



Genome-wide maps of CPD deamination in yeast reveal the impact of DNA sequence context and nucleosome architecture on cytosine deamination rates

Marian F. Laughery, Bastian Stark, Benjamin Morledge-Hampton, et al.

Genome Res. published online December 9, 2025

Access the most recent version at doi:[10.1101/gr.280384.124](https://doi.org/10.1101/gr.280384.124)

P<P Published online December 9, 2025 in advance of the print journal.

Open Access Freely available online through the *Genome Research* Open Access option.

Creative Commons License This article, published in *Genome Research*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Comprehensive immune receptor profiling.
Discover the **DriverMap™ AIR Assay** difference.

LEARN
MORE



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

© 2025 Laughery et al.; Published by Cold Spring Harbor Laboratory Press

Method

Genome-wide maps of CPD deamination in yeast reveal the impact of DNA sequence context and nucleosome architecture on cytosine deamination rates

Marian F. Laughery,¹ Bastian Stark,¹ Benjamin Morledge-Hampton,¹
Steven A. Roberts,^{1,2} and John J. Wyrick¹

¹School of Molecular Biosciences, Washington State University, Pullman, Washington 99164, USA; ²Department of Microbiology and Molecular Genetics, University of Vermont Cancer Center, University of Vermont, Burlington, Vermont 05405, USA

UV light induces cyclobutane pyrimidine dimers (CPDs) and other mutagenic lesions in cellular DNA. Cytosine-containing CPDs can subsequently undergo rapid deamination to uracil, a process that has been linked to UV mutagenesis. However, the impact of genomic context and chromatin architecture on CPD deamination rates in cells remains poorly understood. Here, we develop a method known as dCPD-seq to map deaminated CPDs (dCPDs) across the genome of repair-deficient yeast cells at single-nucleotide resolution. Our dCPD-seq data reveal that sequence context significantly modulates CPD deamination rates in UV-irradiated yeast cells, with CPDs in TCG contexts showing particularly rapid deamination rates. Our analysis indicates that rapid CPD deamination can explain why UV-induced mutations are specifically enriched at TCG sequences, both in UV-irradiated yeast cells and in human skin cancers. CPD deamination is suppressed near the transcription start and end sites of yeast genes, which may in part be mediated by DNA-bound transcription factors. Finally, we show that the wrapping of DNA in nucleosomes modulates CPD deamination in yeast cells. Our data indicate that CPD deamination is elevated at minor-in rotational positions where the DNA minor groove faces the histone octamer, likely owing to increased solvent accessibility of the C4 position of the cytosine base. Moreover, we also observe strand-specific enrichment of CPD deamination at rotational positions where the DNA backbone faces out toward the solvent. Taken together, these findings reveal how DNA sequence context and chromatin architecture modulates CPD deamination rates across a eukaryotic genome.

[Supplemental material is available for this article.]

Ultraviolet (UV) light is the primary causative agent of skin cancers because it induces mutagenic damage to DNA (Pfeifer et al. 2005; Brash 2015). The primary class of UV damage is the cyclobutane pyrimidine dimer (CPD), which forms between adjacent pyrimidine bases (i.e., TT, TC, CT, and CC) (Friedberg et al. 2006). Although CPDs form most frequently at TT dinucleotides (Mao et al. 2016), somatic substitutions at TT dinucleotides are relatively infrequent in sequenced skin cancer genomes (Brash 2015; Bonilla et al. 2016; Hayward et al. 2017). Instead, cytosine-to-thymine (C>T) substitutions in dipyrimidine sequences comprise >80% of the mutations observed in skin cancers, including malignant melanoma (Bonilla et al. 2016; Hayward et al. 2017). These observations suggest that cytosine-containing CPDs are especially mutagenic; however, the exact mechanism involved remains unclear.

Although all CPDs are initially formed through the same UV-induced [2+2] cycloaddition reaction, cytosine-containing CPDs are chemically unstable and undergo subsequent spontaneous deamination of the damaged cytosine base to uracil (Setlow et al. 1965; Taylor 2022). Deamination of cytosine bases also occurs in undamaged DNA, but this reaction is very slow, with a half-life of ~30,000 years in double-stranded DNA and ~200 years in sin-

gle-stranded DNA (Frederico et al. 1990). In contrast, the rate of deamination is vastly accelerated for cytosine bases in CPDs, with a half-life of tens to hundreds of hours (Taylor 2022). Because even error-free replicative bypass of a deaminated CPD (dCPD) would result in a C>T substitution, this mechanism could potentially account for the preponderance of C>T substitutions in human skin cancers (Pfeifer et al. 2005; Ikehata and Ono 2011; Jin et al. 2021). The rate-limiting step in the CPD deamination reaction is the hydrolytic attack of a hydroxide or water molecule on the C4 position of the cytosine base (Lemaire and Ruzsicska 1993; Uddin et al. 2014). Extensive biochemical analyses of 5-methylcytosine (5mC)-containing CPDs with variable flanking bases in vitro has revealed that deamination occurs fastest in cytosines with neighboring guanine bases (Cannistraro and Taylor 2009; Taylor 2022). However, to what extent flanking bases affect CPD deamination and subsequent mutagenesis across the genome of UV-irradiated cells is unclear.

Previous biochemical studies have also suggested that CPD deamination can be modulated by DNA-bound proteins, particularly nucleosomes (Cannistraro and Taylor 2010; Song et al. 2011, 2014). Nucleosomes are the fundamental packaging unit of eukaryotic chromatin and are composed of an octamer of

Corresponding author: jwyrick@wsu.edu

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.280384.124>. Freely available online through the *Genome Research* Open Access option.

© 2025 Laughery et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

histone proteins and ~147 bp of DNA (Luger et al. 1997). As the nucleosomal DNA wraps around the histone octamer, the minor groove of the DNA backbone is oriented either in toward the histone octamer (minor-in) or away from the histone octamer (minor-out) at 5 bp intervals. We and others have previously shown that both UV-induced CPD formation and somatic mutations in skin cancers are elevated at minor-out and suppressed at minor-in rotational settings (Gale et al. 1987; Mao et al. 2016; Brown et al. 2018; Pich et al. 2018). To what extent the orientation of the DNA minor groove effects CPD deamination is unknown. In between the minor-in and minor-out positions, one DNA strand is oriented with its backbone facing out away from the histone octamer (backbone-out), whereas the complementary backbone faces inward (backbone-in). Previous studies analyzing deamination of a 5mC base in a TCG sequence context found that CPD deamination is significantly elevated at backbone-out orientations in nucleosomes and suppressed at backbone-in orientations in vitro (Song et al. 2011, 2014). However, to what extent nucleosomes modulate CPD deamination in other sequence contexts and across the genome of intact cells remains unclear.

Genome-wide sequencing methods have emerged as powerful tools to understand how different genomic and chromatin contexts impact DNA damage formation and repair (Bohm and Wyrick 2023; Pfeifer and Jin 2024). Recently, Pfeifer and colleagues developed a sequencing method known as circle-damage-seq to measure deamination of CPDs in UV-radiated human cells (Jin et al. 2021). They found that CPD deamination is suppressed near the transcription start site (TSS) and transcription end site (TES) of human genes and that the sequence pattern of CPD deamination shows strong similarity to the COSMIC signature 7 mutation pattern that occurs in melanoma and other skin cancers (Jin et al. 2021). Although these results revealed a promising correlation between deamination and skin cancer mutations, there were some limitations. Specifically, the human cells used in this study were repair-competent, and thus, ongoing CPD repair could complicate analysis of deamination rates, and detection of lesions via cleavage by uracil DNA glycosylase (UDG) precludes the detection of deaminated cytosines in methylated CpG contexts (i.e., TCG and CCG). Moreover, the impact of chromatin organization on deamination rates was largely unexplored, possibly owing to high levels of background deamination during library preparation.

Here, we develop a new method called dCPD sequencing (dCPD-seq) and use it to map dCPDs across the genome of repair-deficient yeast cells in order to characterize the impact of DNA sequence, genomic context, and chromatin organization on CPD deamination.

Results

To develop a method to map the deamination of cytosine-containing CPDs, we initially analyzed CPD deamination in yeast genomic DNA that was UV-irradiated in vitro. To quantify the kinetics of CPD deamination in naked DNA in vitro, isolated yeast genomic DNA was UV-irradiated (~600 J/m² of UVC based on our previous calibration) and subsequently incubated for various times at 37°C to promote deamination (Fig. 1A). This relatively high UV dose was used to increase the number of cytosine-containing CPDs, which comprise only a subset of all CPDs (TT CPDs being the most frequent) (Friedberg et al. 2006). To detect dCPDs, CPDs were repaired using CPD photolyase (CPD PL), and residual uracil lesions resulting from CPD deamination were cleaved with UDG. The resulting cleavage products were analyzed by alkaline

gel electrophoresis (Fig. 1B). Although little, if any, DNA cleavage was detectable in the 0 h or no UV samples, we observed increasing levels of DNA cleavage in samples incubated for 6 h, 24 h, and 48 h at 37°C (Fig. 1B). Little if any DNA cleavage by UDG was detected in the absence of CPD PL at these time points (Fig. 1B), indicating that UDG was almost exclusively cleaving uracil lesions present in CPDs. Quantification of these data indicated that uracil formation arising from dCPDs increased at later time points, reaching a frequency of 0.404 ± 0.081 uracil lesions per kilobase of genomic DNA after 48 h incubation at 37°C (Fig. 1C). Parallel experiments using T4 endonuclease V digestion followed by alkaline gel analysis to measure overall CPD induction indicated that this UV dose resulted in 1.76 ± 0.16 CPDs per kilobase in yeast genomic DNA, indicating that ~23% of all CPDs are deaminated in vitro by 48 h.

To map dCPDs at single-nucleotide resolution across the yeast genome, we developed the dCPD-seq method (Fig. 1D). Yeast genomic DNA (naked DNA) containing dCPDs were blocked on the 3' end by incubating the DNA with terminal transferase and dideoxyATP or dideoxyGTP (Fig. 1D). The time allotted to this step was relatively short (i.e., 1 h at 37°C) to minimize in vitro deamination of CPDs during library preparation. The resulting DNA fragments were then coincubated with CPD PL and UDG and photo-reactivated with UVA light for 2 h to both reverse the CPD lesions and cleave uracils associated with them. Apurinic/aprimidinic endonuclease 1 (APE1) was then used to generate a free 3' hydroxyl (3'OH) immediately upstream of the uracil lesion. Single-stranded DNA with a free 3'OH generated from the UDG and APE1 digestions was ligated to a biotinylated first adapter (Fig. 1D, green), and ligation products were sonicated and purified using streptavidin beads. The complementary strand was synthesized by polymerase extension using a primer complementary to the first adapter, and the other DNA end was ligated to the second adapter (Fig. 1D, blue). Following PCR amplification using primers complementary to both adapters, the resulting libraries were submitted for high-throughput sequencing. The resulting sequencing reads were mapped to the yeast genome, and the location of the uracil lesion was inferred as the DNA base on the opposite strand immediately upstream of the 5' end of the sequencing read originating from the first adapter.

