Supplemental Methods, Data and Figures for:

Alexandre Esadze¹, Gaddiel Rodriguez², Brian P. Weiser¹, Philip A. Cole¹ and James T. Stivers¹,*

¹Department of Pharmacology and Molecular Sciences and ²Department of Molecular Biophysics, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205-2185, USA
* To whom correspondence should be addressed.

Sequences of DNA substrates (prime mark indicates the complementary strand sequence).

U10
5' - (FAM) - ATG GC G ATC CGCT AGT CAC AAT TCC ACA CAA TGC TGA GGA ATC GA U AG CTA AT CGA U AGC TAA GCT GAG GCA TAC A A CG GTA - (FAM) - 3'

U10’
5' - TAC CGT TGT ATG CCT CAG CT T AGC TAT CGA TTA GCT ATC GAT TCC TCA GCA TTG TGT GGA ATT GTG ACT AGC GGA TCG CCA T - 3'

T10
5' - (FAM) - ATG GC G ATC CGCT AGT CAC AAT TCC ACA CAA TGC TGA GGA ATC GA T AG CTA AT CGA T AGC TAA GCT GAG GCA TAC A ACG GTA - (FAM) - 3'

T10’
5' - TAC CGT TGT ATG CCT CAG CT T AGC TAT CGA TTA GCT ATC GAT TCC TCA GCA TTG TGT GGA ATT GTG ACT AGC GGA TCG CCA T - 3'

U40
5' - (FAM) - ATG GC GGT CGC TCA CAA TT ATG CTG AGG AAT CGA U AGC TAA GTG AAT CTC TCA ACT AGC ACA TGG AAT GAA TCG A U AGC TAA GCT GAG GCA TAG C CG GTA - (FAM) - 3'

U40’
5' - TAC CGG CTA TGC CTC AGC TTA GCT ATC GAT TCA TTC CAT GTG CTA GGT AGA TTC ACT TAG CTA TCG ATT CCT CAG CAT AAT TGT GAG CGA CCG CCA T -3'

U80-OH
5' - (FAM) - ATG GC GGT ATC CGC TCA CAA TT ATG CTG AGG AAT CGA U AGC TAA GTG AAT CTC TCA ACT AGC ACA TGG AAT GAA TCG A U AGC TAA GCT GAG GCA TAG C CG GTA - (OH) - 3'

U80-phos
5' - (PO₄) - TC GAG CC ATA GAC GTA CAC TCA GTC CGC ACT AGC ACA TGG AAT GAA TCG A U AGC TAA GCT GAG GCA TAC A CG GTA - FAM - 3'

U80’
5' - TAC CGT GTA TGC CTC AGC TTA TCA TTC CAT GTG CTA GTG CGG ACT GAG TGT ACG TCT ATG GCT CGA CAC GAT GTG ACG TGA GAG ATT CAC TTA GCT ATC GAT TCC TCA GCA TTG TGT GGA ATT GTG AGC GGA TAC CGC CAT - 3'

HP18
5' - (FAM) - GCA UUA AGA AGA AG-(PEG)6-CUU CUU AAT TGC - (BHQ) - 3'

Supplemental Methods and Data

Generation of Hap1ihUNG2 cell line
After transfection with pCW57.1 and the packaging plasmids pRRE, PRSV-rev and pMD2.g (see Methods), the supernatant medium was harvested, treated with Turbo DNAse (2 units/mL) for 15 min. to remove residual plasmid DNA and then filtered through a 0.45 µm filter. The lentiviral particles were purified via ultracentrifugation at 25000 g (Optima XL-100K ultracentrifuge, Beckman) at 4 °C using a 20% sucrose bed (20% sucrose, 80% buffer NL [12 mM Tris HCl pH 7.5, 12 mM NaCl]). The virus particles were resuspended in 0.5 mL of NL buffer and serial dilutions were made in the range of 10⁻¹ to 10⁻⁷. One-hundred microliters of each dilution was used for spinoculation of Hap1UNG cells (2 h at 260 g and 37 °C) using a 96 well microplate format. The plating density was 8 x 10⁴ cells/well and the spinoculation media was DMEM-10 (DMEM + 1% Pen/Strep + 10% dialyzed, tetracycline free, FBS to avert residual induction of hUNG2 by trace levels of tetracycline). The spinoculation medium was replaced with fresh 200 µL DMEM-10 and the cells were cultured for another 48 h at 37 °C, 5% CO₂ and 80% humidity. The cells were released from the plate with 0.5% Trypsin-EDTA (Gibco) and washed twice with DMEM-10-P (DMEM-10 + 1 µg/mL puromycin). The cells were then cultured in T-175 flasks for an additional 9 days. By day 9, the cells transduced with viral dilutions in the range 10⁻¹ to 10⁻⁵ had generated dense colonies. Inducible expression of hUNG2 was tested by supplementing DMEM-10-P medium with varying concentrations of doxycycline (DMEM-10-P-D). Cell lines generated from 10⁻¹ to 10⁻⁵ viral dilution produced noticeable over-expression of hUNG2 as judged by western blotting and 10⁻² was selected as the working cell line.

