pubs.acs.org/acschemicalbiology

Articles

Bifacial PNAs Destabilize MALAT1 by 3′ A-Tail Displacement from the U-Rich Internal Loop

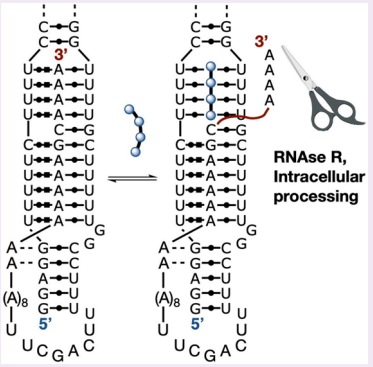
Shiqin Miao, Debmalya Bhunia, Shekaraiah Devari, Yufeng Liang, Oliver Munyaradzi, Sarah Rundell, and Dennis Bong*

Cite This: *ACS Chem. Biol.* 2021, 16, 1600–1609

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We report herein a new class of synthetic reagents for targeting the element for nuclear expression (ENE) in MALAT1, a long noncoding RNA upregulated in many cancers. The *cis*-acting ENE contains a U-rich internal loop (URIL) that forms an 11 base UAU-rich triplex stem with the truncated 3′ oligo-A tail of MALAT1, protecting the terminus from exonuclease digestion and greatly extending transcript lifetime. Bifacial peptide nucleic acids (bPNAs) similarly bind URILs *via* base triple formation between two uracil bases and a synthetic base, melamine. We synthesized a set of low molecular weight bPNAs composed of α-linked peptide, isodipeptide, and diketopiperazine backbones and evaluated their ENE binding efficacy *in vitro via* oligo-A strand displacement and consequent exonuclease sensitivity. Degradation was greatly enhanced by bPNA treatment in the presence of exonucleases, with ENE half-life plunging to 6 min from >24 h. RNA digestion kinetics could clearly distinguish between bPNAs with similar URIL affinities, highlighting the utility of functional assays for evaluating synthetic RNA binders. *In vitro* activity was mirrored by a 50% knockdown of MALAT1 expression in pancreatic cancer (PANC-1) cells upon treatment with bPNAs, consistent with intracellular digestion triggered by a similar ENE A-tail displacement mechanism. Pulldown from PANC-1 total RNA with biotinylated bPNA enriched MALAT1 > 4000× , supportive of bPNA-URIL selectivity. Together, these experiments establish the feasibility of native transcript targeting by bPNA in both *in vitro* and intracellular contexts. Reagents such as bPNAs may be useful tools for the investigation of transcripts stabilized by *cis*-acting poly(A) binding RNA elements.

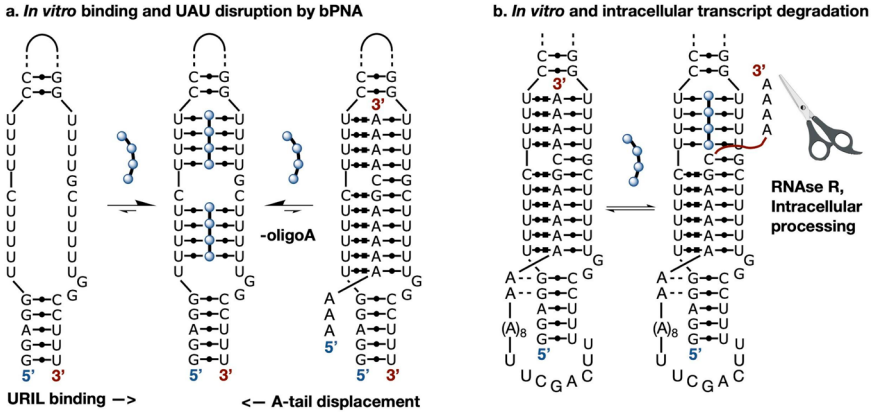


INTRODUCTION

Numerous studies have indicated that *cis*-acting RNA elements such as the element for nuclear expression (ENE) are important in the stabilization of polyadenylated RNAs.[1,2] The ENE extends RNA lifetime by complexation of the 3′ oligo-A terminus[3] in a proximal U-rich internal loop (URIL), thereby protecting the transcript from exonuclease degradation.[4−7] The resulting UAU-rich triplex stem can be bolstered by A-minor contacts between the oligo-A domain and flanking duplex regions, raising the intriguing possibility that *cis*-acting triplex motifs could generally serve to stabilize polyadenylated transcripts, perhaps including those lacking a well-defined 3′ URIL.[1,3] The ENE is found in the 3′ UTR of mRNAs and the 3′ proximal domains of lncRNAs, including metastasis associated lung adenocarcinoma transcript 1 (MALAT-1), a well-studied lncRNA important in the regulation of alternative splicing,[8] cellular proliferation in multiple cancers,[6,9] tumor suppression,[10,11] neuroregeneration[12] as well as other biological functions.[8,13] Structure−function studies by the Steitz lab demonstrated that the C+GC base triple punctuating the UAU triplex stem (Scheme 1) contributes significantly to triplex stability,[6,14] though Brown and co-workers successfully engineered a mutant that folds with an uninterrupted 10 UAU base-triple stem.[15] The unique ENE triplex architecture has been targeted by small molecule library screens,[16,17] a useful method to identify synthetic binders *via* the displacement of RNA-binding fluorophores.[18,19] Binders were found that can both disrupt[17] and enhance[20] ENE triplex stability, resulting in increased and decreased sensitivity to exonuclease digestion, respectively.[17,20] Though powerful for lead generation, these data also show that library screens relying on intercalative binding to RNA targets can sometimes have unpredictable and opposite functional outcomes. For U-rich RNA targets, bPNAs achieve uridine-specific base-triple docking while maintaining both cellular permeability and water solubility. Thus, bPNAs combine appealing aspects of conventional antisense oligonucleotides and small molecules while minimizing common drawbacks of each strategy. A URIL-targeting strategy centered on well-defined base-recognition was expected to be tunable *via* structural iteration; we explore this notion herein with bPNA variants. The ENE itself is an ideal native substrate as

Received: July 26, 2021
Accepted: August 2, 2021
Published: August 12, 2021

Scheme 1. *In Vitro* Functional Assays for bPNA Binding MALAT1 ENE[a]



[a](a) Direct bPNA binding to the URIL alone; bPNA displacement of a 14 nt oligo-A tail from a triplex heterodimer. (b) Enhanced exonuclease degradation of the MALAT1 ENE+A triplex by bPNA displacement of the 3′ oligo-A domain that was carried out both *in vitro* with RNase R and intracellularly (PANC-1).

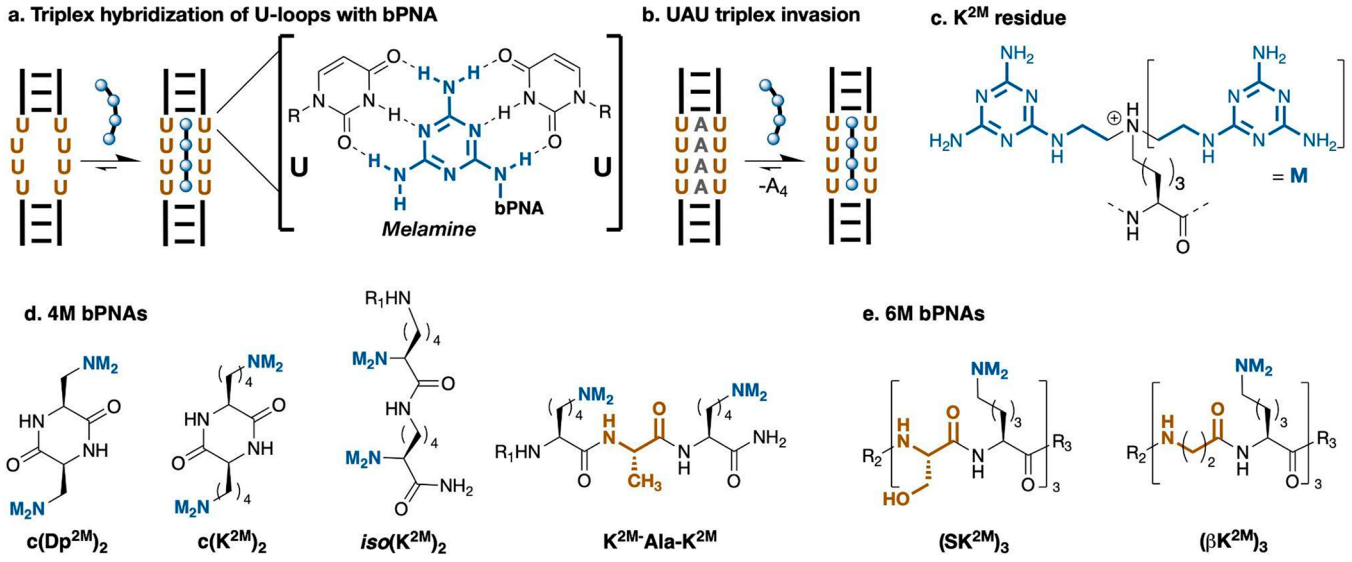


Figure 1. (a) Illustration of bPNA (blue strand) binding to a U_4XU_4 internal bulge in structured RNA *via* UMU base triple formation. (b) Displacement of a oligo-A strand from a UAU triplex to form a bPNA triplex hybrid. (c) K^{2M} lysine derivative used in bPNAs that displays two melamine bases. (d) Backbone variants of 4 M bPNA, from left: diketopiperazines of bismelamine diaminopropionic acid (Dp^{2M}) and K^{2M}; isodipeptide of α-modified K^{2M} and normal α-backbone-linked tripeptide $K^{2M}AK^{2M}$ (R_1 = 4-acetamidobenzamide, acetyl). The isotripeptide (with three residues) iso$(K^{2M})_3$ was also prepared and studied. (e) Normal peptide backbone 6 M bPNAs (R_2 = (5,6)-carboxyfluorescein, Cy5; R_3 = G-$CONH_2$).

