SI Materials and Methods

CPD-seq library preparation and sequencing. Yeast cells (BY4741; WT and mutant strains) were grown to mid-log phase and irradiated with 125J/m^2 or 100J/m^2 UVC light (254 nm). After UV treatment, cells were incubated in pre-warmed fresh YPD medium for repair. Cells were collected before UV irradiation (No UV), immediately after UV (0hr), and at different repair time points (e.g., 1hr and 2hr), and cell pellets were stored at the -80 ^{0}C freezer until genomic DNA isolation. Detailed protocols for genomic DNA extraction, CPD-seq library preparation and quality control, sequencing with an Ion Proton sequencer, CPD-seq reads alignment and data processing, can be found in (1). CPD damage site was identified after reads alignment by extracting the dinucleotide sequence immediately upstream of the 5' end of the read but on the opposite strand. Only sequencing reads associated with dipyrimidines (i.e., TT, TC, CT, and CC), which are likely CPD damage sites, were included for repair analysis.

Generation of Rad26 ATPase-dead and Rad3 helicase-dead mutant strains. The Rad26 ATPase-dead (K328R) and Rad3 helicase-dead (K48R) mutants were generated using the yeast CRISPR-Cas9 gene editing system (2). The guide RNA plasmid was constructed by inserting a double-stranded oligonucleotide into the expression vector pTO40 (2). Yeast cells were transformed with Cas9 expression plasmid (pTO22), guide RNA expression plasmid, and donor DNA containing the designed mutation, for gene editing. An additional silent mutation (i.e., no amino acid change) targeting the protospacer adjacent motif (PAM) was also designed in the donor DNA to prevent Cas9 binding after gene editing. The presence of the designed mutations was confirmed by Sanger sequencing and whole-genome CPD-seq data.

Genome-wide analysis of TC-NER with CPD-seq data. To compare GG-NER and TC-NER in transcribed genes, the transcribed strand (TS) and non-transcribed strand (NTS) of each yeast gene were divided into six equal sized bins. Yeast gene annotation file was obtained from a published study (3). Three additional bins were tabulated for the promoter and the terminator, respectively. The number of CPD-seq reads was counted in each bin and summed for all genes.

To analyze repair, CPD-seq count after repair (e.g., 2hr) was normalized by CPD-seq count at 0hr to calculate the fraction of remaining CPDs in each bin of each strand (see Fig 1).

To analyze TC-NER at single-nucleotide resolution, yeast genes (3) were aligned at their TSS and the number of CPD-seq reads was counted at each position from 200 bp upstream of the TSS to 650 bp downstream of the TSS. After normalizing CPDs after repair (i.e., 2hr) to CPDs at 0hr, the resulting fraction of remaining CPDs at each position was plotted for both TS and NTS. CPD-seq data generated in the Rad3-K48R mutant was normalized to the spiked pUC19 CPD count for quantitative analysis of the fraction of remaining CPDs. For high-resolution analysis of TC-NER in the termination region, genes were aligned at their polyadenylation sites (PAS) (3). Fraction of remaining CPDs at each position around the PAS was plotted for each DNA strand.

ChIP-seq and ChIP-exo data analysis. TFIIH ChIP-seq data, including Input-WT, Rad3-WT-IP (2 datasets) and Kin28-WT-IP, were obtained from a published study (4) and downloaded from ArrayExpress (Accession number: E-MTAB-1595). TFIIH ChIP-seq reads and the matched input sequencing reads were aligned to the yeast reference genome SacCer3 and duplicate sequencing reads were removed. The resulting BAM files were analyzed with the MACS2 software to call ChIP-seq peaks, using input reads as the control. Significantly enriched Rad3 and Kin28 peaks were collected and their occupancy around gene TSS was analyzed using a custom Python script. Spt5 ChIP-seq raw data were obtained from a published study (5) and downloaded from NCBI Gene Expression Omnibus (GEO; accession number GSE79222). Spt5 peak calling and its occupancy around gene TSS were performed the same as described above.

TFIIH ChIP-exo sequencing data were obtained from a published study (6), downloaded from GEO (accession number GSE98573), and aligned to the yeast reference genome SacCer3. The resulting BAM files were sorted and analyzed with the MACE software (7) to identify significantly enriched TFIIH peaks. Peak density around gene TSS was analyzed using a custom Python script.