The optimal concentration of doxycycline was 1 µg/mL and the best induction time was 6 h. Based on activity measurements of extracts (see main text), the level of hUNG2 produced at 6 h post-induction was 10 to 15 times greater than the endogenous expression in Hap1mt. No hUNG2 expression or hUNG2 activity was detected in absence of doxycycline based on Western blotting and activity measurements (Fig. 2A, 2B, 2C), indicating tight regulation of hUNG2 expression by the TET-on system. This cell line was cultured in large scale in DMEM-10-P and cryopreserved at -135 °C in 1 mL aliquots (5 x 10⁶ cells/mL) in 90% DMEM-10-P containing 10% DMSO.

Western blot based detection of hUNG2 from Hap1 cell lines
We cultured Hap1 cells in an appropriate medium (DMEM-10 or DMEM-10-P) in T-75 flasks for 48 h prior to inducing the Hap1ihUNG2 cells with DMEM-10-P-D medium over 16 h or just continuous incubation in DMEM-10 for other cell lines. Extracts were obtained from 2 x 10⁶ cells quantified as described in the main text. Ten µg of each cell extract (20 µl) was mixed with 6 x SDS loading buffer (Sigma), heated for 10 min. at 95 °C and loaded on a 4 - 12 % gradient SDS denaturing gel (Invitrogen). The gel was run for 90 min. at 150 V in 1x MES-SDS running buffer (Invitrogen). The gel was removed from the cassette, rinsed with water, the bottom 1 mm thick portion was carefully removed with razor blade and the gels were transferred onto Immun-Blot PVDF pre-cut membranes (Bio-Rad) in a 1-step transfer buffer (Thermo) using a Pierce Power Blot Station (Thermo). The successful transfer of hUNG2 (≈ 35 kDa) was confirmed by full transfer of the 35 kDa and 40 kDa pre-stained Rainbow markers (Sigma). The membrane was then blocked over 36 h with slow shaking in blocking buffer TBST (50 mM Tris-HCl pH
7.6, 150 mM NaCl, 0.1% Tween-20) + 5 % BSA at 4 °C. Extended blocking was required as the hUNG2 antibody is notorious for high non-specific background. The membrane was then stained with primary antibody for hUNG2 (ab62520, Abcam) and α tubulin loading control (ab15246, Abcam) with 1:10000 dilution in TBST for 24 h at 4 °C with slow shaking. The membrane was then washed in TBST for 30 min. three times with slow shaking and stained with 1:2000 dilution of the secondary antibody ab97080 (Abcam) in TBST for 1 h at room temperature with shaking. The membrane was then washed two times in TBST for 10 min. and developed with SuperSignal West Pico Chemiluminescent Substrate kit (Thermo) for 1 min. The membrane was then mounted on a clean glass plate and imaged on GelDock (Bio-Rad) in chemiluminescence signal acquisition mode, with automatic signal accumulation time. The image was then processed using ImageJ software. Western blots for hUNG2 expression in Hap1ΔhUNG and Hap1ΔhUNG2 (induced/uninduced conditions) are shown in Fig. 2A.

Purification of B-hUNG2 from bacterial expression system

Full-length hUNG2 was overexpressed and purified from a bacterial protein expression system as previously described (B-hUNG2)(4). Briefly, hUNG2 was expressed with an 8xHis-tag and the SUMO protein fused to its N-terminus. After Ni²⁺ column purification, the His-SUMO tag was removed with the SUMO protease, and B-hUNG2 was further purified chromatographically with size exclusion chromatography. B-hUNG2 protein was >95% pure based on SDS-PAGE and Coomassie blue staining, and the enzyme was flash frozen in liquid nitrogen and stored at -80°C as 40 μL single use samples.

