nature structural & molecular biology

Dynamic opening of DNA during the enzymatic search for a damaged base

Chunyang Cao[1], Yu Lin Jiang[1], James T Stivers[1] & Fenhong Song[2]

Uracil DNA glycosylase (UDG) removes uracil from U•A or U•G base pairs in genomic DNA by extruding the aberrant uracil from the DNA base stack. A question in enzymatic DNA repair is whether UDG and related glycosylases also use an extrahelical recognition mechanism to inspect the integrity of undamaged base pairs. Using NMR imino proton exchange measurements we find that UDG substantially increases the equilibrium constant for opening of T-A base pairs by almost two orders of magnitude relative to free B-DNA. This increase is brought about by enzymatic stabilization of an open state of the base pair without increasing the rate constant for spontaneous base pair opening. These findings indicate a passive search mechanism in which UDG uses the spontaneous opening dynamics of DNA to inspect normal base pairs in a rapid genome-wide search for uracil in DNA.

A long-standing mystery in enzymatic DNA repair is how enzymes efficiently locate and excise rare damaged bases in DNA before the genomic integrity is irreversibly altered by misguided DNA replication. Such reactions are carried out by a family of DNA glycosylase enzymes that cleave the *N*-glycosidic bond connecting the deoxyribose sugar to the damaged base[1]. Because damaged bases may be present in the human genome at a frequency of less than one in ten million normal base pairs, these enzymes are, by thermodynamic default, nearly always bound nonspecifically to undamaged DNA. Given the overwhelming prevalence of these nonspecific complexes, it is of importance to ask if they exhibit structural or dynamic features that are relevant to the location of damaged sites in DNA. Such questions require new methods for addressing the dynamic nature of undamaged DNA bound to these enzymes.

Some insight into the basis of damaged base recognition by DNA glycosylases has been provided by structural and kinetic approaches. Structural studies have revealed an evolutionarily conserved conformational transformation known as 'base flipping' involving a 180° rotation of the damaged base and sugar from the DNA base stack into the enzyme active site[1,2]. Unifying features of flipped-base structures are enzyme-induced DNA bending by phosphodiester backbone interactions[3–5], aromatic stacking and hydrogen bonding interactions with specific features of damaged bases[6–8], and enzyme side chain intercalation into the DNA duplex to fill the void left by the departed base[9–11]. Although these static structures provide satisfying explanations for how enzymes recognize the specific features of damaged bases, they provide little insight into how the process of base flipping is initiated, and whether normal base pairs can also be opened to ascertain their integrity. In this regard, rapid kinetic studies have detected a nonspecific enzyme-DNA intermediate on the pathway for base flipping by the enzyme uracil DNA glycosylase, a powerful enzyme that removes unwanted uracil bases from DNA[12,13]. This intermediate, which is kinetically and thermodynamically indistinguishable from the complex of UDG with undamaged DNA[13], is characterized by an intrahelical uracil base and weak nonspecific interactions with the enzyme[2]. The detection of this intermediate on the uracil flipping pathway raises the question of whether UDG can promote DNA base pair opening before any damage-specific interactions are formed, and the role of such nonspecific interactions in uracil flipping and recognition.

In the absence of direct experimental measurements that elucidate the dynamics of very early intermediates on the reaction trajectory for base flipping, two general mechanisms have been put forward, termed passive and active (**Fig. 1**). The passive mechanism envisions that enzymes identify damaged bases when they are exposed extrahelically owing to the spontaneous breathing dynamics of DNA base pairs[13–15]. In this model, the intrinsic instability of the damaged base pair facilitates specific recognition because the damaged base, owing to its weakened interaction within the duplex, spends more time in a productive extrahelical conformation than a normal base. Thus, the key distinguishing feature of the passive mechanism is that the enzyme does not increase the rate constant for base pair opening as compared with the free DNA ($k_{op}^{free} = k_{op}^{enz}$; **Fig. 1**). The passive mechanism has two possible pathways depending on whether the enzyme encounters the extrahelical base in a bimolecular encounter event (passive-bi, lower pathway in **Fig. 1**), or when it is already preassociated at the site, and flipping occurs within a unimolecular complex (passive-uni, upper pathway in **Fig. 1**). Which of these two pathways is followed depends on the rate constant for base pair closing, or equivalently, the lifetime of the extrahelical base (**Fig. 1**)[16]. The distinguishing feature of the active mechanism is that the enzyme increases the rate constant for base pair opening by forming interactions with the DNA (upper red pathway, where $k_{op}^{enz} >> k_{op}^{free}$)[15]. Although

[1]Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, USA. [2]Center for Advanced Research in Biotechnology of the National Institute of Standards and Technology and the University of Maryland Biotechnology Institutes, 9600 Gudelsky Drive, Rockville, Maryland 20850, USA.