We sequenced dCPD-seq libraries for two independent deamination time courses, consisting of yeast genomic DNA incubated for 0 h, 6 h, 24 h, and 48 h at 37°C after UVC irradiation. Mapping of the resulting dCPD-seq reads to the yeast genome indicated a significant enrichment of lesions associated with cytosine bases. At later time points (i.e., 6 h, 24 h, and 48 h), >80% of the reads were associated with a cytosine lesion (Fig. 1E). We also observed a weak enrichment of dCPD-seq reads at cytosine bases at the 0 h time point, although their enrichment was not significantly different than that of adenine-associated reads. Moreover, there was relatively little signal associated with 0 h libraries compared with the later time points. Analysis of DNA bases immediately flanking these cytosine lesions indicated that nearly all of these were associated with CPD-forming dipyrimidine sequences (Fig. 1F; Supplemental Fig. S1). Taken together, these findings suggest that the vast majority of dCPD-seq reads are associated with deaminated cytosine bases derived from UV-induced CPD lesions.

Genome-wide map of CPD deamination in repair-deficient yeast cells

Although our aim was to analyze CPD deamination over a 48 h time course in nucleotide excision repair (NER)-deficient (i.e.,

Genomic maps of CPD deamination in yeast chromatin

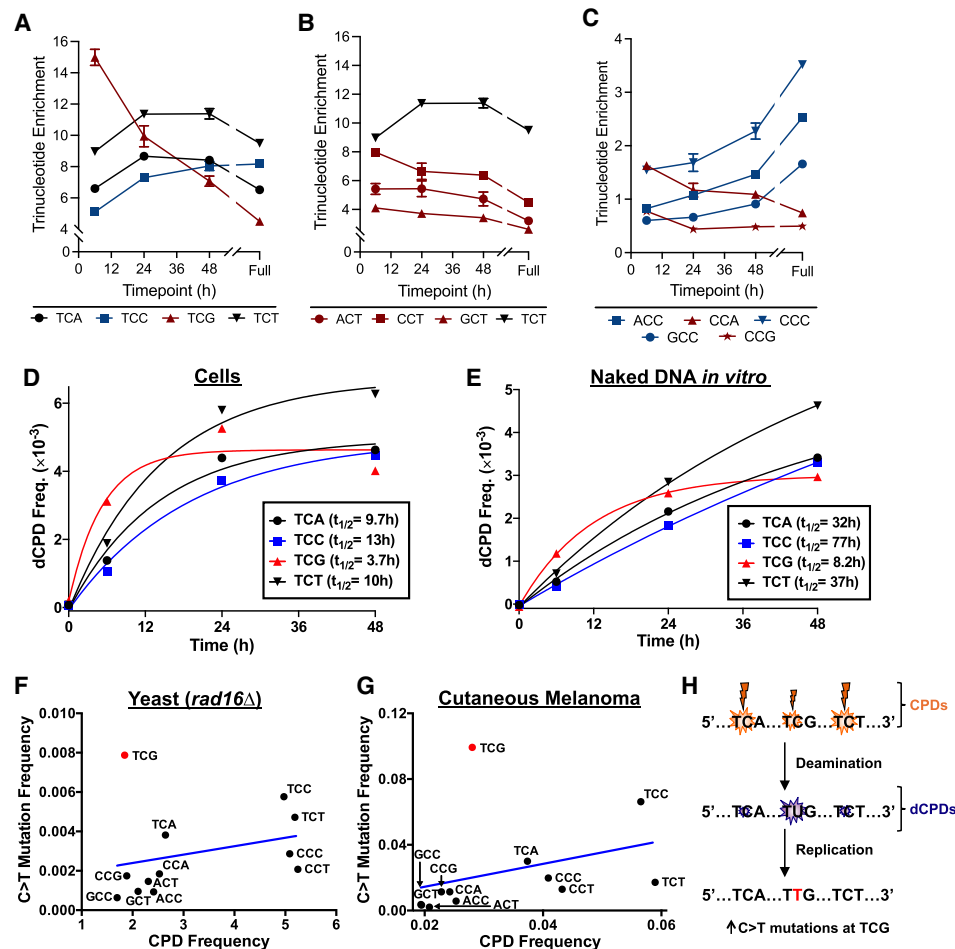


Figure 3. DNA sequence context significantly modulates cellular CPD deamination, likely impacting UV mutagenesis. (A–C) The fraction of total dCPD-seq reads mapping to each cytosine-central trinucleotide context was normalized by the fraction of the individual trinucleotide sequences in the yeast genome (resulting in “trinucleotide enrichment”) and plotted as a function of 37°C incubation time. Red lines/symbols indicate fast deaminating sequence contexts; blue indicates relatively slow deaminating sequence contexts; and black is intermediate. Results are grouped by TCN (A), NCC/CCN (B), and NCT contexts (C). The 6 h, 24 h, and 48 h results represent the average and SEM from three independent experiments. The 0 h results are from two experiments, and the “full” deamination time point is derived from a single experiment. (D) Quantitation of dCPD lesions in cellular deamination samples across the yeast genome at the indicated trinucleotide sequence contexts. dCPD frequency was calculated from frequency of dCPD-seq reads at each trinucleotide context and scaled using the alkaline gel data for the cellular deamination time course. The calculated half-life ($t_{1/2}$) of a single exponential fit to each trinucleotide deamination time course is indicated. (E) Same as panel D, except for UV-irradiated naked yeast genomic DNA that was deaminated *in vitro* (see Fig. 1). (F) The frequency of C>T single-base substitutions in different trinucleotide contexts (e.g., TCG, etc.) is plotted relative to the frequency of UV-induced CPD lesions in each context. Mutation data are derived from whole-genome sequencing of UV-irradiated *rad16Δ* yeast cells (Laughery et al. 2020), whereas CPD frequency is derived from CPD-seq data from UV-irradiated *rad16Δ* yeast cells (0 h time point) (Mao et al. 2020). Blue line indicates the linear regression fit to the data. The mutation frequency at TCG (0.00788; shown in red) is a significant outlier relative to other contexts and lies outside the 95% confidence interval of the predicted value based on its CPD frequency (0.00235 ± 0.00539). (G) Same as panel F, except mutation frequency is derived from published whole-genome sequencing data from 140 cutaneous melanomas (Hayward et al. 2017), and CPD frequency is derived from published CPD-seq data from UV-irradiated melanoma cells (Lindberg et al. 2019). Again, the mutation frequency at the TCG (0.0993; shown in red) is a significant outlier and lies outside the 95% confidence interval of the value predicted from its CPD frequency (0.0201 ± 0.0683). (H) Model depicting potential mechanism for elevated C>T mutations at TCG contexts in yeast and human skin cancers. Although the frequency of CPD lesions is somewhat lower at TCG than other contexts (e.g., TCA or TCT), its deamination occurs much more rapidly, leading to mutation enrichment.

This analysis indicated that trinucleotide sequence contexts with higher levels of initial CPD formation tended to be associated with higher levels of UV-induced C>T substitutions (Fig. 3F). However, the frequency of UV-induced mutations in the TCG sequence context was a significant outlier compared with other trinucleotide sequence contexts, being 3.4-fold more frequent in TCG sequences than expected from their associated frequency of CPDs based on a linear regression model (Fig. 3F). If the TCG sequence context was excluded, then there was a significant positive correlation between CPD levels and UV-induced C>T muta-

tions ($r=0.70$, $P<0.05$), but this correlation was not significant if the TCG context was included ($P>0.05$).

To further test this hypothesis, we analyzed a compendium of somatic mutations derived from whole-genome sequencing of 140 cutaneous melanomas (Hayward et al. 2017). Unlike in budding yeast, CG dinucleotides in the human genome are frequently methylated, which is known to affect CPD formation when exposed to UVB wavelengths present in sunlight (e.g., Pfeifer et al. 2005; Lindberg et al. 2019; Wilson and Wyrick 2024). To account for this potential confounding factor, we used a published CPD-

seq data set derived from melanoma cells that had been irradiated with UVB light (Lindberg et al. 2019), and compared the frequency of CPDs in this data set to the frequency of somatic mutations across different trinucleotide contexts.

This analysis indicated that the TCG sequence context was again a significant outlier to other trinucleotide sequence contexts, with 4.9-fold more mutations than expected based on its CPD levels using the linear regression model (Fig. 3G). There was a significant correlation between CPD levels and somatic mutation frequency across different trinucleotide contexts if the TCG sequence context was excluded ($r=0.72$, $P<0.05$), but this correlation was insignificant if the TCG context was included ($P>0.05$). These findings indicate that UV-induced mutations in both *rad16Δ* mutant yeast and human melanomas are elevated at fast-deaminating TCG sequences, occurring about three- to fivefold more frequently than expected from the levels of damage formation (Fig. 3H).

CPD deamination is suppressed near the transcription start and end sites of yeast genes

To determine how CPD deamination is modulated within genes in yeast cells, we analyzed the frequency of dCPDs from our cellular dCPD-seq data between the TSS and TES of about 5000 yeast genes, based on published gene coordinates (Park et al. 2014). Each gene was divided into six equally sized bins, and the number of dCPDs in each bin was divided by the number of cytosines in a CPD-forming dipyrimidine sequence context in each bin to yield the normalized dCPD frequency. Additionally, we also analyzed the normalized dCPD frequency in six flanking bins, each 167 bp in length, three upstream of the TSS and three downstream from the TES.

