BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

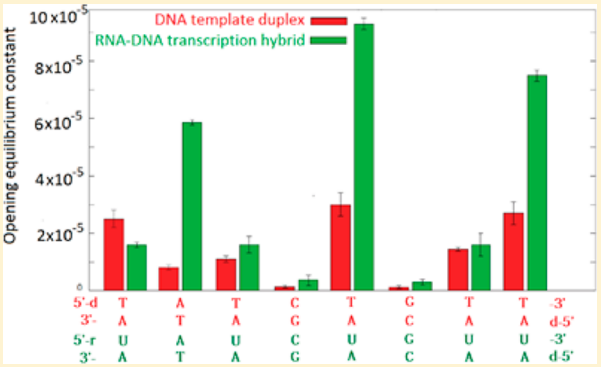
Site-Resolved Structural Energetics of the T7 Concatemer Junction

Jie Zhang and Irina M. Russu*

Department of Chemistry and Molecular Biophysics Program, Wesleyan University, Middletown, Connecticut 06459, United States

Supporting Information

ABSTRACT: The concatemer junction is a conserved sequence of 8 bp, which is strategically located at the junction between the head-to-tail repeats of genomic DNA in T7 and related bacteriophages. The RNA polymerase pauses at this site to recruit the machinery necessary for cleavage of the concatemer into single genome DNA. During pausing, the transcription bubble collapses and the transcription RNA−DNA hybrid is shortened to only 3 bp. This work addresses the question of the role of the nucleic acid components of the transcription elongation complex in this collapse of the transcription bubble. The nucleic acid structures investigated are the DNA−DNA duplex structure present at the concatemer junction when the DNA is not transcribed and the RNA−DNA hybrid formed when the concatemer junction is transcribed. The structural energetics of each base pair in the two structures is characterized using imino proton exchange and nuclear magnetic resonance spectroscopy. The results show that 5 bp in the DNA−DNA duplex at the concatemer junction site are significantly more stable than the corresponding base pairs in the RNA−DNA hybrid that forms when the site is transcribed. Because of their energetic preference for the DNA−DNA duplex, these 5 bp favor the collapse of the transcription bubble. Four of the 5 bp with enhanced stability in the DNA−DNA duplex are located in the downstream half of the concatemer junction site. This location suggests that only after the entire concatemer junction is transcribed can the RNA−DNA hybrid accumulate sufficient structural destabilization to trigger the dissociation of the RNA and the switch of the DNA template strand from the hybrid structure to the DNA−DNA double-helical structure.



In T7 and related bacteriophages, the DNA is replicated as head-to-tail repeats of the genome, also called concatemers.[1,2] To form mature phage particles, the concatemers are cut into single genomes and then packaged. This process relies upon a conserved sequence of 8 bp that is located at the junctions between adjacent genomes (also called concatemer junction or the CJ site). The processing of the DNA also requires transcription by T7 RNA polymerase because the polymerase pauses downstream of the CJ site for tens of seconds.[3] It has been proposed that, during the time window provided by the pause, the enzyme recruits the proteins necessary for cleaving the concatemer into single-genome DNA.[4]

The base sequence in the CJ site is HATCTGTT (5′ to 3′ in the DNA nontemplate strand). At the position labeled H, the allowed bases are A, C, and T. The remaining base sequence is strictly conserved; i.e., any base substitutions or mismatches abolish the pausing of T7 RNA polymerase.[5] Transcription can also terminate at a position 6−8 bp downstream of the CJ site. The extent of termination is determined by the intervening base sequence; namely, termination is enhanced as the AT content of the sequence immediately downstream of the CJ site increases.[3] The base sequence HATCTGTT is also found in other systems, such as the *Escherichia coli rrnB* operon, the human prepro-parathyroid hormone gene, and intergenic junctions of the vesicular stomatitis virus DNA. In these cases, transcription terminates near the sequence HATCTGTT even in the absence of an AT-rich sequence downstream from it.[6−9] The base sequence HATCTGTT differs from those in other transcription terminators in an important way. The most common transcription terminators, generally called class I terminators, consist of a base sequence with inverted symmetry, which forms a hairpin in the RNA transcript, followed by an A-rich sequence in the DNA template strand. In contrast, the base sequence HATCTGTT does not exhibit any features that may result in the formation of a secondary structure in the RNA. To highlight this difference and to distinguish them from hairpin-forming class I terminators, the CJ and its related sequences are generally called class II transcription terminators.

The molecular mechanisms involved in transcription pausing and/or termination at the CJ site and other class II transcription terminators have been extensively investigated by several laboratories.[3,10,11] The largest structural change that occurs during transcription through the CJ site is the collapse of the transcription bubble. Generally, the transcription bubble encompasses ∼8−10 bp with ∼8 bases of the DNA template therein being involved in the RNA−DNA hybrid.[5,12] This geometry is maintained for thousands of bases, as the

Received: March 31, 2014
Revised: July 2, 2014
Published: July 7, 2014

dx.doi.org/10.1021/bi500393q | *Biochemistry* 2014, 53, 4806−4813

A

```
                a   b   c   1   2   3   4   5   6   7   8   c′  b′  a′
5′ - d(G    C   G   T   A   T   C   T   G   T   T   G   C   G)   - 3′
3′ -  (C    G   C   A   T   A   G   A   C   A   A   C   G   C)d  - 5′

                a   b   c   1   2   3   4   5   6   7   8   c′  b′  a′
5′ - r(G    C   G   U   A   U   C   U   G   U   U   G   C   G)   - 3′
3′ -  (C    G   C   A   T   A   G   A   C   A   A   C   G   C)d  - 5′
```

B

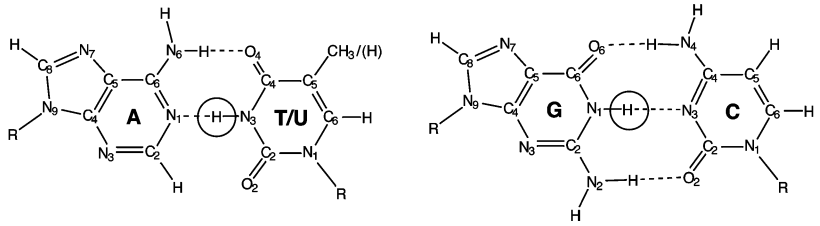


Figure 1. (A) Base sequences and numbering of base pairs in the DNA−DNA (top) and RNA−DNA hybrid (bottom) duplexes investigated. The base pairs in the CJ site (bold) are numbered 1−8. The G-C/C-G base pairs added at the ends of each duplex to prevent fraying are labeled a−c/a′−c′. (B) Structures of A-T/A-U and G-C base pairs. The imino proton in each base pair is circled.

