Supplementary Information for

Intramolecular DNA Translocation by Human Uracil DNA Glycosylase: The Case of ssDNA and Clustered Uracils

Joseph D. Schonhoft and James T. Stivers

Methods:

Oligonucleotide Sequences:

Oligonucleotide substrates are named as follows, with the number indicating the spacing of the uracil sites and the superscript "ss" indicating single-stranded DNA; all other substrates were used in the duplex form. The letters F indicate a tetrahydrofuran abasic site mimic. Importantly, the single stranded substrates were designed to have no more than two possible adjacent Watson-Crick pairings as determined from the hybridization prediction program mFold (*1*).

$S20^{ss}$_90mer:

CAC AAT AAC ACA TAC A CTA A TCAT ACA TCA CAC AAA ACA U ACA AAA CAC AA CAC AAA ACAUACA AAA CAC ACT AATC C ACAC AC ATA ACA

$S40^{ss}$_90mer:

CAC AAT AAC ACA TAC A CTA A TCAT ACA TCA CAC AAA ACA U ACA AAA CAC AA CAC AAA ACA ACA AAA CAC AA CAC AAA ACA U ACA AAA CAC ACT AATC C ACAC AC ATA ACA

S5F_90mer:

GGT ATC CGC TGA AGT AGT CAC AAT TCC ACA CAA TGC TGA GGA ATC GA U AG F GA U AGC TAA GCT GAG GCA TAC AGG ATC AAT TGT CGA GCC

S11F_90mer:

GGT ATC CGCT AGT CAC AAT TC ACA CAATGC TGA GG A AT CGA U AG C F ATA F CGA U AGC TAA GCT GAG GCATAC AGG ATC AAT TGT CGA GCC

(F = Tetrahydrofuran abasic mimic)

S19F_90mer:

GGT ATC CGCT AGT CAC AAGT TC AATGC TGA GG A AT CGA U AG C F ATA F TGT F ATA F CGA U AGC TAA GCT GAG GCATC AGG AT TGT CGA GCC

(F = Tetrahydrofuran abasic mimic)

$1XU^{ss}$_90mer:

CAC AAT AAC ACA TAC ACT AAT CAT ACA TCA CAC AAA ACA U ACA AAA CAC AAC ACA AAA CAT ACA AAA CAC ACT AAT CCA CAC ACA TAA CA

Chase duplex (chDNA)

5′ - GCG GCC AAA F AA AAA GCG C

3′ - CGC CGG TTT A TT TTT CGC G

(F = Tetrahydrofuran abasic mimic)

Non-specific ssDNA (nsDNA^{ss})

CGC GTG TGC C – FAM

Michaelis-Menten kinetics of hUNG2 reaction on ssDNA

1XU^{ss} ssDNA was labeled at the 5′ end with ^{32}P using g^{32}P ATP and polynucleotide kinase. Concentrations of 1XU^{ss} from 0-800 nM were reacted with hUNG (4-15 pM) and at time points in the initial rate regime 4 μL aliquots were quenched with 5 μL 0.5 M NaOH. The quenched samples were then heated to 95 °C for 10 minutes to cleave abasic sites. Formamide loading buffer was added to 50% final concentration and substrate and product fragments resolved by denaturing gel electrophoresis. Band intensities were quantified and initial rates determined by linear regression. The resulting rates were normalized to enzyme concentration and plotted against substrate concentration. The resulting data was fitted to the Michaelis-Menten equation to determine V_{max} and the K_m (Supplementary Fig. S6).

Calculation of the mean square distance between target sites using the worm like chain model for ssDNA

For analysis of hUNG hopping, P_{trans} values were plotted against the mean square distance <r> of the uracil target sites as determined from the worm like chain model (*2*)(eq S1).

$$\langle r \rangle = \left\{ 2PL \left[1 - \left(\frac{P}{L} \right) \left(1 - e^{-\frac{L}{P}} \right) \right] \right\}^{-1/2} \qquad \text{(eq S1)}$$