Published online 21 November 2004; doi:10.1038/nsmb864

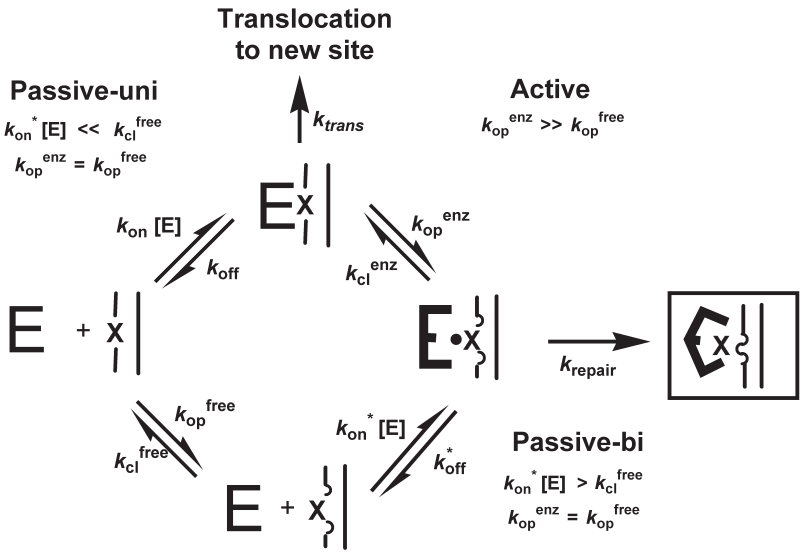

Figure 1 Active and two passive mechanisms for base flipping. Two general pathways are shown for the formation of a complex (boxed) between an enzyme and an extrahelical base (X). The lower pathway (passive-bi) involves bimolecular encounter of the enzyme with the base that has been exposed extrahelically owing to passive dynamic excursions of the base from the DNA base stack without any assistance from the enzyme. The upper pathway involves preassociation of the enzyme with the site while the base is still intrahelical, followed by base flipping. For the upper pathway, if the enzyme actively participates in accelerating the rate of base pair opening ($k_{op}^{enz} >> k_{op}^{free}$), the mechanism is termed active. If the enzyme plays no role in facilitating base pair opening ($k_{op}^{enz} = k_{op}^{free}$), the mechanism is termed passive-uni, because the spontaneous opening event occurs within a unimolecular complex. The kinetic conditions favoring each mechanism are shown (see text). The rate constants pertain to the following events: base pair opening and closing in the free DNA (k_{op}^{free} and k_{cl}^{free}) and in the enzyme complex (k_{op}^{enz} and k_{cl}^{enz}), enzyme association with the intrahelical site (k_{on} [E] and k_{off}) and extrahelical base (k_{on}^{*} [E] and k_{off}^{*}), and the rate of site repair if X is a damaged base (k_{repair}). For completeness, a kinetic step for enzyme translocation along the DNA is also shown (k_{trans}, see Discussion). For UDG, $k_{on} \approx 10^8$ M^{-1} s^{-1}; $k_{off} \approx 1{,}000$ s^{-1}.

a variation of the passive and active mechanisms has been proposed that involves processive flipping of normal bases as the enzyme translocates along the DNA[17,18], the pathway that is followed will ultimately be determined by two factors: how fast the base emerges from the duplex, and its lifetime in the extrahelical state (see **Fig. 1** legend). Although there has been much speculation that DNA dynamics plays a role in distinguishing between damaged and undamaged sites in DNA[19,20], no direct experimental tests have been developed. This requires the detection of a small population of high-energy open states concealed in a much larger population of stable closed states.

Here we report the use of NMR imino proton exchange experiments to measure the dynamic opening of normal DNA base pairs in the presence and absence of the DNA repair enzyme UDG, which locates and hydrolytically excises uracil bases from U•G or U•A base pairs in genomic DNA[1]. This approach allows detection of rapid DNA opening events that result in open states populated at a frequency of <1 part in 10^5, yet are likely to play an important role in the enzymatic search for genomic alterations. These kinetic measurements provide the information required to distinguish between active and passive mechanisms for opening of normal base pairs by UDG.

RESULTS

Nonspecific and specific UDG–DNA complexes

The earliest step in the base flipping pathway is the formation of a nonspecific encounter complex between the enzyme and DNA where the enzyme interacts weakly using nonspecific electrostatic interactions[12,13]. Structural studies of nonspecific complexes are intrinsically difficult because the enzyme may bind in several different registers along the DNA. However, if the number of binding modes can be minimized, it is feasible to measure average properties of the bound DNA molecule. To minimize the number of binding modes of UDG to duplex DNA, we used two short palindromic sequences that differed by a central T-A or G-C base pair (**Fig. 2**). The T-A and G-C duplexes allow the investigation of the effect of UDG binding on the opening dynamics of undamaged base pairs within an identical sequence context, and provide a useful model system for understanding the earliest step in the process of base flipping. UDG interacts equivalently with both duplexes because the binding dissociation constants are equal (K_d = 6.7 ± 0.2 μM at 10 °C) (see **Supplementary Fig. 1** online).

The imino proton spectra of the T-A and G-C duplexes show excellent chemical shift dispersion in the free and bound states, and the absence of any overlap from proton resonances of the enzyme (**Fig. 2a**,**b**). The imino proton resonances for both duplexes show similar chemical shifts and intensities for the free and bound forms, indicating that the enzyme does not induce a large distortion in the DNA leading to disrupted base-pairing. Although it is likely that UDG binds the T-A and G-C duplexes in several loose binding modes that are thermodynamically biased toward the center of the duplex, the observation of one chemical shift for each imino proton indicates that all of the binding modes are in rapid exchange on the chemical shift timescale. The presence of multiple bound complexes that are in dynamic equilibrium implies that any observed NMR property of the enzyme-bound DNA will be a population-weighted average. That is, any measured property provides a lower limit to its true magnitude.

In strong contrast with these observations, an identical DNA sequence in which the central T was replaced with a nonhydrolyzable deoxyuridine analog (2′-β-fluoro-2′-deoxyuridine, 2′-FU) shows extensive changes in the imino proton shifts and line broadening upon binding of the catalytically inactive D64N UDG mutant (**Fig. 2c**). Owing to the extensive exchange line broadening in this complex, we could not assign the residual imino peaks in the spectrum using the NOESY method (**Fig. 2c**). Nevertheless, these NMR results with the U•A duplex indicate that substantial DNA conformational changes occur upon formation of the specific complex, with disruption of the uracil Watson-Crick hydrogen bonds, as would be expected from an enzyme that flips the uracil base into its active site through several enzyme-bound intermediate states[4,12]. Under such conditions, the imino proton exchange methodology is not useful for investigating base pair dynamics.

UDG accelerates T-A imino exchange

The similar imino proton chemical shifts of the free and enzyme-bound DNA in **Figure 2** indicate that if enzyme-induced opening of normal T-A and G-C base pairs is occurring, the process is highly dynamic, with the equilibrium strongly biased toward the closed state. To explore whether dynamic opening of normal base pairs was induced by UDG

