Title

Transcriptional Perturbations of 2,6-Diaminopurine and 2-Aminopurine.

Journal

ACS Chemical Biology, 17(7)

Authors

Tan, Ying

You, Changjun

Park, Jiyeong

et al.

Publication Date

2022-07-15

DOI

10.1021/acschembio.2c00369

Peer reviewed

HHS Public Access

Author manuscript

ACS Chem Biol. Author manuscript; available in PMC 2023 July 15.

Published in final edited form as:

ACS Chem Biol. 2022 July 15; 17(7): 1672–1676. doi:10.1021/acschembio.2c00369.

Transcriptional Perturbations of 2,6-Diaminopurine and 2-Aminopurine

Ying Tan,
Environmental Toxicology Graduate Program, University of California Riverside, Riverside, California 92521-0403, United States

Changjun You,
Department of Chemistry, University of California Riverside, Riverside, California 92521-0403, United States

Jiyeong Park,
Center for Genomic Integrity, Institute for Basic Science, Ulsan 44919, Republic of Korea

Hyun Suk Kim,
Center for Genomic Integrity, Institute for Basic Science, Ulsan 44919, Republic of Korea

Su Guo,
Environmental Toxicology Graduate Program, University of California Riverside, Riverside, California 92521-0403, United States

Orlando D. Schärer,
Center for Genomic Integrity, Institute for Basic Science, Ulsan 44919, Republic of Korea; Department of Biological Sciences, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea

Yinsheng Wang
Environmental Toxicology Graduate Program and Department of Chemistry, University of California Riverside, Riverside, California 92521-0403, United States

Abstract

2,6-Diaminopurine (Z) is a naturally occurring adenine (A) analog that bacteriophages employ in place of A in their genetic alphabet. Recent discoveries of biogenesis pathways of Z in bacteriophages have stimulated substantial research interest in this DNA modification. Here, we systematically examined the effects of Z on the efficiency and fidelity of DNA transcription. Our results showed that Z exhibited no mutagenic yet substantial inhibitory effects on transcription mediated by purified T7 RNA polymerase and by human RNA polymerase II in HeLa nuclear

Corresponding Author: Yinsheng Wang − Environmental Toxicology Graduate Program and Department of Chemistry, University of California Riverside, Riverside, California 92521-0403, United States; Yinsheng.Wang@ucr.edu.

Supporting Information

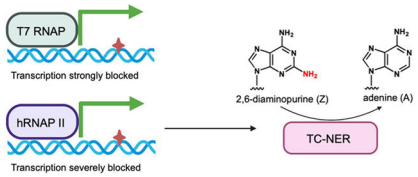
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.2c00369.
Materials and cell lines used, Sanger sequencing and Western-blot for verifying the successful knockout of *XPC* and *CSB* genes, plasmid construction, *in vitro* and cellular transcription assay, supplementary mass spectrum and gel images (PDF)

Complete contact information is available at: https://pubs.acs.org/10.1021/acschembio.2c00369

The authors declare no competing financial interest.

extracts and in human cells. A structurally related adenine analog, 2-aminopurine (2AP), strongly blocked T7 RNA polymerase but did not impede human RNA polymerase II *in vitro* or in human cells, where no mutant transcript could be detected. The lack of mutagenic consequence and the presence of a strong blockage effect of Z on transcription suggest a role of Z in transcriptional regulation. Z is also subjected to removal by transcription-coupled nucleotide-excision repair (TC-NER), but not global-genome NER in human cells. Our findings provide new insight into the effects of Z on transcription and its potential biological functions.

Graphical Abstract



2,6-Diaminopurine (Z), or 2-aminoadenine, is a naturally occurring adenine (A) analog (Figure 1). In the S-2L cyanophage genome, A is completely substituted with Z.[1] The Z-containing genome is resistant to bacterial restriction endonucleases, providing S-2L cyanophage a selective advantage.[2] The biogenesis pathways of base Z were recently elucidated,[3–6] and it was found that the Z-containing genome is more widespread in bacteriophages than previously thought.[3]

Multiple DNA modifications discovered in bacteriophages were also detected in bacterial or human genomes, albeit with different biological functions. For example, to dodge the host defense system, some bacteriophages replace cytosine (C) with 5-methylcytosine (5mC) in their genomes.[7] On the other hand, 5mC is an important epigenetic regulator in human cells.[7] Interestingly, while bacteriophages benefit from Z-substituted genomes, introducing Z into the genome is toxic to bacteria,[8] and Z has not yet been detected in the human genome. Z can readily base pair with thymine (T) or uracil (U) by forming three hydrogen bonds.[9] However, not much is known about how Z substitution modulates the efficiency or fidelity of transcription and whether Z is subjected to repair in mammalian cells.

To shed light on the above questions, we first explored whether Z in template DNA affects the efficiency and fidelity of transcription by a bacteriophage RNA polymerase *in vitro*, and by human RNA polymerase II *in vitro* and in human cells. We also investigated whether Z could be repaired in human cells and compared the transcriptional outcomes of Z with those of another structurally related adenine analog, 2-aminopurine (2AP).