Measurements of hUNG2 activity in cell extracts

We measured hUNG2 activity in Hap1 cell extracts using a continuous fluorescence based assay that employed substrate HP18. This substrate contains six uracils paired with adenines forming a hairpin that brings together FAM fluorophore with the BHQ (Black Hole) quencher. The close proximity of quencher-fluorophore pair reduces the fluorescence of the probe ≈ 10-fold. In the presence of hUNG2, the uracils are removed and the hairpin unfolds, producing a large increase in fluorescence. Using LS buffer and 10 μg of cell extract (see main text) we measured the UNG activity present in Hap1ΔhUNG2, Hap1ΔUNG and Hap1wt cells (Fig. 2B). The complete absence of activity for Hap1ΔUNG indicated that no other glycosylase activity was responsible for the excision of uracils from the U/A base pairs. Thus, the uracil excision activities observed in Hap1wt and Hap1ΔhUNG2 cell lines can be assigned to hUNG2. This conclusion was further supported by the finding that the uracil excision activity in extracts was completely inhibited by the potent UNG specific inhibitor UGI (Fig. 2C). (1)

Immunofluorescence staining of hUNG2 expressed in Hap1ΔhUNG2 cells

Hap1ΔhUNG2 cells were grown on 22 x 22 mm glass cover slips (Corning) in 6 well plates at 5 % confluence to form well isolated and defined colonies. The cells were then either induced with 2 mL of DMEM-10-P-D or cultured in DMEM-10-P in parallel over 6 h. Two mL of 4% Paraformaldehyde-PBS solution was slowly added to each well as to not perturb cells and incubated at 37 °C, 5% CO₂, 80% humidity for 15 min. and then another 30 min. at room temperature to fix the cells. From this step forth 3 times 10 min. PBS wash was implemented before and after each procedure and slow rocking was used during all steps including washes. All procedures are done at room temperature. Cells were incubated at room temperature in blocking buffer (3% BSA 0.1% Triton X 100 in PBS) for 30 min. stained with anti-UNG primary antibody (TA503563, OriGene) 1:2000 in blocking buffer and incubated with secondary Alexa Fluor 488 antibody (A11011, Invitrogen) 1:4000 in blocking buffer for 1 h. To stain with the DAPI, 100 nM PBS solution of DAPI was added and slowly rocked for 30 min. Mounting slides were vigorously cleaned of impurities by soaking in 5 M NaOH for 10 min. followed by ddI water and spectroscopic grade methanol rinse. A droplet of Fluorsave Reagent (milipore) was used to mount the cover slips with stained cells onto the slides. Before the mounting lint free wax paper was used to remove any residual cells from the reverse of the cover slip. Mounting was done for 30 min. at room temperature in the dark after which the mounting
medium hardened and samples were stored at 4 °C in the dark until use. The cells were imaged with 20x magnification on a Olympus fluorescence microscope at the JHMI Microscopy Core Facility. The raw images were processed using ImageJ software. hUNG2 was exclusively localized in the nucleus as judged by co-localization with DAPI staining in induced Hap1^{hUNG2} cells (Fig. 2D). In contrast, no staining was observed in uninduced Hap1^{hUNG2} cells (Fig. 2D).

Detection of Successful cellular delivery of probe DNA via Confocal Microscopy. Hap1^{hUNG2} induced or non induced cells were grown on 15 mm glass cover slips in 6 well plates to form a 60 - 70 % confluent monolayer and then transfected with U10 DNA probe (see main text). The cover slip was then taken out and vigorously washed with appropriate medium and then incubated in 1 mL of appropriate medium + 2 units of Turbo DNase. The cover slips were then washed vigorously with PBS and reverse side was carefully cleaned with cotton swabs to remove any contaminating cells. Staining for DAPI was done at room temperature, overnight in the dark with ProLong Gold antifade reagent with DAPI (life technologies). The Hap1 cells were immediately imaged on the second day on ZEISS LSM700 confocal microscope under 40 fold magnification with oil immersion, in JHMI microscope core facility. The images were processed via ImageJ with application of linear brightness enhancement for the DAPI channel as some of the nuclear staining gave weak signals. In all cases the green fluorescence coming from the transfected U10 probes was not excluded from the volume occupied via nuclear fraction of the cell (blue stain) indicating a uniform delivery of the transfected probe DNA. We note that for the non induced cell lines the transfection was more uniform while for induced cell lines we observed more speckled distribution.

Cellular internalization and stability of duplex DNA in transfected Hap1 cells and cell extracts

The internalization of transfected DNA was confirmed by a DNase protection assay. Hap1^{hUNG2} cells were transfected for 4 hours and then induced with Dox for 6 h. Prior to harvesting the intracellular DNA, TB-DNase was added to half of the cell culture for 15 min (2 units/ml), followed by washing with media and then DNA extraction. Compared to the fraction of the culture that was not treated with DNase, only a 20% reduction in recovered DNA was measured, which is within the variability of the recoveries in the DNA extraction protocol (see main text). This result establishes that our estimates of intracellular DNA concentration are not biased by extracellular DNA that persisted after transfection. Nevertheless, our standard protocol involves treatment with DNase prior to cellular DNA extraction to reduce any possible contribution from external probe DNA (Fig. S1A).