the melamine base of bPNA has been rigorously established to engage with two uracil bases to form UMU base triples[21−23] in many contexts and successfully compete with oligo-A domains for oligo-U binding.[24−27] The triazine family of bases, including melamine (M), are known for their assembly properties[28−32] with native bases[32−35] and have been used as artificial base pairs[36−38] in PNA,[39] while melamine can form UMU triples[40] when displayed on small-molecule,[27,40] peptide,[24−26,41−43] peptoid,[44] and polymer[35,41] scaffolds, analogous to native UAU base triples.[45−47] Oligouridine-bPNA triplex hybrid stems exhibit nanomolar dissociation and high thermal stability[26] and can functionally replace native stem elements.[25] Consistent with this prior work, we found that bPNAs could also displace the 3′ oligo-A strand from the MALAT1 triplex to form a triplex hybrid stem with the URIL. In addition, we tested the extent to which the base-driven binding mode could sustain backbone optimization without catastrophic loss of function and found that RNA binding efficacy varied as bPNA backbone topology was changed from acyclic, cyclic, normal, and isopeptide. A much lower tolerance was observed for variation in side chain length, with RNA binding ablated upon shortening the linkage of the melamine base to the normal peptide backbone. To evaluate the backbone variants, we developed *in vitro* assays that follow the displacement of the native oligo-A sequence from the ENE U-loop as well as RNase degradation experiments. These functional assays enabled sensitive discrimination between the backbone variants and provided support for the proposed mechanism of intracellular MALAT1 silencing *via* 3′ A-tail displacement. The ability of bPNA to selectively target native MALAT1 in heterogeneous milieu was further supported by MALAT1 pulldown from total RNA using biotinylated bPNA. Taken together, this study establishes the efficacy of bPNA in targeting native intracellular RNAs and highlights the utility of

https://doi.org/10.1021/acschembio.1c00575
ACS Chem. Biol. 2021, 16, 1600−1609

functional, enzyme-linked assays in the evaluation of synthetic RNA binders to the MALAT1 ENE.

RESULTS

Design of bPNAs. To target the 4 and 5 UAU base triple stems of the ENE, we synthesized 4 M bPNA backbone variants containing two modified residues, K^{2M} or Dp^{2M}, for RNA binding; these bPNAs are the most compact to-date. These lysine (K) and diaminopropionic acid (Dp) derivatives were synthesized *via* double reductive alkylation with melamine acetaldehyde (Figure 1), resulting in 4 M bPNAs with 4 melamine bases and 2 cationic charges when protonated. Variants were designed to probe the effect of backbone conformation and length of the base-triple side chain on URIL binding. Most analogous to prior work is the α-linked tripeptide, $K^{2M}AK^{2M}$, a truncated version of the 6 M scaffold $(XK^{2M})_3$ previously reported.[48] A novel diketopiperazine scaffold was prepared by the cyclization of $(K^{2M})_2$ and $(Dp^{2M})_2$ dipeptides to yield $c[K^{2M}]_2$ and $c[Dp^{2M}]_2$, respectively. The diketopiperazines are expected to adopt a constrained boat-like conformation,[49] projecting the side chains axially into the bPNA−RNA interface, with $(Dp^{2M})_2$ demanding a more intimate contact as a result of a side chain three carbons shorter than lysine. We also prepared an isodipeptide bPNA by installing the melamine bases on the lysine α-nitrogen and linking two residues *via* the side chain ε-amine to yield $iso(K^{2M})_2$. This variant was expected to have a more flexible backbone as a result of the *n*-butyl backbone segment and also features shortened base linkages due to α-*N* modification. It was anticipated that the normal peptide, diketopiperazine, and isopeptide variants would sample distinct backbone conformations space while maintaining identical charge (2+) and base content (4M), enabling the comparison of backbone impact on URIL RNA-targeting. In addition, normal peptide 6 M bPNAs $iso(K^{2M})_3$ and $(XK^{2M})_3$ were prepared with *N*-terminal dye modification (Figure 1) for use in labeling experiments (*vide infra*) and functional assays.[48]

Binding to MALAT1 URIL. We tested binding of this family of bPNA backbone variants to an *in vitro* transcribed MALAT1 URIL duplex that lacked the 3′ A-tail (Scheme 1a). Our data is consistent with bPNA-URIL triplex hybridization *via* UMU base triple formation, similar to the function of the native oligo-A sequence that forms UAU base triples in MALAT1. The MALAT1 URIL duplex has considerable structure even without the oligo-A triplex domain, with a significant thermal transition observed at 82 °C and a weaker broad transition ∼60 °C that may be, respectively, attributed to the canonical duplex and the significantly structured U-rich internal bulge.[50,51] Heterodimerization with an oligo-A strand (Figure 2) forms a UAU triplex, as reflected by two distinct and sharp transitions at 62 and 82 °C.[6] This first transition increases to 68 °C upon hybridization with 4 and 6 M K^{2M} bPNAs (normal peptide, diketopiperazine, isodipeptide) (Figure 2), suggestive of greater thermal stability relative to the UAU triplex. In contrast, the diketopiperazine $c[Dp^{2M}]_2$ did not elicit the emergence of a new transition at 68 °C (Supporting Information).

Clear electrophoretic mobility shifts could be observed by gel upon treatment of the URIL RNA with 4 M bPNA derivatives, even though the RNA is much larger than the bPNAs (Figure 3). It was more difficult to resolve the binding of the 6 M bPNAs, possibly because the increased cationic charge of the 6 M bPNAs narrows an already subtle

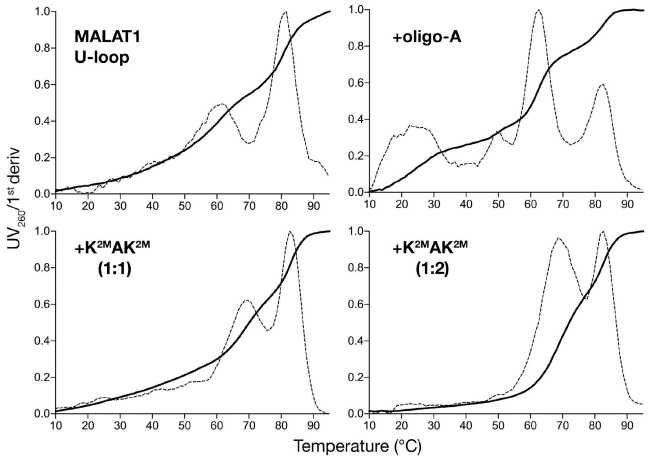


Figure 2. Thermal denaturation studies on URIL duplex alone (top, left), with oligo-A at 1:1 (top, right), with $K^{2M}AK^{2M}$ at 1:1 and 1:2 RNA to bPNA ratios (lower row). Melting curves shown in bold and first derivatives in dashed lines. RNA concentration was held constant at 2 μM, in 20 mM HEPES-KOH and 1 mM Mg^{2+}, pH 7.4.

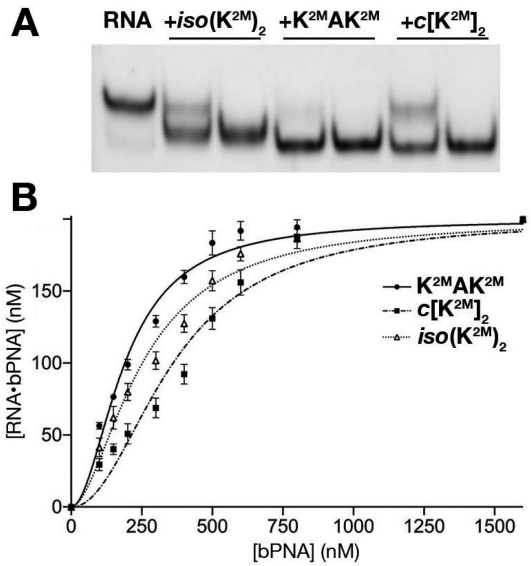


Figure 3. (A) Native gel of MALAT1 U-loop (RNA, 500 nM) treated bPNAs as indicated at 1 and 4 μM from left to right (SYBR Gold). (B) EMSA derived binding curves of the indicated bPNAs.

electrophoretic shift. Nevertheless, it was possible to affirm the good affinity binding of the 4 M bPNA tripeptides to the MALAT1 URIL, with the exception of $c[Dp^{2M}]_2$, again indicative of a loss of binding due to the shortened side chain. As expected from the 9 UAU sites, a 1:2 model (Figure 2) gave a better fit to the binding isotherm than a 1:1 model (Supporting Information), supportive of two 4 M bPNAs binding to the MALAT1 URIL. Though all backbones exhibit good URIL-binding, curve-fitting analysis indicated the highest affinity binding from the Ala tripeptide ($K^{2M}AK^{2M}$) followed by isodipeptide ($iso(K^{2M})_2$ and diketopiperazine ($c[K^{2M}]_2$), with apparent dissociation constants in the 100−200 nM range (Table 1).