processive RNA polymerase transcribes the DNA at speeds of up to 200 bases per second. However, at the CJ site, the transcription bubble undergoes a drastic conformational change: as the elongation complex moves 3 bases past the end of the CJ site, the upstream part of the bubble closes up, leaving behind an RNA−DNA hybrid of only 3 bp.[10] The collapse of the transcription bubble has major consequences for the functional integrity of the elongation complex. For example, the conformation of the complex changes such that the contacts of the enzyme with the upstream DNA template duplex are changed.[11,13] Furthermore, the elongation complex is greatly destabilized, and the RNA transcript is often released.[11,14]

This work addresses the question of the role of the nucleic acid framework in transcription pausing and/or termination at the CJ site. Specifically, this work aims to characterize the energetic properties of the two structures formed at the CJ site, namely, the template DNA double helix and the transcription RNA−DNA hybrid. The goal is to determine if the energetic properties of these two structures may favor the closing of the upstream part of the transcription bubble at the expense of shortening the RNA−DNA hybrid. A role of the nucleic acid scaffold in pausing and/or termination at the CJ site has been first suggested by the finding that both template and nontemplate DNA strands are required for transcription termination.[5,9] Recent genetic and biochemical analyses of various RNA polymerases have reinforced this suggestion. The CJ base sequence is present at concatemeric junctions of the genomes of other phages, such as Sp6.[4,11] The RNA polymerase from the Sp6 phage also pauses and terminates at the CJ base sequence in the same way that T7 RNA polymerase does.[3,14] However, only 20% of the amino acids that could contact the DNA in Sp6 polymerase are the same as those in T7 polymerase.[11] Mutations of these conserved amino acids have no effect on the direct recognition of the CJ site by the enzyme.[11] These findings, therefore, suggest that the pausing of the RNA polymerase at the CJ site does not result solely from specific interactions of protein side chains with nucleic acid bases. Instead, they suggest that features of the nucleic acid scaffold, induced by the CJ base sequence, may also contribute to the pausing of all these RNA polymerases at the site.

The nucleic acid structures investigated in this work contain the transcription termination/pausing CJ base sequence TATCTGTT/UAUCUGUU (Figure 1). The same base sequence has been used in previous biochemical and functional characterizations of the CJ site.[10,11] The first structure investigated is the DNA−DNA duplex present at the CJ site when the DNA is not transcribed. The second structure is the transcription RNA−DNA hybrid duplex that forms at the CJ site. Three G-C/C-G base pairs are added at each end to increase the stability of the duplex structure and to prevent the effects of fraying upon the base pairs within the CJ site.

EXPERIMENTAL PROCEDURES

Materials. The DNA and RNA strands were synthesized on an automated DNA synthesizer using solid support phosphoramidite chemistry. They were purified on Glen-Pack DNA/RNA purification cartridges. The counterions were replaced with sodium ions by repeated centrifugation in Centricon YM-3 tubes using 0.5 M NaCl, followed by repeated centrifugation against water. The concentrations of DNA and RNA strands were determined using extinction coefficients calculated as described by Cantor and co-workers.[15] The strands were annealed by equilibrating the corresponding oligonucleotides in a temperature bath at 90 °C for 10 min and slowly cooling them at room temperature. The concentrations in the final NMR samples were 1.4 mM for the DNA−DNA duplex and 1.1 mM for the RNA−DNA hybrid duplex. All NMR samples were in 10 mM phosphate buffer with 0.5 mM EDTA and 1 mM triethanolamine in a 90% H_2O/10% D_2O mixture at pH 8.2 ± 0.1 (at 20 °C). The proton resonances of triethanolamine were used to measure the pH of the sample directly in the NMR tube as we have described previously.[16]

Methods. *NMR Experiments.* The NMR experiments were performed on a Varian INOVA 500 spectrometer operating at

dx.doi.org/10.1021/bi500393q | *Biochemistry* 2014, 53, 4806−4813

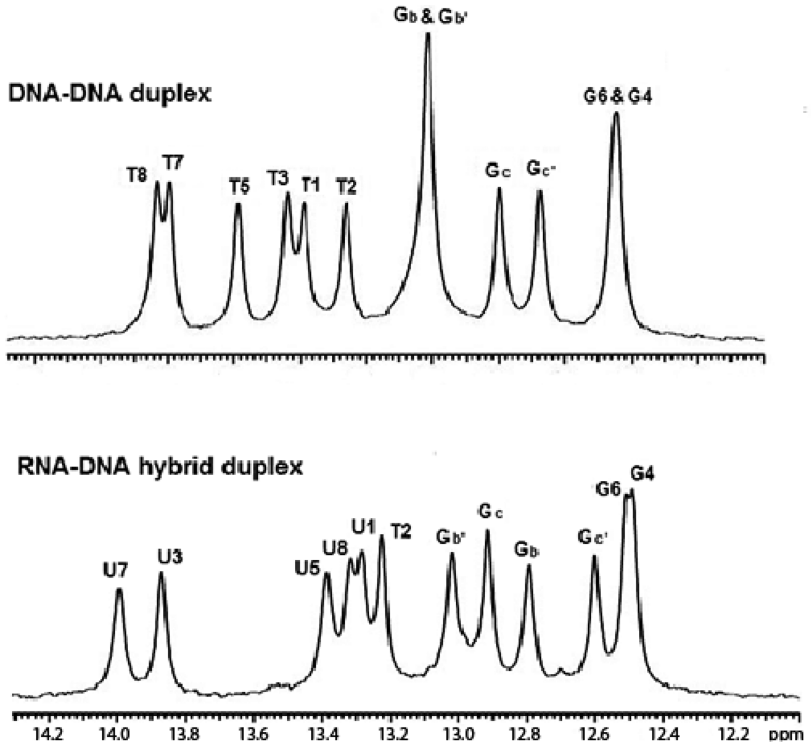


Figure 2. NMR resonances of imino protons in the DNA−DNA duplex (top) and in the RNA−DNA hybrid duplex (bottom) at 20 °C.