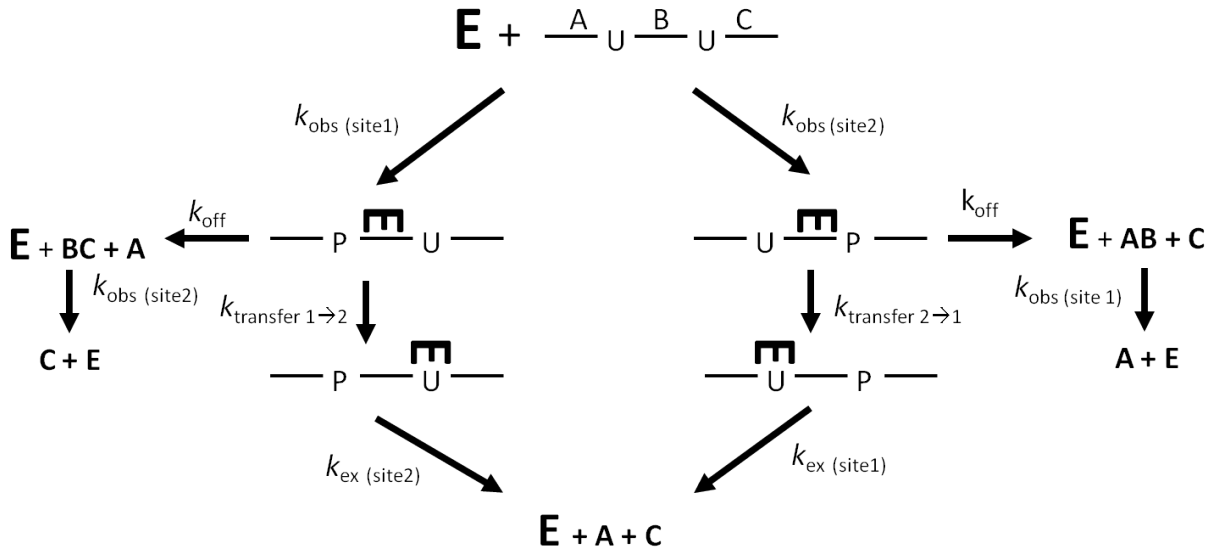
Where the P is the persistence length and L is the contour length of the chain. To estimate <r> for ssDNA we employed experimental parameters obtained from single molecule FRET experiments on ssDNA (*3*, *4*). At the salt concentrations employed here experimental estimates of P were between 2 and 3 nm, while the contour length for ssDNA is estimated as 0.5 – 0.7 nm. Error bars for the *x*-axis in Fig. 1e represent the maximum and minimum values for <r> using the above ranges for P and L.

Kinetic Modeling of the data for S19F

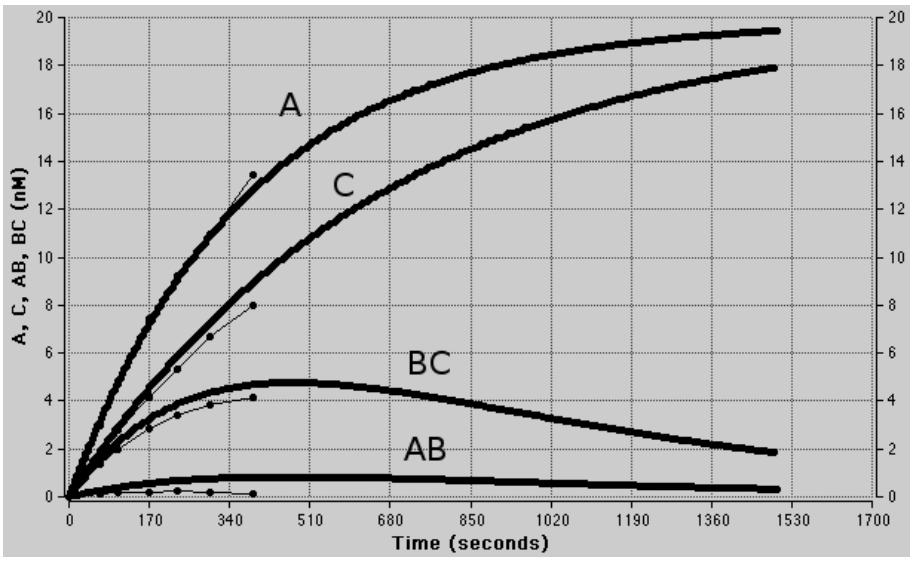
Mechanism 1 (below) was used to simulate the reaction timecourse for S19F. The concentrations of the DNA and enzyme were set to match those of the site transfer assay ([hUNG] = 50 pM, [DNA] = 40 nM). At these concentrations the rate-limiting step is release of the enzyme from the first site after uracil excision (*5*). These constants are all encapsulated within the rate constant k_{obs}. Once cleavage and release from the first product site has occurred the enzyme can transfer and find the second site, or it can fall off of the DNA. The exact

