



OPEN ACCESS

EDITED BY

Sumeet Nayak,
KSQ Therapeutics, United States

REVIEWED BY

Bevin P. Engelward,
Massachusetts Institute of Technology,
United States
Michael G. Kemp,
Wright State University, United States

*CORRESPONDENCE

Natalie Saini

✉ sainina@muscc.edu

RECEIVED 09 August 2024

ACCEPTED 26 August 2024

PUBLISHED 12 September 2024

CITATION

Blouin T and Saini N (2024) Aldehyde-induced DNA-protein crosslinks- DNA damage, repair and mutagenesis. *Front. Oncol.* 14:1478373. doi: 10.3389/fonc.2024.1478373

COPYRIGHT

© 2024 Blouin and Saini. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Aldehyde-induced DNA-protein crosslinks- DNA damage, repair and mutagenesis

Thomas Blouin and Natalie Saini*

Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, SC, United States

Aldehyde exposure has been shown to lead to the formation of DNA damage comprising of DNA-protein crosslinks (DPCs), base adducts and interstrand or intrastrand crosslinks. DPCs have recently drawn more attention because of recent advances in detection and quantification of these adducts. DPCs are highly deleterious to genome stability and have been shown to block replication forks, leading to wide-spread mutagenesis. Cellular mechanisms to prevent DPC-induced damage include excision repair pathways, homologous recombination, and specialized proteases involved in cleaving the covalently bound proteins from DNA. These pathways were first discovered in formaldehyde-treated cells, however, since then, various other aldehydes have been shown to induce formation of DPCs in cells. Defects in DPC repair or aldehyde clearance mechanisms lead to various diseases including Ruijs-Aalfs syndrome and AMeD syndrome in humans. Here, we discuss recent developments in understanding how aldehydes form DPCs, how they are repaired, and the consequences of defects in these repair pathways.

KEYWORDS

DNA protein crosslink, aldehydes, DNA damage, DNA repair, genome instability

Introduction

Aldehydes are reactive carbonyl compounds that are abundantly present in the environment, ambient air, and human diets. Human cells also generate endogenous aldehydes from a variety of sources. Because of their polar carbon-oxygen double bond, they can react with all the major biomolecules of the cell, including DNA. Notably, multiple aldehydes can form direct DNA base adducts and DNA-protein crosslinks (DPC). The sources, types, adducts, and mutagenic effects of aldehydes have all been thoroughly reviewed (1). DPCs are a highly variable class of DNA lesions that can vary in size depending on the identity of the crosslinked protein. These bulky adducts pose a significant threat to DNA-based processes and have the ability to block transcription if the DPC is located on the template strand (2). DPCs can also block replication fork progression when they are located on the leading strand by inhibiting CMG helicase movement (3, 4). Finally, DPCs may inhibit new DNA synthesis by blocking DNA polymerases and have been shown

to interfere with chromatin remodelling (5–7). Overall, various studies have shown that DPCs are an extremely heterogeneous class of DNA damage that can have different impacts on the genome depending on the identity and location of the crosslinked protein (8, 9).

As evidence of the importance of DPC repair in maintaining genomic stability, mutations in *SPRTN*, encoding a specialized DPC protease, have been shown to be the cause of Rujis-Aalfs syndrome. This genetic disorder leads to genomic instability, early-onset hepatocellular carcinoma, and progeria (10, 11). This cancer predisposition phenotype in the absence of efficient DPC processing underscores how common and damaging these lesions are in the genome.

The goal of this review is to summarize which aldehydes generate DPCs and how these adducts are formed. We will also review how these DPCs are repaired, the consequences of improper repair, and avenues for future work in the field.

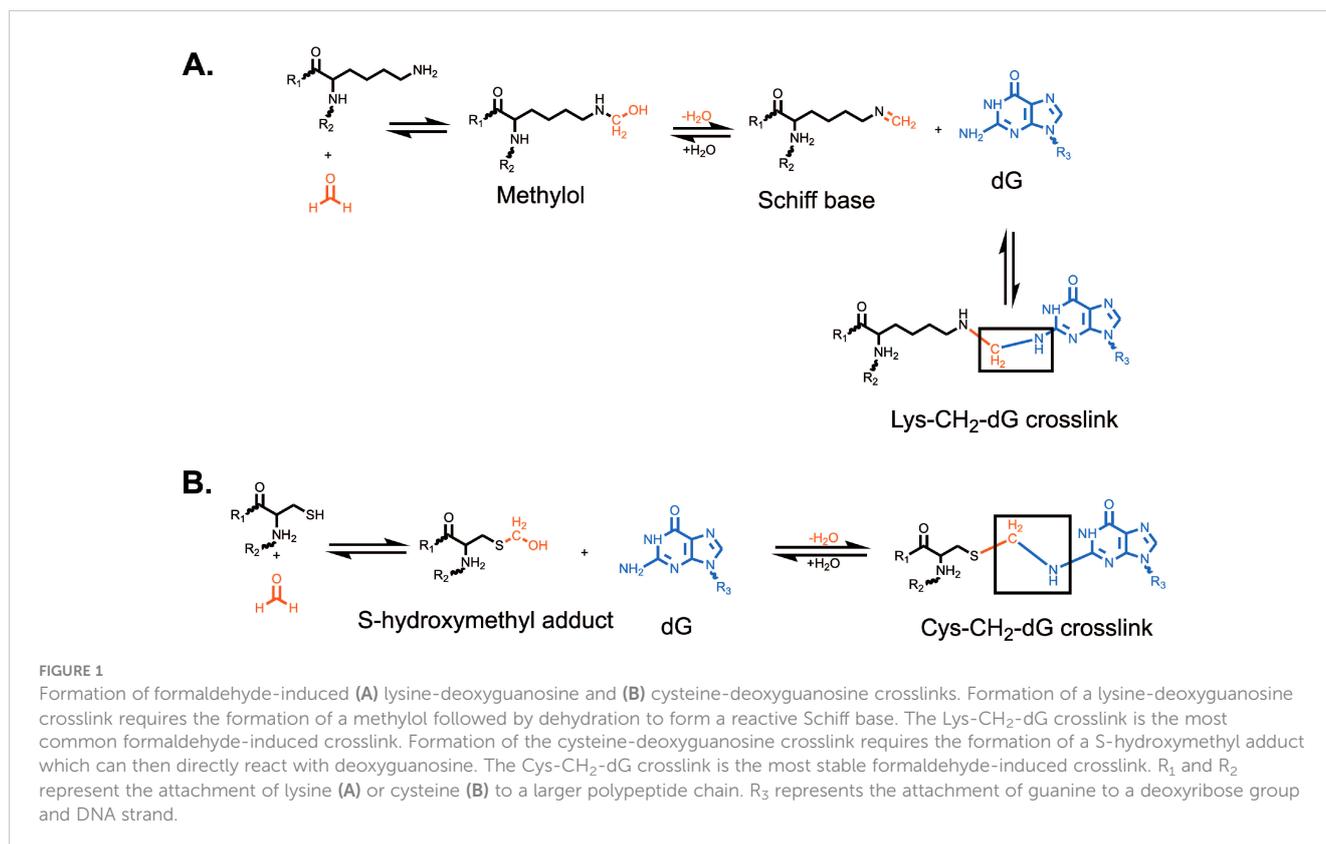
Aldehydes as sources of DPCs—formaldehyde, acetaldehyde, acrolein, methylglyoxal, and malondialdehyde

Formaldehyde

Formaldehyde is environmentally abundant from sources such as vehicle exhaust, factories, cigarettes, nail salons and is formed upon consumption of aspartame or methanol (12, 13). It is also produced endogenously in close proximity to DNA by processes

such as histone demethylation (14, 15) and repair of DNA-methylation damage by AlkB-family dioxygenases (16, 17). Finally, it can be generated via one-carbon metabolism (18). Formaldehyde is highly reactive towards nucleotides and proteins, generating various DNA adducts which contribute to genomic instability (19, 20). Amongst the most studied adducts induced by formaldehyde are DPCs. Formaldehyde generally creates crosslinks by forming a covalent bond with a nucleophilic group, such as amines, amides, thiols, and hydroxyls, which generates a methylol adduct (21). This methylol can then be converted into a Schiff base via a dehydration reaction. These Schiff bases are unstable and either decompose or are stabilized when they react with a nucleophilic group in an adjacent molecule, forming a methylene bridge that links the two molecules (21). Formaldehyde has been shown to react with cysteine, histidine, lysine, tryptophan, and arginine (22). The most common formaldehyde-induced DPC occurs between deoxyguanosine and lysine, while the most stable DPC is between deoxyguanosine and cysteine (21) (Figure 1). It has been suggested that formaldehyde can only crosslink proteins that interact with DNA for greater than 5 seconds (23). For this reason, lysine and arginine-rich histones have been identified as the most common formaldehyde crosslinked proteins (24–26). Other commonly crosslinked proteins include topoisomerases and polymerases (26).

