

Altered Minor-Groove Hydrogen Bonds in DNA Block Transcription Elongation by T7 RNA Polymerase

Marina Tanasova,^[b] Silvan Goeldi,^[a] Fabian Meyer,^[a] Philip C. Hanawalt,^[c] Graciela Spivak,^[c] and Shana J. Sturla^{*[a]}

DNA transcription depends upon the highly efficient and selective function of RNA polymerases (RNAPs). Modifications in the template DNA can impact the progression of RNA synthesis, and a number of DNA adducts, as well as abasic sites, arrest or stall transcription. Nonetheless, data are needed to understand why certain modifications to the structure of DNA bases stall RNA polymerases while others are efficiently bypassed. In this study, we evaluate the impact that alterations in dNTP/rNTP base-pair geometry have on transcription. T7 RNA polymerase was used to study transcription over modified purines and pyr-

Introduction

Chemical modifications to DNA can impede gene transcription. The relationships between altered DNA structures and the progression of RNA polymerases have been characterized for mammalian RNA polymerase II (RNAP II), as well as bacteriophage T7 RNA polymerase (T7 RNAP), which is commonly used as a model RNA polymerase for evaluating transcription in vitro. Helix-distorting lesions induced by UV light (i.e., cyclobutane pyrimidine dimers and 6,4-photoproducts) or by chemotherapy drugs (e.g., cisplatin and psoralen) interrupt the progression of RNA synthesis by stalling RNA polymerase at the site of modification.^[1] Similarly, bulky DNA mono-adducts induced by carcinogens such as benzo[a]pyrene,^[2] malonaldehyde,^[3] aminofluorene,^[4] aflatoxin B1,^[5] and aristolochic acid^[6] pose a strong block to transcription by encumbering the RNA polymerase active site and inhibiting NTP incorporation. Finally, transcription is also stalled at abasic sites due to the loss of a templating base to direct NTP insertion.^[7]

In contrast to bulky adducts, non-bulky DNA modification or changes in H-bonding capacities in the templating DNA nucleobases appear to not have a predictable disruptive effect on

[a]	S. Goeldi, F. Meyer, Dr. S. J. Sturla Department of Health Sciences and Technology, ETH Zürich Schmelzbergstrasse 9, 8092 Zürich (Switzerland) E-mail: shana.sturla@hest.ethz.ch
[b]	Dr. M. Tanasova Department of Chemistry, Michigan Technological University 1400 Townsend Drive, Houghton, MI 49931 (USA)
[c]	P. C. Hanawalt, G. Spivak Department of Biology, Stanford University 371 Serra Mall, Stanford, CA 94305-5020 (USA)
	Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201500077.

imidines with altered H-bonding capacities. The results suggest that introducing wobble base-pairs into the DNA:RNA heteroduplex interferes with transcriptional elongation and stalls RNA polymerase. However, transcriptional stalling is not observed if mismatched base-pairs do not H-bond. Together, these studies show that RNAP is able to discriminate mismatches resulting in wobble base-pairs, and suggest that, in cases of modifications with minor steric impact, DNA:RNA heteroduplex geometry could serve as a controlling factor for initiating transcriptioncoupled DNA repair.

transcription. Thus, RNAP II and T7 RNAP effectively bypass 5hydroxycytosine and thymidine diols.^[8] T7 RNA polymerase also appears to tolerate a lack of H-bonding capacities in templating DNA nucleobase isosteres, as was shown with nonand weakly H-bonding synthetic pyrimidines.^[9] On the other hand, RNAP II and T7 RNAP are stalled by *O*⁶-methylguanine (*O*⁶-MeG) and, in some cases, by 8-oxoguanine (8-oxoG).^[10] These data suggest a relative tolerance of RNA polymerases towards non-bulky DNA modifications, but the possibility that the transcriptional efficiency of RNA polymerases could be diminished by certain alterations in H-bonding patterns. Furthermore, we hypothesized that, in such situations, the stalling might arise from an induced ability of the modified bases to form pairs of non-Watson–Crick geometry, rather than by changes in the steric properties of the template.

To test the impact of alterations in base-pair H-bonding and geometry on the progression of RNA synthesis, we evaluated the propensity of purines and pyrimidines with altered Hbonding capacities to stall T7 RNAP. The mechanism of the induced transcriptional stall was assessed by correlating the DNA:RNA duplex destabilization and the propensity of wobble pair formation upon incorporation of the incoming NTP opposite the modified templating base. The results provide evidence for non-bulky DNA modifications blocking the progress of transcription and support a model in which transcriptional blockages arise from changes related to H-bonding in the nascent base pair.