We compared the integrity of intracellular transfected duplex U10 probe DNA at the time of induction and then 6 h later (Fig. S1B and S1C). DNA was extracted at both times as described in the main methods and equal volumes of the extracts were loaded on a denaturing polyacrylamide gel for visualization and image quantification. Using data from two independent experiments, we summed all of the fluorescence intensities of the individual bands that were detected in each sample (full-length substrate and/or double excision bands A and C), and found that the 6 h post-induction sample had an average decrease in total fluorescence intensity of about 20% compared to the zero-time reference. Although this is not statistically different than the variability in the extraction yields from sample to sample, it establishes that there is no large-scale degradation of the DNA occurring (see text for further discussion). We note that relatively long stability of double-stranded, fluorophore-labeled DNA in mammalian cells has been previously reported (3).

In a further experiment, we tested the time-dependent stability of U10 using isolated Hap1^{hUNG2} cell extracts (Fig. S1D). In these experiments 10 μg/μL of extract was reacted with 100 nM U10 in total volume of 100 μL over six hours using our standard in vitro translocation buffer containing 20% PEG8K and supplemented with 10 mM Mg^{2+} (pH 7.5 and 37 °C). The additional value of this experiment is that
the probe DNA is present in all the possible intermediate states throughout the experiment and analysis of individual and net intensities would allow determination if any intermediate species were preferred targets for degradation. We found that no appreciable degradation occurred over 6 h (Fig. S1E).

Derivation of the equations describing the time averaged concentration of hUNG2 during induction of Hap1\textsuperscript{hUNG2} cells
Meaningful interpretation of the translocation measurements requires that single hit conditions are being followed rather than multiple enzyme molecules acting on the same substrate. A single hit regime is strongly indicated by the combined observations that $P_{\text{trans}}$ is independent of reaction extent in the range 3 to 25%, and the translocation efficiency decreases with site spacing, eventually becoming distributive (Fig. 5C). These observations strongly support single hit kinetics and require that $[\text{Probe}] >> [\text{Free hUNG2}]$.

To investigate this further, we estimated the time averaged concentration of hUNG2 during the 6 h induction of Hap1\textsuperscript{hUNG2} cells. The estimation of intracellular enzyme concentration was based on measurements of in vitro reaction velocities using B-hUNG2 as the activity standard. In this experiment, extracts were processed at several times over the 6 h induction time course (Fig. 2B). The fractional maximal velocity reached at each induction time point $[\text{frac } v(t)]$ reflects the amount of hUNG2 that had accumulated over that time interval as compared to the maximal steady-state level at 6 h (i.e. $\text{frac } v(t) = v_v/v_{6h} = k_{\text{obs}}[E]/k_{\text{obs}}[E]_{\text{max}} = [E]/[E]_{\text{max}}$). Here, $k_{\text{obs}}$ is the intrinsic enzymatic activity of H-hUNG2 under the given conditions of the activity assay. For integration of the hUNG2 concentration over the induction time course, a simple two-state sigmoidal function was used (eq S1), where $[E]_t$ was plotted against time ($t$) (Fig. S4) and $\alpha$ and $\beta$ are fitting parameters.

$$\frac{[E]_t}{[E]_{\text{max}}} = \frac{1}{1 + \alpha \cdot 10^{-7t}} \quad \text{(S1)}$$

The product concentration with respect to time $P(t)$ can therefore be expressed as (eq S2):

$$P(t) = k_{\text{obs}} \cdot [E]_{\text{max}} \cdot \int_0^t \frac{1}{1 + \alpha \cdot 10^{-7t}} \, dt \quad \text{(S2)}$$

Following standard procedures for integration via substitution, we obtain the following expression for $P(t)$ (eq S3):

$$P(t) = k_{\text{obs}} \cdot [E]_{\text{max}} \cdot \left(\frac{(\alpha - \beta)t + 10^{(\alpha - \beta)t} + A}{\ln 10} \right)_0^6 \quad \text{(S3)}$$

where $A$ is an integration constant and the terms $\alpha = 3.5$, $\beta = 0.84$ are determined from non-linear regression fitting of the data in Fig. S4 to eq S1. The calculated value of the integral term is 1.84 h.

To convert the time integrated concentration of hUNG2 into an equivalent steady-state concentration that would produce the same amount of product over the same time period we calculate the conversion coefficient, $\zeta$, to be 0.31:

$$\zeta = \left(\frac{(\alpha - \beta)t + 10^{(\alpha - \beta)t} + A}{\ln 10} \right)_0^6 \quad \text{(S4)}$$

The time averaged nuclear concentration of $\langle [\text{hUNG2}] \rangle$ is then obtained from eq S4:
\[ \langle [\text{hUNG2}] \rangle = \frac{1}{N} \cdot \frac{V}{\frac{4}{3} \pi R^3} \cdot \gamma \cdot s \cdot f^{-1} \cdot \zeta \tag{S5} \]

