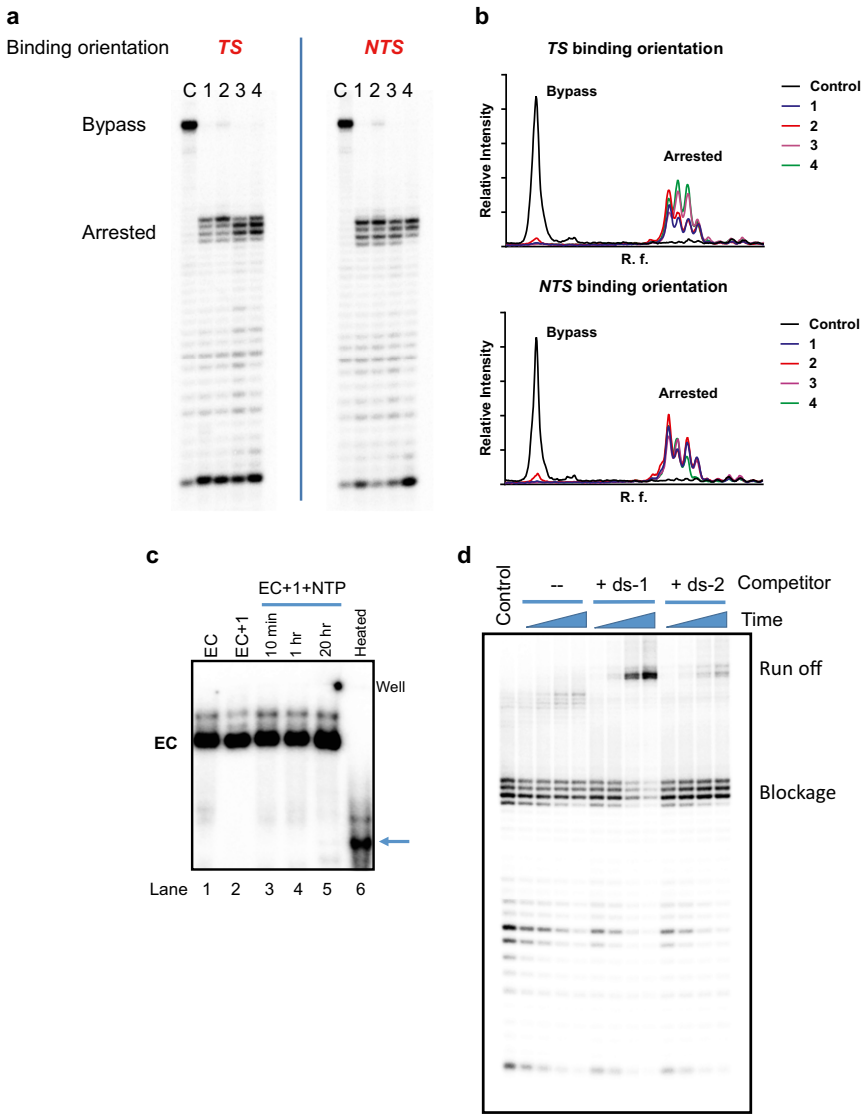

Fig. S1. RNA pol II was persistently arrested by Py-Im polyamides. (*A*) RNA pol II transcription elongation is blocked by Py-Im polyamides in both binding orientations (TS and NTS). C, control lanes in the absence of Py-Im polyamides; **1, 2, 3**, and **4** represent the transcription inhibition results by four compounds listed in Fig. 1*A*. The incubation time for this transcription assay was 20 h. RNA pol II failed to bypass the blockage with prolonged incubation. (*B*) Traces of lanes in the presence of different compounds. (*C*) Pol II elongation complex (EC) was persistently arrested by the polyamides and no obvious free RNA transcript is released. The pol II EC stability experiment was performed in 4% (wt/vol) native PAGE with ^{32}P labeled in the RNA transcript. Lane 1: pol II EC assembled as described in the method in the absence of Py-Im polyamide **1**; lane 2: pol II EC in the presence of Py-Im polyamide **1**; lanes 3–5: Pol II transcription elongation in the presence of **1** after the designated time points (10 min, 1 h, and 20 h, respectively); lane 6: the 20-h time point sample heated at 95 °C for 10 min before loading. Free RNA transcript is released upon heat treatment (indicated by the blue arrow). (*D*) The arrested pol II EC was slowly recovered upon treatment of the DNA duplex competitor. The control lane is the arrested EC in the presence of compound 1 after 1 h transcription elongation. The duplex competitor with 5′-WGGWCW-3′ binding site (ds-1) was subsequently added into the arrested EC mixture. As comparison, the DNA duplex without the designed binding site (ds-2) was incubated with the arrested EC as well. The sample without any competitor (–) was also analyzed. The EC was 50 nM, the compound 1 was 1 μM, and the competitor DNA duplex was 10 μM. Time points were 1, 4, 18, and 42 h.