NET-seq data analysis. NET-seq data in wild-type yeast cells were obtained from a published study (8). Positions of the 3' end of the nascent transcript, which are indicative of Pol II locations, were downloaded from GEO (accession number GSE68484). The number of the 3' ends of nascent transcripts was counted in the coding region (i.e., from TSS to PAS) for each gene and normalized to the gene length to obtain NET-seq density (i.e., NET-seq per base). Genes were sorted by the NET-seq density. The top one third and bottom one third of genes were grouped into high-Pol II and low-Pol II genes, respectively.

Yeast UV sensitivity assay. For spot test, yeast cells were cultured in YPD medium to mid-log phase. Cells were 10-fold serially diluted in fresh YPD broth and spotted on YPD plates. After exposing to UV light, plates were incubated for 3 days. For quantitative survival test, yeast cells were plated on YPD plates and exposed to desired UV doses. The number of colonies on each plate was counted after incubating for 3 days and the survival rate for each dose was calculated.

SI Figures and Figure Legends

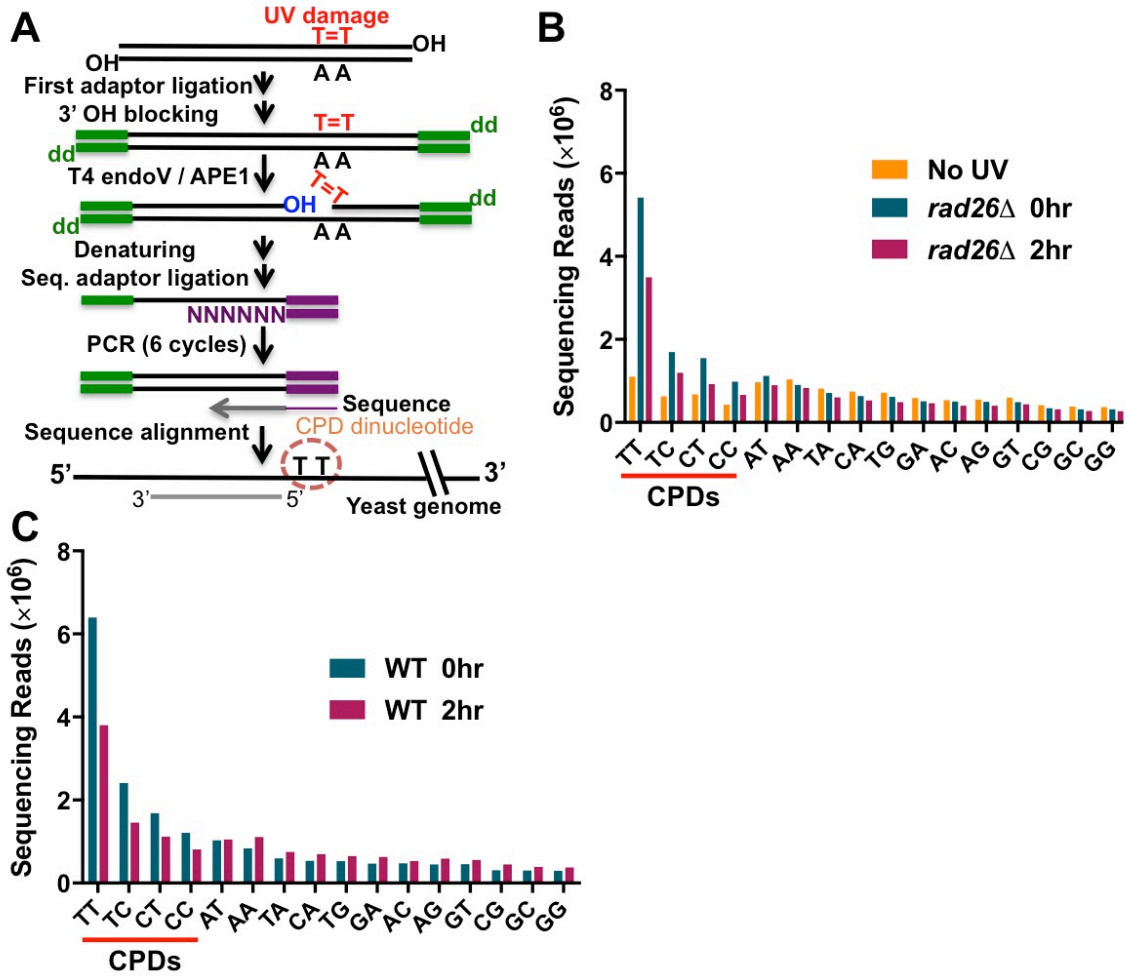

Figure S1: CPD-seq procedure and dinucleotides counting. (A) Schematic showing the CPD-seq procedure [adapted from (1)]. Genomic DNA is sonicated to short fragments and ligated with the first adaptor DNA (green). All free 3'-OHs are blocked by terminal transferase and didoxyATP (dd). CPD is cleaved by T4 endonuclease V (T4 endo) and AP endonuclease (APE1) to generate a new ligatable 3'-OH immediately upstream of the CPD damage. DNA fragments are denatured and ligated to the second adaptor DNA (purple), which is double-stranded with 6 random nucleotides overhang. After streptavidin beads purification and second strand synthesis, the resulting CPD-seq library is amplified with primers complimentary to the first and the second adaptors and sequenced with the Ion Torrent sequencer. Sequencing reads are aligned to the reference genome. The two nucleotides immediately upstream of the 5' end of the read are identified, and the dinucleotide sequence on the opposite strand is recognized as the CPD damage site. (B) Dinucleotide counts reveal high enrichment of dipyrimidines such as TT, TC, CT, and CC in UV-irradiated samples, but not in the No UV control. (C) Same as part B, except dinucleotide counts in WT samples are shown.