Here, \( N \) is the number of extracted cells (2 \( \times 10^6 \)) used in our fluorescence based activity assay, \( \pi = 3.14 \), \( R \) is nuclear radius \( \sim 7 \) µm (based on fluorescence microscopy imaging)(Fig. 2D), \( \gamma \) is the ratio of the initial velocity (\( \nu_{\text{obs}} \)) obtained with the extract compared to the H-hUNG2 standard under the given in vitro conditions (\( \nu_{\text{obs}} = k_{\text{obs}}[E]_{\text{tot}} \)), \( s \) is the concentration of the hUNG2 standard (300 pM), \( V \) is the volume of the reaction cell (200 µL), and \( f \) is the fraction of the total cell extract used in the activity measurement. The term \( \frac{1}{N} \cdot \frac{V}{\frac{4}{3} \pi R^3} \cdot \gamma \cdot s \cdot f^{-1} \) in eq S5 is \([E]_{\text{max}}\). Thus, the time average concentration of total hUNG2 in Hap1<sup>hUNG2</sup> cells is \([E]_{\text{max}} \cdot \alpha = 10 \) µM \( \times 0.31 = 3 \) µM.

The time average copy number of total hUNG2 in the cell nucleus can then be calculated from its time average concentration using eq S6:

\[ \langle [\text{hUNG2}] \rangle \text{ copies} = C \cdot \frac{4}{3} \pi R^3 \cdot N_a \tag{S6} \]

Here \( N_a \) is Avogadro's constant (6.62 \( \times 10^{23} \) particles/mol), \( C \) is 3 µM from eq S6 and \( R \) is the nuclear radius (\( \sim 7 \) µm). Using eq S6, the time averaged copy number is \( 3 \times 10^5 \) copies.

**Calculation of the DNA probe concentration in Hap1 cells**

Estimation of intracellular DNA probe concentration can be directly obtained from eq S7:

\[ [\text{Probe DNA}] = \frac{1}{N} \cdot \frac{V_{\text{sample}}}{\frac{4}{3} \pi R^3} \cdot I \cdot s \tag{S7} \]

Here \( V_{\text{sample}} \) is the total volume of the extracted DNA (typically 50 to 100 µL), \( R \) is the cell radius (about 8 µm), \( I \) is ratio of the fluorescence intensities of the extracted DNA as compared to a standard of known concentration, \( s \) is the concentration of the standard (10 - 25 nM) and \( N \) is the number of cells extracted (\( \sim 2 \times 10^7 \)). This analysis gives a cellular probe concentration of about 0.2 µM corresponding to a copy number of about 20,000 per cell.

**Estimation of \([\text{hUNG2}]_{\text{free}}\) inside cells**

Using the time average \( \langle [\text{hUNG2}] \rangle = 3 \) µM as calculated above, and the observation that \( \approx 25\% \) of the 200 nM intracellular substrate concentration is consumed over a 6 h induction time (50 nM, or about 5,000 uracils removed per cell), we can estimate the fraction of the total enzyme present that is free to perform catalysis (i.e. not bound to DNA, proteins or other possible inhibitors). For this estimate we use our previously determined rate constant (\( k_{\text{obs}} = 60 \text{ h}^{-1} \)) for reaction of hUNG2 in the presence of 20% PEG8K, 150 mM K*, and 1 mM salmon sperm competitor DNA (concentration in DNA nucleotide monomers).\( ^5 \) These conditions provide our best estimate of the rate in the intra-nuclear milieu. Thus for \([P]_{\text{inh}} = 50 \text{ nM} \) and \( t = 6 \) h, \( \nu_{\text{obs}} = 8 \text{ nM h}^{-1} \). Accordingly, \( \langle [\text{hUNG2}] \rangle_{\text{free}} = \frac{\nu}{k_{\text{obs}}} = 8 \text{ nM h}^{-1}/60 \text{ h}^{-1} = 140 \) pM. Comparing the \( \langle [\text{hUNG2}] \rangle_{\text{free}} \) with the total \( \langle [\text{hUNG2}] \rangle \) indicates that only one in 20,000 enzyme molecules is free to react (\( \sim 15 \) copies per nucleus). Although there is uncertainty in estimating the free \( \langle [\text{hUNG2}] \rangle_{\text{free}} \) using the above approach, it is indisputable from the slow reaction of the intracellular probe DNA and the distributive site excisions at longer site spacings, that the effective concentration of the active enzyme in the nucleus is very low. The estimation above is consistent with these qualitative observations.