We employed a competitive transcription and adduct bypass (CTAB) assay[10,11] to assess the effects of Z and 2AP on transcription *in vitro* and in human cells. The modified nucleoside-containing plasmids were prepared by inserting a single Z or 2-AP on the template strand of a nonreplicative double-stranded plasmid downstream of cytomegalovirus (CMV) and T7 promoters (Figure 1). A normal adenine base was placed at the Z or 2AP site in the control plasmid. The competitor plasmid is also modification-free but contains three additional nucleobases next to the modification site not present in the control plasmid

yielding a transcription product with a distinct length. The Z-, 2AP-, or control A-containing plasmid were premixed with the competitor plasmid and used as templates for transcription. The runoff transcripts were subjected to reverse-transcription PCR (RT-PCR), restriction digestion, and polyacrylamide gel electrophoresis (PAGE) as well as liquid chromatography-tandem mass spectrometry (LC-MS/MS) analyses (Figure 1).

Frist, we quantitatively assessed the effects of Z and 2AP on transcription elongation *in vitro* catalyzed by purified T7 bacteriophage RNA polymerase (T7 RNAP) or human RNA polymerase II (hRNAPII) in a HeLa cell nuclear extract. Linearized plasmids were used as templates and the transcription reactions were terminated after a 1-h incubation. PAGE analysis showed that base Z posed a strong blocking effect on hRNAPII with a relative transcriptional bypass efficiency (RBE) of around 1.3%, compared to the normal base A. Z also strongly blocked T7 RNAP, with an RBE of 18.4% (Figure 2). On the other hand, 2AP only modestly impeded transcription elongation by hRNAPII (with a 65.9% RBE) but strongly impaired transcription mediated by T7 RNAP (with a 12.3% RBE; Figure 3). In addition, our LC-MS/MS analysis results revealed that the full-length transcripts from *in vitro* transcription of Z- or 2AP-containing template did not elicit any detectable mutations (Figures S1 and S2).

We further asked how Z and 2AP influence transcription in human cells. To this end, we premixed Z-, 2AP-, and A-containing plasmids individually with the competitor plasmid and cotransfected them into HEK293T and HeLa cells (Figure 1). After 24 h, the transcripts were extracted and analyzed as described above. Our results showed that base Z strongly blocks transcription machinery in human cells, with the RBE values being ~18.6% and ~8.1% in HEK293T and HeLa cells, respectively (Figure 2). By contrast, 2AP was well tolerated by the human transcription machinery, where the RBE values were comparable to that of the unmodified adenine in both HEK293T and HeLa cells (Figure 3). There was no detectable mutant transcript arising from either substrate in human cells, as revealed by PAGE and LC-MS/MS results (Figures 2 and 3 and S1 and S2).

Considering that base Z markedly impedes cellular transcription, we next examined whether it is subjected to repair by the nucleotide excision repair (NER) pathway in human cells. XPA and ERCC1 are essential protein factors in NER.[12,13] We performed a CTAB assay in isogenic human skin fibroblasts that are XPA-deficient (GM04312) or XPA-complemented (GM15876).[14] The RBE value for Z was significantly lower in the XPA-deficient cells than the XPA-complemented cells at 24 h following transfection (Figure 4A), indicating that Z is an NER substrate in human cells. Likewise, the RBE value for Z was markedly lower in ERCC1-deficient CHO-7-27 cells[15] (3.6%) than in repair-competent CHO-AA8 cells (13.2%) at 24 h following transfection (Figure 4B), again supporting the role of NER in removing base Z in mammalian cells.

There are two major subpathways in NER, namely, global-genome NER (GG-NER) and transcription-coupled NER (TC-NER), which require XPC and CSB, respectively, early in the damage recognition process.[12,13] To determine the influence of the two subpathways on the transcriptional alteration induced by Z, we generated *XPC* and *CSB* knockout HEK293T cells by using CRISPR/Cas9. The successful ablations of these two genes were verified

by Sanger sequencing and Western blot analyses (Figure S4). The RBE of Z decreased significantly in *CSB*-knockout cells compared to that in parental HEK293T cells, while genetic ablation of *XPC* did not exert any significant effect on RBE values of Z in human cells (Figure 2). The transcriptional fidelity of Z was not compromised by genetic depletion of *XPC* or *CSB*, as no mutant transcripts were detectable from either PAGE or LC-MS/MS analyses (Figures 2, S1, and S2).

To further delineate the contributions of the two subpathways of NER in the removal of Z from human cells, we conducted the CTAB assay at different time points for HEK293T cells and the isogenic *XPC*- and *CSB*-knockout cells. Our time-dependent results showed that the RBE values of Z, relative to that of normal base A, displayed time-dependent increases in HEK293T and the isogenic *XPC*-knockout cells (Figure 4C). The magnitude of increase in RBE values for base Z, however, was much smaller in *CSB*-knockout cells (Figure 4C). These results are consistent with the active repair of Z in HEK293T and the isogenic XPC-deficient cells and the ensuing formation of modification-free template DNA for transcription. By contrast, this repair is impaired in CSB-deficient HEK293T cells, indicating that the repair occurs by TC-NER. Moreover, *in vitro* experiments with the incubation of recombinant XPC-RAD23B protein complex with 24-merZ-containing double-stranded oligonucleotides (dsODNs) in the same sequence context as cellular experiments showed significantly less efficient binding of the Z-containing dsODN compared to the corresponding unmodified A-containing dsODN (Figure S3). Previous studies have demonstrated that XPC-RAD23B binding requires thermodynamic destabilization of duplex DNA, and that DNA lesions such as adducts formed by aristolochic acid do not destabilize DNA duplexes, fail to specifically bind XPC-RAD23B, and thus escape GG-NER.[16] Together, these results indicated that base Z is not a substrate for GG-NER; however, when situated on the template strand of an actively transcribed gene, Z can be readily repaired by TC-NER.

