Nucleic Acid Analogues Hot Paper

DOI: 10.1002/anie.201307661

A Chemical Perspective on Transcriptional Fidelity: Dominant Contributions of Sugar Integrity Revealed by Unlocked Nucleic Acids**

*Liang Xu, Steven W. Plouffe, Jenny Chong, Jesper Wengel, and Dong Wang**

Nucleic acids (RNA and DNA) consist of three key structural moieties: the nucleobase, the sugar, and the phosphate group. A comprehensive understanding of how these intrinsic structural features are recognized by nucleic acid enzymes during gene expression is not only important for elucidating the chemical basis of the central dogma of biology (genetic information transferred from DNA to RNA to protein), but also has many implications for synthetic biology and nucleic acid based therapeutics.[1–7] Synthetic nucleic acid chemistry has now provided us with powerful tools to advance our understanding of the functional interplay between nucleic acids and nucleic acid enzymes, an advance that cannot be achieved by conventional biological approaches.[1,8]

RNA Polymerase II (Pol II) is responsible for transcribing a DNA template into precursor messenger RNA in all eukaryotic cells. Key questions concern how Pol II recognizes the structural moieties of nucleic acids and how it maintains high transcriptional fidelity. Previous studies have mainly focused on understanding the contributions of the nucleotides' peripheral functional groups (the nucleobase and 2′-OH group) to Pol II transcriptional fidelity.[9–13] However, the contribution of the central structural moiety (the sugar backbone) is unclear. It is worth noting that nature has selected cyclic ribose as the sugar backbone for both RNA and DNA, and modern nucleic acid enzymes, such as Pol II, have adapted to this sugar backbone during the course of evolution. This raises several intriguing questions: how important is an intact sugar backbone to Pol II substrate recognition and transcriptional fidelity? Do modern nucleic acid enzymes, such as Pol II, have an inbuilt capacity to recognize and select for nucleotides and nucleic acids with intact sugar moieties.

To address these questions, we employed a synthetic chemical biology approach that involved comparing the Pol II mediated incorporation of a canonical nucleotide to that of a nucleotide analogue with a disrupted sugar moiety (Figure 1). This synthetic nucleotide analogue, termed

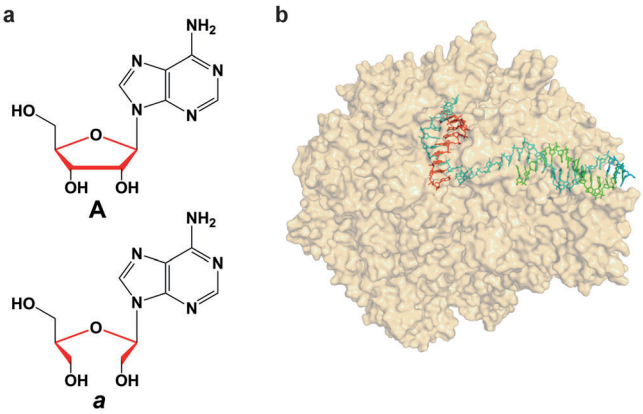

Figure 1. a) Canonical and unlocked nucleosides used in this study. Structures of adenosine (A) and unlocked adenosine (denoted as ***a***). The sugar backbones are highlighted in red. b) Structure of the Pol II elongation complex (PDB: 2E2H).[19] Shown are Pol II (tan), RNA (red), template DNA (cyan), and non-template DNA (green).

unlocked nucleic acid (UNA),[14–17] contains all the peripheral functional groups of the natural nucleotide but has a disrupted ribose ring (Figure 1a).[14] Importantly, the UNA residues remain able to form Watson–Crick base-pairing with RNA or DNA strands without significant disruption to the duplex structure.[18] We systematically and quantitatively dissected the contribution of the sugar backbone to Pol II substrate incorporation, elongation, and the three key checkpoint steps of transcriptional fidelity. This knowledge is not only important for elucidating the molecular basis of sugar recognition by Pol II, but also provides insight into the molecular basis of genetic information storage and transfer during evolution.