ChemBioChem 0000, 00, 0-0 Wiley Online Library

These are not the final page numbers! 77

1



Results and Discussion

Changes in H-bonding geometry diminish the transcriptional efficiency of T7 RNAP

We examined modified purine and pyrimidines with an established capacity to sustain Watson–Crick base pairs, wobble base pairs, or weakly H-bonded base pairs to test whether alterations in H-bonding and base-pair geometry impact transcription. Among the modified adenine (A*) and thymidine (T*) analogues (Scheme 1), 2-aminopurine (2AP) and 5-methyl-2'-



Scheme 1. A) Standard and modified nucleoside-analogue substrates for T7 RNAP as probes to induce Watson–Crick or wobble (B) pairs.

deoxyisocytosine (*isoC*) were selected for their potential to sustain wobble base pairs due to the presence of a single Hdonor at C2.^[11] 2-Aminoadenine (2AA) forms either Watson– Crick or wobble base pairs with an incoming UTP or CTP, respectively, due to the presence of H-donors at C2 and C6.^[11c] *iso*-Guanine (*isoG*) has a H-bond acceptor at C2, and is expected to form Watson–Crick pairs with either UTP or CTP;^[12] purine (P) and zebularine (Zeb) form weakly H-bonded pairs.^[13] Modified bases were incorporated into a 3.6 kDa doublestranded (ds) DNA at defined positions either in the transcribed (TS) or non-transcribed (NTS) strand downstream of the T7 RNA polymerase promoter, and their propensity to stall RNAP was evaluated in a transcription-arrest assay with T7 RNAP.^[3] An unmodified dsDNA template with the same sequence was used as a positive control for effective transcription, and a DNA template containing an abasic site (AP) was used as a negative control because it was expected to stall transcription.^[7]

To construct the 3.6 kDA DNA templates, modified adenine (A*) and thymidine (T*) probes were incorporated into 18-mer oligonucleotides that were complimentary to a portion of a single-strand (ss) covalently closed circular DNA plasmid downstream of the T7 RNAP promoter. After modified 18-mers had been annealed with the ssDNA plasmid, T4 DNA polymerase was used to complete the construction of doublestranded circular DNA plasmids. Subsequent ligation and linearization of the circular dsDNA yielded linear dsDNA templates containing either A* or T* downstream of the T7 RNAP promoter by 256 or 250 nt, respectively (Figure 1 A). A stall of transcription at the site of modification was expected to produce 256-mer or 250-mer RNA transcripts for A*- and T*-containing DNA templates, respectively, whereas complete transcription was expected to yield 472-mer run-off (RO) RNA. The corresponding NTSs contained A* and T* 257 or 263 nt downstream of the T7 RNAP promoter and were expected to produce only RO RNA.



Figure 1. Impact of modified purines (A*) and pyrimidines (T*) on the transcriptional efficiency of T7 RNAP: A) DNA templates containing A* or T* either in the TS or NTS downstream of the T7 RNAP promoter. Transcription proceeds in the direction of the bent arrow. The transcript lengths expected from full-length run-off transcript (RO) or from arrest at A*/T* are depicted by lines below each template. B) T7 RNAP transcription of dsDNA containing an A* (lanes with odd numbers correspond to TS, lanes with even numbers correspond to NTS); "Nick" indicates arrest from a nonligated portion of the template (266 nt); C) Transcription of dsDNA containing T*.

Transcription-arrest assays were carried out in single-round conditions in the presence of an excess of NTPs. Transcription products were extracted from the reaction mixture and separated on the basis of length by gel electrophoresis. Transcription was carried out in triplicate, RNA bands resulting from each transcription were quantified, and the signal intencity was normalized on the basis of C-content ([³²P]CTP was used during transcription). The loss in the efficiency of transcription (or degree of stalling) was evaluated based on the proportion of truncated versus full-length RNAs within each transcription reaction.

When transcription was performed with dsDNA templates that contained A*, we observed that the presence of the H-donor at C2 of adenine stalls transcription, yielding RNA transcripts truncated at the site of modification. Thus, whereas transcription of dsDNA containing adenine as A* generated 100% full-length RNA, 2AA as A* resulted in an RNA mixture with transcripts truncated at the 2AA site ($5\pm0.7\%$ of total RNA, Figure 1B, lane 5). The amount of truncated RNAs increased to 35% when 2AP-containing dsDA was transcribed (Figure 1B, lane 7). The magnitude of the transcriptional stall induced by 2AP was similar to that induced by AP, which is known to stall transcription (36%, Figure 1B, lane 9).^[7] In contrast, purine (P), which has no H-donating sites, and *iso*G, which contains a H-acceptor at C2, did not have any impact on transcription (< 1% transcriptional stall, Figure 2B, lanes 3

A)

