Characterisation of CSB ubiquitylation in response to UV-induced DNA damage

Liam Gaul

University College London

and

The Francis Crick Institute

PhD Supervisor: Jesper Q. Svejstrup

A thesis submitted for the degree of

Doctor of Philosophy

University College London

January 2021

Declaration

I Liam Gaul confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

A variety of DNA repair pathways operate in different cellular contexts to tackle a diversity of DNA lesions and maintain genome stability. Nucleotide excision repair (NER) recognises and removes a variety of helix-distorting lesions, operating via two pathways: a global-genome (GG-NER) and transcription-coupled (TC-NER) pathway. GG-NER repairs damage at any locus in the genome thus promoting genome stability by abrogating replication-mediated stress and mutagenesis. TC-NER on the other hand is restricted to the template strand of actively transcribed genes and provides means to rapidly repair lesions that would otherwise impair transcription. Thus, TC-NER has seemingly evolved to not only maintain genome stability but sustain transcription by removal of RNA polymerase II (RNAPII)-stalling lesions.

The two pathways differ in their mode of recognition of lesions while downstream repair steps of excision and DNA synthesis are mutual. GG-NER relies on XPC and UV-DDB that recognise and bind directly to lesions, initiating repair. TC-NER is initiated by a lesion-stalled RNAPII, which is recognised by Cockayne's syndrome B (CSB) prompting recruitment of further repair factors. Mutations in CSB or CSA result in Cockayne's syndrome, a disease characterised by photosensitivity, neurological deficiencies and progeria.

UV-DDB and CSA reside in ubiquitin ligase complexes highlighting the importance of ubiquitylation in NER. The Svejstrup laboratory previously identified a ubiquitin-binding domain in CSB that was essential for its function as well as several CSA- and UV-dependent ubiquitylation sites on CSB. Building on this work, I have developed a cell system exclusively expressing CSB mutants that are not ubiquitylated in response to UV-irradiation. I present evidence that ubiquitylation of CSB is necessary for the recovery of transcription and cell survival following UV-irradiation. Inhibition of CSB ubiquitylation does not affect its recruitment to chromatin following UV, indicating it is a step downstream of the recognition and binding of a stalled RNAPII. These data support the hypothesis that CSB ubiquitylation is a vital step in TC-NER.

Impact Statement

The repair of DNA damage is a fundamental and ubiquitous process in all cells of the body that continuously protects from debilitating diseases such as cancer. Metabolic processes provide a constant source of DNA damaging agents, as do simple daily activities such as being outside and exposed to the sun. Therefore, there is no escape from DNA damage and all of our health depends on its recognition and repair. Unfortunately, rare genetic diseases exist in the population that cause some people to suffer from defective DNA repair, the consequences of which can be life-threatening and debilitating. It is therefore imperative that research into the causes, consequences, and treatment of these diseases is undertaken to ensure the improvement in quality of life of sufferers. In fact, the development each of these areas is dependent on the others. Development of treatments cannot occur without first understanding the basic biology of what is going wrong.

In this thesis I investigate the function of a key protein, CSB, involved in a DNA repair process called transcription-coupled nucleotide excision repair (TC-NER). When CSB is mutated and non-functional it causes the severe and debilitating disease Cockayne syndrome due to impairment of TC-NER. Patients suffer from premature aging (progeria), severe sensitivity to sunlight, moderate to severe learning delay, and impaired nervous system development. While much progress has been made elucidating the mechanisms of how TC-NER works at the molecular level, much less is known about the master regulator of this process, CSB. This study aims to shed light on the role of CSB in the TC-NER process helping us to understand what goes wrong in the cells of people with cockayne syndrome. This might help to develop better diagnostic tools, or better inform clinicians about the prospects of patients with certain mutations, thus improving care. It should inform and encourage further research, creating a positive feedback cycle that drives progress forward. Advancement in basic biomedical research is essential to the development of drugs and treatments for patients and this study is a small but necessary contribution to that endeavour.

Acknowledgement

This work is the product of encouragement, love, support, and guidance from numerous people throughout my life that helped me to get here. But most of all, I want to thank my parents for giving me all of that in abundance. To my partner who has been with me and supported me through all of this, you have kept me grounded and made life outside of the lab all the more enjoyable.