A-Tail Displacement. With affinity to the URIL confirmed, we turned our attention to the competitive displacement of the oligo-A tail from the ENE triplex. A heterodimeric ENE triplex (Scheme 1) was treated with bPNA

https://doi.org/10.1021/acschembio.1c00575
ACS Chem. Biol. 2021, 16, 1600−1609

Table 1. Biophysical and Functional Data for bPNAs Binding to MALAT1 RNA[a]

bPNA	*app* K_d (nM)	A-tail EC_{50} (μM)	$t_{1/2}$ (h) exo T	$t_{1/2}$ (h) RNase R
none			>24	>24
$K^{2M}AK^{2M}$	109 ± 3	0.23 ± 0.01	0.56 ± 0.05	0.50 ± 0.08
iso$(K^{2M})_2$	136 ± 14	0.92 ± 0.03	n.d.	0.24 ± 0.06
c$[K^{2M}]_2$	214 ± 20	0.48 ± 0.03	0.67 ± 0.25	0.11 ± 0.03
c$[Dp^{2M}]_2$	n.d.	9.24 ± 0.93	>24	>24
$(SK^{2M})_3$	n.d.	0.86 ± 0.03	1.0 ± 0.4	2.15 ± 0.64
$(\beta K^{2M})_3$	n.d.	0.87 ± 0.03	0.51 ± 0.16	0.65 ± 0.17

[a]Data from triplicate runs. Binding and A-tail displacement determined *via* native gels with 5′ ^{32}P-labeled MALAT1 U-loop and oligo A-tail RNA, respectively. Degradation of intramolecular MALAT1 triplex with exo T and RNase R assessed *via* radiogel quantification as described.

variants and the displacement of the ^{32}P-labeled oligo-A strand tracked by native gel. Interestingly, while shorter A-tail oligos were ineffective in displacement (Supporting Information), the backbone library displayed a range of displacement efficacies, with the 4 M tripeptide $K^{2M}AK^{2M}$ exhibiting the lowest EC_{50} (0.23 μM) followed by diketopiperazine c$[K^{2M}]_2$ and isodipeptide, with EC_{50} values of 0.48 and 0.92 μM, respectively (Table 1). Again, side chain length of the base-tripling residue was critical, with a significant loss of displacement observed in the diaminopropionic acid derivative c$[Dp^{2M}]_2$ relative to c$[K^{2M}]_2$. The isotripeptide *iso*$(K^{2M})_3$ exhibited a lower EC_{50} (0.3 μM) than *iso*$(K^{2M})_2$, suggestive of functional gains for extending the isopeptide backbone. However, it remains unclear if these gains come from electrostatic or base interactions (Figure 4).

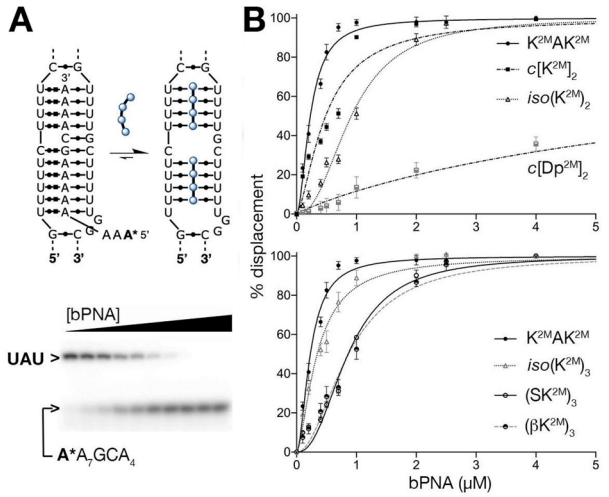


Figure 4. (A) Illustration (top) of oligo-A tail displacement from the UAU triplex by two equivalents of bPNA and (below) a representative native gel with the UAU triplex and free 5′-radiolabeled (A*) oligo-A strand, with $K^{2M}AK^{2M}$ bPNA. (B) Fractional displacement of oligo-A from a 1:1 UAU heterodimeric triplex (100 nM) from the quantification of radiolabeled bands of (top) 4 M bPNAs and (below) 6 M bPNAs as indicated. Displacement conditions: 25 mM sodium cacodylate pH 7.0, 50 mM KCl, 1 mM $MgCl_2$.

Treatment with Exonucleases. We hypothesized that quantitative A-tail displacement may not be needed to diminish MALAT1 lifetime; increased off-rate[51] should increase 3′ exonuclease degradation *via* transient exposure of the 3′ terminus. Indeed, the extended transcript lifetime *in vivo* is attributed to the protection of the 3′ end in the triplex core, and small molecules that disrupt triplex stability have been shown to increase transcript susceptibility to exonucleases.[17]

We tested the extent to which bPNA could increase exonuclease sensitivity by incubation of the MALAT1 triplex core with both bPNA and commercially available nucleases, exo T and RNase R. These experiments were carried out well below A-tail displacement EC_{50} values.

Exonuclease T (exo T) requires high (6−10 mM) Mg^{2+} for good activity; under these non-native conditions, the MALAT1 ENE is highly resistant to degradation.[51] Despite this, we observed a rapid and quantitative exo T degradation of full-length RNA within 30 min at bPNA concentrations below displacement EC_{50} while transcript was stable in the absence of bPNA (Figure 5). Interestingly, degradation was halted at a common cleavage product, which corresponded in length to a loss of ∼30 nt. Reminiscent of prior studies that indicated the limited inability of exo T to read-through bPNA-RNA triplex and secondary structure,[42] the length of truncated RNA corresponds to blockage of exo T at the beginning of the duplex domain. We therefore expected that exo T was

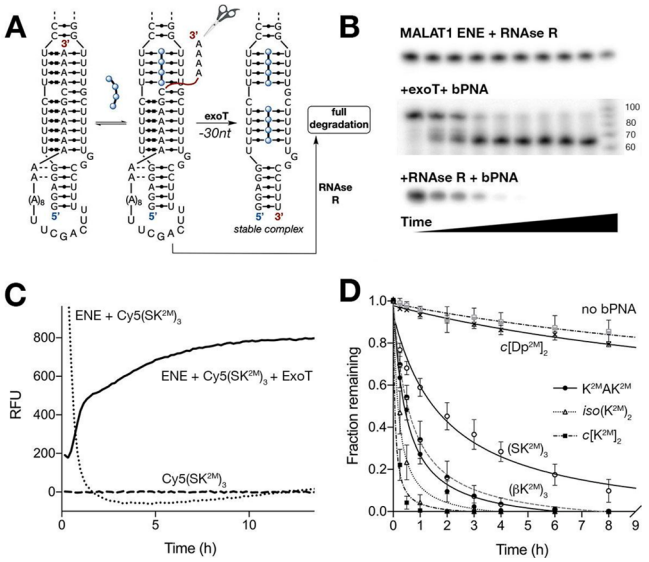


Figure 5. (A) Illustration of intramolecular A-tail displacement by bPNA leading to partial digestion with exo T and full degradation with RNase R. (B) Representative denaturing gels of 5′ ^{32}P-labeled MALAT1 ENE (top) with RNase R, (middle) with exo T and $K^{2M}AK^{2M}$ bPNA, and (bottom) with RNase R and $K^{2M}AK^{2M}$. Lanes from left to right represent time-points taken at 0, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, and 4 h (exo T) and 0, 0.25, 0.5, 1, 2, 3, 4, 5, 6, and 8 h (RNase R). (C) Cy5 emission intensity as a function of time upon treatment of MALAT1 ENE with Cy5$(SK^{2M})_3$, Cy5$(SK^{2M})_3$ alone, and ENE with Cy5$(SK^{2M})_3$ and exo T as indicated. (D) Degradation of MALAT1 ENE upon treatment with RNase R and bPNAs as indicated.

https://doi.org/10.1021/acschembio.1c00575
ACS Chem. Biol. 2021, 16, 1600−1609

facilitating the insertion of bPNAs into the RNA fold, resulting in a bPNA-bound degradation product (Figure 5). This notion was confirmed by the use of Cy5-labeled 6 M bPNA (Cy5-$(SK^{2M})_3$). The triplex hybridization of Cy5-$(SK^{2M})_3$ with RNA can result in a significant increase of Cy5 emission intensity;[48] indeed, Cy5 emission reflected RNA cleavage kinetics, increasing sharply in the first 30 min and then remaining stable for hours afterward, supportive of a stable RNA−bPNA hybrid.

In contrast to exo T, RNase R has a larger binding footprint and higher readthrough capability. While the MALAT1 triplex was again stable to RNaseR treatment ($t_{1/2}$ > 24 h), treatment with bPNAs resulted in a greatly accelerated *full* degradation of transcript, with $t_{1/2}$ dropping as low as 0.1 h (Table 1). Degradation kinetics were similar but clearly distinct for different bPNAs. Consistent with prior experiments, the least efficient bPNA was c$[Dp^{2M}]_2$, affirming the ineffectiveness of the shorter side chain in triplex hybridization.

Silencing of MALAT1 in PANC-1 Cells. Having established the *in vitro* activity of bPNAs against the MALAT ENE in binding, displacement, and exonuclease assays, we set out to test the extent to which bPNAs could trigger the intracellular exonuclease degradation of MALAT1 in pancreatic cancer cells (PANC-1), which exhibit elevated MALAT1 expression.[52] The intracellular transport of cationic bPNAs is known,[53] raising expectations that the targeting of native MALAT1 would be feasible. For reasons of synthetic accessibility, we focused our intracellular studies on the side-chain-modified subset of the bPNA library rather than the α-modified isodipeptides. PANC-1 cells were treated with $(SK^{2M})_3$, $(\beta K^{2M})_3$, $K^{2M}AK^{2M}$, c$[K^{2M}]_2$, and c$[Dp^{2M}]_2$ bPNAs in media and sampled after 1 and 2 days of treatment and MALAT1 levels in isolated total RNA were assessed. Quantitative reverse transcription-polymerase chain reaction (RT-PCR) revealed an approximate 50% reduction in MALAT1 transcript levels relative to the internal standard housekeeping gene GAPDH, with the important exception of c$[Dp^{2M}]_2$ (Figure 6). The reduction in MALAT1 was similar for all active bPNAs, with increased silencing on day 2 vs day 1. In addition to stable housekeeping gene levels, cell viability was unaffected over two day treatment with the same suite of bPNAs (Supporting Information).