To understand the effect of the disrupted sugar backbone on Pol II substrate incorporation, we focused on dissecting the contributions of two components: the incoming substrate and the RNA primer. Two critical enzyme kinetic parameters

[*] Dr. L. Xu,[+] S. W. Plouffe,[+] J. Chong, Prof. Dr. D. Wang
Skaggs School of Pharmacy and Pharmaceutical Sciences
The University of California, San Diego
La Jolla, CA 92093-0625 (USA)
E-mail: dongwang@ucsd.edu
Prof. Dr. J. Wengel
Nucleic Acid Center and Biomolecular Nanoscale Engineering Center, Department of Physics, Chemistry and Pharmacy
University of Southern Denmark
Campusvej 55, 5230 Odense M (Denmark)
[+] These authors contributed equally to this work.

[**] D.W. acknowledges the NIH (GM102362, Kimmel Scholar Award from the Sidney Kimmel Foundation for Cancer Research, and start-up funds from Skaggs School of Pharmacy and Pharmaceutical Sciences, UCSD. S.W.P. was supported in part by the UCSD Graduate Training Program in Cellular and Molecular Pharmacology training grant (T32 GM007752). We thank Dr. E. Peter Geiduschek, Dr. Eric T. Kool, and Dr. Yitzhak Tor for their critical reading of our manuscript.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201307661.

were determined: the substrate specificity constant ($k_{pol}/K_{d,app}$), which is derived from the rate of polymerization (k_{pol}) and apparent substrate dissociation constant ($K_{d,app}$); and discrimination, calculated from the ratio of $k_{pol}/K_{d,app}$ for substrate A over substrate B ($(k_{pol}/K_{d,app})_A/(k_{pol}/K_{d,app})_B$). Discrimination is a quantitative measure of the enzyme's power to discriminate between two substrates. To study the effect of a disrupted sugar in the substrate, we investigated the incorporation of ATP, a process that occurs with the loss of two phosphate groups to give the corresponding nucleoside monophosphate (AMP) in the resulting nucleic acid. We compared the nucleotide incorporation efficiency of canonical ATP substrate and unlocked adenosine triphosphate analogue (denoted as ***a***TP) into a Pol II elongation complex with Scaffold I containing a 10nt RNA primer with an A at the 3′-RNA terminus (termed 10A, Figure S1a in the Supporting Information). Strikingly, the nucleotide incorporation of ***a***TP substrate is substantially slower than that of ATP (Figure 2a), and Pol II exhibits an approximately 10^7-fold discrimination for ATP over ***a***TP ($(k_{pol}/K_{d,app})_{ATP}/(k_{pol}/K_{d,app})_{aTP}$) as revealed by pre-steady-state single-turnover incorporation assays (Figure 2c, Table S1), even though both substrates share the same nucleobase, sugar hydroxy groups, and triphosphate functional groups. This strong discrimination largely results from the difference in catalytic rates (ca. 10^6-fold) rather than a difference in substrate affinity (ca. 10-fold; Figure S2). It was unexpected that the contribution of the sugar backbone (ca. 10^7-fold in $k_{pol}/K_{d,app}$) would be approximately 100-fold stronger than that of nucleobase pairing (matched A:dT pair versus mismatched U:dT measured by $(k_{pol}/K_{d,app})_{ATP}/(k_{pol}/K_{d,app})_{UTP}$) and 10^3–10^4 stronger than that of sugar 2′-OH interactions (NTP versus dNTP).[9,11,19] This result reveals that the integrity of the substrate's sugar backbone plays a dominant role in ensuring efficient Pol II substrate incorporation.

To investigate the role of sugar integrity in the RNA primer, we compared nucleotide incorporation into Pol II elongation complexes with either a scaffold containing a canonical RNA primer (10A) or a scaffold containing an RNA primer with an unlocked adenosine (***a***) substitution at the 3′-RNA terminus (10$\boldsymbol{a}_n$, n indicating the 3′-RNA terminus position; Figure 2a). We found that incorporation of ATP substrate into 10$\boldsymbol{a}_n$ is substantially slower than ATP incorporation into Scaffold 10A, with an approximately (2×10^5)-fold reduction in the $k_{pol}/K_{d,app}$ value (Figure 2c and Table S1 in the Supporting Information). This result reveals an important role for sugar integrity at the 3′-RNA terminus in efficient nucleotide incorporation. No incorporation of ***a***TP substrate into the 10$\boldsymbol{a}_n$ primer was observed even after extensive incubation periods (Figure 2a), thus suggesting that disruption of the sugar backbone of both substrate and 3′-RNA terminal residue together completely abolishes Pol II transcription.