Formaldehyde-induced DPCs have been identified in cultured cells, animals, and humans exposed to formaldehyde. Formaldehyde is a known carcinogen and has been shown to be involved in the etiology of nasopharyngeal cancer in humans and squamous cell carcinomas in the nasal respiratory epithelium of



rodents (27–29). Various aldehyde dehydrogenases (ALDH) and enzymes that process aldehyde-glutathione conjugates have been implicated in aldehyde clearance from cells (30). Mice carrying defects in alcohol dehydrogenase 5 (ADH5) and FANCD2 demonstrated formaldehyde-induced genotoxicity in hepatocytes, nephrons, and hematopoietic stem cells (31). Further, defects in ADH5 and ALDH2 led to elevated formaldehyde-induced DNA damage, mutagenesis, and carcinogenesis. Defective ADH5 and ALDH2 led to decreased clearance of endogenously formed formaldehyde leading to accumulation of this genotoxin (32). While these studies did not directly assay DPCs, it is likely that DPC formation was at least partly responsible for the observed genotoxicity. In an effort to determine the relative contribution of endogenous and exogenous formaldehyde in generating DPCs, rats were exposed to isotope labeled ($^{13}\text{CD}_2$]-formaldehyde) formaldehyde and various tissues were subjected to mass spectrometry (33). The researchers observed an increase in labeled exogenous DPCs in the nasal respiratory tissues but found no labeled DPCs in other tissues (33). Endogenous formaldehyde-induced DPCs were present in all the tested tissue types. This data suggests that endogenous formaldehyde, which is present in human blood at a concentration of $\sim 100\ \mu\text{M}$, could be responsible for a vast majority of DPCs (33, 34). Similarly, DPCs were detected in the respiratory tract of rhesus monkeys exposed to formaldehyde (35). Finally, elevated DPC levels were detected in lymphocytes obtained from workers in hospital pathology departments who had chronic occupational exposure to formaldehyde (36). Such DPCs likely underlie formaldehyde-induced genotoxicity and carcinogenicity.

Acetaldehyde

Humans are exposed to acetaldehyde from cigarette smoke and alcohol consumption. Acetaldehyde can react readily with deoxyguanosine to generate N²-ethylidene-deoxyguanosine (37, 38). While acetaldehyde can form DPCs, they have been shown to be unstable under physiological conditions. Treatment of histone-bound plasmids with acetaldehyde for one hour at 37°C demonstrated an increase in DPCs; however, only 25% of crosslinks persisted past eight hours with an average half-life of 1.5–2 hours (39, 40). Nonetheless, defective DPC repair in fission yeast led to increased acetaldehyde-induced genotoxicity (41). Similarly, we recently demonstrated that deletion of DPC repair proteins in budding yeast leads to increased acetaldehyde-induced cytotoxicity and mutagenesis (42). These data indicate that despite the relatively low stability of acetaldehyde-induced DPCs *in vitro*, these moieties are potential sources of genomic instability *in vivo*.

Acrolein

Acrolein, or 2-propenal, is abundantly present in cooked foods and the environment. Some common sources of acrolein include tobacco smoke, lipid metabolism, and fossil fuel combustion. It can also be produced endogenously from methionine and threonine synthesis (43). Acrolein reacts readily with nucleophiles such as

deoxyguanosine to make mostly α - and γ -hydroxy-1,N²-propano-2'-deoxyguanosine (1, 43). These adducts have been shown to promote the formation of DNA interstrand crosslinks and have been used *in vitro* to generate acrolein-induced peptide adducts (44, 45). Finally, acrolein has been shown to react with the thiol group of cysteine, the imidazole group of histidine, and the amino groups of lysine and arginine to form Schiff bases and crosslinks (43, 46, 47).

Methylglyoxal

Methylglyoxal, also known as pyruvaldehyde or 2-oxo-propanal, is an oxo-aldehyde that is generated endogenously from triose phosphate precursors during respiration (1, 48). Humans are also exposed to methylglyoxal from food and beverages, cigarette smoke, and vehicle exhaust (48). Methylglyoxal generates DNA adducts on deoxyguanosine and deoxyadenosine, and has an ability to react with arginine, lysine, and cysteine (48, 49). The protein and DNA adducts formed by methylglyoxal are referred to as advanced glycation end-products (AGEs), which have been implicated in many diseases such as cancer, diabetes, and various neurodegenerative diseases (1, 50, 51). Methylglyoxal is a bis-electrophile, meaning that it can form crosslinks between two independent nucleophiles through two electrophilic centers. Previously, it has been shown to form crosslinks *in vitro* between DNA and the *E. coli* DNA polymerase I (52). In this study, crosslinks were only observed between 2'-deoxyguanosine and *N*-acetyllysine or *N*-acetylcysteine, suggesting that methylglyoxal specifically makes deoxyguanosine-lysine or deoxyguanosine-cysteine DPCs (52).

Malondialdehyde

Malondialdehyde is produced by polyunsaturated fatty acid peroxidation (53–55). Under physiological conditions, malondialdehyde is an enolate anion and is of relatively low reactivity (56). Despite this, it has been shown experimentally to produce DNA base adducts which are typically found on deoxyguanosine and deoxyadenosine (57, 58). Malondialdehyde has been shown to crosslink histones to DNA (59). Histone H1 was found to have an especially high affinity for malondialdehyde-induced crosslinking, and the authors concluded that this was likely due to the abundance of lysine residues within histone H1 (59). Importantly, these crosslinks formed readily at physiological temperature and pH, and were stable for up to 13 days. Malondialdehyde crosslinking was found to be limited to proteins that physically bind DNA, as a control BSA protein was unable to form crosslinks regardless of the concentration of malondialdehyde (59).

Repair of aldehyde-induced DPCs

Various studies *in vitro* and in model systems have shown that DPCs block DNA polymerases (5, 60–62). Moreover, treatment of V79 Chinese hamster cells with formaldehyde led to elevated levels

of sister chromatid exchange and micronuclei formation, likely due to abundant DPCs (63). These data indicate that DPCs formed by formaldehyde lead to genomic instability, which significantly contributes to formaldehyde-induced carcinogenesis. As such, mechanisms to remove DPCs are conserved from yeast to mammals. These mechanisms will be reviewed in the following section.

