HHS Public Access

Author manuscript

Chembiochem. Author manuscript; available in PMC 2024 May 16.

Published in final edited form as:

Chembiochem. 2023 May 16; 24(10): e202200765. doi:10.1002/cbic.202200765.

Replication Protein A Enhances Kinetics of Uracil DNA Glycosylase on ssDNA and Across DNA Junctions: Explored with a DNA Repair Complex Produced with SpyCatcher/SpyTag Ligation

Sharon N. Greenwood[a], **Rashmi S. Kulkarni**[a], **Michel Mikhail**[a,b], **Brian P. Weiser**[a]

[a]Department of Molecular Biology, Rowan University School of Osteopathic Medicine, Stratford, NJ 08084 (USA)

[b]Department of Internal Medicine, Newark Beth Israel Medical Center, Newark, NJ 07112 (USA)

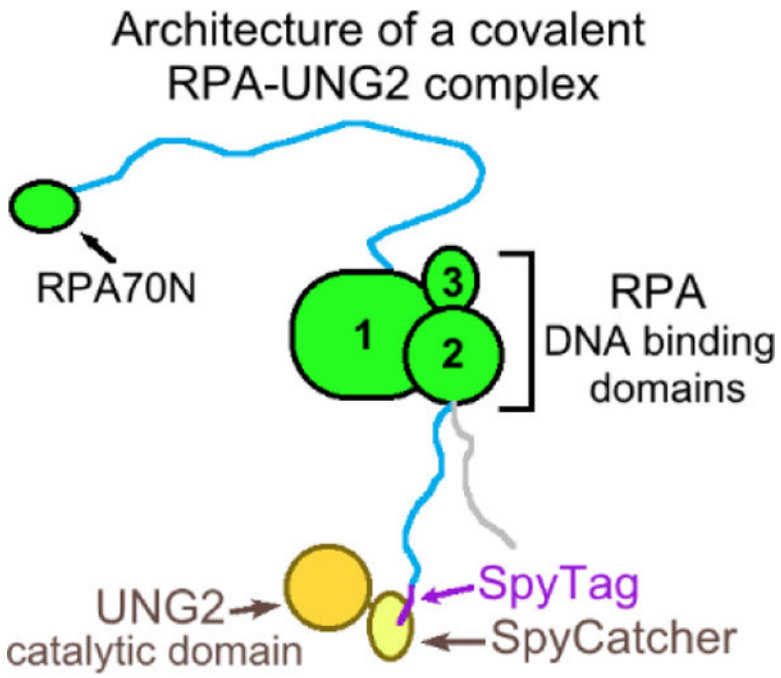
Abstract

DNA repair proteins participate in extensive protein-protein interactions that promote the formation of DNA repair complexes. To understand how complex formation affects protein function during base excision repair, we tested SpyCatcher/SpyTag ligation to produce a covalent complex between human uracil DNA glycosylase (UNG2) and replication protein A (RPA). Our covalent “RPA-Spy-UNG2” complex could identify and excise uracil bases in duplex areas next to ssDNA-dsDNA junctions slightly faster than the wild-type proteins, but this was highly dependent on DNA structure, as the turnover of the RPA-Spy-UNG2 complex slowed at DNA junctions where RPA tightly engaged long ssDNA sections. Conversely, the enzymes preferred uracil sites in ssDNA where RPA strongly enhanced uracil excision by UNG2 regardless of ssDNA length. Finally, RPA was found to promote UNG2 excision of two uracil sites positioned across a ssDNA-dsDNA junction, and dissociation of UNG2 from RPA enhanced this process. Our approach of ligating together RPA and UNG2 to reveal how complex formation affects enzyme function could be applied to examine other assemblies of DNA repair proteins.

Graphical Abstract

weiser@rowan.edu .

Supporting information for this article is given via a link at the end of the document.



SpyCatcher/SpyTag ligation was used to produce a covalent complex containing DNA repair proteins replication protein A (RPA) and uracil DNA glycosylase (UNG2). By comparing its activity to the wild-type proteins, we assigned functions specific to the complex and investigated the mechanisms used by RPA to target UNG2 towards uracil bases in ssDNA and at ssDNA-dsDNA junctions.

Keywords

enzyme; uracil; RPA; UNG2; spycatcher

Introduction

DNA repair processes are critical for cellular health and are involved in disease states such as the formation and treatment of cancers. DNA repair is also essential for programmed mutagenesis in adaptive immunity and targeted mutagenesis including CRISPR gene editing. DNA repair processes are orchestrated by hundreds of proteins with precise spatial and temporal regulation. To complete each process, unique sets of proteins are coordinated through defined pathways. Proteins in similar DNA repair pathways often interact with each other, or else interact with a common scaffold or hub protein that binds multiple other proteins. This led to the finding that proteins organize into distinct replication and repair complexes to enhance the efficiency of their processes.[1–5] One well-examined hub protein is replication protein A (RPA), which binds ssDNA with high affinity and has dozens of protein binding partners.[6,7] RPA recruits other proteins to specific nuclear sites including replication forks,[8–15] and RPA can affect the activity of enzymes and other DNA binding proteins.[16–19]

In the cell, RPA facilitates various DNA repair processes in the crowded nucleus where its many protein binding partners exist and the potential for protein complex formation is high. DNA repair proteins including RPA interact with sufficient strength to remain intact as isolated complexes under various chromatography conditions.[1,3] On the other hand, the interactions of recombinant proteins with RPA *in vitro* typically occur with K_d values in the high nM to low μM range.[17,19–22] These moderate affinities suggest that RPA-containing protein complexes may be dynamic in composition, at least in cell-free systems.[17,21,22] Cellular factors such as macromolecular crowding, post-translational modifications, and changes in protein expression can promote the formation of specific protein-protein interactions.[7,19,22,23] It is unclear how the activity of DNA repair proteins would be affected by forming long-lived protein complexes with their hub proteins such as RPA as opposed to forming transient complexes that can assemble and disband following reversible binding events.

The above characteristics of protein interactions involving RPA apply to its role in uracil base excision repair where it interacts with the enzyme uracil DNA glycosylase (UNG2).[24] UNG2 removes uracil bases from DNA to initiate base excision repair pathways that also involve the hub proteins PCNA and XRCC1, as well as AP endonuclease, DNA polymerases, DNA ligases, and other proteins.[24–27] Multi-protein complexes containing UNG2 have been purified from human cells and are capable of completing uracil base excision repair *in vitro*.[4,25] RPA is thought to bind UNG2 to regulate its localization to the replication fork during normal replication and during DNA damage responses,[2,15,28] and PTMs regulate the interaction of UNG2 with RPA and other proteins including PCNA.[19,22,28,29] Reported K_d values for the interaction between RPA and UNG2 range from 114 nM to 3 μM.[17,19,22] Finally, RPA interacts with ssDNA and is capable of enhancing UNG2's uracil excision activity by targeting the enzyme to nearby DNA regions that contain uracil bases.[16,17,19]

To learn more about the functional effects of protein complex formation in uracil base excision repair, we examined the feasibility of using the SpyCatcher/SpyTag system[30–32] to ligate UNG2 and RPA to form a small covalent protein complex that resembles the architecture of the two proteins that normally bind reversibly. This allowed us to compare the uracil excision activity of an obligate RPA-UNG2 complex to the activity of the wild-type proteins acting together on defined DNA substrates *in vitro* (our covalent RPA-UNG2 complex is referred to here as "RPA-Spy-UNG2"). Consistent with previous reports,[16,17,19] we found that RPA enhanced UNG2 activity when the individual proteins acted together to remove uracil bases from model ssDNA-dsDNA junction substrates. The activity of the covalent RPA-Spy-UNG2 complex that we designed slightly exceeded the uracil excision rate of the individual proteins, but the ability of the complex to search bulk DNA for uracil was dependent on the structure of the DNA substrate. This was because the turnover of the RPA-Spy-UNG2 complex was slowed at ssDNA-dsDNA junctions where RPA engaged long ssDNA sections and dsDNA was available for UNG2 to bind. In contrast, RPA could efficiently handoff long ssDNA substrates to UNG2 in the absence of a DNA junction whether or not the two proteins were joined as a complex. This resulted in RPA significantly enhancing uracil excision by UNG2 in ssDNA. Finally, RPA also enhanced the ability of UNG2 to remove multiple uracils located across a ssDNA-dsDNA junction, and dissociation

of RPA and UNG2 further facilitated this process. In summary, the covalent RPA-Spy-UNG2 complex exhibited substrate-dependent behavior that provided only small kinetic advantages for UNG2 in specific contexts, and the reversible binding of wild-type RPA and UNG2 improved the enzyme's performance on many DNA substrates. This knowledge of UNG2 activity is important for understanding its role during antibody diversification processes that involve genomic uracilation,[33–35] and additionally, for understanding how UNG2 affects the efficacy of thymidylate synthase inhibitors that elevate genomic uracil levels.[36–39] While we used SpyCatcher/SpyTag ligation to explore how RPA and UNG2 function in an obligate complex, we anticipate that larger assemblies of DNA repair proteins could be produced for functional studies with orthogonal genetically-encoded ligation tools.[40,41]

Results and Discussion

Design and Production of a Covalent RPA-Spy-UNG2 Protein Complex

To understand how RPA influences UNG2 activity as part of a protein complex, we aimed to produce an irreversible RPA-UNG2 complex that resembled the complex that forms when the native proteins interact. RPA is an obligate heterotrimer with a DNA binding core and two satellite domains that are connected to the core by long flexible linkers (Figure 1A). The satellite "winged-helix" domain of RPA2 (RPA2-WH) interacts with a small section of the N-terminal domain of UNG2 adjacent to the UNG2 catalytic domain (Figure 1A).[22,24,42] We initially attempted to express recombinant fusion proteins with UNG2 fused to the C-terminus of the RPA2 subunit, but this approach was unsuccessful (see Experimental Section). Instead, we designed recombinant RPA and UNG2 variants for ligation using the SpyCatcher/SpyTag system that relies on a high affinity interaction between the globular SpyCatcher domain and the 13 residue SpyTag; after binding, SpyCatcher and SpyTag rapidly form a covalent isopeptide bond.[30–32] For this approach, we genetically deleted the winged-helix domain from RPA2 on a heterotrimeric RPA expression plasmid and replaced it with the SpyTag sequence (Figure 1B). The winged-helix domain of RPA is not a DNA binding domain and does not affect RPA's ability to bind ssDNA.[17] Recombinant "RPA-SpyTag" containing RPA1, RPA2-SpyTag, and RPA3 was expressed and purified similar to wild-type RPA (Figure 1C, lane 3).[16,17] Separately, we purified a recombinant UNG2 protein called "SpyCatcher-UNG2" (SpyC-UNG2) that contained an N-terminal SpyCatcher domain followed by a small flexible linker and the UNG2 catalytic domain (residues 92–313) (Figure 1B and Figure 1C, lane 2).