I express a great deal of gratitude to my PhD supervisor, Jesper Svejstrup, for giving me my own project to work on and develop, for the opportunity to contribute to the work of his lab, and for the great deal of science and patience I have learnt throughout. I also thank my thesis committee, Simon Boulton, Sue Hadjur, and Richard Treisman for their useful insights and scientific discussions. To all members of the Svejstrup lab during my tenure, I have learnt so much from all of you and I am so grateful to have spent the formative years of my scientific career learning from and alongside you.

A huge and special thank you goes to Lea Gregersen, who has tolerated me with stoicism for more than four years and without whom I would certainly not be here today. I can't imagine learning from a better scientist.

A deep debt of gratitude to Sanjay Chandriani and Amy Kistler, who gave me my first scientific research opportunity and instilled in me the confidence to pursue a PhD. That was an incredible experience for many reasons and I'm eternally grateful.

To my Crick pals, you know who you are. Thanks for making this experience all the more enjoyable and giving me friends for life! I can't wait to celebrate with you if this pandemic ever ends!

Many people at The Crick have made this research possible, not least the STPs and I am grateful for all their work and contributions. This of course wouldn't be possible without charitable donations and I thank all members of the public who have donated to medical research charities.

Table of Contents

Table of figures

List of tables

Abbreviations

ATM	A-T mutated
AP site	Apurinic or apyrimidinic site
ATR	Ataxia telangiectasia and Rad3-related protein
BER	Base excision repair
CTD	C-terminal domain of RNAPII
COFS	Cerebro-oculo-facio-skeletal syndrome
CS	Cockayne syndrome
CV	Column volumes
CSN	COP9 signalosome
CRL	Cullin-RING ligase
DCAF	DDB1-CUL4 associated factors
DNA	Deoxyribose nucleic acid
DUB	Deubiquitylating enzymes
DDR	DNA damage response
DSBs	DNA double strand breaks
LIG1	DNA ligase I
LIG3	DNA ligase III
LIG4	DNA ligase IV
Pol β	DNA polymerase β
Pol δ	DNA polymerase δ
Pol ϵ	DNA polymerase ϵ
Pol η	DNA polymerase η
Pol κ	DNA polymerase κ
DNA-PKs	DNA-dependent protein kinase catalytic subunit
HECT	Homologous to the E6-AP carboxyl terminus
ES	Embryonic stem cells
ERAD	Endoplasmic-reticulum-associated protein degradation
EPR	Enzymatic photoreactivation
FEN1	Flap endonuclease 1
GG-NER	Global genome nucleotide excision repair
gRNA’s	guide RNA’s
HR	Homologous recombination

IP	Immunoprecipitation
IR	Ionising radiation
MEFs	Mouse embryonic fibroblasts
MFD	Mutation frequency decline
NHEJ	Non-homologous end joining
NLS	Nuclear localisation signal
NER	Nucleotide excision repair
PIKKs	Phosphoinositide 3-kinase related kinases
PARP	Poly(ADP-ribose) polymerase
PNKP	Polynucleotide kinase phosphatase enzyme
PCNA	Proliferating cell nuclear antigen
ROS	Reactive oxygen species
RING	Really interesting new gene
RFC	Replication factor C
RFC	Replication factor C
RNA	Ribose nucleic acid
RBR	RING-in-between-RING
RNAP	RNA polymerase (bacterial)
RNAPI	RNA polymerase I
RNAPII	RNA polymerase II
RNAPIII	RNA polymerase III
SCF	Skp1–Cul1–F-box-protein
SUMO	Small ubiquitin-like modifier
SILAC	Stable isotope labelling with amino acids in cell culture
TRCF	Transcription repair coupling factor
TC-NER	Transcription-coupled nucleotide excision repair
TDP1	Tyrosyl-DNA phosphodiesterase 1
UBA	Ubiquitin associated domain
UBD	Ubiquitin-binding domain
UBZ	Ubiquitin-binding ZnF domain
UbL	Ubiquitin-like domain
UIM	Ubiquitin-interacting motif
UV	Ultraviolet radiation

UVSS	UV sensitivity syndrome
WCE	Whole-cell extract
WT	Wildtype
XP	Xeroderma pigmentosum

Chapter 1. Introduction

1.1 The DNA damage response

The ‘DNA damage response’ refers to the multi-pronged cellular response to chemical afflictions of DNA, by way of surveillance, recognition and repair of DNA damage, as well as signalling to, and management of, other cellular processes such as transcription, replication and the cell cycle, to maintain genome integrity.