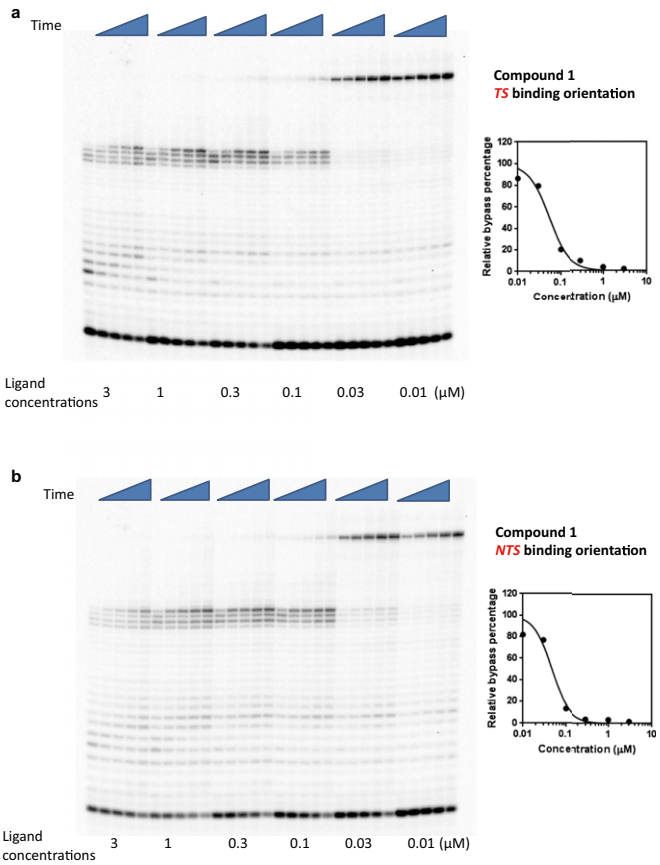

Fig. S2. Concentration-dependent transcriptional inhibition by Py-Im polyamide **1** in TS (*A*) and NTS (*B*) binding orientations. Time points were 1 min, 5 min, 20 min, 1 h, and 3 h.

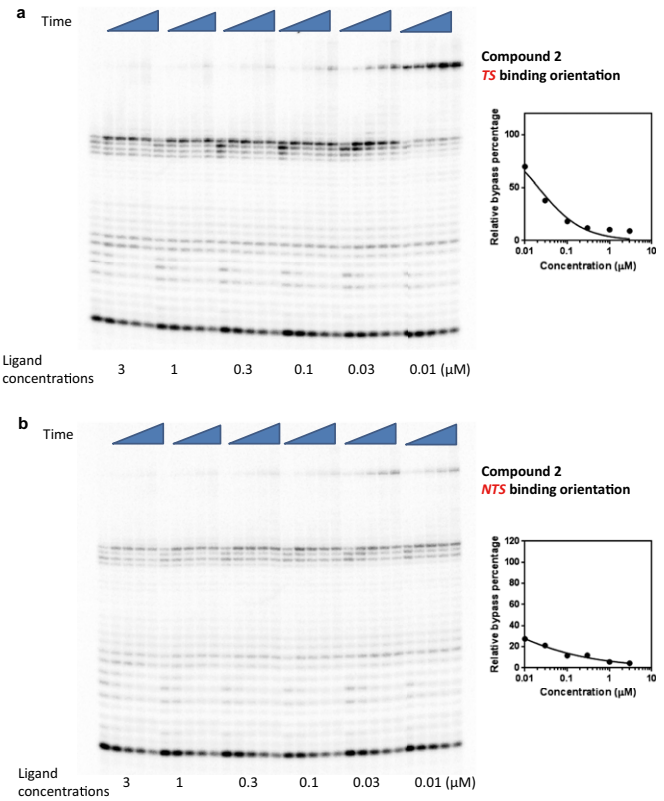

Fig. S3. Concentration-dependent transcriptional inhibition by Py-Im polyamide **2** in TS (*A*) and NTS (*B*) binding orientations. Time points were 1 min, 5 min, 20 min, 1 h, and 3 h.