To test production of the covalent RPA-Spy-UNG2 complex, we mixed different ratios of SpyC-UNG2 and RPA-SpyTag at room temperature for 4 hours. When the proteins were mixed with 1:1 stoichiometry at 15 μM each, >95% of RPA-SpyTag ligated to SpyC-UNG2 (Figure 1C, lane 6). The covalent product of the ligation reaction was visible on SDS-PAGE by the appearance of an appropriately-sized RPA2-SpyTag/SpyC-UNG2 polypeptide (Figure 1C). Analysis of the reactions with size exclusion chromatography (SEC) indicated that the multi-protein complex eluted at a volume that was appropriate compared to wild-type RPA (Figure 1D). Both proteins eluted at slightly higher molecular weights than predicted (Figure 1D), which was likely due to the extended modular architectures of RPA and the

covalent RPA-Spy-UNG2 complex that differs from the tight, globular proteins used as SEC standards. We subsequently reacted RPA-SpyTag and SpyC-UNG2 on a preparative scale and used SEC to produce the RPA-Spy-UNG2 complex with high purity (Figure 1E).

Activity of the RPA-Spy-UNG2 Complex at ssDNA-dsDNA Junctions with Long ssDNA Sections

Prior to its use in ligation reactions, the enzymatic activity of SpyC-UNG2 was measured using a ssDNA-dsDNA junction substrate that contained a 32 nt ssDNA section adjacent to a 55 bp duplex, and a single uracil base was positioned 21 bp from the junction (Figure 2A). This junction substrate contained a fluorescein end-label on the uracilated strand that allowed for in-gel quantification of DNA fragments that resulted from uracil base excision by SpyC-UNG2 followed by chemical cleavage of abasic sites.[16,17] SpyC-UNG2's kinetic parameters (K_m = 3.9 ± 0.7 μM and k_{cat} = 81 ± 8 min^{-1}) were within experimental error of previously reported values for the UNG2 catalytic domain (K_m = 3.2 ± 0.4 μM and k_{cat} = 68 ± 4 min^{-1}) under identical experimental conditions[17] (Figure 2A). Note that the disordered N-terminal domain of UNG2 (residues 1–91) was omitted during the design of SpyC-UNG2 (Figure 1B), and the absence of the N-terminal domain may affect the uracil excision kinetics of free SpyC-UNG2 compared to UNG2.[16,17,43,44] Nonetheless, its kinetic parameters demonstrated that SpyC-UNG2 was folded correctly and functional, and its inclusion in assays below served to ensure that the SpyCatcher domain had no adverse effects on the ability of the enzyme to react on different DNA substrates.

Using the same DNA junction substrate, we compared the uracil excision activity of our proteins of interest (UNG2, UNG2 in the presence of wild-type RPA, SpyC-UNG2, and the covalent RPA-Spy-UNG2 complex). As expected from previous work,[16,17] RPA binds tightly to the 32 nt ssDNA section and significantly enhanced UNG2 excision of the uracil positioned 21 bp from the junction (Figure 2B, middle lane). Although SpyC-UNG2 had ~2-fold higher activity than wild-type UNG2 in this assay, we could barely detect any uracil excision by the RPA-Spy-UNG2 complex when we used equivalent levels of enzyme (Figure 2B). The percent of substrate processed by RPA-Spy-UNG2 suggested that each molecule of covalent complex averaged one turnover or less in the three minute assay (0.5 turnovers/molecule were measured for RPA-Spy-UNG2 compared to 3.9 for UNG2) (see Experimental Section). The activity of the RPA-Spy-UNG2 complex was similarly weak when the uracil base was positioned 9 bp or 33 bp from the junction (Figure S1), and additionally, the activity remained weak when the polarity of the DNA strands on the junction substrate were changed (explored in detail below). Our previous work demonstrated that RPA binds tightly to long ssDNA sections adjacent to DNA junctions and enhances UNG2 activity in the duplex area near the junction regardless of DNA strand polarity or the placement of uracil in either strand,[16,17] and thus the weak activity we measured here was specific to the RPA-Spy-UNG2 complex.

To further examine the activity of the RPA-Spy-UNG2 complex, we prepared a similar DNA junction substrate with a 32 nt ssDNA section that terminated with a 3′ end, and the uracil base was again positioned 21 bp from the junction. In time point assays, the activity of the RPA-Spy-UNG2 complex remained weak compared to SpyC-UNG2, which was used as a

control (Figure 2C). The RPA-Spy-UNG2 complex appeared to exhibit a burst of activity at the onset of reactions followed by a much slower rate of activity that was linear throughout the measured time course, but the linear rate was 20-fold slower than the linear rate of SpyC-UNG2 at equivalent enzyme levels (Figure 2C and Figure S2). The linear fit for RPA-Spy-UNG2 activity in Figure 2C intersected the y-axis at 3.8% substrate processed, which was similar to the molar ratio of enzyme to DNA used in the assay (15 nM/300 nM or 5%). The burst of activity at the onset of reactions indicated a ~stoichiometric amount of rapid uracil excision and product formation by the covalent complex, while subsequent turnover was likely slowed because of rate-limiting post-cleavage steps including product release and product inhibition, which is commonly observed among DNA glycosylases.[45,46]

To further understand how RPA influences UNG2 activity at the onset of enzyme reactions, we used a stopped-flow device and modified our DNA junction substrate for detecting continuous uracil base excision. We prepared a junction substrate with a 32 nt ssDNA section and a 55 bp duplex with a uracil positioned 21 bp from the junction, but in this case, we incorporated a 2-aminopurine base in the duplex opposite the uracil (Figure 2D). When the uracil is removed by UNG2, the 2-aminopurine becomes unstacked from the helix and its fluorescence increases in a proportional manner.[47] Reaction of 400 nM DNA with 200 nM UNG2 in the stopped-flow assays resulted in a monophasic exponential curve that determined a single rate of uracil excision activity for UNG2 ($k = 0.15\ s^{-1}$) (Figure 2D, Table 1, and Figure S3). Analysis of SpyC-UNG2 in this assay determined a similar rate of uracil excision ($k = 0.10\ s^{-1}$) (Figure 2E and Table 1). Although the assay conditions did not satisfy steady-state or single turnover criteria (see Experimental Section), we reasoned that the relative uracil excision rates for our proteins of interest may still reveal effects of RPA on UNG2 alone and as part of the RPA-Spy-UNG2 complex.

In contrast to the monophasic activity of UNG2 and SpyC-UNG2 that we observed in stopped-flow assays, the uracil excision activity of UNG2 was biphasic in the presence of RPA. This biphasic behavior occurred under conditions where UNG2 was introduced into the stopped-flow cuvette with RPA pre-bound to the junction substrate (Figure 2F), and also under conditions where RPA was introduced to the DNA simultaneously with UNG2 (RPA/UNG2 co-injection) (Figure 2G and Table 1). The biphasic reactions had two rates of activity (k_{fast}andk_{slow}) and an additional parameter that estimated the amount of uracil excision that occurred with the faster rate (%fast). The effects of RPA on UNG2 were similar regardless of how the proteins were introduced to the DNA substrate (Figure 2F, Figure 2G, and Table 1). The stimulatory effect of RPA on UNG2 was observed in these assays, as the k_{fast} was ~15-fold higher than the rate of uracil excision by UNG2 alone (Figure 2D, Figure 2F and Figure 2G). The k_{slow} in experiments containing RPA also matched the activity of UNG2 alone, capturing the RPA-independent rate of uracil excision (Figure 2D, Figure 2F and Figure 2G). Finally, it was estimated that 22–26% of UNG2 activity in the presence of RPA occurred with the faster rate (Figure 2F, Figure 2G, and Table 1).

Next, we measured the uracil excision activity of the RPA-Spy-UNG2 complex in stopped-flow assays, which also had biphasic behavior (Figure 2H). Remarkably, the k_{fast} for the RPA-Spy-UNG2 complex was 185-fold higher than its k_{slow} (Figure 2H and Table 1). The k_{fast}

for RPA-Spy-UNG2 was 2-fold higher than the k_{fast} for wild-type RPA stimulating UNG2, but the k_{slow} for the RPA-Spy-UNG2 complex was reduced 5-fold compared to UNG2 alone. Even though its k_{fast} exceeded the rates of activity for all other protein/enzyme combinations, the RPA-Spy-UNG2 complex became less efficient than the individual proteins over time because the majority of its activity was attributed to its k_{slow} (see $t_{50\%}$ in Table 1). The fast initial activity followed by slow turnover and low overall activity was consistent with the initial endpoint assays using RPA-Spy-UNG2 (Figures 2B and 2C).

Interestingly, the k_{fast} measured in stopped-flow experiments appeared to originate from RPA being present to interact with ssDNA, and because RPA binds tightly to the 32 nt ssDNA section, the fraction of substrate processed with the faster rate should not exceed the molar ratio of RPA to DNA. Stopped-flow assays used a ~two-fold molar excess of substrate compared to protein, and the junction substrates were prepared with a slight excess of unlabeled oligo that contributed non-productive RPA binding sites (see Experimental Section); thus, we estimated a maximum %fast parameter of 43% if each RPA protein interacted tightly with a ssDNA section and stimulated UNG2 for exactly one turnover. However, the %fast was consistently 22% to 26% across all samples containing RPA including the RPA-Spy-UNG2 complex (Figure 2 and Table 1), suggesting some inefficiency in RPA's ability to target UNG2 near a junction. Accordingly, the %fast parameter was also reduced by half in RPA/UNG2 co-injection assays when the protein concentrations were reduced by half (Figure S4).