The faithful duplication and propagation of DNA from mother to daughter cell marks the completion of the cell cycle. The ultimate goal of any organism is to complete this process successfully and faithfully enough times that DNA from the germline is passed from parent to offspring, requiring that the organism maintain genome stability for life. For this to occur, throughout the organism’s life cycle, DNA must be protected from chemical alterations that would interfere with fitness, viability and fertility. Thus, repair pathways have evolved in all three domains of life that work to keep DNA intact and undamaged from the enormous assortment of endogenous and exogenous sources of DNA damage.

1.1.1 Endogenous DNA damage

DNA is an intrinsically unstable molecule due to its labile glycosidic bond between the deoxyribose sugar and nucleoside base that is susceptible to hydrolysis in aqueous solution resulting in depurination more commonly, or depyrimidination at a 20-fold lower rate, leaving abasic or AP sites (Lindahl, 1993). It is estimated from bacterial studies that human cells may have some 2,000–10,000 depurination events every day, and these must, of course, be repaired (Lindahl & Nyberg, 1972). DNA bases themselves are subject to hydrolytic deamination, with the most common target being 5-methylcytosine and then cytosine deamination to thymine and uracil, respectively. However, the double helical structure of DNA provides much protection from this process, in contrast to providing little protection against hydrolysis of the glycosidic bond (Lindahl, 1993). Processes such as transcription

and replication—that melt the double helix—increase the susceptibility of bases to deamination (Bhagwat et al., 2016; Jinks-Robertson & Bhagwat, 2014). Deaminated cytosine is more rapidly repaired due to excision by the abundant uracil-DNA glycosylase, generating an abasic site, which is corrected by base excision repair. Uracil-DNA glycosylase does not recognise deaminated 5-methylcytosine (thymine); the guanine–thymine base pair is recognised and corrected by mismatch repair, which is much slower (Lindahl, 1993). Thus, although 5-methylcytosines are far less abundant than cytosine, at 4%, they are disproportionately a source of mutagenesis (C·G → T·A) due to their greater risk of deamination and subsequent slower rate of repair (Breiling & Lyko, 2015).

Reactive oxygen species (ROS), such as hydroxyl radicals, superoxide, hydroperoxyl, hydrogen peroxide, and nitric oxide, are just some of the natural products of aerobic metabolism, enzymatic reactions and intercellular signalling pathways (M. D. Evans et al., 1997). These endogenously produced ROS present a major source of DNA damage (as well as to damage of other organic molecules) in cells and the abundance of DNA lesions they can generate are constantly being discovered (Cooke et al., 2003). The major and most well-studied lesion produced from oxidative attack on DNA is 8-hydroxyguanine due to guanine possessing the lowest redox potential of the nitrogenous bases. 8-hydroxyguanine preferentially pairs with adenine over cytosine, potentially leading to G·C → T·A transversion mutations following replication (Shibutani et al., 1991). Oxidation of guanine to 8-hydroxyguanine and hydrolytic deamination of cytosine to uracil are likely the most abundant endogenously-produced DNA lesions leading to mutagenesis in cells (Lindahl, 1993).

Other sources of endogenous DNA damage arise from lipid peroxidation forming aldehydes that generate exocyclic DNA adducts, oestrogen metabolites, alkylating agents such as *S*-adenosylmethionine that methylate DNA, and formaldehyde as a by-product of the one carbon cycle that can form DNA crosslinks (Burgos-Barragan et al., 2017; De Bont & van Larebeke, 2004). The number of endogenous DNA damaging agents produced in cells is innumerable, as is the number of different DNA lesions they can generate, and new discoveries are constantly being made.

Those outlined above are merely some of the most well studied and most abundant lesions, and thus likely the most important for human disease.