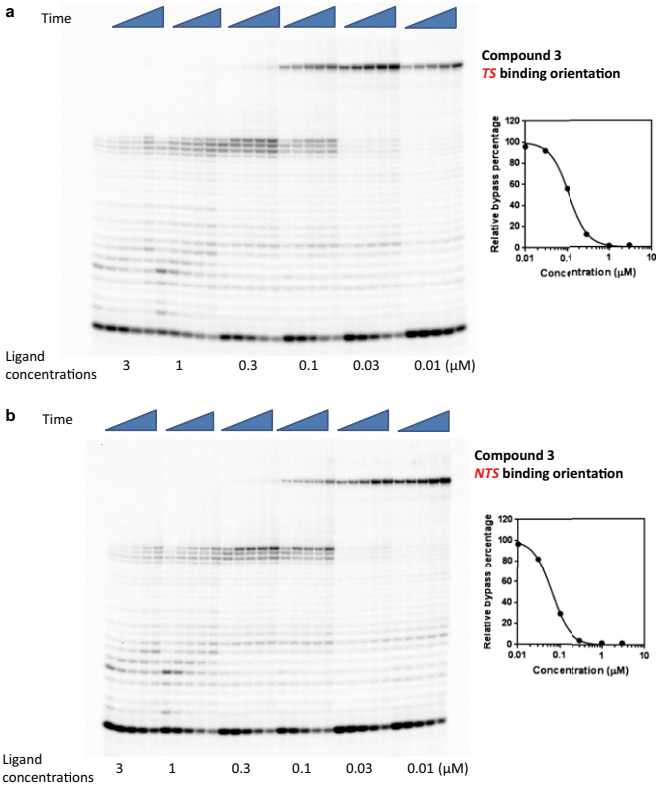

Fig. S4. Concentration-dependent transcriptional inhibition by Py-Im polyamide **3** in TS (*A*) and NTS (*B*) binding orientations. Time points were 1 min, 5 min, 20 min, 1 h, and 3 h.

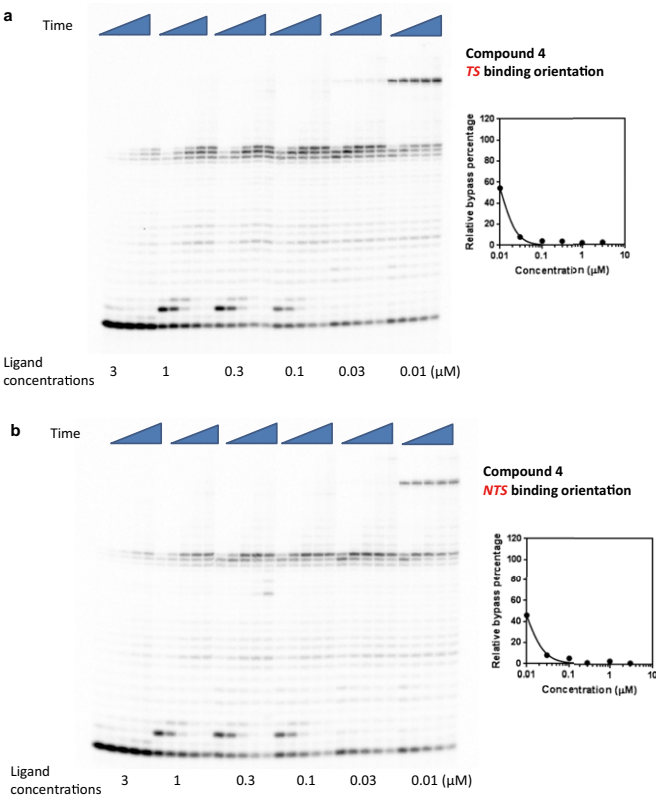

Fig. S5. Concentration-dependent transcriptional inhibition by Py-Im polyamide **4** in TS (*A*) and NTS (*B*) binding orientations. Time points were 1 min, 5 min, 20 min, 1 h, and 3 h.