Proteolysis-dependent DNA-protein crosslink repair

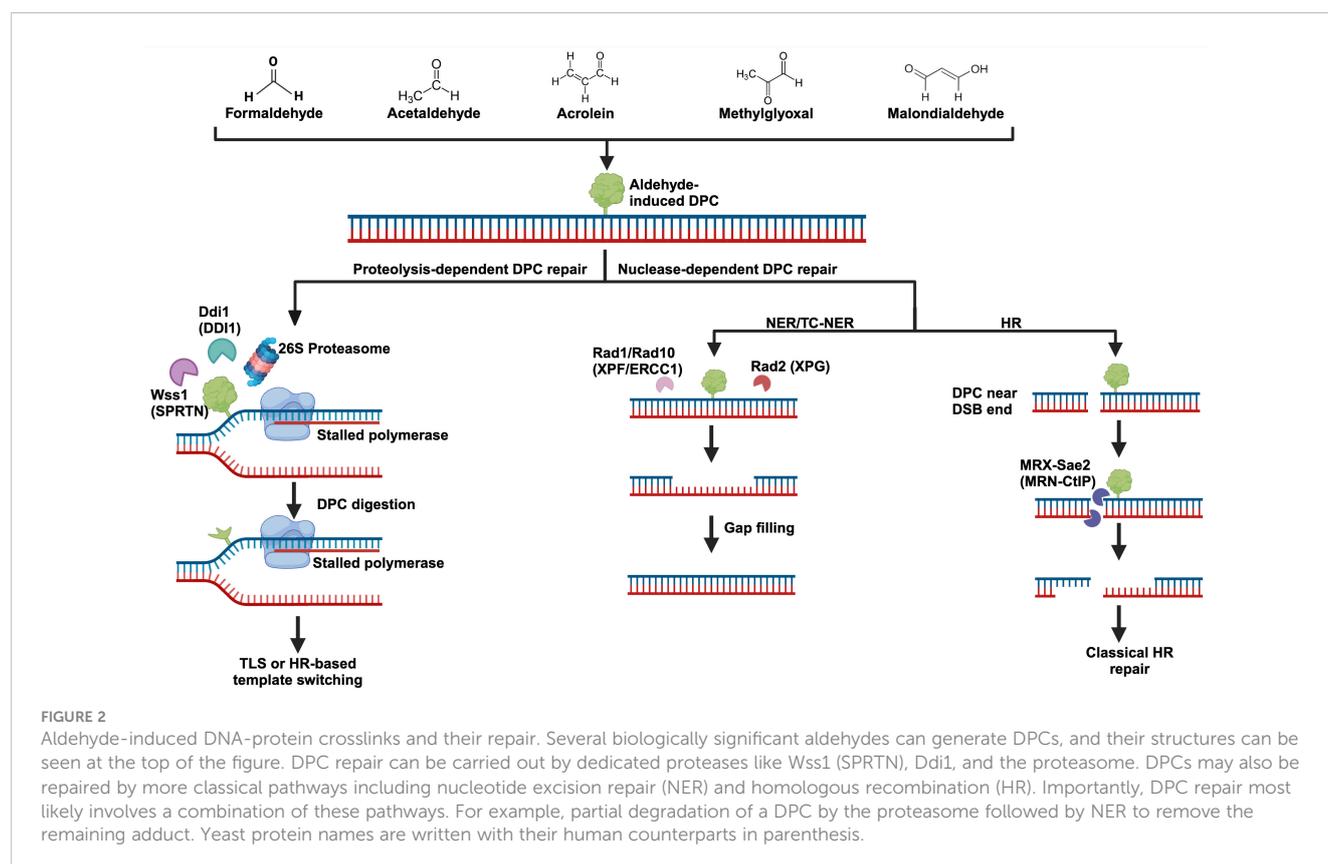
Yeast Wss1 and Ddi1

In 2014, Julian Stingle and Stefan Jentsch discovered and characterized the protease weak suppressor of Smt3 (Wss1) (64). This finding was significant as it was the first protease discovered to process DPCs regardless of the identity of the crosslinked protein in S-phase (Figure 2). Wss1 has a compact protease domain with an active site that contains few specificity-generating features (65). This general protease domain likely explains why Wss1 can process such a large variety of protein substrates. Wss1 becomes activated by DNA, and activation is most robust in the presence of single-stranded DNA (64). The requirement of DNA for Wss1 activation is an important regulator of function, as it limits Wss1's protease activity to exclusively DPCs. Once activated, Wss1 targets SUMOylated DPCs via interactions with its SUMO-binding domain (8, 9). Finally, Wss1 function is dependent on its

interaction with the chaperone-like enzyme Cdc48 which likely regulates Wss1 access to substrates (66). In the absence of this interaction, which is mediated by Ubx5, *in vivo* function of Wss1 is eliminated (66).

Enzymatic DPCs are common and can occur without any exogenous influence. For this reason, specific repair of commonly crosslinked DNA-interacting enzymes has evolved. Tyrosyl-DNA phosphodiesterase 1 and 2 (Tdp1 and Tdp2) have evolved to process topoisomerase crosslinks (67). Interestingly, Wss1 can process both enzymatic and non-enzymatic DPCs, and synthetic sickness has been observed in *Δtdp1Δwss1* yeast (64). This suggests that Wss1 may function in a secondary pathway to Tdp1 and Tdp2 in processing enzymatic DPCs. The *Candida albicans* Wss1 homolog, *CaWss1*, has been shown to protect against formaldehyde-induced DPCs (68). Finally, fission yeast with Wss1 or Wss2 deletions had an increased sensitivity to acetaldehyde and formaldehyde (41, 69). Together, these data suggest that Wss1's role in the processing of DPCs is highly conserved.

Ddi1 is an aspartic protease that can interact with ubiquitylated proteins and the proteasome via its Ub-like domain (UBL) (70, 71). Humans possess a Ddi1 ortholog, and the sequence identity between human and yeast Ddi1 is 55% (72, 73). In humans, it is thought to promote proteasome-dependent replication fork restart following replication stress by degrading the replication termination factor 2 (72, 74). A role for the human DDI1 in DPC repair has not yet been established. In yeast however, Ddi1 has been shown to aid in the processing of camptothecin and formaldehyde-induced DPCs (Figure 2) (71, 75). While Ddi1 has a similar function to



Wss1 in yeast, DPC processing by Wss1 seems to be the primary pathway. In the absence of the Ubx5-Cdc48-Wss1 axis, yeast become more dependent on Ddi1, and deletion of Wss1 results in an increase in Ddi1 protein levels (66, 71). Accordingly, $\Delta wss1\Delta ddi1$ yeast have been shown to be extremely sensitive to crosslinking agents (71). Deletion of either Wss1 or Ddi1 leads to increased acetaldehyde mutagenesis, and $\Delta wss1\Delta ddi1$ strains had a further increase in mutagenesis (42).

SPRTN (DVC1) and other proteases in higher eukaryotes

Shortly after the discovery of the role of Wss1 in yeast, a similar DPC repair mechanism was identified *in vitro* using *Xenopus laevis* egg extracts (76). In this study, Duxin et al. found that replication coupled proteolysis could prevent persistent replication fork stalling at a DPC (76). In 2015, reciprocal BLAST searches were performed to find proteins with sequence similarity to Wss1 (9). The human protein SPRTN was found to be a match, and a phylogenetic analysis based on sequence similarity and domain organization found that they shared N-terminal protease domains and C-terminal tails containing motifs for binding Cdc48/P97/VCP (9). Further, both SPRTN and Wss1 can bind ubiquitin or SUMO, and SPRTN contains a PIP box for interactions with PCNA, supporting the idea that SPRTN associates at the replication fork (9). Despite the obvious similarities of these two proteins, they share very little sequence homology, and it is most likely that their similarity is due to convergent evolution. As such, SPRTN is generally considered to be a functional homolog of Wss1. In 2016, SPRTN was finally shown experimentally to be the mammalian protease responsible for processing DPCs in cells (77, 78).

To prevent promiscuous cleavage of DNA bound proteins, SPRTN function is tightly regulated. This regulation is dependent on ubiquitylation, the type of DNA substrate, and autocleavage (78). Firstly, mono-ubiquitinated SPRTN is excluded from chromatin (78). Upon induction of DPCs by formaldehyde treatment, SPRTN gets deubiquitinated and it can now return to chromatin (78). SPRTN activity is also strongly regulated by DNA-structure specificity. More specifically, it can process crosslinks at or around disruption in the double-stranded DNA helix (79). Examples of these disruptions include short ssDNA gaps, double-stranded breaks, mismatched DNA bubbles, or ssDNA-dsDNA junctions (79). Notably, SPRTN is unable to process crosslinks on strictly dsDNA or ssDNA substrates. This specificity is determined by SPRTN containing both a basic DNA-binding region (BR) and a zinc-binding domain (ZBD) (79). The BR domain preferentially binds dsDNA while the ZBD domain binds ssDNA, and proper SPRTN association and activation on DNA is dependent on both regions binding (79). Finally, SPRTN undergoes autocleavage as a negative regulator to its activity (78).

SPRTN function is essential to genome maintenance in higher eukaryotes. *Sprtn*^{-/-} MEF cells have higher levels of γ H2Ax, 53BP1, and RAD51 foci, as well as increased CHK2 activation (77, 78, 80). These results illustrate the importance of SPRTN in processing DPCs to prevent replication stress and subsequent double-stranded

breaks (Figure 2). SPRTN expression is tightly regulated to S and G2 phases of the cell cycle where it can associate and travel with replication forks to prevent stalling at DPCs (81). Recently, it was shown that the helicase FANCI is essential in promoting SPRTN function by binding to ssDNA downstream of the DPC and unfolding the protein adduct (82). This unfolding exposes the DNA beneath the crosslink, enabling SPRTN cleavage of the DPC (82). This cleavage frees most of the crosslink from the DNA but leaves a short peptide tail which must be bypassed by translesion synthesis. During G1, SPRTN gets degraded by APC/Cdh1 (81). It is thought that outside of S and G2, the proteasome may dominate in processing DPCs (83).

Mutations within *SPRTN* have been shown to be the cause of Ruijs-Aalfs syndrome, a genetic disorder that leads to genomic instability, early-onset hepatocellular carcinoma, and progeria (10, 11). Unsurprisingly, lymphoblastoid cells derived from patients with Ruijs-Aalfs syndrome had increased levels of γ H2Ax staining compared to control cells following formaldehyde exposure (77). The characterization of Ruijs-Aalfs syndrome has underscored the importance of SPRTN in maintaining genomic stability and the effect of aldehyde-induced DPCs when they cannot be properly repaired.

Higher eukaryotes also contain the serine protease FAM111A which can repair TOP1 and PARP1-DNA crosslinks (84). Knockout of *FAM111A* resulted in only a slight sensitivity to formaldehyde, suggesting that either it only repairs a small portion of DPCs or that it works in a secondary pathway to SPRTN (84). In support of this idea, work from another group found that SPRTN, and not FAM111A, was responsible for processing PARP1-DNA crosslinks (85). Saha et al. concluded that these differing results were likely due to differential expression of FAM111A among different cell types where a lower SPRTN expression in one cell type may be compensated with higher FAM111A expression. Importantly, FAM111A expression is highest in late S phase (86).

Finally, the acidic repeat containing protein (ACRC/GCNA) was found in higher eukaryotes and was determined to contain a SprT domain like SPRTN and Wss1, and is similar to SPRTN based on phylogenetic analysis (83, 87, 88). ACRC/GCNA has been shown to interact with polySUMO chains and localizes to formaldehyde-induced damage foci (89). ACRC/GCNA is required for germ cell genomic stability in *Drosophila*, *C. elegans*, and zebrafish (90). ACRC/GCNA seems to be especially well suited to resolving topoisomerase II DPCs (90, 91). Despite these discoveries, the roles of ACRC/GCNA in DPC repair remains largely unexplored.