We initially used this strategy to analyze the 48 h cellular deamination samples. This analysis revealed high levels of dCPDs within genes but lower levels of dCPDs immediately upstream of the TSS and flanking the TES (Fig. 4A). To test whether this was merely a consequence of differences in DNA sequence composition in these genomic regions, we performed similar analysis of the dCPD-seq data derived from 48 h in vitro deamination of naked DNA that had been isolated from UV-irradiated yeast cells (i.e., naked DNA deamination control) (see Fig. 2A). These DNA samples showed nearly constant levels of normalized dCPDs across yeast genes and flanking DNA (Fig. 4A). To quantify the enrichment of CPD deamination in cells, we calculated the ratio of dCPD reads in the cellular samples relative to the naked DNA control at each deamination time point. This ratio was standardized by the overall number of dCPD-seq reads associated with cytosines in dipyrimidine sequences in each sample to yield a dCPD enrichment value (Fig. 4B). This analysis indicated that dCPD enrichment after 48 h of deamination was low (i.e., <1.0) immediately upstream of the TSS and flanking the TES of yeast genes, confirming that CPD deamination is specifically suppressed in regions neighboring the TSS and TES of genes in yeast cells. Similar analysis of dCPD enrichment at the 6 h and 24 h samples again showed suppression of CPD deamination in cells near the TSS and TES of genes (Fig. 4C,D).

To validate these findings, we analyzed the dCPD-seq data for the cellular and naked DNA 48 h samples that had subsequently undergone full deamination in vitro (i.e., 16 h at 67°C). This analysis revealed that the dCPD enrichment for the full deamination samples showed nearly constant deamination across yeast genes and no suppression neighboring the TSS or TES (Fig. 4E). These re-

sults confirm that there are similar levels of cytosine-containing CPDs across yeast genes in both the 48 h cellular and in vitro deaminated samples, because complete deamination results in similar levels of dCPDs (i.e., enrichment value of about 1.0).

One possible explanation of this finding is that protein–DNA interactions associated with the TSS and TES of yeast genes may suppress CPD deamination in cells, as suggested by previous studies (Cannistraro and Taylor 2010; Jin et al. 2021; Duan et al. 2024). To test this hypothesis, we used our dCPD-seq data to analyze patterns of CPD deamination at yeast transcription factor binding sites (TFBSs), because these are among the most frequent protein–DNA interactions in regulatory regions like promoters. We used a published yeast ChIP-exo data set that contained 1893 TFBSs derived from 78 different transcription factors (Rossi et al. 2021). Analysis of dCPD-seq data from the 6 h cellular deamination sample indicated that the normalized dCPD levels were decreased at yeast TFBSs (Fig. 4F). To determine if this was because of a decrease in deamination rates at TFBSs in yeast cells, we analyzed the cellular deamination samples relative to the in vitro deamination control. This analysis revealed lower levels of dCPD enrichment near the center of TFBSs in aggregate at the 6 h and 24 h time points (Fig. 4G,H). In contrast, there were no significant difference in dCPD enrichment associated with TFBSs in the fully deaminated control (Fig. 4I). These findings suggest that inhibition of CPD deamination associated with transcription factor binding may contribute to lower levels of dCPDs near the TSS (and potentially TES) of yeast genes.

CPD deamination is elevated at minor-in rotational settings in yeast nucleosomes

Previous studies have suggested that CPD deamination can be modulated in nucleosomes (Song et al. 2011, 2014), but to what extent this occurs across the genome was previously unclear. To address this question, we used our dCPD-seq data to analyze patterns of CPD deamination in about 10,000 strongly positioned nucleosomes identified in a previously published map of nucleosome positioning in yeast (Brogaard et al. 2012). Analysis of normalized dCPDs in the 6 h cellular deamination sample revealed a roughly periodic pattern of CPD deamination, particularly at distal positions (i.e., -73 to -20 and $+20$ to $+73$) in the nucleosome DNA (Supplemental Fig. S9, top). In general, normalized dCPD levels were lower at minor-out rotational settings (Supplemental Fig. S9, dashed lines) and higher at minor-in positions (in between dashed lines). A similar pattern of dCPDs was observed at the 24 h and 48 h cellular deamination samples, although the periodicity was somewhat reduced at the 48 h time point and largely absent from the full deamination control (Supplemental Fig. S9).

To control for the effects of DNA sequence composition in nucleosomal DNA on deamination rates, we also analyzed dCPD enrichment following deamination in yeast cells relative to the in vitro deamination (i.e., naked DNA) control. Importantly, both of these samples were UV-irradiated in cells prior to deamination because nucleosomes affect initial CPD formation (Gale et al. 1987; Mao et al. 2016; Brown et al. 2018). This analysis confirmed that dCPD enrichment in yeast cells relative to the naked DNA control displays a periodic pattern in strongly positioned nucleosomes across the deamination time course (Fig. 5A). Following 6 h deamination, dCPD enrichment was decreased at minor-out rotational settings (Fig. 5A, dotted lines) and elevated at minor-in rotational settings. A similar pattern was observed at the 24 h and 48 h deamination time points, although the magnitude of

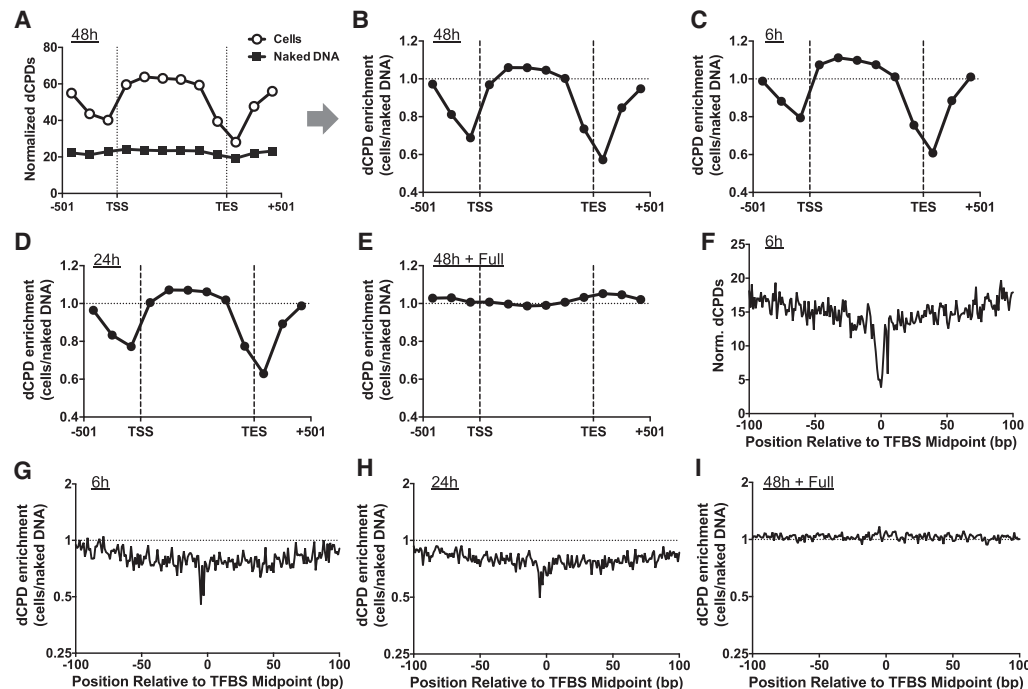


Figure 4. CPD deamination is suppressed near the TSS and TES of yeast genes and adjacent to transcription factor binding sites. (A) Normalized frequency of dCPDs following 48 h of deamination after UV irradiation of yeast cells. Lines with open circles and gray squares indicated data for cellular and in vitro (control) experiments, respectively. Data are plotted for about 5000 yeast genes in aggregate. Each gene was divided into six equally sized bins, and the number of dCPD lesions, derived from the count of dCPD-seq reads associated with cytosines in dipyrimidine sequences, was tallied for each bin. dCPDs associated with six 167 bp flanking bins, three upstream of the TSS and three downstream of the TES for each gene, were also tallied. The frequency of dCPD-seq reads in each bin was normalized using the count of cytosine-containing dipyrimidines in each bin. TSS and TES coordinates are from Park et al. (2014). (B) Same as panel A, except the dCPD enrichment, which is defined as the number of dCPDs in the cellular deamination experiments relative to the in vitro deamination control, is depicted. dCPD enrichment was normalized by the total number of dCPDs in each sample. (C–E) Same as panel B, except for the 6 h (C), 24 h (D), and full (E) deamination samples. (F) Normalized dCPD frequency following 6 h deamination in UV-irradiated yeast cells at 1893 binding sites for 78 different yeast transcription factors. The dCPD-seq data at each position were normalized by the frequency of cytosine bases in a dipyrimidine sequence context. TFBS data are from Rossi et al. (2021). (G) Same as panel F, except plotting dCPD enrichment at yeast TFBSs, which is calculated from the normalized ratio of cellular dCPD-seq reads at each position relative to the in vitro deamination control. (H,I) Same as panel G, except dCPD enrichment for the 24 h (H) and full (I) deamination sample is plotted.

the differences in dCPD enrichment between the minor-in and minor-out rotational positions was reduced at later time points, including in the full deamination control (Fig. 5A). Statistical analysis revealed significantly higher levels of dCPD enrichment at the minor-in rotational settings relative to the minor-out settings across the deamination time course (Fig. 5B–D).

We also compared dCPD enrichment in strongly positioned nucleosomes with initial CPD formation using our published CPD-seq data (Mao et al. 2016). The results indicated that although CPD and dCPD enrichment both showed clear periodicity in nucleosomes, they had the opposite phasing (Fig. 5E,F). CPD enrichment in UV-irradiated yeast cells relative to the naked DNA control is elevated at minor-out rotational settings (Fig. 5E, dashed lines), whereas dCPD enrichment is suppressed at these same positions. In contrast, at minor-in positions, dCPD enrichment is elevated, whereas CPD enrichment was reduced (Fig. 5E).

We wondered what molecular mechanism was responsible for elevated CPD deamination at minor-in rotational settings in nucleosomes. Previous biophysical and computational studies (Lemaire and Ruzsicska 1993; Uddin et al. 2014) have indicated that the rate-limiting step in the CPD deamination reaction is the hydrolytic attack of a water (or hydroxide) molecule on the C4 position of the cytosine base (Fig. 6A). We reasoned that differences in the solvent accessible surface area (SASA) of these C4

atoms could potentially account for the modulation in CPD deamination. To test this hypothesis, we calculated the SASA of C4 positions ($SASA_{C4}$) to water molecules (radius ~ 1.4 Å) for all cytosine bases in a published structure of the yeast (*Saccharomyces cerevisiae*) nucleosome (White et al. 2001). Although C4 atoms in most cytosine bases had zero SASA, a subset had a non-zero $SASA_{C4}$ (Supplemental Fig. S10A), indicating that there are significant differences in C4 solvent accessibility depending upon the location of the cytosine base within the nucleosome. We compared the calculated $SASA_{C4}$ between cytosines in minor-out and minor-in rotational settings in the yeast nucleosome structure and found that minor-in rotational settings had significantly higher $SASA_{C4}$ ($P < 0.01$) (Fig. 6B). Multiple cytosines at minor-in rotational settings had nonzero $SASA_{C4}$, whereas none of the cytosines at minor-out positions had nonzero $SASA_{C4}$ (Fig. 6B). The $SASA_{C4}$ of minor-in positions was also significantly elevated relative to minor-out positions in other high-resolution nucleosome structures (Supplemental Fig. S10B,C).