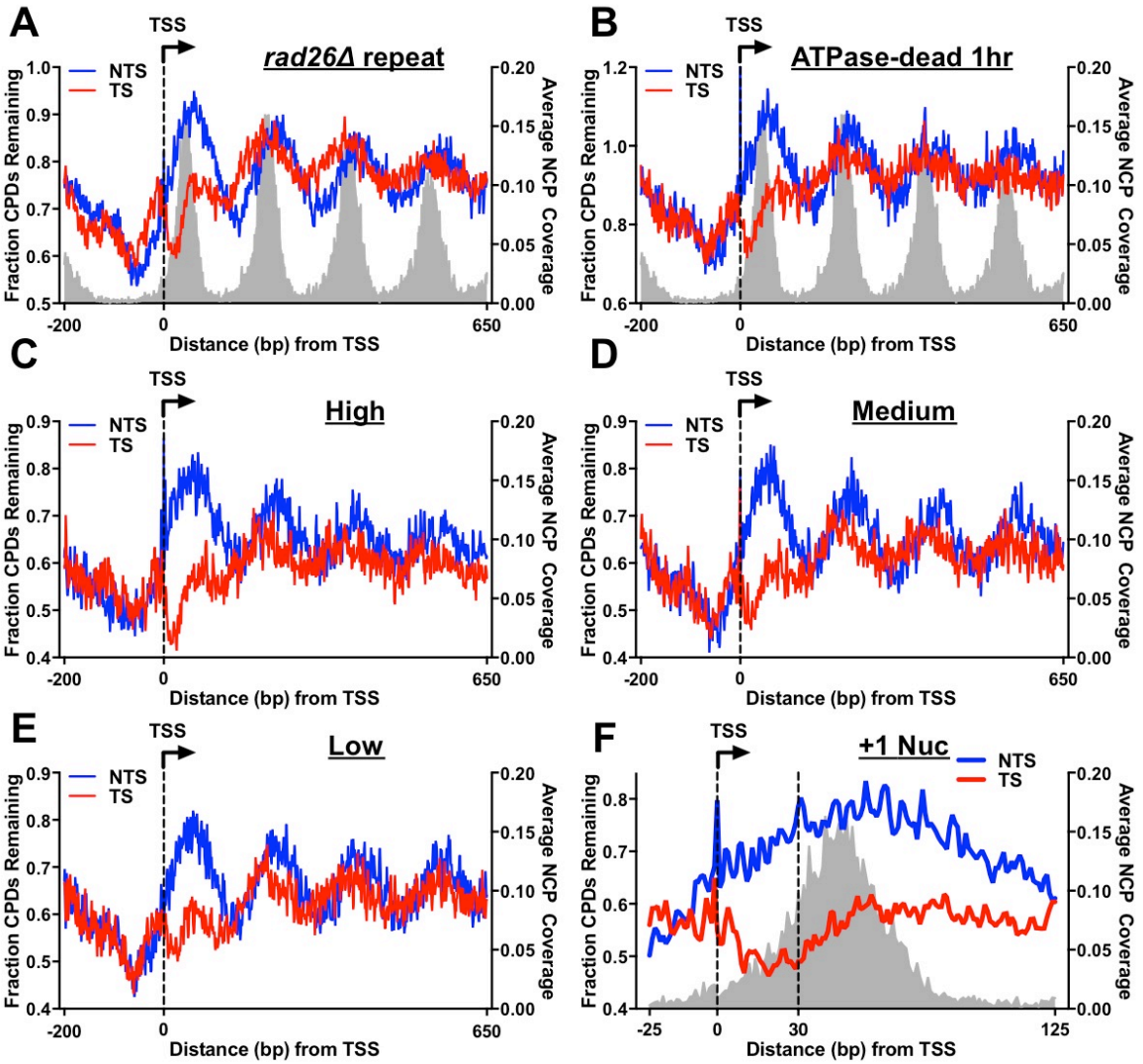

Figure S2: Rad26 is required for TC-NER in regions downstream of the +1 nucleosome across the yeast genome. (A) Biological repeat of the CPD-seq experiment in the *rad26Δ* mutant. Yeast genes were aligned at the transcription start site (TSS) and remaining CPDs at 2hr (normalized to 0hr) were plotted around the TSS at each position. (B) CPD-seq data analysis following 1hr repair in the Rad26 ATPase-dead strain. Remaining CPDs at 1hr (normalized by 0hr) were plotted around the TSS of transcribed genes. (C) to (E) High-resolution repair analysis showing that Rad26 is generally required for TC-NER in downstream nucleosomes (e.g., +2, +3, and +4) for high, medium, and low expression genes. Genes were stratified by their transcription frequency (9). CPD-seq data in *rad26Δ* was analyzed in gene groups with different transcription levels. (F) Rad26-independent TC-NER was particularly prominent for the first ~ 30 nt downstream of the TSS. CPD-seq data in *rad26Δ* was analyzed in the +1 nucleosome, which consists of DNA ranging from -23 nt to +123 nt relative to the TSS. Gray background depicts the +1 nucleosome.