11.75 T. One-dimensional NMR spectra were obtained using the Jump-and-Return pulse sequence.[17] The imino proton resonances were assigned from ^{1}H−^{1}H NOESY spectra obtained with the Watergate pulse sequence[18] and a mixing time of 150 ms. The exchange rates of imino protons were measured using the method of magnetization transfer from water. In this method, the exchange is initiated by inverting selectively the water magnetization using a Gaussian 180° pulse (6.8 ms). A variable delay is allowed following water inversion for exchange to occur. A weak gradient (0.21 G/cm) is applied during this delay to prevent the effects of radiation damping upon the recovery of water magnetization to equilibrium. At the end of the exchange delay, a second Gaussian pulse (2.2 ms) is applied to bring the water magnetization back onto the *z*-axis, and the observation is made with the Jump-and-Return pulse sequence. We used 24 values of the exchange delay in the range from 1 to 800 ms. The exchange rates were calculated from the dependence of the intensity of the proton resonance of interest on the exchange delay as we have described previously.[19] The highest exchange rates that we could measure reliably using this method are ∼60−70 s^{-1}.

Imino Proton Exchange in Nucleic Acids. Our characterization of the structural energetics in the two nucleic acid duplexes relies upon the exchange of imino protons (Figure 1B) with solvent protons. The exchange occurs in two steps. The first step is a base pair opening reaction in which the base containing the imino proton moves out of the double-helical structure into an open state. In this state, the hydrogen bond holding the imino proton is broken and the imino proton is accessible to proton acceptors present in the solvent, such as NH_3.[20] The second step is the actual transfer of the imino proton to the acceptor. The exchange rate observed experimentally depends on the rates of base pair opening and closing (k_{op} and k_{cl}, respectively) as[20,21]

$$k_{ex} = \frac{k_{op}k_{ex,open}}{k_{cl} + k_{ex,open}} \quad (1)$$

where $k_{ex,open}$ is the rate of exchange of the imino proton from the open state of the base pair. This rate is proportional to the concentration of proton acceptor B:

$$k_{ex,open} = k_0 + k_B[B] \quad (2)$$

where k_B is the rate constant for transfer of the imino proton to the acceptor B in the open state of the base pair and k_0 is the rate of exchange from the open state in the absence of the proton acceptor. The dependence of the exchange rate observed experimentally on the concentration of proton acceptor B is obtained by combining eqs 1 and 2 as

$$k_{ex} = \frac{k_{op}(k_0 + k_B[B])}{k_{cl} + k_0 + k_B[B]} \quad (3)$$

The rate of exchange from the open state in the absence of proton acceptor, k_0, is much smaller than the rate of base pair closing, k_{cl} (Supporting Information).[19] Hence, when the concentration of the proton acceptor is low such that $k_B[B] \ll k_{cl}$, eq 3 simplifies to

$$k_{ex} = K_{op}(k_0 + k_B[B]) \quad (4)$$

In this so-called EX2 regime, the exchange rate is simply proportional to the concentration of the proton acceptor, with the equilibrium constant of the opening reaction defined as

$$K_{op} = \frac{k_{op}}{k_{cl}} \quad (5)$$

In this work, we have used ammonia base (NH_3) as the acceptor in imino proton exchange. The rate constants for the transfer of imino protons to NH_3 were calculated from the pK_a

dx.doi.org/10.1021/bi500393q | *Biochemistry* 2014, 53, 4806−4813

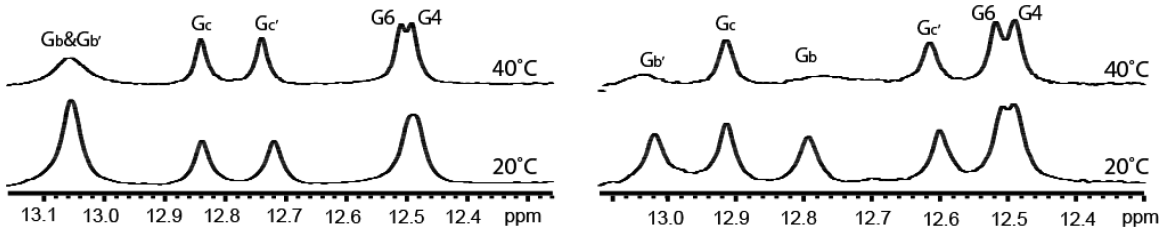


Figure 3. Dependence of guanine imino proton resonances of the DNA−DNA duplex (left) and RNA−DNA hybrid duplex (right) on temperature.

value of the imino group (pK_{NH}) and the pK_a value of ammonia (pK_B) on the basis of the following equation:[22]

$$k_B = k_{coll} \times \frac{1}{1 + 0.27 \times 10^{pK_{NH} - pK_B + 0.8}} \quad (6)$$

where k_{coll} is the rate of diffusion-controlled collision between the imino group and ammonia in the open state of the base pair (e.g., 2.6×10^9 M^{-1} s^{-1} at 20 °C). The following pK_a values of the imino groups were used: pK_a = 9.39 for uracil at 20 °C, pK_a = 9.88 for thymine at 20 °C, and pK_a = 8.93 for guanine at 40 °C.[23] The following rate constants were obtained: $k_B = 9.8 \times 10^8$ M^{-1} s^{-1} for uracil at 20 °C, $k_B = 4.2 \times 10^8$ M^{-1} s^{-1} for thymine at 20 °C, and $k_B = 1.3 \times 10^9$ M^{-1} s^{-1} for guanine at 40 °C. Increasing concentrations of ammonia base were obtained by titrating the sample with small volumes of a stock ammonia solution (0.5 or 3 M). The concentration of ammonia base was calculated from the total ammonia concentration C_0, the pH, and the pK_a value of ammonia (pK_B = 9.4 at 20 °C, and pK_B = 8.8 at 40 °C[24]) as

$$[NH_3] = C_0 \times \frac{10^{-pK_B}}{10^{-pH} + 10^{-pK_B}} \quad (7)$$

The pH was measured at each ammonia concentration, directly in the NMR tube, using the proton resonances of triethanolamine.[16]

