A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA

A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA

Paul C. Blainey*, Antoine M. van Oijen*†, Anirban Banerjee*, Gregory L. Verdine*§¶, and X. Sunney Xie*¶

Departments of *Chemistry and Chemical Biology and §Molecular and Cellular Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138

Edited by Peter H. von Hippel, University of Oregon, Eugene, OR, and approved February 14, 2006 (received for review November 8, 2005)

A central mystery in the function of site-specific DNA-binding proteins is the detailed mechanism for rapid location and binding of target sites in DNA. Human oxoguanine DNA glycosylase 1 (hOgg1), for example, must search out rare 8-oxoguanine lesions to prevent transversion mutations arising from oxidative stress. Here we report high-speed imaging of single hOgg1 enzyme molecules diffusing along DNA stretched by shear flow. Salt-concentration-dependent measurements reveal that such diffusion occurs as hOgg1 slides in persistent contact with DNA. At near-physiologic pH and salt concentration, hOgg1 has a subsecond DNA-binding time and slides with a diffusion constant as high as 5×10^6 bp^2/s. Such a value approaches the theoretical upper limit for one-dimensional diffusion and indicates an activation barrier for sliding of only 0.5 kcal/mol (1 kcal = 4.2 kJ). This nearly barrierless Brownian sliding indicates that DNA glycosylases locate lesion bases by a massively redundant search in which the enzyme selectively binds 8-oxoguanine under kinetic control.

nonspecific binding | shear flow stretching | single-molecule imaging | DNA glycosylase mechanism | hopping

DNA glycosylases initiate base-excision DNA repair by catalyzing excision of damaged bases from the genome. Some of these lesion bases differ only subtly from native bases, rendering the proteins' task of locating these lesions amidst the overwhelming excess of native bases a particularly challenging search problem (1, 2). The problem is exacerbated by the fact that DNA glycosylases need to completely extrude their substrates from the DNA helix to catalyze base excision in an extrahelical enzyme active site (2, 3). This finding raises the intriguing question of whether DNA glycosylases are able to distinguish lesions from normal bases while they are situated within the DNA helix, or, alternatively, whether the enzymes need to extrude each base for presentation to the extrahelical lesion-recognition pocket on the enzyme (1).

The mechanism of lesion recognition has received extensive attention in the case of human oxoguanine DNA glycosylase 1 (hOgg1) (Fig. 1*A*), the enzyme responsible for removal of the highly mutagenic base 8-oxoguanine (oxoG) from the human genome (4–7). Arising from the attack of reactive oxygen species on guanine (G) residues, oxoG differs from G at only two atomic positions (Fig. 1*B*). Despite its structural similarity to G, oxoG mispairs almost exclusively with adenine (A) during replication, causing G:C→T:A transversion mutations. Interrogation of DNA by hOgg1 to locate oxoG residues is made difficult not only by the vast excess of structurally similar G residues but also because oxoG causes no discernible alteration to the helical structure of DNA (8, 9) and only slightly destabilizes the duplex (10). It has recently been demonstrated that the extrahelical lesion-recognition pocket of hOgg1 has remarkable selectivity for oxoG versus G (7 kcal/mol; 1 kcal = 4.18 kJ) (7), but it remains unknown whether hOgg1 also possesses the ability to distinguish intrahelical oxoG:C pairs from G:C. Here we report direct real-time observations of the behavior of single hOgg1 enzyme molecules moving along undamaged DNA that provide insight into the strategy used by the enzyme to search out lesion bases.

It has been recognized that the rate at which DNA-binding proteins associate with their specific sites on DNA can exceed the diffusion-limited association rate (11). The ability of some proteins to approach or exceed the diffusion-limited rate is critical to their biological function, for example, allowing the rapid modulation of gene expression levels by activators and repressors of transcription, restriction of viral DNA inside bacteria, and repair of DNA lesions before they become fixed as mutations or chromosomal defects (12, 13).

These proteins' ability to bind non-target (nonspecific) DNA is key to their rapid target-binding activity. The law of mass action dictates that even proteins with modest nonspecific binding constants spend the majority of time bound to nonspecific DNA under the *in vivo* condition of high DNA concentration (14, 15). Multistep target location involving initial association with nonspecific DNA and subsequent binding to the target is called "facilitated diffusion." It is generally accepted that rapid target search *in vivo* is made possible through a combination of one-dimensional diffusion along DNA segments and three-dimensional transfer among DNA segments (12, 13, 16); the work reported here concerns the one-dimensional process.

Although facilitated diffusion has been inferred from the kinetic behavior of several classes of DNA-binding proteins (11, 17–20), direct real-time observation of this phenomenon has proven difficult and has only been reported in the case of *Escherichia coli* RNA polymerase (21–23). This lack of direct experimental data has hindered progress toward a detailed understanding of facilitated diffusion.

For proteins that do not consume biochemical energy to assist their diffusion, target search along DNA is a thermally activated and directionally unbiased process. Two distinct mechanisms for one-dimensional diffusion along DNA, sliding and hopping, have been proposed (11, 12). Sliding is a process wherein the protein undergoes diffusion while remaining bound to the DNA. Hopping refers to the protein repeatedly dissociating from the DNA and rebinding at a new location on the DNA. The likelihood of proximate dissociation and rebinding sites, particularly on stretched DNA, makes distinguishing the two mechanisms difficult. Innovative biochemical studies with restriction enzymes have demonstrated hopping and suggested that sliding, if it occurs, only contributes to movement on length scales of <173 bp (17, 24). The question remains whether sliding, in addition to hopping, contributes to one-dimensional movement in the context of facilitated diffusion.

The ability to record trajectories of single molecules moving on DNA afforded by recent advances in single-molecule microscopy (21) in combination with high-throughput data collection

---

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: oxoG, 8-oxoguanine; hOgg1, human oxoguanine DNA glycosylase 1.

†Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115.

¶To whom correspondence may be addressed. E-mail: verdine@chemistry.harvard.edu or xie@chemistry.harvard.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0509723103

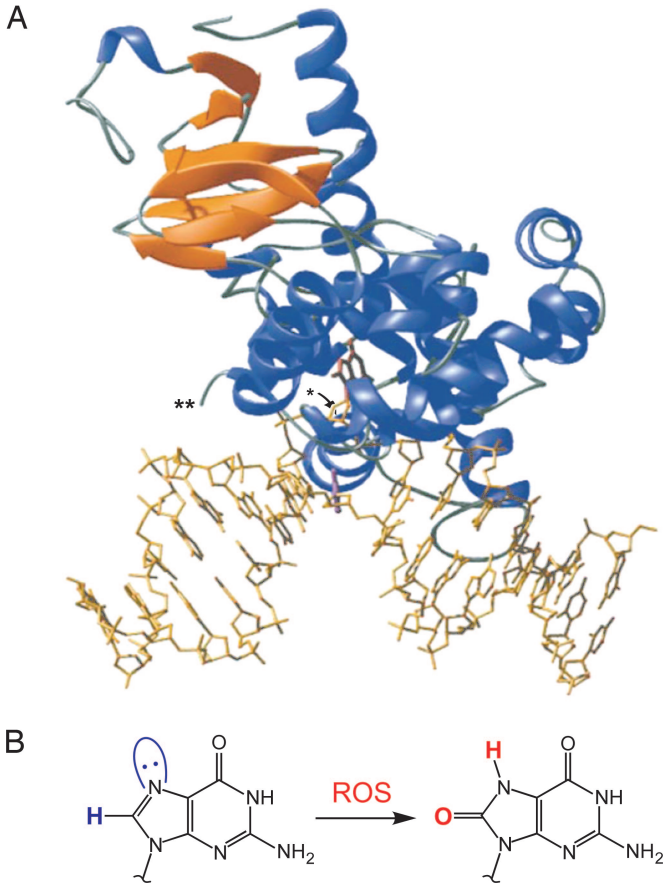

Fig. 1. hOgg1 and its cognate lesion oxoG. (*A*) hOgg1 cocrystal structure with oxoG:C-containing DNA (5). The location of His-270 is indicated by an asterisk and arrow. The C terminus is indicated by a double asterisk. (*B*) Scheme depicting the formation of the oxoG lesion by attack of reactive oxygen species (ROS) on guanine.

and analysis allows us to directly address this question. Our *in vitro* assay (Fig. 2), based on flow-stretching DNA, (25, 26) makes it possible to track the one-dimensional diffusion of many individual enzyme molecules along DNA with high spatial and temporal resolution across various solution conditions.

It has been proposed that sliding and hopping can be distinguished by evaluating the dependence of facilitated diffusion on salt concentration, which modulates the affinity of many proteins for nonspecific DNA (12, 27). Our salt-concentration-dependent single-molecule data show that hOgg1 can move along local stretches of DNA by sliding in persistent contact with the duplex. In addition, we find that the one-dimensional diffusion constant is strongly pH-dependent, indicating that specific molecular interactions within the protein-DNA interface contribute to hOgg1's rapid sliding activity. The new data provide constraints on models of lesion recognition by oxoG glycosylases.

Results

We observe bidirectional movement of hOgg1 molecules over kilobases of undamaged DNA on the time scale of seconds (Fig. 3*A*) (see *Supporting Text*, which is published as supporting information on the PNAS web site, for control showing that hOgg1 recognizes oxoG in the single-molecule assay). The net displacements of individual molecules are distributed symmetrically around zero (Fig. 3*B*), indicating that movement is not biased in the flow direction (see *Supporting Text*). The mean-square displacement in the direction of DNA extension (x) for