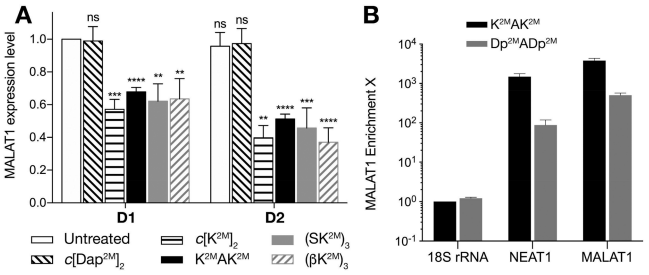


Figure 6. (A) Expression levels of MALAT1 in PANC-1 cells from qRT-PCR analysis following treatment of cells in media with 1% FBS and 1 μM bPNA and harvested after 1 (D1) and 2 days (D2). MALAT1 transcript levels were normalized against GAPDH housekeeping gene and triplicated with error bars and *P*-values as indicated. Non-significant (ns) differences from untreated cells were found for the diketopiperazine of Dp^{2M} (c$[Dp^{2M}]_2$. (B) Fold enrichment shown in the log scale of MALAT1 and NEAT1 from PANC-1 total RNA following pulldown with biotin-modified $K^{2M}AK^{2M}$ and $Dp^{2M}ADp^{2M}$, normalized against 18S rRNA,.

Affinity Pulldown Enrichment of MALAT1 from PANC-1 Total RNA. To further substantiate MALAT1 targeting, we carried out affinity pulldown experiments using *N*-terminally biotinylated Ala tripeptide (Btn-$K^{2M}AK^{2M}$). As whole cell treatment with bPNA depletes MALAT1 (Figure 6), we instead treated total RNA isolated from PANC-1 cell lysate with Btn-$K^{2M}AK^{2M}$. Following incubation with bPNA, total RNA samples were subject to affinity pulldown with streptavidin-agarose beads. Beads were washed and qRT-PCR was carried out directly on beads according to published protocols.[54] Remarkably, Btn-$K^{2M}AK^{2M}$ pulldown yielded a >4000× enrichment of native MALAT1 over GAPDH or 18S rRNA normalized background in the heterogeneous milieu of total RNA. A broader selectivity study was carried out by subjecting PANC1 total RNA to pulldown using Btn-$Dp^{2M}ADp^{2M}$ in addition to Btn-$K^{2M}AK^{2M}$ and quantifying the enrichment of 80 RNAs in addition to MALAT1. These included 71 high abundance RNAs (rRNAs, mRNAs, tRNAs, and small RNAs),[58] 9 noncoding U-rich transcripts and MALAT1. Among the 9 U-rich RNAs, three selected from the NONCODE database (http://noncode.org/download.php) contain a 5′-UUUUUCUUUUZUUUU-3′ sequence (Lnc-MAST4-5:14, ZNF674-AS1:10, PTCSC-AS1:10) and have been previously studied[55] as triplex forming RNAs. In addition, four mRNAs (HIST1H3E, POLR3K, ATP6V0B, and TNFα) containing AU-rich elements[56] were selected from the ARED database (https://brp.kfshrc.edu.sa/ared), along with an mRNA (FAU)[57] containing U-rich repeats $(UCU)_6$. Importantly, we also included NEAT1 (MENβ), which contains an ENE triplex motif and a URIL homologous to that found in MALAT1.[6] The $Dp^{2M}ADp^{2M}$ bPNA exhibits a much weaker melting than $K^{2M}AK^{2M}$ with U-rich RNAs, like the diketopiperazine analogue (Supporting Information). As expected, the Btn-$Dp^{2M}ADp^{2M}$ probe is ∼10× less efficient at MALAT1 pulldown relative to Btn-$K^{2M}AK^{2M}$, normalized against 18S rRNA (Figures 6B and 7 inset). Consistent with URIL-targeted binding, both bPNA probes significantly enriched MALAT1 and NEAT1 from total RNA (Figure 7) over 18S rRNA (Figures 6B and 7). Furthermore, a markedly lower enrichment was observed with all other 79 transcripts, including high abundance, triplex-forming, and U-rich RNAs, strongly supportive of URIL targeting selectivity. Overall, this result validates the intracellular relevance of *in vitro* targeting experiments performed on the ENE+A core domain.

DISCUSSION

Thermal denaturation, gel shift, and A-tail displacement studies clearly establish the expected binding of bPNAs to the URIL duplex and invasion of the ENE. Effective displacement is observed, even though bPNAs do not address the register-defining native CG/C+GC base pair/base triple motif nested within the URIL (Scheme 1). Steitz and co-workers reported this motif to be critical,[14,59] with a significant loss of ENE stability upon substitution of the GC dyad with AA in the tail or replacement the CG/C+GC motif with AU/UAU.[6] In the heterodimeric RNA triplex, we found that the abasic substitution of one or both A-tail GC positions decreases affinity ∼5× and 50×, respectively, consistent with a loss of cooperative stabilizing interactions (Supporting Information). We speculated that the similar pK_as of protonated melamine (5.0) and cytidine (4.4) might allow M+ to take the place of C+. As the ENE domain has 4 and 5 UAU base triple runs separated by the CG/C+GC motif, the opportunity to test this

https://doi.org/10.1021/acschembio.1c00575
ACS Chem. Biol. 2021, 16, 1600−1609

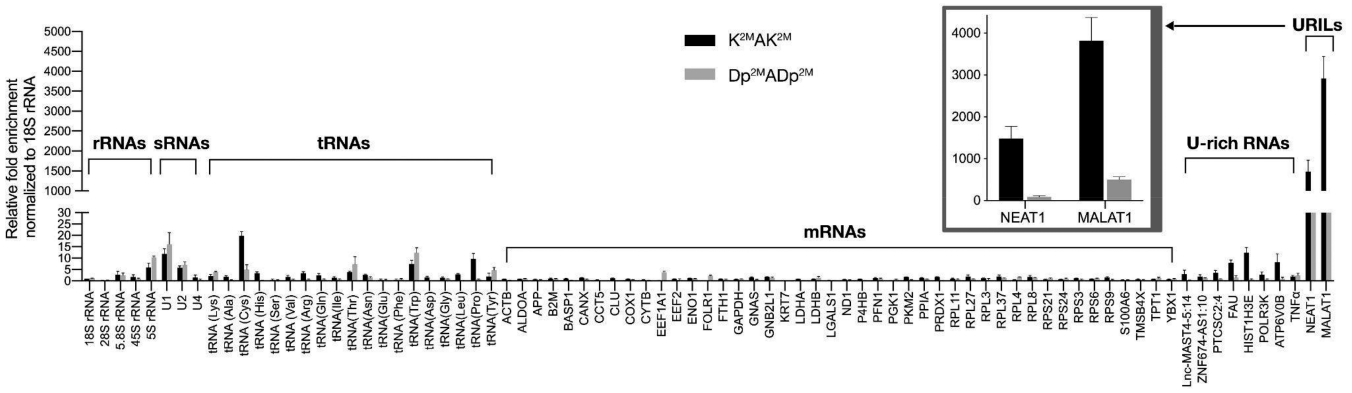


Figure 7. Pulldown selectivity using biotinylated bPNAs $K^{2M}AK^{2M}$ and $Dp^{2M}ADp^{2M}$ among selected RNAs in PANC-1 total RNA, from qRT-PCR experiments. Enrichment is normalized to 18S rRNA, with distinct classes of RNAs as indicated. Enrichment of URIL transcripts NEAT1 and MALAT1 are enlarged and shown inset.

notion arose with 6 M bPNAs that were not expected to be uridine-saturated. However, despite one additional cationic charge and at least one "unused" M base, 6 M bPNAs do not exhibit significant advantages over 4 M bPNAs in binding or functional assays. This may be due in part to steric issues raised by the larger *N*-terminal dyes on the 6 M bPNAs used for labeling; we thus focus our comparative discussion on 4 M variants with acetyl, benzamide, or no capping group (diketopiperazine).

Though all 4 M backbones show significant binding, the 4 M tripeptide $K^{2M}AK^{2M}$ generally outperforms all bPNAs, suggesting a preference for the normal peptide backbone in binding the ENE URIL. In exonuclease degradation, the 4 M isodipeptide *iso*-$[K^{2M}]_2$, diketopiperazine c$[K^{2M}]_2$, and $K^{2M}AK^{2M}$ tripeptide bPNAs all decreased $t_{1/2}$ to less than 0.5 h (Table 1). The dye-labeled 6 M bPNAs, in particular $(SK^{2M})_3$, were less efficient in all assays than the 4 M bPNAs; again, possibly due to steric issues caused by the dyes. As expected, the combination treatment of RNA with bPNAs and exonucleases can trigger rapid degradation; in the case of exo T, this facilitates bPNA insertion into the URIL. These effects are readily observed even under high magnesium conditions where RNA folds are stabilized and at bPNA concentrations well above displacement EC_{50} values. General trends for the binding and displacement, exo T and RNase R assays were similar, attesting to the consistency of binding and functional measurements. Deviations were also observed: while the tripeptide $K^{2M}AK^{2M}$ exhibited the best binding and displacement activity, diketopiperazine *c*$[K^{2M}]_2$ and isodipeptide *iso*$(K^{2M})_2$ triggered faster RNase R degradation. We speculate that weaker binders enable more efficient RNase R read-through and degradation, resulting in a reversal of the activity ranking (Table 1). Furthermore, though $t_{1/2}$ values are close, degradation kinetics for each bPNA are clearly distinct (Figure 5). Discrimination between binders with relatively similar affinity and subtle reversals of ranking underscore the need for context-specific functional assays for library screening.