39-mer transcription templates

ChemPubSoc

	+13							
5'-ATA ATA CGA CTC ACT A	TA GGG CAT ACC GGC T GC ACG GAC-3							
3'-TAT TAT GCT GAG TGA T								
promoter	transcriptional direction							
RO	5'-GGG CAC ACC GGC UGC ACG GAC-3'							
Arrest at A	5'-GGG CAC ACC GGC C -3'							
Arrest at A-1	5'-GGG CAC ACC GGC-3'							
control 30-mer template								
5'-ATA ATA CGA CTC ACT ATA GGG CAT ACC GGC-3' 3'-TAT TAT GCT GAG TGA TAT CCC GTG TGG CCG-5'								



Figure 2. Impact of 2AP:C vs. 2AP:U on transcription by T7 RNAP. A) 39-mer dsDNA templates. Transcription proceeds in the direction of the bent arrow. The transcript lengths expected from full-length run-off transcripts (RO, 21 nt) or arrest at A*/T* are depicted by sequences below the template. B) C-containing RNAs (-UTP, +[³²P]CTP); C) U-containing RNAs (+all four NTPs, +[³²P]UTP); D) Total RNA products (+all four NTPs, +[³²P]CTP). A*: transcripts arrested at a modified purine (13 nt); A*-1: transcripts arrested just prior to the modification site (12 nt); A**: transcripts possibly arising from backtracking of the polymerase.^[17]

and 11, respectively), thus emphasizing the role of H-donating sites in the stalling of T7 RNAP. The observed transcription elongation at the weakly H-bonding sites agrees with previous observations for different non- or weakly H-bonding base surrogates.^[9]

When placed in the NTS strand, the nucleoside analogues did not influence transcription by T7 RNAP and only full-length transcripts of the complementary TSs were detected (Figure 1B). Likewise, transcription over an unmodified template resulted in RO products (Figure 1B, lane 1). Finally, a minor stall at a site of incomplete ligation of the constructed template also was detected (Figure 1B, "Nick").

Data obtained from transcribing dsDNA templates containing modified pyrimidines (T*) also supported a model suggesting that transcriptional stalling occurs at modifications that alter base-pair geometry. Thus, whereas controls T, C, or the weakly H-bonding Zeb^[14] (Figure 1 C, lanes 2–4) had no impact on transcription, isoC (H donor at C2) induced a severe stalling of T7 RNAP, generating 94% truncated RNA (Figure 1 C, lane 1). It is noteworthy that only 29% of the total truncated transcripts observed for isoC corresponded to the 250 nt-long RNA expected from transcriptional termination at the site of modification. Shorter transcripts were formed in 23% (240-245 nt) and 29% (230 nt) abundance, along with several shorter transcripts upstream of the modification. The base sequences around these sites have no features expected to stall the polymerase, therefore a different process, such as polymerase backtracking, which is known to occur upon polymerase stalling and is largely promoted by a strong destabilization of the nascent DNA:RNA heteroduplex,^[15] might account for these short transcripts.

Wobble pairs as transcriptional stall inducers

To directly assess the role of Watson-Crick type A*:U versus wobble type A*:C pairing geometries of the nascent base pair in stalling T7 RNAP, we performed transcription experiments with 39-mer DNA templates that contained a G-C-T-rich 21-ntlong transcription sequence suitable for transcription in the absence or in the presence of UTP (Figure 2A). Templates contained A* at position +13 of the transcription sequence and, upon transcriptional stall at A*, were expected to produce 13mer RNA, whereas RO transcripts were expected to be 21 nt long. A 30-mer template, generating 12-mer RNA transcripts was used as a control. The RNAs produced after transcribing these templates in the absence of UTP, but in presence of CTP, ATP, GTP, and [³²P]CTP, were expected to indicate how a nascent A*:C base pair impacts the progression of transcription; RNAs obtained in the presence of [32P]UTP and all four NTPs would indicate how a nascent A*:U base pair impacts the progression of transcription.

Transcription with 39-mer templates containing A* modifications in the absence of UTP and in the presence of CTP, ATP, GTP, and [^{32}P]CTP resulted in a mixture of RNA transcripts that were truncated to various extents at A*, A*–1, and A** sites, depending on the identity of A* (Figure 2B). Truncated 13-mer RNA transcripts resulting from failed transcriptional elongation



after the incorporation of C opposite A* were detected with all templates, thus suggesting that A*:C impedes transcriptional elongation. The amounts of truncated RNAs correlate with the propensity of A*:C to be a wobble pair, with P producing the lowest amount and 2AP producing the largest amount of the 13-mer truncated RNA. A*-1 RNAs, possibly resulting from the failed incorporation of C opposite A*, were prevalent with Ade-, P- and 2AP-containing templates, whereas A** transcripts were prevalent with 2AA-containing template (Figure 2). In addition to truncated RNAs, small amounts of the full-length transcripts were observed for all templates, thereby suggesting that the polymerase could bypass A*:C wobble pairs to a small extent. Alternatively, RO transcripts might arise from the formation of and extension from low-stability A*:C pairs of the Watson-Crick type, as has been suggested for high-affinity DNA polymerases.^[16]