Proteasomal degradation of aldehyde-induced DPCs

The 26S eukaryotic proteasome is an extremely large ATP-dependent protease that is recruited to several cellular locations to degrade proteins (92). For efficient degradation, proteins must pass into the hollow 20S proteasomal core (7). Essential to the regulation and function of the proteasome is its ability to recognize

polyubiquitylated protein substrates using specialized ubiquitin binding sites (92). The proteasome is intimately related to the DPC-repair response (Figure 2). Inhibition of the 26S proteasome in A549 cells resulted in a significant reduction in formaldehyde-induced DPC repair (93), and sensitized cells to formaldehyde treatment (94). In *Xenopus laevis* egg extract, proteasome components were found to accumulate on replicating plasmids containing DPCs, and degradation of DPCs was dependent on polyubiquitylation by the E3 ligase TRAIP (95). Proteasomal and SPRTN mediated processing of DPCs are non-redundant pathways. When SPRTN was depleted in *Xenopus* extract, translesion synthesis across the DPC was impaired even when the proteasome was unchanged, suggesting that SPRTN was the preferred pathway to process the crosslink (95). Given that DPCs can not readily enter the proteasomal 20S core, there are a few models for how it may contribute to DPC proteolysis. Firstly, it is possible that the proteasome partially degrades a DPC into a smaller peptide crosslink which may be removed by nucleotide excision repair (7). It is also likely that proteasomal degradation depends on specialized proteases such as SPRTN or TDP1 to cleave the bulk of a DPC off of DNA, with the now freed protein being a more manageable substrate for the proteasome (7). Because aldehydes can generate a heterogenous class of DPCs, it is most likely that coordination between different repair pathways are required to efficiently resolve aldehyde-induced DPCs (Figure 2).

Nuclease-dependent DNA-protein crosslink repair

Global nucleotide excision repair

Nucleotide excision repair (NER) can recognize helix distorting lesions and involves the removal of the lesion and surrounding nucleotides, leaving a single-stranded DNA gap (Figure 2). Early studies in bacteria showed that NER was involved in processing formaldehyde-induced DNA damage as *uvrA* mutants showed an increased sensitivity (96). Studies in yeast have shown that NER defective strains (*Δrad1* and *Δrad4*) were highly sensitive to acute formaldehyde treatment (60 mM for 15 minutes), but were not especially sensitive to chronic formaldehyde treatment (1.5 mM for 48 hours) (97). Similarly, deletion of the NER pathway in fission yeast led to increased sensitivity of cultures to formaldehyde and acetaldehyde (41, 69). These data were echoed by our recent study demonstrating a role for NER in the repair of acetaldehyde-induced lesions (42). These results suggest that NER is an important 'first line of defense' to aldehyde-induced DPCs, but if given adequate time, other repair pathways can compensate for loss of NER. In human cells, NER was able to repair partially digested DPCs but was unable to repair certain full sized DPCs (98). In fact, mammalian NER is limited to relatively small 8-10 kDa (~70-90 amino acids) crosslinks (99). For this reason, it is likely that NER works cooperatively with the DPC proteases described earlier in this review, where the DPCs are cleaved and the remaining peptide tail is then removed by NER in a subsequent step.

Transcription-coupled repair of aldehyde-induced DPCs

Recently, it was shown that transcription-coupled repair is involved in processing formaldehyde-induced DPCs. Cells were treated with 600 μM formaldehyde for one-hour followed by zero- and four-hour recovery times. After four hours, there was a marked reduction in DPC-seq derived sequencing reads within gene bodies (100). This result suggests that DPCs within transcribed regions are being efficiently repaired. To further support the hypothesis that transcription promotes DPC repair, the RNA polymerase II transcribed genes *PKM*, *TK1*, *AFMID*, and *TCF7L2* were analyzed in cells treated with an RNA polymerase II inhibitor. Upon inhibition of RNA polymerase II, there was a reduction in DPC processing within the targeted gene bodies, suggesting that crosslink repair was dependent on active transcription (100). Immunoprecipitation of elongating RNA polymerase II in the presence of formaldehyde followed by mass spectrometry showed that RNA polymerase II was associated with the transcription-coupled nucleotide excision repair factors CSA, CSB, UVSSA, and TFIIH (100). In the absence of CSB, cells were unable to recover RNA synthesis, suggesting that transcription-coupled repair is required for rescuing DPC-stalled transcription machinery (100). Overall, this work establishes a novel role for transcription-coupled nucleotide excision repair in the processing of formaldehyde-induced DPCs. Whether other aldehyde-induced DPCs can be repaired via a similar transcription-coupled mechanism remains to be seen. In the future, it would be interesting to test if proteasomal or SPRTN function is required for the DPC-processing role of transcription-coupled nucleotide excision repair, and whether this pathway is under the same DPC size restrictions as global nucleotide excision repair.

Homologous recombination in DPC repair

Early studies in bacteria showed that *recA* mutants were sensitive to formaldehyde treatment, establishing a role for homologous recombination (HR) in the repair of aldehyde-induced damage (96). Interestingly, HR in yeast was shown to aid DPC processing following chronic, low dose exposure to formaldehyde while not contributing significantly to the repair of acute, high dose treatment (97). HR was also shown to be required for tolerance to acetaldehyde and formaldehyde in fission yeast (41, 69). HR has been implicated in the processing of DPCs in higher eukaryotes as well. A study in chicken DT40 cells that were defective in the BRCA/FANC pathway and HR were hypersensitive to low doses of formaldehyde and acetaldehyde, but were not sensitive to acrolein, crotonaldehyde, glyoxal, or methylglyoxal (101). There are a few ways in which HR may contribute to DPC repair. Firstly, HR may assist in the processing of DPCs in close proximity to double-stranded breaks with the help of the MRE11-RAD50-NBS1 (Xrs2 in yeast) complex (Figure 2) (102). The MRN/MRX complex has been shown to remove Spo11-generated crosslinks from the ends of a double-stranded break in yeast and in mouse testis extract,

confirming that DPCs can be efficiently processed from double-stranded break ends (103). Finally, *Xenopus* egg extracts have been used to show that MRN, CtIP, and BRCA1 work together in the removal of TOP2-DPCs at double-strand break ends (104), and very similar results have been observed in mammalian cells (105, 106). HR may also assist in DPC tolerance via the template switching pathway. When forks stall at a DPC, the sliding clamp PCNA can be polyubiquitinated which leads to template switching initiation (107). In this model, the stalled polymerase uses the nascent sister strand as a template, allowing for extension beyond the adduct. The resulting structure can then be resolved via a classic HR dissolution reaction (107, 108). It is important to note that template switching is a DNA damage tolerance pathway, and following bypass the DPC will persist and must be repaired later. Given that aldehydes generate DPCs which can lead to double-stranded breaks if they persistently block replication forks, it is likely that homologous recombination plays a significant role in processing DPCs both at break ends and via template switching to restore chromosomal integrity.

In summary, efficient nucleotide excision repair depends on the recruitment of various bulky protein complexes. As such, DPCs can only be repaired by NER if they do not impede the recruitment of the NER factors. It is possible that NER works in concert with DPC proteases or the proteasome, wherein the proteases partially process the DPC, enabling NER to excise the remaining peptide bound DNA. Importantly, DPC processing by NER can likely occur throughout the cell cycle. HR does not face the same size constraints as NER, but generally can only resolve crosslinks near double-stranded breaks using the nuclease activity of the MRN complex. Repair of DPCs by HR requires the presence of a double-strand break and homologous DNA, likely limiting this repair to S/G2 phases of the cell cycle. Given that aldehydes can crosslink a variety of proteins to DNA, it is likely that an elaborate combination of protease- and nuclease-based repair pathways are required for efficient processing of aldehyde-induced DPCs. A graphical summary of protease- and nuclease-dependent DPC repair pathways can be seen in Figure 2.

DNA-protein crosslink-induced genotoxicity in cancers

ALDH2 is a key enzyme that removes acetaldehyde and various other endogenous aldehydes from the cell. Defects in ALDH2 have been linked to increased alcohol-associated cancers (109). Interestingly, low ALDH2 levels are also predictive for poor survival in lung and liver cancer patients, likely due to elevated cellular aldehydes. Inhibition of ALDH2 has been shown to cause an increase in DPCs in cancer cell lines (110). As such, DPC-induced genome instability from endogenous aldehydes is common in cancers with defective aldehyde clearance pathways.

Germline mutations in SPRTN have been seen in patients with Ruijjs-Aalfs syndrome which is characterized by increased risk of hepatocellular carcinoma (10, 111). SPRTN hypomorphic mice were also shown to have elevated levels of DNA damage, DPCs, and spontaneous development of liver tumors (112). These studies

demonstrate that the liver is the primary site of carcinogenesis in patients with SPRTN deficiencies. However, why these patients develop liver cancers has remained unclear. We hypothesize that lipid peroxidation by cytochrome P450 enzymes in the liver lead to elevated cellular aldehydes and DPCs.