Whereas all backbones showed distinct but comparable activity, the side chain-shortened diketopiperazine c$[Dp^{2M}]_2$ showed weak or undetectable activity in every assay, strongly suggesting that RNA binding is more sensitive to side chain sterics than backbone conformation.[41] Thus, *c*$[Dp^{2M}]_2$, with identical functional groups and charge as *c*$[K^{2M}]_2$ and other 4 M bPNAs, serves as an effective negative control. Importantly, the lack of *c*$[Dp^{2M}]_2$ silencing activity against MALAT1 in PANC-1 cells was consistent with *in vitro* outcomes, lending credence to the 50% MALAT1 knockdown observed with K^{2M}-derived normal, isodipeptide, and diketopiperazine bPNAs (Figure 6). Together with a significant (~4000×) MALAT1 enrichment from PANC-1 total RNA upon biotin-bPNA pulldown, the data converge on a model for intracellular silencing that mirrors *in vitro* assays: bPNA variants are transported into the cellular compartment, where they target native MALAT1 ENE and displace the 3′ oligo A-tail, resulting in native exonuclease transcript degradation. We note that the ENE URIL is an ideal binding bPNA site;[53] not all T-/U-rich sequences are effectively bound.[53] Though we anticipate that there may be capture of a U-rich subset[60] of the transcriptome in addition to MALAT1 and NEAT1, our targeted selectivity study (Figure 7) indicates that U-rich transcripts lacking the URIL are not significantly pulled down, suggestive of motif-selectivity. It is intriguing that NEAT1 (MENβ) is not enriched to the same extent as MALAT1, despite its nearly identical ENE motif; this may be due to differential expression levels.[6,61,62] Further experiments on transcriptome-wide bPNA selectivity are ongoing that may identify novel URIL and ENE-like motifs[4,63] in transcripts that are excluded from poly-A pulldown.[64]

CONCLUSIONS

Overall, the studies herein establish bPNAs as useful tools for intracellular targeting of URIL-containing motifs in native RNAs such as the ENE in MALAT1. The bPNA strategy combines appealing aspects of antisense and small molecule approaches with sequence-specific base interactions, low molecular weight (<1 kDa), scalable synthesis, and intracellular penetration. Though the selective targeting of the ENE is clear from *in vitro* binding studies, competitive 3′ A-tail displacement, pulldown enrichment and MALAT1 knockdown in PANC-1 cells, transcriptome-wide effects, and cellular distribution of bPNAs remain unknown. In particular, though MALAT1 is nuclear retained,[65,66] it is transported throughout the cell and found in exosome fractions;[12] thus, targeting does not imply nuclear localization of bPNA.[12,67] Despite these caveats and remaining questions, the data presented indicate a practical application of bPNAs as a new class of URIL-targeting reagents. Furthermore, functional assays indicate that good RNA binding activity can be attained with 4 M bPNAs with diverse peptide backbone topology, enabling future work to be carried out on compact RNA URIL motifs with synthetically

https://doi.org/10.1021/acschembio.1c00575
ACS Chem. Biol. 2021, 16, 1600−1609

accessible 2−3 residue peptides. Given the diversity of U-/A-/AU-rich elements in regulation,[68,69] we anticipate that the further engineering of bPNAs will lead to useful reagents for the modulation of critical processes such as transcript stabilization.[1]

MATERIALS AND METHODS

General. All chemicals were used without further purification from commercial sources unless otherwise noted. DNAs and RNAs were purchased from Integrated DNA Technologies (IDT), and SYBR gold was purchased from Thermo Fisher Scientific. Stock solutions (DNA, RNA) were serially diluted in deionized water, and concentrations were determined by UV (260 nm) with a Thermo Fisher Nanodrop 2000. RNA constructs were prepared by *in vitro* transcription except where noted.

Solid Phase bPNA Synthesis and Purification. The peptide synthesis of bPNAs was performed manually using Rink-amide resin (100−200 mesh, 0.3 mmol/g, 150 mg per batch) and standard Fmoc chemistry protocols. Coupling steps were carried out at a 0.2 M amino acid final concentration with 1 equiv of PyAOP and DIEA in 2 mL of NMP. Dyes were coupled on-resin using three equivalents of dye, 3.3 equiv of HBTU, and 3.3 equiv of DIEA in 2 mL of DMF. Fmoc cleavage was performed with 2 mL of piperidine/NMP (1:1) with 3% DBU. Acetyl capping was performed with 2 mL of acetic anhydride/pyridine (3:2) in NMP. Coupling reagents were prereacted (15 min) before addition to resin. Peptides were cleaved from solid support (95% trifluoroacetic acid (TFA), 5% H_2O) over 2 h, and the peptide was precipitated and washed with cold Et_2O two times and dried over a vacuum. Crude peptides were dissolved in solvent A and purified to homogeneity by high performance liquid chromatography (HPLC) (semiprep C18, 8 mL/min, 230/238 nm, solvent A = 0.1% TFA in water, solvent B = 0.07% TFA, 90% acetonitrile, 10% water), as judged by MALDI-TOF and analytical HPLC. Peptides were biotinylated in solution (0.02 M purified bPNA, 0.2 M biotin-PEG-NHS ester, 0.2 M DIPEA, 500 μL of DMF, ON, RT) and purified by HPLC.

Preparation of RNAs. MALAT1 triplex and URIL duplex RNA were synthesized by T7 *in vitro* transcription from PCR-amplified DNA templates.[16] Typical PCR protocol: The final concentration template was 1 ng/μL with 0.2 mM dNTP mix, 0.5 μM forward and reverse primer, 0.01 unit/μL Pfu,[70] which were mixed in 1X Pfu buffer (20 mM Tris-HCl, pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100 and 0.1 mg mL^{-1} nuclease-free BSA) and thermocycled (98 °C/30 s, 40x (98 °C/30 s, 48 °C/15 s (URIL duplex) or 50 °C/15 s (ENE+A), 72 °C/30 s, 72 °C/2 min, hold 4 °C). Samples were combined and purified using a MinElute PCR Purification Kit (Qiagen) and quality-checked (1.5% agarose gel, EtBr, 50 bp DNA ladder, Invitrogen). The concentration of PCR products was determined by UV. Typical transcription protocol: RNA transcription buffer 10X: 1 M HEPES-KOH pH 7.5, 20 mM Spermidine-HCl, and 400 mM DTT. Transcription was performed with 1X buffer with an additional 10 mM DTT, 50 mM $MgCl_2$, 20 mM rNTP, 10 ng/μL PCR product, 2 μL of 100 U/mL inorganic yeast pyrophosphatase (New England Biolabs), and 5 μL of T7 polymerase[71] stock per 100 μL of transcription reaction solution (conditions subject to optimization on the basis of sequences). Transcription was incubated at 37 °C overnight and then treated with DNase 1 (New England Biolabs, 30 min, 37 °C), quenched with EDTA (1.5 equiv), and extracted (phenol/chloroform/isoamyl alcohol, 25:24:1, pH 6.7/8.0 saturated liquid, Fisher BioReagents). An equal volume of 2X TBE/urea loading buffer was added, and the sample was then heated (95 °C, 10 min), cooled in ice, and loaded onto 20% (19:1 acrylamide/bis(acrylamide)) 8 M denaturing gel (8.3 cm × 7.3 cm × 1.5 mm). The RNA was visualized by UV shadowing, and desired gel bands were cut and RNA electroeluted from gel slices (Spectra/Por-3 Standard RC Dry Dialysis Tubing, Spectrum Laboratories, 1X TBE buffer, 150 V, 2 h). RNA solution was removed from dialysis tubing and precipitated (NaOAc, 0.3 M final concentration, 2.5 volumes of 200 proof ethanol, −20 °C, ≥ 2 h). RNA was centrifuged down at high speed (20 min, 4 °C), and the pellet was washed (500 μL of cold 70% ethanol) and centrifuged again at a high speed (10 min, 4 °C). RNA concentration was determined by UV (Nanodrop) using extinction coefficients determined by the IDT OligoAnalyzer. Sequence length and purity were checked by gel (20% PAGE TBE-urea).

Thermal Denaturation (UV). Melts were carried out on a Cary-100 UV−vis spectrometer with Peltier temperature control (1 °C/min) and monitored at 260 nm. Samples were freshly annealed (95 °C, 5 min followed by slow cool) in 20 mM HEPES-KOH, pH 7.4, 1 mM Mg^{2+} before measurement in the same.

Radiolabeling. RNA was dephosphorylated using CIAP (Calf Intestinal Alkaline phosphatase) following Invitrogen protocol (150 pmol of RNA, 1.5 units CIAP, 75 μL of 1X CIAP buffer/1 M diethanolamine, 10 mM 4-nitrophenyl phosphate, 0.25 mM $MgCl_2$, pH 9.8). The reaction was incubated (37 °C, 15 min), heat-inactivated (65 °C, 20 min), extracted with phenol/chloroform/isoamyl alcohol (25:24:1, v/v) (Invitrogen) and ethanol precipitated to remove CIAP. End labeling was performed according to NEB protocol: 150 pmol of RNA, 30 pmol [γ-^{32}P] of ATP (250 μCi, 3000 Ci/mmol 10 mCi/ml, PerklinElmer), 3 μL of T4 PNK kinase (10,000 units/mL), and 50 μL of 1X T4 PNK Reaction Buffer (70 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM DTT). The reaction was incubated (1.5 h, 37 °C) and heated (20 min, 72 °C), and then, excess radiolabel was removed using a P30 spin column. Stock concentrations were determined by UV (260 nm).