Activity of the RPA-Spy-UNG2 Complex at ssDNA-dsDNA Junctions with Short ssDNA Sections and on dsDNA

In our experiments using ssDNA-dsDNA junction substrates containing a 32 nt ssDNA section, we found that the RPA-Spy-UNG2 complex was able to identify and excise uracil bases faster than UNG2 or SpyC-UNG2 at the onset of reactions, but that the subsequent turnover of RPA-Spy-UNG2 was significantly slower. We hypothesized that the slower phase of uracil excision by the covalent complex was caused by RPA's high affinity for ssDNA, which could hold the covalently tethered UNG2 near a single junction, thereby limiting the enzyme's ability to unbind product DNA and search for additional uracil sites. Likewise, the significantly slower rate of activity would not occur after wild-type RPA stimulated UNG2 near a junction because the enzyme was able to unbind both RPA and product DNA, quickly freeing the enzyme from a single junction. For perspective, DNA binding rates (k_{on}) for RPA and UNG2 are both fast and diffusion-limited,[48,49] but RPA unbinding from 32 nt ssDNA (k_{off}) is ~100-fold slower than UNG2's off-rate from DNA.[49–51] UNG2 has relatively weak interactions with both ssDNA and dsDNA to facilitate its search for rare uracil bases that may be present among large amounts of DNA.[48,52]

We hypothesized that the k_{slow} for the RPA-Spy-UNG2 complex would accelerate by shortening the ssDNA section next to the DNA junction because RPA affinity would reduce, and the covalent complex would not interact so strongly at a single DNA junction. Shortening the ssDNA from 32 nt to 10 nt reduces the affinity of RPA by ~200-fold which is primarily caused by a faster off-rate.[49,53] As measured with our stopped-flow assay, the activity of the RPA-Spy-UNG2 complex remained biphasic on a junction substrate with a

10 nt ssDNA section and 55 bp duplex (Table 1 and Figure S5). Its k_{slow} quickened and was comparable to free UNG2, which was not significantly affected by ssDNA length (Table 1 and Figure S5). Shortening the ssDNA had only a minor effect on the k_{fast} for the RPA-Spy-UNG2 complex, which remained higher than the rates for other combinations of RPA and UNG2, and the %fast parameter for the complex increased to 37% (Table 1 and Figure S5). The RPA-Spy-UNG2 complex was in fact more efficient over time in this assay compared to UNG2 with or without wild-type RPA, suggesting a substrate-dependent advantage for the protein complex (Table 1 and Figure S5). Interestingly, the activity of UNG2 remained biphasic when co-injected with RPA, but was monophasic when RPA was pre-equilibrated with the DNA substrate (Table 1). The distinct fast and slow rates were apparently eliminated by allowing RPA to reach a dynamic equilibrium with the shorter ssDNA section prior to the introduction of UNG2.

Lastly, we examined the uracil excision activity of the enzymes using stopped-flow with a 55 bp duplex substrate without a ssDNA section. As expected, the rates for UNG2 and SpyC-UNG2 were comparable to those measured on the other substrates that contained a ssDNA section (Table 1 and Figure S6). UNG2 activity was also monophasic in the presence of RPA, but was slightly reduced compared to free UNG2 (Table 1 and Figure S6). The activity of the RPA-Spy-UNG2 complex was the weakest on the duplex substrate, and its rate matched the k_{slow} that we measured for the complex on the junction substrate with a 32 nt ssDNA section (Table 1 and Figure S6). These results were not surprising and were consistent with RPA's role of facilitating repair processes that involve ssDNA.

Following the results from the stopped-flow assays, we returned to our endpoint assay using fluorescein end-labeled DNA substrates to confirm the preference of the covalent RPA-Spy-UNG2 complex for DNA junctions with short ssDNA sections. We prepared a DNA junction substrate with a 10 nt ssDNA section and 55 bp duplex to examine the uracil excision activity of our enzymes 21 bp from the junction. The RPA-Spy-UNG2 complex had markedly improved activity on this junction substrate compared to the substrate with a 32 nt ssDNA section (Figure 3A, for comparison to Figure 2B). A key difference in this experiment compared to stopped-flow assays was that the concentration of the RPA-Spy-UNG2 complex was two orders of magnitude lower than the substrate. The percent of uracil removed by the RPA-Spy-UNG2 complex during the three minute endpoint assay indicated that each molecule of the covalent complex averaged 3.2 turnovers, and its ability to identify and excise uracil bases was similar to UNG2 with or without RPA (Figure 3A). Interestingly, RPA had little effect on UNG2 activity on the junction substrate with the short ssDNA section (Figure 3A). This finding was consistent with the stopped-flow assays where RPA conferred no advantage for UNG2 when pre-equilibrated with the DNA junction substrate containing a 10 nt ssDNA section (Table 1 and Figure S5). This was in stark contrast to RPA's stimulatory effects on UNG2 when RPA was bound to long ssDNA sections next to ssDNA-dsDNA junctions (Figure 2B).[16,17] The high rate of activity for SpyC-UNG2 in Figure 3A also revealed important effects of enzyme and substrate concentration on relative enzyme activities because SpyC-UNG2 and UNG2 were much more similar in the stopped-flow assays (Table 1 and Figure S5). Likewise, the overall activity of the RPA-Spy-UNG2

complex exceeded UNG2 at lower concentrations of DNA substrate (≤ 400 nM), but not higher concentrations (Table 1, Figure 3B, and Figure S5).

Activity of RPA, UNG2, and the RPA-Spy-UNG2 Complex on ssDNA

The activity of RPA, UNG2, and the RPA-Spy-UNG2 complex were examined on fluorescein end-labeled ssDNA substrates of 11 nt, 21 nt, or 31 nt where a uracil base was positioned in the center of the oligonucleotide. UNG2 alone at a concentration of 1 nM had similar uracil excision activity on all of the ssDNA substrates, removing 15–17% of the uracil bases in three minutes (Figure 4). The activity of SpyC-UNG2 was within 2-fold of UNG2 (Figure 4). RPA significantly enhanced UNG2 and allowed nearly complete excision of uracil from ssDNA when RPA was present at a concentration that was molar equivalent to the DNA (0.5 μM) (Figure 4). These stimulatory effects of RPA were somewhat dampened but remained significant when the DNA was in molar excess compared to RPA (Figure S7). Finally, the covalent RPA-Spy-UNG2 complex had similar rates of activity on the ssDNA substrates of different lengths, removing 6–9% of the uracil bases (Figure 4). In these cases, each molecule of RPA-Spy-UNG2 complex turned over 29–44 times over the three minute reaction compared to 75–87 turnovers for UNG2 alone.

An interesting question is why tight binding of RPA to 31 nt ssDNA enhanced the ability of UNG2 to remove uracils in ssDNA as opposed to sterically preventing access to the uracil sites.[19,54] A plausible mechanism that has been extended to other proteins involves the dynamic nature of RPA's four DNA binding domains which allow the protein to handoff ssDNA to other proteins.[55] Even though its DNA binding domains together facilitate strong binding of RPA to ssDNA, the four domains individually undergo rapid dissociation and re-binding events.[51,56–58] The small ssDNA sections that become available when individual DNA binding domains of RPA locally dissociate may be accessed by other proteins such as UNG2.[55] Consistent with this idea, the catalytic domain of UNG2 binds small ssDNA sections that are similar in size to sites occupied by individual DNA binding domains of RPA (4–6 nt) and with similar association rates.[48,49,55,59,60] We speculate that the relatively high turnover by the covalent complex on 31 nt ssDNA indicated that RPA-Spy-UNG2 could participate in a handoff of ssDNA from the RPA portion of the complex to its UNG2 portion. This would imply that the two proteins do not need to physically dissociate for the handoff to occur, and that the winged-helix domain of RPA2 is not strictly required for the handoff other than to serve as a means for RPA to spatially recruit UNG2 to the ssDNA substrate.

Dissociation of RPA and UNG2 Facilitates UNG2 Excision of Multiple Uracil Sites Across ssDNA-dsDNA Junctions

Complex formation and dissociation with RPA could be important for the ability of UNG2 to remove multiple uracil bases located at different places in the same DNA molecule. After UNG2 binds DNA and removes a uracil base, it could translocate down the same DNA strand to remove a different uracil base, or else it could unbind from the DNA substrate and diffuse away to randomly find another.[43,61] It would also be possible that uracil at one site—or the product that arises from its excision—would affect the ability of the enzyme to excise uracil at a different site. This is relevant because multiple uracils can be found clustered in DNA during somatic hypermutation and class-switch recombination, and also

during treatments with thymidylate synthase inhibitors.[43,48,61,62] Excision of uracil sites across a ssDNA-dsDNA junction could be measured with a junction substrate containing uracil in both the ssDNA and dsDNA sections (Figure 5). Excision of the uracil bases produces a pattern of DNA fragments visible by fluorescein end-labels, and their relative intensities reflect the fraction of DNA molecules that had either one or two uracils removed in the presence and absence of RPA (Figure 5).[63–65] This substrate also measured the selectivity of UNG2 for the two uracil sites in a single assay. Regarding the experiment in Figure 5, selectivity for the ssDNA site would result in an abundance of fragments AB and C relative to BC and A, and the opposite would be observed if the dsDNA site was preferred. Concurrently, excision of both uracil sites in the junction DNA would result in an abundance of smaller fragments A and C relative to fragments AB and BC, the latter resulting from single excision. For comparative purposes, we calculated the fraction of product that resulted from double uracil excision events ($F_{\text{double excision}}$) in our endpoint assays (Figure 5).

As expected from previous assays, the enzymes had a strong preference for uracil bases in the ssDNA section compared to the dsDNA section (Figure 5 and Figure S8). UNG2 alone produced a low fraction of DNA with both uracils removed ($F_{\text{double excision}} = 0.02$) which was not surprising considering its strong preference for ssDNA and the spacing between the uracil sites (37 nt), and it is well characterized that shorter distances between uracils would favor double excision events.[43,61] However, RPA enhanced the likelihood of UNG2 producing double uracil excision across the ssDNA-dsDNA junction ($F_{\text{double excision}} = 0.41$) (Figure 5). The activity of 4 nM UNG2 at the ssDNA site was within two-fold of 1 nM UNG2 with 0.5 μM RPA, yet the appearance of fragment A in the presence of RPA also revealed efficient excision at the dsDNA site. In contrast, the fraction of substrate with both uracils removed was 0.08 for the RPA-Spy-UNG2 complex, which was four-fold better UNG2 alone, but this remained much lower than the combination of wild-type RPA and UNG2. Thus, excision of multiple uracil sites by UNG2 across ssDNA-dsDNA junctions was specifically enhanced by RPA and was facilitated by dissociation of UNG2 from RPA.