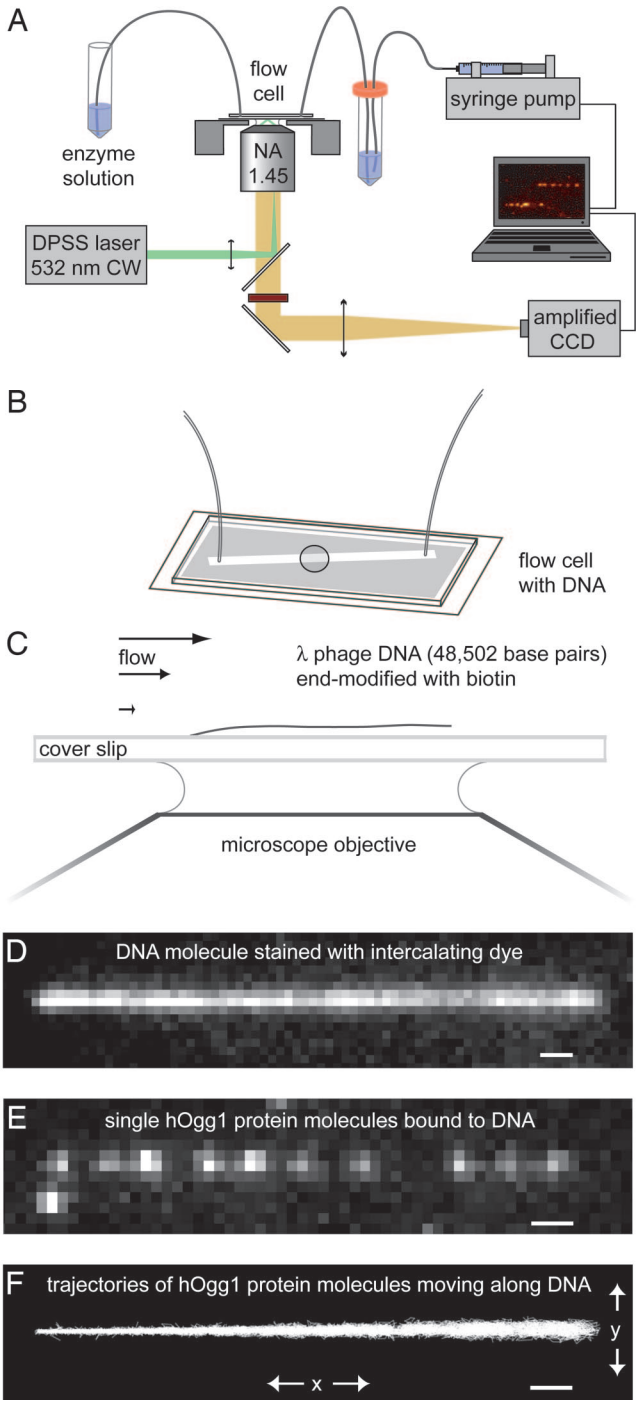

Fig. 2. Single-molecule assay for protein translocation along DNA. (*A* and *B*) Inverted microscope fitted for total internal reflection fluorescence imaging (*A*) with mounted flow cell (*B*). (*C*) Schematic of flow-stretched DNA molecule (not to scale). Buffer solution flows over the glass coverslip to which double-stranded λ DNA, 16 μm in length, is attached by one end. (*D*) Image of SYTOX orange-stained λ DNA molecule stretched by flow (integration time, 0.040 s). (*E*) Image of 10 single hOgg1 enzymes bound to an undamaged DNA molecule (integration time, 0.050 s). We selected an image from an experiment with a higher than usual protein concentration (0.2 nM) for the given buffer condition (0.01 M NaCl, pH 7.5) to obtain the larger-than-usual number of DNA-bound enzyme molecules. Noncollinear signals arise from surface-adsorbed enzymes. (*F*) Trajectories of 700 enzyme molecules diffusing on an undamaged λ DNA molecule over the course of 150 s (note broad distribution of transverse positions at the free end due to fluctuations of the DNA position). Buffer flows toward the right in *A* and *C–F*. The DNA extension in *D* is greater because of the intercalating dye. (Scale bars, 1.0 μm.)

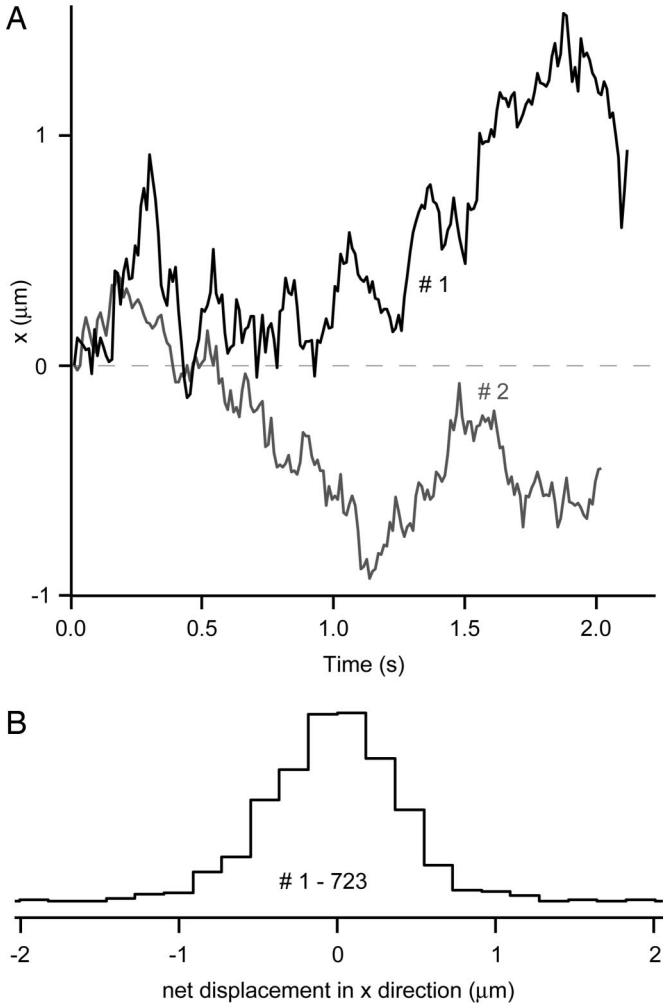

Fig. 3. Diffusion of single hOgg1 enzyme molecules along DNA. (*A*) Trajectories of two hOgg1 molecules diffusing along an undamaged DNA molecule at pH 7.5 with 0.01 M NaCl (molecule no. 1, black solid line; molecule no. 2, gray solid line) in the direction of DNA extension/flow (toward positive *x*-axis direction). Starting times and positions are defined as zero (starting position indicated by gray dashed line). (*B*) Histogram of net displacement ($= x_{final} - x_{initial}$) for 723 trajectories collected in one experiment on an undamaged DNA molecule. The histogram mean is 0.013 μm.

