Published in final edited form as:

DNA Repair (Amst). 2018 November ; 71: 101–107. doi:10.1016/j.dnarep.2018.08.013.

DNA scanning by base excision repair enzymes and implications for pathway coordination

Michael J. Howard and Samuel H. Wilson

Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC

Abstract

Site-specific DNA binding proteins must search the genome to locate their target sites, and many DNA modifying enzymes have the ability to scan along DNA in search of their substrates. This process is termed processive searching, and it serves to decrease the search time by effectively increasing the DNA binding footprint of a protein. The repertoire of proteins capable of processive searching is expanding, highlighting the need to understand the governing principles behind this fundamental process. Many of the enzymes in the base excision DNA repair pathway are capable of processive searching. Here, we briefly summarize methodology for determining if a protein can scan DNA and highlight the discovery that the base excision repair DNA polymerase β performs a processive search. Elucidation of physical models for DNA searching has also provided a plausible mechanism for pathway coordination during repair. The ability of BER enzymes to transiently sample adjacent DNA sites while bound to their product confers accessibility to downstream enzymes and does not require proteinprotein interactions for coordination.

Keywords

Facilitated diffusion; Processive searching; DNA scanning; DNA Polymerase; Base excision repair; Coordination

1. The search problem

For a biochemical reaction to occur, two or more molecules must first interact. When a substrate for an enzyme is a small molecule, both molecules eventually collide through random three-dimensional (3D) diffusion. However, many substrates are embedded in large polymers. This alters the diffusive properties of the substrate, and the enzyme's target site will have some similarity to non-target sites within the polymer which leads to unproductive binding events. This is the case for DNA binding proteins that interact with specific sites in

Corresponding Author: Samuel H. Wilson wilson5@niehs.nih.gov.

Conflict of interest

The authors declare no conflict of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

the genome. The process that these proteins must undergo to find their target sites embedded in a polymer of nontarget DNA is referred to as searching (1,2).

2. Processive searching: A general biochemical view

Searching can be described as the binding of a protein to DNA repeatedly, until the target site for the protein is eventually located. Generally, once a target site is located the protein will undergo a conformational change that enhances the binding affinity for the site (3,4). If the encountered site does not contain the target, the protein will dissociate. It is in dissociation where the protein can proceed through two different trajectories. The first, and most simplistically, the protein completely diffuses away from the DNA in an event that can be described as macroscopic dissociation. In the second dissociation model, the protein undergoes a microscopic dissociation, where it remains within the domain of the DNA and can rebind the DNA non-specifically near the first encountered site (Fig. 1A, 2). Cycling of these microscopic dissociation/reassociation events gives rise to facilitated diffusion (termed processive searching or DNA scanning) via a hopping mechanism (Fig. 2, Table 1) (1,5). Alternatively, the protein may translocate to adjacent DNA sites using a sliding mechanism (1). In a sliding mechanism, the protein makes continual contact with a single DNA strand's backbone, tracking along the helical backbone of DNA (Fig. 2) (6). In contrast, a protein using a hopping mechanism can interrogate both DNA strands (1,5,7). Both hopping and sliding are stochastic and are driven by thermal motion (1).

The DNA binding footprint can be defined as the number of base pairs that a protein encompasses when statically bound to DNA. In contrast, the DNA search footprint is the number of base pairs a protein encompasses due to the dynamic nature of protein-DNA interactions. Processive searching increases a protein's search footprint beyond the static DNA binding footprint thereby increasing the probability of locating a specific site (Fig. 1B).

Another strategy to decrease search time that does not rely on processive searching involves increasing the concentration of a searching protein. This increases the probability of the protein finding its target site by simply leading to more DNA binding events. There is likely a delicate balance that is struck between these two strategies, processive searching and increasing the concentration of a searching protein, as the latter can have deleterious consequences for the cell (8) and increasing a search footprint to longer lengths can make a search too redundant significantly hindering the process (7).

3. Methodology

The two most widely used methods to characterize the ability of a protein to processively scan DNA are singlemolecule fluorescence and ensemble biochemical processivity assays. In this review, we will highlight processivity-based assays and refer readers to excellent reviews on single-molecule fluorescence imaging techniques (9–11). However, it is worth discussing the advantages and limitations of both single-molecule and bulk biochemical assays and how the two techniques can complement each other.

3.1 Single-molecule fluorescence imaging

Searching proteins can be directly observed scanning DNA molecules by using single-molecule fluorescence microscopy setups. Typically, proteins are visualized by attachment of quantum dots or small molecule fluorophores, through a variety of strategies (11–15). The DNA is visualized by an intercalating fluorescent molecule called YOYO-1 or in some cases not at all. The advantages of such a technique include observing single molecules and transient intermediates, measuring trajectories and kinetics, and calculation of observed diffusion coefficients and dwell times on DNA.

There are some disadvantages with single-molecule fluorescent studies as compared to bulk biochemical processivity assays. The disadvantages include temporal and spatial resolution (~100–200 bp) limits and the challenge of assigning mechanistic understanding to observed diffusive states (10,15). Single-molecule studies performed with DNA binding proteins often observe longer searching lengths (kilobases) compared to the lengths measured by processivity biochemical assays (6,12,16–18). This difference may arise because processivity assays consider every active molecule of the enzyme in solution. In contrast, single-molecule experiments are more likely to observe proteins that have long binding-lifetimes on DNA, due to limitations of current setups (10). Therefore, it is possible that long scanning lengths observed in single-molecule experiments represent the behavior of only a small fraction of the total protein in solution and may be biased against the shorter binding encounters.