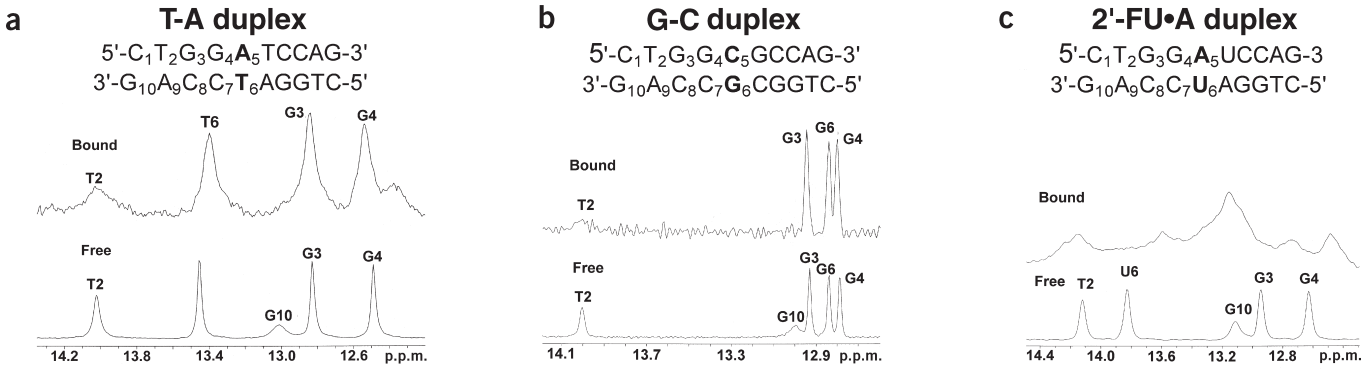

Figure 2 Imino proton spectra and sequences of DNA duplexes used in these studies (pH 8.0, 10 °C). (**a**–**c**) NMR spectrum (500 MHz) of the free and bound DNA duplex containing a central T-A base pair (**a**), a central G-C base pair (**b**), and a central 2′-FU•A base pair (**c**). The imino proton assignments were made using a standard ^{1}H-^{1}H NOESY experiment. The free DNA samples contained ~1 mM DNA duplex and the samples of the complex contained 0.8 mM DNA and 1 mM UDG.

binding, we carried out NMR hydrogen exchange experiments using magnetization transfer from water. In these experiments, which have been used previously to probe base pair opening in free duplex DNA, water magnetization is selectively inverted and the rate of exchange of these magnetically labeled protons into the imino proton sites of the DNA bases is monitored[21]. We first measured the imino proton exchange rates for the free and enzyme-bound T-A duplex in the absence of added buffer catalyst, and observed that UDG binding produced distinctive effects on the exchange rates for T-A and G-C base pairs (**Fig. 3a**,**b**). Most notably, the resonance assigned to the central T6 base was found to exchange 25-fold faster in the complex than in the free DNA (**Fig. 3a**), and the other thymine imino resonance near the end of the DNA sequence (T2) showed a five-fold increase in its exchange rate upon UDG binding (**Fig. 3b**). In contrast, the exchange rate for the imino resonance assigned to G3 was similar in the free and bound DNA, whereas the imino proton of the other observed guanine base (G4) showed a four-fold increase in its exchange rate in the complex. The increased exchange rate for G4 may be explained by a direct effect or by a local perturbation of the duplex when UDG interacts with the adjacent T6 base pair. An indirect effect on the exchange rate of G4 is supported by the experiments presented below.

To test whether the large increase in the exchange rate for the T6 imino proton was due to the specific nature of the T-A base pair, we then examined the exchange behavior of the G-C duplex in which T6 is replaced with guanine (**Fig. 3c**). In stark contrast with T6, the G6 imino resonance showed no increase in its exchange rate upon UDG binding, and the other imino resonances of the G-C duplex showed similar small changes upon formation of complex as the T-A duplex (**Fig. 3c**). The most straightforward explanation for these results is that UDG preferentially opens T-A base pairs, but adjacent G-C base pairs are also affected by proximity to the opened T-A pair. It is possible that T-A base pairs are opened more readily than G-C base pairs because of the weaker T-A hydrogen bonds[22], increased flexibility at T-A sites[19], or alternatively, because the enzyme does not have the means to stabilize an unpaired guanine because of its steric properties. Previous studies have shown that UDG can cleave the glycosidic bond of thymidine and cytosine when the active site of the enzyme has been altered to accommodate the hydrogen bonding and steric features of these bases[7,23]. Thus, the dynamic excursions of thymidine detected here are likely to be on the productive pathway for base flipping.

Imino proton exchange catalysis

Does UDG enhance the exchange rate for T-A base pairs by increasing the opening rate (k_{op}) of base pairs, or by decreasing the closing rate of the open state (k_{cl})? To address this question, we examined the imino proton exchange rates of the free and enzyme-bound T-A and G-C duplexes at increasing concentrations of difluoroethylamine (DFEA) base catalyst at pH 8.0 and 10 °C (**Fig. 4**). In the two-step exchange mechanism shown in **Figure 4a**, the base pair imino proton (NH) is depicted in two states:

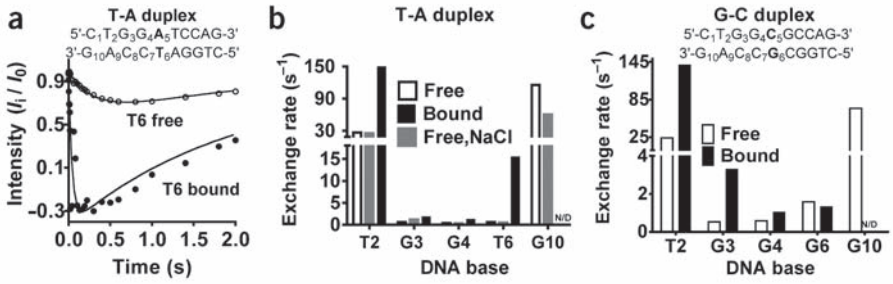

Figure 3 Water magnetization transfer kinetics at pH 8.0 and 10 °C in the absence of DFEA buffer catalyst (see **Supplementary Methods** online). (**a**) Intensity of the imino proton resonance of residue T6 of the T-A duplex as a function of transfer delay time. The data for the free duplex and its complex with UDG are shown. The rapid initial decrease in intensity reflects the exchange of inverted water protons into the H3 imino site, and the slower increase in intensity arises from longitudinal relaxation of the water and imino proton magnetization. The data were fitted to the equation $I_z\,(t_{mix}) / I_{z,eq} = 1 + E\,k_{ex}\,[\exp(-R_{1i}\,t_{mix}) - \exp(-R_{1w}\,t_{mix}) / (R_{1w} - R_{1i})]$. In this expression, R_{1w} is the water longitudinal relaxation time, R_{1i} is the sum of the imino proton longitudinal relaxation rate and k_{ex}, the chemical exchange rate, and $I_z\,(t_{mix})$ and $I_{z,eq}$ are the intensities of the imino peaks at a given value of t_{mix} and at equilibrium, respectively. E is the efficiency of water inversion, which is equal to –2 for an inversion efficiency of 100%. This fitting procedure gives k_{ex}^{free} (T6) = 0.63 s^{-1}, and k_{ex}^{bound} (T6) = 15.4 s^{-1}. (**b**) Exchange rate constants for each observed imino proton of the T-A duplex in the free and bound state. The gray bars show the exchange rate constants for the free DNA in the presence of 1.035 M NaCl, which differ little from the rates measured in the presence of 35 mM NaCl (open bars). Thus, changes in k_{ex} arising from ionic strength effects are negligible[12]. (**c**) Exchange rate constants for each observed imino proton of the G-C duplex in the free and bound state in the absence of added buffer catalyst.