The discovery of a third purine base as a genetic code in place of A in bacteriophages has attracted substantial interest recently;[3–6] however, it is unclear how this replacement affects transcription. In this study, we systematically investigated the effects of this adenine analog on the process of DNA transcription *in vitro* and in human cells and found that Z poses a significant impediment to transcription by RNAPII in human cells.

Sturla and co-workers[17] showed that altered minor-groove hydrogen bonding in DNA can block transcription elongation by T7 RNA polymerase and that the 2AP-U wobble base pair constitutes stronger blockage to transcription elongation than the Z−U base pair, which maintains Watson−Crick-type base pairing. In agreement with this previous study, our results showed that, when located on the transcribed strand, Z conferred less blockage than 2AP to transcription mediated by T7 RNA polymerase, though the Z−U base pair displays better thermal stability.[17] It is worth noting that the aforementioned study[17] used about a 16-fold higher T7 RNA polymerase to template ratio than in this study, which resulted in different absolute transcript yields between the two studies; nevertheless, the magnitude of change was similar, i.e., T7 RNA polymerase transcribed through Z about 1.5-fold more efficiently than 2AP. Contrary to single-subunit T7 RNA polymerase, the multisubunit human RNA polymerase II in HeLa nuclear extract and cultured human cells

efficiently transcribed 2AP-containing templates but was severely blocked by Z, indicating differential recognition mechanisms by RNA polymerases from different organisms.

Apart from inducing mutations during DNA replication, modified nucleosides in template DNA can give rise to mutations during transcription, a process called transcriptional mutagenesis (TM).[18,19] Several lines of evidence showed that TM of simple DNA modifications such as 8-oxo-7,8-dihydroguanine (8-oxoG) or O^6-methylguanine (O^6-meG) could lead to pathogenic protein production and contribute to neurodegenerative disease and cancer.[20,21] We did not observe any mutations arising from transcription across Z or 2AP *in vitro* or in human cells. In theory, Z or 2AP can not only pair with T and U via three hydrogen bonds but also form a wobble base pair with C. A previous study showed that, in the absence of UTP, T7 RNA polymerase can misincorporate CTP opposite Z or 2AP; however, the ensuing wobble base pair hampered transcription elongation, resulting in truncated transcripts.[17] On the grounds that our experiments were conducted in the presence of all four natural ribonucleoside triphosphates, the absence of mutations is not surprising. Similarly, DNA replication across 2AP mediated by wild-type T4 DNA polymerase did not yield any mutations.[22] With mutant bacteriophage DNA polymerase, 2AP was marginally mutagenic, pairing with T about 100 times more frequently than with C, and it has been shown that these mutagenic properties preferentially occur in AT-rich regions.[23] In addition, the imino tautomer formation is a primary mechanism for base mispairing. However, solution-phase NMR studies revealed that both bases in the 2AP-T base pair adopt the canonical amino tautomeric forms, ruling out mutagenic base pairing.[24]

We also found that the transcription efficiency of Z was lower in NER-deficient cells than in NER-competent cells. Moreover, genetic depletion of CSB, but not XPC, strongly impedes transcription across Z. Our previous observations from the CTAB assay revealed similar findings for N^2-(1-carboxyethyl)-2′-deoxyguanosine located on the template strand, the transcription efficiency of which is exacerbated by a loss of TC-NER.[25] Similarly, aristolochic acid-induced aristolactam (AL)-DNA adducts are efficiently repaired by TC-NER but are resistant to GG-NER, resulting in a mutation spectrum almost exclusively on the nontranscribed strand.[16,26] Recent structural studies showed that CSB could rescue stalled RNA polymerase II in two ways, i.e., to push Pol II forward to bypass less bulky adducts or to recruit TFIIH to the damage sites and initiate TC-NER,[27,28] a NER subpathway that selectively repairs DNA lesions on the transcribed strand of active genes.[12] Our results showed that depletion of CSB or downstream NER factors (XPA or ERCC1) substantially attenuated the transcription bypass efficiency of Z, suggesting a role of CSB in initiating TC-NER to rescue the Z-induced RNPII stalling in human cells. Along this line, TC-NER has been shown to sense transcription blockage rather than specific structural features of a lesion itself.[13] A number of other nonbulky, yet potent transcription stalling DNA modifications, including abasic sites and 5-guanidinohydantoin (Gh), are efficient substrates for TC-NER.[29–31]

In conclusion, we demonstrated that base Z did not forge transcriptional mutagenesis, but it strongly attenuated the efficiencies of transcription mediated by single-subunit T7 RNA polymerase *in vitro* and multisubunit human RNA polymerase II *in vitro* and in human cells. This result indicates that, like 5mC, the noncanonical DNA base Z, first

discovered in bacteriophages as a defense mechanism and later found to be an important epigenetic regulator, could potentially play a role in transcriptional regulation, albeit through different mechanisms. In this context, 5mC recruits methyl-CpG-binding proteins to block transcription,[32] whereas Z attenuates transcription directly by impeding RNA polymerase-mediated transcription elongation. In addition, Z was discovered in meteorites among other nucleobases[33] and was often formed during prebiotic synthesis reactions that generate adenine and guanine.[33,34] It was suggested that Z could have existed on early earth alongside A and G.[35] Furthermore, Z exhibits excellent capacities in nonenzymatic oligomerization,[36] template-directed ribonucleotide incorporation[37] and damage repair.[35] The discovery that Z strongly impedes transcription and is readily repaired by TC-NER in human cells expands our knowledge of this naturally occurring nucleobase.