It has also been suggested that in some conditions hydroxide ions may perform the hydrolytic attack on the C4 position of cytosine (Taylor 2022). Hence, we also calculated the $SASA_{C4}$ in nucleosome structures to hydroxide (radius ~ 1.1 Å) (Supplemental Fig. S10D). Again, we observed significantly higher $SASA_{C4}$ (for hydroxide) at minor-in positions relative to minor-out in the yeast

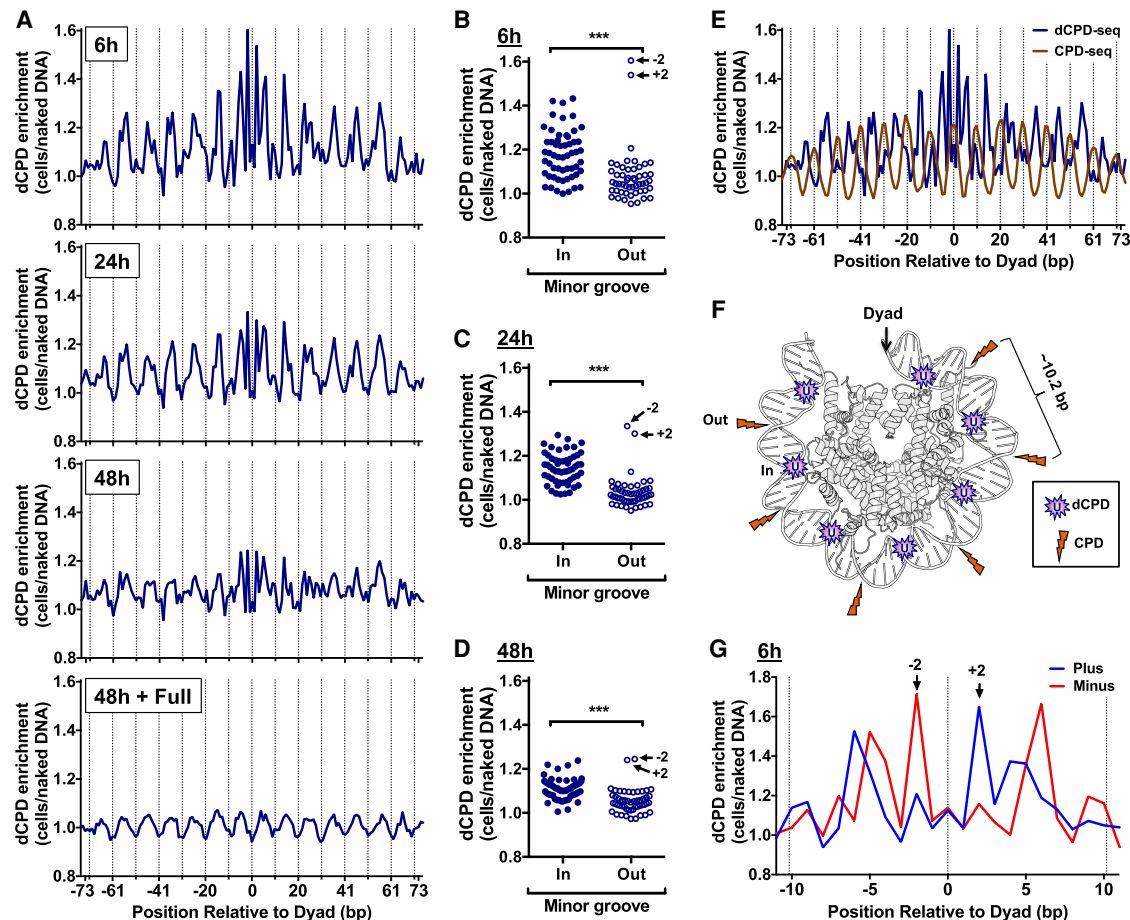


Figure 5. CPD deamination is elevated at minor-in rotational settings and suppressed at minor-out rotational settings in nucleosomes. (A) dCPD enrichment in UV-irradiated cells relative to the in vitro deamination control is plotted at each position relative to the central dyad of about 10,000 strongly positioned nucleosomes for the indicated deamination time points. Dashed lines indicate minor-out rotational settings. Nucleosome positioning data are from Brogaard et al. (2012). (B–D) Plot of dCPD enrichment at minor-in and minor-out rotational settings for 6 h (B), 24 h (C), and 48 h (D) deamination samples. (***) $P < 0.0001$ based on a Mann–Whitney U test. (E) Comparison of dCPD (6 h time point) and CPD (0 h, immediately following damage) enrichment in strongly positioned yeast nucleosomes. Dashed lines indicate minor-out rotational settings. CPD enrichment was calculated relative to a UV-irradiated naked DNA control and is from Mao et al. (2016). (F) Model showing structure of the yeast nucleosome (PDB ID: 1ID3) (White et al. 2001) and highlighting locations of elevated CPD enrichment (orange lightning bolts) and dCPD enrichment (purple U [uracil]). CPD enrichment is elevated at minor-out positions, but CPD deamination is elevated at minor-in positions in nucleosomes. Only one DNA gyre is shown. Image generated using PyMOL. (G) Close-up of strand-specific dCPD enrichment in UV-irradiated cells following 6 h deamination relative to the in vitro deamination control. dCPD enrichment is depicted for nucleosome positions within 11 bp of the central dyad of about 10,000 strongly positioned nucleosomes.

nucleosome structure and other high-resolution nucleosome structures (Fig. 6C; Supplemental Fig. S10E,F). At minor-in rotational settings, the minor groove of the DNA faces toward the histone octamer, whereas the major groove faces the solvent. The C4 position of a cytosine base is located on the major groove side of a G–C base pair (Fig. 6D), which can potentially explain why C4 solvent accessibility is elevated at minor-in rotational settings.

CPD deamination is elevated at backbone-out positions in nucleosomes

Closer inspection of dCPD enrichment data indicated that although most minor-out positions in strongly positioned nucleosomes showed relatively low levels of dCPD enrichment, two positions immediately flanking the central dyad axis (i.e., -2 and $+2$) displayed the highest levels of dCPD enrichment (Fig. 5A–D). Analysis of strand-specific dCPD enrichment neighboring the central nucleosome dyad indicated that the peak at the -2 po-

sition specifically occurred on the minus strand, whereas the peak at the $+2$ position specifically occurred on the plus strand (Fig. 5G). Although our previous analysis indicated that CPD deamination is elevated at minor-in positions using the aggregate of both DNA strands, these data suggest that dCPD enrichment is also modulated in a strand-specific manner.

To investigate strand-specific patterns of CPD deamination in nucleosomes, we analyzed the dCPD-seq enrichment data for each DNA strand (i.e., plus or minus) separately. This analysis revealed that dCPD enrichment in cells relative to the naked DNA control was elevated near minor-in rotational settings (Fig. 7A, solid lines) in strongly positioned nucleosomes in both DNA strands, consistent with our previous analysis. However, dCPD enrichment for the plus strand was generally elevated immediately to the left of the minor-in position, whereas minus strand enrichment was elevated to the right of the minor-in rotational setting (Fig. 7A). Because the plus and minus DNA strands have an antiparallel orientation, these findings suggest that CPD deamination for each

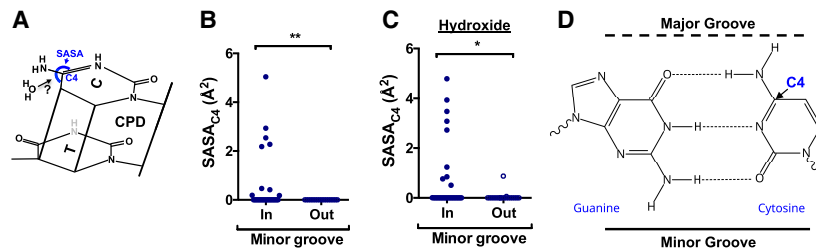


Figure 6. Molecular mechanism underlying elevated CPD deamination at minor-in rotational settings in nucleosomes. (A) Chemical structure of a thymine–cytosine CPD. Previous studies indicate that the rate-limiting step of cytosine deamination is hydrolytic attack by a water (H₂O) molecule on the cytosine C4 position. The accessibility of these C4 position to hydrolytic attack can be quantified using the solvent accessible surface area (SASA). (B) Quantification of the SASA_{C4} for cytosine bases in the yeast nucleosome structure (1ID3) for bases at minor-in or minor-out rotational settings. (**) $P < 0.01$ based on a Mann–Whitney U test. (C) Same as panel B, except the SASA_{C4} is calculated for hydroxide (radius ~ 1.1 Å) instead of water. (*) $P < 0.05$ based on a Mann–Whitney U test. (D) Chemical structure of the guanine–cytosine base pair, highlighting the fact that the cytosine C4 position is located near the major groove.

DNA strand is elevated on the 5′ side of the minor-in rotational setting and suppressed on the 3′ side. These strand-specific differences in dCPD enrichment were most pronounced in the 6 h sample, were diminished at 24 h and 48 h, and were essentially absent from the full deaminated control (48 h+Full) (see Fig. 7A).

To quantify these strand-specific differences in CPD deamination, we plotted the average dCPD enrichment values at positions adjacent to each minor-out rotational setting in the nucleosome (Fig. 7B). This analysis confirmed that dCPD enrichment is elevated at minor-in rotational settings (solid lines) and suppressed at minor-out rotational settings (Fig. 7B, dotted lines). It also confirmed strand-specific differences in CPD deamination, as dCPD enrichment for each DNA strand was elevated on the 5′ sides of minor-in rotational settings and reduced on the 3′ side (Fig. 7B).

Closer inspection of the nucleosome structure (Fig. 7C) indicates that between each minor-in and minor-out rotational setting, the sugar–phosphate backbone of one strand faces out toward the solvent (backbone-out), whereas the DNA backbone of the complementary strand is facing in toward the histone octamer (backbone-in). Comparison of these structural features with our dCPD-seq data revealed that dCPD enrichment was specifically elevated at backbone-out orientations (Fig. 7B). Statistical analysis confirmed that strand-specific dCPD enrichment is significantly higher at backbone-out positions relative to backbone-in positions ($P < 0.0001$) (Fig. 7D–F).