The parameters and observations obtained using single-molecule techniques also can be lacking in mechanistic understanding. The observed diffusion coefficients of scanning proteins do not report on a single-state of diffusion but rather at least two states: bound and unbound (10); the spatial resolution of current setups cannot detect a microscopically dissociated protein. Furthermore, a protein may display a broad range of diffusive behavior that can be classified into different diffusive modes. However, determining the underlying processes that give rise to these different states can be difficult, given the ambiguity of the observed diffusion coefficients, not knowing when an enzyme is bound to its substrate, and not knowing the active fraction of fluorescently labeled enzyme.

More obvious limitations include the possibility of introducing artifacts from labeling strategies. For instance, the stretching and intercalation of YOYO-1 dye into DNA may change how the protein will interact with the DNA. Tethering a quantum dot may additionally change the fraction of active enzyme, and how the protein interacts with DNA and protein binding partners.

3.2 Biochemical processivity assays

Ensemble biochemical processivity assays can be used to determine if an enzyme can perform a processive search for DNA sites. In these assays, a DNA molecule is constructed to contain two substrate sites, and upon a single DNA binding event the enzyme can catalyze product formation at one or both sites (Fig. 3). Assays must be designed such that each site gives a unique product DNA length and the concentration of substrate must be significantly higher than the enzyme concentration to obtain single-binding conditions (1:200 – 1:1000,

enzyme: substrate (17–21)). Because energy-independent diffusion is a stochastic process, processive searching is not an all or nothing event. Therefore, to quantify processivity the field uses the term fraction processive (F_p or P_{trans}) (17,19–21). The F_p value is the probability of an enzyme to perform a productive processive search (i.e., the probability of reaching and catalyzing product formation at the second site after catalyzing product formation at the first encountered site). The fraction processive is calculated by dividing the initial velocity for processive product formation by all initial velocity events (Fig. 3). A F_p value of unity means the enzyme will always locate and catalyze product formation at both sites within a single DNA binding event, whereas a F_p value of zero means the enzyme will only catalyze product formation at one site and then macroscopically dissociates from the DNA.

The fraction processive is the product of two probabilities: the probability of an enzyme to scan to the second site and the probability of catalyzing product formation once the second site is encountered (17). The latter probability is called the efficiency (E), and it is the probability of how committed toward catalysis an enzyme is once it has bound to its substrate (17). This efficiency (E) is different than catalytic efficiency (k_{cat}/K_M), which is a second order rate constant that measures productive associations between enzyme and substrate. The efficiency of an enzyme can be determined by pulse-chase or partition assays (17,18,22,23). In these assays, a competent enzyme complex with substrate(s) bound is challenged with competitor DNA (trap). Typically, the trap is an unlabeled DNA substrate. Once the trap is added, any enzyme that dissociates from the DNA substrate will bind the excess competitor DNA. The trapping of the dissociated enzyme results in less product formation and thus the amount of product formed is used to calculate the efficiency. In general, processive enzymes tend to have high efficiency values (>0.7), and this is what one would expect for a searching enzyme.

There are at least three methods used to determine if a protein uses a sliding or a hopping mechanism. The first method measures the ionic strength dependence of diffusion and is typically used in single-molecule fluorescence experiments. Diffusion coefficients that are ionic strength dependent are associated with a hopping mechanism. This is because the electrostatic interactions made between DNA and a protein can be disrupted by counter ions during the protein's micro-dissociated state. This leads to a reduction in the fraction of time a protein is in contact with DNA, presumably allowing for more time in the more diffusive microdissociated state, resulting in a faster observed diffusion coefficient. In contrast, the diffusion coefficient for sliding is ionic strength independent (1). However, this approach assumes that counterion rearrangements with the DNA backbone are faster than the sliding rate constant, an assumption that may not be true at high ionic strengths (1). Another method, that is used in processivity assays, involves placing two substrates on opposite DNA strands within one DNA molecule (17,18,24). If sliding is operational, the searching enzyme will dissociate off the end of the DNA molecule before reaching the second site. In another similar experiment, a roadblock (large molecule such as fluorescein or a protein) is placed between two sites on a single DNA backbone (16,18,24). Decreases in F_p values with these substrates as compared to a DNA molecule that has both substrates on the same DNA strand or lacking a roadblock would be consistent with a sliding mechanism. Finally, measuring the length dependence of F_p can provide insight into a sliding vs hopping mechanism. The F_p

value in a sliding mechanism has a stronger dependence on site spacing than a hopping mechanism (17,19).

Processive assays with the restriction enzyme BbvCI and uracil DNA glycosylase (UNG) have revealed that sliding likely occurs only over short distances and at low ionic strengths (17,25). For BbvCI, as the ionic strength or site spacing is increased the mechanism of translocation switches from sliding to hopping (25). *A priori* it would seem like sliding alone would be an inefficient strategy: for example, consider a situation where a protein is only angstroms away from its target site, but alas the target site is on the opposite DNA strand to which the protein resides.