magnitudes of the rate constants following the initial rate-limiting step are not critical in this analysis, because it is the partitioning of the enzyme-F DNA complex between dissociation and transfer to the second site that determines the outcome of the simulation (i.e $P_{transfer} = k_{transfer}/(k_{transfer} + k_{off})$. For the simulation, we chose realistic values for the reaction steps following the initial cleavage at the first site: (i) the off-rate from the intervening DNA containing F sites (k_{off}) was set at 20 s^{-1} obtained from stopped flow measurements of enzyme dissociation from F site containing DNA (*5*), (ii) the rate of cleavage at the second step is set to the single turnover rate for uracil excision (k_{ex} = 100 s^{-1} at 25 °C (*6*) and estimated to be 200 s^{-1} at 37 °C).

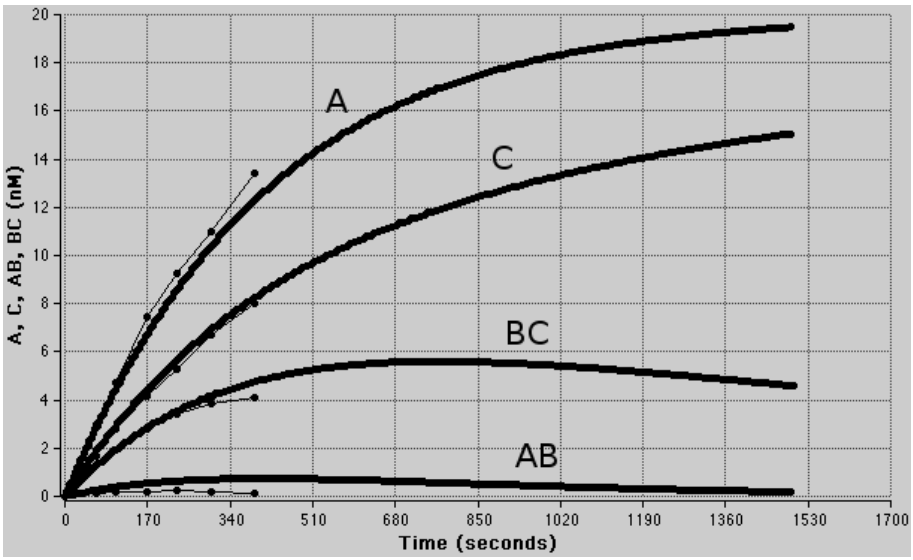
The results of the simulation I and II below show that the site preference observed for S19F can be attributed to either differences in cleavage rates at each site, differences in transfer rates, or a combination of both effects. In the simulations below, the raw data for S19F is plotted along with the simulated progress curves.