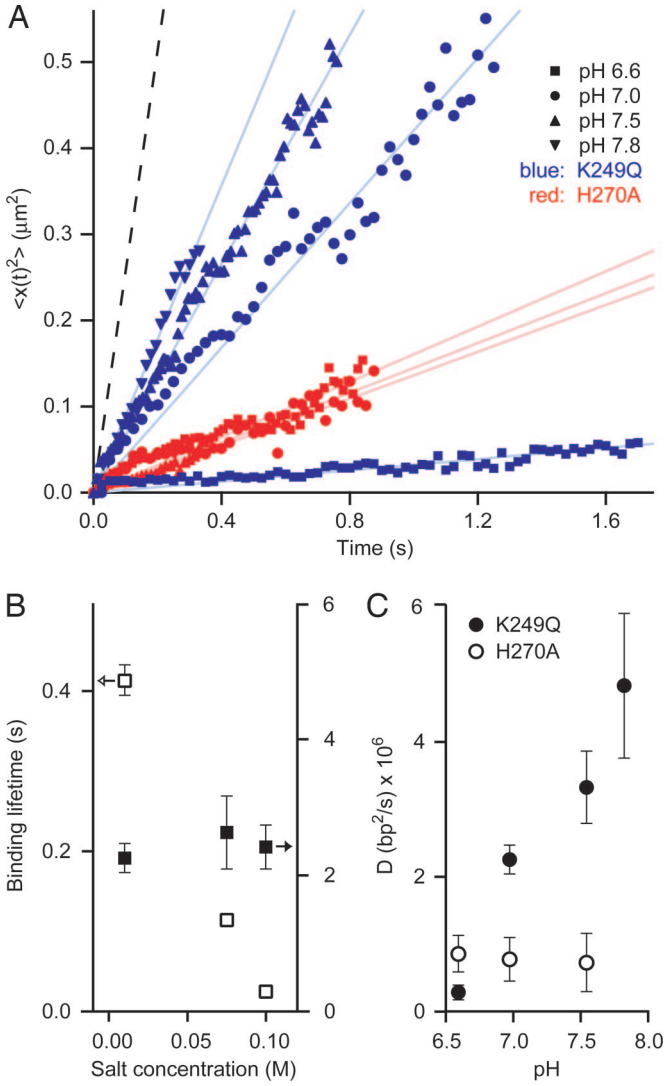

Fig. 4. pH and salt concentration-dependence of hOgg1's sliding activity. (*A*) Mean-square displacement of hOgg1 K249Q (blue points) and hOgg1 H270A (red points) along undamaged DNA at various indicated pH values. Data for each measurement represent an average over >100 binding events (except for hOgg1 H270A at pH 7.5, for which the average is over eight events). Solid lines are least-squares fits to the data. The case for free sliding is shown, indicated by the dashed line with slope 2.46 $\mu m^2/s$. The salt concentration for all plots in *A* is 0.010 M. (*B*) Mean binding lifetime (□, left axis as indicated) and diffusion constant (■, right axis as indicated) of the K249Q mutant at pH 7.0 as a function of salt concentration along undamaged DNA. (*C*) Diffusion constant versus pH for the two mutants along undamaged DNA with salt concentration 0.010 M. Error bars represent SD. Where no error bar appears in *B* and *C*, the uncertainty is comparable to the dimension of the plotted symbol.

hOgg1 at pH 7.8, averaged over hundreds of molecules, is shown as a function of time (t) in Fig. 4*A* (blue trace with greatest slope). The linear relationship, characteristic of Brownian motion, is described by a one-dimensional diffusion constant (D) of $4.8 \pm 1.1 \times 10^6$ bp^2/s according to $D = \langle x^2 \rangle / 2t$. All data for Fig. 4 *A* and *C* were taken at a salt concentration of 0.010 M.

Varying the salt concentration in our experiments allows us to determine whether the linear diffusion is caused by sliding or hopping. Were the observed one-dimensional diffusion due to hopping, faster diffusion would be predicted at higher salt concentration. This dependence is expected because stronger electrostatic screening lowers the nonspecific binding affinity, increasing the fraction of time that the enzyme is unbound and mobile. However, when increasing the salt concentration to 0.1 M, the highest achievable because of hOgg1's decreasing affinity for DNA, we observe no change in the diffusion constant despite the marked decrease in mean binding lifetimes (Fig. 4*B*). Our finding that hOgg1's one-dimensional diffusion constant is invariant with salt concentration demonstrates that sliding, rather than hopping, underlies the one-dimensional diffusion observed in our assay.

We calculate the upper limit of hOgg1's diffusion constant to be 1.2×10^7 bp^2/s at 25°C, based on a hydrodynamic model (28) developed for a protein sliding freely along the DNA helix (Fig. 4*A*, dashed line). The measured diffusion constant is smaller than this theoretical value because of the presence of an activation barrier for translocation along DNA (12, 28). To determine the magnitude of this energy barrier, we consider that the protein undergoes a random walk along DNA with a step size of 1 bp. The random walk stepping rate corresponding to the free sliding diffusion constant is $k_{lim} = 2.4 \times 10^7$ steps per second, and the rate corresponding to the measured diffusion constant is $k = 1.0 \pm 0.2 \times 10^7$ steps per second. According to the

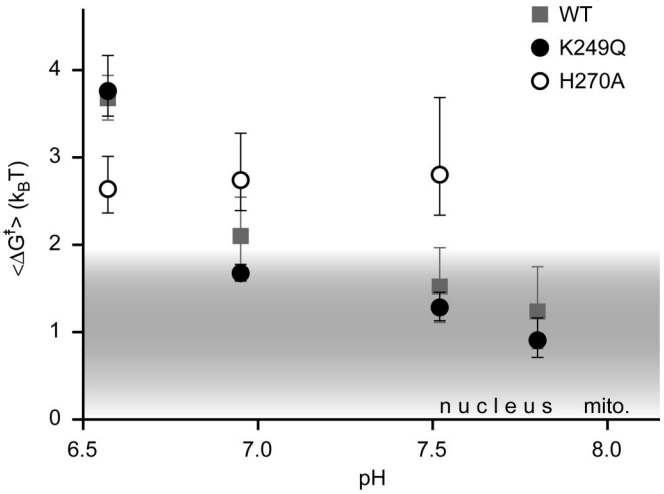

Fig. 5. Mean activation energy for hOgg1 sliding on undamaged DNA. Shown is the mean activation energy for sliding versus pH. The shaded area indicates activation energies at which fast searching is possible (29). Also indicated are the range of intranuclear pH values measured in mammalian cells (7.55–7.79) (33) and the pH of the mitochondrial matrix (mito) (pH 8). Error bars represent SD.