To determine whether forming an A*:U pair leads to the observed transcriptional stall at A*, 39-mer templates were used in the presence of all four nucleosides, and [^{32}P]UTP was used to assess the impact of U incorporation on the RNA pool. In this case, only RO transcripts (Figure 2C) were observed, thus suggesting that, unlike A*:C, A*:U pairs do not impede transcription. It is noted that, in addition to RO transcripts, a RO + 1 band from the incorporation of an additional U was observed, most significantly for P- and 2AP-containing templates; this product was not detected in

the absence of UTP.

Finally, RNA was synthesized in the presence of all four nucleotides, [³²P]CTP and [³²P]UTP (Figure 2D). As a result, a significant prevalence for RO transcripts was observed; this suggested a preferential incorporation of U opposite A* under conditions. these Truncated RNAs were assumed to result from the incorporation of C opposite A* and, in agreement with previous experiment, the largest amount of product truncated at A* was detected with 2AP-containing dsDNA template. The observed formation of truncated RNAs in the presence of UTP suggests that C could compete with U for incorporation opposite 2AP.

Alterations in H-bonding prevent transcriptional elongation

The transcriptional data presented above suggest that T7 RNAP is stalled when it encounters base pairs with non-Watson-Crick geometry, such as wobble

pairs. Furthermore, the location of the stall indicates that it occurs primarily after incorporation of an NTP opposite a modified A* or T* template, thus impairing transcriptional elongation rather than base incorporation at the site of modification. With respect to transcriptional elongation, structural data suggest that proper alignment of the newly incorporated NTP, particularly the 3'-OH of the ribose ring, in the enzyme active site is required for the formation of a phosphodiester linkage with the subsequent nucleoside and, therefore, transcriptional elongation.^[18] To visualize how nascent wobble pairs might alter the position of the ribose ring with respect to the natural Watson-Crick pair and thus impair transcriptional elongation, we performed a molecular modeling analysis (MMFF) of the corresponding modified base pairs and compared them with A:T. Base pairs were constructed in Spartan14; angles and Hbond lengths between the bases for A:T, T:A and 2AP:C were constrained based on literature data.^[9c, 13, 19]

Overlaying the modified base pairs with A:T (or T:A) indicated that the position of the ribose ring is significantly altered if nucleoside incorporation opposite 2AP results in a wobble 2AP:C pair (Figure 3 A). Likewise, alterations in the position of the ribose ring were predicted for wobble *iso*C:G and *iso*C:A, or *iso*C:C (Figure 3 B–D). The models also suggest the ribose to be displaced in Watson–Crick type pyrimidine–pyrimidine pairs *iso*C:C and *iso*C:U, the later discerned on the basis of modeling



Figure 3. Overlay of Watson–Crick-like (magenta) and wobble (green) pairs depicting changes in the position of the ribose of the incoming NTP. Magenta: natural A:T base pairs, green: modified base pairs. Modeling was performed with Spartan14 (molecular mechanics force field), and the figures were rendered with PyMol.



the DNA base pair isoC:T, pairs (Figure 3E and F) due to the change in the overall size of the base pairs. Although there is no direct evidence for such base pairs to be formed, the possibility for transcription over *iso*C to be impaired by virtually every base pair (except the Watson–Crick type *iso*C:A pair) is consistent with the striking 94% stall of T7 RNAP after incorporation of NTP(s) opposite *iso*C.

Impact of A* and T* modifications on DNA:RNA heteroduplex stability

Mismatched base pairs with A* or T* may be expected to decrease the stability of the DNA:RNA heteroduplex,^[20] potentially impeding transcription.^[15] Thus, we evaluated the relative stability of DNA:RNA duplexes containing base pairs relevant to the observed stalls in transcription. Melting temperatures $(T_m [^{\circ}C])$ were measured by variable temperature UV analysis and determined for DNA:RNA 18-mer heteroduplexes containing A*:U and T*:U pairs (5'-ATCGG CGCCG (A*/T*)CGGT GTG-3' DNA paired with the complimentary 3'-UAGCC GCGGC UGCCA CAC-5' RNA) and A*:C and T*:C pairs (3'-UAGCC GCGGC CGCCA CAC-5' as complimentary RNA). Changes in duplex stability with A* analogues did not appear to account for transcriptional stalling, as pairing P with U or C was destabilizing but there was no stalling, whereas 2AP appeared to stabilize DNA:RNA duplexes, but induced a 35% stall. Likewise, the relative stabilities of duplexes containing T* analogues did not appear to correlate with the degree of transcriptional stall; this suggests that destabilization of the duplex does not account for stalling of the polymerase at isoC. However, duplex destabilization might contribute to the formation of shorter RNA transcripts (prior to T*) after transcriptional stall, and, in the case of isoC, might induce polymerase backtracking, a hypothesis that is consistent with proofreading in mismatch excision^[16] and the propensity for polymerase backtracking to increase with duplex destabilization.[17]