Analysis of strand-specific dCPD enrichment, in which both the plus and minus DNA strands in strongly positioned nucleosomes were oriented in the same 5′→3′ orientation, revealed very similar patterns of dCPD enrichment between the two strands (Fig. 8A). Combining the dCPD-seq data for the two aligned strands confirmed that dCPD enrichment in cells relative to the naked DNA control was generally elevated at backbone-out positions and suppressed at backbone-in positions (Fig. 8B).

We wondered whether differences in the SASA_{C4} between backbone-out and backbone-in positions of nucleosomal DNA might also explain the observed patterns of dCPD enrichment in nucleosomes. However, our analysis of the yeast nucleosome structure indicated that there was no significant difference in SASA_{C4} between backbone-in and backbone-out positions ($P > 0.05$) (Fig. 8C).

It has been hypothesized that differences in DNA mobility could also promote CPD deamination (Song et al. 2014). To test this hypothesis, we analyzed the B-factor of the DNA backbone

for the yeast nucleosome structure, because the B-factor is a measure of DNA mobility (Sun et al. 2019). Inspection of the yeast nucleosome structure, color-coded to visualize the B-factor, indicated that the B-factor was generally low at backbone-in positions and elevated at backbone-out locations (Fig. 8D). Quantification of these data indicated that the DNA backbone B-factor was significantly elevated at backbone-out positions in the nucleosome structure and reduced at backbone-in positions ($P < 0.0001$) (Fig. 8E,F).

Discussion

The rapid deamination of cytosine bases in CPDs to uracil is thought to play a critical role in promoting C>T mutations in UV-exposed cells and skin cancers (Pfeifer et al. 2005; Ikehata and Ono 2011; Jin et al. 2021). We and others have previously shown that DNA sequence context and packaging of DNA into nucleosomes significantly alters the formation of UV-induced CPDs (Gale et al. 1987; Mao et al. 2016; Brown et al. 2018; Pich et al. 2018; Lu et al. 2021; Wilson and Wyrick 2024), but the impact of these features on CPD deamination across the genome was previously unclear. Here, we use a new method known as dCPD-seq to map genome-wide patterns of CPD deamination. Our data indicate that CPD deamination is significantly modulated by flanking sequence context, occurring most rapidly in TCG sequences in repair-deficient yeast cells and isolated genomic DNA. This propensity for rapid deamination can potentially explain our finding that TCG sequences show a much higher frequency of UV-induced C>T substitutions in both *rad16Δ* yeast and human skin cancers than expected from the frequency of CPDs in this sequence context. Our dCPD-seq data also indicate that nucleosomes significantly modulate CPD deamination rates, as dCPDs are generally elevated at minor-in positions. Structural analysis indicates that these positions are more accessible to hydrolytic attack at the key C4 atom of cytosine bases, suggesting a mechanism for elevated deamination rates at these positions. Our data also revealed strand-specific differences in CPD deamination in nucleosomal DNA, with elevated rates of CPD deamination at backbone-out positions. Structural analysis suggests that such positions have elevated DNA mobility (and flexibility), which has been previously suggested to promote CPD deamination (Song et al. 2014).

Our experimental strategy relies on arresting the UV-irradiated repair-deficient cells by shifting the *cdc13-1 rad14Δ* strain to 37°C, which results in telomere uncapping and resection of a small subset of telomeric regions (Booth et al. 2001). Because single-stranded DNA promotes CPD deamination (Cannistraro and Taylor 2009; Taylor 2022), we tested whether dCPD levels were elevated adjacent to yeast telomeres. We observed a small increase in dCPDs (normalized to cytosine-containing dipyrimidine sequence contexts) within ~ 5 kb of telomere ends, particularly in the 48 h time point in the cellular samples, but not in the naked DNA controls (Supplemental Fig. S11A–C), consistent with the hypothesis that CPD deamination is elevated in single-stranded DNA regions induced immediately adjacent to yeast telomeres. These findings also suggest that dCPD-seq can map uracils in single-stranded DNA, consistent with previous reports that CPD PL, UDG, and

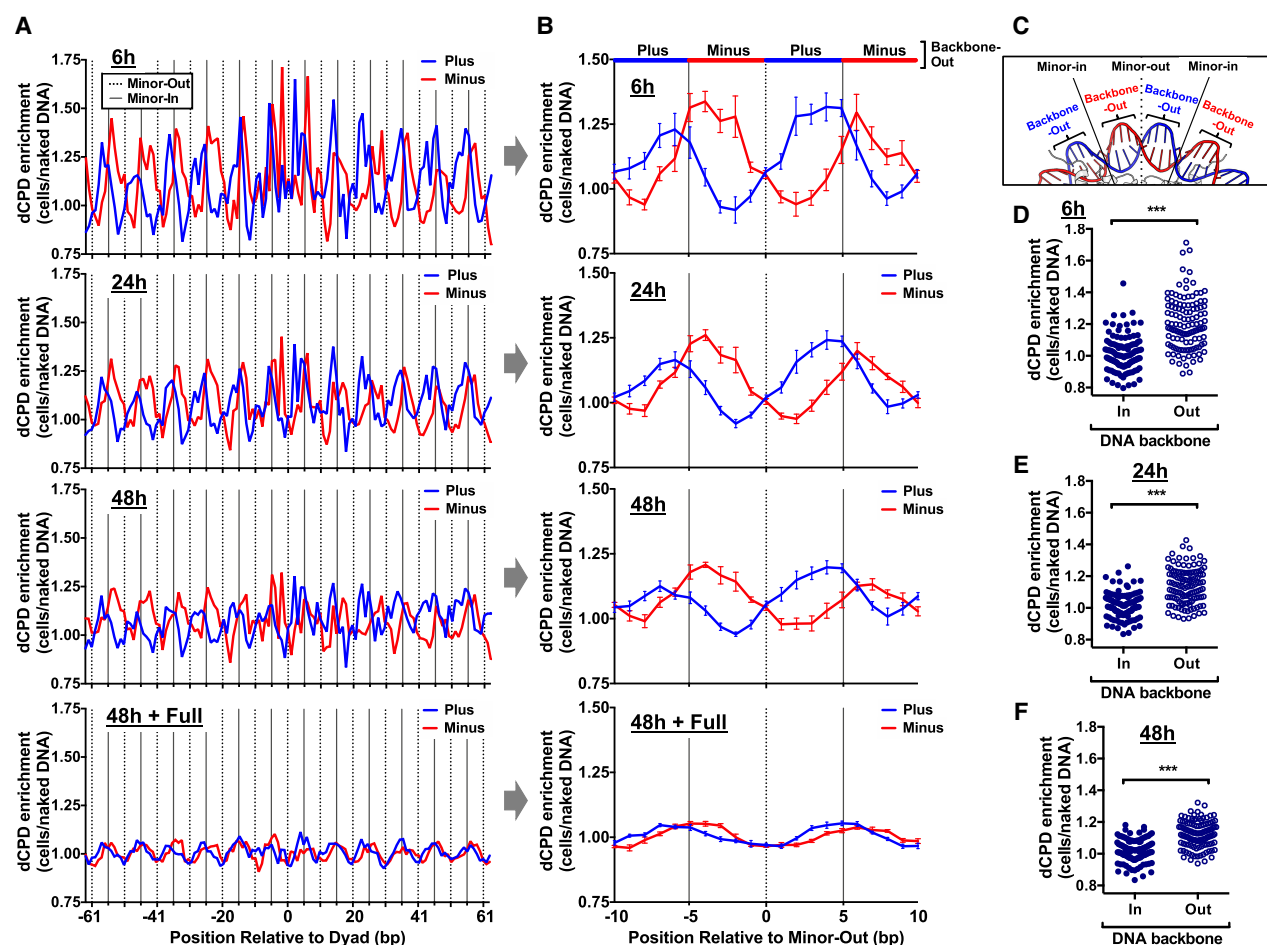


Figure 7. CPD deamination is elevated at DNA backbone-out positions and suppressed at backbone-in positions in nucleosomes. (A) Strand-specific dCPD enrichment in UV-irradiated cells relative to the in vitro deamination control is plotted at each position relative to the central dyad of about 10,000 strongly positioned nucleosomes. dCPD enrichment data are plotted for the indicated deamination time points. Dashed lines indicate minor-out rotational settings, and solid lines depict minor-in rotational settings. The mean and SEM depicted for the plus strand is in blue and minus strand in red. Nucleosome positioning data are from Brogaard et al. (2012). (B) Same as panel A, except showing strand-specific average dCPD enrichment for the 10 bp flanking minor-out rotational settings. Dashed line indicates minor-out, and solid lines indicate flanking minor-in positions. The alternating red and blue colored line at the top of the panel indicates which DNA strand is backbone-out (i.e., blue indicates the plus strand is backbone-out, and red indicates the minus strand is backbone-out) at that position in the nucleosome structure. (C) Portion of the yeast nucleosome structure (1ID3) highlighting the locations of the minor-in and minor-out rotation settings and DNA backbone-out and backbone-in positions. The plus strand is colored blue and the minus strand red. Image generated using PyMOL. (D–F) Plot of dCPD enrichment at DNA backbone-in and backbone-out rotational settings in about 10,000 strongly positioned nucleosomes for 6 h (D), 24 h (E), and 48 h (F) deamination samples. (***) $P < 0.0001$ based on a Mann-Whitney U test.

APE1 can each act on lesions in single-stranded DNA, albeit in some cases with reduced efficiency (Mahaputra Wijaya et al. 2015; Hoitsma et al. 2023). Excluding subtelomeric regions (i.e., regions within 5 kb from a telomere end) did not noticeably impact our analysis of CPD deamination in nucleosomes (Supplemental Fig. S12) or gene coding regions (Supplemental Fig. S13).

Our data indicate that genomic context also regulates CPD deamination, as dCPDs are reduced near the TSSs and TESs of yeast genes. A previous report mapping dCPDs in human cells (Jin et al. 2021) also found decreased CPD deamination near the TSSs and TESs of human genes. This prior report suggested that reduced dCPDs at the TSS could be a consequence of more rapid repair or the impact of DNA-bound transcription factors. Our yeast dCPD-seq data are derived from repair-deficient cells, indicating that differential repair is not responsible for lower levels of dCPDs near the TSSs, at least in yeast. Instead, our analysis suggests that transcrip-

tion factors and potentially other DNA-bound proteins may play a role in suppressing CPD deamination, consistent with previous reports (Cannistraro and Taylor 2010; Duan et al. 2024). The previous study also suggested that lower levels of dCPDs at the TES might simply be a consequence of differences in sequence composition (i.e., C/G poor sequences) (Jin et al. 2021). However, control dCPD-seq experiments in which the UV-irradiated DNA was deaminated in vitro did not show a decrease in normalized dCPD levels near the TESs of yeast genes, indicating that sequence composition alone is not responsible for this effect. Instead, we propose that DNA-binding by transcription factors or other proteins, such as those involved in transcription termination, suppresses CPD deamination near the TES.