RESULTS

The NMR resonances of the imino protons in the two nucleic acid duplexes investigated are shown in Figure 2. The resonances were assigned to specific bases in the two structures using $^1H-^1H$ NOESY experiments. The $^1H-^1H$ NOESY spectra and the procedure used for assignments are presented in the Supporting Information. As shown in Figure 2, at 20 °C, the imino proton resonances of thymines and uracils are well-resolved in both structures while those of guanines are not all resolved. Among the unresolved resonances are those from G4 and G6, which are of special interest because these two bases are part of the CJ site (Figure 1). The resonances of G4 and G6 can, however, be resolved by increasing the temperature to 40 °C (Figure 3). Given these observations, we have measured the exchange rates of thymine and uracil imino protons at 20 °C and the exchange rates of guanine imino protons at 40 °C. The obtained dependence of the exchange rates on the concentration of proton acceptor (NH_3) is shown in Figure 4 for thymine and uracil imino protons and in Figure 5 for guanine imino protons. The exchange rates of uracil imino protons are expected to be higher than those of thymine imino protons. This is because the pK_a value of the imino group in uracil (9.39 at 20 °C[23]) is lower than that in thymine (9.88 at 20 °C[23]) and thus favors the exchange of the imino proton. This fact is reflected in the rate constants for the exchange of the imino proton from the open state of the base pair (see Methods): the rate constant for uracil ($k_B = 9.8 \times 10^8$ M^{-1} s^{-1} at 20 °C) is ~2-fold higher than the rate constant for thymine ($k_B = 4.2 \times 10^8$ M^{-1} s^{-1} at 20 °C). Therefore, the exchange rate of a uracil imino proton should be ~2 times higher than that of a thymine imino proton (eq 4). This difference in the chemistry of imino proton exchange between thymine and uracil was taken into account in our analysis as described below.

The exchange data were analyzed according to the model for imino proton exchange in nucleic acids discussed in Methods. The exchange rates were fit as a function of the ammonia base concentration to eq 3 using the following rate constants for the exchange from the open state of the base pairs: $k_B = 9.8 \times 10^8$ M^{-1} s^{-1} for rU-dA base pairs, $k_B = 4.2 \times 10^8$ M^{-1} s^{-1} for dT-dA and dT-rA base pairs, and $k_B = 1.3 \times 10^9$ M^{-1} s^{-1} for G-C base pairs. The fits provided the rates of base pair opening and closing (k_{op} and k_{cl}, respectively). The equilibrium constants of the opening reaction, K_{op}, were calculated from the rates based on eq 5 and are reported in Table 1. The rates of opening and closing for these base pairs are reported in the Supporting Information. This analysis was used for the dT-dA base pairs in the DNA−DNA duplex, the $(rU\text{-}dA)_3$ and $(rU\text{-}dA)_7$ base pairs in the RNA−DNA hybrid duplex, and the G-C base pairs in the CJ site of both duplexes. For the remaining base pairs, namely, $(rU\text{-}dA)_1$, $(rA\text{-}dT)_2$, $(rU\text{-}dA)_5$, and $(rU\text{-}dA)_8$, the exchange rates are very high even at low ammonia base concentrations (Figure 4). As the ammonia concentration increases, the exchange rates become so high that they cannot be measured accurately by the NMR transfer of magnetization technique. For these base pairs, the data were analyzed only in the EX2 regime by fitting them to eq 4. The obtained opening equilibrium constants are listed in Table 1.

DISCUSSION

Our results provide a high-resolution characterization of the energetics of individual base pairs in the DNA template duplex and the transcription RNA−DNA hybrid at the CJ site. The exchange of the imino proton in each base pair monitors the transition of the base pair from its closed state inside the double helix to the open, solvent accessible state. The equilibrium constant for this transition, K_{op}, provides a measure of the stability of the base pair in its native intrahelical state. Base pairs with high structural stabilities have small equilibrium constants. In contrast, larger K_{op} values identify the base pairs with lower stabilities. An alternative way to present this correlation is the free energy change in the opening reaction, ΔG_{op}, which is related to equilibrium constant K_{op} by

$$\Delta G_{op} = -RT \ln K_{op} \quad (8)$$

where T is the absolute temperature and R is the universal gas constant. An increase in base pair stability corresponds to an increase in the free energy change ΔG_{op} for opening of the base pair.