The relative stabilities of duplexes containing $A^*/T^*:U$ versus $A^*/T^*:C$ were also evaluated. DNA:RNA duplexes were more stable with $A^*:U$, thus suggesting that, on the basis of nucleic acid interactions only, incorporating U opposite A^* could be favored over incorporating C (Tables 1 and 2). In contrast, the 2AP-containing duplexes were equally stable regardless of

Table 1. Relative stability of DNA:RNA duplexes containing purine analogues A* paired with U or $C_{\cdot}^{\rm [a]}$								
DNA 5'-A RNA 3'-U A*	TCGGCGCCG A* AGCCGCGGC U <i>T</i> m (A*:U) [°C]	CGGTGTG-3' GCCACAC-5' T _m (A*:C) [°C]	Transcriptio	nal stall [%]				
А	80.0	74.2	0					
2AA	81.1	76.1	5					
isoG	78.3	74.8	0					
2AP	78.4	77.6	35					
Р	76.4	74.7	2					
AP	69.8	70.2	36					
[a] 500 m duplex.	и NaCl, 0.5 mм	1 EDTA, 20 mм s	odium phosphat	е, 2 µм in				

Table 2. Relative stability of DNA:RNA duplexes containing pyrimidine analogues T* paired with U, A or $C^{[a]}_{}$								
DNA 5'-AT RNA 3'-UA T*	CGGCGC GCCGCG(N	CG T* CGGTGTG-3' GC U GCCACAC-5' T_m (T*:N) [°C]	Transcriptional stall [%]					
т —	A	80.2	0					
Zeb	Α	76.0	0					
isoC	Α	79.6						
isoC	С	74.2	94					
isoC	U	71.6						
[a] 500 mм duplex.	NaCl, ().5 mм EDTA, 20 m	м sodium phosphate, 2µм in					

whether U or C was paired with 2AP. Although A*:U was shown to be efficiently bypassed (Figure 2B), A*:C, which has a wobble base-pair geometry, accounted for the observed 35% transcriptional stall.

Conclusion

Transcription reactions performed with DNA templates containing base analogues with systematically altered H-bonding capacities indicated how alterations in base-pair geometry and size prevents transcriptional elongation by T7 RNA polymerase. Combined with molecular modeling and duplex stability data, the new findings suggest that the capacity to sustain wobbletype nascent base pairs allows nucleotide incorporation to occur but prevents transcriptional elongation, possibly due to improper alignment of the ribose of the newly inserted nucleotide. In contrast, forming nascent mismatches that either lack H-bonds or have a Watson-Crick-type geometry does not appear to have an impact on transcription. The propensity of RNAP to stall after synthesizing a base pair that includes a modified base but does not adopt a standard Watson-Cricktype geometry suggests the possibility for minor alterations in DNA bases to be substrates for transcription-coupled repair, which is initiated by RNAP stalling.

Experimental Section

Reagents: All enzymes, ribonucleoside and deoxyribonucleoside triphosphates, helper phage M13K07, and proteinase K were from New England Biolabs; RNase inhibitor (RNasin Plus) and the competent *Escherichia coli* cells MV1121 and MV1122 were from Promega. [α -³²P]CTP and [α -³²P]UTP were from PerkinElmer. All phosphoramidites for DNA synthesis were purchased from Glen Research (Sterling, VA, USA). RNA sequences were purchased from Thermo Scientific Custom Biopolymers. The gel extraction kit and the RNeasy Mini CleanUp Kit were purchased from QlAgen.

DNA templates for transcription: To test the propensity of modified purines to stall transcription, linearized 3.6 kDa dsDNA plasmids containing modified nucleobases at a defined position paired opposite dT, either in the TS or NTS downstream of the T7 promoter, were constructed as described previously.^[3] Separately, 3.6 kDa DNA plasmids with modified nucleobases opposite dC were constructed to evaluate the impact of noncoding strands on transcrip-