Conclusion

DNA repair proteins are thought to form complexes in the nucleus, and this work revealed how the activity of UNG2 is affected by forming short- and long-lived complexes with its binding partner, RPA. In experiments that measured the ability of UNG2 to turn over multiple times and continuously identify uracil bases among bulk DNA, the kinetics of uracil excision were generally faster when wild-type RPA and UNG2 interacted reversibly as opposed to when they were joined as a complex in RPA-Spy-UNG2. However, protein complex formation may confer advantages in specific cellular contexts where multiple turnovers are not essential such as when a rare uracil base is present next to a ssDNA-dsDNA junction. The RPA-Spy-UNG2 complex also had much slower turnover at DNA junctions with specific structures. If this reflected slow unbinding from the junction, this could theoretically have kinetic advantages for the downstream processing of the abasic sites created by UNG2 if other proteins joined their complex. The small DNA repair complex that we designed and assembled using SpyCatcher/SpyTag ligation had properties that were an accurate composite of the individual UNG2 and RPA proteins. Orthogonal ligation systems such as the SnoopLigase/DogTag system might be employed to create larger assemblies of

proteins involved in DNA repair or other processes to understand how proteins affect each other as part of complexes.[40,41]

Interestingly, the covalent RPA-Spy-UNG2 complex that we produced also proved to be a useful tool for understanding the different mechanisms used by RPA to target UNG2 activity towards uracil bases in ssDNA compared to its ability to target UNG2 activity to duplex areas adjacent to ssDNA-dsDNA junctions. RPA had a remarkable stimulatory effect on UNG2 during excision of uracil from ssDNA. This requires a means for RPA to recruit UNG2 to ssDNA, then RPA must allow UNG2 access to the ssDNA that RPA also binds with higher affinity. RPA also stimulated UNG2 activity when uracils were present on both sides of a ssDNA-dsDNA junction, but this was weakened when the proteins were unable to dissociate as in the covalent complex. Our study illustrates how engineered protein complexes containing RPA and other hub proteins involved in DNA repair have the potential to improve our mechanistic understanding of a variety of processes that occur on different DNA structures found in the nucleus.

Experimental Section

Production of Recombinant Proteins

A detailed protocol for the bacterial expression and purification of full-length human UNG2 was described elsewhere.[22] This method expressed UNG2 with a 8xHis-SUMO tag fused to its N-terminus, and the 8xHis-SUMO tag was removed during purification of the enzyme.[22] BL21(DE3)pLysS cells were used for UNG2 expression, and the gene was encoded in a pET21a vector. SpyC-UNG2 was cloned into the same vector using restriction free/megaprimer cloning as follows.[66] N-terminal domain residues 1–91 were deleted from the UNG2 gene without disrupting the 8xHis-SUMO tag or the UNG2 catalytic domain (residues 92–313). Following the 8xHis-SUMO tag, we inserted a four residue spacer followed by SpyCatcher residues 22–104 (numbering according to PDB code 4MLI),[32] which was followed by a ten residue spacer and UNG2 residues 92–313. The final sequence was verified with Sanger sequencing and is reported in the Supporting Information. The protein expressed in BL21(DE3)pLysS cells and was purified using the same procedures as wild-type UNG2, which included the removal of the 8xHis-SUMO tag.[22] The yield for purified SpyC-UNG2 was 4.5 mg of protein per L of bacterial culture.

Recombinant wild-type human RPA was expressed in BL21(DE3)CodonPlus cells using the p11d-tRPA plasmid,[67,68] which was a generous gift from Dr. Marc Wold. Purification of RPA followed our standard chromatography procedures using Affi-Gel Blue, hydroxyapatite, Mono Q, and gel filtration columns.[17,67,68] To produce RPA-SpyTag, we used restriction free/megaprimer cloning with the p11d-tRPA plasmid to delete RPA2 residues 205–270 (essentially, its winged-helix domain), and we replaced it with the SpyTag sequence SGAHIVMVDAYKPTK. RPA1 and RPA3 remained as wild-type sequences in the p11d-tRPA plasmid and in the heterotrimeric RPA-SpyTag protein, which was expressed in bacteria and purified identically to wild-type RPA. The final sequence of RPA2 modified with SpyTag was verified with Sanger sequencing and is reported in the Supporting Information.

As mentioned in the main text, *initial* attempts to create an obligate RPA-UNG2 complex without the SpyCatcher/SpyTag system involved genetically fusing the UNG2 catalytic domain (amino acids 92–313) to the C-terminus of RPA2. This was performed by genetically inserting the UNG2 catalytic domain into the p11d-tRPA plasmid immediately after the RPA2 winged-helix domain, or alternatively, we replaced the RPA2 winged-helix domain with the UNG2 catalytic domain. Neither fusion protein expressed under conditions used for RPA and RPA-SpyTag.

Production of the Recombinant Covalent RPA-Spy-UNG2 Complex

For preparative scale reactions, 9.1 μM of RPA-SpyTag was gently mixed with 8.2 μM of SpyC-UNG2 in a buffer with final concentrations of 7% glycerol, 15 mM Tris-Cl (pH 7.4), 10 mM sodium phosphate, 200 mM NaCl, 0.1 mM EDTA, and 1 mM DTT. The protein components were left to react at room temperature for 16 hours, then the covalent RPA-Spy-UNG2 complex was purified with SEC using a mobile phase of 5% glycerol, 10 mM Tris-Cl (pH 8.0), 200 mM NaCl, 0.1 mM EDTA, and 1 mM DTT. SEC was performed using a Bio-Rad NGC chromatography system with a Bio-Rad Enrich SEC 650 column (10 × 300 mm). The elution of the RPA-Spy-UNG2 complex on SEC was compared to RPA and SEC standards under identical chromatography conditions, and the SEC standards were from Sigma-Aldrich (catalog #69385). After SEC, the purified covalent RPA-Spy-UNG2 complex was concentrated to 12 μM with an Amicon 10 kDa MW cutoff centrifugal filter, then the protein was aliquoted for single use, snap frozen in liquid nitrogen, and stored at −80°C.

Preliminary ligation reactions used different protein concentrations (9 μM-15 μM), different buffer conditions (5%−10% glycerol, 10 mM-20 mM Tris, 0 mM-10 mM sodium phosphate, and 0 mM-0.1 mM EDTA), and different reaction times (4 hours-24 hours). The reaction efficiency was identical under all conditions and depended on the molar ratio of RPA-SpyTag to SpyC-UNG2. The ligation reaction occurred with >95% efficiency when the protein components were present at a 1:1 molar ratio; however, a slight molar excess of RPA-SpyTag was used in the preparative reactions to eliminate the possibility of residual activity from unreacted SpyC-UNG. Regardless, the RPA-Spy-UNG2 complex could be purified from the unreacted components (including RPA-SpyTag) during the final SEC step.

Uracil Excision Assays

Uracil excision assays were performed with synthetic oligonucleotides that were purchased from Integrated DNA Technologies and purified by denaturing (Urea-TBE) PAGE. The full sequences of all oligonucleotide substrates can be found in the Supporting Information. Where applicable, fluorescein end-labels were attached to the phosphate of the terminal nucleotide as phosphodiester linkages, and a six-carbon spacer was between the phosphate and fluorescein. For annealing, complementary strands were mixed, heated to 95°C for five minutes, then slowly cooled to room temperature. In all cases, the unlabeled oligonucleotide lacking either fluorescein or 2-aminopurine was used during annealing with an 8% molar excess over the labeled oligonucleotide to ensure that all of the labeled strand, which reported enzymatic activity, was in duplex form. The annealing buffer contained 10 mM Tris-Cl (pH 8.0), 100 mM NaCl, and 0.1 mM EDTA.

Endpoint UNG2 assays using fluorescein end-labeled DNA substrates and denaturing PAGE to analyze quenched reactions were performed as previously described.[16,17] These reactions were performed at 22°C in a buffer of 10 mM Tris-Cl (pH 8.0), 100 mM NaCl, 0.1 mM EDTA, and 1 mM DTT, and the concentration of enzyme and substrate used in each reaction was reported in the appropriate Figure and/or Figure Legend. An Azure c400 imager was used to visualize the fluorescein-labeled DNA oligonucleotides, and band intensity was quantified using Fiji/ImageJ.[69] This methodology was also used in endpoint assays that used DNA substrate containing two uracil bases where we measured the fraction of product that resulted from double uracil excision. Note that the assays and equation in Figure 5 resembled experiments that measure DNA translocation or correlated cleavage, which describes the probability of UNG2 excising uracil bases at different sites upon the same encounter with the DNA molecule.[43,61,63–65] DNA translocation experiments require time point analyses and accounting for the intrinsic preference of the enzyme for one site over the other, which was not investigated here.[63–65] Finally, the number of turnovers per enzyme molecule in an assay was calculated by dividing the total moles of substrate processed in the reaction by the total moles of enzyme that were present.

Stopped-flow enzyme assays were performed with an Applied Photophysics RX2000 device in two-syringe mode with its cuvette and trigger connected to a Horiba Fluoromax 4 Spectrofluorometer. The fluorometer was set to excitation/emission wavelengths of 311 nm/360 nm with slit widths of 2 nm/4 nm. Assays were performed at 22°C in a buffer of 10 mM Tris-Cl (pH 8.0), 100 mM NaCl, 0.1 mM EDTA, 5 mM $MgCl_2$, and 1 mM DTT. The stopped-flow assays used DNA substrates containing a uracil/2-aminopurine bp placed within a 55 bp duplex. The 55 bp duplex was adjacent to a 10 nt or 32 nt ssDNA section that terminated with a 3′ end, or alternatively, the ssDNA section was omitted to examine a dsDNA substrate. The uracil base in the duplex was positioned 21 bp from the ssDNA-dsDNA junction and was placed in the strand that was opposite from the ssDNA section.