dx.doi.org/10.1021/bi500393q | Biochemistry 2014, 53, 4806−4813

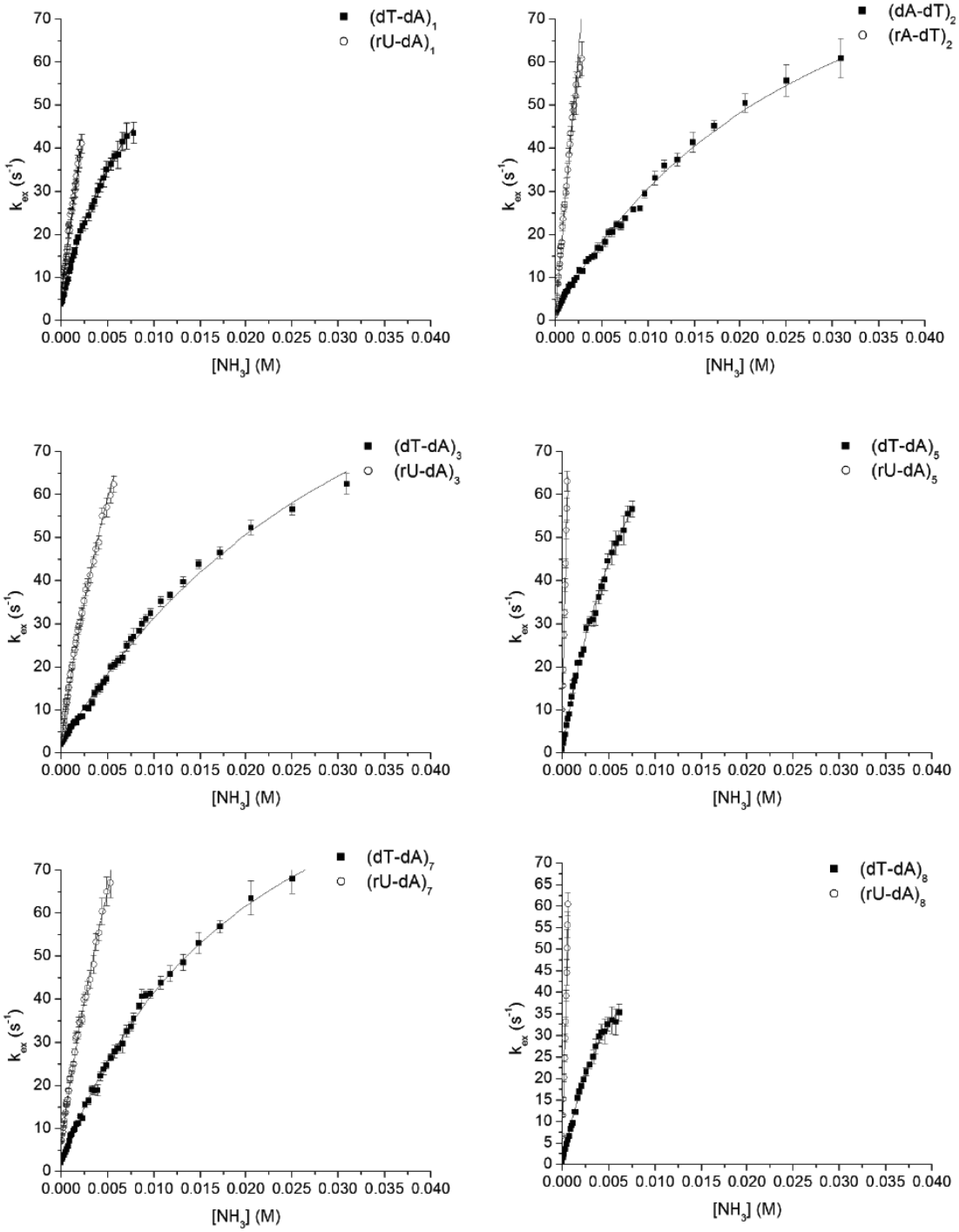


Figure 4. Dependence of the imino proton exchange rates for thymines and uracils in the CJ site on the ammonia base concentration at 20 °C. The filled squares represent the results for the DNA−DNA duplex and the empty circles the results for the RNA−DNA hybrid duplex. The rates for the DNA−DNA duplex and those for $(rU\text{-}dA)_3$ and $(rU\text{-}dA)_7$ in the RNA−DNA hybrid duplex were fit as a function of ammonia base concentration to eq 3. The rates for $(rU\text{-}dA)_1$, $(rA\text{-}dT)_2$, $(rU\text{-}dA)_5$, and $(rU\text{-}dA)_8$ in the RNA−DNA hybrid were fit as a function of the ammonia base concentration to eq 4.

The equilibrium constants K_{op} in the two structures investigated range from ~1×10^{-6} to 93×10^{-6} (Table 1). They correspond to free energy changes in base pair opening ranging from 5.4 to 8.5 kcal/mol. Interestingly, for five of the base pairs in the CJ site, the stabilities in the RNA−DNA hybrid are significantly lower than those in the DNA−DNA duplex (Table 1 and Figure 6). For example, for the base pair in the second position in the CJ site, the opening equilibrium constant in the hybrid is ~7-fold higher than that in the DNA duplex (Table 1). This difference corresponds to a destabilization of this base pair in the hybrid of 1.15 kcal/mol (Table 1 and Figure 7). Similarly, for the base pair in the fifth position, the opening equilibrium constant in the hybrid is ~3-fold higher than that in the DNA duplex, corresponding to a decrease in stability of 0.66 kcal/mol. The exceptions to this trend are the base pairs in the first and seventh positions. For

dx.doi.org/10.1021/bi500393q | *Biochemistry* 2014, 53, 4806−4813

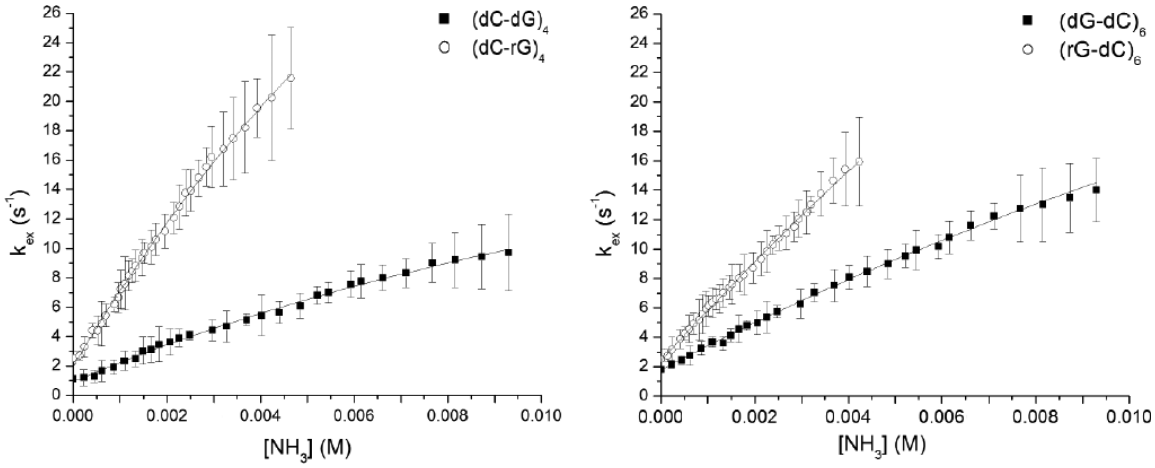


Figure 5. Dependence of the imino proton exchange rates for guanines in the CJ site on the ammonia base concentration at 40 °C. The filled squares represent the results for the DNA−DNA duplex and the empty circles the results for the RNA−DNA hybrid duplex. The rates were fit as a function of the ammonia base concentration to eq 3.