Our data indicate that the DNA sequence immediately flanking a CPD significantly impacts deamination rates, both in vitro and in intact yeast cells. Our data are roughly consistent with

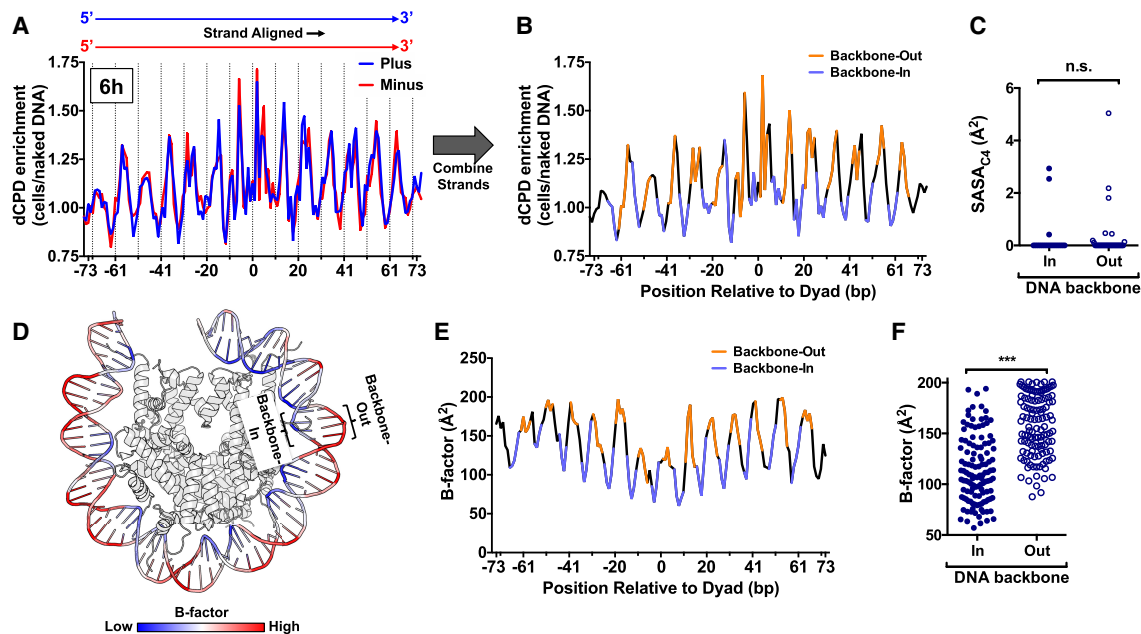


Figure 8. Structural analysis indicates that increased DNA mobility at DNA backbone-out positions promotes CPD deamination. (A) Strand-specific dCPD enrichment after 6 h deamination in cells relative to the in vitro control in about 10,000 strongly positioned nucleosomes. Data are the same as Figure 7A (top panel), except both DNA strands are aligned in the same 5'–3' orientation. dCPD enrichment data for the plus strand is in blue and for the minus strand in red. Dashed lines indicate minor-out rotational settings. (B) Same as panel A, except showing the aggregate dCPD enrichment of both DNA strands each aligned in the 5'–3' orientation. DNA backbone-out positions are colored orange, and backbone-in positions are in light purple. Intermediate positions are in black. (C) Quantification of the C4 solvent accessible surface area (SASA_{C4}) for cytosine bases in the yeast nucleosome structure (PDB ID: 1ID3) for bases at minor-in or minor-out rotational settings. (n.s.) Not significant, $P > 0.05$ based on a Mann–Whitney U test. (D) Visualization of the DNA backbone B-factor for the yeast nucleosome structure (White et al. 2001). Higher B-factors are colored red and lower B-factors in blue. Image generated using PyMOL. (E) Quantification of DNA backbone B-factors calculated from the yeast nucleosome structure (PDB ID: 1ID4). Both DNA strands were aligned in the 5'–3' orientation, and the B-factor values for both strands were averaged. DNA backbone-out positions are colored orange, backbone-in positions are in light purple, and intermediate positions are in black. (F) Plot of average DNA backbone B-factors calculated from the yeast nucleosome structure (PDB ID: 1ID4) for DNA backbone-in and backbone-out positions. (***) $P < 0.0001$ based on a Mann–Whitney U test.

previous in vitro studies that characterized CPD deamination rates in specific sequence contexts (Cannistraro and Taylor 2009; Taylor 2022). Consistent with these studies, we observe rapid deamination at TCG sequences, both in yeast cells and in vitro. It has been proposed that the O6 (and possibly N7) functional groups of the neighboring guanine base likely help to promote the hydrolytic reaction of the C4 atom in a 5mC-containing CPD (Cannistraro and Taylor 2009). Our data suggest that a similar mechanism operates for unmethylated cytosines in yeast DNA. Furthermore, analysis of mutations derived from whole-genome sequencing of melanoma tumors or UV-irradiated yeast cells revealed a striking enrichment of C>T substitutions at TCG sequences, despite relatively low damage formation in this same context. This suggests that the presence of a flanking guanine base has opposing effects on mutagenesis, as it suppresses CPD formation by quenching the excited state following UV absorption (Cannistraro and Taylor 2009; Lu et al. 2021) but promotes CPD deamination, which is inherently mutagenic. Although C>T mutations in TCG are the most frequent substitution (after normalizing for the abundance of different sequence contexts) in a data set of 140 cutaneous melanomas (Hayward et al. 2017), this mutation class does not show the same degree of enrichment in the UV-associated COSMIC signatures SBS7a and SBS7b (Forbes et al. 2017). Instead, pyrimidine-associated contexts like TCC have a larger contribution. However, many melanomas and other skin cancers included in the COSMIC database were also determined by nonnegative matrix factorization (NMF) to have enrichment

of clock-like SBS1 mutations, which overlaps SBS7a,b primarily at TCG (and CCG) sequence contexts. Hence, it is possible that C>T substitutions in TCG sequence contexts are somewhat unrepresented in SBS7a,b because in some cases they are misassigned to SBS1.

Our dCPD-seq data also indicate that nucleosomes modulate CPD deamination in UV-irradiated cells via two distinct mechanisms. First, dCPDs are enriched at minor-in rotational settings, likely because DNA bending into the minor groove at these positions may facilitate hydrolytic attack of the cytosine C4 position, which at such rotational settings is located in the solvent-facing major groove. We observed residual levels of this pattern even in the full deamination control, potentially reflecting incomplete deamination in nucleosomal DNA or, alternatively, the impact of residual CPD repair by other repair pathways. Notably, CPD formation is elevated at minor-out rotational settings, in which CPD deamination rates are low, and is suppressed at minor-in settings, in which deamination rates are high (Mao et al. 2016). These findings suggest that the same structural features that promote CPD formation may suppress subsequent CPD deamination and vice versa.

Our data indicate that nucleosomes also modulate CPD deamination via a second, strand-specific mechanism. Here, CPD deamination is elevated at backbone-out positions, where the DNA backbone of one strand faces out from the histone octamer toward the solvent, and is suppressed at backbone-in positions. These findings are consistent with previous in vitro studies, which indicated that deamination of a 5mC base in a CPD is stimulated

by nucleosomes at backbone-out positions and suppressed at backbone-in positions neighboring the nucleosome dyad (Song et al. 2011, 2014). The magnitude of the difference in dCPD enrichment that we observe in strongly positioned nucleosomes in cells is smaller than that observed for a single artificial nucleosome in vitro, consistent with a previous report (Cannistraro et al. 2015). This discrepancy could be in part because of the proximity of a CPD from the nucleosome dyad. Prior in vitro studies specifically characterized CPDs immediately adjacent to the dyad (Song et al. 2011, 2014), and the biggest effect observed in our cellular studies was similarly at backbone-out positions immediately flanking the dyad (i.e., positions -2 and $+2$) (see Fig. 5G). This dyad-proximal region is the only segment of the nucleosome in which a single DNA gyre is present (Luger et al. 1997). It is possible that the lack of a neighboring DNA gyre near the dyad might facilitate CPD deamination. It has been previously suggested that enhanced DNA flexibility/mobility at the backbone-out positions in nucleosomes may be responsible for elevated rates of CPD deamination (Song et al. 2014). Consistently, our structural analysis indicates that the B-factor in the yeast nucleosome structure is significantly elevated at backbone-out positions. Taken together, these findings suggest that two distinct biophysical mechanisms promote cytosine deamination in nucleosomes.

Although previous in vitro studies indicated that CPD deamination is elevated at backbone-out positions, these studies did not observe an increase in deamination rates at minor-in rotational settings (Song et al. 2011, 2014). One possible explanation for this discrepancy is that these in vitro studies only analyzed deamination of 5mC in a single sequence context (i.e., TCG). Our dCPD-seq data revealed elevated dCPD enrichment in TCG sequences at backbone-out positions in nucleosomes, consistent with prior in vitro studies (Song et al. 2011, 2014), but no significant enrichment at minor-in positions (Supplemental Fig. S14). One possible explanation is that for TCG sequences, DNA flexibility/mobility may have a dominant effect on deamination rates, as it can regulate the orientation of the CPD relative to the O6 (and possibly N7) functional groups of the neighboring guanine base.

A previous study used comparative genomics and analysis of spontaneous mutations in yeast strains lacking UDG (i.e., *ung1Δ*) to conclude that C>T and other spontaneous mutation classes (i.e., G>T and A>T) are suppressed in nucleosomes, being around twofold lower than in nucleosome-depleted DNA (Chen et al. 2012). Their conclusion was that the packaging of DNA into nucleosomes suppresses spontaneous cytosine deamination. In contrast, analysis of our dCPD-seq data indicate that CPD deamination is enriched in nucleosomes relative to adjacent linker DNA ($P < 0.01$) (Supplemental Fig. S15A,B). This is not simply because of differences in CPD levels, as there is no significant difference between nucleosomes and linker DNA in the fully deaminated samples (Supplemental Fig. S15C,D). These findings indicate that cytosine deamination in CPDs is elevated in yeast nucleosomes and suggest that deamination of CPDs is modulated differently by chromatin than spontaneous cytosine deamination.