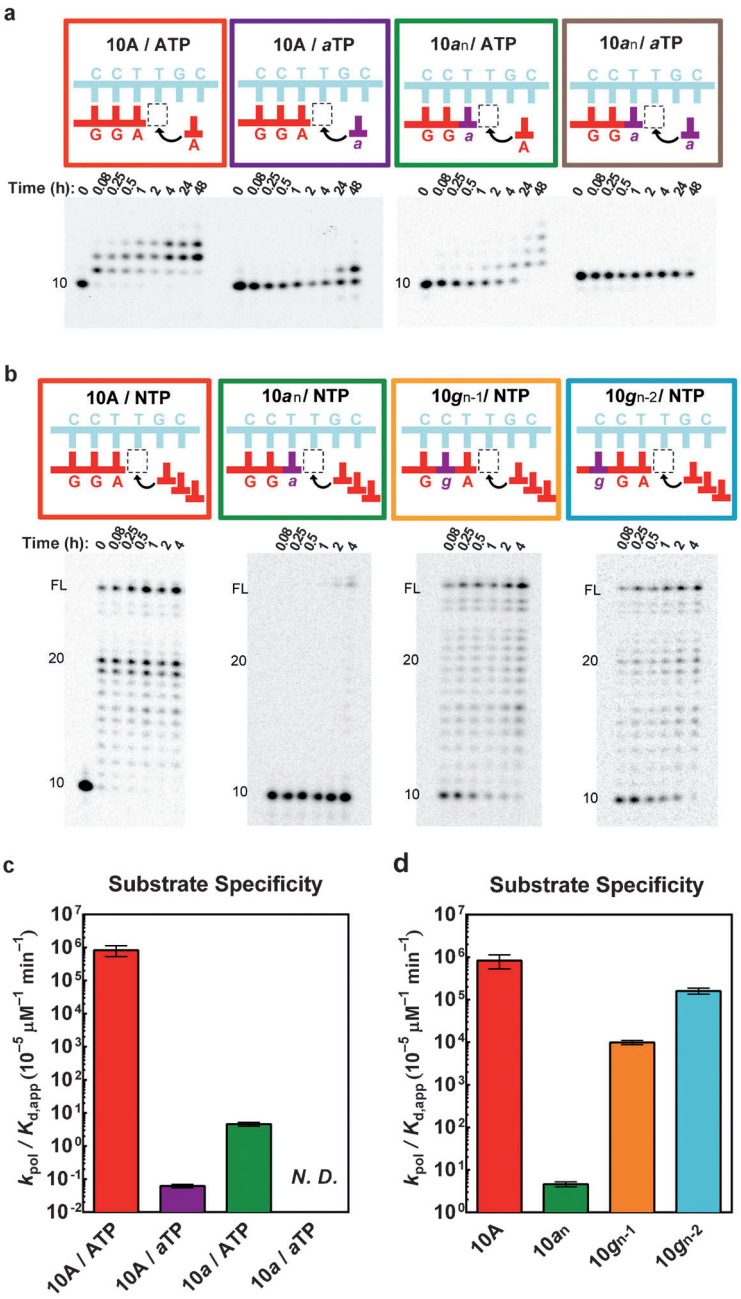

Figure 2. Sugar backbone integrity plays an important role in ensuring efficient Pol II transcription. a) ATP or ***a***TP incorporation into 10A or 10$\boldsymbol{a}_n$ scaffold by Pol II. b) Effects of an unlocked nucleotide substitution in the RNA primer on Pol II elongation. NTP, ATP, and ***a***TP were used at 1 mM concentration. c) Loss of sugar backbone integrity greatly reduces the substrate specificity constants. N.D.: not determined. d) Pol II elongation efficiency is recovered as the unlocked nucleoside is moved upstream.