Together, our study provides important insights into the impact of Z and 2AP on the efficiency and fidelity of DNA transcription. In the future, it will be pivotal to examine the structural basis underlying Z's substantial transcriptional blockage effect and the differential recognition of Z and 2AP by different RNA polymerases. It will also be interesting to investigate how placement of Z and 2-AP on nontemplate strand affects transcription *in vitro* and in cells.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (R35 ES031707 to Y.W. and P01 CA092584 to ODS) and the Korean Institute for Basic Science (IBS-R022-A1 to ODS).

REFERENCES

(1). Kirnos MD; Khudyakov IY; Alexandrushkina NI; Vanyushin BF 2-Aminoadenine Is an Adenine Substituting for a Base in S-2L Cyanophage DNA. Nature 1977, 270 (5635), 369–370. [PubMed: 413053]

(2). Szekeres M; Matveyev AV Cleavage and Sequence Recognition of 2,6-Diaminopurine-Containing DNA by Site-Specific Endonucleases. FEBS Lett. 1987, 222 (1), 89–94. [PubMed: 2820801]

(3). Zhou Y; Xu X; Wei Y; Cheng Y; Guo Y; Khudyakov I; Liu F; He P; Song Z; Li Z; Gao Y; Ang EL; Zhao H; Zhang Y; Zhao S A Widespread Pathway for Substitution of Adenine by Diaminopurine in Phage Genomes. Science 2021, 372, 512–516. [PubMed: 33926954]

(4). Sleiman D; Garcia PS; Lagune M; Loc'h J; Haouz A; Taib N; Röthlisberger P; Gribaldo S; Marlière P; Kaminski PA A Third Purine Biosynthetic Pathway Encoded by Aminoadenine-Based Viral DNA Genomes. Science 2021, 372, 516–520. [PubMed: 33926955]

(5). Pezo V; Jaziri F; Bourguignon P-Y; Louis D; Jacobs-Sera D; Rozenski J; Pochet S; Herdewijn P; Hatfull GF; Kaminski P-A; Marliere P Noncanonical DNA Polymerization by Aminoadenine-Based Siphoviruses. Science 2021, 372, 520–524. [PubMed: 33926956]

(6). Czernecki D; Legrand P; Tekpinar M; Rosario S; Kaminski P-A; Delarue M How Cyanophage S-2L Rejects Adenine and Incorporates 2-Aminoadenine to Saturate Hydrogen Bonding in Its DNA. Nat. Commun 2021, 12 (1), 2420. [PubMed: 33893297]

(7). Colot V; Rossignol J-L Eukaryotic DNA Methylation as an Evolutionary Device. BioEssays 1999, 21 (5), 402–411. [PubMed: 10376011]

(8). Czernecki D; Bonhomme F; Kaminski P-A; Delarue M Characterization of a Triad of Genes in Cyanophage S-2L Sufficient to Replace Adenine by 2-Aminoadenine in Bacterial DNA. Nat. Commun 2021, 12, 4710. [PubMed: 34354070]

(9). Piccirilli JA; Benner SA; Krauch T; Moroney SE; Benner SA Enzymatic Incorporation of a New Base Pair into DNA and RNA Extends the Genetic Alphabet. Nature 1990, 343 (6253), 33–37. [PubMed: 1688644]

(10). You C; Wang Y Quantitative Measurement of Transcriptional Inhibition and Mutagenesis Induced by Site-Specifically Incorporated DNA Lesions in Vitro and in Vivo. Nat. Protoc 2015, 10 (9), 1389–1406. [PubMed: 26292071]

(11). Tan Y; Guo S; Wu J; Du H; Li L; You C; Wang Y DNA Polymerase eta promotes the transcriptional bypass of N2-alkyl-2'-deoxyguanosine adducts in human cells. J. Am. Chem. Soc 2021, 143, 16197–16205. [PubMed: 34555898]

(12). Hanawalt PC; Spivak G Transcription-Coupled DNA Repair: Two Decades of Progress and Surprises. Nat. Rev. Mol. Cell Biol 2008, 9 (12), 958–970. [PubMed: 19023283]

(13). Marteijn JA; Lans H; Vermeulen W; Hoeijmakers JHJ Understanding Nucleotide Excision Repair and Its Roles in Cancer and Ageing. Nat. Rev. Mol. Cell Biol 2014, 15 (7), 465–481. [PubMed: 24954209]

(14). Kitsera N; Rodriguez-Alvarez M; Emmert S; Carell T; Khobta A Nucleotide Excision Repair of Abasic DNA Lesions. Nucleic Acids Res. 2019, 47 (16), 8537–8547. [PubMed: 31226203]

(15). Rolig RL; Layher SK; Santi B; Adair GM; Gu F; Rainbow AJ; Nairn RS Survival, Mutagenesis, and Host Cell Reactivation in a Chinese Hamster Ovary Cell ERCC1 Knock-out Mutant. Mutagenesis 1997, 12 (4), 277–283. [PubMed: 9237774]

(16). Sidorenko VS; Yeo J-E; Bonala RR; Johnson F; Schärer OD; Grollman AP Lack of Recognition by Global-Genome Nucleotide Excision Repair Accounts for the High Mutagenicity and Persistence of Aristolactam-DNA Adducts. Nucleic Acids Res. 2012, 40 (6), 2494–2505. [PubMed: 22121226]

(17). Tanasova M; Goeldi S; Meyer F; Hanawalt PC; Spivak G; Sturla SJ Altered Minor-Groove Hydrogen Bonds in DNA Block Transcription Elongation by T7 RNA Polymerase. Chem-BioChem. 2015, 16 (8), 1212–1218.