EMSA: Binding and Displacement. Samples were prepared by mixing varying concentrations of bPNA with a constant concentration of RNA in 1X DPBS buffer. Samples were annealed (95 °C, 5 min, slow cool), run on 15% native acrylamide gel (150 V, 4 °C), stained with SYBR gold, and imaged. Radiolabeled oligo-A strands (see Supporting Information for sequences) were mixed in varying concentrations with a constant concentration of MALAT1 URIL duplex (50 nM, 1X binding buffer: 25 mM sodium cacodylate, pH 7.0, 50 mM KCl, 1 mM $MgCl_2$). Samples were heated (95 °C, 5 min), snap cooled, and run on 10% native acrylamide gel (120 V, 4 °C). Gels were exposed, imaged, and quantified using ImageJ to determine percentage bound A-strand. For displacement gels, MALAT URIL and radiolabeled oligo-A (WT) were mixed at 1:1 stoichiometry (3 μM) in 1X binding buffer (25 mM sodium cacodylate pH 7.0, 50 mM KCl, 1 mM $MgCl_2$), heated (95 °C, 5 min), and slow cooled. The complex was diluted into varying concentrations of bPNAs in the binding buffer (up to 4 μM final concentration) to yield a final RNA concentration of 100 nM, incubated (RT, 30 min), and analyzed on 10% native acrylamide gel (120 V, ice). EC_{50} values were determined by fitting[72] to an [agonist] vs response−variable slope model:

$$Y = \text{bottom} + (X^{\text{hillslope}}) \times (\text{top} - \text{bottom})/(X^{\text{hillslope}} + \text{EC}_{50}^{\text{hillslope}})$$

Exonuclease Assays. For exo T, radiolabeled (γ-^{32}P) MALAT1 triplex RNA (100 nM) was annealed as described and reacted with bPNA (2 μM) in 0.5 U/μL exo T (NEB BioLabs) in 1X NEB buffer 4 (50 mM KOAc, 20 mM Tris-acetate, 10 mM $Mg(OAc)_2$, 1 mM DTT, pH 7.9) at 25 °C. For time-course studies, 15 μL aliquots were taken at 0, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, and 4 h, overnight. Reaction was quenched by the addition of 2X TBE urea loading dye (Biorad) and heating (95 °C, 3 min). Samples were analyzed by gel and quantified as described above. Exo T guided binding was followed by fluorescence studies with Cy5-bPNA (1.5 μM), ENE+A triplex RNA (300 nM), and exo T (0.5 U/μL) in 1X NEB buffer 4 at 25 °C, with Cy5 fluorescence measured every 15 min (Bio-Rad CFX96). For RNase R, radiolabeled (γ-^{32}P) triplex RNA was annealed by dilution to 3 μM in 1X RNase R buffer (20 mM HEPES-KOH, pH 7.4 with 1 mM Mg^{2+}), heating (95 °C, 3 min), and slow cooling to RT (1 h). Digestion reactions were carried out (100 nM RNA, 400 nM bPNA) with RNase R (0.1 U/μL, Lucigen) in 1X RNase R buffer (37 °C), and 15 μL aliquots taken at the following time points: 0, 0.25, 0.5, 1, 2, 3, 4, 6, and 8 h, overnight. Reactions were quenched by the addition of 2X TBE-urea loading dye (Biorad) and heating (95 °C, 3 min). Reactions with $K^{2M}AK^{2M}$ bPNA were also carried out at half concentration: 200 nM bPNA and 0.05 U/μL RNase R, with time

https://doi.org/10.1021/acschembio.1c00575
ACS Chem. Biol. 2021, 16, 1600−1609

points taken at 0, 0.25, 0.5, 1, 2, 3, 4, 5, 6, and 8 h. Samples were run 10% denaturing PAGE, imaged, and quantified to determine [MALAT1] over time. Half-life was determined by fitting to a single-phase exponential decay: $Y = (Y_0 - \text{plateau}) \times \exp(-K \times X) + \text{plateau}$. For concentration-dependence studies, bPNAs were serially diluted and reactions were carried out with 0.1 U/μL RNase R. All reactions were quenched and analyzed at 3 h as described.

Native MALAT1 Pulldown from Total RNA with Biotin-bPNA. Streptavidin-agarose resin (EMD Millipore) was spun down (500*g*) and buffer exchanged to 1X DPBS. To 1 mL of resin slurry (500 μL settled, 42.5 nmol of free biotin binding capacity), 212.5 nmol of biotin-PEG$_3$-K^{2M}AK^{2M} was added. The mixture was spin-mixed (4 h, RT), the unbound btn-bPNA in supernatant was removed, and the resin was washed (1X wash with 5X DPBS, 0.5% Triton X-100; 3X wash with 10 mM sodium phosphate, pH 7) and resuspended in 500 μL of 1X DPBS. Total RNA was extracted from PANC-1 cells with TRIzol (Invitrogen protocol) and diluted to a final concentration of 100 ng/μL with 1X DPBS to final volume of 1 mL. The mixture was spin-mixed for 1 h at 4 °C. Unbound total RNA in the supernatant was removed, and the resin was washed (2X wash, ice cold 1X DPBS, 3X wash, water) and resuspended in 1 mL of water. Resin samples (20 μL obtained from 40 μL of resin slurry upon removal of 20 μL of water), 20 μL of supernatant, and total RNA (no treatment) were prepared. cDNA synthesis was carried out with M-MuLV Reverse Transcriptase (NEB protocol) with random hexamer priming (50 μL, Thermo Fisher Scientific). The freshly synthesized cDNAs were heat-treated (65 °C, 20 min) and directly subjected to qPCR analysis (Bio-Rad CFX96 touch Real-Time PCR) with Power SYBR Green PCR Master Mix (Applied Biosystems). Relative abundance of MALAT1/GAPDH in resin-released RNA, and total RNA was compared.

MALAT1 Knockdown in PANC-1 Cells with bPNA. PANC-1 cells were purchased from the American Type Culture Collection (ATCC) with cell line authentication provided and cultured in Dulbecco's modified eagle medium (DMEM) supplemented with 10% FBS, passaged at 60% confluency with trypsin/EDTA, and seeded on a 24-well culture plate to 1×10^5 cells/mL the day prior to treatment. Treatment protocol: media was removed and cells were washed (1X DPBS, fresh media with 1% FBS) and bPNA was added (final concentration 1 μM). Cells were harvested after 1 and 2 days of bPNA incubation with bPNA by washing (1X DPBS twice), treatment with TRIzol (ThermoFisher), RNA extraction by chloroform, and sequential precipitation according to manufacturer protocols. Reverse transcription to obtain cDNA was performed using M-MuLV Reverse Transcriptase (NEB), and 500 nM gene specific primers for both target gene (MALAT1) and housekeeping gene (GAPDH) and qRT-PCR were carried out as described above.

Cell Viability. PANC-1 cells were seeded onto a 96-well tissue culture plate with 10^5 cell/mL and 100 μL of total volume. Cells were treated with a series diluted bPNA and incubated for 2 days in a humidified tissue culture incubator (5% CO_2). Cell viability reagent alamarBlue (Thermo Fisher Scientific) was added to each well and incubated at 37 °C for 3 h. Fluorescence was measured (560/590 nm) and plotted.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.1c00575.

Discussions of materials, instrumentation, and sequences used, general notes, experimental and synthetic procedures, and solid phase peptide synthesis, figures of secondary structures, UV spectra, representative EMSA and fittings, fitting of A-tail binding to the MALAT1 duplex RNA, fitting of A-tail displacement, fitting of bPNA-assisted MALAT1 degradation, condition screening for RNase R assay, cell viability, synthesis pathways, NMR spectra, ESI spectra, HPLC traces, and MALDI spectra, and tables of transcripts for selectivity study and qPCR primers, T_m of major transitions, binding affinity and stoichiometry, A-tail binding affinity, EC_{50} of displacement, and $t_{1/2}$ of MALAT1 with exo T and RNase R (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Dennis Bong − *Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States;* orcid.org/0000-0003-3778-9183; Phone: (614) 247-8404; Email: bong.6@osu.edu

Authors

Shiqin Miao − *Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States*

Debmalya Bhunia − *Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States*

Shekaraiah Devari − *Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States*

Yufeng Liang − *Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States*

Oliver Munyaradzi − *Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States*

Sarah Rundell − *Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acschembio.1c00575

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by NIH (GM111995-01A1), NSF (DMR 1802432) and NASA (GRT00044810) grants (to D.B.) and the Center for RNA Biology at OSU.

■ REFERENCES

(1) Torabi, S.-F.; Vaidya, A. T.; Tycowski, K. T.; DeGregorio, S. J.; Wang, J.; Shu, M.-D.; Steitz, T. A.; Steitz, J. A. RNA Stabilization by a poly(A) Tail 3′-End Binding Pocket and Other Modes of poly(A)-RNA Interaction. *Science* **2021**, *371* (6529), eabe6523.

(2) Akiyama, B. M.; Eiler, D.; Kieft, J. S. Structured RNAs That Evade or Confound Exonucleases: Function Follows Form. *Curr. Opin. Struct. Biol.* **2016**, *36*, 40−47.

(3) Mitton-Fry, R. M.; DeGregorio, S. J.; Wang, J.; Steitz, T. A.; Steitz, J. A. Poly(A) Tail Recognition by a Viral RNA Element through Assembly of a Triple Helix. *Science* **2010**, *330* (6008), 1244−1247.

(4) Tycowski, K. T.; Shu, M.-D.; Borah, S.; Shi, M.; Steitz, J. A. Conservation of a Triple-Helix-Forming RNA Stability Element in Noncoding and Genomic RNAs of Diverse Viruses. *Cell Rep.* **2012**, *2* (1), 26−32.

(5) Conrad, N. K.; Steitz, J. A. A Kaposi's Sarcoma Virus RNA Element That Increases the Nuclear Abundance of Intronless Transcripts. *EMBO J.* **2005**, *24* (10), 1831−1841.