To further test the effect of UNA incorporation into the RNA primer on Pol II elongation, we systematically designed primers containing UNA residues at either the n, $n-1$, or $n-2$ positions (termed 10$\boldsymbol{a}_n$, 10$\boldsymbol{g}_{n-1}$, or 10$\boldsymbol{g}_{n-2}$, respectively, Figure 2b). These scaffolds allowed us to compare the effect on

Pol II elongation of a disrupted sugar backbone at different positions along the RNA primer. We observed that UNA at the 3′-RNA terminus (10$\boldsymbol{a}_n$) greatly inhibits Pol II elongation, whereas the inhibitory effects were much lower with 10$\boldsymbol{g}_{n-1}$ and 10$\boldsymbol{g}_{n-2}$ (Figure 2b). The $k_{pol}/K_{d,app}$ value for the incorporation of ATP into 10$\boldsymbol{g}_{n-2}$ is only about 5-fold lower than that of incorporation into the 10A scaffold; in comparison, there was a (2×10^5)-fold reduction in specificity when using 10$\boldsymbol{a}_n$ (Figure 2d and Table S1). These results indicate that the effect of RNA primer sugar integrity on Pol II transcription is position dependent, and the inhibitory effect on Pol II transcription is greatly reduced as the UNA residue moves further upstream from the 3′-RNA terminus.

Pol II transcriptional fidelity is maintained by three checkpoint steps: 1) specific nucleotide selection and incorporation at the 3′-RNA terminus, 2) discrimination between matched and mismatched nucleotides at the 3′-RNA terminus during elongation, and 3) proofreading activity through preferential cleavage of misincorporated nucleotides.[8,11] We assembled five active RNA Pol II elongation complexes with either an unmodified RNA primer or one containing a site-specific UNA modification (Figure S1, Scaffolds I and II). This system allowed us to systematically and quantitatively measure the contributions of sugar integrity at each of the three Pol II transcriptional fidelity checkpoint steps.

To measure the contribution of sugar integrity to the first Pol II transcriptional fidelity checkpoint step (nucleotide selection and incorporation), we measured the kinetic parameters for matched ATP, mismatched UTP, sugar-disrupted ***a***TP, and mismatched and sugar-disrupted ***u***TP incorporation (Figure 3a, Table 1, and Figure S2).

Comparison of the $k_{pol}/K_{d,app}$ values among the four substrates reveals two very striking results (Figure 3a, Table 1). First, Pol II's discrimination for sugar integrity ($(k_{pol}/K_{d,app})_{ATP}/(k_{pol}/K_{d,app})_{\boldsymbol{a}TP}$, ca. 10^7-fold) is approximately 100-fold higher than its nucleobase discrimination ($(k_{pol}/K_{d,app})_{ATP}/(k_{pol}/K_{d,app})_{UTP}$, ca. 10^5-fold). Second, Pol II's nucleobase discrimination between ATP and mismatched UTP is completely abolished following the replacement of the canonical ribose with a UNA sugar (dropped from ca. 10^5-fold to ca. 1). This clearly indicates that disruption of the sugar backbone integrity not only greatly reduces Pol II transcriptional efficiency but also abolishes nucleobase recognition during incorporation, thus indicating that sugar backbone integrity is a prerequisite for correct nucleotide selection during Pol II transcription.