(18). Saxowsky TT; Doetsch PW RNA Polymerase Encounters with DNA Damage: Transcription-Coupled Repair or Transcriptional Mutagenesis? Chem. Rev 2006, 106 (2), 474–488. [PubMed: 16464015]

(19). Bregeon D; Doetsch PW Transcriptional Mutagenesis: Causes and Involvement in Tumour Development. Nat. Rev. Cancer 2011, 11 (3), 218–227. [PubMed: 21346784]

(20). Morreall JF; Petrova L; Doetsch PW Transcriptional Mutagenesis and Its Potential Roles in the Etiology of Cancer and Bacterial Antibiotic Resistance. J. Cell. Physiol 2013, 228 (12), 2257–2261. [PubMed: 23696333]

(21). Saxowsky TT; Meadows KL; Klungland A; Doetsch PW 8-Oxoguanine-Mediated Transcriptional Mutagenesis Causes Ras Activation in Mammalian Cells. Proc. Natl. Acad. Sci. U A 2008, 105 (48), 18877–18882.

(22). Fidalgo da Silva E; Mandal SS; Reha-Krantz LJ Using 2-Aminopurine Fluorescence to Measure Incorporation of Incorrect Nucleotides by Wild Type and Mutant Bacteriophage T4 DNA Polymerases. J. Biol. Chem 2002, 277 (43), 40640–40649. [PubMed: 12189135]

(23). Reha-Krantz LJ; Hariharan C; Subuddhi U; Xia S; Zhao C; Beckman J; Christian T; Konigsberg W Structure of the 2-Aminopurine-Cytosine Base Pair Formed in the Polymerase Active Site of the RB69 Y567A-DNA Polymerase. Biochemistry 2011, 50 (46), 10136–10149. [PubMed: 22023103]

(24). Sowers LC; Fazakerley GV; Eritja R; Kaplan BE; Goodman MF Base Pairing and Mutagenesis: Observation of a Protonated Base Pair between 2-Aminopurine and Cytosine in an Oligonucleotide by Proton NMR. Proc. Natl. Acad. Sci. U. S. A 1986, 83 (15), 5434–5438. [PubMed: 3461441]

(25). You C; Dai X; Yuan B; Wang J; Wang J; Brooks PJ; Niedernhofer LJ; Wang Y A Quantitative Assay for Assessing the Effects of DNA Lesions on Transcription. Nat. Chem. Biol 2012, 8, 817–822. [PubMed: 22902614]

(26). Hoang ML; Chen C-H; Sidorenko VS; He J; Dickman KG; Yun BH; Moriya M; Niknafs N; Douville C; Karchin R; Turesky RJ; Pu Y-S; Vogelstein B; Papadopoulos N; Grollman AP; Kinzler KW; Rosenquist TA Mutational Signature of Aristolochic Acid Exposure as Revealed by Whole-Exome Sequencing. Sci. Transl. Med 2013, 5 (197), 197ra102.

(27). Yan C; Dodd T; Yu J; Leung B; Xu J; Oh J; Wang D; Ivanov I Mechanism of Rad26-Assisted Rescue of Stalled RNA Polymerase II in Transcription-Coupled Repair. Nat. Commun 2021, 12 (1), 7001. [PubMed: 34853308]

(28). Kokic G; Wagner FR; Chernev A; Urlaub H; Cramer P Structural Basis of Human Transcription-DNA Repair Coupling. Nature 2021, 598 (7880), 368–372. [PubMed: 34526721]

(29). Kim N; Jinks-Robertson S Abasic Sites in the Transcribed Strand of Yeast DNA Are Removed by Transcription-Coupled Nucleotide Excision Repair. Mol. Cell. Biol 2010, 30 (13), 3206–3215. [PubMed: 20421413]

(30). Oh J; Fleming AM; Xu J; Chong J; Burrows CJ; Wang D RNA Polymerase II Stalls on Oxidative DNA Damage via a Torsion-Latch Mechanism Involving Lone Pair-π and CH-π Interactions. Proc. Natl. Acad. Sci. U. S. A 2020, 117 (17), 9338–9348. [PubMed: 32284409]

(31). You C; Wang P; Dai X; Wang Y Transcriptional Bypass of Regioisomeric Ethylated Thymidine Lesions by T7 RNA Polymerase and Human RNA Polymerase II. Nucleic Acids Res. 2014, 42 (22), 13706–13713. [PubMed: 25404131]

(32). Klose RJ; Bird AP Genomic DNA Methylation: The Mark and Its Mediators. Trends Biochem. Sci 2006, 31 (2), 89–97. [PubMed: 16403636]

(33). Callahan MP; Smith KE; Cleaves HJ; Ruzicka J; Stern JC; Glavin DP; House CH; Dworkin JP Carbonaceous Meteorites Contain a Wide Range of Extraterrestrial Nucleobases. Proc. Natl. Acad. Sci. U. S. A 2011, 108 (34), 13995–13998. [PubMed: 21836052]

(34). Hartel C; Göbel MW Substitution of Adenine by Purine-2,6-Diamine Improves the Nonenzymatic Oligomerization of Ribonucleotides on Templates Containing Thymidine. Helv. Chim. Acta 2000, 83 (9), 2541–2549.