**Correction of apparent \( P_{\text{trans}} \) values to extent reaction**
Equation 1 in the main text defines an apparent value for $P_{\text{trans}}$ that increases depending on the extent of reaction because the A and C products progressively increase with time:

$$P_{\text{trans}} = \frac{([A] + [C]) - ([AB] + [BC])}{[A] + [C] + [AB] + [BC]}$$

Thus, the true value of $P_{\text{trans}}$ is only obtained when the extent reaction is infinitely small. At all other times, the $P_{\text{trans}}$ measurement is increased by secondary encounters of the enzyme with the fraction of substrate molecules that had been cleaved once in an initial encounter with enzyme. This reduces the AB and BC band intensities and increases A and C band intensities. We have generally found that the apparent $P_{\text{trans}}$ values increase approximately linearly in the range of zero to forty percent reaction with a very gentle slope (4.6-8). Accordingly, measurements done under conditions of <25% reaction provide reasonably good estimates of the true value of $P_{\text{trans}}$. Because the cellular $P_{\text{trans}}$ measurements can only be practically obtained at a single time point and percent reaction, we wished to derive a mathematical correction that would allow estimation of the true value of $P_{\text{trans}}$.

**Derivation.** We define $P'_{\text{trans}}$ as the true corrected value of the apparent $P_{\text{trans}}$ that is measured over some reaction time ($t$), and $\varrho$ as the correction factor that converts $P_{\text{trans}}$ into $P'_{\text{trans}}$ (eq S8):

$$P'_{\text{trans}} = P_{\text{trans}} \cdot \varrho$$

The correction factor $\varrho$ is the fraction of substrate molecules over a given reaction time that have come into contact with the enzyme only once. Calculating $\varrho$ requires knowledge of the instantaneous probability that any enzyme molecule re-encounters a singly-excised substrate $[P(t_1)]$. This probability is simply equal to the concentration of the *singly-reacted substrate molecules* relative to *unreacted substrate* and the doubly-excised substrate (i.e. substrate where both sites were excised in a single encounter) at time $t_1$. However, such treatment requires an accurate population measurement of the singly-reacted substrate over a period of time, particularly in the early stages of reaction, which is not easily obtained. Instead, we make the approximation that $P(t_1)$ can be replaced by the total substrate molecules reacted (single and double cleavage products) and then validate this approximation against actual data. The probability of single encounter with the substrate at a time point $t_1$ and is given by eq S9:

$$P(t_1) = \frac{F(t_1)}{F(\infty)}$$

Here $F(t_1)$ is the amount of singly-excised substrate at time point $t_1$ and $F(\infty)$ is the maximum amount of product attainable (two times the substrate concentration for two site substrates). Thus, $P(t_1)$ is simply the fraction reaction at $t_1$, reflecting that the encounter of the enzyme with a substrate that has already reacted once and is equal to its relative concentration in the mixture. Accordingly, the probability of encountering a substrate that has yet to react is given by eq S10:

$$\varrho_1 = 1 - \frac{F(t_1)}{F(\infty)}$$

The parameter $\varrho_1$ is the contribution of first-time encounters to $P_{\text{trans}}$ at time point $t_1$, and correction factor $\varrho$ is obtained by integration over time as shown graphically in Figure S5A and B. Thus, when the reaction progresses for a period of time $t_1$ to $t_2$ (eq S11):

$$\varrho = \int_{t_1}^{t_2} \frac{F(t)}{F(\infty)} \, dt = \int_{t_1}^{t_2} \frac{[1 - F(t)]}{t_2 - t_1} \, dt$$

(S11)
Dropping $F(\infty)$ from the integral because it is a constant equal to one, setting $t_1 = 0$, and inserting eq S11 into eq S8 gives (eq S12):

$$P'_\text{trans} = P_{\text{trans}} \cdot \frac{\int_0^{t_2} [1 - F(t)] \, dt}{t_2}$$  \hspace{1cm} (S12)

**Validation of Eq S12**

To establish the validity of the above method we compared the calculated $P'_\text{trans}$ values with the standard linear correction method using published data with B-hUNG2 on a substrate with 20 bp uracil spacing in the presence of 20% PEG8K (Fig. S6A)(6). For this data, the $P'_\text{trans}$ value from linear extrapolation was $0.46 \pm 0.05$. To calculate $P'_\text{trans}$ using eq S12, we need an appropriate function for $F(t)$, which is arbitrary as long as the data are fit reasonably well. In this case we choose to fit the fraction reaction data corresponding to the experimental $P_{\text{trans}}$ measurements in Figure S6A to an exponential (eq S13):

$$F(t) = 1 - e^{-k \cdot t}$$  \hspace{1cm} (S13)

Where $k$ is the fitting parameter, $t$ is time and $F(\infty) = 1$ and $F(0) = 0$. Inserting eq 13 into eq 12, and performing integration we obtain eq S14, which provides a correction to $P_{\text{trans}}$ measured over any given time period.