SPRTN was shown to be involved in the unfolded protein response (UPR) and interacts with GRP78 in HepG2 liver cancer cells (111). SPRTN depleted cells exhibit increased sensitivity to ER stress and the UPR. Increased ER stress was shown to lead to DNA damage that was epistatic to DNA damage induced by SPRTN deficiency (111). It could thus be possible that increased ER stress leads to DNA damage that is repaired via SPRTN-induced proteolysis. ER stress and the UPR also results in fatty acid accumulation in cells (113). Cytochrome P450 enzymes have been shown to directly lead to oxidation of lipids forming lipid hydroperoxides (114). Moreover, activity of the cytochrome P450 enzymes leads to elevated free radical generation that can further lead to lipid peroxidation (115). Lipid peroxidation in liver cells is expected to generate reactive aldehydes like malondialdehyde, 3-hydroxynonenal, acrolein, and acetaldehyde (54, 55, 116). It is likely that the reactive aldehydes produced upon cytochrome P450 mediated lipid peroxidation in hepatocytes induce DPCs which are targets for SPRTN processing. In the absence of SPRTN, these persistent DPCs lead to DNA damage and mutagenesis, likely contributing to the development of hepatocellular carcinoma in patients with Ruijjs-Aalfs syndrome.

Alternately, livers from *SPRTN* hypomorphic mice demonstrate an accumulation of topoisomerase I cleavage complexes (112). Such adducts also lead to acute DNA damage and genome instability culminating in carcinogenesis. Interestingly, lipid peroxidation products including 4-hydroxynonenal have been shown to modify TOP1 active site residue C630 leading to TOP1 entrapment on DNA (117). In the absence of SPRTN, TOP1-DPCs would accumulate leading to genome instability.

These studies demonstrate the importance of SPRTN in mitigating DNA damage from aldehyde-induced DPCs in liver cancers.

Conclusions

Aldehydes are present ubiquitously in the environment, and several are generated endogenously by cells. These aldehydes have been shown to generate a variety of DNA damage including base adducts, interstrand and intrastrand DNA crosslinks, and DPCs. DPCs can have a variety of deleterious effects on genome integrity; inhibiting DNA synthesis, transcription, and chromatin remodeling. As such mechanisms for repair of DPCs are evolutionarily conserved. These include proteolysis-dependent repair with the specialized proteases Wss1/SPRTN and Ddi1, or more general proteolysis using the proteasome. Cells can also employ nuclease-dependent repair of DPCs which include global and transcription-coupled nucleotide excision repair, as well as homologous recombination. Defects in any of these repair pathways greatly sensitize cells to aldehyde-induced DPCs, and evidence of

this can be seen in Rujis-Aalfs syndrome, a cancer predisposition disorder that results from hypomorphic mutations in *SPRTN*.

Research on reactive aldehydes has been challenged by several technical limitations. Firstly, many of these aldehydes are highly volatile so maintaining consistent experimental conditions can be challenging. DPC research has long been limited by the ability to detect DPCs. Relatively new detection methods such as rapid approach to DNA adduct discovery (RADAR) and purification of x-linked proteins (PxP) promise to increase reliability and sensitivity of DPC detection (26, 118, 119). Finally, these reactive aldehydes can generate a variety of adducts, making it difficult to determine the impact of DPCs specifically. While the field of DPC repair has seen significant advances in recent years, several questions remain. Much of the work done on aldehyde-induced DPCs has used formaldehyde. It is still unknown if the damage generated by other aldehydes are processed by *SPRTN* in similar ways. While many studies have looked at double-stranded breaks following the loss of DPC repair, the mutagenic consequences of losing DPC repair are relatively unknown. Finally, it is not well understood why reactive aldehydes and *SPRTN* function are drivers for liver cancer specifically. It remains to be seen if aldehyde levels are higher in *SPRTN* deficient cells or if *SPRTN* deficiency in conjunction with non-alcoholic fatty liver disease cause increased DPC formation and hepatocellular carcinoma development.

Author contributions

TB: Writing – original draft, Writing – review & editing. NS: Writing – original draft, Writing – review & editing.

References

- Vijayraghavan S, Saini N. Aldehyde-associated mutagenesis—Current state of knowledge. *Chem Res Toxicol.* (2023) 36:983–1001. doi: 10.1021/acs.chemrestox.3c00045
- Nakano T, Ouchi R, Kawazoe J, Pack SP, Makino K, Ide H. T7 RNA polymerases backed up by covalently trapped proteins catalyze highly error prone transcription *. *J Biol Chem.* (2012) 287:6562–72. doi: 10.1074/jbc.M111.318410
- Nakano T, Miyamoto-Matsubara M, Shoukamy MI, Salem AMH, Pil Pack S, Ishimi Y, et al. Translocation and stability of replicative DNA helicases upon encountering DNA-protein cross-links *. *J Biol Chem.* (2013) 288:4649–58. doi: 10.1074/jbc.M112.419358
- Fu YV, Yardimci H, Long DT, Vinh Ho T, Guainazzi A, Bermudez VP, et al. Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell.* (2011) 146:931–41. doi: 10.1016/j.cell.2011.07.045
- Yeo JE, Wickramaratne S, Khatwani S, Wang YC, Vervacke J, Distefano MD, et al. Synthesis of site-specific DNA–protein conjugates and their effects on DNA replication. *ACS Chem Biol.* (2014) 9:1860–8. doi: 10.1021/cb5001795
- Ji S, Shao H, Han Q, Seiler CL, Tretyakova NY. Reversible DNA–protein cross-linking at epigenetic DNA marks. *Angew Chem Int Ed Engl.* (2017) 56:14130–4. doi: 10.1002/anie.201708286
- Pachva MC, Kisselev AF, Matkarimov BT, Saparbaev M, Groisman R. DNA-histone cross-links: formation and repair. *Front Cell Dev Biol.* (2020) 8:607045. doi: 10.3389/fcell.2020.607045
- Stingle J, Jentsch S. DNA–protein crosslink repair. *Nat Rev Mol Cell Biol.* (2015) 16:455–60. doi: 10.1038/nrm4015
- Stingle J, Habermann B, Jentsch S. DNA–protein crosslink repair: proteases as DNA repair enzymes. *Trends Biochem Sci.* (2015) 40:67–71. doi: 10.1016/j.tibs.2014.10.012
- Lessel D, Vaz B, Halder S, Lockhart PJ, Marinovic-Terzic I, Lopez-Mosqueda J, et al. Mutations in *SPRTN* cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nat Genet.* (2014) 46:1239–44. doi: 10.1038/ng.3103
- Ruijs MW. Atypical progeroid syndrome: An unknown helicase gene defect? *Am J Med Genet Part A - Wiley Online Library.* (2003) 116A:295–9. doi: 10.1002/ajmg.a.10730
- Jensen RP, Luo W, Pankow JF, Strongin RM, Peyton DH. Hidden formaldehyde in E-cigarette aerosols. *N Engl J Med.* (2015) 372:392–4. doi: 10.1056/NEJMc1413069
- Trocho C, Pardo R, Rafecas I, Virgili J, Remesar X, Fernández-López JA, et al. Formaldehyde derived from dietary aspartame binds to tissue components *in vivo*. *Life Sci.* (1998) 63:337–49. doi: 10.1016/S0024-3205(98)00282-3
- Shi Y, Lan F, Matsun C, Mulligan P, Whetstone JR, Cole PA, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell.* (2004) 119:941–53. doi: 10.1016/j.cell.2004.12.012
- Tsukada Yi, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, et al. Histone demethylation by a family of JmjC domain-containing proteins. *Nature.* (2006) 439:811–6. doi: 10.1038/nature04433
- Trewick SC, Henshaw TF, Hausinger RP, Lindahl T, Sedgwick B. Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature.* (2002) 419:174–8. doi: 10.1038/nature00908
- Shen L, Song CX, He C, Zhang Y. Mechanism and function of oxidative reversal of DNA and RNA methylation. *Annu Rev Biochem.* (2014) 83:585–614. doi: 10.1146/annurev-biochem-060713-035513
- Burgos-Barragan G, Wit N, Meiser J, Dingler FA, Pietzke M, Mulderrig L, et al. Mammals divert endogenous genotoxic formaldehyde into one-carbon metabolism. *Nature.* (2017) 548:549–54. doi: 10.1038/nature23481
- Sauerbier W. Reaction of formaldehyde with deoxyribonucleic acid in phage T1. *Nature.* (1960) 188:327–9. doi: 10.1038/188327a0
- Fraenkel-Conrat H, Olcott Hs. REACTION OF FORMALDEHYDE WITH PROTEINS: VI. CROSS-LINKING OF AMINO GROUPS WITH PHENOL, IMIDAZOLE, OR INDOLE GROUPS. *J Biol Chem.* (1948) 174:827–43. doi: 10.1016/S0021-9258(18)57292-6

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The authors of this review are supported by funding from the National Institute of General Medical Sciences (R35GM151021 to NS) and Cellular, Biochemical and Molecular Sciences training grant (5T32GM132055 to TB).