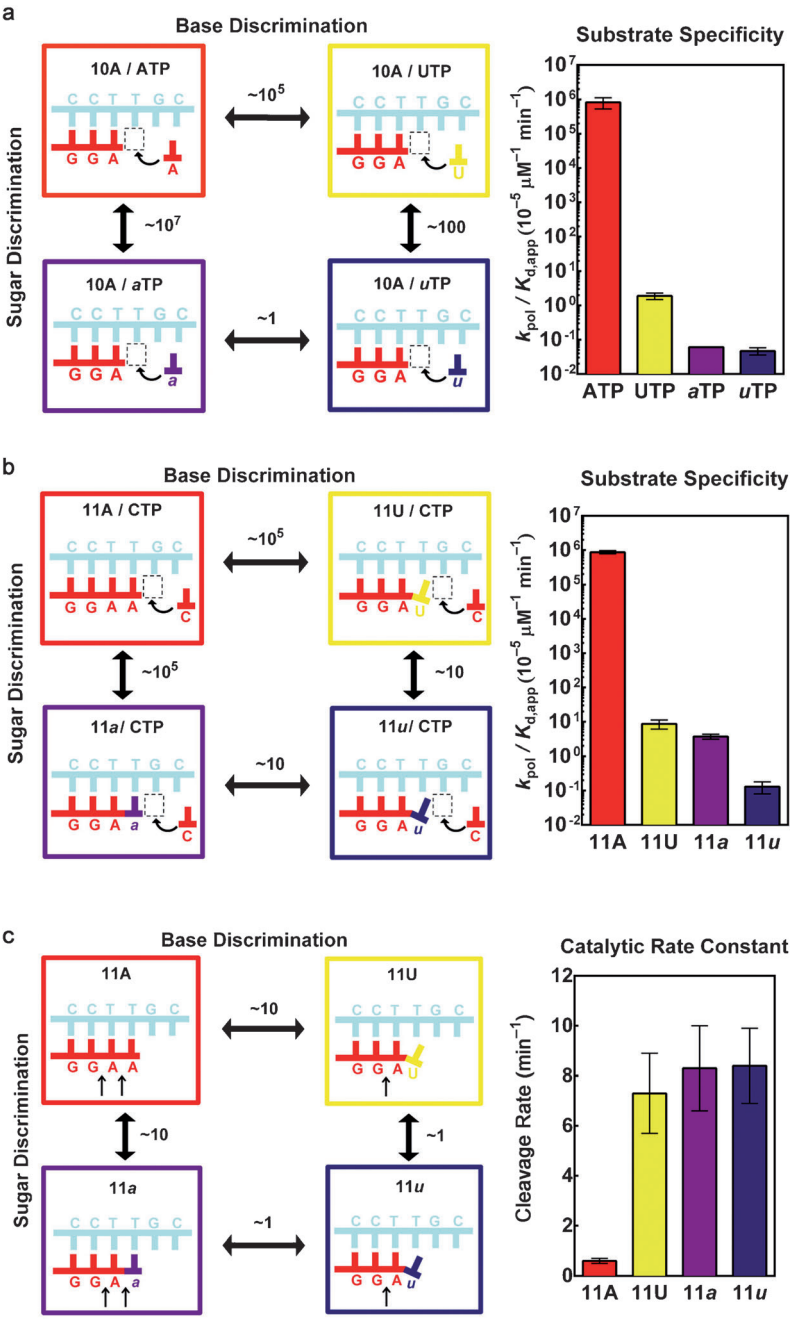

Figure 3. Sugar backbone integrity plays a dominant role in Pol II transcriptional fidelity at all three fidelity checkpoint steps. a) Loss of sugar backbone integrity completely abolishes nucleobase discrimination at the first checkpoint step. b) Loss of sugar backbone integrity significantly undermines primer nucleobase discrimination at the second checkpoint step. c) Loss of sugar integrity abolishes the difference in TFIIS-mediated cleavage between matched and mismatched 3′-RNA:DNA terminus at the third checkpoint step. Arrows indicate observed cleavage products.

To investigate the effect of a disrupted sugar backbone on the second checkpoint step (transcript extension), we assembled four Pol II elongation complexes with Scaffold II containing the following RNA:DNA pairings at the 3′-RNA terminus: A:dT, U:dT, ***a***:dT, and ***u***:dT (termed 11A, 11U, 11***a***, and 11***u*** respectively, Figure 3b and Figure S1). These four

Table 1: Effect of sugar backbone integrity on the three fidelity checkpoint steps of RNA Pol II transcription.

Fidelity checkpoint	Scaffold/Substrate	k_{pol} [min^{-1}]	$K_{d,app}$ [μM]	$k_{pol}/K_{d,app}$ [$\mu M^{-1} min^{-1}$]	Relative efficiency[a]	Nucleobase discrimination[b]
nucleotide incorporation	10A/ATP	750 ± 210	90 ± 20	8.3 ± 3.0	1	$(4.4 \pm 1.8) \times 10^5$
	10A/UTP	0.015 ± 0.003	800 ± 60	$(1.9 \pm 0.4) \times 10^{-5}$	$(2.3 \pm 1.0) \times 10^{-6}$	
	10A/***a***TP	$(5.8 \pm 0.2) \times 10^{-4}$	940 ± 100	$(6.2 \pm 0.7) \times 10^{-7}$	$(7.5 \pm 2.8) \times 10^{-8}$	1.3 ± 0.3
	10A/***u***TP	$(2.9 \pm 0.3) \times 10^{-4}$	620 ± 110	$(4.7 \pm 1.0) \times 10^{-7}$	$(5.7 \pm 2.4) \times 10^{-8}$	
nucleotide extension	11A/CTP	450 ± 20	52 ± 5	8.7 ± 0.9	1	$(1.0 \pm 0.3) \times 10^5$
	11U/CTP	0.26 ± 0.05	3000 ± 700	$(8.7 \pm 2.6) \times 10^{-5}$	$(1.0 \pm 0.3) \times 10^{-5}$	
	11***a***/CTP	$(3.4 \pm 0.1) \times 10^{-3}$	55 ± 5	$(6.2 \pm 0.6) \times 10^{-5}$	$(7.1 \pm 1.0) \times 10^{-6}$	6.2 ± 1.4
	11***u***/CTP	$(4.9 \pm 0.2) \times 10^{-4}$	47 ± 7	$(1.0 \pm 0.2) \times 10^{-5}$	$(1.1 \pm 0.3) \times 10^{-6}$	
	Scaffold	Cleavage rate (k_{obs} [min^{-1}])			Relative efficiency[c]	Nucleobase discrimination[d]
proofreading	11A	0.6 ± 0.1			1	12 ± 3
	11U	7.3 ± 1.6			12 ± 3	
	11***a***	8.3 ± 1.7			14 ± 4	1 ± 0.3
	11***u***	8.4 ± 1.5			14 ± 3	