Varying the site spacing between the two substrates can also provide information about the distance a protein can scan DNA. A typical question asked is how far can a protein scan DNA before it falls off. Because DNA scanning is probabilistic, a value that describes the mean distance traveled or the distance at which the protein has a 50% chance of traveling is useful here. This has also been referred to as the targeting radius: the distance from the binding site where there are equal probabilities of exiting the target radius region or binding to the target site (7). These values may only be informative from a comparative stand point, less so from a *in vivo* perspective. This is because F_p values are highly dependent on solution conditions (ionic strength and crowding agents) and estimating the physiological concentrations of these conditions is complex. However, understanding the concentration dependent effects of these components can provide mechanistic insights into the search process and clues about what is happening *in vivo* (see *in vivo* section below).

4. Comparative searching efficiencies of X family polymerases

The base excision repair pathway maintains genome integrity by repairing damaged DNA bases, and DNA polymerase (Pol) β is a key enzyme in this pathway. We recently developed a biochemical processivity assay to determine if Pol β can perform a processive search for 1-nt gaps in DNA (18). Using this assay, we demonstrated that Pol β can perform a processive search for DNA damage and it does so using a hopping mechanism. Mutation of three lysines to alanine within the lyase domain of Pol β abolished the observed processive searching ability while having only a modest effect on the efficiency (E). This revealed a novel function of the lyase domain and lead us to propose a model where the lyase domain interacts non-specifically with DNA during the search process.

Site-specific DNA binding proteins also bind DNA non-specifically, a consequence of overlapping interactions in specific and non-specific binding modes (10). This leads to the question of whether processive searching evolved to decrease search time or if it is just an inevitable consequence of DNA binding proteins. The prevailing view is that many, if not all, DNA repair enzymes use processive searching for substrate localization. Indeed, examples in the literature of DNA repair enzymes not performing processive searching are rare (*E. coli* proteins RNAP (26) and FtsK (27); however, these studies could not rule out short scanning events (<115 bp). To provide insight into this question, we measured the processive searching abilities of the DNA polymerase X family members that are capable of catalyzing 1-nt gap filling DNA synthesis (Pols β , μ , and λ) (23).

Pol X family members (Pol β , μ , and λ) exhibit a range of processive searching abilities for DNA gaps. Pol β demonstrates the highest searching ability with F_p values ~2-fold higher and at least 4-fold higher than for Pol λ and Pol μ , respectively. The 8 kDa domain of Pol β is instrumental for searching, and we showed that processive searching correlates with the net positive charge of the 8 kDa domains among the X family polymerases, consistent with our model of the 8 kDa domain making non-specific electrostatic interactions with DNA during the search process (18). Pols β , μ , and λ can all perform DNA gap filling, but have greatly variable processive searching abilities, possibly indicating that Pol β has specifically evolved an efficient searching mechanism to find DNA gaps.

5. Processive searching *in vivo*

A major question that remains is, to what extent does processive searching occur *in vivo*? To begin to answer this question we need first to understand the environment (solution conditions) and organization (DNA compaction) of the nucleus. The nuclear environment will include metabolites, inorganic ions, an increased viscosity and molecular crowding as compared to *in vitro* conditions. However, estimating the *in vivo* ionic strength is likely a complex endeavor given that macromolecules such as RNA, DNA, and proteins can alter the effective concentration of ions. However, recent studies using *in vivo* sensors have estimated ionic strength in mammalian cells to be ~ 140 mM (28). Nonetheless, knowing the exact concentration of a component may be challenging or can even vary between cell types, and therefore it can be more informative to understand the concentration dependence of the components on the ability of an enzyme to perform a processive search. Since the inception of facilitated diffusion research, increasing the ionic strength was shown to decrease processive searching (1). This concept is intuitive given that almost all DNA binding proteins make electrostatic interactions with DNA. However, the effects of other solution conditions have not gained as much attention until relatively recently. Stivers and coworkers have demonstrated that increasing molecular crowding, using agents such as PEG8000 (PEG 8K), increases processive searching of two DNA glycosylases by decreasing their dissociation rate constant (k_{off}) from DNA (29). In the presence of 20% (w/v) PEG 8K, human UNG2 and OGG1 fraction processive values increase from 0.16 to 0.53 and from zero to 0.26 at 150 mM potassium for hUNG2 and OGG1 at site spacing of 20 bp, respectively (4,30). In conclusion, although increasing the ionic strength decreases processive searching, the crowded cellular environment promotes it.

Most of the mammalian genome is wrapped around histone proteins forming nucleosomes. Depending on the location of the substrate in the nucleosome, DNA glycosylases and APE1 have comparable activity on nucleosome substrates as with DNA, whereas some sites are completely inhibitory (31–34). The gap-filling activity of Pol β with nucleosomal substrates is inhibitory (18,31,35–38). A question remains if BER enzymes can scan nucleosomal DNA to locate substrates and/or scan past them. To provide insight into this question, we created a processive nucleosomal substrate by positioning two 1-nt gaps into the 601 Widom strong positioning DNA sequence (18,39). The gaps were positioned 16 bp away from each other near the entry/exit site of the DNA. We observed no processive searching with this nucleosome substrate. Examination of the gap-filling kinetics revealed that DNA synthesis catalyzed by Pol β was significantly inhibited (~500-fold under singleturnover conditions as

compared to free DNA). Since the F_p value depends on searching and efficiency, and the efficiency value is near zero at this site (observed $k_{pol} \sim 0.006 \text{ s}^{-1}$), we could not determine if Pol β is able to scan nucleosomal DNA. However, Pol β DNA scanning within a nucleosomal substrate would seem futile because once a substrate is encountered it is more likely to dissociate than catalyze product formation.