1.1.2 Exogenous DNA damage

There are two main sources of environmental DNA damage, which are ionising radiation (IR) and ultraviolet radiation (UV). Ionising radiation, in the form of alpha, beta, gamma, neutrons and X-rays is produced from diverse sources such as cosmic and terrestrial radiation, radon gas, soil, food and medical devices (Chatterjee & Walker, 2017). Ionising radiation can damage DNA directly, causing single strand breaks with unique 3′ ends of 3′-phosphate or 3′-phosphoglycolate, rather than 3′-OH (Henner et al., 1982). If two single strand breaks are produced in close proximity on opposite strands, this will lead to a double-strand break in the phosphodiester backbone, which is probably the most serious and lethal of DNA lesions. IR can also be absorbed by water, producing hydroxyl radicals that induce oxidative damage to DNA and this may account for 65% of DNA damage by IR (Vignard et al., 2013).

UV light is one of the most abundant sources of DNA damage as most humans are exposed to it for several hours every day as sunlight. It is categorised into three classes based on wavelength: UV-A (320–400 nm), UV-B (290–320 nm), and UV-C (190–290 nm). UV-C is most damaging to DNA, due to greater photo-absorption of energy, which dramatically decreases for wavelengths above 260 nm. However, UV-C is almost entirely absorbed by the ozone layer and sunlight reaching the earth's surface is 5.1% UV-A, 0.3% UV-B, 62.7% visible light, and 31.9% infrared (Chatterjee & Walker, 2017). Direct absorption of UV by DNA results in the covalent linkage of two adjacent pyrimidines, resulting in either cyclobutane pyrimidine dimers (CPDs) or 6-4 photoproducts (6–4 PPs) that disrupt Watson–Crick base pairing and result in in distortion of the DNA helix (Davies, 1995).

UV of longer wavelengths, such as UV-A, although not absorbed by DNA, can damage it indirectly via photosensitisers. These are cellular chromophores that can absorb UV-A and transfer energy to DNA or other molecules producing ROS, which

induce oxidative damage (Brem et al., 2017). As mentioned above, the most common DNA lesion from oxidative damage is 8-hydroxyguanine, which produces G→T transversion mutations, but these are rarely observed in UV-A irradiated cells or in the p53 gene of skin tumours, where T→G (a signature of CPD-induced mutagenesis) predominate. It was shown that in cells exclusively irradiated with UV-A, in contrast to cells irradiated with UV-C (the most commonly used wavelength in the laboratory), CPDs were preferentially formed at TT sites over other dipyrimidines sites (Rochette et al., 2003). Due to the extremely low absorption of UV-A by DNA, it was presumed that an as-yet unidentified photosensitiser must transfer energy to DNA to produce CPDs. However, more recent studies using highly purified DNA suggest that CPDs can be produced directly by DNA absorption of UV-A (Jiang et al., 2009; Mouret et al., 2010). In contrast to what was previously believed, several lines of evidence now indicate that the majority of natural UV damage to DNA probably results in CPDs, not just from UV-B, but also from UV-A irradiation. However, oxidative damage is common too and many DNA lesions are still being discovered. The mutational signatures sequenced from UV-A, B, or C irradiated cells, or from skin tumours, show heterogeneity and suggest several different DNA lesions are produced by UV (Moreno et al., 2020).

Other exogenous sources of DNA damage are chemical agents in the environment that are innumerable and diverse. Some examples are alkylating agents, aromatic amines, polycyclic aromatic hydrocarbon, and microorganismal toxins. They can come from diverse sources such as food, smoking, cooking, fossil fuel combustion, chemotherapy and other medications, industrial chemicals, pesticides, chemical weapons and microorganisms, among others (Chatterjee & Walker, 2017). The scope of this thesis (and for brevity) does not demand these be reviewed in great detail — much of the damage they inflict is similar or is at least repaired by the different repair pathways in the same way as the endogenous sources outlined above.