[a] Relative Efficiency $= (k_{pol}/K_{d,app})_{ATP\ or\ UTP\ or\ aTP\ or\ uTP}/(k_{pol}/K_{d,app})_{ATP}$; $(k_{pol}/K_{d,app})_{11A\ or\ 11U\ or\ 11a\ or\ 11u}/(k_{pol}/K_{d,app})_{11A}$; [b] Discrimination $= (k_{pol}/K_{d,app})_{ATP\ or\ aTP}/(k_{pol}/K_{d,app})_{UTP\ or\ uTP}$; $(k_{pol}/K_{d,app})_{11A\ or\ 11U}/(k_{pol}/K_{d,app})_{11a\ or\ 11u}$; [c] Relative Efficiency $= (k_{obs})_{11A\ or\ 11U\ or\ 11a\ or\ 11u}/(k_{obs})_{11A}$; [d] Discrimination $= (k_{obs})_{11A\ or\ 11a}/(k_{obs})_{11U\ or\ 11u}$.

Scaffold II variants represent the four potential product states of Pol II elongation of Scaffold 10A after the first checkpoint step (Figure 3b). This approach allowed us to evaluate the influence of both sugar integrity and base pairing at the 3′-RNA terminus of the primer on incorporation of the next nucleotide. We measured the specificity constants for the incorporation of CTP substrate into all four Pol II elongation complexes. The efficiency of nucleotide incorporation following the correct A:dT base pair is consistently approximately 10^5-fold higher than that following the mismatched base pair U:dT (Figure 3b, Table 1, and Figure S3).[8] Strikingly, in the absence of sugar integrity, Pol II's ability to discriminate between a terminal matched pair over a terminal mismatched pair was significantly reduced from 10^5-fold (extension of A:dT over U:dT) to 6.2-fold (extension of ***a***:dT over ***u***:dT; Figure 3b, Table 1, and Figure S3). These results suggest that sugar backbone integrity is also a dominant factor in controlling the second checkpoint step of Pol II transcriptional fidelity.

Finally, to investigate how the presence of incorporated UNA in the RNA primer affects the third transcriptional fidelity checkpoint step (proofreading), we measured transcription factor IIS (TFIIS)-simulated cleavage rates for four Pol II elongation complexes containing the following RNA:DNA pairs at the 3′-RNA terminus: A:dT, U:dT, ***a***:dT, and ***u***:dT (Scaffold II 11A, 11U, 11***a***, and 11***u***, Figure 3c and Figure S1). The cleavage rate for 11A, which has a matched A:dT pair, is about 12-fold slower than that for 11U, which has a mismatched U:dT pair (Figure 3c and Table 1).[8] Intriguingly, we found that replacing the canonical sugar with an unlocked sugar also resulted in a large increase (ca. 14-fold) in cleavage rate for 11***a*** in comparison with that for 11A, even if the nucleobase was correctly matched to the dT template. Thus, Pol II is unable to discriminate between a matched end (11***a***:dT) and a mismatched end (11***u***:dT) in the absence of terminal sugar integrity. Notably, the sugar substitution did not change the corresponding cleavage pattern (Figure S4). Taken together, these results indicate that UNA sugar substitution at the 3′-RNA terminus abolishes Pol II proofreading activity.