tional efficiency over modified purines. In brief, the constructs were transformed into the F' E. coli strains MV1121 and MV1122 for replication of the plus strand by the M13K07 phage. Covalently closed circular DNA was prepared by T4 DNA polymerase secondstrand synthesis on the purified single-stranded plasmid by using the appropriate 18-mer oligonucleotide as a primer. Annealing was done by heating an 18-mer and single-stranded plasmid at 75 °C for 3 min, then cooling to room temperature over 3 h. Secondstrand synthesis was performed in a 50 µL reaction mixture containing T4 DNA polymerase (six units) and T4 DNA ligase (800 cohesive end units), 10× ligase reaction buffer (5 μL), 100× BSA (0.5 μ L) and dNTP mix (5 μ L, 100 mm). Polymerization was allowed to proceed for 2 h at 37 °C, followed by ligation at 16 °C overnight. After heat inactivation of the enzymes, the plasmids were digested with HindIII (40 units) for 1 h at 37 °C in the same reaction mixture. The DNA fragments were separated with a 1% agarose gel at 100 V, and the 3.6 kb linearized fragments were purified with a gel extraction kit (Qiagen) and used in the transcription-arrest assay.

To evaluate the fidelity of transcription over modified bases and identify which nucleobases were inserted opposite A* during transcription, 39-mer dsDNA templates containing T7 RNAP promoter and a modified base downstream from the promoter were constructed. The 39-mer coding strand (template, 1 μ L of a 100 μ M solution, final concentration 10 μ M) and 39-mer noncoding strand (1 μ L of a 100 μ M solution, final concentration 10 μ M) oligonucleotides were annealed in a Tris/NaCl solution (1 μ L, 10 mM, final concentration 10 μ M) by heating at 95 °C for 3.5 min and cooling to room temperature over 3 h. The annealed DNA templates were directly used in the transcription-arrest assay.

Synthetic oligonucleotides: Oligonucleotides were synthesized by standard solid-phase chemical DNA synthesis (Mermade DNA synthesizer, DMT off mode) with modified purine phosphoramidites (Glen Research). For 3.6 kDa DNA plasmid construction, synthetic purines were incorporated into the 5'-ATCGG CGCCG A*CGGT GTG-3' sequence in place of adenine at position +8 from the 3'-end (transcriptional direction). An oliogonucleotide containing an abasic site (abasic II phosphoramidite, AP) incorporated into the 18-mer at position +8 from the 3'-end served as a positive control for T7 transcription arrest,^[7] while an oligonucleotide containing adenine acted as a negative control. Synthetic pyrimidines were incorporated into 5'-ATCGG CGCCG ACGGT* GTG-3' in place of adenine at position +4 from the 3'-end (transcriptional direction).

For thermal stability studies, 18-mer oligonucleotides containing all modifications at position +9 from the 3'-end and the complimentary RNAs were purchased from Thermo Fisher Scientific.

For 39-mer DNA templates, modified nucleotides (A*) were incorporated into the T7 RNAP promoter-containing sequence 3'-TATTA TGCTG AGTGA TATCC CGTGT GGCCG A*CGTG CCTG-5' at position +31 from the 3'-end (transcriptional direction).

Cleavage from the CPG support and removal of the protecting groups was carried out in concentrated aqueous NH₄OH for 6 h at 55 °C. Aqueous NH₄OH/MeNH₂ (1:1, *v/v*) was used with 2AA-containing oligonucleotides for complete deprotection. The oligonucleotides were purified by reversed-phase HPLC (Agilent 100) on an Agilent Eclipse XDB-C18 5 μ m 4.6 × 150 mm column. The mobile phases were as follows: solvent A: triethylammonium acetate (50 mM) and solvent B: acetonitrile. The solvent gradient for 18-mer oligonucleotides was increased linearly from 5 to 15% B over 19 min; for 39-mer oligonucleotides it was increased linearly from 8 to 13% B over 19 min. Fractions corresponding to the oligonucleotide peak were collected, lyophilized, resuspended in water,

and stored at -20 °C. The presence of each lesion was confirmed by MS analysis (Agilent MSD SL ion trap mass spectrometer with electrospray ionization). Characterization data (HPLC traces, ESI-MS) for all oligonucleotides are provided in the Supporting Information. Before use, the 18-mer oligonucleotides were phosphorylated at their 5'-ends for 1 h at 37 °C by using T4 polynucleotide kinase.

In vitro transcription of dsDNA templates: To transcribe plasmidderived dsDNA, 3.6 kDa DNA templates (1 ng) were incubated at 37 °C for 1 h in a reaction mixture (10 μL) containing T7 RNAP (50 units), a 10× RNAP reaction buffer (1 μL; containing 400 mm Tris·HCl, pH 8.0, MgCl₂ (80 mM), spermidine (20 mM) and ATP (100 mM), $[\alpha^{-32}P]$ CTP (1 μL, 3000 Cimmol⁻¹), NTP mix (1 μL, 200 μM), and RNase inhibitor (20 units)). After 1 h, proteinase K (80 μg) was added, and the mixture was incubated for another 30 min at 37 °C. After the addition of a formamide loading dye, samples were subjected to denaturing gel electrophoresis (20% polyacrylamide, 7 m urea, and 1×TBE). Gels were autoradiographed with intensifying screens and visualized with a Bio-Rad GS-363 phosphorimaging device. Quantification of arrest relative to runoff was done with Bio-Rad software by manually selecting the bands. Signal intensity was normalized on the basis of C content.