Arrhenius relation, the mean activation free energy for hOgg1 to slide is $\langle \Delta G^{\ddagger} \rangle / k_B T = \ln (k_{lim}/k) = 0.9 \pm 0.2\, k_B T$ (pH 7.8). Note that because k and k_{lim} vary in the same manner for different chosen step sizes, the value of $\langle \Delta G^{\ddagger} \rangle$ does not depend on the step size specified. The low barrier height we have determined is consistent with previous theoretical predictions that $\approx 2\, k_B T$ is the maximum roughness of the energy landscape tolerated for fast search (29) and that sliding with a rate close to the hydrodynamic limit is possible (12).

Surprisingly, hOgg1's diffusion constant on DNA grows by more than an order of magnitude between pH 6.6 and 7.8 (Fig. 4 *A* and *C*). The only amino acid with a side-chain pK_a close to this range is histidine (pK_a = 6.1) (30), which becomes positively charged at low pH. The crystal structures of hOgg1–DNA complexes (5, 7) reveal a conserved histidine (His-270) contacting the DNA that could explain the pH dependence. We replaced His-270 with alanine and found that the diffusion constant for the H270A mutant is reduced to $0.8 \pm 0.4 \times 10^6$ bp^2/s, and is independent of pH across the entire range tested (Fig. 4 *A* and *C*). This result proves that His-270 is indeed responsible for the observed pH dependence and indicates that hOgg1's specialized DNA-binding interface contributes to its rapid sliding activity. Recent observation of DNA groove tracking by a nonspecifically bound enzyme supports the notion of highly complementary protein–DNA interactions in such complexes (31). It appears that native interactions within the hOgg1–DNA interface can lower the activation barrier for protein translocation along DNA in a manner analogous to the reduction of activation barriers for chemical reactions by enzymes (32). In this sense, hOgg1 seems to be fine-tuned by nature to play a catalytic role in its own translocation along DNA by making specific interactions within the DNA-binding interface that reduce the activation energy for sliding.

The mean activation energy for sliding by hOgg1 is plotted as a function of pH in Fig. 5. It is interesting to note that the measured activation barriers for sliding of WT and His-270-containing hOgg1 fall below 2 $k_B T$ at the intranuclear pH of mammalian cells, pH 7.55–7.79 (33). We further note that a low barrier for sliding at the pH of the mitochondrial matrix, pH 8, would allow hOgg1 to search rapidly for lesions on mitochondrial DNA, which experiences high oxidative stress.

Discussion

Rapid search by facilitated diffusion has been attributed to the combination of one-dimensional (along DNA) and three-dimensional (sequence-distal transfer) processes. That hopping contributes to intersegment and local transfer has been proven experimentally (17, 24). Our experiments provide definitive proof that sliding dominates one-dimensional movement along DNA at a length scale of hundreds of bases over a broad range of salt concentrations.

A protein molecule's search behavior may differ in cells compared with the *in vitro* single-molecule assay. For instance, search *in vivo* may be directly or indirectly coupled to DNA-dependent molecular motors. Intracellular solution conditions and the state of cellular DNA are significantly different from those of the single-molecule assay. In the assay, we approach cellular salt concentrations and replicate physiological pH but not the compact DNA configuration, occlusion of binding sites by DNA-bound proteins, high total concentration of searching protein molecules, and macromolecular crowding known to occur *in vivo*.

From measurements of hOgg1's copy number in cells (34), we know that the physiological concentration is orders of magnitude higher than is used in the single-molecule assay. However, the stoichiometry of bound glycosylase to base pairs of DNA is remarkably similar in cells and in the *in vitro* assay: approximately one bound protein molecule per 10^4 bp of DNA. *In vivo*, much of the DNA is inaccessible due to bound proteins including histones, which constitute at least partial blockades to the motion of a sliding protein molecule. We can estimate hOgg1's *in vivo* sliding length based on the observed diffusion constant at physiological pH and binding lifetime measured at high salt concentration. Taking $D = 5 \times 10^6$ bp^2/s and the mean binding lifetime of 0.025 s at the highest salt concentration tested, we compute the mean sliding length at 440 bp (see *Supporting Text*). Such short sliding lengths seem compatible with searching the short segments of DNA accessible in a cellular environment. It has been noted previously that selective occlusion of non-target DNA could significantly accelerate target location (29).

If searching proteins depended solely on translational diffusion to move along DNA (e.g., by only hopping), we would expect the diffusion constant along DNA to be much lower in the crowded interior of cells. However, the rate at which a protein slides along the DNA helix is limited primarily by the protein's rotational diffusion constant (28). Because macromolecular crowding slows rotational diffusion only weakly compared with translational diffusion (35), we propose that hOgg1's diffusion along DNA in cells is slowed by a factor much smaller than the increase in solution viscosity.

These considerations predict efficient search of DNA *in vivo* through the synergistic interplay of hopping and sliding. The relative insensitivity of the sliding rate to macromolecular crowding allows fast local one-dimensional searching. Frequent and rapid hops, arising from the short mean binding time and rapid rebinding overcome the redundancy of sliding on long length scales and constitute an efficient mechanism for obstacle bypass.