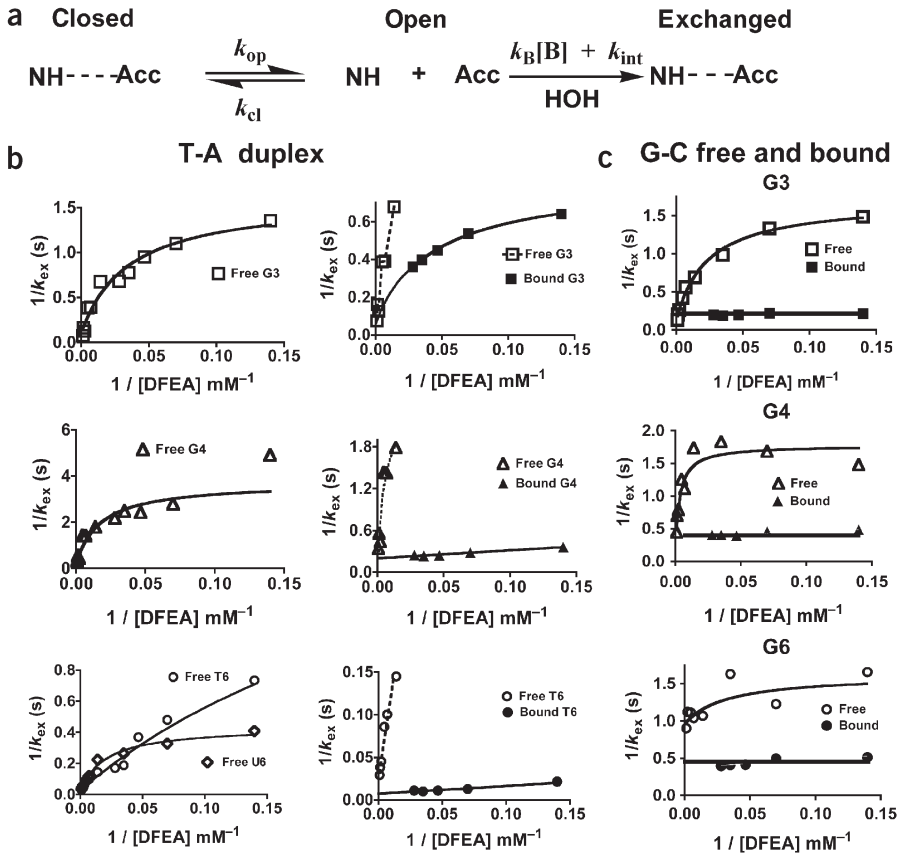

Figure 4 Imino proton exchange mechanism and double-reciprocal plots of the observed exchange rates against DFEA catalyst concentration (see equation (1)). (**a**) Two-step imino proton (NH) exchange mechanism (see text for definitions). (**b**) Imino proton exchange catalysis for residues G3, G4 and T6 of the T-A duplex in the free state (left column), and UDG-bound state (middle column). The exchange data for the U6 residue of the free U•A duplex is shown on the same plot with that of residue T6 of the T-A duplex. The exchange data for the free DNA all show downward curvature that is indicative of a change in rate-limiting step from general base catalysis by added DFEA (k_B [B], equation (1)) to intramolecular catalysis by N1 and N3 of the A and C bases (k_{int}, equation (1))[24]. For comparison, the curves and data for the free T-A DNA are expanded and superimposed on the graphs for the complex (dashed lines). For the bound T-A duplex, only residue G3 had a substantial k_{int} term, which was identical to the value in the free DNA. The curves were obtained from a nonlinear least-squares fit to equation (1), and the fitted parameters (k_{op}, k_{cl}, αK_{op} and k_{int}) are reported in **Supplementary Table 1** online. (**c**) Imino proton exchange catalysis for residues G3, G4 and G6 of the G-C duplex in the free and bound states. The bound G-C duplex showed no buffer catalysis, and thus, only k_{op}^{enz} was obtained from the data (see **Supplementary Table 3** online).

a closed state where it is hydrogen-bonded to its Watson-Crick acceptor (Acc), and an open state where it may exchange with solvent protons (red) in a step that is accelerated by increasing concentrations of a buffer catalyst ([B]). Thus, at high buffer concentrations the observed exchange rate (k_{ex}) will reach a maximal value limited by the opening rate (k_{op}), and at low buffer concentrations k_{ex} approaches a finite rate ($k_o = k_{op}\, k_{int} / k_{cl}$) that reflects the buffer independent exchange rate. This finite rate is facilitated by intramolecular general base catalysis from the N1 and N3 nitrogens of adenine and cytosine, and this process is represented by the k_{int} term in equation (1) (ref. 22):

$$1 / k_{ex} = 1 / k_{op} + \{(k_{cl} / \alpha) / [\, k_{op}\, (k_B\, [B] + k_{int})]\} \qquad (1)$$

The α term in equation (1) is a constant that takes into account that the imino proton may not be as accessible to the base catalyst in a macromolecule as in the free nucleoside ($\alpha < 1$) (ref. 21).

At intermediate concentrations of DFEA catalyst, the exchange rate depends strongly on catalyst concentration. Using this approach, the rate constants for base pair opening and closing (k_{op} and k_{cl}) and the apparent equilibrium constant for base pair opening ($\alpha K_{op} = k_{op} / k_{cl}$) can be determined by fitting the DFEA concentration-dependence of the exchange rates to equation (1).