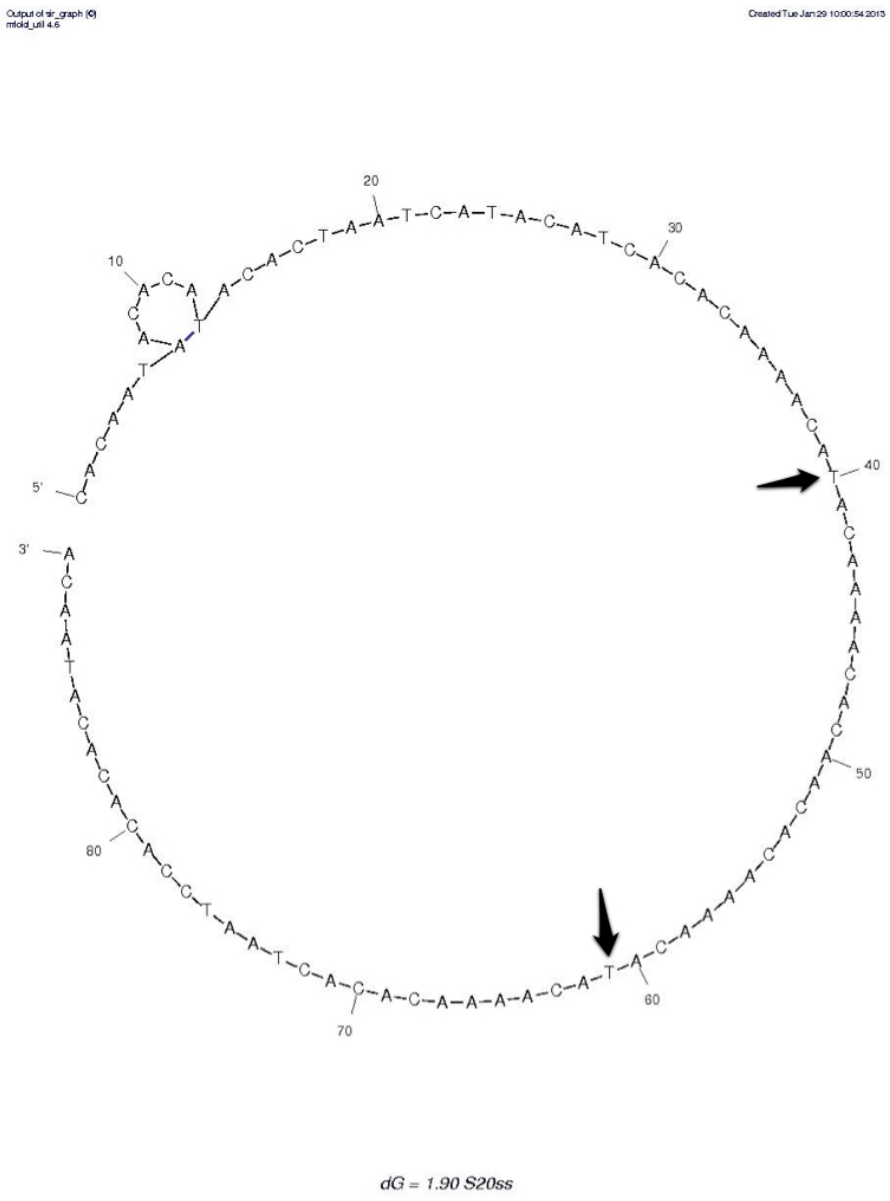
Mechanism 1: Site transfer mechanism used in numerical simulations. k_{obs} at both sites is the initial binding and cleavage steps of the hUNG with a uracil site. k_{off} is the off-rate for the intervening DNA between uracil sites and $k_{transfer}$ is the rate at which the enzyme transfers between uracil sites. k_{ex} is the rate of uracil excision once the enzyme has found the uracil site.



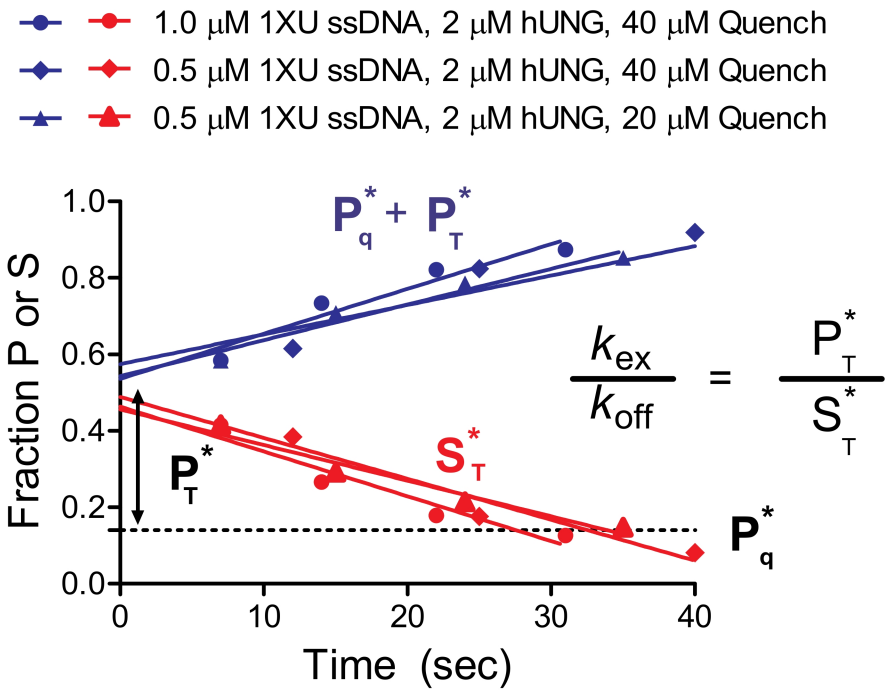
Simulation I: Differential transfer rates. We first tested a model involving differential rates of transfer between each site. In this simulation, all cleavage rates at sites 1 and 2 were set to be equal ratio (k_{obs} site 1 and 2 = 0.03 nM^{-1} s^{-1}), while the ratio of $k_{transfer2\rightarrow1}$ = 100 s^{-1} was set to be 5 times faster than $k_{transfer1\rightarrow2}$ = 20 s^{-1}.