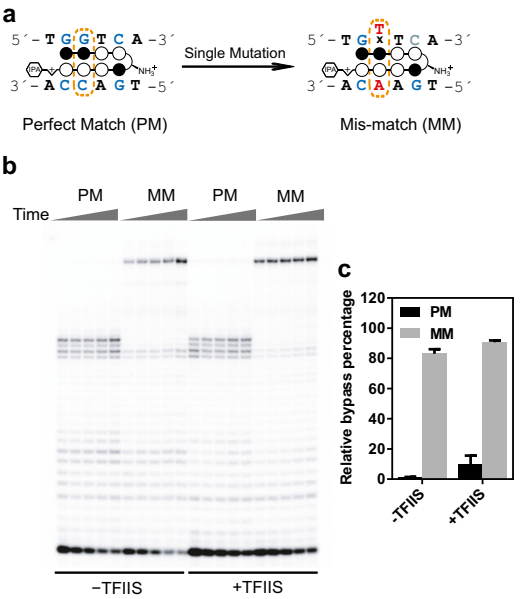

Fig. S6. Py-Im polyamide cause sequence-specific inhibition of pol II transcription elongation. The polyamide binding site is in a NTS orientation. (*A*) Scheme of single mutation of the designed target sequence. The PM sequence is 5′-WGGWCW-3′, whereas the single base MM sequence is 5′-WGWWCW-3′ (mutation underlined). (*B*) Comparison of inhibition of polyamide **1** on pol II transcription elongation from the PM sequence or MM sequence. The effects of TFIIS were also investigated as shown in PAGE gel. Time points were 10 min, 20 min, 30 min, 1 h, and 2 h. Py-Im polyamide concentration was 0.5 μM. (*C*) Quantitative analysis of selective transcriptional inhibition by polyamide **1**.

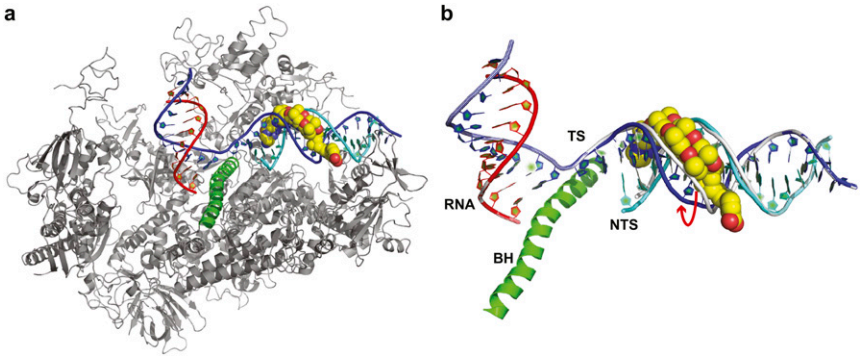

Fig. S7. Molecular modeling of RNA pol II transcriptional pausing and arrest by Py-Im polyamide. (*A*) Overall view of an arrested pol II complex in the presence of Py-Im polyamide. This model view is derived from 10 subunit RNA pol II elongation complex with Rpb2 omitted. The protein residues are colored gray; RNA, template DNA, and nontemplate DNA are shown in red, blue, and cyan, respectively. The small DNA binding molecule is highlighted in yellow CPK mode. (*B*) Py-Im polyamide binding causes a widened DNA duplex minor groove. The Py-Im polyamide bound downstream duplex (blue and cyan) was aligned with regular downstream DNA duplex with pol II elongation complex (gray). The arrow indicates the shifted strand caused by Py-Im polyamide binding.

Table S1. DNA thermal stabilization analysis of compounds 1–4 on TS match and mismatch sequence

	5′-GACT **TGGTCA** TACA -3′		5′-GACT **TGTTCA** TACA -3′	
Polyamide	Tm/°C	ΔTm/°C	Tm/°C	ΔTm/°C
—	56.4 ± 0.7	—	52.9 ± 0.5	—
1	71.3 ± 0.8	14.9	53 ± 0.8	0.1
2	66.6 ± 0.7	10.2	55.9 ± 0.2	3.0
3	71.1 ± 0.8	14.7	54.2 ± 1.0	1.3
4	71.8 ± 0.9	15.4	53.1 ± 0.3	0.2

The compound binding site is indicated in bold font.