In addition to histones acting as a potential barrier for DNA scanning, other DNA binding proteins could also act as roadblocks. However, proteins that diffuse along DNA using a hopping mechanism can bypass such roadblocks. The DNA glycosylase AAG can bypass a bound EcoRI dimer using a hopping mechanism (24), and covalent attachment of fluorescein in between two substrate sites on the same DNA molecule has no effect on processive searching for OGG1 and DNA Pol β (16,18).

DNA is flexible and elastic in regions of the genome that are histone depleted. The persistence length is a mechanical property that influences the stiffness of a polymer and quantifies the length at which the DNA polymer begins to fold back on itself. The persistence length of DNA is approximately 150 bps, and recent single molecule studies suggest it could be even shorter (40). A facilitated diffusion mechanism that is likely to be operational with long and high concentrations of DNA, such as that found in the nucleus (accessible DNA $\sim 0.5 \text{ mM}$ (41,42)), is intersegmental transfer. Intersegmental transfer is the ability of a protein to directly transfer from one DNA chain to another via a bridging intermediate or from a microscopically dissociated state (Fig. 2, Table 1) (43,44). During a microscopic dissociation event, the protein can either reassociate with the same strand (a hop (intramolecularly)) or transfer to a neighboring DNA (a jump (intermolecularly)). Early conceptions of intersegmental transfer postulated that only proteins with at least two DNA-binding domains could utilize intersegmental transfer because the mechanism of bridging was envisioned (1). However, recent studies support a mechanism of intersegmental transfer that involves a microscopically dissociated protein jumping to a neighboring DNA segment (44). Using a substrate that was previously utilized to characterize the intersegmental transfer ability of the DNA glycosylase AAG, we recently demonstrated that Pol β can perform intersegmental transfer (23,44). Considering the high concentration of DNA in the nucleus, intersegmental transfer may be a highly operational mode of facilitated diffusion inside the cell. This would allow a protein to escape a redundant search on a single DNA segment, while optimizing its search time on DNA.

Evidence for *in vivo* processive searching comes from studies performed with the Lac repressor (LacI), EcoRV, and some DNA repair enzymes (22,45–51). Measurements of the association kinetics of single-molecules of yellow fluorescent protein-labeled LacI in live *E. coli* cells have suggested that the LacI repressor uses DNA scanning *in vivo* to locate target sites (45,46). A scanning length of 45 bp (presumably reported as the mean search length) was calculated for LacI *in vivo* (46). Mutants of T4 endonuclease, EcoRV, MutY, and AAG that reduce processive searching *in vitro* show biological consequences when expressed in cells (22,49,50). Human uracil DNA glycosylase 2 (UNG2) can excise two closely spaced uracil lesions (10–80 bp apart) on DNA molecules transiently transfected into live cells in a single-binding encounter (51). These data provide evidence that the search process is most

likely the rate limiting step *in vivo*, and *in vitro* studies are useful for understanding DNA scanning processes *in vivo*.

6. Processive searching: a common theme in the BER pathway

DNA repair pathways require the use of multiple enzymes that act in sequential order. The product of some of these enzymes can be reactive or deleterious DNA repair intermediates that, if not repaired by the downstream enzymes in a timely manner, can lead to mutations or cell death. In these situations, it has been proposed that pathway enzymes act in a coordinated fashion—the product of the first reaction is bound by the second enzyme within the pathway before the first enzyme fully dissociates (52). This has been proposed for the base excision repair pathway that requires at least four enzymes: a DNA glycosylase, APE1, DNA polymerase β , and a DNA ligase. The available data thus far have revealed that DNA scanning is a conserved activity of at least 3 of the 4 required enzyme activities (the DNA scanning ability of DNA ligase has not yet been examined) (Fig. 4) (4,23,24,30,53).