Acknowledgments

We would like to thank Dr. Sriram Vijayraghavan and Yashvi Patel for their critical reading of the manuscript and helpful feedback.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

21. Lu K, Ye W, Zhou L, Collins LB, Chen X, Gold A, et al. Structural characterization of formaldehyde-induced cross-links between amino acids and deoxynucleosides and their oligomers. *J Am Chem Soc*. (2010) 132:3388–99. doi: 10.1021/ja908282f
22. Metz B, Kersten GFA, Hoogerhout P, Brugghe HF, Timmermans HAM, Jong AD, et al. Identification of formaldehyde-induced modifications in proteins: REACTIONS WITH MODEL PEPTIDES *. *J Biol Chem*. (2004) 279:6235–43. doi: 10.1074/jbc.M310752200
23. Schmiedeberg L, Skene P, Deaton A, Bird A. A temporal threshold for formaldehyde crosslinking and fixation. *PLoS One*. (2009) 4:e4636. doi: 10.1371/journal.pone.0004636
24. Siomian Y, Simonov VV, Poverenny AM. The reaction of formaldehyde with deoxynucleotides and DNA in the presence of amino acids and lysine-rich histone. *Biochim Biophys Acta BBA - Nucleic Acids Protein Synth*. (1973) 331:27–32. doi: 10.1016/0005-2787(73)90415-2
25. O'Connor PM, Fox BW. Isolation and characterization of proteins cross-linked to DNA by the antitumor agent methylene dimethanesulfonate and its hydrolytic product formaldehyde. *J Biol Chem*. (1989) 264:6391–7. doi: 10.1016/S0021-9258(18)83362-2
26. Weickert P, Li HY, Götz MJ, Dürauer S, Yaneva D, Zhao S, et al. SPRTN patient variants cause global-genome DNA-protein crosslink repair defects. *Nat Commun*. (2023) 14:352. doi: 10.1038/s41467-023-35988-1
27. Kerns WD, Pavkov KL, Donofrio DJ, Gralla EJ, Swenberg JA. Carcinogenicity of formaldehyde in rats and mice after long-term inhalation exposure. *Cancer Res*. (1983) 43:4382–92.
28. Monticello TM, Swenberg JA, Gross EA, Leininger JR, Kimbell JS, Seilkop S, et al. Correlation of regional and nonlinear formaldehyde-induced nasal cancer with proliferating populations of cells. *Cancer Res*. (1996) 56:1012–22.
29. Swenberg JA, Kerns WD, Mitchell RI, Gralla EJ, Pavkov KL. Induction of squamous cell carcinomas of the rat nasal cavity by inhalation exposure to formaldehyde vapor. *Cancer Res*. (1980) 40:3398–402.
30. Jackson B, Brocker C, Thompson DC, Black W, Vasiliou K, Nebert DW, et al. Update on the aldehyde dehydrogenase gene (ALDH) superfamily. *Hum Genomics*. (2011) 5(4):283–303. doi: 10.1186/1479-7364-5-4-283
31. Pontel LB, Rosado IV, Burgos-Barragan G, Garaycochea JI, Yu R, Arends MJ, et al. Endogenous formaldehyde is a hematopoietic stem cell genotoxin and metabolic carcinogen. *Mol Cell*. (2015) 60:177–88. doi: 10.1016/j.molcel.2015.08.020
32. Dingler FA, Wang M, Mu A, Millington CL, Oberbeck N, Watcham S, et al. Two aldehyde clearance systems are essential to prevent lethal formaldehyde accumulation in mice and humans. *Mol Cell*. (2020) 80:996–1012.e9. doi: 10.1016/j.molcel.2020.10.012
33. Lai Y, Yu R, Hartwell HJ, Moeller BC, Bodnar WM, Swenberg JA. Measurement of endogenous versus exogenous formaldehyde-induced DNA-protein crosslinks in animal tissues by stable isotope labeling and ultrasensitive mass spectrometry. *Cancer Res*. (2016) 76:2652–61. doi: 10.1158/0008-5472.CAN-15-2527
34. Nakamura J, Nakamura M. DNA-protein crosslink formation by endogenous aldehydes and AP sites. *DNA Repair*. (2020) 88:102806. doi: 10.1016/j.dnarep.2020.102806
35. Casanova M, Morgan KT, Steinhagen WH, Everitt JJ, Popp JA, Heck Hd'A. Covalent binding of inhaled formaldehyde to DNA in the respiratory tract of rhesus monkeys: Pharmacokinetics, rat-to-monkey interspecies scaling, and extrapolation to man. *Fundam Appl Toxicol*. (1991) 17:409–28. doi: 10.1016/0272-0590(91)90230-2
36. Shaham J, Bomstein Y, Gurvich R, Rashkovsky M, Kaufman Z. DNA-protein crosslinks and p53 protein expression in relation to occupational exposure to formaldehyde. *Occup Environ Med*. (2003) 60:403–9. doi: 10.1136/oem.60.6.403
37. Vaca CE, Fang JL, Schweda EKH. Studies of the reaction of acetaldehyde with deoxynucleosides. *Chem Biol Interact*. (1995) 98:51–67. doi: 10.1016/0009-2797(95)03632-V
38. Wang M, McIntee EJ, Cheng G, Shi Y, Villalta PW, Hecht SS. Identification of DNA adducts of acetaldehyde. *Chem Res Toxicol*. (2000) 13:1149–57. doi: 10.1021/tx000118t
39. Kuykendall JR, Bogdanffy MS. Reaction kinetics of DNA–histone crosslinking by vinyl acetate and acetaldehyde. *Carcinogenesis*. (1992) 13:2095–100. doi: 10.1093/carcin/13.11.2095
40. Kuykendall JR, Bogdanffy MS. Formation and stability of acetaldehyde-induced crosslinks between poly-lysine and poly-deoxyguanosine. *Mutat Res Mol Mech Mutagen*. (1994) 311:49–56. doi: 10.1016/0027-5107(94)90072-8
41. Noguchi C, Grothusen G, Anandarajan V, Martínez-Lage G, Terlecky D, Corzo K, et al. Genetic controls of DNA damage avoidance in response to acetaldehyde in fission yeast. *Cell Cycle*. (2017) 16:45–58. doi: 10.1080/15384101.2016.1237326
42. Porcher L, Vijayraghavan S, McCollum J, Mieczkowski PA, Saini N. Multiple DNA repair pathways prevent acetaldehyde-induced mutagenesis in yeast. (2024). 2024.01.07.574575. doi: 10.1101/2024.01.07.574575
43. Stevens JF, Maier CS. Acrolein: Sources, metabolism, and biomolecular interactions relevant to human health and disease. *Mol Nutr Food Res*. (2008) 52:7–25. doi: 10.1002/mnfr.200700412
44. Stone MP, Cho YJ, Huang H, Hye-Young K, Kozekov ID, Kozekova A, et al. Interstrand DNA cross-links induced by α,β -unsaturated aldehydes derived from lipid peroxidation and environmental sources. *Acc Chem Res*. (2008) 41:793–804. doi: 10.1021/ar700246x
45. Minko IG, Kozekov ID, Kozekova A, Harris TM, Rizzo CJ, Lloyd RS. Mutagenic potential of DNA–peptide crosslinks mediated by acrolein-derived DNA adducts. *Mutat Res*. (2008) 637:161–72. doi: 10.1016/j.mrfmmm.2007.08.001
46. Zarkovic N, Cipak A, Jaganjac M, Borovic S, Zarkovic K. Pathophysiological relevance of aldehydic protein modifications. *J Proteomics*. (2013) 92:239–47. doi: 10.1016/j.jprot.2013.02.004
47. Moghe A, Ghare S, Lamoreau B, Mohammad M, Barve S, McClain C, et al. Molecular mechanisms of acrolein toxicity: relevance to human disease. *Toxicol Sci*. (2015) 143:242–55. doi: 10.1093/toxsci/kfu233
48. Petrova KV, Millsap AD, Stec DF, Rizzo CJ. Characterization of the deoxyguanosine–lysine cross-link of methylglyoxal. *Chem Res Toxicol*. (2014) 27:1019–29. doi: 10.1021/tx500068v
49. Zeng J, Davies MJ. Evidence for the formation of adducts and S-(Carboxymethyl)cysteine on reaction of α -dicarbonyl compounds with thiol groups on amino acids, peptides, and proteins. *Chem Res Toxicol*. (2005) 18:1232–41. doi: 10.1021/tx050074u
50. Thornalley PJ. Protein and nucleotide damage by glyoxal and methylglyoxal in physiological systems - role in ageing and disease. *Drug Metabol Drug Interact*. (2008) 23:125–50. doi: 10.1515/DMDI.2008.23.1-2.125
51. Ahmed N, Babaei-Jadidi R, Howell SK, Beisswenger PJ, Thornalley PJ. Degradation products of proteins damaged by glycation, oxidation and nitration in clinical type 1 diabetes. *Diabetologia*. (2005) 48:1590–603. doi: 10.1007/s00125-005-1810-7
52. Murata-Kamiya N, Kamiya H. Methylglyoxal, an endogenous aldehyde, crosslinks DNA polymerase and the substrate DNA. *Nucleic Acids Res*. (2001) 29:3433–8. doi: 10.1093/nar/29.16.3433
53. Rio DD, Stewart AJ, Pellegrini N. A review of recent studies on malondialdehyde as toxic molecule and biological marker of oxidative stress. *Nutr Metab Cardiovasc Dis*. (2005) 15:316–28. doi: 10.1016/j.numecd.2005.05.003
54. Martín-Fernández M, Arroyo V, Carnicero C, Sigüenza R, Busta R, Mora N, et al. Role of oxidative stress and lipid peroxidation in the pathophysiology of NAFLD. *Antioxidants*. (2022) 11:2217. doi: 10.3390/antiox1112217
55. Pizzimenti S, Ciamporcerio E, Daga M, Pettazzoni P, Arcaro A, Cetrangolo G, et al. Interaction of aldehydes derived from lipid peroxidation and membrane proteins. *Front Physiol*. (2013) 4:242. doi: 10.3389/fphys.2013.00242
56. Esterbauer H, Schaur RJ, Zollner H. Chemistry and biochemistry of 4-hydroxynonenal, malonaldehyde and related aldehydes. *Free Radic Biol Med*. (1991) 11:81–128. doi: 10.1016/0891-5849(91)90192-6
57. Marnett LJ. Lipid peroxidation—DNA damage by malondialdehyde. *Mutat Res Mol Mech Mutagen*. (1999) 424:83–95. doi: 10.1016/S0027-5107(99)00010-X
58. Chaudhary AK, Reddy GR, Blair IA, Marnett LJ. Characterization of an N 6 -oxopropenyl'-deoxyadenosine adduct in malondialdehyde-modified DNA using liquid chromatography/electrospray ionization tandem mass spectrometry. *Carcinogenesis*. (1996) 17:1167–70. doi: 10.1093/carcin/17.5.1167
59. Voitkun V, Zhitkovich A. Analysis of formaldehyde-induced DNA-protein crosslinking activity of malondialdehyde in vitro. *Mutat Res*. (1999) 424:97–106. doi: 10.1016/s0027-5107(99)00011-1
60. Kuo HK, Griffith JD, Kreuzer KN. 5-Azacytidine induced methyltransferase-DNA adducts block DNA replication in vivo. *Cancer Res*. (2007) 67:8248–54. doi: 10.1158/0008-5472.CAN-07-1038
61. Chválová K, Brabec V, Kašpárková J. Mechanism of the formation of DNA–protein cross-links by antitumor cisplatin. *Nucleic Acids Res*. (2007) 35:1812–21. doi: 10.1093/nar/gkm032
62. Ji S, Fu I, Naldiga S, Shao H, Basu AK, Brody S, et al. 5-Formylcytosine mediated DNA–protein cross-links block DNA replication and induce mutations in human cells. *Nucleic Acids Res*. (2018) 46:6455–69. doi: 10.1093/nar/gky444
63. Merk O, Speit G. Significance of formaldehyde-induced DNA-protein crosslinks for mutagenesis. *Environ Mol Mutagen*. (1998) 32:260–8. doi: 10.1002/(sici)1098-2280(1998)32:3<260::aid-em9>3.0.co;2-m
64. Stingle J, Schwarz MS, Bloemeke N, Wolf PG, Jentsch S. A DNA-dependent protease involved in DNA-protein crosslink repair. *Cell*. (2014) 158:327–38. doi: 10.1016/j.cell.2014.04.053
65. Yang X, Li Y, Gao Z, Li Z, Xu J, Wang W, et al. Structural analysis of Wss1 protein from *Saccharomyces cerevisiae*. *Sci Rep*. (2017) 7:8270. doi: 10.1038/s41598-017-08834-w
66. Noireterre A, Serbyn N, Bagdiul I, Stutz F. Ubx5-Cdc48 assists the protease Wss1 at DNA-protein crosslink sites in yeast. *EMBO J*. (2023) 42:e113609. doi: 10.15252/embj.2023113609
67. Pommier Y, Huang S yin N, Gao R, Das BB, Murai J, Marchand C. Tyrosyl-DNA-phosphodiesterases (TDP1 and TDP2). *DNA Repair*. (2014) 19:114–29. doi: 10.1016/j.dnarep.2014.03.020
68. Homchan A, Sukted J, Mongkolsuk S, Jeruzalmi D, Matangkasombut O, Pakotitrapha D. Wss1 homolog from *Candida albicans* and its role in DNA–protein crosslink tolerance. *Mol Microbiol*. (2020) 114:409–22. doi: 10.1111/mmi.14518