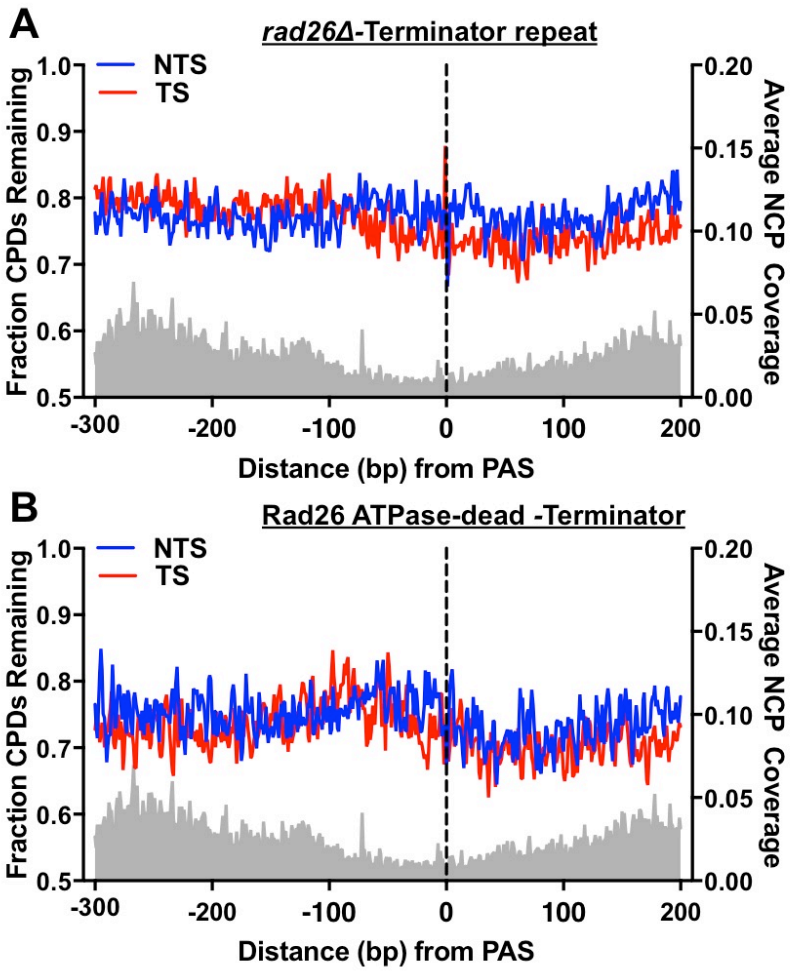

Figure S3: Rad26 and its ATPase activity is essential for TC-NER in the termination region. (A) Biological repeat of the CPD-seq experiment in the *rad26Δ* mutant. Graph shows high-resolution analysis of remaining CPDs at 2hr (normalized by 0hr) in the *rad26Δ* mutant around the polyadenylation site (PAS) of all yeast genes. (B) CPD-seq data in Rad26 ATPase-dead strain was analyzed in the termination region. Remaining CPDs at at 2hr (normalized by 0hr) were analyzed around the PAS and base-resolution data is presented.

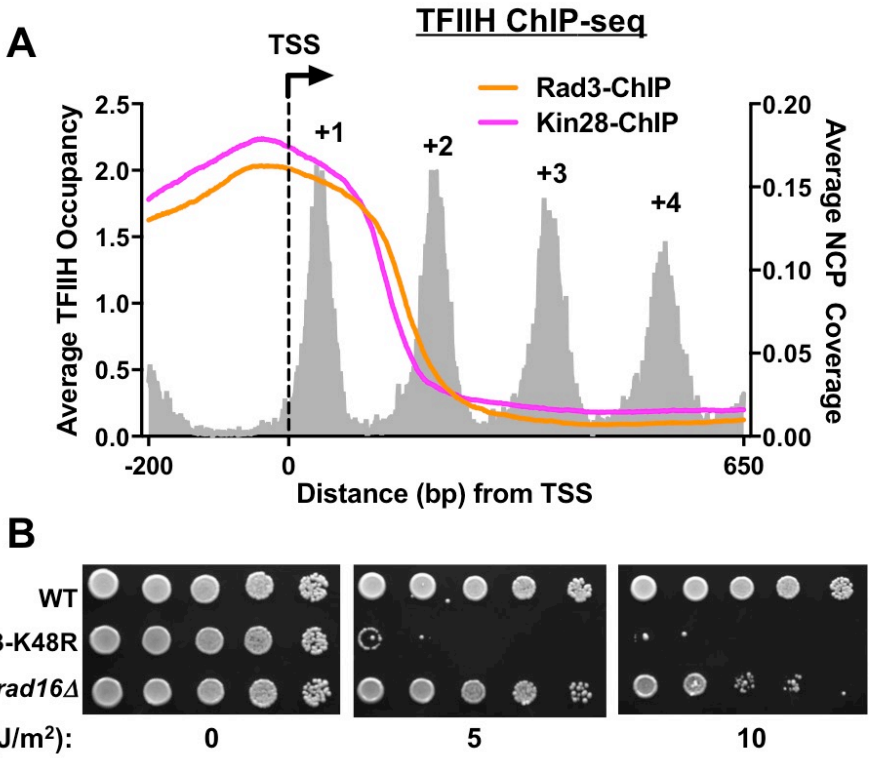

Figure S4: ChIP-seq analysis showing high TFIIH occupancy in the +1 nucleosome. (A) Analysis of Rad3 and Kin28 (two subunits of TFIIH) ChIP-seq data revealed high occupancy of TFIIH in the promoter and the +1 nucleosome, and TFIIH occupancy decreases in the downstream nucleosomes. Rad3 and Kin28 ChIP-seq and the matched input data were derived from a published study (4). Rad3 and Kin28 occupancy is shown between -200 bp and +650 bp relative to the TSS. Nucleosomes are shown by the gray background (right Y axis). (B) Loss of the Rad3 DNA helicase activity leads to high UV sensitivity in yeast. The Rad3 helicase-dead mutant (i.e., Rad3-K48R) cells (10-fold serial dilution) were spotted on YPD plates and UV irradiated at the desired dose. Pictures were taken 72hr after UV treatment.