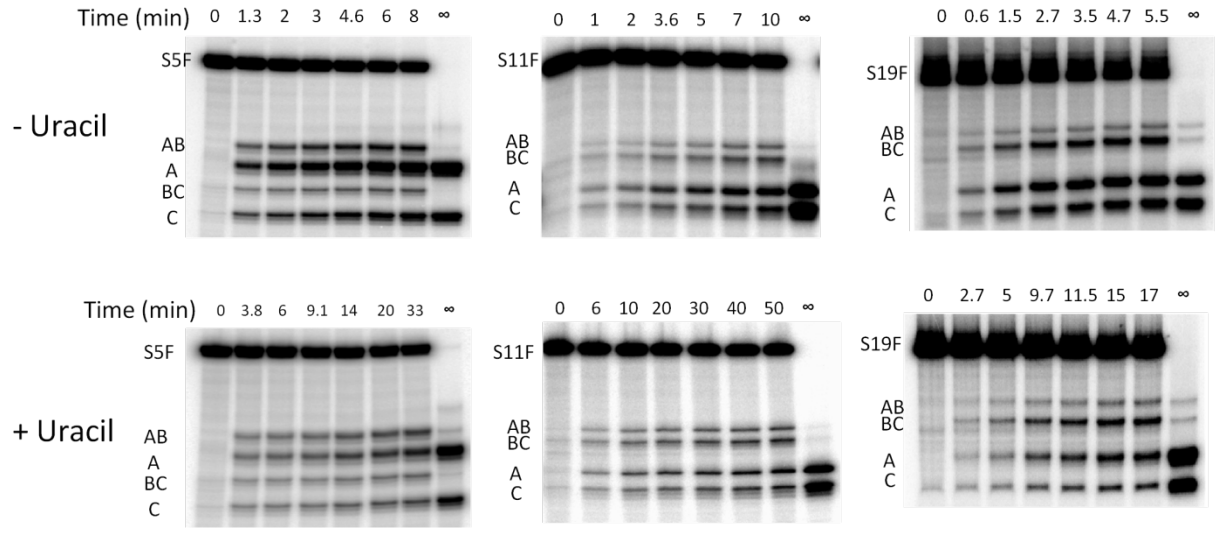
Simulation II: Differential rates of cleavage. In the second simulation we tested a model involving differential rates of cleavage between each uracil site. $k_{obs\ (site\ 1)}$ was set to be 0.05 nM^{-1} s^{-1}, $k_{obs\ (site\ 2)}$ was set to be 0.01 nM^{-1} s^{-1} and the transfer rates were set to be equal (20 s^{-1}).