The final DNA substrate concentration in stopped-flow assays was 0.4 μM, and the final concentration of included protein components (UNG2, RPA, SpyC-UNG2, or RPA-Spy-UNG2 complex) was 0.2 μM each. For example, in experiments where RPA and DNA were equilibrated prior to the assays, these components were mixed at room temperature (i.e., 0.4 μM RPA and 0.8 μM DNA) for at least ten minutes prior to placing their solution in a stopped-flow syringe. The other syringe contained 0.4 μM UNG2, and the 1:1 mixture of the syringe components in the cuvette to initiate the reaction resulted in final concentrations of 0.2 μM of each protein and 0.4 μM of DNA. The experimental design ensured that the same number of UNG2 catalytic sites were present in each assay which was critical for comparing activity rates (k, k_{fast}, and k_{slow}). Likewise, the concentration of RPA when included was equivalent to the concentration of RPA-Spy-UNG2 complex used in its assays to ensure that the same number of RPA proteins were present for each condition. Upon initiating the enzyme reaction with the stopped-flow device, fluorescence intensity data was collected at 10 millisecond intervals for 60 seconds (the stopped-flow deadtime was 8 milliseconds). The hyberbolic fluorescence intensity data was plotted on the y axis versus time on the x axis using GraphPad Prism 7, and the data was fit with a curve using

either a one-phase exponential equation or a two-phase exponential equation. The one-phase exponential equation was

$$y = y_{min} + (y_{max} - y_{min}) * \left(1 - e^{-k * x}\right) \quad (1)$$

where y_{min} was initial background fluorescence intensity, $y_{max}$was the maximum fluorescence intensity when all the substrate was processed, and k was the rate constant. The two-phase exponential equation was

$$y = y_{min} + SpanFast * \left(1 - e^{-k_{fast} * x}\right) + SpanSlow * \left(1 - e^{-k_{slow} * x}\right) \quad (2)$$

$$SpanFast = (y_{max} - y_{min}) * \%fast * 0.01$$

$$SpanSlow = (y_{max} - y_{min}) * (100 - \%fast) * 0.01$$

where y_{min} was again the initial background fluorescence intensity, $y_{max}$was the maximum fluorescence of all processed substrate, SpanFast represented the fluorescence intensity change attributed to the fast phase, $k_{fast}$was the rate constant for the fast phase, %fast was the percent of the total fluorescence intensity change $(y_{max} - y_{min})$ attributed to fast phase, $k_{slow}$was the rate constant for the slow phase, and SpanSlow represented the fluorescence intensity change attributed to the slow phase. To convert the data from a change in fluorescence intensity over time to μM substrate processed over time, the background fluorescence y_{min} was subtracted from the data, and the fluorescence intensity values were multiplied by $(0.4\mu M/(y_{max} - y_{min}))$ (note that this conversion of y axis units has no effect on the curve fitting or the parameters from the fitted curves). All data sets were tested with both one-phase and two-phase exponential equations. For data that we report as monophasic, the %fast values were negligible (less than 2%) if we attempted to fit two-phase exponential curves to the data. The quality of curves fit to stopped-flow data was manually confirmed and was easily distinguished by examining the first ~500 ms of the reaction (for example, see Figure S9). The amount of enzyme and DNA in stopped-flow experiments were chosen such that complete excision of uracil sites would occur over the experimental time frame for all tested protein combinations, which was essential to obtain y_{max} and rate constants. Note that the assay conditions used here in stopped-flow experiments (0.4 μM DNA substrate and 0.2 μM enzyme) do not satisfy conditions for steady state kinetics (E << S) or single-turnover kinetics (E >> S) which should be considered when interpreting the relevance of the rate constants to other systems.

As mentioned above and in the Results and Discussion, an 8% excess of the unlabeled oligonucleotide without 2-aminopurine was used during annealing of the substrates used for stopped-flow assays. Because these assays used 400 nM DNA junction substrate, this resulted in 32 nM of free unlabeled oligonucleotide being present in the enzyme reactions. This unlabeled oligonucleotide was 55 nt long which was sufficient for two high-affinity RPA binding sites.[53] Thus, the reactions using the DNA junction substrate containing the

32 nt ssDNA section contained 464 nM of high-affinity RPA sites. If all of the 200 nM RPA or 200 nM RPA-Spy-UNG2 complex used in the assays interacted tightly with ssDNA, then 43% of the DNA junction substrate would be expected to interact with the RPA proteins (200 nM/464 nM * 100%). Thus, the maximum amount of uracil that could be removed with the fast rate (%fast) driven by a high-affinity interaction between RPA and ssDNA next to a junction was 43%.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by National Institutes of Health grant R01GM135152, the Osteopathic Heritage Foundation, and a Rowan University IAAME award.

Corresponding Author Biography

Brian Weiser received his PhD in Pharmacology at the University of Pennsylvania where he studied the pharmacology of anesthetics under Dr. Roderic Eckenhoff. This was followed by a postdoc at Johns Hopkins University under the joint mentorship of Drs. Philip Cole and James Stivers. The Weiser Lab, which started at Rowan University in 2018, has diverse interests including how protein-protein interactions affect DNA repair machinery, understanding the molecular pharmacology of sirtuin deacylase enzymes, and developing statistical methods for analyzing non-canonical dose-response relationships.

References

[1]. Li CM, Miao Y, Lingeman RG, Hickey RJ, Malkas LH, PLOS ONE 2016, 11, e0169259. [PubMed: 28036377]

[2]. Sirbu BM, McDonald WH, Dungrawala H, Badu-Nkansah A, Kavanaugh GM, Chen Y, Tabb DL, Cortez D, J. Biol. Chem 2013, jbc.M113.511337.

[3]. Parlanti E, Locatelli G, Maga G, Dogliotti E, Nucleic Acids Res 2007, 35, 1569–1577. [PubMed: 17289756]

[4]. Akbari M, Solvang-Garten K, Hanssen-Bauer A, Lieske NV, Pettersen HS, Pettersen GK, Wilson DM, Krokan HE, Otterlei M, DNA Repair (Amst.) 2010, 9, 785–795. [PubMed: 20466601]

[5]. Kumar A, Priya A, Ahmed T, Grundström C, Negi N, Grundström T, The Journal of Immunology 2018, ji1701586.

[6]. Fanning E, Klimovich V, Nager AR, Nucleic Acids Res 2006, 34, 4126–4137. [PubMed: 16935876]

[7]. Maréchal A, Zou L, Cell Research 2015, 25, 9–23. [PubMed: 25403473]

[8]. Bhat KP, Bétous R, Cortez D, J. Biol. Chem 2015, 290, 4110–4117. [PubMed: 25552480]

[9]. Witosch J, Wolf E, Mizuno N, Nucleic Acids Res 2014, 42, 12912– 12927. [PubMed: 25348395]

[10]. Lee Y-C, Zhou Q, Chen J, Yuan J, Curr. Biol 2016, 26, 3257–3268. [PubMed: 27818175]

[11]. Bass TE, Luzwick JW, Kavanaugh G, Carroll C, Dungrawala H, Glick GG, Feldkamp MD, Putney R, Chazin WJ, Cortez D, Nat. Cell Biol 2016, 18, 1185–1195. [PubMed: 27723720]

[12]. Haahr P, Hoffmann S, Tollenaere MAX, Ho T, Toledo LI, Mann M, Bekker-Jensen S, Räschle M, Mailand N, Nat. Cell Biol 2016, 18, 1196–1207. [PubMed: 27723717]

[13]. Gibb B, Ye LF, Kwon Y, Niu H, Sung P, Greene EC, Nat. Struct. Mol. Biol 2014, 21, 893–900. [PubMed: 25195049]

[14]. Ma CJ, Kwon Y, Sung P, Greene EC, J. Biol. Chem 2017, 292, 11702–11713. [PubMed: 28551686]

[15]. Torseth K, Doseth B, Hagen L, Olaisen C, Liabakk N-B, Græsmann H, Durandy A, Otterlei M, Krokan HE, Kavli B, Slupphaug G, DNA Repair (Amst.) 2012, 11, 559–569. [PubMed: 22521144]

[16]. Weiser BP, Rodriguez G, Cole PA, Stivers JT, Nucleic Acids Res 2018, 46, 7169–7178. [PubMed: 29917162]

[17]. Weiser BP, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 2020, 1868, 140347. [PubMed: 31866506]

[18]. de Laat WL, Appeldoorn E, Sugasawa K, Weterings E, Jaspers NG, Hoeijmakers JH, Genes Dev 1998, 12, 2598–2609. [PubMed: 9716411]

[19]. Kavli B, Iveland TS, Buchinger E, Hagen L, Liabakk NB, Aas PA, Obermann TS, Aachmann FL, Slupphaug G, Nucleic Acids Res 2021, 49, 3948–3966. [PubMed: 33784377]

[20]. Mer G, Bochkarev A, Gupta R, Bochkareva E, Frappier L, Ingles CJ, Edwards AM, Chazin WJ, Cell 2000, 103, 449–456. [PubMed: 11081631]

[21]. Xie S, Lu Y, Jakoncic J, Sun H, Xia J, Qian C, FEBS J 2014, 281, 3382–3396. [PubMed: 24910198]

[22]. Weiser BP, Stivers JT, Cole PA, Biophys. J 2017, 113, 393–401. [PubMed: 28746850]

[23]. Richter K, Nessling M, Lichter P, Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 2008, 1783, 2100–2107. [PubMed: 18723053]

[24]. Otterlei M, Warbrick E, Nagelhus TA, Haug T, Slupphaug G, Akbari M, Aas PA, Steinsbekk K, Bakke O, Krokan HE, EMBO J 1999, 18, 3834–3844. [PubMed: 10393198]

[25]. Akbari M, Otterlei M, Peña-Diaz J, Aas PA, Kavli B, Liabakk NB, Hagen L, Imai K, Durandy A, Slupphaug G, Krokan HE, Nucleic Acids Res 2004, 32, 5486–5498. [PubMed: 15479784]

[26]. Hanssen-Bauer A, Solvang-Garten K, Sundheim O, Peña-Diaz J, Andersen S, Slupphaug G, Krokan HE, Wilson DM, Akbari M, Otterlei M, Environ Mol Mutagen 2011, 52, 623–635. [PubMed: 21786338]

[27]. Krokan HE, Bjørås M, Cold Spring Harb Perspect Biol 2013, 5, a012583. [PubMed: 23545420]

[28]. Hagen L, Kavli B, Sousa MML, Torseth K, Liabakk NB, Sundheim O, Pena-Diaz J, Otterlei M, Hørning O, Jensen ON, Krokan HE, Slupphaug G, EMBO J 2008, 27, 51–61. [PubMed: 18079698]

[29]. Lu X, Bocangel D, Nannenga B, Yamaguchi H, Appella E, Donehower LA, Molecular Cell 2004, 15, 621–634. [PubMed: 15327777]

[30]. Zakeri B, Fierer JO, Celik E, Chittock EC, Schwarz-Linek U, Moy VT, Howarth M, Proc. Natl. Acad. Sci. U.S.A 2012, 109, E690–697. [PubMed: 22366317]

[31]. Reddington SC, Howarth M, Curr Opin Chem Biol 2015, 29, 94–99. [PubMed: 26517567]

[32]. Li L, Fierer JO, Rapoport TA, Howarth M, J Mol Biol 2014, 426, 309–317. [PubMed: 24161952]