7. Processive searching provides a mechanistic explanation for BER pathway coordination.

When an enzyme catalyzes product formation it may remain bound to the product, and for BER enzymes this is typically the longest step of the reaction cycle *in vitro* (i.e., rate limited by product release). In a coordination model, the downstream enzyme must bind the product from the first reaction before the first enzyme completely dissociates, yet the first enzyme is bound to the product. So, how can the downstream enzyme ever “see” its substrate, or how can two proteins bind to the same site at the same time? Mechanistic considerations of processive searching provide the answer. When a processive enzyme is bound to its product, it is not necessarily bound to the catalytic product site. Instead, the enzyme spends some amount of time sampling adjacent DNA sites. It is during these microscopic excursions when the substrate becomes available to the downstream enzyme. This scenario has been termed the dynamic exchange model (54) or facilitated dissociation (55–57) (Fig. 5). This model also explains the observed increase in the steady-state turnover number k_{cat} of many BER enzymes in reactions that contain a downstream enzyme (54,58), which is often interpreted as a readout for coordination. As stated earlier, the k_{cat} value for many DNA repair enzymes often measures k_{off} (product release): complete dissociation of the enzyme from the DNA product. How can binding of the downstream enzyme to the product site change the kinetics of the first reaction? Again, this can be reconciled by considering that there are at least two forms of dissociation: macro and micro. Addition of the downstream enzyme leads to the occupation of the product site when the first enzyme is in a microdissociated state, thus leading to the dissociation of the first enzyme from undamaged DNA. Dissociation from undamaged DNA occurs faster than from the product leading to an increase in k_{cat} . The ability of BER enzymes to transiently sample adjacent DNA sites while bound to their product, allows accessibility to downstream enzymes and does not require protein-protein interactions for coordination. Furthermore, in the absence of a downstream enzyme, the product-bound enzyme will most likely dissociate into bulk solution by first

binding nonspecifically to DNA that is adjacent to the product rather than the direct dissociation from the product.

8. Conclusions and perspectives

DNA binding proteins that bind specific sites in the genome undergo a search process that involves repeated DNA binding events until the target site is encountered. To increase the probability of finding a target site, many DNA binding proteins scan DNA nonspecifically, in a processive search. The discovery of processive searching has provided a possible explanation to the search problem, provides insights into the mechanisms of substrate recognition, and reveals the dynamic nature of enzyme-product complexes. The available data on processive searching suggest: a hopping mechanism appears to be the most likely mechanism by which a protein translocates appreciable distances at physiological ionic strengths, whereas sliding likely only occurs over short distances (at least for proteins that do not encircle DNA); molecular crowding increases processive searching; whereas, increasing the ionic strength decreases it; scanning distances near physiological ionic strength are not kilobase long (consistent with theoretical calculations of 100 bp being the optimal scan distance (7)); and processive searching may be operational *in vivo*.

Understanding how BER enzymes can accomplish genome-wide repair and how these enzymes coordinate in the BER pathway are major questions in the field. Processive searching provides a solution to both these questions by revealing the dynamic nature of protein-DNA complexes. In particular, the model of a DNA binding protein undergoing microscopic dissociation/reassociation with DNA is possibly the most important concept that has emerged from these studies. Future studies should try to address the properties of a microdissociated protein by measuring rate constants that describe these events and by determining the distance (targeting radius) over which a microdissociated protein can travel from DNA before irreversibly diffusing into bulk solution. Furthermore, studies that address the subtle molecular interactions that are required for processive searching are needed. Observationally, there is a zone around DNA where microscopically dissociated enzyme has a much higher probability of rebinding along the target DNA than binding to a trap, even though it is in vast excess. The explanation for this difference in association energies is currently lacking, but is of fundamental importance in understanding the behavior of DNA binding proteins and enzymes.

Acknowledgements

This research was supported by grants Z01-ES050158 and Z01-ES050159 from the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences. We thank members of the Wilson lab, Mark Hedglin, and Patrick O'Brien for helpful comments.

Abbreviations:

BER	base excision repair
Pol B	DNA polymerase beta
APE1	apurinic/apyrimidinic endonuclease