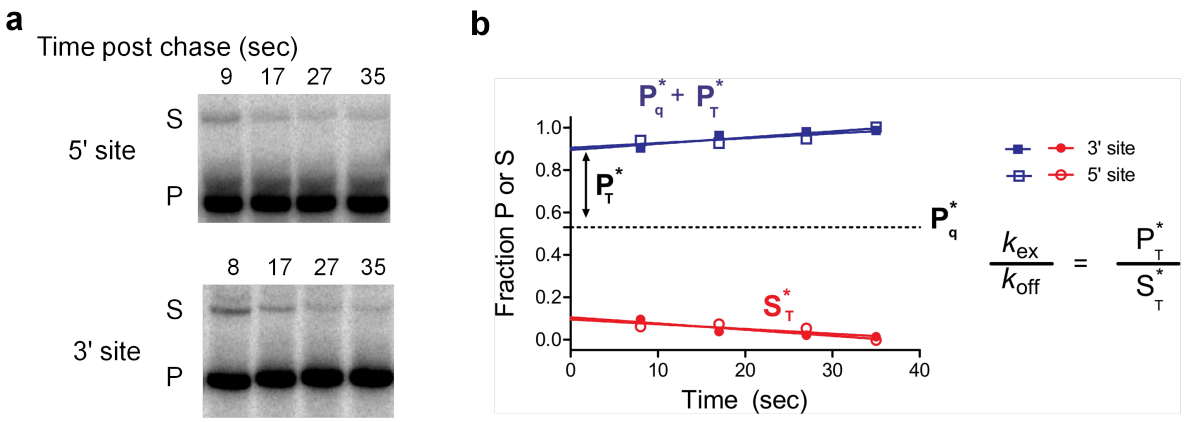
Supplementary Figure S1: Predicted lowest energy secondary structure (20 mM NaCl, 37 °C) as determined by the mFold web server (*1*) for $S20^{ss}$ showing the absence of any secondary structure based on WC pairing. Uracil positions are marked with the arrow and were changed to T for folding.



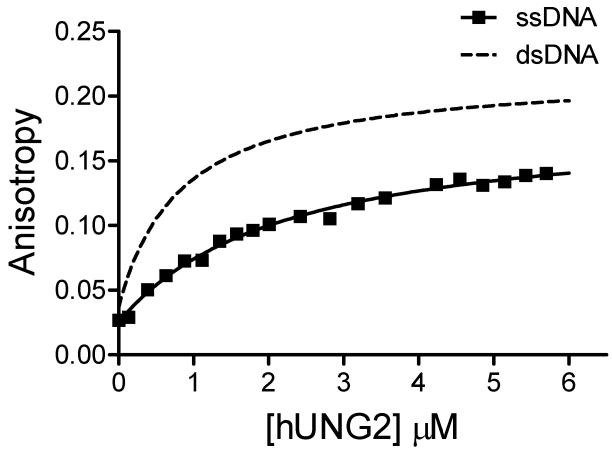
Supplementary Figure S2: Control experiments for determination of the excision efficiency (E) for ssDNA showing the independence of the DNA/hUNG ratio and also independence of the concentration of the quench (chDNA) indicating zero order trapping of the enzyme after dissociation from the substrate DNA. See Figure 2 for a more detailed description.



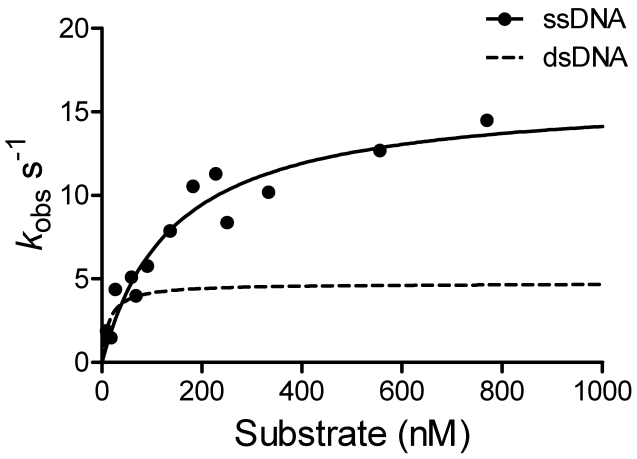
Supplementary Figure S3: Gel images for substrates S5F, S11F, and S19F (from left to right) in the absence (top row) and presence (bottom row) of uracil.