$$P'_\text{trans} = P_{\text{trans}} \cdot \frac{\int_0^{t_2} [1 - (1 - e^{-k \cdot t})] \, dt}{t_2} = P_{\text{trans}} \cdot \frac{\int_0^{t_2} e^{-k \cdot t} \, dt}{t_2} = P_{\text{trans}} \cdot \frac{\int_0^{t_2} e^{-k \cdot t} \, dt}{t_2} \bigg|_{t=0}^{t=t_2}$$  \hspace{1cm} (S14)

The corrected $P'_\text{trans}$ values obtained using eq S14 are listed in Table S1 for each reaction extent shown in Figure S6A. The corrected values are in excellent agreement with the $P'_\text{trans} = 0.46 \pm 0.05$ obtained from linear extrapolation of the data (Fig. S6A).

**Correction of cellular $P_{\text{trans}}$ values**

To correct the cellular $P_{\text{trans}}$ values measured for U10, U40 and U80, we fit the fraction reactions to a sigmoid function (eq S1). This function was substituted into eq S12 and provided the correction factors that are reported in Table S2.
Supplemental Figure S1. Internalization and persistence of double Fam-labeled probes in cells. All experiments are done with U10 substrate. (a) The substrate DNA is internalized in Hap1 cells after transfection. Denaturing polyacrylamide gel analysis of the U10 substrate DNA recovered from Hap1\textsuperscript{inhUNG2} cell extracts under conditions where the cells were treated (or not) with Turbo DNase (TB-DNase) prior to cell lysis. Protection of the full-length DNA and the uracil excision product bands indicates that the DNA is internalized in the cells. (b) Stability of U10 duplex DNA in transfected Hap1\textsuperscript{inhUNG2} cells. Following standard transfection and DNA extraction procedures (see Methods), DNA samples collected at the time of induction and 6 h post-induction were analyzed on 14% denaturing polyacrylamide gel. Markers M1 and M2 are full-length and fully digested U10 probes, respectively. As shown, duplicate experiments were performed. (c) Quantification of total DNA Fam label retained at zero and 6 h post-induction. (d) Stability of U10 substrate DNA in Hap1\textsuperscript{inhUNG2} cell extracts. Reactions contained 10 µg/µL cell extract and 100 nM U10 in a standard translocation reaction buffer containing 20% PEG8K and supplemented with 10 mM MgCl\(_2\) (pH 7.5 and 37 °C). Samples were analyzed on 14% denaturing polyacrylamide gel and imaged. (e) Quantification of total U10 fluorescent label retained over a 6 h incubation with cell lysate. The y-axis is the sum of the intensities of all detectible bands in each lane. Over the length of the experiment no appreciable degradation was detected. Experiments were performed in triplicate and the bars represent standard errors of measurements.
Supplemental Figure S2. U40 probe is in the duplex form after isolation from Hap1<sup>HUNG2</sup> cells. Duplex probe U40 was isolated from cells at 6 h post-induction as described and its size compared with that of the pure duplex standard (ds) and single-stranded form (ss). The DNAs were resolved using a native 8% polyacrylamide gel and the emitted FAM fluorescence was imaged.
Supplemental Figure S3. Transfecting Hap1 cell lines with T10 DNA does not produce the product bands observed with the U10 substrate DNA. The extracted T10 DNA (which lacks U/A base pairs) was analyzed by denaturing polyacrylamide gel electrophoresis with or without post-extraction treatment of UNG and base/heat.
Supplemental Figure S4. The induction time course of hUNG2 in Hap1^{hUNG2} cells using 1 µg/mL of doxycycline. The experimental conditions are the same as in Figure 1B. Each time point is the average of three individual in vitro velocity measurements reflecting the enzyme concentration present at that time ([E]_t). The observed velocities were normalized to the maximal velocity at t = 6 h (v_t = k_{obs}[E]/k_{obs}[E]_{max}), where the constant k_{obs} is the observed rate constant under the given reaction conditions. The curve is the best fit to eq S1. The concentration [E]_{max} was calculated based on the amount of purified B-hUNG2 that gave the same velocity.
Figure S5. Graphical representation of $\varrho$ and $\varrho_1$ values. (a) The parameter $\varrho_1$ is the contribution of first-time encounters to $P_{\text{trans}}$ at time point $t_1$, and correction factor $\varrho$ is obtained by integration over time. (b) when the reaction progresses for a period of time $t_1$ to $t_2$, $\varrho$ is the integrated contribution to encountering a substrate for the first time during this period (green dashed area) divided over the time span (eq S12).
**Figure S6.** Linear extrapolation and new analytical method for correction of apparent $P_{\text{trans}}$ values to the true value ($P'_{\text{trans}}$). (a) Linear extrapolation method to obtain $P'_{\text{trans}}$ values. The line is drawn through the first two data points. The previously published data was from in vitro measurements of hUNG2 using substrate U20 in the presence of 20% PEG8K (6). The extrapolated value is $P'_{\text{trans}} = 0.46 \pm 0.05$. (b) The data in panel a were fit to an exponential function (eq S13) to obtain the parameter $k = 0.3$. This exponential function was used for correction of $P_{\text{trans}}$ to extent reaction using eq 12. The output values for $P'_{\text{trans}}$ and $\varphi$ as function of reaction extent and time are listed in **Table S1**.
Figure S7. Fitting cellular reaction extents for substrates U10, U40 and U80 using eq S1. The use of the hyperbolic function is warranted due to measured time dependence of hUNG2 induction. (a) U10 time course, (b) U40 time course, (c) U80 time course. These optimized functions were substituted into eq 12 to obtain the corrected $P_{\text{trans}}$ values for each substrate reported in Table 2.
Table S1. Correction of apparent $P_{\text{trans}}$ values at various fraction reactions using eq S14\textsuperscript{a}