69. Anandarajan V, Noguchi C, Oleksak J, Grothusen G, Terlecky D, Noguchi E. Genetic investigation of formaldehyde-induced DNA damage response in *Schizosaccharomyces pombe*. *Curr Genet.* (2020) 66:593–605. doi: 10.1007/s00294-020-01057-z
70. Trempe JF, Šašková KG, Sívá M, Ratcliffe CDH, Veverka V, Hoegl A, et al. Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family. *Sci Rep.* (2016) 6:33671. doi: 10.1038/srep33671
71. Serbyn N, Noireterre A, Bagdiul I, Plank M, Michel AH, Loewith R., et al. The aspartic protease ddi1 contributes to DNA-protein crosslink repair in yeast. *Mol Cell.* (2020) 77:1066–1079.e9. doi: 10.1016/j.molcel.2019.12.007
72. Ruggiano A, Ramadan K. DNA-protein crosslink proteases in genome stability. *Commun Biol.* (2021) 4:1–11. doi: 10.1038/s42003-020-01539-3
73. Sirkis R, Gerst JE, Fass D. Ddi1, a eukaryotic protein with the retroviral protease fold. *J Mol Biol.* (2006) 364:376–87. doi: 10.1016/j.jmb.2006.08.086
74. Kottemann MC, Conti BA, Lach FP, Smogorzewska A. Removal of RTF2 from stalled replisomes promotes maintenance of genome integrity. *Mol Cell.* (2018) 69:24–35.e5. doi: 10.1016/j.molcel.2017.11.035
75. Svoboda M, Konvalinka J, Trempe JF, Grantz Saskova K. The yeast proteases Ddi1 and Wss1 are both involved in the DNA replication stress response. *DNA Repair.* (2019) 80:45–51. doi: 10.1016/j.dnarep.2019.06.008
76. Duxin JP, Dewar JM, Yardimci H, Walter JC. Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell.* (2014) 159:346–57. doi: 10.1016/j.cell.2014.09.024
77. Lopez-Mosqueda J, Maddi K, Prgomet S, Kalayil S, Marinovic-Terzic I, Terzic J, et al. SPTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks. *eLife.* (2016) 5:e21491. doi: 10.7554/eLife.21491
78. Stinglele J, Bellelli R, Alte F, Hewitt G, Sarek G, Maslen SL, et al. Mechanism and regulation of DNA-protein crosslink repair by the DNA-dependent metalloprotease SPTN. *Mol Cell.* (2016) 64:688–703. doi: 10.1016/j.molcel.2016.09.031
79. Reinking HK, Kang HS, Götz MJ, Li HY, Kieser A, Zhao S, et al. DNA structure-specific cleavage of DNA-protein crosslinks by the SPTN protease. *Mol Cell.* (2020) 80:102–113.e6. doi: 10.1016/j.molcel.2020.08.003
80. Vaz B, Popovic M, Newman JA, Fielden J, Aitkenhead H, Halder S, et al. Metalloprotease SPTN/DVC1 orchestrates replication-coupled DNA-protein crosslink repair. *Mol Cell.* (2016) 64:704–19. doi: 10.1016/j.molcel.2016.09.032
81. Mosbech A, Gibbs-Seymour I, Kagias K, Thorslund T, Beli P, Povlsen L, et al. DVC1 (Clorf124) is a DNA damage-targeting p97 adaptor that promotes ubiquitin-dependent responses to replication blocks. *Nat Struct Mol Biol.* (2012) 19:1084–92. doi: 10.1038/nsmb.2395
82. Yaneva D, Sparks JL, Donsbach M, Zhao S, Weickert P, Bezalel-Buch R, et al. The FANCD1 helicase unfolds DNA-protein crosslinks to promote their repair. *Mol Cell.* (2023) 83:43–56.e10. doi: 10.1016/j.molcel.2022.12.005
83. Fielden J, Ruggiano A, Popović M, Ramadan K. DNA protein crosslink proteolysis repair: From yeast to premature ageing and cancer in humans. *DNA Repair.* (2018) 71:198–204. doi: 10.1016/j.dnarep.2018.08.025
84. Kojima Y, Machida Y, Palani S, Caulfield TR, Radisky ES, Kaufmann SH, et al. FAM111A protects replication forks from protein obstacles via its trypsin-like domain. *Nat Commun.* (2020) 11:1318. doi: 10.1038/s41467-020-15170-7
85. Saha LK, Murai Y, Saha S, Jo U, Tsuda M, Takeda S, et al. Replication-dependent cytotoxicity and Spartan-mediated repair of trapped PARP1-DNA complexes. *Nucleic Acids Res.* (2021) 49:10493–506. doi: 10.1093/nar/gkab777
86. Fine DA, Rozenblatt-Rosen O, Padi M, Korkhin A, James RL, Adelman G, et al. Identification of FAM111A as an SV40 host range restriction and adenovirus helper factor. *PLoS Pathog.* (2012) 8(10):e1002949. doi: 10.1371/journal.ppat.1002949
87. Carmell MA, Dokshin GA, Skaletsky H, Hu YC, van Wolfswinkel JC, Igarashi KJ, et al. A widely employed germ cell marker is an ancient disordered protein with reproductive functions in diverse eukaryotes. *eLife.* (2016) 5:e19993. doi: 10.7554/eLife.19993
88. Delabaere L, Orsi GA, Sapey-Triomphe L, Horard B, Couble P, Loppin B. The spartan ortholog maternal haploid is required for paternal chromosome integrity in the *Drosophila* zygote. *Curr Biol.* (2014) 24:2281–7. doi: 10.1016/j.cub.2014.08.010
89. Borgermann N, Ackermann L, Schwertman P, Hendriks IA, Thijssen K, Liu JCY, et al. SUMOylation promotes protective responses to DNA-protein crosslinks. *EMBO J.* (2019) 38:e101496. doi: 10.15252/embj.2019101496
90. Bhargava V, Goldstein CD, Russell L, Xu L, Ahmed M, Li W, et al. GCNA preserves genome integrity and fertility across species. *Dev Cell.* (2020) 52:38–52.e10. doi: 10.1016/j.devcel.2019.11.007
91. Dokshin GA, Davis GM, Sawle AD, Eldrige MD, Nicholls PK, Gourley TE, et al. GCNA interacts with spartan and topoisomerase II to regulate genome stability. *Dev Cell.* (2020) 52:53–68.e6. doi: 10.1016/j.devcel.2019.11.006
92. Livneh I, Cohen-Kaplan V, Cohen-Rosenzweig C, Avni N, Ciechanover A. The life cycle of the 26S proteasome: from birth, through regulation and function, and onto its death. *Cell Res.* (2016) 26:869–85. doi: 10.1038/cr.2016.86
93. Quiévryn G, Zhitkovich A. Loss of DNA-protein crosslinks from formaldehyde-exposed cells occurs through spontaneous hydrolysis and an active repair process linked to proteasome function. *Carcinogenesis.* (2000) 21:1573–80. doi: 10.1093/carcin/21.8.1573