(35). Szabla R; Zdrowowicz M; Spisz P; Green NJ; Stadlbauer P; Kruse H;Šponer J; Rak J 2,6-Diaminopurine Promotes Repair of DNA Lesions under Prebiotic Conditions. Nat. Commun 2021, 12 (1), 3018. [PubMed: 34021158]

(36). Jiménez EI; Gibard C; Krishnamurthy R Prebiotic Phosphorylation and Concomitant Oligomerization of Deoxynucleosides to Form DNA. Angew. Chem 2021, 133 (19), 10870–10878.

(37). Kozlov IA; Orgel LE Nonenzymatic Oligomerization Reactions on Templates Containing Inosinic Acid or Diaminopurine Nucleotide Residues. Helv. Chim. Acta 1999, 82 (11), 1799–1805. [PubMed: 11543571]

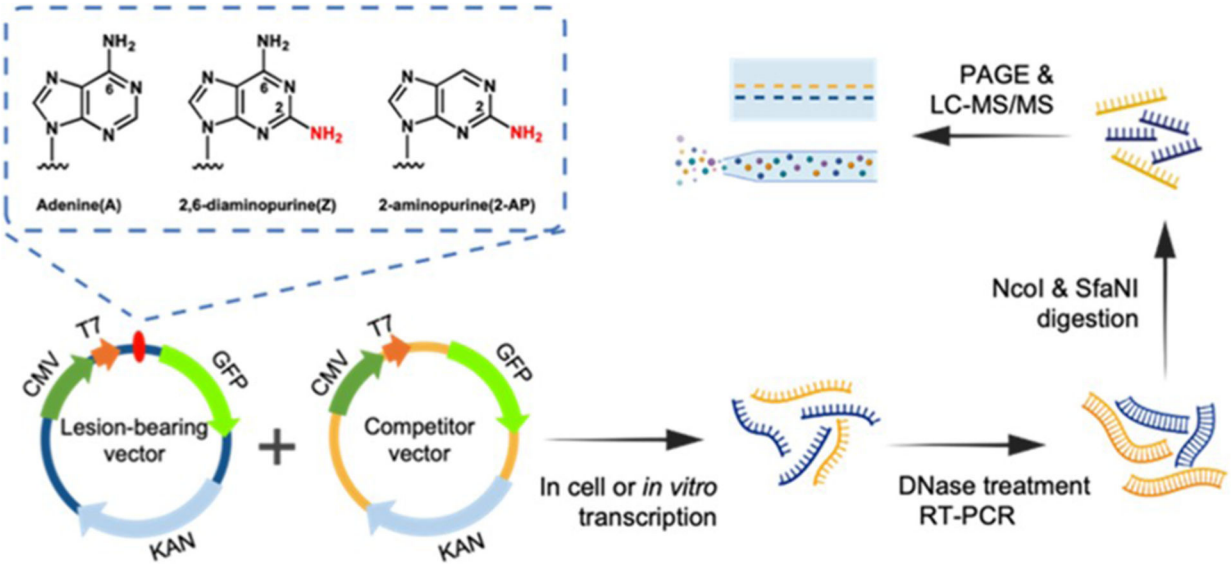


Figure 1.

A schematic diagram depicting the CTAB assay and the chemical structures of adenine and its analogs investigated in this study.