In summary, our dCPD-seq data indicate that both sequence context and chromatin architecture significantly modulate CPD deamination rates across a eukaryotic genome. Our data also suggest that rapid deamination of CPDs at TCG sequences results in elevated mutation rates in both UV-irradiated yeast and human melanomas. These findings have potentially important ramifications for understanding mechanisms of UV mutagenesis in skin cancers.

Methods

Yeast strains

In vitro UV irradiation and deamination experiments were performed using yeast genomic DNA isolated from strain BY4741. Cellular experiments and naked DNA controls used the NER-deficient (i.e., *rad14Δ*) yeast strain YML461 (Supplemental Table S1), which contains the *cdc13-1* allele (Booth et al. 2001). Details of YML461 construction can be found in the Supplemental Methods.

Alkaline gel electrophoresis assays of CPD deamination

To quantify the frequency of uracil bases in genomic DNA resulting from CPD deamination, ~ 50 μ g of DNA based on nanodrop measurements from each of the prepared time points and controls was coincubated with *Escherichia coli* CPD PL and UDG (NEB) in 0.9 $\times$ CPD PL reaction buffer for 2 h under UVA (~ 365 nm) lamps (Spectroline EA-160) at room temperature. Digested samples were analyzed by alkaline gel electrophoresis essentially as previously described (Hodges et al. 2019). A detailed description of UV-irradiated sample preparation and alkaline gel analysis can be found in the Supplemental Methods. Alkaline gel quantifications of dCPDs are reported as the average and SEM resulting from three independent experiments.

dCPD-seq library construction

The dCPD-seq protocol was adapted from the CPDseq 2.0 method (Duan et al. 2024). Preparation of UV-irradiated genomic DNA and UV-irradiated cellular samples is detailed in the Supplemental Methods. Briefly, ~ 160 – 400 μ g of yeast gDNA (measured by NanoDrop spectrophotometry) was incubated with TdT (NEB) and either ddATP or ddGTP to block free 3'OH ends. After phenol:chloroform:isoamyl alcohol extraction and ethanol precipitation, the DNA was resuspended in water and 10 $\times$ CPD PL buffer (100 mM NaCl, 50 mM Tris HCl at pH 7.5, 20 mM DTT, 1 mM EDTA, 50% glycerol) to a final concentration of 0.7 $\times$ –1 $\times$ buffer, and *E. coli* CPD PL and UDG were added to reverse CPDs and excise uracil bases. Reactions were incubated under 365 nm light (Spectroline EA-160) for 2 h at room temperature, and then phenol:chloroform:isoamyl alcohol was extracted and ethanol-precipitated. Because deamination of cytosines in undamaged DNA is dramatically slower than it is in CPDs, performing the photoreversal step early in the dCPD-seq protocol was critical for minimizing unwanted in vitro deamination. APE1 (NEB) was then used to cleave the abasic sugar backbone to yield free 3'OH ends that were ligated to a splint adaptor labeled with biotin on the complementary 5' end. Ligated DNA was then sonicated (Diagenode Bioruptor 300) to achieve DNA fragments ~ 200 – 400 bp in length, and captured with streptavidin beads (Invitrogen), and washed with 1 $\times$ saline sodium citrate (SSC) buffer to remove nonspecific binding, and nonbiotinylated strand containing the ligated 3' adaptor was eluted with 0.15 M NaOH. After synthesis of the complementary strand with EconoTaq (Lucigen) and preparation of DNA ends with NEBNext End Repair and NEBNext dA-Tailing modules, a second adaptor ligation was performed followed by eight to 10 cycles of PCR amplification. Libraries were submitted for either Ion Torrent (WSU LBB1 Core Facility) or Illumina (Azenta) sequencing.

dCPD-seq data analysis

For analysis of dCPD-seq samples sequenced on the Ion Torrent sequencer, the barcode and 3' base of the read were removed using custom Perl scripts, as previously described (Mao et al. 2020),

and the sequencing reads were aligned to the standard high-quality sacCer3 reference genome using Bowtie 2 (Langmead and Salzberg 2012). The resulting SAM file was converted to a BAM file using SAMtools (Li et al. 2009) and to a BED file using BEDTools (Quinlan and Hall 2010), as previously described (Mao et al. 2016). For dCPD-seq samples sequenced using an Illumina sequencer, adapter trimming sequences were performed by Trimmomatic v0.39 (Bolger et al. 2014), with the parameters “ILLUMINACLIP:[XR-seq_adapter_file]:2:30:10.” Subsequent processing steps were the same as for Ion Torrent sequencing reads. Additional details can be found in the Supplemental Methods.

Analysis of CPD deamination rates in different sequence contexts

Enrichment of dCPD-seq reads associated with different trinucleotide sequence contexts was calculated from the fraction of total dCPD-seq reads associated with lesions occurring in that particular trinucleotide context (e.g., TCG) divided by the fraction of total trinucleotide sequences in the yeast genome associated with that particular trinucleotide context. This calculation yielded the trinucleotide enrichment score, using the equation for each trinucleotide (e.g., TCG) shown below:

$$\text{TCG trinucleotide enrichment score} = \frac{(\text{fraction of dCPD-seq reads in TCG context})}{(\text{frac. all trinucleotides in genome in TCG context})}$$

To calculate the absolute number of dCPDs in each trinucleotide sequence context, we multiplied the trinucleotide enrichment score calculated above by the frequency of uracil lesions in bulk genomic DNA measured by the alkaline gel assay. This yielded an absolute frequency of dCPDs per nucleotide in each trinucleotide context. To calculate the deamination rate for each trinucleotide context, we fit the aggregate dCPD frequency data to a single exponential curve using one-phase association kinetics in the Graphpad Prism software.

Analysis of structural parameters in nucleosomes

Software for calculating the SASA (*dr_sasa*) was downloaded from GitHub (Ribeiro et al. 2019). The high-resolution X-ray crystal structures of the *Xenopus* (1KX5) and yeast (1ID3) nucleosomes (White et al. 2001; Davey et al. 2002) were obtained from PDB. We analyzed high-resolution structures (3LZ0 and 3LZ1) of the *Xenopus* nucleosome with the strong 601 nucleosomes positioning sequence (Vasudevan et al. 2010). *dr_sasa* was used to calculate the SASA of all atoms for each structure, either using the default parameter for a spherical solvent (i.e., radius of ~1.4 Å) to simulate a water molecule or using a custom radius of 1.1 Å to simulate a hydroxide ion. A custom Python script was then used to filter output data for DNA cytosine C4 atoms. Nucleosomal DNA position offsets derived from a previous publication (Stark et al. 2022) were then used to align measurements across each structure, using position 0 as the midpoint/dyad axis of the nucleosomal DNA. For the analysis, data were divided into minor groove in/out positions, as previously described (Stark et al. 2022). Analysis of DNA mobility/flexibility using B-factor values from the yeast nucleosome structure (1ID3) was performed, as previously described, for both DNA strands aligned in the 5'→3' orientation. B-factor values were averaged for all DNA backbone atoms associated with each nucleotide position and averaged across the two aligned DNA strands.

Analysis of CPD deamination at yeast genes, TFBSs, and nucleosomes

Analysis of dCPD-seq patterns in yeast genes, TFBSs, and nucleosomes was performed essentially as previously described, using published maps of yeast TSSs and TESs (Park et al. 2014), TFBSs (Rossi et al. 2021), and nucleosome positions (Brogaard et al. 2012). Details can be found in the Supplemental Methods.

Data access

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) under accession numbers GSE284099, GSE284778, GSE285152, and GSE285153. Custom scripts are available as Supplemental Code. Source data for the graphs shown in the figures are available at Zenodo (<https://zenodo.org/records/17245989>) and as Supplemental Data.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank Jamison Haueter for technical assistance and Debra Mitchell for making the YDM14 strain. We thank Drs. Peng Mao and Mingrui Duan for helpful suggestions. This research was supported by grants from the National Institute of Environmental Health Sciences (R01ES028698 to J.J.W.), R01ES032814 (S.A.R. and J.J.W.), and R21ES035139 (J.J.W.).

Author contributions: M.F.L. performed the experiments. M.F.L., B.S., B.M-H., and J.J.W. performed the computational analysis. S.A.R. provided purified CPD photolyase enzyme and yeast strain. M.F.L., B.S., and J.J.W. compiled the figures and wrote the manuscript. All authors read and/or edited the manuscript.

References

- Bohm KA, Wyrick JJ. 2023. Damage mapping techniques and the light they have shed on canonical and atypical UV photoproducts. *Front Genet* **13**: 1102593. doi:10.3389/fgene.2022.1102593
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120. doi:10.1093/bioinformatics/btu170
- Bonilla X, Parmentier L, King B, Bezrukov F, Kaya G, Zoete V, Seplyarskiy VB, Sharpe HJ, McKee T, Letourneau A, et al. 2016. Genomic analysis identifies new drivers and progression pathways in skin basal cell carcinoma. *Nat Genet* **48**: 398–406. doi:10.1038/ng.3525
- Booth C, Griffith E, Brady G, Lydall D. 2001. Quantitative amplification of single-stranded DNA (QAOS) demonstrates that cdc13-1 mutants generate ssDNA in a telomere to centromere direction. *Nucleic Acids Res* **29**: 4414–4422. doi:10.1093/nar/29.21.4414
- Brash DE. 2015. UV signature mutations. *Photochem Photobiol* **91**: 15–26. doi:10.1111/php.12377
- Brogaard K, Xi L, Wang JP, Widom J. 2012. A map of nucleosome positions in yeast at base-pair resolution. *Nature* **486**: 496–501. doi:10.1038/nature11142
- Brown AJ, Mao P, Smerdon MJ, Wyrick JJ, Roberts SA. 2018. Nucleosome positions establish an extended mutation signature in melanoma. *PLoS Genet* **14**: e1007823. doi:10.1371/journal.pgen.1007823
- Cannistraro VJ, Taylor JS. 2009. Acceleration of 5-methylcytosine deamination in cyclobutane dimers by G and its implications for UV-induced C-to-T mutation hotspots. *J Mol Biol* **392**: 1145–1157. doi:10.1016/j.jmb.2009.07.048
- Cannistraro VJ, Taylor JS. 2010. Methyl CpG binding protein 2 (MeCP2) enhances photodimer formation at methyl-CpG sites but suppresses dimer deamination. *Nucleic Acids Res* **38**: 6943–6955. doi:10.1093/nar/gkq582
- Cannistraro VJ, Pondugula S, Song Q, Taylor JS. 2015. Rapid deamination of cyclobutane pyrimidine dimer photoproducts at TCG sites in a