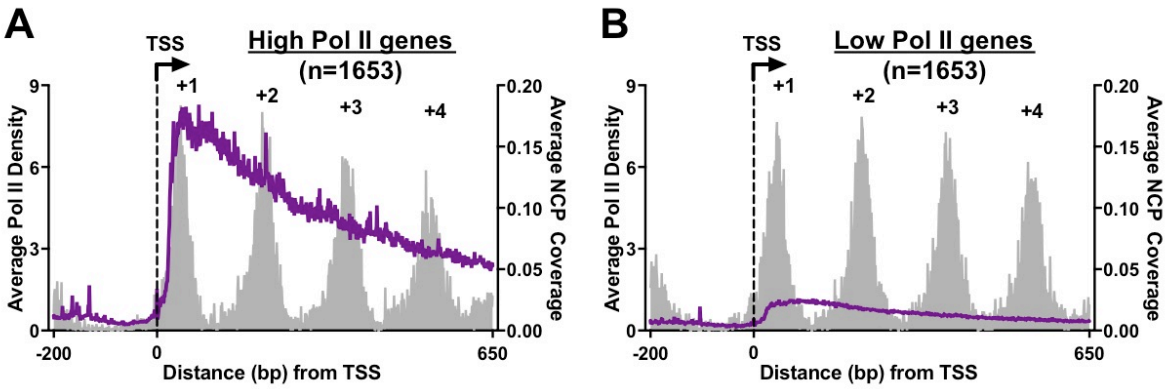

Figure S5: Pol II density in high-Pol II and low-Pol II genes. Yeast genes were sorted by their Pol II density, using NET-seq data derived from the published study (8). The average Pol II density for the top one third of genes (n=1,653) (A) and the bottom 1,653 genes (B) was plotted around the transcription start site (TSS), respectively.