| Fraction reaction | 0.26 | 0.34 | 0.56 | 0.72 | 0.82 | 0.9 | 0.92 |
|-------------------|------|------|------|------|------|-----|------|
| $P_{\text{trans}}$ | 0.57 | 0.68 | 0.73 | 0.78 | 0.8  | 0.87| 0.9  |
| $\varrho$         | 0.86 | 0.75 | 0.66 | 0.59 | 0.52 | 0.47| 0.43 |
| $P'_{\text{trans}}$ | 0.49 | 0.51 | 0.48 | 0.46 | 0.42 | 0.41| 0.4  |

\textsuperscript{a} For comparison, the linearly extrapolated $P'_{\text{trans}} = 0.46 \pm 0.05$ (see Fig. S6A).

Table S2. Correction of cellular $P_{\text{trans}}$ measurements to extent reaction\textsuperscript{a}

| Fraction reaction | U10          | U40          | U80\textsuperscript{b} |
|-------------------|--------------|--------------|-------------------------|
|                   | 0.23 ± 0.04  | 0.31 ± 0.04  | 0.21 ± 0.06             |
| $P_{\text{trans}}$ | 1            | 0.6 ± 0.05   | 0.12 ± 0.02             |
| $\varrho$         | 0.94         | 0.9          | 0.9                     |
| $P'_{\text{trans}}$ | 0.94         | 0.54 ± 0.05  | 0.11 ± 0.02             |

\textsuperscript{a} A sigmoidal function (eq S1) was used in eq S12. \textsuperscript{b}Measurements were done in biological duplicate at 5 and 5.5 h post induction, giving average fraction reaction of 17 and 26%. Reported value is the average of these values.

References

1. Bennett, S.E. and Mosbaugh, D.W. (1992) Characterization of the Escherichia coli uracil-DNA glycosylase inhibitor protein complex. J Biol Chem, 267, 22512-22521.
2. Chu, B.C. and Orgel, L.E. (1992) The stability of different forms of double-stranded decoy DNA in serum and nuclear extracts. Nucleic Acids Res, 20, 5857-5858.
3. Tomita, T., Takeuchi, E., Tomita, N., Morishita, R., Kaneko, M., Yamamoto, K., Nakase, T., Seki, H., Kato, K., Kaneda, Y. \textit{et al.} (1999) Suppressed severity of collagen-induced arthritis by in vivo transfection of nuclear factor kappaB decoy oligodeoxynucleotides as a gene therapy. Arthritis Rheum, 42, 2532-2542.
4. Weiser, B.P., Stivers, J.T. and Cole, P.A. (2017) Protein Semi-Synthesis to Characterize Phospho-Regulation of Human UNG2. Biophys J, 112, in press.
5. Cravens, S.L. and Stivers, J.T. (2016) Comparative Effects of Ions, Molecular Crowding, and Bulk DNA on the Damage Search Mechanisms of hOGG1 and hUNG. Biochemistry, 55, 5230-5242.
6. Rodriguez, G., Esadze, A., Weiser, B.P., Schonhoff, J.D., Cole, P.A. and Stivers, J.T. (2017) Disordered N-Terminal Domain of Human Uracil DNA Glycosylase (hUNG2) Enhances DNA Translocation. Acs Chemical Biology (In Press).
7. Cravens, S.L., Schonhoft, J.D., Rowland, M.M., Rodriguez, A.A., Anderson, B.G. and Stivers, J.T. (2015) Molecular crowding enhances facilitated diffusion of two human DNA glycosylases. *Nucleic Acids Res*, **43**, 4087-4097.

8. Schonhoft, J.D. and Stivers, J.T. (2012) Timing facilitated site transfer of an enzyme on DNA. *Nature Chemical Biology*, **8**, 205-210.