- translationally and rotationally positioned nucleosome in vivo. *J Biol Chem* **290**: 26597–26609. doi:10.1074/jbc.M115.673301
- Chen X, Chen Z, Chen H, Su Z, Yang J, Lin F, Shi S, He X. 2012. Nucleosomes suppress spontaneous mutations base-specifically in eukaryotes. *Science* **335**: 1235–1238. doi:10.1126/science.1217580
- Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J Mol Biol* **319**: 1097–1113. doi:10.1016/S0022-2836(02)00386-8
- Duan M, Song S, Wasserman H, Lee PH, Liu KJ, Gordân R, He Y, Mao P. 2024. High UV damage and low repair, but not cytosine deamination, stimulate mutation hotspots at ETS binding sites in melanoma. *Proc Natl Acad Sci* **121**: e2310854121. doi:10.1073/pnas.2310854121
- Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting L, et al. 2017. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res* **45**: D777–D783. doi:10.1093/nar/gkw1121
- Frederico LA, Kunkel TA, Shaw BR. 1990. A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry* **29**: 2532–2537. doi:10.1021/bi00462a015
- Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T. 2006. *DNA repair and mutagenesis*. ASM Press, Washington, DC.
- Gale JM, Nissen KA, Smerdon MJ. 1987. UV-induced formation of pyrimidine dimers in nucleosome core DNA is strongly modulated with a period of 10.3 bases. *Proc Natl Acad Sci* **84**: 6644–6648. doi:10.1073/pnas.84.19.6644
- Hayward NK, Wilmott JS, Waddell N, Johansson PA, Field MA, Nones K, Patch A-M, Kakavand H, Alexandrov LB, Burke H, et al. 2017. Whole-genome landscapes of major melanoma subtypes. *Nature* **545**: 175–180. doi:10.1038/nature22071
- Hodges AJ, Roberts SA, Wyrick JJ. 2019. Using yeast as a model organism to study the functional roles of histone acetylation in DNA excision repair. *Methods Mol Biol* **1983**: 175–190. doi:10.1007/978-1-4939-9434-2_10
- Hoitsma NM, Norris J, Khoang TH, Kaushik V, Chadda R, Antony E, Hedglin M, Freudenthal BD. 2023. Mechanistic insight into AP-endonuclease 1 cleavage of abasic sites at stalled replication fork mimics. *Nucleic Acids Res* **51**: 6738–6753. doi:10.1093/nar/gkad481
- Ikehata H, Ono T. 2011. The mechanisms of UV mutagenesis. *J Radiat Res* **52**: 115–125. doi:10.1269/jrr.10175
- Jin SG, Pettinga D, Johnson J, Li P, Pfeifer GP. 2021. The major mechanism of melanoma mutations is based on deamination of cytosine in pyrimidine dimers as determined by circle damage sequencing. *Sci Adv* **7**: eabi6508. doi:10.1126/sciadv.abi6508
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359. doi:10.1038/nmeth.1923
- Laughery MF, Brown AJ, Bohm KA, Sivapragasam S, Morris HS, Tschmola M, Washington AD, Mitchell D, Mather S, Malt EP, et al. 2020. Atypical UV photoproducts induce non-canonical mutation classes associated with driver mutations in melanoma. *Cell Rep* **33**: 108401. doi:10.1016/j.celrep.2020.108401
- Lemaire DG, Ruzsicska BP. 1993. Kinetic analysis of the deamination reactions of cyclobutane dimers of thymidyl-3',5'-2'-deoxycytidine and 2'-deoxycytidyl-3',5'-thymidine. *Biochemistry* **32**: 2525–2533. doi:10.1021/bi00061a009
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup, et al. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079. doi:10.1093/bioinformatics/btp352
- Lindberg M, Boström M, Elliott K, Larsson E. 2019. Intragenomic variability and extended sequence patterns in the mutational signature of ultraviolet light. *Proc Natl Acad Sci* **116**: 20411–20417. doi:10.1073/pnas.1909021116
- Lu C, Gutierrez-Bayona NE, Taylor JS. 2021. The effect of flanking bases on direct and triplet sensitized cyclobutane pyrimidine dimer formation in DNA depends on the dipyrimidine, wavelength and the photosensitizer. *Nucleic Acids Res* **49**: 4266–4280. doi:10.1093/nar/gkab214
- Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–260. doi:10.1038/38444
- Mahaputra Wijaya IM, Iwata T, Yamamoto J, Hitomi K, Iwai S, Getzoff ED, Kandori H. 2015. FTIR study of CPD photolyase with substrate in single strand DNA. *Biophysics (Nagoya-shi)* **11**: 39–45. doi:10.2142/biophysics.11.39
- Mao P, Smerdon MJ, Roberts SA, Wyrick JJ. 2016. Chromosomal landscape of UV damage formation and repair at single-nucleotide resolution. *Proc Natl Acad Sci* **113**: 9057–9062. doi:10.1073/pnas.1606667113
- Mao P, Smerdon MJ, Roberts SA, Wyrick JJ. 2020. Asymmetric repair of UV damage in nucleosomes imposes a DNA strand polarity on somatic mutations in skin cancer. *Genome Res* **30**: 12–21. doi:10.1101/gr.253146.119
- Park D, Morris AR, Battenhouse A, Iyer VR. 2014. Simultaneous mapping of transcript ends at single-nucleotide resolution and identification of widespread promoter-associated non-coding RNA governed by TATA elements. *Nucleic Acids Res* **42**: 3736–3749. doi:10.1093/nar/gkt1366
- Pfeifer GP, Jin SG. 2024. Methods and applications of genome-wide profiling of DNA damage and rare mutations. *Nat Rev Genet* **25**: 846–863. doi:10.1038/s41576-024-00748-4
- Pfeifer GP, You YH, Besaratinia A. 2005. Mutations induced by ultraviolet light. *Mutat Res* **571**: 19–31. doi:10.1016/j.mrfmmm.2004.06.057
- Pich O, Muiños F, Sabarinathan R, Reyes-Salazar I, Gonzalez-Perez A, Lopez-Bigas N. 2018. Somatic and germline mutation periodicity follow the orientation of the DNA minor groove around nucleosomes. *Cell* **175**: 1074–1087.e18. doi:10.1016/j.cell.2018.10.004
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842. doi:10.1093/bioinformatics/btq033
- Ribeiro J, Ríos-Vera C, Melo F, Schüller A. 2019. Calculation of accurate interatomic contact surface areas for the quantitative analysis of non-bonded molecular interactions. *Bioinformatics* **35**: 3499–3501. doi:10.1093/bioinformatics/btz062
- Rossi MJ, Kuntala PK, Lai WKM, Yamada N, Badjatia N, Mittal C, Kuzu G, Bocklund K, Farrell NP, Blanda TR, et al. 2021. A high-resolution protein architecture of the budding yeast genome. *Nature* **592**: 309–314. doi:10.1038/s41586-021-03314-8
- Setlow RB, Carrier WL, Bollum FJ. 1965. Pyrimidine dimers in UV-irradiated poly dI:dC. *Proc Natl Acad Sci* **53**: 1111–1118. doi:10.1073/pnas.53.5.1111
- Song Q, Cannistraro VJ, Taylor JS. 2011. Rotational position of a 5-methylcytosine-containing cyclobutane pyrimidine dimer in a nucleosome greatly affects its deamination rate. *J Biol Chem* **286**: 6329–6335. doi:10.1074/jbc.M110.183178
- Song Q, Cannistraro VJ, Taylor JS. 2014. Synergistic modulation of cyclobutane pyrimidine dimer photoproduct formation and deamination at a TmCG site over a full helical DNA turn in a nucleosome core particle. *Nucleic Acids Res* **42**: 13122–13133. doi:10.1093/nar/gku1049
- Stark B, Poon GMK, Wyrick JJ. 2022. Molecular mechanism of UV damage modulation in nucleosomes. *Comput Struct Biotechnol J* **20**: S393–S400. doi:10.1016/j.csbj.2022.08.071
- Sun Z, Liu Q, Qu G, Feng Y, Reetz MT. 2019. Utility of B-factors in protein science: interpreting rigidity, flexibility, and internal motion and engineering thermostability. *Chem Rev* **119**: 1626–1665. doi:10.1021/acs.chemrev.8b00290
- Taylor JS. 2022. Deamination of C-containing cyclobutane pyrimidine dimers and its role in C to T and CC to TT signature mutations caused by UV light. In *DNA photodamage: from light absorption to cellular responses and skin cancer* (ed. Improta R, Douki T), pp. 164–196. Royal Society of Chemistry, CPI Group (UK) Ltd., Croydon, UK.
- Tu Y, Dammann R, Pfeifer GP. 1998. Sequence and time-dependent deamination of cytosine bases in UVB-induced cyclobutane pyrimidine dimers in vivo. *J Mol Biol* **284**: 297–311. doi:10.1006/jmbi.1998.2176
- Uddin KM, Flinn CG, Poirier RA, Warburton PL. 2014. Comparative computational investigation of the reaction mechanism for the hydrolytic deamination of cytosine, cytosine butane dimer and 5,6-saturated cytosine analogues. *Comput Theor Chem* **1027**: 91–102. doi:10.1016/j.comptc.2013.10.027
- Vasudevan D, Chua EY, Davey CA. 2010. Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. *J Mol Biol* **403**: 1–10. doi:10.1016/j.jmb.2010.08.039
- Verhage R, Zeeman AM, de Groot N, Gleig F, Bang DD, De Putte PV, Brouwer J. 1994. The RAD7 and RAD16 genes, which are essential for pyrimidine dimer removal from the silent mating type loci, are also required for repair of the nontranscribed strand of an active gene in *Saccharomyces cerevisiae*. *Mol Cell Biol* **14**: 6135–6142. doi:10.1128/mcb.14.9.6135-6142.1994
- White CL, Suto RK, Luger K. 2001. Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J* **20**: S207–S218. doi:10.1093/emboj/20.18.S207
- Wilson HE, Wyrick JJ. 2024. Genome-wide impact of cytosine methylation and DNA sequence context on UV-induced CPD formation. *Environ Mol Mutagen* **65 Suppl 1**: 14–24. doi:10.1002/em.22569

Received December 24, 2024; accepted in revised form October 22, 2025.