To transcribe 39-mer dsDNA, dsDNA solution (1 µL, 1 µм) was incubated for 30 min at 37 °C in a reaction mixture (10 µL) containing T7 RNAP (50 units), a $10 \times$ RNAP reaction buffer (1 μ L), NTP mix (1 μL, 200 μм), the corresponding $[\alpha^{-32}P]NTP$ (1 uL. 3000 Cimmol⁻¹), and RNAsin (20 units). After 30 min, the reaction was terminated by heating the mixture at 75 °C for 3 min. Unreacted NTPs were removed by passing the mixture through Micro Bio-Spin columns (Bio-Rad, Bio-Gel P-30). After the addition of loading buffer (10 μ urea in 1 \times TBE, 1:1) samples were subjected to denaturing gel electrophoresis (20% polyacrylamide, $7 \, \text{m}$ urea, and $1 \times$ TBE). All gels were imaged and quantified on a Phosphorimager (Molecular Dynamics) with ImageQuant.

Molecular modeling: Molecular modeling was performed with Spartan 14. The natural base pairs were constructed, and the distances between bases in the pairs were constrained in accordance with the literature.^[21] Base pairs were minimized (Merck molecular force field), and minimized structures were overlaid.

Thermal stability of DNA:RNA duplexes: Duplex melting temperatures (*T*_m) were determined from thermal denaturation curves obtained by recording absorbance at 260 nm as a function of temperature (five heating–cooling cycles in the temperature range of 50– 90 °C, temperature gradient of 0.5 °C min⁻¹) on a Cary UV/Vis spectrophotometer. Samples consisted of 18-mer DNA/RNA heteroduplexes (2 μM in duplex) dissolved in a buffered aqueous solution containing sodium cacodylate (100 mM), EDTA (0.1 mM), and sodium phosphate (8 mM). Melting temperatures were determined by the derivative method with the Cary thermal application software (version 3.0). Data are presented as the mean of five heating cycles for each sample. The maximum intercycle variation was 0.5 °C.

Acknowledgements

We acknowledge the Susan G. Komen for the Cure Foundation (postdoctoral fellowship to M.T.), the Walter Hochstrasser Stiftung (travel grant to S.G.), and the Swiss National Science Foundation (156280, (Research grant 156280 to S.J.S.)) for support of this research. We warmly thank Dr. Boris Belotserkovskii (Stanford University) for expert suggestions and help with transcriptional

ChemBioChem 0000, 00, 0-0



CHEMBIOCHEM Full Papers

assays, and Drs. Todor Angelov and Heidi Dahlmann (ETH Zürich) for helpful insights and suggestions.

Keywords: DNA modifications • mismatches • RNA polymerase • transcription • wobble base pairs

- [1] a) S. Tornaletti, P. C. Hanawalt, *Biochimie* 1999, *81*, 139–146; b) S. Tornaletti, *Mutat. Res.* 2005, *577*, 131–145.
- [2] D. J. Choi, D. J. Marinoalessandri, N. E. Geacintov, D. A. Scicchitano, *Bio-chemistry* 1994, 33, 780-787.
- [3] S. D. Cline, J. N. Riggins, S. Tornaletti, L. J. Marnett, P. C. Hanawalt, Proc. Natl. Acad. Sci. USA 2004, 101, 7275–7280.
- [4] B. A. Donahue, R. P. P. Fuchs, D. Reines, P. C. Hanawalt, J. Biol. Chem. 1996, 271, 10588–10594.
- [5] Y. O. Alekseyev, M. L. Hamm, J. M. Essigmann, Carcinogenesis 2004, 25, 1045–1051.
- [6] V. S. Sidorenko, J. E. Yeo, R. R. Bonala, F. Johnson, O. D. Scharer, A. P. Grollman, Nucleic Acids Res. 2012, 40, 2494–2505.
- [7] S. Tornaletti, L. S. Maeda, P. C. Hanawalt, Chem. Res. Toxicol. 2006, 19, 1215–1220.
- [8] a) N. Charlet-Berguerand, S. Feuerhahn, S. E. Kong, H. Ziserman, J. W. Conaway, R. Conaway, J. M. Egly, *EMBO J.* 2006, *25*, 5481–5491; b) S. Tornaletti, L. S. Maeda, D. R. Lloyd, D. Reines, P. C. Hanawalt, *J. Biol. Chem.* 2001, *276*, 45367–45371; c) S. Cui, Y. H. Kim, C. H. Jin, S. K. Kim, M. H. Rhee, O. S. Kwon, B. J. Moon, *BMB Rep.* 2009, *42*, 373–379.
- [9] a) S. Ulrich, E. T. Kool, *Biochemistry* 2011, *50*, 10343 10349; b) M. W. Kellinger, S. Ulrich, J. N. Chong, E. T. Kool, D. Wang, *J. Am. Chem. Soc.* 2012, *134*, 8231 8240; c) L. Xu, K. V. Butler, J. Chong, J. Wengel, E. T. Kool, D. Wang, *Nucleic Acids Res.* 2014, *42*, 5863 5870.
- [10] a) A. Dimitri, J. A. Burns, S. Broyde, D. A. Scicchitano, *Nucleic Acids Res.* 2008, *36*, 6459–6471; b) S. Tornaletti, L. S. Maeda, R. D. Kolodner, P. C. Hanawalt, *DNA Repair* 2004, *3*, 483–494.