UNG	Uracil DNA glycosylase
AAG	Alkyladenine DNA glycosylase

References

1. Berg OG, Winter RB, and von Hippel PH (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* 20, 6929–6948 [PubMed: 7317363]
2. Halford SE (2009) An end to 40 years of mistakes in DNA-protein association kinetics? *Biochem Soc Trans* 37, 343–348 [PubMed: 19290859]
3. Cravens SL, Hobson M, and Stivers JT (2014) Electrostatic properties of complexes along a DNA glycosylase damage search pathway. *Biochemistry* 53, 7680–7692 [PubMed: 25408964]
4. Cravens SL, and Stivers JT (2016) Comparative Effects of Ions, Molecular Crowding, and Bulk DNA on the Damage Search Mechanisms of hOGG1 and hUNG. *Biochemistry* 55, 5230–5242 [PubMed: 27571472]
5. Berg OG, and von Hippel PH (1985) Diffusion-controlled macromolecular interactions. *Annu Rev Biophys Biophys Chem* 14, 131–160 [PubMed: 3890878]
6. Blainey PC, van Oijen AM, Banerjee A, Verdine GL, and Xie XS (2006) A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc Natl Acad Sci U S A* 103, 5752–5757 [PubMed: 16585517]
7. Halford SE, and Marko JF (2004) How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res* 32, 3040–3052 [PubMed: 15178741]
8. Broustas CG, and Lieberman HB (2014) DNA damage response genes and the development of cancer metastasis. *Radiat Res* 181, 111–130 [PubMed: 24397478]
9. Silverstein TD, Gibb B, and Greene EC (2014) Visualizing protein movement on DNA at the single-molecule level using DNA curtains. *DNA Repair (Amst)* 20, 94–109 [PubMed: 24598576]
10. Redding S, and Greene EC (2013) How do proteins locate specific targets in DNA? *Chem Phys Lett* 570
11. Hughes CD, Simons M, Mackenzie CE, Van Houten B, and Kad NM (2014) Single molecule techniques in DNA repair: a primer. *DNA Repair (Amst)* 20, 2–13 [PubMed: 24819596]
12. Nelson SR, Dunn AR, Kathe SD, Warshaw DM, and Wallace SS (2014) Two glycosylase families diffusively scan DNA using a wedge residue to probe for and identify oxidatively damaged bases. *Proc Natl Acad Sci U S A* 111, E2091–2099 [PubMed: 24799677]
13. Liu L, Kong M, Gassman NR, Freudenthal BD, Prasad R, Zhen S, Watkins SC, Wilson SH, and Van Houten B (2017) PARP1 changes from three-dimensional DNA damage searching to one-dimensional diffusion after auto-PARYlation or in the presence of APE1. *Nucleic Acids Res* 45, 12834–12847 [PubMed: 29121337]
14. Gorman J, Plys AJ, Visnapuu ML, Alani E, and Greene EC (2010) Visualizing one-dimensional diffusion of eukaryotic DNA repair factors along a chromatin lattice. *Nat Struct Mol Biol* 17, 932–938 [PubMed: 20657586]
15. Myler LR, Gallardo IF, Soniat MM, Deshpande RA, Gonzalez XB, Kim Y, Paull TT, and Finkelstein IJ (2017) Single-Molecule Imaging Reveals How Mre11-Rad50-Nbs1 Initiates DNA Break Repair. *Mol Cell* 67, 891898 e894
16. Rowland MM, Schonhoft JD, McKibbin PL, David SS, and Stivers JT (2014) Microscopic mechanism of DNA damage searching by hOGG1. *Nucleic Acids Res* 42, 9295–9303 [PubMed: 25016526]
17. Porecha RH, and Stivers JT (2008) Uracil DNA glycosylase uses DNA hopping and short-range sliding to trap extrahelical uracils. *Proc Natl Acad Sci U S A* 105, 10791–10796 [PubMed: 18669665]
18. Howard MJ, Rodriguez Y, and Wilson SH (2017) DNA polymerase beta uses its lyase domain in a processive search for DNA damage. *Nucleic Acids Res* 45, 3822–3832 [PubMed: 28119421]

19. Stanford NP, Szczelkun MD, Marko JF, and Halford SE (2000) One- and three-dimensional pathways for proteins to reach specific DNA sites. *EMBO J* 19, 6546–6557 [PubMed: 11101527]
20. Terry BJ, Jack WE, and Modrich P (1985) Facilitated diffusion during catalysis by EcoRI endonuclease. Nonspecific interactions in EcoRI catalysis. *J Biol Chem* 260, 13130–13137 [PubMed: 2997157]
21. Hedglin M, and O'Brien PJ (2008) Human alkyladenine DNA glycosylase employs a processive search for DNA damage. *Biochemistry* 47, 11434–11445 [PubMed: 18839966]
22. Zhang Y, and O'Brien PJ (2015) Repair of Alkylation Damage in Eukaryotic Chromatin Depends on Searching Ability of Alkyladenine DNA Glycosylase. *ACS Chem Biol* 10, 2606–2615 [PubMed: 26317160]
23. Howard MJ, and Wilson SH (2017) Processive searching ability varies among members of the gap-filling DNA polymerase X family. *J Biol Chem*
24. Hedglin M, and O'Brien PJ (2010) Hopping enables a DNA repair glycosylase to search both strands and bypass a bound protein. *ACS Chem Biol* 5, 427–436 [PubMed: 20201599]
25. Gowers DM, Wilson GG, and Halford SE (2005) Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. *Proc Natl Acad Sci U S A* 102, 15883–15888 [PubMed: 16243975]
26. Wang F, Redding S, Finkelstein IJ, Gorman J, Reichman DR, and Greene EC (2013) The promoter-search mechanism of Escherichia coli RNA polymerase is dominated by three-dimensional diffusion. *Nat Struct Mol Biol* 20, 174–181 [PubMed: 23262491]
27. Lee JY, Finkelstein IJ, Crozat E, Sherratt DJ, and Greene EC (2012) Single-molecule imaging of DNA curtains reveals mechanisms of KOPS sequence targeting by the DNA translocase FtsK. *Proc Natl Acad Sci U S A* 109, 6531–6536 [PubMed: 22493241]
28. Liu B, Poolman B, and Boersma AJ (2017) Ionic Strength Sensing in Living Cells. *ACS Chem Biol* 12, 2510–2514 [PubMed: 28853549]
29. Cravens SL, Schonhoft JD, Rowland MM, Rodriguez AA, Anderson BG, and Stivers JT (2015) Molecular crowding enhances facilitated diffusion of two human DNA glycosylases. *Nucleic Acids Res* 43, 4087–4097 [PubMed: 25845592]
30. Rodriguez G, Esadze A, Weiser BP, Schonhoft JD, Cole PA, and Stivers JT (2017) Disordered N-Terminal Domain of Human Uracil DNA Glycosylase (hUNG2) Enhances DNA Translocation. *ACS Chem Biol* 12, 2260–2263 [PubMed: 28787572]
31. Beard BC, Wilson SH, and Smerdon MJ (2003) Suppressed catalytic activity of base excision repair enzymes on rotationally positioned uracil in nucleosomes. *Proc Natl Acad Sci U S A* 100, 7465–7470 [PubMed: 12799467]
32. Ye Y, Stahley MR, Xu J, Friedman JI, Sun Y, McKnight JN, Gray JJ, Bowman GD, and Stivers JT (2012) Enzymatic excision of uracil residues in nucleosomes depends on the local DNA structure and dynamics. *Biochemistry* 51, 6028–6038 [PubMed: 22784353]
33. Bilotti K, Kennedy EE, Li C, and Delaney S (2017) Human OGG1 activity in nucleosomes is facilitated by transient unwrapping of DNA and is influenced by the local histone environment. *DNA Repair (Amst)* 59, 1–8 [PubMed: 28892740]
34. Olmon ED, and Delaney S (2017) Differential Ability of Five DNA Glycosylases to Recognize and Repair Damage on Nucleosomal DNA. *ACS Chem Biol* 12, 692–701 [PubMed: 28085251]
35. Rodriguez Y, and Smerdon MJ (2013) The structural location of DNA lesions in nucleosome core particles determines accessibility by base excision repair enzymes. *J Biol Chem* 288, 13863–13875 [PubMed: 23543741]
36. Menoni H, Gasparutto D, Hamiche A, Cadet J, Dimitrov S, Bouvet P, and Angelov D (2007) ATP-dependent chromatin remodeling is required for base excision repair in conventional but not in variant H2A.Bbd nucleosomes. *Mol Cell Biol* 27, 5949–5956 [PubMed: 17591702]
37. Odell ID, Barbour JE, Murphy DL, Della-Maria JA, Sweasy JB, Tomkinson AE, Wallace SS, and Pederson DS (2011) Nucleosome disruption by DNA ligase III-XRCC1 promotes efficient base excision repair. *Mol Cell Biol* 31, 4623–4632 [PubMed: 21930793]
38. Rodriguez Y, Howard MJ, Cuneo MJ, Prasad R, and Wilson SH (2017) Unencumbered Pol beta lyase activity in nucleosome core particles. *Nucleic Acids Res* 45, 8901–8915 [PubMed: 28911106]