UDG stabilizes an open state of DNA

Two main observations can be made from inspection of the buffer-catalyzed exchange data for the observable imino resonances of the T-A duplex in the free and bound states (**Fig. 4b**). First, the exchange data for the free DNA all show downward curvature indicative of a change in mechanism from general base catalysis by added DFEA (k_B [B], equation (1)) to intramolecular catalysis by N1 and N3 of adenine and cytosine bases ($k_o = k_{op}\, k_{int} / k_{cl}$, equation (1)) (ref. 24). Second, the extrapolated exchange times at infinite buffer concentration ($1 / k_{op}$, equation (1)) are nearly identical for free and bound DNA (compare gray and solid curves in the middle panels of **Fig. 4b**), indicating that the increases in exchange rate brought about by UDG binding do not arise from substantial increases in the opening rate constants (k_{op}). Owing to inhibition of UDG activity at high DFEA concentrations (see Methods), only a limited catalyst concentration range was accessible for the imino exchange study of the UDG–DNA complexes. Therefore, it is impossible to exclude the possibility that an additional mode of base pair opening might be detectable for the complex if higher catalyst concentrations could be used[25]. Nevertheless, if another mode does exist, the conclusions obtained here from comparisons of the enzyme-bound and free DNA exchange rates are still valid.

The relative values for k_{op}, k_{cl} and αK_{op} for the free and enzyme-bound T-A duplex reveal the nature of the extrahelical recognition mechanism used by UDG (**Fig. 5**), and the complete data are reported as **Supplementary Table 1** online. Although the kinetic and equilibrium exchange parameters for the free T-A DNA fell within the range of values observed previously for B-DNA[21], comparison of αK_{op} for the free DNA with that of the complex reveals that UDG increases the apparent opening equilibrium for the T6 and G4 bases by 75- and 280-fold, respectively, whereas the other observed imino resonance (G3) shows only a small 2-fold increase. A key finding is that the apparent closing rates for T6 and the adjacent G4 are identical in the complex ($k_{cl}^{enz} = 90{,}000\ s^{-1}$), even though G4 exchanges 25-fold more slowly than T6 in the free DNA. Thus a localized enzymatic stabilization of the DNA that holds both of these bases in a solvent-exposed state is indicated. We conclude that UDG acts as a kinetic trap of the open state of the T-A base pair, but does not substantially change the rate of base pair opening

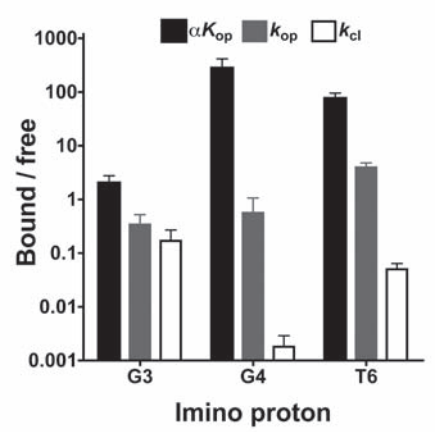

Figure 5 Relative imino proton exchange parameters for the free and UDG-bound T-A duplex. Owing to rapid imino exchange of T2 and G10 as the DFEA concentration was increased, magnetization transfer measurements were possible for only the three internal imino resonances in the complex (G3, G4 and T6). The parameters correspond to the opening equilibrium (αK_{op}), and the opening and closing rates of the individual base pairs (k_{op} and k_{cl}) as shown in equation (1).

(that is, UDG increases αK_{op} by reducing k_{cl}). Similar opening dynamics were observed for the U6 site in the free U•A duplex as compared with T6 (**Fig. 4a** and **Supplementary Table 2** online). Although it was not possible to collect exchange data for the bound U•A duplex owing to complete uracil flipping and the resultant DNA distortion (**Fig. 2c**), the observation of similar exchange dynamics for both U6 and T6 in the free DNA indicates that similar intrinsic opening dynamics prevail when UDG encounters these structurally similar base pairs. Given the strong kinetic evidence for a nonspecific encounter complex on the uracil flipping pathway in which the uracil is still within the base stack[12,13], it may be inferred that the nonspecific T-A complex is a structural and dynamic mimic of this transient intermediate. The structural basis for why thymine is arrested at an early stage of the base flipping process is not known, but investigation of mutational effects on the exchange kinetics should ultimately provide insights into this question.

As anticipated from the initial studies in the absence of buffer catalyst (**Fig. 3c**), the observable imino protons of the G-C duplex showed only small exchange rate enhancements upon formation of complex (**Fig. 4c**). The largest enhancement (more than five-fold) was for G3 adjacent to T2, which may reflect a 'thymine bystander' effect as observed above for G3 and G4 of the T-A duplex. The remaining two observable residues of the G-C duplex (G4 and G6) show less than three-fold exchange rate enhancements upon UDG binding. Notably, all of the imino proton exchange rates for the bound G-C duplex were independent of catalyst concentration and indistinguishable from the k_{op} values measured for the free G-C duplex (**Fig. 4c** and **Supplementary Table 3** online). These results further substantiate that UDG does not increase the spontaneous opening rate of normal base pairs, especially for position G6. Moreover, for these more stable G-C base pairs in the presence of UDG, the overall rate-limiting step for exchange is base pair opening at all catalyst concentrations.