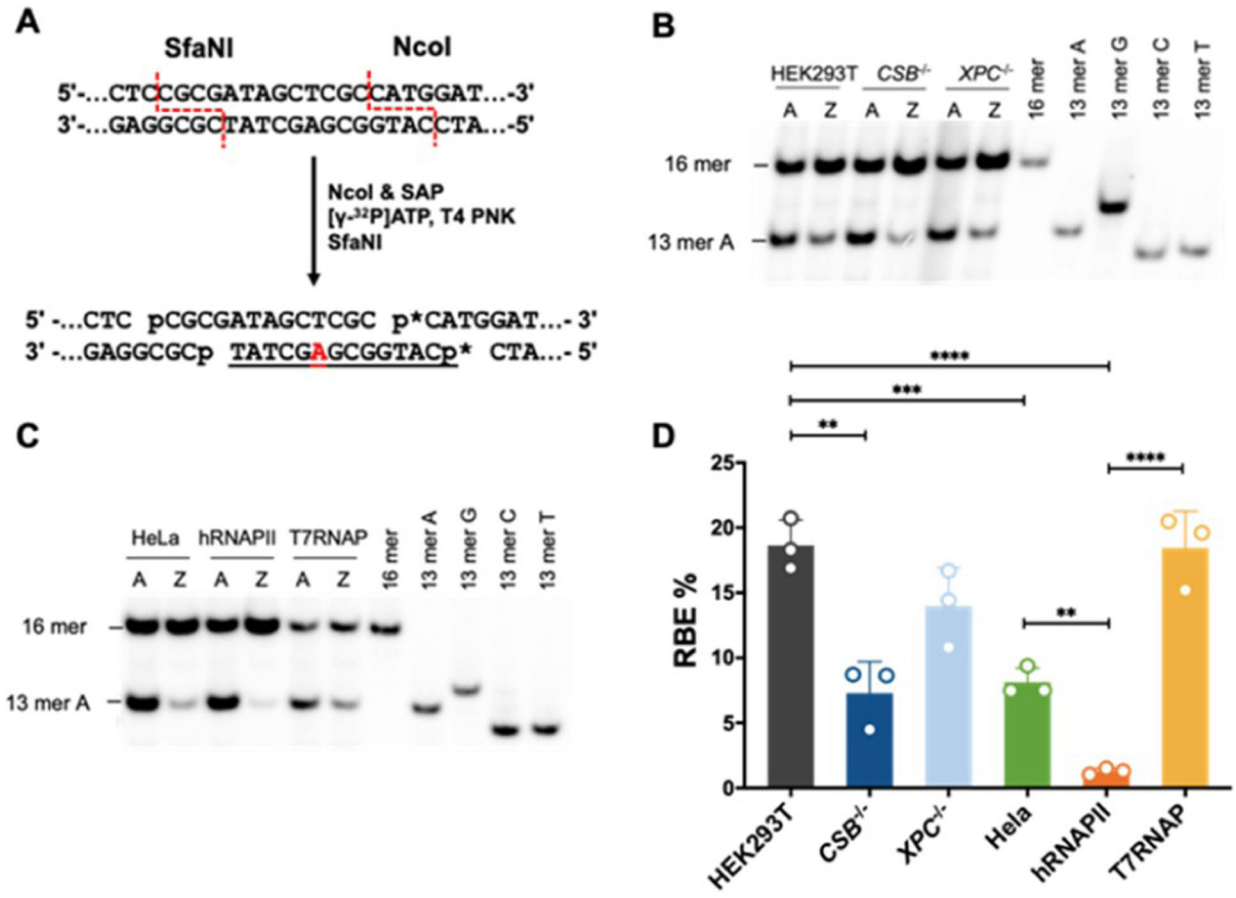


Figure 2.

Restriction digestion and postlabeling method for determining the transcriptional efficiencies of base Z *in vitro* and in human cells. (A) A schematic diagram depicting the selective labeling of the template strand via sequential digestion of the RT-PCR products. A "p*" denotes a ^{32}P-labeled phosphate group. A red colored "A" represents the site where base Z was initially installed. Representative gel images for monitoring the restriction fragments of interest in parental HEK293T cells and the isogenic cells where the *CSB* or *XPC* gene is knocked out (B), in HeLa cells and *in vitro* reactions with human RNAPII or T7 RNA polymerase (C). For all of the transfection experiments, the Z- and control A-containing plasmids were premixed with the competitor plasmid at molar ratios of 2:1 and 1:1, respectively. The transcripts were isolated from cells at 24 h following transfection. In this and other figures, the synthetic ODNs representing the restriction fragment arising from the competitor vector, i.e., d(CATGGCGATAGGCTAT), is designated as "16mer"; "13 mer G", "13 mer A", "13 mer C", and "13 mer T" represent the standard synthetic ODNs d(CATGGCGXGCTAT), where "X" is G, A, C, and T, respectively. (D) Quantitative results of RBE values of Z in different cell lines and *in vitro*. The data represent the mean ± SD of results from three independent experiments. **0.001 < P < 0.01; ***0.0001 < P < 0.001; **** P < 0.0001; ns, not significant. The P values in this and other figures were calculated by using one-way ANOVA and Tukey's multiple comparisons test.