94. Ortega-Atienza S, Green SE, Zhitkovich A. Proteasome activity is important for replication recovery, CHK1 phosphorylation and prevention of G2 arrest after low-dose formaldehyde. *Toxicol Appl Pharmacol.* (2015) 286:135–41. doi: 10.1016/j.taap.2015.03.018
95. Larsen NB, Gao AO, Sparks JL, Gallina I, Wu RA, Mann M, et al. Replication-coupled DNA-protein crosslink repair by SPTN and the proteasome in xenopus egg extracts. *Mol Cell.* (2019) 73:574–588.e7. doi: 10.1016/j.molcel.2018.11.024
96. Takahashi K, Morita T, Kawazoe Y. Mutagenic characteristics of formaldehyde on bacterial systems. *Mutat Res Toxicol.* (1985) 156:153–61. doi: 10.1016/0165-1218(85)90058-8
97. de Graaf B, Clore A, McCullough AK. Cellular pathways for DNA repair and damage tolerance of formaldehyde-induced DNA-protein crosslinks. *DNA Repair.* (2009) 8:1207–14. doi: 10.1016/j.dnarep.2009.06.007
98. Baker DJ, Wuenschell G, Xia L, Termini J, Bates SE, Riggs AD, et al. Nucleotide excision repair eliminates unique DNA-protein cross-links from mammalian cells *. *J Biol Chem.* (2007) 282:22592–604. doi: 10.1074/jbc.M702856200
99. Nakano T, Katafuchi A, Matsubara M, Terato H, Tsuboi T, Masuda T, et al. Homologous recombination but not nucleotide excision repair plays a pivotal role in tolerance of DNA-protein cross-links in mammalian cells *. *J Biol Chem.* (2009) 284:27065–76. doi: 10.1074/jbc.M109.019174
100. Oka Y, Nakazawa Y, Shimada M, Ogi T. Endogenous aldehyde-induced DNA-protein crosslinks are resolved by transcription-coupled repair. *Nat Cell Biol.* (2024) 26:784–96. doi: 10.1038/s41556-024-01401-2
101. Ridpath JR, Nakamura A, Tano K, Luke AM, Sonoda E, Arakawa H, et al. Cells deficient in the FANCD1/BRCA pathway are hypersensitive to plasma levels of formaldehyde. *Cancer Res.* (2007) 67:11117–22. doi: 10.1158/0008-5472.CAN-07-3028
102. Zhang H, Xiong Y, Chen J. DNA-protein cross-link repair: what do we know now? *Cell Biosci.* (2020) 10:3. doi: 10.1186/s13578-019-0366-z
103. Neale MJ, Pan J, Keeney S. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature.* (2005) 436:1053–7. doi: 10.1038/nature03872
104. Aparicio T, Baer R, Gottesman M, Gautier J, MRN, ChIP, and BRCA1 mediate repair of topoisomerase II-DNA adducts. *J Cell Biol.* (2016) 212:399–408. doi: 10.1083/jcb.201504005
105. Lee KC, Padgett K, Curtis H, Cowell IG, Moiani D, Sondka Z, et al. MRE11 facilitates the removal of human topoisomerase II complexes from genomic DNA. *Biol Open.* (2012) 1:863–73. doi: 10.1242/bio.20121834
106. Hoa NN, Shimizu T, Zhou ZW, Wang ZQ, Deshpande RA, Pault TT, et al. Mre11 is essential for the removal of lethal topoisomerase 2 covalent cleavage complexes. *Mol Cell.* (2016) 64:580–92. doi: 10.1016/j.molcel.2016.10.011
107. Ashour ME, Mosammaparast N. Mechanisms of damage tolerance and repair during DNA replication. *Nucleic Acids Res.* (2021) 49:3033–47. doi: 10.1093/nar/gkab101
108. Giannattasio M, Zwicky K, Follonier C, Foiani M, Lopes M, Branzei D. Visualization of recombination-mediated damage bypass by template switching. *Nat Struct Mol Biol.* (2014) 21:884–92. doi: 10.1038/nsmb.2888
109. Yukawa Y, Muto M, Hori K, Nagayoshi H, Yokoyama A, Chiba T, et al. Combination of ADH1B*2/ALDH2*2 polymorphisms alters acetaldehyde-derived DNA damage in the blood of Japanese alcoholics. *Cancer Sci.* (2012) 103:1651–5. doi: 10.1111/j.1349-7006.2012.02360.x
110. Chen X, Legrand AJ, Cunniffe S, Hume S, Poletto M, Vaz B, et al. Interplay between base excision repair protein XRCC1 and ALDH2 predicts overall survival in lung and liver cancer patients. *Cell Oncol Dordr.* (2018) 41:527–39. doi: 10.1007/s13402-018-0390-8
111. Batel A, Polović M, Glumac M, Šuman O, Jadrijević S, Lozić B, et al. SPTN is involved in hepatocellular carcinoma development through the ER stress response. *Cancer Gene Ther.* (2024) 31:376–86. doi: 10.1038/s41417-023-00708-w
112. Maskey RS, Flatten KS, Sieben CJ, Peterson KL, Baker DJ, Nam HJ, et al. Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis. *Nucleic Acids Res.* (2017) 45:4564–76. doi: 10.1093/nar/gkx107
113. Cook KL, Soto-Pantoja DR, Clarke PAG, Crux MI, Zwart A, Wärrri A, et al. Endoplasmic reticulum stress protein GRP78 modulates lipid metabolism to control drug sensitivity and antitumor immunity in breast cancer. *Cancer Res.* (2016) 76:5657–70. doi: 10.1158/0008-5472.CAN-15-2616
114. Niki E, Yoshida Y, Saito Y, Noguchi N. Lipid peroxidation: Mechanisms, inhibition, and biological effects. *Biochem Biophys Res Commun.* (2005) 338:668–76. doi: 10.1016/j.bbrc.2005.08.072
115. Zangar RC, Davydov DR, Verma S. Mechanisms that regulate production of reactive oxygen species by cytochrome P450. *Toxicol Appl Pharmacol.* (2004) 199:316–31. doi: 10.1016/j.taap.2004.01.018
116. Ayala A, Muñoz MF, Argüelles S. Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. *Oxid Med Cell Longev.* (2014) 2014:360438. doi: 10.1155/2014/360438
117. Flor A, Wolfgeher D, Li J, Hanakahi LA, Kron SJ. Lipid-derived electrophiles mediate the effects of chemotherapeutic topoisomerase I poisons. *Cell Chem Biol.* (2021) 28:776–787.e8. doi: 10.1016/j.chembiol.2020.11.011
118. Kiianitsa K, Maizels N. A rapid and sensitive assay for DNA-protein covalent complexes in living cells. *Nucleic Acids Res.* (2013) 41:e104. doi: 10.1093/nar/gkt171
119. Weickert P, Dürauer S, Götz MJ, Li HY, Stinglele J. Electro-elution-based purification of covalent DNA-protein cross-links. *Nat Protoc.* (2024), 1–24. doi: 10.1038/s41596-024-01004-z