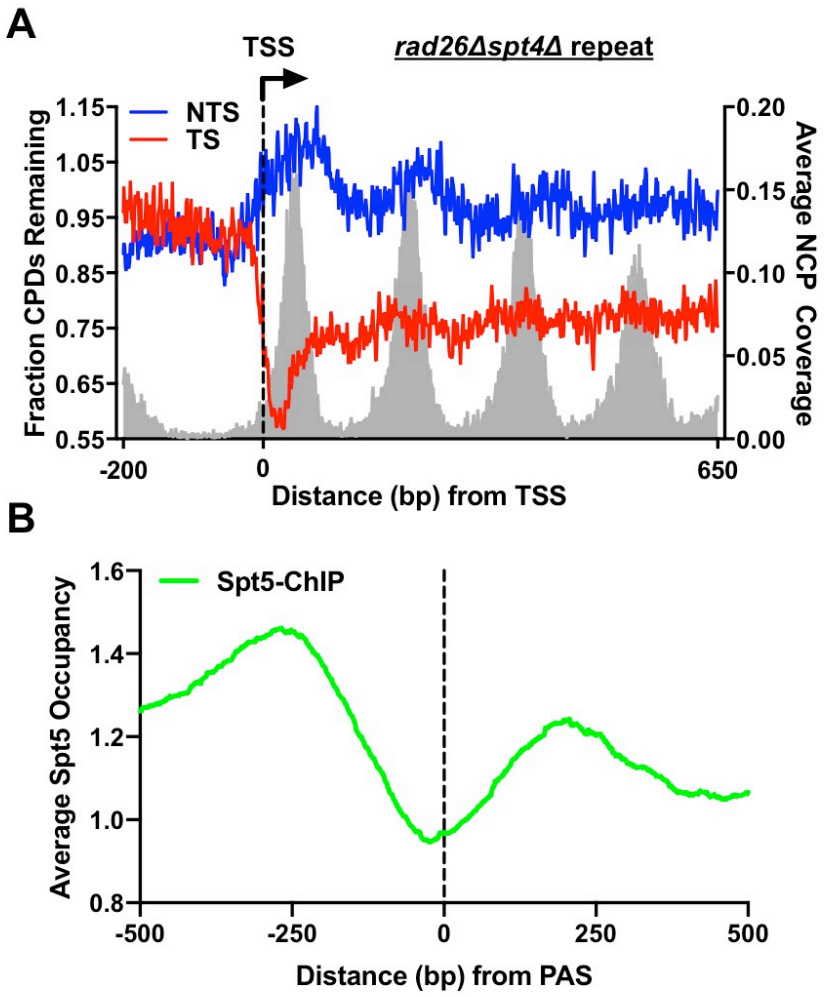

Figure S6: *SPT4* deletion significantly rescues TC-NER in the *rad26Δ* mutant. (A) Biological repeat of the CPD-seq experiment in the *rad26*Δ*spt4*Δ double mutant. CPD-seq data in *rad26*Δ*spt4*Δ was analyzed and the remaining CPDs at 2hr (normalized by 0hr) were plotted around the TSS at base-resolution for both transcribed strand (TS) and non-transcribed strand (NTS). (B) ChIP-seq analysis showing high occupancy of Spt5 upstream of the polyadenylation site (PAS), which is consistent with the observation that TC-NER is also restored in the transcription termination region by *SPT4* deletion. Spt5 ChIP-seq data and the matched input data were derived from the published study (5).