To test the generality of rapid one-dimensional diffusion among oxoG glycosylases, we observed single molecules of hOgg1's prokaryotic counterpart, MutM, bound to undamaged DNA. Although MutM is functionally similar to hOgg1, the two proteins have completely different folds and their DNA complexes are dissimilar in overall shape (5, 36). We find that MutM slides along DNA with $D = 3.5 \pm 0.6 \times 10^5$ bp^2/s (Fig. 6, which is published as supporting information on the PNAS web site), which corresponds to an activation energy for sliding of 3.3 ± 0.2 $k_B T$. Although the sliding rate measured for MutM is lower than that for hOgg1 at physiological pH (all measurements made at room temperature; compare Fig. 3 with Fig. 6), the bacterial

enzyme has only 1/1,000th the number of base pairs to patrol, and MutM's diffusion constant will be significantly higher at the 55–65°C growth temperature of *Bacillus stearothermophilus* than we measured at room temperature.

The sliding exhibited by these glycosylases has implications for the mechanism of lesion recognition. Sliding along DNA with $D = 5 \times 10^6$ bp^2/s, hOgg1 can traverse 1,000 bp of DNA in 0.1 seconds according to $\langle x \rangle = (2Dt)^{1/2}$. This distance is the expected displacement after 1,000,000 random walk steps of length 1 bp along the DNA or, alternatively, 1,000,000 base pair sampling events. Such rapid sliding, with an activation barrier of only ≈0.5 kcal/mol (0.9 − 1.3 k_BT, pH 7.5–7.8, 298 K), all but excludes the possibility that hOgg1 extrudes each base from the DNA helix and tests that base for equilibration into the enzyme's extrahelical active site (1). Furthermore, rapid sliding excludes the possibility that hOgg1 captures a spontaneously formed extrahelical oxoG, a species expected to appear at low frequency (37–39) (no measurements of oxoG:C opening rates are available). Consequently, hOgg1 must possess some mechanism of recognizing intrahelical oxoG residues and facilitating their extrusion from the DNA helix.

Unlike unidirectional movement along DNA by molecular motors, the Brownian target search is extraordinarily redundant. With this search mechanism, it is not important that hOgg1 recognize oxoG upon each encounter, because the protein will have many additional encounters while interrogating a particular segment of DNA. However, it is crucially important that hOgg1 minimize the frequency at which it selects normal bases for detailed interrogation because of the large number of encounters with nonlesion base pairs. Direct evidence for such "negative selection" against a time-consuming search of G:C base pairs has recently been gleaned from cocrystal structures of MutM bound to non-lesion-containing DNA (40). Collectively, these considerations argue that hOgg1 extrudes oxoG lesions into the active site with a higher rate than it does for undamaged bases, thus discriminating lesions by kinetic control.

We find that DNA glycosylases make good use of the time they spend bound to DNA, forming nonspecific complexes that support rapid sliding to search out defect bases. Given these proteins' success in solving a difficult search problem, a highly efficient mechanism for target identification, such as selective kinetically controlled lesion capture, must be expected. This study highlights the promise that single-molecule tracking measurements can complement structural studies to form a deeper understanding of complex protein-DNA interactions.

Materials and Methods

Translocation Assay Overview. hOgg1 enzyme molecules were labeled specifically and quantitatively with a fluorophore at the C terminus (Fig. 1*A*). λ DNA was tethered to a glass surface at one end and stretched by a laminar flow of buffer (Fig. 2 *A–D*) (25, 26). Single hOgg1 molecules that bind to and diffuse along the DNA were illuminated by a laser beam and imaged with a fluorescence microscope (Fig. 2*A*). Fluorescent particles are assigned as hOgg1 monomers because they demonstrate an emission rate consistent with a single dye molecule, a unimodal intensity distribution, and photobleach in one step (see the light-scattering measurements in *Supporting Text*). We recorded 20–100 images per second with a charge-coupled device camera (Fig. 2 *A* and *E*). In a typical movie, we observed hundreds to thousands of hOgg1 molecules bind, move along, and dissociate from individual DNA molecules (Movie 1, which is published as supporting information on the PNAS web site). For each experiment, the excitation power was adjusted such that photobleaching of the label occurred on a longer time scale than the duration of DNA-binding events.

Each image of single molecules was fit with multiple (one for each molecule) two-dimensional Gaussian functions (41) by using DIATRACK software (Semasopht, North Epping, Australia). The centroid positions of the fits from a series of images were reconstructed as spatial trajectories of each protein molecule appearing in the sequence (Fig. 2*F*). The accuracy of a measured centroid (10–50 nm under typical conditions) was primarily limited by the number of photons collected (see *Supporting Text*). Because we were initially concerned about the effect of dynamic conformational fluctuations of the DNA (26) on measurements of one-dimensional diffusion, a dedicated effort was made to investigate their potential effect. Brownian dynamics simulations (Fig. 7*A*, which is published as supporting information on the PNAS web site) and observations of λ DNA stained with a fluorescent intercalating dye (Fig. 7*B*) indicated that conformational fluctuations of the DNA in the flow direction were sufficiently fast such that they did not interfere with measurements of protein translocation under our experimental conditions.

DNA Preparation and Flow Stretching. λ DNA (New England Biolabs) was enzymatically ligated to a modified oligo [5′-ggg cgg cga cct $(aaa)_4$-biotin] and incubated at 1 pM in a flow cell constructed as described in ref. 25 (Fig. 2*B*). The poly(ethylene glycol) brush coating the coverslip was critical to suppress adsorption of labeled protein molecules to the glass coverslip. Buffer solution was drawn through the flow cell at a rate of 10.0 ml/hr by a syringe pump (model no. PHD2000, Harvard Apparatus), creating shear flow near the coverslip surface and stretching the DNA (Fig. 2*C*) (26). The flow velocity at the DNA was 100 μm/s (see *Supporting Text*).

Single-molecule measurements were made with 0.05–1.0 nM hOgg1–Cy3B in 0.01 M Tris or phosphate buffer/0.01–0.1 M NaCl/0.005 M 2-mercaptoethanol/0.5 mg/ml BSA/0.0001 M EDTA/5% glycerol at final pH 6.6–7.8 or with 0.01–0.10 nM MutM–Cy3B in similar buffers (lacking EDTA) at final pH 7.0–7.5.