Table 1. Equilibrium Constants and Free Energy Changes for Base Pair Opening in the CJ Site of the Two Nucleic Acid Structures Investigated[a]

DNA−DNA duplex			RNA−DNA hybrid duplex			
base pair	K_{op} (×10^6)	ΔG_{op} (kcal/mol)	base pair	K_{op} (×10^6)	ΔG_{op} (kcal/mol)	$\Delta\Delta G_{op}$ (kcal/mol)
$(dT\text{-}dA)_1$	25 ± 3	6.17 ± 0.07	$(rU\text{-}dA)_1$	15.5 ± 0.05	6.45 ± 0.02	0.28 ± 0.09
$(dA\text{-}dT)_2$	8.1 ± 0.9	6.82 ± 0.06	$(rA\text{-}dT)_2$	58.6 ± 0.9	5.67 ± 0.01	−1.15 ± 0.07
$(dT\text{-}dA)_3$	11 ± 1	6.67 ± 0.06	$(rU\text{-}dA)_3$	16 ± 3	6.4 ± 0.1	−0.3 ± 0.2
$(dC\text{-}dG)_4$	1.2 ± 0.2	8.5 ± 0.1	$(rC\text{-}dG)_4$	4 ± 2	7.7 ± 0.3	−0.8 ± 0.4
$(dT\text{-}dA)_5$	30 ± 4	6.06 ± 0.07	$(rU\text{-}dA)_5$	93 ± 2	5.40 ± 0.01	−0.66 ± 0.08
$(dG\text{-}dC)_6$	1.2 ± 0.4	8.5 ± 0.2	$(rG\text{-}dC)_6$	3 ± 1	8.0 ± 0.2	−0.5 ± 0.4
$(dT\text{-}dA)_7$	14.4 ± 0.7	6.49 ± 0.03	$(rU\text{-}dA)_7$	16 ± 4	6.4 ± 0.1	−0.1 ± 0.1
$(dT\text{-}dA)_8$	27 ± 4	6.12 ± 0.08	$(rU\text{-}dA)_8$	75 ± 2	5.53 ± 0.02	−0.6 ± 0.1

[a]The data for (dT-dA), (rU-dA), and (rA-dT) base pairs were recorded at 20 °C. The data for (G-C) and (C-G) base pairs were recorded at 40 °C.

the base pair in the first position, the K_{op} equilibrium constant in the RNA−DNA hybrid is smaller than that of the same base pair in the DNA−DNA duplex. Hence, the $(rU\text{-}dA)_1$ base pair in the hybrid is more stable than the $(dT\text{-}dA)_1$ base pair in the DNA duplex by 0.28 kcal/mol. For the base pair in the seventh position, the K_{op} equilibrium constants in the two structures are the same within experimental error, indicating that the two base pairs, $(dT\text{-}dA)_7$ and $(rU\text{-}dA)_7$, have the same structural stabilities.

The stabilities of individual base pairs measured in this work provide new insights into the energetic forces that drive the conformational changes at the CJ site during transcription pausing. The changes in the RNA−DNA hybrid are schematically presented in Figure 8. Upstream of the CJ site, and within this site, the transcription hybrid encompasses ∼8 bp.[12] As the elongation complex moves 3 bases past the end of the CJ site (bottom part of Figure 8), the upstream part of the bubble collapses, leaving behind an RNA−DNA hybrid of only 3 bp.[10] Transcription may continue beyond this position despite the short RNA−DNA hybrid. As the elongation complex moves forward, the hybrid is gradually restored to its original size. This restoration is directional in that the length of the hybrid increases only on its downstream side (as indicated by the arrows in Figure 8).[10] To correlate these changes in the transcription bubble to the base pair opening results obtained in this work, we have used an energetic model for the nucleic acid framework of the elongation complex. The model analyzes the RNA−DNA and DNA−DNA duplexes studied, in the absence of RNA polymerase, and aims to identify the intrinsic energetic properties of these structures that may contribute to the collapse of the transcription bubble at the CJ site.

The energetic changes that occur in the nucleic acid framework of the elongation complex when the transcription bubble collapses are (i) the favorable free energy change for the re-formation of the DNA−DNA double helix and (ii) the unfavorable free energy cost for the disruption of the RNA−DNA hybrid. Accordingly, for an individual base pair, the energetic change during the collapse of the transcription bubble is calculated as the total free energy change that occurs when the base pair in the RNA−DNA hybrid opens and the corresponding base pair re-forms in double-helical DNA:

$$\Delta\Delta G_{op} = \Delta G_{op}(\text{RNA–DNA}) - \Delta G_{op}(\text{DNA–DNA}) \quad (9)$$

The results of the calculation are listed in Table 1 (last column) and shown in Figure 7. The model predicts that the energetic contribution of a base pair to the collapse of the transcription bubble depends on its location in the CJ site. The largest favorable contributions are made by the base pairs at positions 2, 4−6, and 8, and they range from −0.5 to −1.1 kcal/mol. For two of the base pairs in the CJ site, namely, the base pairs in the third and seventh positions, the contributions are negligible within experimental error. Moreover, the base pair in the first position of the CJ site "opposes" the collapse of the bubble by making an unfavorable, positive contribution to the transition