DFEA catalysis is not due to a medium effect

Because imino proton exchange studies unavoidably involve a change in ionic strength as the catalyst concentration is increased, it is important to confirm that the changes in exchange rate are due to general base catalysis and not structural changes in the free DNA or the enzyme–DNA complex arising from changes in solution conditions. To eliminate these possible alternative interpretations, we carried out three control experiments. In the first experiment, the dependence of k_{ex} on the concentration of NaCl was determined for the free T-A duplex, and the exchange rates at 1,035 mM NaCl are plotted as gray bars (**Fig. 3b**). These rates are almost identical to those measured using 35 mM NaCl, indicating that ionic strength effects on the exchange rates are insubstantial for the free DNA. To establish that general base catalysis by the basic form of DFEA gave rise to the buffer-dependent exchange rates, we carried out exchange rate measurements for the bound T-A duplex in the presence of 0–76 mM trifluoroethylamine (TFEA, $pK_a = 6.3$ at 10 °C). TFEA is a close structural congener of DFEA, but it is a much weaker base that cannot efficiently catalyze imino proton transfer to solvent (see equation (3) in Methods). We found that TFEA provided no substantial increase in the imino proton exchange rates in this concentration range (<20% change in k_{ex} at pH 8.0, data not shown). This result contrasts considerably with DFEA, where 100–700% increases in k_{ex} for the complex were observed as the concentration of the basic form of DFEA was increased from 0 to 36 mM. Finally, in further support of general base catalysis by DFEA, the exchange rates for the complex were independent of pH in the pH range 7–8.6 in the absence of DFEA, but decreased by ninefold as the pH was lowered from 8 to 7 in the presence of 50 mM total [DFEA]. This effect can only be reasonably attributed to protonation of the reactive basic form of the DFEA catalyst (data not shown). These results strongly indicate that DFEA operates as a general base catalyst to accelerate the proton transfer step, as indicated in equation (1).

DISCUSSION

Assessing active and passive flipping mechanisms

These imino proton exchange studies provide the relevant data to evaluate the energetic and kinetic feasibility of active and passive mechanisms for extrahelical inspection of normal T-A base pairs by UDG (**Fig. 1**, X = T). By definition, an active mechanism may be excluded because UDG does not substantially increase the rate of base pair opening ($k_o^{free} \approx k_{op}^{enz}$, **Fig. 1**). In addition, a passive-bi mechanism in which UDG binds in a bimolecular encounter event to an already extrahelical thymidine is rendered highly improbable because the lifetime of the extrahelical base is so short (50 ns) that there is not enough time for random diffusion to bring the enzyme and DNA together (see **Supplementary Table 1** online). Another more quantitative way of stating the same conclusion is that the passive-bi mechanism requires an unrealistically fast association rate constant of $k_{on}^{*} = 10^{10}\ M^{-1}\ s^{-1}$ to satisfy the thermodynamic box defined by the K_d for nonspecific DNA, and the measured values of k_{op}^{free}, k_{cl}^{free} and k_{off} (refs. 12,13) in **Figure 1**. Likewise, a highly processive flipping mechanism involving sequential flipping of bases as the enzyme translocates along the DNA is rendered improbable because T-A base pair opening is slower than enzyme dissociation from a nonspecific site ($k_{off} \approx 1{,}000\ s^{-1} > k_{op}^{enz}$) (ref. 13), and for a processive mechanism to be operative, k_{trans} must be even faster than k_{off}[1,18]. Thus, the most probable mechanism for extrahelical inspection of T-A base pairs is the passive-uni mechanism, in which UDG traps the thymidine base as it spontaneously emerges from the base stack when the enzyme is already preassociated at the site.

A passive-uni mechanism for uracil flipping

Are these findings with thymidine relevant to the uracil flipping mechanism? Strong support for such an extrapolation is provided by the finding that both human and *Escherichia coli* UDG are catalytically active against thymidine when their active sites have been enlarged to accommodate the 5-methyl group[23,26]. Such activity would almost certainly require flipping into the UDG active site, most likely by a similar path-

way as followed by uracil. Thus these imino proton exchange findings reflect the dynamic nature of the highly transient nonspecific complex between UDG and uracil containing-DNA that has been detected previously using rapid kinetic approaches[12,13,27]. According to this model, when the enzyme initially encounters a uracil site in the correct register for base flipping, the spontaneous dynamic excursions of the uracil base from the DNA base stack are used to initiate the flipping process, and generate a high-energy intermediate as detected for the thymine base observed here. UDG then interacts weakly and nonspecifically with the DNA backbone adjacent to this nascently unpaired uracil, giving it a sufficient lifetime to allow dynamic motions of the enzyme to guide the base further along the flipping pathway.

What are the probable structural features of the high-energy open complex detected by NMR? Computational studies have estimated that the minimum rotation of a cytosine or guanine base that is required to allow hydrogen exchange with solvent is in the range of 35–45° toward the major groove[20]. This angle is only 20–25% of the entire 180° rotation required to completely flip a uracil base into the UDG active site. Thus, the motions observed here probably reflect very early movement along the reaction coordinate for base flipping[4], indicating a very unstable open intermediate. Because DNA backbone distortion has been implicated as an important aspect of base flipping[20], we speculate that UDG stabilizes the open state by nonspecific phosphodiester interactions leading to dynamic bending of the DNA backbone at the T-A base pair[3,5]. Interaction of UDG at the T-A pair may be favored because of its weaker hydrogen bonding or stacking in the duplex[28], which would be expected to increase its flexibility as compared with G-C pairs[19]. Consistent with this idea, UDG accelerates imino proton exchange of thymine, but not guanine, suggesting that UDG takes advantage of increased dynamic flexibility of T-A and U•A base pairs in its search for uracil in DNA. These ideas are consistent with recent findings that UDG binding to uracil sites is markedly enhanced when the partner base lacks hydrogen bond donor and acceptor groups[15].

The approach described here demonstrates how an old tool in NMR spectroscopy can be used to uncover hidden aspects of enzyme-DNA recognition. Although many factors, including DNA sequence, may influence how UDG recognizes and stabilizes extrahelical bases, a useful experimental tool now exists to discover the functional groups of the enzyme that are involved in extrahelical detection of damaged and normal bases.

METHODS

DNA duplexes and enzymes. The decamer T-A duplex (5′-CTGGATCCAG-3′) and G-C duplex (5′-CTGGCGCCAG-3′) were purchased from Integrated DNA Technologies. The decamer 2′-fluoro-2′-deoxyuridine DNA (5′-CTGGAUCCAG-3′) was synthesized in our laboratory as described[13]. They were purified by anion-exchange HPLC (Zorbax) and desalted using a C-18 reverse-phase HPLC column (Phenomenex Aqua column). The purity of the oligonucleotides was assessed by matrix-assisted laser desorption mass spectroscopy and denaturing PAGE. Because the DNA duplexes were self-complementary, the duplex extinction coefficients were determined by obtaining UV absorption measurements of the duplex stock solutions, which were then digested to nucleosides using nuclease P1 and calf intestinal alkaline phosphatase, followed by separation of the nucleosides using reverse-phase HPLC. Employing calibration standards for each nucleoside, and the known stoichiometry of each nucleoside in the duplex, the HPLC peak areas were used to determine the concentration of the duplex stock solutions[29]. The extinction coefficients of the T-A and G-C duplexes were 144 and 135 mM^{-1} cm^{-1}, respectively, at 260 nm. The overexpression and purification of wild-type and D64N UDG have been described[30]. The D64N mutant is ideal for these studies because it has full activity with respect to DNA binding and base flipping[31], but is defective in stabilization of the transition state for glycosidic bond cleavage.