1.1.3 DNA damage repair pathways

There are several DNA damage repair pathways that each contend with distinct sets of DNA lesions via different mechanisms. Evidence of the simplest of DNA repair mechanisms emerged before DNA was even known to be the molecule carrying genetic information. Albert Kelner was trying to generate mutants of *Streptomyces griseus* using UV, in the hope of identifying antibiotics, when he noticed that fluorescent lights (of visible wavelengths) in the laboratory were responsible for discrepancies in the survival of colonies on agar plates (KELNER, 1949). Concomitantly, Renato Dulbecco too had noticed that agar plates of phage-infected bacteria stacked on his bench displayed different levels of survival after UV irradiation depending on their position in the stack (Dulbecco, 1949). These were the first results describing 'enzymatic photoreactivation' (EPR), a process by which light of longer wavelengths (300–500 nm) activates photolyase enzymes that can directly reverse UV-induced pyrimidine dimers to their original form (Davidson, 2006). Dulbecco's colleague, the young graduate student James Watson, who would later go on to elucidate the structure of DNA, wrote his thesis on how photoreactivation was not observed in cells exposed to ionizing radiation (Friedberg, 2008). Coincidentally, EPR is thought to also be the first enzymatic DNA repair process to evolve, signifying the ubiquity and threat to life from UV light – but EPR is absent from placental mammals.

Several more complex DNA repair pathways involving sometimes dozens of proteins that contend with a great number of different kinds of DNA lesions have since been identified and well characterised, which I review below. Strictly, DNA repair is the enzymatic reversion of chemically altered nucleotides (lesions) to their original state, restoring genome integrity. However, as the DNA repair field has evolved and research has led to important discoveries in adjacent areas that impinge on DNA repair, it has come to encompass a broader variety of biological processes, which I refer to as the DNA damage response, which is also reviewed below.

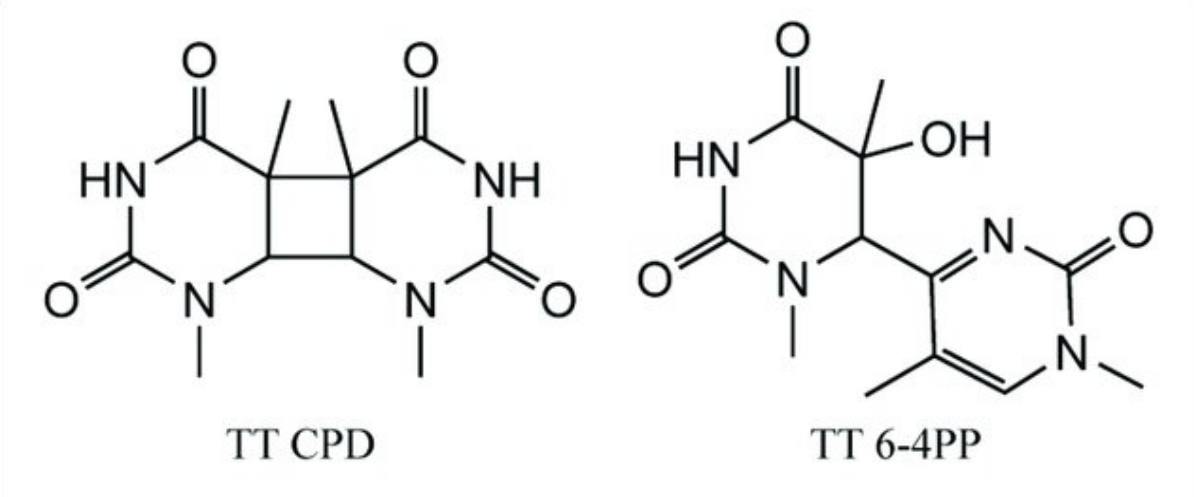


Figure 1-1 Skeletal structural formula of CPD and 6–4PP formed at thymine dimer

1.1.3.1 Base excision repair

Base excision repair (BER) is probably the most widely employed DNA repair mechanism due to the prevalence of DNA lesions that it recognises and repairs. It corrects forms of oxidation, deamination, alkylation and abasic damage that commonly arise due to natural degradative processes of DNA in aqueous solution or from attack by metabolic by-products (Chatterjee & Walker, 2017). The lesions are all common in that they are not 'bulky'; that is, they do not significantly distort the DNA helix, but are structurally diverse. Therefore, a number of enzymes called DNA glycosylases, are needed to recognise the diversity of lesions. Eleven such glycosylases have been identified in humans, each specific for different types of lesions, but with some redundancy. They all cleave the N-glycosidic bond linking the ribose sugar and nucleobase, leaving an abasic (AP) site. In the classic pathway, performed by monofunctional glycosylases (e.g., UNG1) that can't cleave the phosphodiester bond, the resulting AP site is then recognised by an AP endonuclease (APE1 in humans), which cleaves 5′ of the AP site, generating a single strand break (SSB or nick)and leaving a 5′-deoxyribose phosphate and 3′-hydroxyl. DNA polymerase β, with its AP lyase activity can then resolve the 5′-deoxyribose phosphate that is blocking nucleotide incorporation, by cleaving 3′ of the deoxyribose sugar, releasing it and leaving a 5′-phosphate (Ide & Kotera, 2004). DNA polymerase β can then gap fill one nucleotide using the opposite