(6) Brown, J. A.; Valenstein, M. L.; Yario, T. A.; Tycowski, K. T.; Steitz, J. A. Formation of Triple-Helical Structures by the 3′-End

Sequences of MALAT1 and MENβ Noncoding RNAs. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (47), 19202−19207.
(7) Wilusz, J. E.; JnBaptiste, C. K.; Lu, L. Y.; Kuhn, C.-D.; Joshua-Tor, L.; Sharp, P. A. A Triple Helix Stabilizes the 3′ Ends of Long Noncoding RNAs That Lack Poly (A) Tails. *Genes Dev.* **2012**, *26* (21), 2392.
(8) Wang, K. C.; Chang, H. Y. Molecular Mechanisms of Long Noncoding RNAs. *Mol. Cell* **2011**, *43* (6), 904−914.
(9) Tripathi, V.; Ellis, J. D.; Shen, Z.; Song, D. Y.; Pan, Q.; Watt, A. T.; Freier, S. M.; Bennett, C. F.; Sharma, A.; Bubulya, P. A.; Blencowe, B. J.; Prasanth, S. G.; Prasanth, K. V. The Nuclear-Retained Noncoding RNA MALAT1 Regulates Alternative Splicing by Modulating SR Splicing Factor Phosphorylation. *Mol. Cell* **2010**, *39* (6), 925−938.
(10) Sun, Y.; Ma, L. New Insights into Long Non-Coding RNA MALAT1 in Cancer and Metastasis. *Cancers* **2019**, *11* (2), 216.
(11) Kim, J.; Piao, H.-L.; Kim, B.-J.; Yao, F.; Han, Z.; Wang, Y.; Xiao, Z.; Siverly, A. N.; Lawhon, S. E.; Ton, B. N.; Lee, H.; Zhou, Z.; Gan, B.; Nakagawa, S.; Ellis, M. J.; Liang, H.; Hung, M.-C.; You, M. J.; Sun, Y.; Ma, L. Long Noncoding RNA MALAT1 Suppresses Breast Cancer Metastasis. *Nat. Genet.* **2018**, *50* (12), 1705−1715.
(12) Patel, N. A.; Moss, L. D.; Lee, J.-Y.; Tajiri, N.; Acosta, S.; Hudson, C.; Parag, S.; Cooper, D. R.; Borlongan, C. V.; Bickford, P. C. Long Noncoding RNA MALAT1 in Exosomes Drives Regenerative Function and Modulates Inflammation-Linked Networks Following Traumatic Brain Injury. *J. Neuroinflammation* **2018**, *15* (1), 204.
(13) Li, Z.-X.; Zhu, Q.-N.; Zhang, H.-B.; Hu, Y.; Wang, G.; Zhu, Y.-S. MALAT1: A Potential Biomarker in Cancer. *Cancer Manage. Res.* **2018**, *10*, 6757−6768.
(14) Brown, J. A.; Kinzig, C. G.; DeGregorio, S. J.; Steitz, J. A. Hoogsteen-Position Pyrimidines Promote the Stability and Function of the MALAT1 RNA Triple Helix. *RNA* **2016**, *22* (5), 743−749.
(15) Ruszkowska, A.; Ruszkowski, M.; Hulewicz, J. P.; Dauter, Z.; Brown, J. A. Molecular Structure of a U•A-U-Rich RNA Triple Helix with 11 Consecutive Base Triples. *Nucleic Acids Res.* **2020**, *48* (6), 3304−3314.
(16) Donlic, A.; Morgan, B. S.; Xu, J. L.; Liu, A.; Roble, C.; Hargrove, A. E. Discovery of Small Molecule Ligands for MALAT1 by Tuning an RNA-Binding Scaffold. *Angew. Chem., Int. Ed.* **2018**, *57* (40), 13242−13247.
(17) Abulwerdi, F. A.; Xu, W.; Ageeli, A. A.; Yonkunas, M. J.; Arun, G.; Nam, H.; Schneekloth, J. S., Jr; Dayie, T. K.; Spector, D.; Baird, N.; et al. Selective Small Molecule Targeting of a Triple Helix Encoded by the Long Non-Coding RNA, MALAT1. *ACS Chem. Biol.* **2019**, *14* (2), 223.
(18) Ursu, A.; Vézina-Dawod, S.; Disney, M. D. Methods to Identify and Optimize Small Molecules Interacting with RNA (SMIRNAs). *Drug Discovery Today* **2019**, *24* (10), 2002−2016.
(19) Disney, M. D.; Winkelsas, A. M.; Velagapudi, S. P.; Southern, M.; Fallahi, M.; Childs-Disney, J. L. Inforna 2.0: A Platform for the Sequence-Based Design of Small Molecules Targeting Structured RNAs. *ACS Chem. Biol.* **2016**, *11* (6), 1720−1728.
(20) Donlic, A.; Zafferani, M.; Padroni, G.; Puri, M.; Hargrove, A. E. Regulation of MALAT1 Triple Helix Stability and in Vitro Degradation by Diphenylfurans. *Nucleic Acids Res.* **2020**, *48* (14), 7653−7664.
(21) Chen, H.; Meena; McLaughlin, L. W. A Janus-Wedge DNA Triplex with A-W1-T and G-W2-C Base Triplets. *J. Am. Chem. Soc.* **2008**, *130*, 13190−13191.
(22) Branda, N.; Kurz, G.; Lehn, J.-M. JANUS WEDGES: A New Approach towards Nucleobase-Pair Recognition. *Chem. Commun.* **1996**, No. 21, 2443.
(23) Marchi-Artzner, V.; Jullien, L.; Gulik-Krzywicki, T.; Lehn, J.-M. Molecular Recognition Induced Aggregation and Fusion between Vesicles Containing Lipids Bearing Complementary Hydrogen Bonding Head-Groups. *Chem. Commun.* **1997**, 117−118.
(24) Zeng, Y.; Pratumyot, Y.; Piao, X.; Bong, D. Discrete Assembly of Synthetic Peptide-DNA Triplex Structures from Polyvalent Melamine-Thymine Bifacial Recognition. *J. Am. Chem. Soc.* **2012**, *134* (2), 832−835.
(25) Xia, X.; Piao, X.; Bong, D. Bifacial Peptide Nucleic Acid as an Allosteric Switch for Aptamer and Ribozyme Function. *J. Am. Chem. Soc.* **2014**, *136* (20), 7265−7268.
(26) Piao, X.; Xia, X.; Bong, D. Bifacial Peptide Nucleic Acid Directs Cooperative Folding and Assembly of Binary, Ternary, and Quaternary DNA Complexes. *Biochemistry* **2013**, *52*, 6313−6323.
(27) Mao, J.; DeSantis, C.; Bong, D. Small Molecule Recognition Triggers Secondary and Tertiary Interactions in DNA Folding and Hammerhead Ribozyme Catalysis. *J. Am. Chem. Soc.* **2017**, *139* (29), 9815−9818.
(28) Whitesides, G. M.; Simanek, E. E.; Mathias, J. P.; Seto, C. T.; Chin, D.; Mammen, M.; Gordon, D. M. Noncovalent Synthesis: Using Physical-Organic Chemistry To Make Aggregates. *Acc. Chem. Res.* **1995**, *28*, 37−44.
(29) Oshovsky, G. V.; Reinhoudt, D. N.; Verboom, W. Supramolecular Chemistry in Water. *Angew. Chem., Int. Ed.* **2007**, *46*, 2366−2393.
(30) Ma, M.; Bong, D. Determinants of Cyanuric Acid and Melamine Assembly in Water. *Langmuir* **2011**, *27* (14), 8841−8853.
(31) Cafferty, B. J.; Gállego, I.; Chen, M. C.; Farley, K. I.; Eritja, R.; Hud, N. V. Efficient Self-Assembly in Water of Long Noncovalent Polymers by Nucleobase Analogues. *J. Am. Chem. Soc.* **2013**, *135* (7), 2447−2450.
(32) Avakyan, N.; Greschner, A. A.; Aldaye, F.; Serpell, C. J.; Toader, V.; Petitjean, A.; Sleiman, H. F. Reprogramming the Assembly of Unmodified DNA with a Small Molecule. *Nat. Chem.* **2016**, *8* (4), 368−376.
(33) Lange, R. F. M.; Beijer, F. H.; Sijbesma, R. P.; Hooft, R. W. W.; Kooijman, H.; Spek, A. L.; Kroon, J.; Meijer, E. W. Crystal Engineering of Melamine−Imide Complexes; Tuning the Stoichiometry by Steric Hindrance of the Imide Carbonyl Groups. *Angew. Chem., Int. Ed. Engl.* **1997**, *36* (9), 969−971.
(34) Cafferty, B. J.; Fialho, D. M.; Khanam, J.; Krishnamurthy, R.; Hud, N. V. Spontaneous Formation and Base Pairing of Plausible Prebiotic Nucleotides in Water. *Nat. Commun.* **2016**, *7* (1), 11328.
(35) Zhou, Z.; Bong, D. Small-Molecule/Polymer Recognition Triggers Aqueous-Phase Assembly and Encapsulation. *Langmuir* **2013**, *29*, 144−150.
(36) Kool, E. T. Hydrogen Bonding, Base Stacking, and Steric Effects in DNA Replication. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30* (1), 1−22.
(37) Benner, S. A. Understanding Nucleic Acids Using Synthetic Chemistry. *Acc. Chem. Res.* **2004**, *37* (10), 784−797.
(38) Hirao, I.; Kimoto, M.; Yamashige, R. Natural versus Artificial Creation of Base Pairs in DNA: Origin of Nucleobases from the Perspectives of Unnatural Base Pair Studies. *Acc. Chem. Res.* **2012**, *45* (12), 2055−2065.
(39) Mittapalli, G. K.; Reddy, K. R.; Xiong, H.; Munoz, O.; Han, B.; De Riccardis, F.; Krishnamurthy, R.; Eschenmoser, A. Mapping the Landscape of Potentially Primordial Informational Oligomers: Oligodipeptides and Oligodipeptoids Tagged with Triazines as Recognition Elements. *Angew. Chem., Int. Ed.* **2007**, *46*, 2470−2477.
(40) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. A Simple Ligand That Selectively Targets CUG Trinucleotide Repeats and Inhibits MBNL Protein Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 16068−16073.
(41) Zhou, Z.; Xia, X.; Bong, D. Synthetic Polymer Hybridization with DNA and RNA Directs Nanoparticle Loading, Silencing Delivery, and Aptamer Function. *J. Am. Chem. Soc.* **2015**, *137* (28), 8920−8923.
(42) Xia, X.; Piao, X.; Fredrick, K.; Bong, D. Bifacial PNA Complexation Inhibits Enzymatic Access to DNA and RNA. *ChemBioChem* **2014**, *15*, 31−36.
(43) Piao, X.; Xia, X.; Mao, J.; Bong, D. Peptide Ligation and RNA Cleavage via an Abiotic Template Interface. *J. Am. Chem. Soc.* **2015**, *137* (11), 3751−3754.