DNA-binding measurements. The K_d values for binding of the T-A and G-C DNA duplexes to UDG were measured using competitive inhibition analysis of the steady-state reaction with product separation by HPLC, as described[32]. A minor modification was that the abasic product (AϕAp) and reactant (AUAp) were separated using isocratic conditions with 10% (v/v) CH_3CN for the T-A decamer DNA duplex and 9% (v/v) CH_3CN for the G-C decamer DNA duplex. Reactions (35 μl) contained 1 mM Tris-HCl, pH 8.0, at 10 °C, 35 mM NaCl, 12.5 μg ml^{-1} BSA, 1 μM AUAp, 6 nM UDG, and a variable amount of the T-A or G-C decamer DNA duplexes. Reactions were incubated at 10 °C for 5 min before direct injection onto the reverse-phase HPLC column. The K_d for each duplex was determined by fitting to equation (2), where k_i is the inhibited rate and k_0 is the rate in the absence of the competitor DNA.

$$k_i / k_o = 1 / (1 + [DNA] / K_d) \quad (2)$$

NMR sample preparation. The two decamer DNA duplexes were hybridized in NMR buffer (1 mM Tris-d_6-HCl, 35 mM NaCl, pH 8.0 at 10 °C). When preparing samples of the complex, 1 mM UDG and 0.8 mM decamer T-A (or G-C) DNA duplex were combined in a total volume of 500 μl at high salt concentration (~500 mM NaCl) to avoid UDG precipitation, and then diluted with NMR buffer several times using a Microcon YM-3 centrifugal concentrator (Millipore) to reduce the NaCl concentration to 35 mM.

NMR spectroscopy. All NMR spectra were acquired using either a Varian Inova 500 or 600 MHz spectrometer equipped with 5-mm triple-resonance probes with *z*-axis-pulsed magnetic field gradients. The resonance of assignments of the DNA imino protons were obtained from 2D ^1H-^1H NOESY experiments with a mixing time of 200 ms, a proton carrier frequency of 4.96 p.p.m. and a spectral width of 13192.6 Hz at 5 °C. For this experiment, a WATERGATE pulse sequence was used to suppress the water signal. These data were processed using NMRPipe[33] using sine-bell or shifted sine-bell window functions, and analyzed using SPARKY version 3 (http://www.cgl.ucsf.edu/home/sparky/) on an Apple G4 computer running OS 10.2. Imino proton exchange rates were determined using NMR magnetization transfer from water using a gradient enhanced pulse sequence as described[34]. The details of these experiments are described in **Supplementary Methods** online.

The imino proton exchange catalyst. An ideal exchange catalyst should be sterically small, have a high enough pK_a value such that exchange catalysis can be observed, and be largely in the neutral basic form at the pH of the experiments (pH 8.0) to prevent large changes in ionic strength that might affect association of the enzyme and DNA. Based on these criteria DFEA was selected as the base catalyst for these studies. The pK_a values for DFEA (pK_a = 7.6) and TFEA (pK_a = 6.3) were measured at 10 °C using 1D ^1H NMR spectroscopy by following the chemical shifts of the methylene protons (data not shown). Because UDG activity was inhibited by 18% at a total DFEA concentration of 300 mM (not shown), concentrations of DFEA were limited to the range 0–50 mM in the exchange studies of the enzyme–DNA complex.

Determination of second-order rate constants for exchange catalysis (k_B). The data analysis requires calculation of the second-order rate constant for base catalysis (k_B), which may be obtained using Eigen proton transfer theory:

$$k_B = k_D / (1 + 10^{pK_a^{nu} - pK_a^{B}}) \quad (3)$$

In this equation, k_D is the bimolecular collision rate, which was estimated as 1.5×10^9 s^{-1} M^{-1} at 10 °C from literature values for several amine base catalysts[21], and pK_a^{nu} and pK_a^{B} are the pK_a values for the imino proton in the DNA duplex, and the DFEA base catalyst, respectively. The pK_a values of guanosine (9.5), thymidine (9.9) and uridine (9.5) at 10 °C were obtained from the literature[21]. No corrections were made for the small anticipated changes in these pK_a values when these groups are present in a DNA environment[35]. Using the measured pK_a = 7.6 for DFEA at 10 °C, and the above values for k_D and pK_a^{nu}, k_B values of 7.3×10^6 M^{-1} s^{-1}, 1.6×10^7 M^{-1} s^{-1} and 1.6×10^7 M^{-1} s^{-1} for general base catalysis of imino proton exchange for thymidine, 2′-fluoro-2′-deoxyuridine and guanosine were obtained, respectively. To investigate the ionic strength–dependence of the exchange rates for the free T-A duplex, small portions of a 5 M NaCl solution

were added to a duplex sample that contained 1 mM Tris-HCl. Magnetization transfer experiments were then done on samples that contained 35, 535, and 1,035 mM NaCl. The highest concentration of NaCl exceeds by over two-fold the highest concentration of difluoroethyl ammonium ion present in the exchange experiments with the free T-A duplex.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

This work was supported by US National Institutes of Health grant GM56834-09 to J.T.S.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 8 September; accepted 20 October 2004
Published online at http://www.nature.com/nsmb/