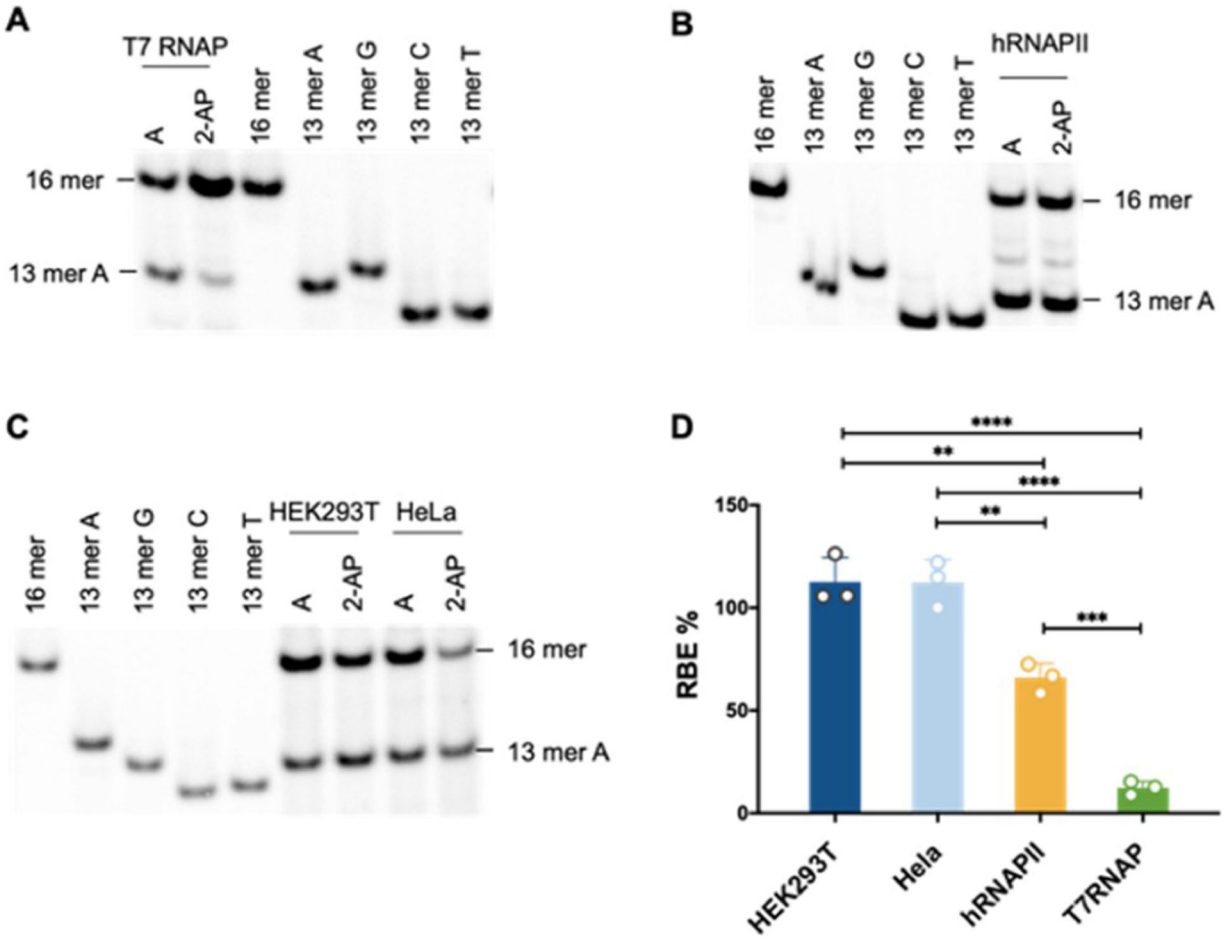


Figure 3.

Restriction digestion and postlabeling method for determining the transcriptional efficiencies of 2AP in human cells and *in vitro*. (A−C) Representative gel images showing the 16 mer (from the competitor genome) and 13 mer (from the control or lesion-containing genome) digestion products from the original template strand. For all of the transfections, the 2AP- and control A-containing plasmids were premixed with the competitor plasmid at molar ratios of 2:1 and 1:1, respectively. The transcripts were isolated from cells at 24 h following transfection. (D) The RBE values of 2AP in HEK293T cells, HeLa cells, and in *in vitro* transcription mediated by hRNAPII and T7 RNAP. The data represent the mean ± SD of results from three independent transcription experiments. **0.01 < P < 0.001; ***0.0001 < P < 0.001; **** P < 0.0001.

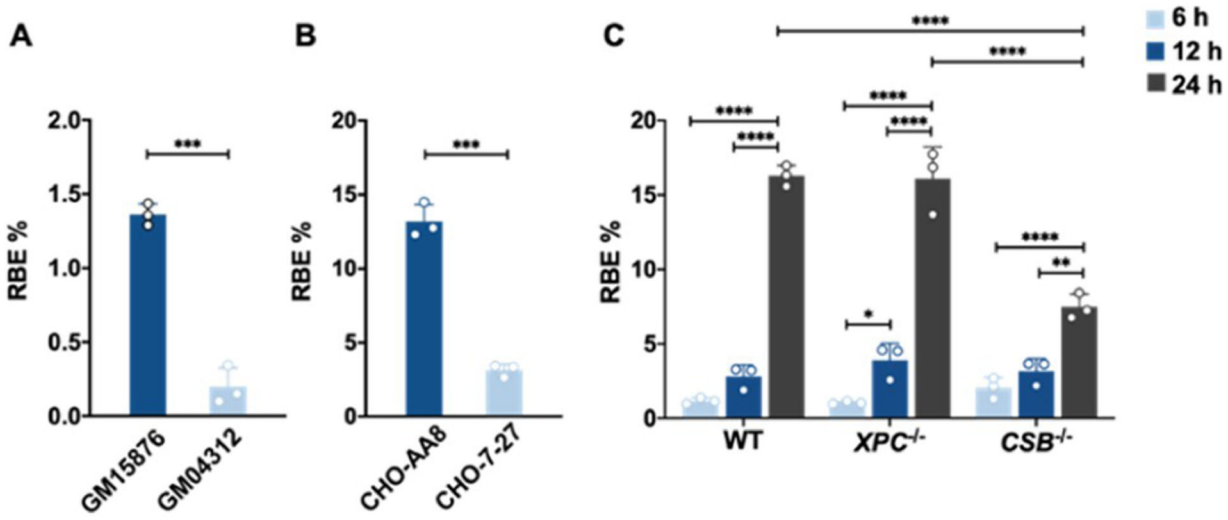


Figure 4.

Transcriptional bypass efficiencies of base Z in NER-deficient and proficient cells. The RBE values of base Z in XPA-deficient cells (GM04312) and the isogenic cells complemented with XPA cDNA (GM15876) (A) and in ERCC1-deficient (CHO-7-27) and wild-type (CHO-AA8) CHO cells (B). The RBE values in parental HEK293T cells and the isogenic CSB- and XPC-knockout cells at 6, 12, and 24 h after transfection (C). The corresponding gel images are shown in Supporting Information Figure S5. The data represent the mean ± SD of results from three independent experiments. *0.01 < P < 0.05; **0.001 < P < 0.01; ***0.0001 < P < 0.001; **** P < 0.0001.