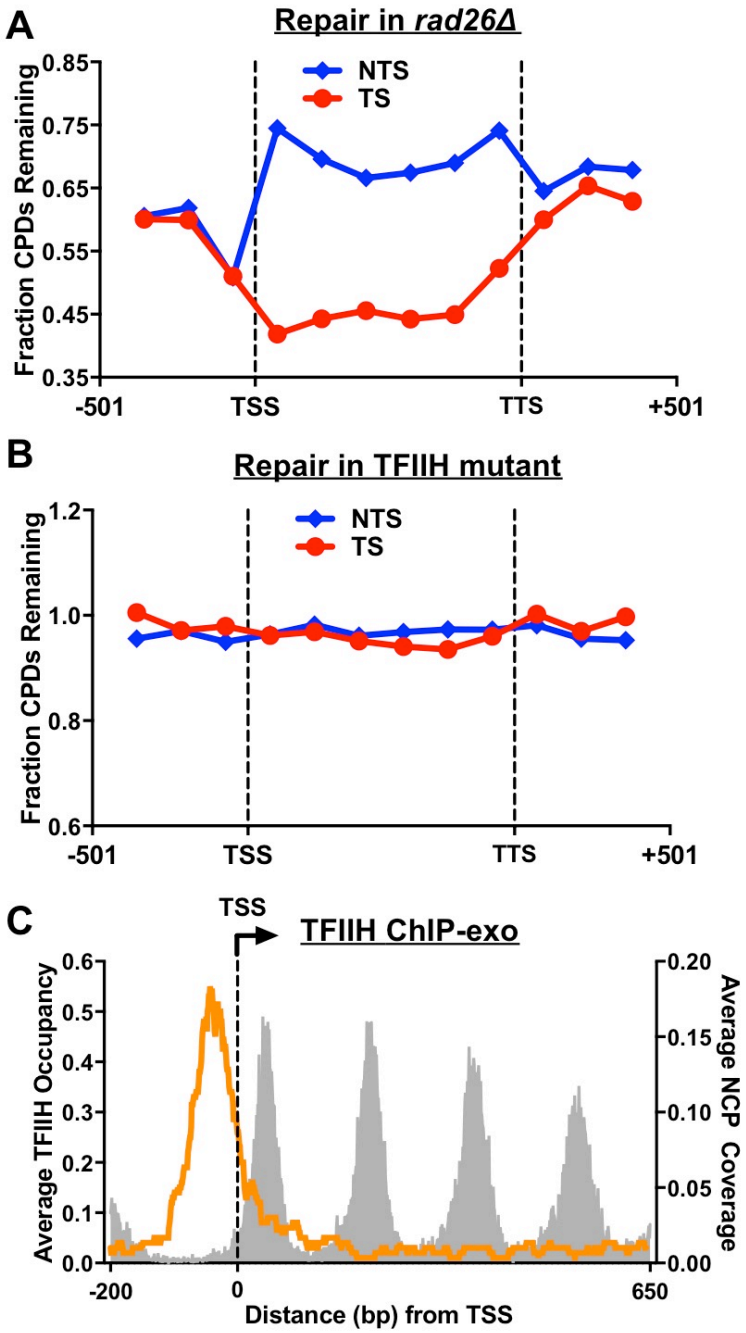

Figure S7: Repair and TFIIH occupancy in Rad26-independent genes. (A) High Rad26-independent TC-NER in a small number of highly transcribed genes. The top 100 Rad26-independent genes were chosen and each gene was separated into 6 equal-sized bins with additional bins in the promoter and terminator. CPD-seq data generated in *rad26*Δ was analyzed in these genes. Remaining CPDs at 2hr (normalized to 0hr) were analyzed in each bin and plotted for both DNA strands. NTS: non-transcribed strand. TS: transcribed strand. (B) No TC-NER was seen in the top 100 Rad26-independent genes in the TFIIH helicase-dead mutant (i.e., Rad3-K48R). Same as part A, except CPD-seq data from the Rad3-K48R mutant was analyzed. (C) Ssl2 ChIP-exo analysis showing low TFIIH occupancy in the coding region downstream of the +1 nucleosome in the top 100 Rad26-independent genes.

Systematic Name	Standard Name	Function
YLR110C	CCW12	Cell wall mannoprotein
YBR118W	TEF2	Translational elongation factor
YGR063C	SPT4	Transcription elongation factor
YNL145W	MFA2	Mating pheromone a-factor
YPR080W	TEF1	Translational elongation factor
YGR192C	TDH3	Glycolysis enzyme
YLL039C	UBI4	Ubiquitin
YKR085C	MRPL20	Mitochondrial ribosomal protein
YLR043C	TRX1	Cytoplasmic thioredoxin isoenzyme
YFL026W	STE2	Receptor for alpha-factor pheromone
YCL040W	GLK1	Glucokinase
YNL281W	HCH1	Heat shock protein regulator
YAL005C	SSA1	ATPase involved in protein folding
YNR014W	N/A	Unknown
YNL322C	KRE1	Cell wall glycoprotein
YKL152C	GPM1	Phosphoglycerate mutase
YOR052C	TMC1	AN1-type zinc finger protein