1. Stivers, J.T. & Jiang, Y.L. A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem. Rev.* **103**, 2729–2759 (2003).
2. Stivers, J.T. Site-specific DNA damage recognition by enzyme-induced base flipping. *Prog. Nucleic Acid Res. Mol. Biol.* **77**, 37–65 (2004).
3. Parikh, S.S. *et al.* Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *EMBO J.* **17**, 5214–5226 (1998).
4. Parikh, S.S. *et al.* Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc. Natl. Acad. Sci. USA* **97**, 5083–5088 (2000).
5. Werner, R.M. *et al.* Stressing-out DNA? The contribution of serine-phosphodiester interactions in catalysis by uracil DNA glycosylase. *Biochemistry* **39**, 12585–12594 (2000).
6. Drohat, A.C. & Stivers, J.T. NMR evidence for an unusually low N1 pK_a for uracil bound to uracil DNA glycosylase: Implications for catalysis. *J. Am. Chem. Soc.* **122**, 1840–1841 (2000).
7. Kwon, K., Jiang, Y. & Stivers, J. Rational engineering of a DNA glycosylase specific for an unnatural cytosine:pyrene base pair. *Chem. Biol.* **10**, 1–20 (2003).
8. Cao, C., Kwon, K., Jiang, Y.L., Drohat, A.C. & Stivers, J.T. Solution structure and base perturbation studies reveal a novel mode of alkylated base recognition by 3-methyladenine DNA glycosylase I. *J. Biol. Chem.* 48012–48021 (2003).
9. Slupphaug, G. *et al.* A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA [see comments]. *Nature* **384**, 87–92 (1996).
10. Lau, A.Y., Scharer, O.D., Samson, L., Verdine, G.L. & Ellenberger, T. Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell* **95**, 249–258 (1998).
11. Lau, A.Y., Wyatt, M.D., Glassner, B.J., Samson, L.D. & Ellenberger, T. Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proc. Natl. Acad. Sci. USA* **97**, 13573–13578 (2000).
12. Jiang, Y.L. & Stivers, J.T. Mutational analysis of the base flipping mechanism of uracil DNA glycosylase. *Biochemistry* **41**, 11236–11247 (2002).
13. Stivers, J.T., Pankiewicz, K.W. & Watanabe, K.A. Kinetic mechanism of damage site recognition and uracil flipping by *Escherichia coli* uracil DNA glycosylase. *Biochemistry* **38**, 952–963 (1999).
14. Roberts, R.J. On base flipping. *Cell* **82**, 9–12 (1995).
15. Krosky, D.J., Schwarz, F.P. & Stivers, J.T. Linear free energy correlations for enzymatic base flipping: how do damaged base pairs facilitate specific recognition? *Biochemistry* **43**, 4188–4195 (2004).
16. Jencks, W.P. When is an intermediate not an intermediate—enforced mechanisms of general acid-base catalyzed, carbocation, carbanion, and ligand-exchange reactions. *Acc. Chem. Res.* **13**, 161–169 (1980).
17. Chen, L., Haushalter, K.A., Lieber, C.M. & Verdine, G.L. Direct visualization of a DNA glycosylase searching for damage. *Chem. Biol.* **9**, 345–350 (2002).
18. Verdine, G.L. & Bruner, S.D. How do DNA repair proteins locate damaged bases in the genome? *Chem. Biol.* **4**, 329–334 (1997).
19. Seibert, E., Ross, J.B. & Osman, R. Role of DNA flexibility in sequence-dependent activity of uracil DNA glycosylase. *Biochemistry* **41**, 10976–10984 (2002).
20. Banavali, N.K. & MacKerell, A.D. Jr. Free energy and structural pathways of base flipping in a DNA GCGC containing sequence. *J. Mol. Biol.* **319**, 141–160 (2002).
21. Gueron, M. & Leroy, J.-L. Studies of Base Pair Kinetics by NMR Measurement of Proton Exchange. *Methods Enzymol.* **261**, 383–413 (1995).
22. Guéron, M. *et al.* Applications to imino proton exchange to nucleic acid kinetics and structure. In *Structure and Methods* (eds. Sarma, R.H. & Sarma, M.H.) 113–137 (Adenine Press, Guilderland, New York, 1990).
23. Kavli, B. *et al.* Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. *EMBO J.* **15**, 3442–3447 (1996).
24. Gueron, M., Kochoyan, M. & Leroy, J.L. A single mode of DNA base-pair opening drives imino proton exchange. *Nature* **328**, 89–92 (1987).
25. Warmlander, S., Sen, A. & Leijon, M. Imino proton exchange in DNA catalyzed by ammonia and trimethylamine: evidence for a secondary long-lived open state of the base pair. *Biochemistry* **39**, 607–615 (2000).
26. Handa, P., Acharya, N. & Varshney, U. Effects of mutations at tyrosine 66 and asparagine 123 in the active site pocket of *Escherichia coli* uracil DNA glycosylase on uracil excision from synthetic DNA oligomers: evidence for the occurrence of long-range interactions between the enzyme and substrate. *Nucleic Acids Res.* **30**, 3086–3095 (2002).
27. Jiang, Y.L., Song, F. & Stivers, J.T. Base flipping mutations of uracil DNA glycosylase: substrate rescue using a pyrene nucleotide wedge. *Biochemistry* **41**, 11248–11254 (2002).
28. Bommarito, S., Peyret, N. & SantaLucia, J. Jr. Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.* **28**, 1929–1934 (2000).
29. Stivers, J.T., Jagadeesh, G.J., Nawrot, B., Stec, W.J. & Shuman, S. Stereochemical outcome and kinetic effects of Rp- and Sp- phosphorothioate substitutions at the cleavage site of vaccinia type I DNA topoisomerase. *Biochemistry* **39**, 5561–5572 (2000).
30. Drohat, A.C., Jagadeesh, J., Ferguson, E. & Stivers, J.T. Role of electrophilic and general base catalysis in the mechanism of *Escherichia coli* uracil DNA glycosylase. *Biochemistry* **38**, 11866–11875 (1999).
31. Jiang, Y.L., Kwon, K. & Stivers, J.T. Turning on uracil-DNA glycosylase using a pyrene nucleotide switch. *J. Biol. Chem.* **276**, 42347–42354 (2001).
32. Jiang, Y.L., Ichikawa, Y. & Stivers, J.T. Inhibition of uracil DNA glycosylase by an oxacarbenium ion mimic. *Biochemistry* **41**, 7116–7124 (2002).
33. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
34. Snoussi, K. & Leroy, J.L. Imino proton exchange and base-pair kinetics in RNA duplexes. *Biochemistry* **40**, 8898–9904 (2001).
35. Benight, A.S., Schurr, J.M., Flynn, P.F., Reid, B.R. & Wemmer, D.E. 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 (1988).