- [11] a) L. J. Reha-Krantz, C. Hariharan, U. Subuddhi, S. Xia, C. Zhao, J. Beckman, T. Christian, W. Konigsberg, *Biochemistry* 2011, *50*, 10136–10149;
 b) A. Ronen, *Mutat. Res.* 1980, *75*, 1–47; c) S. A. Strobel, T. R. Cech, N. Usman, L. Beigelman, *Biochemistry* 1994, *33*, 13824–13835; d) J. N. Patro, M. Urban, R. D. Kuchta, *Biochemistry* 2009, *48*, 180–189; e) F. Le Page, A. Guy, J. Cadet, A. Sarasin, A. Gentil, *Nucleic Acids Res.* 1998, *26*, 1276–1281.
- [12] A. M. Bukowska, J. T. Kusmierek, Acta Biochim. Pol. 1996, 43, 247–254.
 [13] C. L. Dowd, B. T. Sutch, I. S. Haworth, R. Eritja, V. E. Marquez, A. S. Yang,
- Nucleosides Nucleotides Nucleic Acids 2008, 27, 131 145. [14] A. K. Oyelere, J. R. Kardon, S. A. Strobel, *Biochemistry* 2002, 41, 3667 – 3675.
- [15] a) M. L. Kireeva, N. Komissarova, D. S. Waugh, M. Kashlev, J. Biol. Chem. 2000, 275, 6530–6536; b) P. Cramer, D. A. Bushnell, R. D. Kornberg, Science 2001, 292, 1863–1876.
- [16] J. J. Warren, L. J. Forsberg, L. S. Beese, Proc. Natl. Acad. Sci. USA 2006, 103, 19701–19706.
- [17] E. Nudler, A. Mustaev, E. Lukhtanov, A. Goldfarb, Cell 1997, 89, 33-41.
- [18] a) M. D. McCain, A. S. Meyer, S. S. Schultz, A. Glekas, T. E. Spratt, *Biochemistry* 2005, 44, 5647–5659; b) D. T. Minnick, K. Bebenek, W. P. Osheroff, R. M. Turner, Jr., M. Astatke, L. Liu, T. A. Kunkel, C. M. Joyce, *J. Biol. Chem.* 1999, 274, 3067–3075; c) L. G. Brieba, R. Sousa, *Biochemistry* 2000, 39, 919–923.
- [19] a) L. C. Sowers, G. V. Fazakerley, R. Eritja, B. E. Kaplan, M. F. Goodman, Proc. Natl. Acad. Sci. USA 1986, 83, 5434–5438; b) M. F. Goodman, Proc. Natl. Acad. Sci. USA 1997, 94, 10493–10495.
- [20] I. Sidorenkov, N. Komissarova, M. Kashlev, Mol. Cell 1998, 2, 55-64.
- [21] C. Fonseca Guerra, F. M. Bickelhaupt, J. G. Snijders, E. J. Baerends, J. Am. Chem. Soc. 2000, 122, 4117–4128.

Received: February 11, 2015 Published online on ■■ ■, 0000

7

FULL PAPERS

M. Tanasova, S. Goeldi, F. Meyer, P. C. Hanawalt, G. Spivak, S. J. Sturla*

Altered Minor-Groove Hydrogen **Bonds in DNA Block Transcription Elongation by T7 RNA Polymerase**



Stalled transcription

The propensity of RNAP to stall on the basis of alterations in base-pair geometry within the DNA:RNA heteroduplex was evaluated. The data show that the progression of T7 RNAP in RNA synthesis is disrupted by nascent wobble base pairs but not impacted by Watson-Crick type or non-H-bonded mismatches.