dx.doi.org/10.1021/bi500393q | Biochemistry 2014, 53, 4806−4813

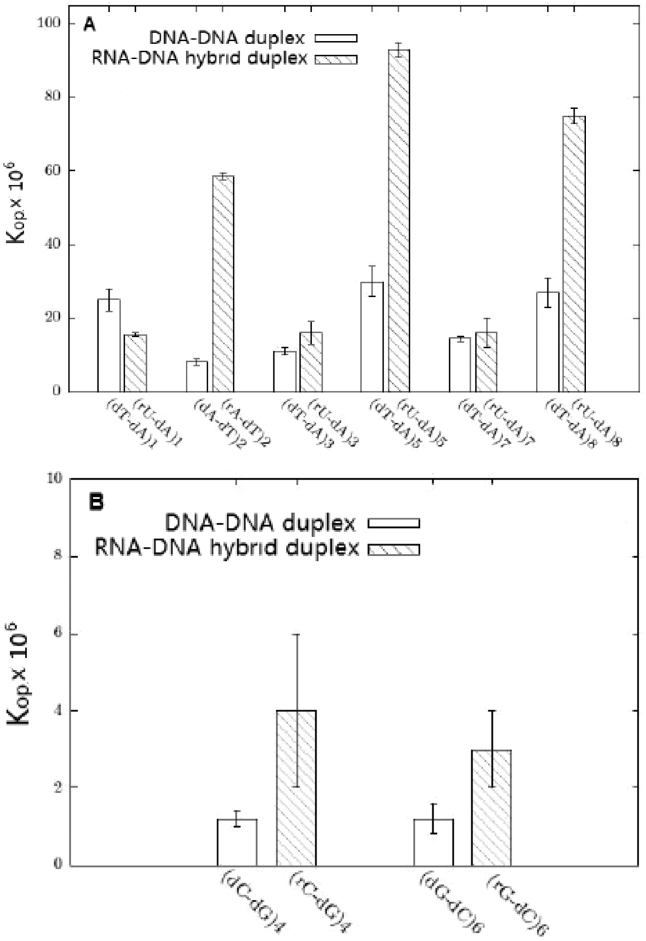


Figure 6. Equilibrium constants for opening of individual base pairs in the DNA–DNA duplex and the RNA–DNA hybrid duplex: (A) A-T/A-U base pairs at 20 °C and (B) G-C base pairs at 40 °C.

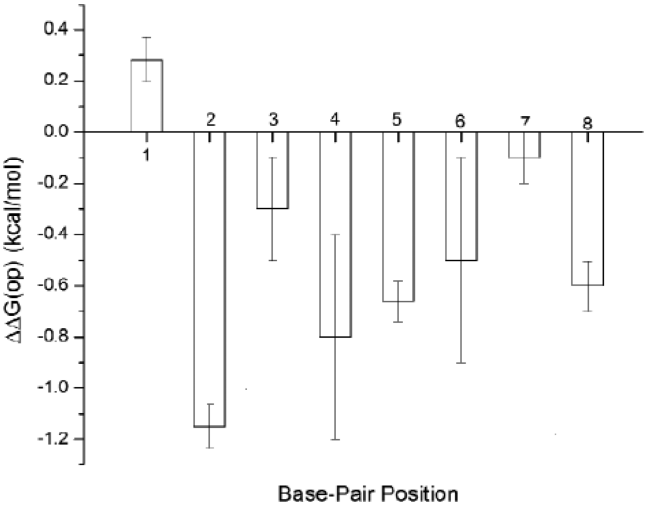


Figure 7. Free energy contributions ($\Delta\Delta G_{op}$) of individual base pairs to the collapse of the transcription bubble, predicted by our model.

from the hybrid to the DNA double helix. These findings suggest an explanation for the fact that the collapse of the transcription bubble occurs only when the elongation complex has moved downstream of the CJ sequence. When the first half of the CJ site is transcribed, the hybrid would include the (rU-dA) base pairs in the first and third positions (top, Figure 8).

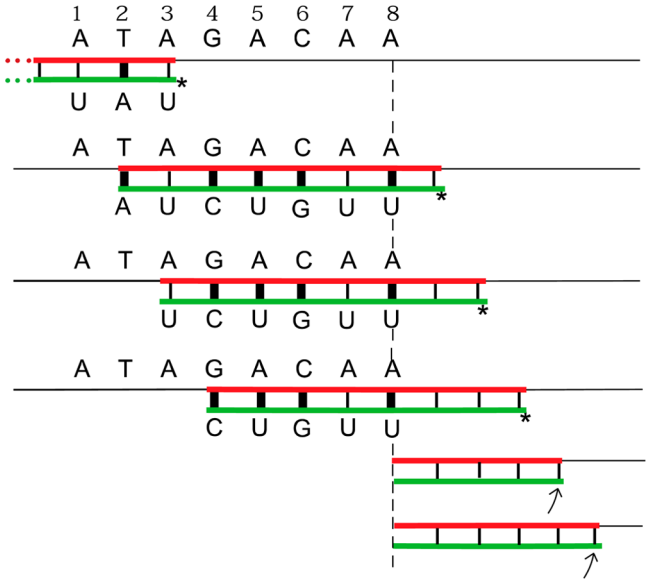


Figure 8. Schematic representation of the progression of transcription through the CJ site. Only the transcription hybrid formed between the DNA template strand (red) and the RNA (green) is shown. The vertical lines represent base pairs in this hybrid. Thicker vertical lines identify the base pairs whose stabilities in the RNA–DNA hybrid are significantly lower than those in the DNA–DNA duplex. The asterisk denotes the last base that has been transcribed; the arrow denotes the base that is going to be transcribed.

This may prevent the spontaneous collapse of the bubble because (i) the breaking of the $(rU\text{-}dA)_1$ base pair in the hybrid is unfavorable relative to the formation of the $(dT\text{-}dA)_1$ base pair in the DNA template duplex and (ii) the adenine in the third position on the DNA template strand does not have a significant energetic preference for the (dT-dA) base pair in the DNA–DNA duplex over the (rU-dA) base pair in the RNA–DNA hybrid. As the elongation complex moves past the CJ site, the transcription hybrid would incorporate more of the base pairs in the CJ site that have an energetic preference for the DNA template duplex, for example, base pairs at positions 4–6 and 8. Thus, the position where the bubble collapses, i.e., 3 bases past the end of the CJ site, may represent a configuration in which sufficient destabilization of the hybrid accumulates to trigger the switch of the DNA template strand from the hybrid to the DNA double-helical structure.

In summary, in this work, we have shown that, for the majority of base pairs in the CJ site, the stabilities in the DNA–DNA template duplex are higher than those in the transcription RNA–DNA hybrid. These distinct energetic properties favor the collapse of the transcription bubble that occurs downstream from the CJ site. A similar energetic role of the nucleic acid framework has been demonstrated recently by our laboratory for class I transcription terminators.[25] The DNA template in these terminators consists of a base sequence with inverted symmetry, which forms a hairpin in the RNA transcript, followed by a tract of adenines. Transcription of the adenine tract into a tract of uracils triggers destabilization of the elongation complex and the release of nascent RNA. We have previously shown that the central base pairs in the (rU-dA) tract of this RNA–DNA transcription hybrid are less stable than the corresponding (dT-dA) base pairs in the DNA template duplex by 3–4.7 kcal/mol. On the basis of these findings, we have proposed that the central (rU-dA)/(dT-dA)

dx.doi.org/10.1021/bi500393q | Biochemistry 2014, 53, 4806–4813

base pairs in the adenine tract contribute to transcription termination by promoting dissociation of the RNA−DNA hybrid and re-formation of the template DNA double helix. The results presented in this work suggest that a similar mechanism may be involved in the collapse of the transcription bubble during pausing of T7 polymerase at the CJ site. Thus, it appears that, in spite of the significant structural differences between class I and class II transcription terminators, the nucleic acid structures in the two classes share similar energetic properties.