39. Segal E, Fondufe-Mittendorf Y, Chen L, Thastrom A, Field Y, Moore IK, Wang JP, and Widom J (2006) A genomic code for nucleosome positioning. *Nature* 442, 772–778 [PubMed: 16862119]
40. Vafabakhsh R, and Ha T (2012) Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization. *Science* 337, 1097–1101 [PubMed: 22936778]
41. Valouev A, Johnson SM, Boyd SD, Smith CL, Fire AZ, and Sidow A (2011) Determinants of nucleosome organization in primary human cells. *Nature* 474, 516–520 [PubMed: 21602827]
42. Esadze A, Kemme CA, Kolomeisky AB, and Iwahara J (2014) Positive and negative impacts of nonspecific sites during target location by a sequence-specific DNA-binding protein: origin of the optimal search at physiological ionic strength. *Nucleic Acids Res* 42, 7039–7046 [PubMed: 24838572]
43. Iwahara J, and Clore GM (2006) Direct observation of enhanced translocation of a homeodomain between DNA cognate sites by NMR exchange spectroscopy. *J Am Chem Soc* 128, 404–405 [PubMed: 16402815]
44. Hedglin M, Zhang Y, and O'Brien PJ (2013) Isolating contributions from intersegmental transfer to DNA searching by alkyladenine DNA glycosylase. *J Biol Chem* 288, 24550–24559 [PubMed: 23839988]
45. Elf J, Li GW, and Xie XS (2007) Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* 316, 1191–1194 [PubMed: 17525339]
46. Hammar P, Leroy P, Mahmutovic A, Marklund EG, Berg OG, and Elf J (2012) The lac repressor displays facilitated diffusion in living cells. *Science* 336, 1595–1598 [PubMed: 22723426]
47. Dowd DR, and Lloyd RS (1989) Biological consequences of a reduction in the non-target DNA scanning capacity of a DNA repair enzyme. *J Mol Biol* 208, 701–707 [PubMed: 2681789]
48. Dowd DR, and Lloyd RS (1990) Biological significance of facilitated diffusion in protein-DNA interactions. Applications to T4 endonuclease V-initiated DNA repair. *J Biol Chem* 265, 3424–3431 [PubMed: 2406255]
49. Brinkmeyer MK, Pope MA, and David SS (2012) Catalytic contributions of key residues in the adenine glycosylase MutY revealed by pH-dependent kinetics and cellular repair assays. *Chem Biol* 19, 276–286 [PubMed: 22365610]
50. Jeltsch A, Wenz C, Stahl F, and Pingoud A (1996) Linear diffusion of the restriction endonuclease EcoRV on DNA is essential for the in vivo function of the enzyme. *EMBO J* 15, 5104–5111 [PubMed: 8890184]
51. Esadze A, Rodriguez G, Weiser BP, Cole PA, and Stivers JT (2017) Measurement of nanoscale DNA translocation by uracil DNA glycosylase in human cells. *Nucleic Acids Res* 45, 12413–12424 [PubMed: 29036472]
52. Wilson SH, and Kunkel TA (2000) Passing the baton in base excision repair. *Nat Struct Biol* 7, 176–178 [PubMed: 10700268]
53. Carey DC, and Strauss PR (1999) Human apurinic/aprimidinic endonuclease is processive. *Biochemistry* 38, 16553–16560 [PubMed: 10600117]
54. Baldwin MR, and O'Brien PJ (2012) Defining the functional footprint for recognition and repair of deaminated DNA. *Nucleic Acids Res* 40, 11638–11647 [PubMed: 23074184]
55. Graham JS, Johnson RC, and Marko JF (2011) Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded DNA. *Nucleic Acids Res* 39, 2249–2259 [PubMed: 21097894]
56. Giuntoli RD, Linzer NB, Banigan EJ, Sing CE, de la Cruz MO, Graham JS, Johnson RC, and Marko JF (2015) DNA-Segment-Facilitated Dissociation of Fis and NHP6A from DNA Detected via Single-Molecule Mechanical Response. *J Mol Biol* 427, 3123–3136 [PubMed: 26220077]
57. Kamar RI, Banigan EJ, Erbas A, Giuntoli RD, Olvera de la Cruz M, Johnson RC, and Marko JF (2017) Facilitated dissociation of transcription factors from single DNA binding sites. *Proc Natl Acad Sci U S A* 114, E3251–E3257 [PubMed: 28364020]
58. Esadze A, Rodriguez G, Cravens SL, and Stivers JT (2017) AP-Endonuclease 1 Accelerates Turnover of Human 8-Oxoguanine DNA Glycosylase by Preventing Retrograde Binding to the Abasic-Site Product. *Biochemistry* 56, 1974–1986 [PubMed: 28345889]