Supplementary Figure S4: Determination of the efficiency of uracil excision (E) for single uracil containing substrates analogous to S19F. The experiment is identical to the scheme depicted in Fig. 2a where hUNG was rapidly mixed with DNA substrate and chased at 2 milliseconds with 20 μM F site containing DNA (chDNA). See Fig. 2 legend and main text for detailed description. The data depicted here were determined using a final concentration of hUNG at 4 μM, and the substrate concentrations were 280 nM (shown above) or 140 nM. Identical results were achieved with each Enzyme/DNA ratio. The HCl quench to determine P_q* is shown in Fig. 6 as part of the single turnover measurements. The excision efficiency, $E = k_{ex}/(k_{off} + k_{ex}) = P_T^*/(P_T^* + S_T^*)$ for the 5′ site was E = 0.92 ± 0.12 and for the 3′ site E = 0.86 ± 0.04. The reported values are the average ± 1 SD.



Supplementary Figure S5: Non-specific binding of hUNG2 to ssDNA. hUNG2 was titrated into a cuvette containing 50 nM of a 10mer FAM labeled oligo and the anisotropy was recorded. Using a one-site binding isotherm, the K_D was determined as 2.0 ± 0.3 μM with a maximal anisotropy of 0.16 ± 0.01. The least-squares fit for non-specific hUNG2 binding for the identical DNA in the duplex form is depicted by the dotted line as determined in (*7*).



Supplementary Figure S6: Steady-state reaction kinetics for a ssDNA substrate containing a single uracil (1XU_90mer). The K_m and $V_{max}/[E_{tot}]$ were determined by fitting to the Michaelis-Menten equation: K_m = 141 ± 34 nM and $V_{max}/[E_{tot}]$ = 16 ± 2 s^{-1}. The errors are the standard error from the least squares fit to the data in Graphpad Prism 5. The fit for dsDNA under identical conditions is shown by the dotted line (*7*).

Supplementary Table 1: Comparison of extrapolation and initial rate methods for determination of site transfer properties[a].

Substrate	P_{trans} extrapolation	P_{trans} initial rates	P_{slide} extrapolation	P_{slide} initial rates
S5F	0.47 ± 0.04	0.49 ± 0.01	0.19 ± 0.02	0.23 ± 0.05
S11F	0.44 ± 0.05	0.46 ± 0.12	0.18 ± 0.05	0.23 ± 0.04
S19F	0.41 ± 0.04	0.46 ± 0.03	0.09 ± 0.05	0.12 ± 0.04

[a]For the intitial rates method the rate of formation was measured for each product (v_A, v_C, v_{BC}, v_{AB}) and P_{trans} was then calculated using eq 5 of the main text. For the extrapolation method, P_{trans}^{obs} was calculated using the band densities ([A], [C], [AB], [BC]) at each time point (eq 1 of the main text) and then linearly extrapolated to zero to obtain P_{trans}. Here both methods are compared in calculating site transfer probabilities for F containing substrates with uracil (P_{slide}) and without uracil (P_{trans}) under the standard reaction conditions (T=37 °C in 20 mM HEPES pH 7.5, 0.002% Brij 35 detergeant, 3 mM EDTA, and 1 mM DTT). For both methods identical results were obtained within error. The reported errors represent ± 1 s.d.

Supplementary References

1. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res. 31*, 3406–3415.
2. Mazur, A. K. (2006) Evaluation of elastic properties of atomistic DNA models., *Biophys J 91*, 4507–4518.
3. Murphy, M. C., Rasnik, I., Cheng, W., Lohman, T. M., and Ha, T. (2004) Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy., *Biophys J 86*, 2530–2537.
4. Chen, H., Meisburger, S. P., Pabit, S. A., Sutton, J. L., Webb, W. W., and Pollack, L. (2012) Ionic strength-dependent persistence lengths of single-stranded RNA and DNA., *Proc. Natl. Acad. Sci. U.S.A. 109*, 799–804.
5. Jiang, Y. L., Ichikawa, Y., and Stivers, J. T. (2002) Inhibition of uracil DNA glycosylase by an oxacarbenium ion mimic., *Biochemistry 41*, 7116–7124.
6. Parker, J. B., and Stivers, J. T. (2008) Uracil DNA glycosylase: revisiting substrate-assisted catalysis by DNA phosphate anions., *Biochemistry 47*, 8614–8622.
7. Schonhoft, J. D., and Stivers, J. T. (2012) Timing facilitated site transfer of an enzyme on DNA., *Nat. Chem. Biol. 8*, 205–210.