Protein Preparation and Labeling. We overproduced α-hOgg1 with an N-terminal hexahistidine tag (α indicates the hOgg1 splice variant targeted to the nucleus) including amino acids 12–345 in *E. coli*. All hOgg1 proteins contained an engineered cysteine at the C terminus for labeling; some had the point mutation K249Q, which inactivates glycosylase activity, or H270A. Single-molecule experiments performed with WT and K249Q hOgg1 demonstrated that the K249Q mutation does not affect hOgg1's pH-dependent translocation activity (Fig. 5).

Single-molecule experiments using hOgg1 with a C-terminal truncation (at residue 322) reveal that neither the presence nor the absence of the amino acids excluded from the C terminus in crystallographic studies (residues 327–345) affect the protein's pH-dependent translocation activity (data not shown). From the same measurements, we conclude that the translocation dynamics observed are independent of the position of the engineered C-terminal cysteine and dye molecule label.

Purified hOgg1 was labeled by reaction with Cy3B–maleimide (Amersham Pharmacia Biosciences) at the C-terminal engineered cysteine. Conjugation under native conditions (0.02 M Tris or phosphate buffer, pH 7.5/0.5 M NaCl/0.001 M EDTA/10% glycerol) with excess reactive dye depended on the presence of the terminal cysteine, indicating that none of the eight native cysteine residues are labeled under the conditions used. Reaction with the sterically accessible C-terminal cysteine proceeded to completion over 60 min at room temperature under argon, assayed by visible (for quantitation of the label) and UV (for quantitation of the protein) absorption after stringent removal of the free dye by protein-affinity purification of the labeled sample. The selectivity of labeling was confirmed by mass spectrometry.

hOgg1 was also labeled with tetramethylrhodamine (Molec-

ular Probes), Rhodamine red (Molecular Probes), and Cy3 (Amersham Pharmacia Biosciences). Because the translocation activity observed did not vary with either the position or the charge of the label (data not shown), we concluded that the dye molecule label does not significantly influence the translocation dynamics.

WT MutM from *B. stearothermophilus* was cloned with a C-terminal cysteine residue and overproduced in *E. coli*. Similarly to hOgg1, the unmodified protein does not react with Cy3B–maleimide, but that containing the engineered cysteine reacts readily at room temperature.

Fluorescence Imaging. An inverted microscope (Olympus IX70) was used for wide-field imaging. A 40-μm diameter area was illuminated with 80–300 W/cm^2 at the critical angle for total internal reflection by focusing the beam of a 532-nm diode laser (UGA-250; LambdaPro, Beijing) into the back aperture of a $\times$60 objective lens (numerical aperture 1.45; PlanApo, Olympus) (Fig. 2*A*). Fluorescent emission was collected through a dichroic mirror (model no. Z532RDC; Chroma Technology, Rockingham, VT) and bandpass filter (model no. HQ575/50; Chroma Technology) to be imaged on a back-illuminated and amplified charge-coupled device camera (Cascade:512B; Roper Scientific, Trenton, NJ). The protein concentration was chosen for each experiment such that, on average, fewer than three hOgg1 molecules were bound per DNA molecule.

Calculation of Diffusion Constant for Free Sliding Proteins on DNA. The calculation was performed according to

$$D_1 = \frac{k_B T}{6\pi\eta a[1 + (4/3)(2\pi)^2(a/3.4 \times 10^{-9})^2]},$$

as described in the literature for the *E. coli lac* repressor (28), where the denominator corresponds to the "effective friction" and η is the solution viscosity. We used a = hOgg1's hydrodynamic radius of 3.2 nm (see *Supporting Text*) and applied two improvements to the model to accommodate aspects of hOgg1–DNA complexes visible in cocrystal structures (5, 7). First, we adjusted for the fact that the center of mass of the protein is displaced 2.5 nm from the axis of the DNA. Incorporating the 5.6-fold-longer translational path implied by the offset increases the effective friction for sliding by 10%. Second, the aspheric shape of hOgg1 apparent from the crystal structure was taken into account. We approximated hOgg1 as an oblate ellipsoid of revolution with axial ratio 2.2 and chose rotation about the short axis as that coupled to sliding, increasing the effective friction by an additional 20% compared with a sphere of equal volume (42). The total adjustment of 30% reduced the theoretical upper limit on the diffusion constant for free sliding to 1.2 μm^2/s, 62 times lower than that measured for free diffusion of hOgg1 in solution (see *Supporting Text*). The calculation for MutM was carried out in a similar manner but for modeling the protein as a prolate ellipsoid of revolution. We found the sensitivity of the free sliding rate to the size, shape, and orientation of the protein (through rotational friction) notable.

We thank C. M. Schroeder for stochastic simulations of flow-stretched DNA, M. Slutsky and L. Mirny for discussions on the mechanism of one-dimensional diffusion, S. Jiralerspong for discussions regarding unpublished MutM cocrystal structures, C. Chenshaw for DLS measurements on hOgg1, C. Radom (Harvard University) for plasmid DNA containing hOgg1 H270A, C. Zong (Harvard University) for preparation of plasmid DNA, B. H. King for help with labeling hOgg1, and M. Ward for assistance in data analysis and hydrodynamic calculations. This work is supported by National Institutes of Health Grants CA100742 (to G.L.V.), GM061577 (to X.S.X.), ST32 GM07598-25 (to P.C.B. through the Harvard University's Molecular, Cellular, and Chemical Biology training program) and a National Institutes of Health Director's Pioneer Award (to X.S.X.). A.M.v.O. was supported by a Human Frontier Science Program fellowship.