In summary, we systematically dissected the contribution of the nucleic acid sugar backbone to Pol II transcription. Surprisingly, we discovered that during Pol II transcription, the contribution of the sugar backbone to enzymatic efficiency is 10^3–10^4 stronger than that of the 2′-OH group on the ribose and around 100-fold stronger than that of base pairing. Even more strikingly, we found that sugar integrity is a dominant factor in controlling all three checkpoint steps of Pol II transcriptional fidelity. These results provide novel insights into the molecular basis of sugar backbone recognition by Pol II, which has been underappreciated. Sugar backbone integrity is a prerequisite for ensuring correct nucleotide selection, incorporation, and extension during Pol II transcription as revealed by nucleobase discrimination being completely abolished following replacement with an unlocked sugar. The nucleotide's peripheral functional groups in themselves are insufficient for efficient Pol II transcription. Rather, correct spatial arrangement of the functional groups, which is guided by the sugar backbone, is key to maintaining high Pol II catalytic activity and fidelity. This study may also provide some insight into the way in which nucleic acid enzymes evolved the capability to recognize the key structural features of nucleic acids.

Received: August 31, 2013
Published online: October 24, 2013

Keywords: nucleic acid analogues · RNA polymerase II · sugar integrity · synthetic biology · transcriptional fidelity

[1] S. A. Benner, A. M. Sismour, *Nat. Rev. Genet.* **2005**, *6*, 533–543.
[2] J. B. Opalinska, A. M. Gewirtz, *Nat. Rev. Drug Discovery* **2002**, *1*, 503–514.

[3] S. A. Benner, *Acc. Chem. Res.* **2004**, *37*, 784 – 797.
[4] S. A. Benner, *Science* **2004**, *306*, 625 – 626.
[5] A. T. Krueger, E. T. Kool, *Chem. Biol.* **2009**, *16*, 242 – 248.
[6] T. Carell, C. Brandmayr, A. Hienzsch, M. Muller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs, M. Wagner, *Angew. Chem.* **2012**, *124*, 7220 – 7242; *Angew. Chem. Int. Ed. Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 7110 – 7131.
[7] V. B. Pinheiro, P. Holliger, *Curr. Opin. Chem. Biol.* **2012**, *16*, 245 – 252.
[8] M. W. Kellinger, S. Ulrich, J. Chong, E. T. Kool, D. Wang, *J. Am. Chem. Soc.* **2012**, *134*, 8231 – 8240.
[9] C. D. Kaplan, K.-M. Larsson, R. D. Kornberg, *Mol. Cell* **2008**, *30*, 547 – 556.
[10] M. L. Kireeva, Y. A. Nedialkov, G. H. Cremona, Y. A. Purtov, L. Lubkowska, F. Malagon, Z. F. Burton, J. N. Strathern, M. Kashlev, *Mol. Cell* **2008**, *30*, 557 – 566.
[11] Y. Yuzenkova, A. Bochkareva, V. R. Tadigotla, M. Roghanian, S. Zorov, K. Severinov, N. Zenkin, *BMC Biol.* **2010**, *8*, 54.
[12] D. A. Erie, O. Hajiseyedjavadi, M. C. Young, P. H. von Hippel, *Science* **1993**, *262*, 867 – 873.
[13] D. W. Su Zhang, *Isr. J. Chem.* **2013**, *53*, 442 – 449.
[14] N. Langkjær, A. Pasternak, J. Wengel, *Bioorg. Med. Chem.* **2009**, *17*, 5420 – 5425.
[15] D. Werk, J. Wengel, S. L. Wengel, H. P. Grunert, H. Zeichhardt, J. Kurreck, *FEBS Lett.* **2010**, *584*, 591 – 598.
[16] M. B. Laursen, M. M. Pakula, S. Gao, K. Fluiter, O. R. Mook, F. Baas, N. Langklaer, S. L. Wengel, J. Wengel, J. Kjems, J. B. Bramsen, *Mol. BioSyst.* **2010**, *6*, 862 – 870.
[17] M. A. Campbell, J. Wengel, *Chem. Soc. Rev.* **2011**, *40*, 5680 – 5689.
[18] A. Pasternak, J. Wengel, *Nucleic Acids Res.* **2010**, *38*, 6697 – 6706.
[19] D. Wang, D. A. Bushnell, K. D. Westover, C. D. Kaplan, R. D. Kornberg, *Cell* **2006**, *127*, 941 – 954.