■ ASSOCIATED CONTENT

Supporting Information

^{1}H−^{1}H NOESY spectra of the DNA−DNA and RNA−DNA duplexes, assignments of imino proton resonances, and rates of opening and closing (k_{op} and k_{cl}, respectively) for base pairs in the two nucleic acid duplexes investigated. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: irussu@wesleyan.edu. Phone: (860) 685-2428. Fax: (860) 685-2211.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

CJ, concatemer junction; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy.

■ REFERENCES

(1) Carlson, K. (1968) Intracellular fate of deoxyribonucleic acid from T7 bacteriophages. *J. Virol. 2*, 1230−1233.
(2) Watson, J. (1972) Origin of concatemeric T7 DNA. *Nat. New Biol. 239*, 197−201.
(3) Lyakhov, D. L., He, B., Zhang, X., Studier, F. W., Dunn, J. J., and McAllister, W. T. (1998) Pausing and termination by bacteriophage T7 RNA polymerase. *J. Mol. Biol. 280*, 201−213.
(4) Zhang, X., and Studier, F. W. (2004) Multiple roles of T7 RNA polymerase and T7 lysozyme during bacteriophage T7 infection. *J. Mol. Biol. 340*, 707−730.
(5) He, B., Kukarin, A., Temiakov, D., Chin-Bow, S. T., Lyakhov, D., Rong, M., Durbin, R. K., and McAllister, W. T. (1998) Characterization of an unusual, sequence-specific termination signal for T7 RNA polymerase. *J. Biol. Chem. 273*, 18802−18811.
(6) Mead, D. A., Skorupa, E. S., and Kemper, B. (1986) Single-stranded DNA 'blue' promoter plasmids: A versatile tandem promoter system for cloning and protein engineering. *Protein Eng. 1*, 67−74.
(7) Macdonald, L. E., Durbin, R. K., Dunn, J. J., and McAllister, W. T. (1994) Characterization of two types of termination signal for bacteriophage T7 RNA polymerase. *J. Mol. Biol. 238*, 145−158.
(8) Whelan, S. P. J., Ball, A., Barr, J. N., and Wertz, G. T. W. (1995) Efficient recovery of infectious vesicular stomatitis virus entirely from cDNA clones. *Proc. Natl. Acad. Sci. U.S.A. 92*, 8388−8392.
(9) Hartvig, L., and Christiansen, J. (1996) Intrinsic termination of T7 RNA polymerase mediated by either RNA or DNA. *EMBO J. 15*, 4767−4774.
(10) Sohn, Y., and Kang, C. (2005) Sequential multiple functions of the conserved sequence in sequence-specific termination by T7 RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A. 102*, 75−80.
(11) Nayak, D., Siller, S., Guo, Q., and Sousa, R. (2008) Mechanism of T7 RNAP pausing and termination at the T7 concatemer junction: A local change in transcription bubble structure drives a large change in transcription complex architecture. *J. Mol. Biol. 376*, 541−553.
(12) Yin, Y. W., and Steitz, T. A. (2002) Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. *Science 298*, 1387−1395.
(13) Mukherjee, S., Brieba, L. G., and Sousa, R. (2003) Discontinuous movement and conformational change during pausing and termination by T7 polymerase. *EMBO J. 22*, 6483−6493.
(14) Kwon, Y.-S., and Kang, C. (1999) Bipartite modular structure of intrinsic, RNA hairpin-independent termination signal for phage RNA polymerase. *J. Biol. Chem. 274*, 29149−29155.
(15) Cantor, C. R., Warshaw, M. M., and Shapiro, H. (1970) Oligonucleotide Interactions. III. Circular Dichroism Studies of the Conformation of Deoxyoligonucleotides. *Biopolymers 9*, 1059−1077.
(16) Chen, C., and Russu, I. M. (2004) Sequence-Dependence of the Energetics of Opening of AT Base Pairs in DNA. *Biophys. J. 87*, 2545−2551.
(17) Plateau, P., and Gueron, M. (1982) Exchangeable proton NMR without base-line distortion, using new strong-pulse sequences. *J. Am. Chem. Soc. 104*, 7310−7311.
(18) Lippens, G., Dhalluin, C., and Wieruszeski, J.-M. (1995) Use of a water flip-back pulse in the homonuclear NOESY experiment. *J. Biomol. NMR 5*, 327−331.
(19) Russu, I. M. (2004) Probing Site-Specific Energetics in Proteins and Nucleic Acids by Hydrogen Exchange and NMR Spectroscopy. *Methods Enzymol. 379*, 152−175.
(20) Englander, S. W., and Kallenbach, N. R. (1984) Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys. 16*, 521−655.
(21) Gueron, M., and Leroy, J.-L. (1995) Studies of Base Pair Kinetics by NMR Measurement of Proton Exchange. *Methods Enzymol. 261*, 383−413.
(22) Benight, A. S., Schurr, J. M., Flynn, P. F., Reid, B. R., and Wemmer, D. E. (1988) Melting of a Self-complementary DNA Minicircle. Comparison of Optical Melting Theory with Exchange Broadening of the Nuclear Magnetic Resonance Spectrum. *J. Mol. Biol. 200*, 377−399.
(23) Ts'o, P. O. P. (1974) *Basic Principles in Nucleic Acid Chemistry*, Vol. *I*, Academic Press, New York.
(24) Weast, R. C. (1987) *CRC Handbook of Chemistry and Physics*, 67th ed., CRC Press, Boca Raton, FL.
(25) Huang, Y., Weng, X., and Russu, I. M. (2010) Structural Energetics of the Adenine Tract from an Intrinsic Transcription Terminator. *J. Mol. Biol. 397*, 677−688.