YLL024C	SSA2	HSP70 family ATP-binding protein
YGR275W	RTT102	Chromatin remodeling factor
YLR206W	ENT2	Epsin-like protein

Table S1: Name and function of the top 20 yeast genes that are independent or partially independent of Rad26 to activate TC-NER.

References

1. P. Mao, M. J. Smerdon, S. A. Roberts, J. J. Wyrick, Chromosomal landscape of UV damage formation and repair at single-nucleotide resolution. *PNAS* **113**, 9057–9062 (2016).

2. M. F. Laughery, *et al.*, New vectors for simple and streamlined CRISPR-Cas9 genome editing in *Saccharomyces cerevisiae*: Vectors for simple CRISPR-Cas9 genome editing in yeast. *Yeast* **32**, 711–720 (2015).

3. D. Park, A. R. Morris, A. Battenhouse, V. R. Iyer, 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 (2014).

4. F. Eyboulet, *et al.*, Mediator links transcription and DNA repair by facilitating Rad2/XPG recruitment. *Genes Dev.* **27**, 2549–2562 (2013).

5. C. Baejen, *et al.*, Genome-wide Analysis of RNA Polymerase II Termination at Protein-Coding Genes. *Molecular Cell* **66**, 38-49.e6 (2017).

6. V. Vinayachandran, *et al.*, Widespread and precise reprogramming of yeast protein–genome interactions in response to heat shock. *Genome Res.* **28**, 357–366 (2018).

7. L. Wang, *et al.*, MACE: model based analysis of ChIP-exo. *Nucleic Acids Res.* **42**, e156 (2014).

8. K. M. Harlen, *et al.*, Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. *Cell Reports* **15**, 2147–2158 (2016).

9. F. C. Holstege, *et al.*, Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**, 717–728 (1998).