1. Verdine, G. L. & Bruner, S. D. (1997) *Chem. Biol.* **4,** 329–334.
2. Zharkov, D. O. & Grollman, A. P. (2005) *Mutat. Res.* **577,** 24–54.
3. Fromme, J. C. & Verdine, G. L. (2004) *Adv. Protein Chem.* **69,** 1–41.
4. Bjoras, M., Luna, L., Johnsen, B., Hoff, E., Haug, T., Rognes, T. & Seeberg, E. (1997) *EMBO J.* **16,** 6314–6322.
5. Bruner, S. D., Norman, D. P. G. & Verdine, G. L. (2000) *Nature* **403,** 859–866.
6. Chen, L., Haushalter, K. A., Lieber, C. M. & Verdine, G. L. (2002) *Chem. Biol.* **9,** 345–350.
7. Banerjee, A., Yang, W., Karplus, M. & Verdine, G. L. (2005) *Nature* **434,** 612–618.
8. Grollman, A. P. (1992) in *Structure and Function: Proceedings of the Seventh Conversation in the Discipline Biomolecular Stereodynamics*, eds. Sarma, R. H. & Sarma, M. H. (Adenine, Schenectady, NY), pp. 165–170.
9. Lipscomb, L. A., Peek, M. E., Morningstar, M. L., Verghis, S. M., Miller, E. M., Rich, A., Essigmann, J. M. & Williams, L. D. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 719–723.
10. Plum, G. E., Grollman, A. P., Johnson, F. & Breslauer, K. J. (1995) *Biochemistry* **34,** 16148–16160.
11. Riggs, A. D., Bourgeois, S. & Cohn, M. (1970) *J. Mol. Biol.* **53**, 401–417.
12. Berg, O. G., Winter, R. B & von Hippel, P. H. (1981) *Biochemistry* **20,** 6929–6948.
13. Halford, S. E. & Marko, J. F. (2004) *Nucleic Acids Res.* **32,** 3040–3052.
14. von Hippel, P. H., Revzin, A., Gross, C. A. & Wang, A. C. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 4808–4812.
15. Bruinsma, R. F. (2002) *Physica A* **313,** 211–237.
16. Hu, T., Grosberg, A. Y. & Shklovskii, B. I. (2005) *Biophys. J.* **90,** 2731–2744.
17. Gowers, D. M. & Halford, S. E. (2003) *EMBO J.* **22**, 1410–1418.
18. Nickell, C. & Lloyd, R. S. (1991) *Biochemistry* **30,** 8638–8648.
19. McKinney, K., Mattia, M., Gottifredi, V. & Prives, C. (2004) *Mol. Cell* **16,** 413–424.
20. Francis, A. W. & David, S. S. (2003) *Biochemistry* **42,** 801–810.
21. Harada, Y., Funatsu, T., Murakami, K., Nonoyama, Y., Ishihama, A. & Yanagida, T. (1999) *Biophys. J.* **76,** 709–715.
22. Kabata, H., Kurosawa, O., Arai, I., Washizu, M., Margarson, S. A., Glass, R. E. & Shimamoto, N. (1993) *Science* **262,** 1561–1563.
23. Guthold, M., Zhu, X., Rivetti, C., Yang, G., Thomson, N. H., Kasas, S., Hansma, H. G., Smith, B., Hansma, P. K. & Bustamante, C. (1999) *Biophys. J.* **77,** 2284–2294.
24. Gowers, D. M., Wilson, G. G. & Halford, S. E. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 15883–15888.
25. van Oijen, A. M., Blainey, P. C., Crampton, D. J., Richardson, C. C., Ellenberger, T. & Xie, X. S. (2003) *Science* **301,** 1235–1238.
26. Doyle, P. S., Ladoux, B. & Viovy, J. L. (2000) *Phys. Rev. Lett.* **84,** 4769–4772.
27. Halford, S. E. & Szczelkun, M. D. (2002) *Eur. Biophys. J.* **31,** 257–267.
28. Schurr, J. M. (1979) *Biophys. Chem.* **9,** 413–414.
29. Slutsky, M. & Mirny, L. A. (2004) *Biophys. J.* **87,** 4021–4035.
30. Ellenbogen, E. (1952) *J. Am. Chem. Soc.* **74,** 5198–5201.
31. Sakata-Sogawa, K. & Shimamoto, N. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 14731–14735.
32. Pauling, L. (1946) *Chem. Eng. News* **24,** 1375–1377.
33. Seksek, O. & Bolard, J. (1996) *J. Cell Sci.* **109,** 257–262.
34. Cappelli, E., Hazra, T., Hill, J. W., Slupphaug, G., Bogliolo, M. & Frosina, G. (2001) *Carcinogenesis* **22,** 387–393.
35. Kuttner, Y. Y., Kozer, N., Segal, E., Schreiber, G. & Haran, G. (2005) *J. Am. Chem. Soc.* **127,** 15138–15144.
36. Fromme, J. C. & Verdine, G. L. (2002) *Nat. Struct. Biol.* **9,** 544–552.
37. Gueron, M. & Leroy, J. L. (1995) *Methods Enzymol.* **261,** 383–413.
38. Bhattacharya, P. K., Cha, J. & Barton, J. K. (2002) *Nucleic Acids Res.* **30,** 4740–4750.
39. Stivers, J. T. (2004) *Prog. Nucleic Acid Res. Mol. Biol.* **77,** 37–65.
40. Banerjee, A., Santos, W. L. & Verdine, G. L. (2006) *Science* **311,** 1153–1157.
41. Thompson, R. E., Larson, D. R. & Webb, W. W. (2002) *Biophys. J.* **82,** 2775–2783.
42. Perrin, F. (1934) *J. Phys. Radium Ser.* **7,** 497–511.