[33]. Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T, Neuberger MS, Curr. Biol 2002, 12, 1748–1755. [PubMed: 12401169]

[34]. Kavli B, Andersen S, Otterlei M, Liabakk NB, Imai K, Fischer A, Durandy A, Krokan HE, Slupphaug G, J. Exp. Med 2005, 201, 2011–2021. [PubMed: 15967827]

[35]. Zan H, White CA, Thomas LM, Mai T, Li G, Xu Z, Zhang J, Casali P, Cell Rep 2012, 2, 1220–1232. [PubMed: 23140944]

[36]. Seiple L, Jaruga P, Dizdaroglu M, Stivers JT, Nucleic Acids Res 2006, 34, 140–151. [PubMed: 16407331]

[37]. Fischer JA, Muller-Weeks S, Caradonna SJ, Cancer Res 2006, 66, 8829–8837. [PubMed: 16951200]

[38]. Weeks LD, Zentner GE, Scacheri PC, Gerson SL, Cell Death Dis 2014, 5, e1045. [PubMed: 24503537]

[39]. Huehls AM, Huntoon CJ, Joshi PM, Baehr CA, Wagner JM, Wang X, Lee MY, Karnitz LM, Mol. Pharmacol 2016, 89, 53–62. [PubMed: 26494862]

[40]. Veggiani G, Nakamura T, Brenner MD, Gayet RV, Yan J, Robinson CV, Howarth M, Proc. Natl. Acad. Sci. U.S.A 2016, 113, 1202–1207. [PubMed: 26787909]

[41]. Buldun CM, Jean JX, Bedford MR, Howarth M, J. Am. Chem. Soc 2018, 140, 3008–3018. [PubMed: 29402082]

[42]. Nagelhus TA, Haug T, Singh KK, Keshav KF, Skorpen F, Otterlei M, Bharati S, Lindmo T, Benichou S, Benarous R, Krokan HE, J. Biol. Chem 1997, 272, 6561–6566. [PubMed: 9045683]

[43]. Rodriguez G, Esadze A, Weiser BP, Schonhoft JD, Cole PA, Stivers JT, ACS Chem. Biol 2017, 12, 2260–2263. [PubMed: 28787572]

[44]. Rodriguez G, Orris B, Majumdar A, Bhat S, Stivers JT, DNA Repair (Amst) 2020, 86, 102764. [PubMed: 31855846]

[45]. Coey CT, Drohat AC, in Methods in Enzymology (Ed.: Eichman BF), Academic Press, 2017, pp. 357–376.

[46]. Sassa A, Beard WA, Shock DD, Wilson SH, J Vis Exp 2013, e50695. [PubMed: 23995844]

[47]. Stivers JT, Nucl. Acids Res 1998, 26, 3837–3844. [PubMed: 9685503]

[48]. Schonhoft JD, Stivers JT, Biochemistry 2013, 52, 2536–2544. [PubMed: 23506270]

[49]. Patrick SM, Turchi JJ, J Biol Chem 2001, 276, 22630–22637. [PubMed: 11278662]

[50]. Guo S, Zhang Y, Yuan F, Gao Y, Gu L, Wong I, Li G-M, J Biol Chem 2006, 281, 21607–21616. [PubMed: 16731533]

[51]. Chen R, Subramanyam S, Elcock AH, Spies M, Wold MS, Nucleic Acids Res 2016, 44, 5758–5772. [PubMed: 27131385]

[52]. Cravens SL, Hobson M, Stivers JT, Biochemistry 2014, 53, 7680–7692. [PubMed: 25408964]

[53]. Kim C, Paulus BF, Wold MS, Biochemistry 1994, 33, 14197–14206. [PubMed: 7947831]

[54]. Fan J, Pavletich NP, Genes Dev 2012, 26, 2337–2347. [PubMed: 23070815]

[55]. Caldwell CC, Spies M, Crit Rev Biochem Mol Biol 2020, 55, 482–507. [PubMed: 32856505]

[56]. Pokhrel N, Origanti S, Davenport EP, Gandhi D, Kaniecki K, Mehl RA, Greene EC, Dockendorff C, Antony E, Nucleic Acids Res 2017, 45, 9413–9426. [PubMed: 28934470]

[57]. Gibb B, Ye LF, Gergoudis SC, Kwon Y, Niu H, Sung P, Greene EC, PLoS ONE 2014, 9, e87922. [PubMed: 24498402]

[58]. Wieser TA, Wuttke DS, Biochemistry 2022, 61, 2592–2606. [PubMed: 36278947]

[59]. Bochkarev A, Pfuetzner RA, Edwards AM, Frappier L, Nature 1997, 385, 176–181. [PubMed: 8990123]

[60]. Chembazhi UV, Patil VV, Sah S, Reeve W, Tiwari RP, Woo E, Varshney U, Nucleic Acids Res 2017, 45, 5863–5876. [PubMed: 28369586]

[61]. Schonhoft JD, Stivers JT, Nat. Chem. Biol 2012, 8, 205–210. [PubMed: 22231272]

[62]. Cravens SL, Schonhoft JD, Rowland MM, Rodriguez AA, Anderson BG, Stivers JT, Nucleic Acids Res 2015, 43, 4087–4097. [PubMed: 25845592]

[63]. Sidorenko VS, Mechetin GV, Nevinsky GA, Zharkov DO, FEBS Lett 2008, 582, 410–414. [PubMed: 18201572]

[64]. Hedglin M, O'Brien PJ, Biochemistry 2008, 47, 11434–11445. [PubMed: 18839966]

[65]. Porecha RH, Stivers JT, Proc Natl Acad Sci U S A 2008, 105, 10791–10796. [PubMed: 18669665]

[66]. Bond SR, Naus CC, Nucleic Acids Res 2012, 40, W209–213. [PubMed: 22570410]

[67]. Binz SK, Dickson AM, Haring SJ, Wold MS, Meth. Enzymol 2006, 409, 11–38.

[68]. Henricksen LA, Umbricht CB, Wold MS, J. Biol. Chem 1994, 269, 11121–11132. [PubMed: 8157639]

[69]. Cardona A, Schmid B, Rueden C, White DJ, Frise E, Arganda- Carreras I, Tinevez J-Y, Schindelin J, Eliceiri K, Longair M, Tomancak P, Preibisch S, Saalfeld S, Pietzsch T, Kaynig V, Hartenstein V, Nature Methods 2012, 9, 676. [PubMed: 22743772]

[70]. Esadze A, Rodriguez G, Weiser BP, Cole PA, Stivers JT, Nucleic Acids Res 2017, 45, 12413–12424. [PubMed: 29036472]

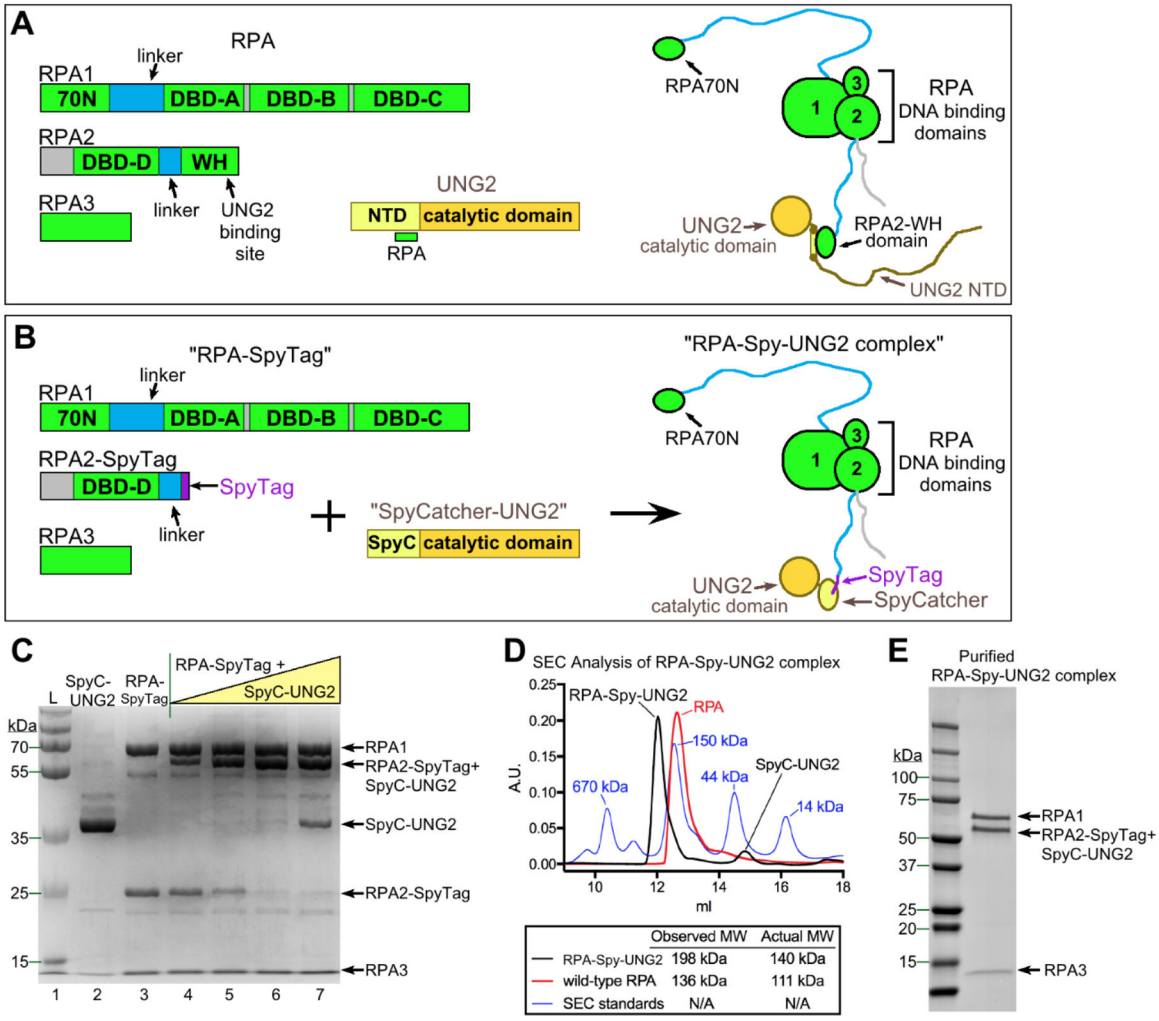


Figure 1.