Figure 1. Processive searching (facilitated diffusion) increases a protein's DNA binding footprint. A. Comparison of random 3D diffusion and facilitated diffusion (FD). In the 3D diffusion binding model, the protein binds nonspecifically to DNA and then completely dissociates. In facilitated diffusion, the protein binds nonspecifically but then through a series of microscopic dissociation and reassociation events scans the adjacent DNA sites before macro-dissociating. The light green cartoon protein represents a scanning protein. B. Gaussian distribution representing the probability of a protein scanning a given distance from the initial encounter site. The initial encounter site is represented by the area between the dotted lines and also represents the binding footprint of a protein that does not use facilitated diffusion, but just binds and dissociates.

Figure 2. Mechanisms of facilitated diffusion.
See Table 1 for definitions.

Figure 3. Biochemical processivity assay design.

A. The processive substrate contains two substrates on one DNA molecule. The correlated enzymatic activity at both sites in one DNA binding event by an enzyme will lead to the processive product (ABC). If the enzyme dissociates after catalysis at the first encounter site, it will lead to the distributive products (AB or BC). The rate of product formation for the distributive products will be the same if they are encountered equally. B. Representative time course for a processivity assay. The total product is ABC+BC+AB. Processive searching can be quantified by using the term fraction processive (F_p). This value represents the probability of an enzyme performing a productive search in a single-DNA binding event.

Figure 4. Fraction processive values for some DNA scanning enzymes.

The green bars represent F_p values measured at low ionic strengths (~50 mM) and grey bars at higher ionic strength (~150 mM). Note that AP endonuclease 1 (APE1) has a processive searching ability, but a quantitative analysis has not been reported (53). The site spacing is reported below the enzyme: 25, 20, 20, 20, 20, and 54 bp for AAG (24), UNG2 (30), OGG1 (4), Pol β (23), Pol μ (23), and EcoRI (25) respectively. Two values have not been measured/reported (n.m.). It should be noted that addition of the molecular crowding agent PEG8K increased the fraction processive values of UNG2 and OGG1 at 150 mM ionic strength from 0.16 and 0.0 to 0.53 and 0.26, respectively. Pol μ is involved in the non-homologous end joining pathway and is shown here along with EcoRI for comparison purposes.

Figure 5. Processive searching provides a mechanistic explanation for BER pathway coordination: The dynamic exchange model.

Biochemical processivity assays have revealed that the product bound state of an enzyme can be dynamic. The enzyme performs micro-dissociation/association events where it samples adjacent DNA sites. During these transient excursions, the product site becomes accessible for binding of the downstream protein in a repair pathway without the need for a protein-protein interaction.

Table 1.

Definitions.

Facilitated diffusion	The ability of a DNA binding protein to sample multiple sites on DNA in a single binding encounter and is driven by thermal fluctuations and non-specific DNA binding. Synonyms: processive searching, DNA scanning, intramolecular transfer, and translocation.
Sliding	An intramolecular mode of facilitated diffusion where a protein diffuses along a single-DNA strand.
Hopping	An intramolecular mode of facilitated diffusion where a protein diffuses along DNA via microscopic dissociation/reassociation events, allowing sampling of both DNA strands.
Intersegmental transfer	An intermolecular mode of facilitated diffusion where a DNA binding protein transfers from one DNA segment to another via a jumping or a bridging mechanism.
Jumping	Similar to a hopping mechanism, but the site of reassociation is much farther from the site of dissociation or on a different DNA molecule.
Bridging	An intermolecular mode of facilitated diffusion where a protein with more than one DNA binding domain (or a dimeric form) transfers from one DNA segment to another via an intermediate where both DNAs are bound simultaneously.
Processivity	The ability of a DNA modifying enzyme to catalyze more than one reaction on a DNA molecule in a single binding encounter.