https://doi.org/10.1021/acschembio.1c00575
ACS Chem. Biol. 2021, 16, 1600−1609

(44) Mao, J.; Bong, D. Synthesis of DNA-Binding Peptoids. *Synlett* **2015**, *26* (11), 1581−1585.
(45) Malnuit, V.; Duca, M.; Benhida, R. Targeting DNA Base Pair Mismatch with Artificial Nucleobases. Advances and Perspectives in Triple Helix Strategy. *Org. Biomol. Chem.* **2011**, *9*, 326−336.
(46) Soto, A. M.; Loo, J.; Marky, L. A. Energetic Contributions for the Formation of TAT/TAT, TAT/CGC(+), and CGC(+)/CGC(+) Base Triplet Stacks. *J. Am. Chem. Soc.* **2002**, *124*, 14355−14363.
(47) Noeske, J.; Richter, C.; Grundl, M. A.; Nasiri, H. R.; Schwalbe, H.; Wöhnert, J. An Intermolecular Base Triple as the Basis of Ligand Specificity and Affinity in the Guanine- and Adenine-Sensing Riboswitch RNAs. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (5), 1372−1377.
(48) Miao, S.; Liang, Y.; Marathe, I.; Mao, J.; DeSantis, C.; Bong, D. Duplex Stem Replacement with bPNA+ Triplex Hybrid Stems Enables Reporting on Tertiary Interactions of Internal RNA Domains. *J. Am. Chem. Soc.* **2019**, *141* (23), 9365−9372.
(49) Borthwick, A. D. 2,5-Diketopiperazines: Synthesis, Reactions, Medicinal Chemistry, and Bioactive Natural Products. *Chem. Rev.* **2012**, *112* (7), 3641−3716.
(50) Yonkunas, M. J.; Baird, N. J. A Highly Ordered, Nonprotective MALAT1 ENE Structure Is Adopted prior to Triplex Formation. *RNA* **2019**, *25* (8), 975−984.
(51) Ageeli, A. A.; McGovern-Gooch, K. R.; Kaminska, M. M.; Baird, N. J. Finely Tuned Conformational Dynamics Regulate the Protective Function of the lncRNA MALAT1 Triple Helix. *Nucleic Acids Res.* **2019**, *47* (3), 1468−1481.
(52) Jiao, F.; Hu, H.; Yuan, C.; Wang, L.; Jiang, W.; Jin, Z.; Guo, Z.; Wang, L. Elevated Expression Level of Long Noncoding RNA MALAT-1 Facilitates Cell Growth, Migration and Invasion in Pancreatic Cancer. *Oncol. Rep.* **2014**, *32* (6), 2485−2492.
(53) Liang, Y.; Miao, S.; Mao, J.; DeSantis, C.; Bong, D. Context-Sensitive Cleavage of Folded DNAs by Loop-Targeting bPNAs. *Biochemistry* **2020**, *59* (26), 2410−2418.
(54) Tost, J.; Gut, I. G. DNA Methylation Analysis by Pyrosequencing. *Nat. Protoc.* **2007**, *2* (9), 2265−2275.
(55) Kunkler, C. N.; Hulewicz, J. P.; Hickman, S. C.; Wang, M. C.; McCown, P. J.; Brown, J. A. Stability of an RNA• DNA–DNA Triple Helix Depends on Base Triplet Composition and Length of the RNA Third Strand. *Nucleic Acids Res.* **2019**, *47* (14), 7213−7222.
(56) Bakheet, T.; Hitti, E.; Khabar, K. S. A. ARED-Plus: An Updated and Expanded Database of AU-Rich Element-Containing mRNAs and Pre-mRNAs. *Nucleic Acids Res.* **2018**, *46*, D218−D220.
(57) Zhou, Z.; Giles, K. E.; Felsenfeld, G. DNA·RNA Triple Helix Formation Can Function as a Cis-Acting Regulatory Mechanism at the Human β-Globin Locus. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (13), 6130−6139.
(58) Yang, W.-Y.; Gao, R.; Southern, M.; Sarkar, P. S.; Disney, M. D. Design of a Bioactive Small Molecule That Targets r(AUUCU) Repeats in Spinocerebellar Ataxia 10. *Nat. Commun.* **2016**, *7*, 11647.
(59) Brown, J. A.; Bulkley, D.; Wang, J.; Valenstein, M. L.; Yario, T. A.; Steitz, T. A.; Steitz, J. A. Structural Insights into the Stabilization of MALAT1 Noncoding RNA by a Bipartite Triple Helix. *Nat. Struct. Mol. Biol.* **2014**, *21* (7), 633−640.
(60) Bakheet, T.; Williams, B. R. G.; Khabar, K. S. A. ARED 3.0: The Large and Diverse AU-Rich Transcriptome. *Nucleic Acids Res.* **2006**, *34*, D111−D114.
(61) Clemson, C. M.; Hutchinson, J. N.; Sara, S. A.; Ensminger, A. W.; Fox, A. H.; Chess, A.; Lawrence, J. B. An Architectural Role for a Nuclear Noncoding RNA: NEAT1 RNA Is Essential for the Structure of Paraspeckles. *Mol. Cell* **2009**, *33* (6), 717−726.
(62) Wilusz, J. E. Long Noncoding RNAs: Re-Writing Dogmas of RNA Processing and Stability. *Biochim. Biophys. Acta, Gene Regul. Mech.* **2016**, *1859* (1), 128−138.
(63) Quinn, J. J.; Chang, H. Y. Unique Features of Long Non-Coding RNA Biogenesis and Function. *Nat. Rev. Genet.* **2016**, *17* (1), 47−62.
(64) Zhao, W.; He, X.; Hoadley, K. A.; Parker, J. S.; Hayes, D. N.; Perou, C. M. Comparison of RNA-Seq by Poly (A) Capture, Ribosomal RNA Depletion, and DNA Microarray for Expression Profiling. *BMC Genomics* **2014**, *15*, 419.
(65) Galganski, L.; Urbanek, M. O.; Krzyzosiak, W. J. Nuclear Speckles: Molecular Organization, Biological Function and Role in Disease. *Nucleic Acids Res.* **2017**, *45* (18), 10350−10368.
(66) Nakagawa, S.; Ip, J. Y.; Shioi, G.; Tripathi, V.; Zong, X.; Hirose, T.; Prasanth, K. V. Malat1 Is Not an Essential Component of Nuclear Speckles in Mice. *RNA* **2012**, *18* (8), 1487−1499.
(67) Gezer, U.; Özgür, E.; Cetinkaya, M.; Isin, M.; Dalay, N. Long Non-Coding RNAs with Low Expression Levels in Cells Are Enriched in Secreted Exosomes. *Cell Biol. Int.* **2014**, *38* (9), 1076−1079.
(68) Ripin, N.; Boudet, J.; Duszczyk, M. M.; Hinniger, A.; Faller, M.; Krepl, M.; Gadi, A.; Schneider, R. J.; Šponer, J.; Meisner-Kober, N. C.; Allain, F. H.-T. Molecular Basis for AU-Rich Element Recognition and Dimerization by the HuR C-Terminal RRM. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (8), 2935−2944.
(69) Hitti, E.; Bakheet, T.; Al-Souhibani, N.; Moghrabi, W.; Al-Yahya, S.; Al-Ghamdi, M.; Al-Saif, M.; Shoukri, M. M.; Lánczky, A.; Grépin, R.; Győrffy, B.; Pagès, G.; Khabar, K. S. A. Systematic Analysis of AU-Rich Element Expression in Cancer Reveals Common Functional Clusters Regulated by Key RNA-Binding Proteins. *Cancer Res.* **2016**, *76* (14), 4068−4080.
(70) Dąbrowski, S.; Kur, J. Cloning and Expression inEscherichia Coliof the Recombinant His-Tagged DNA Polymerases fromPyrococcus furiosusandPyrococcus Woesei. *Protein Expression Purif.* **1998**, *14* (1), 131−138.
(71) Lyon, S.; Gopalan, V. A T7 RNA Polymerase Mutant Enhances the Yield of 5′-Thienoguanosine-Initiated RNAs. *ChemBioChem* **2018**, *19* (2), 142−146.
(72) Sebaugh, J. L. Guidelines for Accurate EC50/IC50 Estimation. *Pharm. Stat.* **2011**, *10* (2), 128−134.

https://doi.org/10.1021/acschembio.1c00575
ACS Chem. Biol. 2021, 16, 1600−1609