Architecture and generation of a covalent RPA-Spy-UNG2 protein complex. (A) Schematic of wild-type RPA subunits, wild-type UNG2, and a model of the protein complex that forms upon their interaction. DBD, DNA binding domain; WH, winged-helix domain; NTD, N-terminal domain. (B) Schematic of RPA-SpyTag subunits, SpyC-UNG2, and a model of the covalent RPA-Spy-UNG2 protein complex that forms upon their interaction. For perspective, the SpyCatcher domain is 9 kDa and the RPA2-WH domain is 7 kDa. (C) Coomassie-stained gel of recombinant proteins (lanes 2–3) and the ligation of 15 μM RPA-SpyTag with increasing amounts of SpyCatcher-UNG2 (6 μM-22 μM, lanes 4–7). The proteins were combined with 1:1 molar stoichiometry in lane 6, and all reactions were 4 hours. The ligation was evident by the shift of RPA2-SpyTag to a higher MW band that also contained SpyC-UNG2. Lane 1 contained a MW ladder. (D) SEC trace showing the elution of the covalent RPA-Spy-UNG2 complex following a test reaction with excess SpyC-UNG2. The elution of RPA and SEC standards are shown for comparison. (E) Coomassie-stained gel of purified RPA-Spy-UNG2 complex.

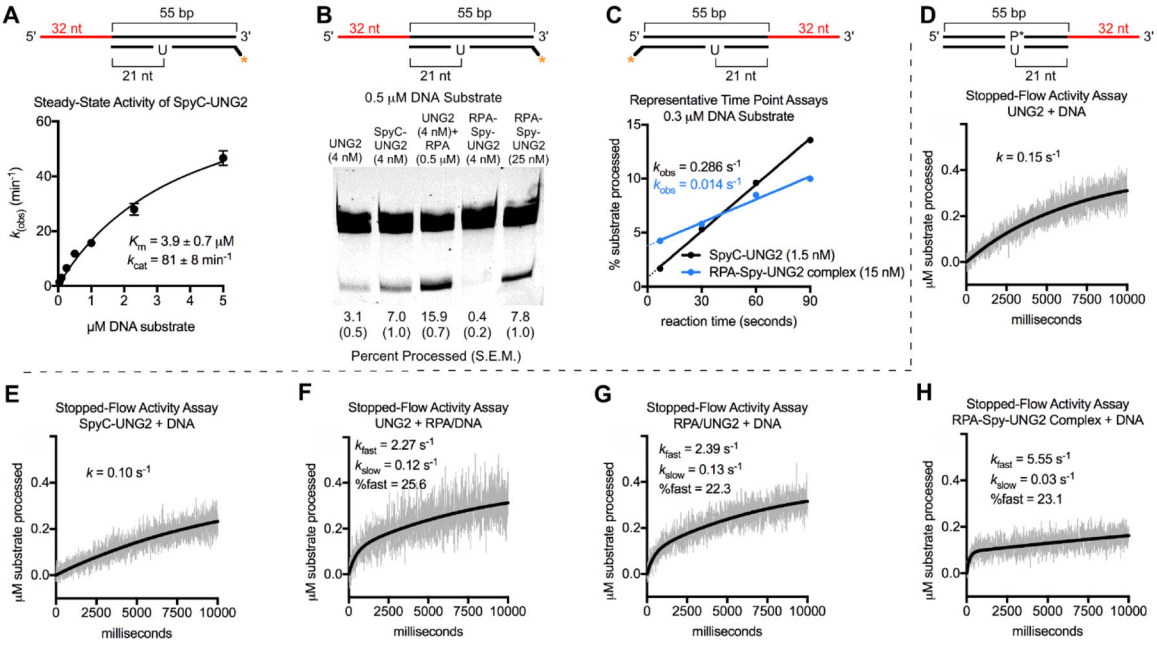


Figure 2.

Activity of UNG2 enzymes with or without RPA, and the activity of the covalent RPA-Spy-UNG2 complex. (A) Steady-state kinetics of uracil excision for SpyC-UNG2 using the substrate shown in the panel. The standard error of the kinetic parameters is also shown. (B) The relative uracil excision activity for the indicated protein combinations was visualized with fluorescence following Urea-TBE PAGE. The top band is unreacted substrate, and the bottom band is a product that results from uracil excision by UNG2 and cleavage of abasic sites.[16,70] The reaction time was three minutes, and the percent processed is an average from three independent assays; standard error of the mean (S.E.M.) is shown in parenthesis. When added, RPA was pre-equilibrated with the substrate prior to UNG2 addition. (C) Time point assays measuring the activity of SpyC-UNG2 or the covalent RPA-Spy-UNG2 complex. The dashed lines in the graphs extend to the y-intercept at values of 3.80% (RPA-Spy-UNG2) and 0.86% (SpyC-UNG2). Similar time point assays are shown in Figure S2. (D) Stopped-flow uracil excision assay using UNG2 and the indicated substrate containing 2-aminopurine (P*). The UNG2 concentration was 200 nM, and the DNA concentration was 400 nM. The activity was monophasic (one rate). (E) Stopped-flow uracil excision assay using SpyC-UNG2 (200 nM) and the substrate shown in panel D (400 nM). The activity was monophasic. (F) Stopped-flow assay where RPA (200 nM) was pre-equilibrated with the DNA substrate in panel D (400 nM) prior to the introduction of UNG2 (200 nM). The activity was biphasic (two rates). (G) Stopped-flow assay where RPA (200 nM) and UNG2 (200 nM) were simultaneously introduced to the DNA substrate shown in panel D (400 nM). The activity was biphasic. (H) Stopped-flow assay using RPA-Spy-UNG2 (200 nM) and the substrate shown in panel D (400 nM). The activity was biphasic. In panels D-H, the gray bars represent standard error for the average amount of DNA processed at every 10 millisecond time point, and the black line is the curve that was fit to the data. At least three independent replicates were performed for each panel. Standard error for the kinetic values

can be found in Table 1. This data represents the first 10 seconds of a 60 second time course, which can be found in Figure S3.

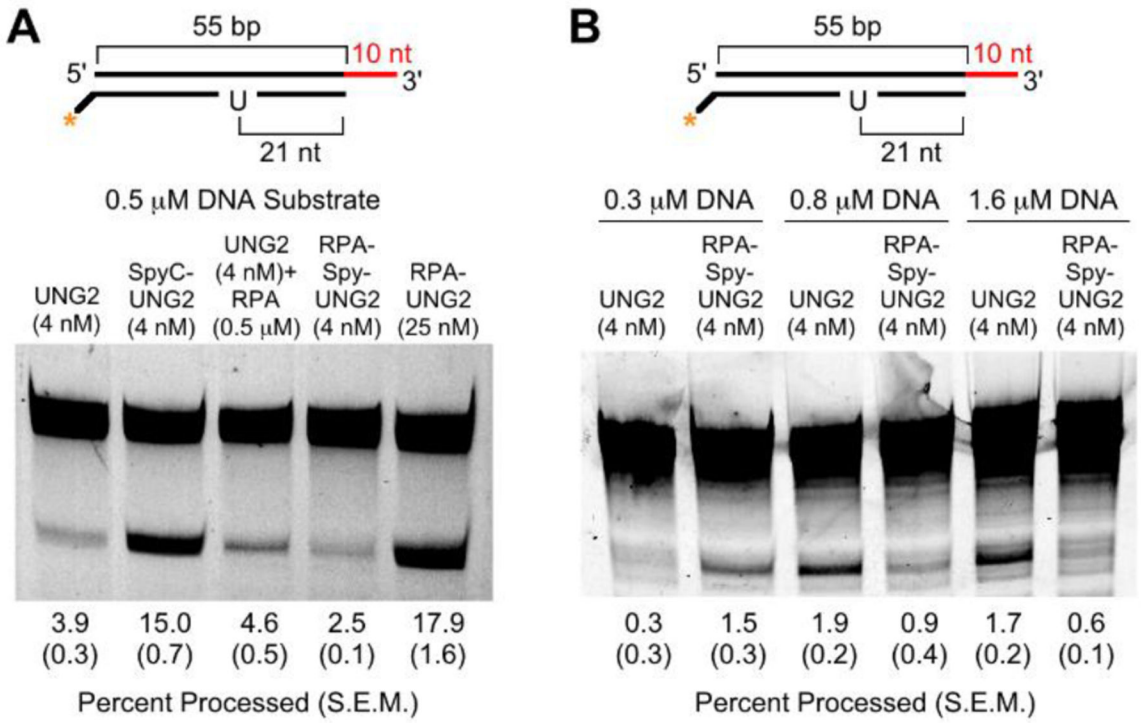


Figure 3.

Uracil excision activity for UNG2 enzymes including the covalent RPA-Spy-UNG2 complex using endpoint assays and a DNA junction substrate containing a 55 bp duplex and a 10 nt ssDNA section with uracil positioned 21 bp from the junction. (A) The relative activity for the indicated protein combinations was visualized with in-gel fluorescence following Urea-TBE PAGE. The reaction time was three minutes, and the percent processed is an average from three independent assays; standard error of the mean (S.E.M.) is shown in parenthesis. When added, RPA was pre-equilibrated with the substrate prior to UNG2 addition, and "RPA-Spy-UNG2" refers to the covalent protein complex. (B) Relative activity for UNG2 and the covalent RPA-Spy-UNG2 complex at different DNA substrate concentrations. For the gel, the same total amount of DNA was run in each lane. The reaction time was three minutes, and the percent processed is an average from three independent assays; standard error of the mean (S.E.M.) is shown in parenthesis.

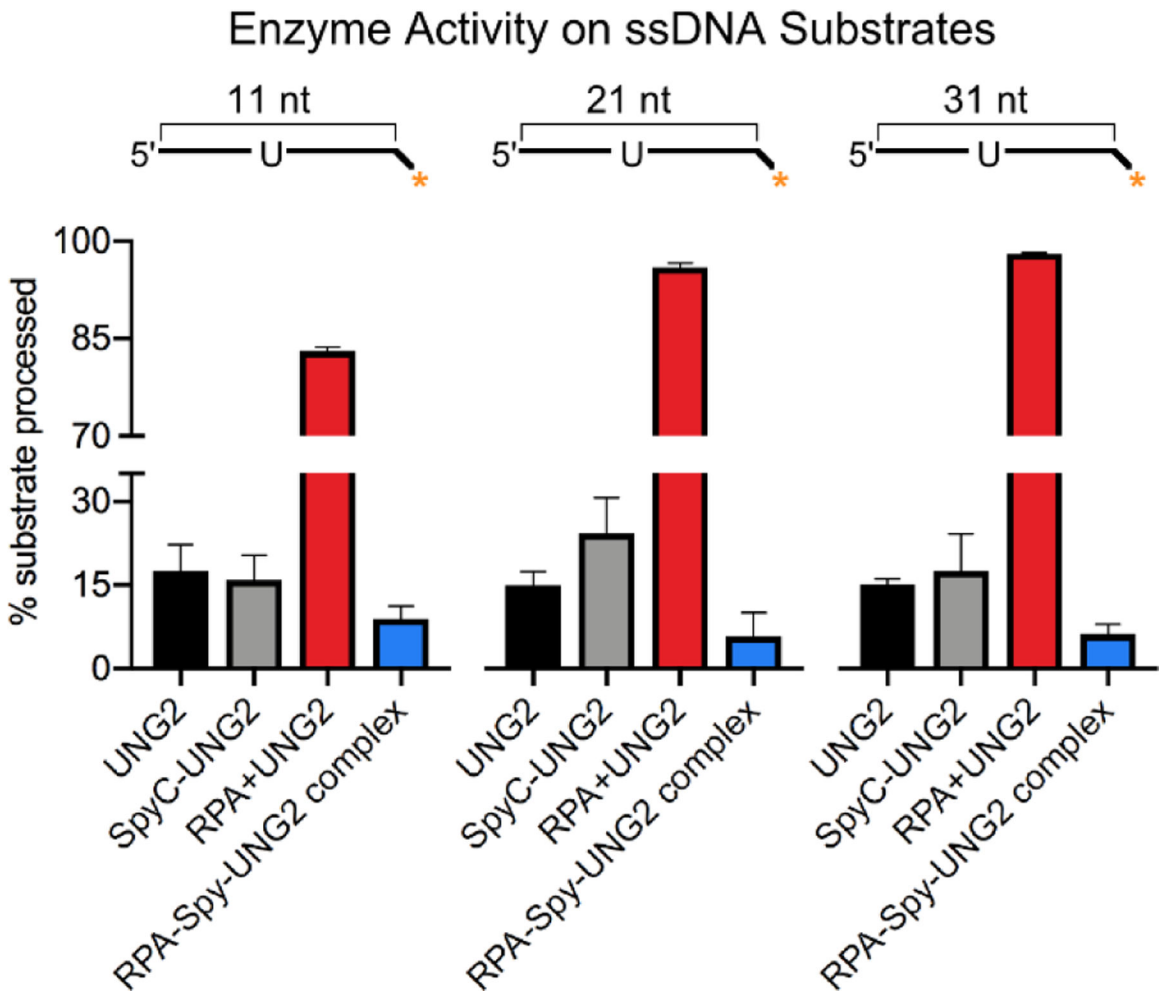


Figure 4.

Uracil excision activity for UNG2 enzymes using endpoint assays and ssDNA substrates of 11 nt, 21 nt, or 31 nt with a uracil base in the center of the ssDNA. All reactions used 0.5 μM ssDNA substrate and 1 nM of UNG2, SpyC-UNG2, or RPA-Spy-UNG2 complex; the reaction time was three minutes. When added, RPA was used at a concentration of 0.5 μM and was pre-equilibrated with the substrate prior to UNG2 addition. The average percent processed from at least three independent assays is shown with standard error.

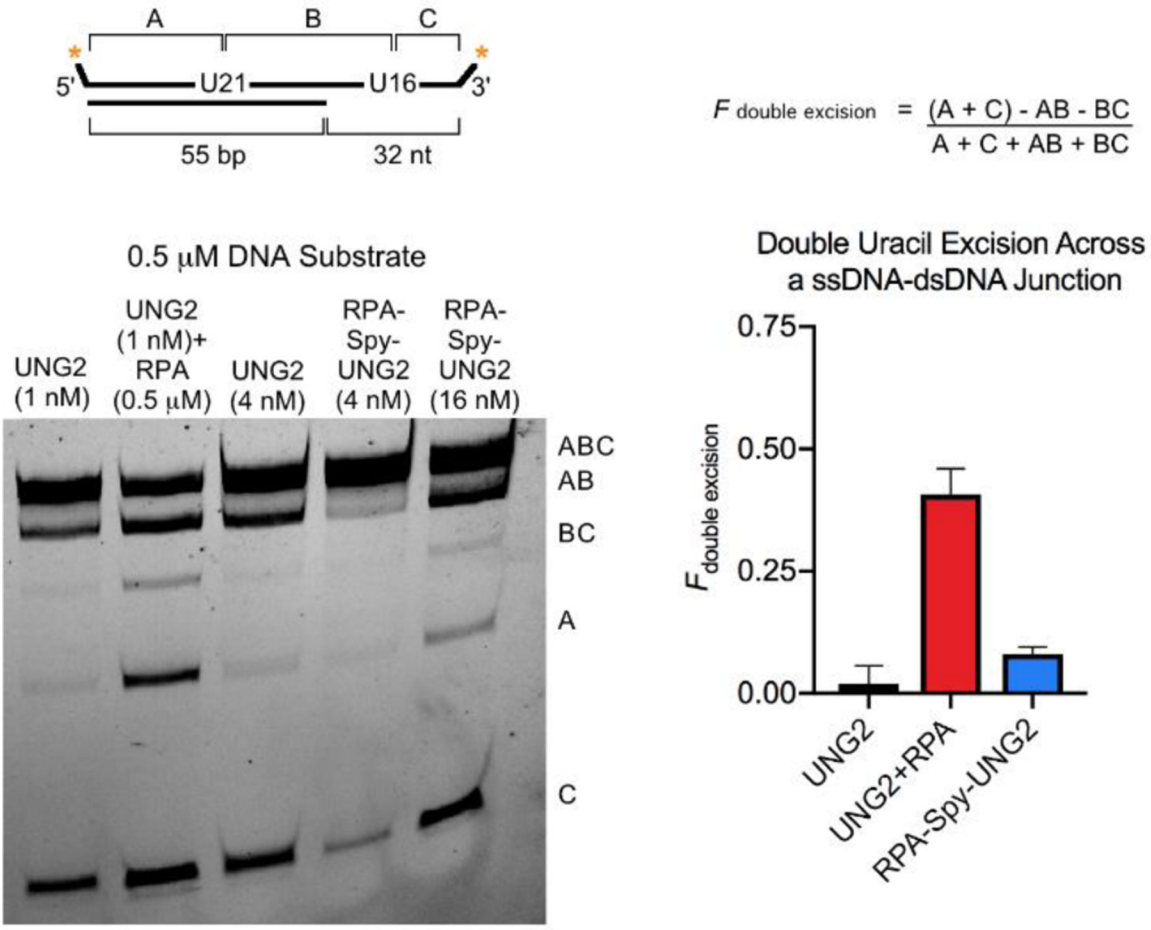


Figure 5.

Uracil excision activity for UNG2 enzymes in endpoint assays using the uracilated substrate shown. The reaction time was three minutes. Excision of uracil in the duplex at 21 bp from the junction resulted in fragments A and BC, while excision of uracil in the ssDNA located 16 nt from the junction resulted in fragments AB and C. Excision of both uracil sites elevated fragments A and C relative to AB and BC. The gel shows that RPA facilitated double excision of the uracil sites located on either side of the junction, whereas UNG2 alone strongly favored single excision in the ssDNA section. The RPA-Spy-UNG2 complex was less efficient at producing double excision compared to RPA and UNG2 under conditions where their activity on ssDNA was similar. For the graph, we calculated the fraction of product that resulted from DNA substrate having both uracils removed ($F_{\text{double excision}}$) using the 4 nM UNG2 and the 16 nM RPA-Spy-UNG2 assays, which produced similar levels of uracil excision at the ssDNA site as the combination of wild-type RPA and 1 nM UNG2. At least three independent replicates were performed for each condition, and the average values with standard error are shown in the graph.

Table 1.

Kinetics of uracil excision for UNG2 ± RPA, SpyC-UNG2, and the RPA-Spy-UNG2 complex measured on different DNA substrates with a stopped-flow device

Protein and Substrate Conditions	Substrate with a Uracil / 2-Aminopurine bp Positioned 21 bp from the DNA Junction or 21 bp from the Duplex End (for dsDNA)												
	32 nt ssDNA-55 bp dsDNA					10 nt ssDNA-55 bp dsDNA					55 bp dsDNA		
	k or k_{slow} (s^{-1})	k_{fast} (s^{-1})	%fast	$t_{20\%}$ (s)[c]	$t_{50\%}$ (s)[d]	k or k_{slow} (s^{-1})	k_{fast} (s^{-1})	%fast	$t_{20\%}$ (s)	$t_{50\%}$ (s)	k (s^{-1})	$t_{20\%}$ (s)	$t_{50\%}$ (s)
UNG2 + DNA	0.15 (0.002) [e]	N/A[f]	N/A	1.495	4.644	0.16 (0.002)	N/A	N/A	1.432	4.448	0.16 (0.002)	1.397	4.339
SpyC-UNG2 + DNA	0.10 (0.001)	N/A	N/A	2.551	7.924	0.14 (0.001)	N/A	N/A	1.545	4.800	0.16 (0.002)	1.368	4.249
RPA/UNG2 + DNA[a]	0.13 (0.002)	2.39 (0.482)	22.3 (2.0)	0.483	3.387	0.19 (0.001)	1.85 (0.876)	20.7 (4.2)	0.510	2.429	0.08 (0.001)	3.111	9.634
UNG2 + RPA/ DNA[b]	0.12 (0.003)	2.27 (0.641)	25.6 (3.1)	0.439	3.277	0.13 (0.002)	N/A	N/A	1.716	5.330	0.08 (0.001)	2.638	8.193
RPA-Spy-UNG2 complex + DNA	0.03 (0.001)	5.55 (1.169)	23.1 (2.5)	0.328	16.738	0.12 (0.002)	2.18 (0.214)	36.7 (1.0)	0.302	2.002	0.03 (0.001)	7.189	22.329

[a] RPA and UNG2 were equilibrated prior to DNA introduction

[b] RPA and DNA were equilibrated prior to UNG2 introduction

[c] Time required for 20% of substrate to be processed as determined from the curve fit to the data

[d] Time required for 50% of substrate to be processed as determined from the curve fit to the data

[e] Numbers in parenthesis represent standard error

[f] N/A, not applicable