strand as template. Finally, the XRCC1-DNA ligase IIIα then seals the nick to complete the repair (Dianov & Hübscher, 2013).

These last steps are also called single-strand break (SSB) repair, which is a sub-portion of the BER pathway, employed where SSBs are generated by cleavage of the phosphodiester backbone, usually from oxidative damage or abortive DNA topoisomerase 1 activity (Caldecott, 2008). SSB formation can result in several different DNA ends, which must be processed to a suitable substrate for DNA polymerase gap-filling. Some glycosylases, termed 'bifunctional' (NEIL1, NEIL2, NTH1, OGG1), possess intrinsic AP lyase activity as well as glycosylase activity and cleave the phosphodiester backbone 5′ of the AP site. This can leave a 3′-α,β-unsaturated aldehyde in the case of β lyases, or a 3′-phosphate in the case of β,δ-eliminations, which need to be further processed for gap-filling (Dianov & Hübscher, 2013). The five enzymes known to process SSB ends are: DNA polymerase β (Pol β), which removes 5′-deoxyribose phosphates (Matsumoto & Kim, 1995); APE1, which removes 3′- deoxyribose phosphates (Demple & Harrison, 1994); Polynucleotide Kinase Phosphatase (PNKP), which phosphorylates or dephosphorylates 5′-hydroxyl or 3′-ends, respectively (Weinfeld et al., 2011); Aprataxin, which removes 5′-termini blocked by abortive ligation reactions (Ahel et al., 2006); and tyrosyl DNA phosphodiesterase (TDP1), which repairs SSBs generated by abortive DNA topoisomerase reactions (Ledesma et al., 2009). Once these ends are resolved, Pol β can gap fill one nucleotide, which is the short patch BER pathway.

If the DNA ends can't be resolved to leave a 5′-phosphate or 3′-OH necessary for Pol β, then the long patch BER pathway is used. Here, Pol δ or Pol ϵ synthesise several nucleotides starting at the gap and proceeding by synthesising several nucleotides by displacement of the 5′-deoxyribose phosphate-containing downstream strand in concert with proliferating cell nuclear antigen (PCNA) and replication factor C (RFC). The results in a flap of displaced nucleotides which is cut by flap endonuclease 1 (FEN1), after which the nick is sealed by DNA ligase I (Dianov & Hübscher, 2013; Krokan & Bjørås, 2013).

Both the short patch- and long patch pathways of BER have been completely reconstituted with purified proteins *in vitro,* demonstrating that at least the minimal protein repertoire for BER has been identified and characterised (Klungland & Lindahl, 1997; Kubota et al., 1996). Of course, there can and does exist certain cellular contexts where BER cannot operate efficiently without help from additional enzymes that are not always essential. For example, chromatin remodelling is sometimes needed for access of lesions in nucleosomes (Odell et al., 2013).

1.1.3.2 Mismatch repair

Mismatch repair (MMR) doesn't repair DNA lesions as such, but as the name suggests, corrects for mismatched DNA bases that arise as a result of DNA replication. Although DNA replication is largely a highly accurate process, due to the nucleotide specificity of DNA polymerases as well as their proofreading capabilities, mistakes invariably arise. These mistakes can occur in the form of single base pair mismatches, or insertion/deletion mispairs, both of which must be identified and corrected to avoid mutations (G. M. Li, 2008).

As MMR has been extensively studied and well characterised since first being discovered in *E. coli*, it is useful to outline the mechanism of the prokaryotic pathway as a foundation to understand the further complexities of the mammalian one.

Mismatches in DNA are recognised by the MutS protein, which acts as a homodimer when binding DNA mismatches, but each monomer have different conformations that essentially makes MutS a heterodimer at the structural level (Lamers et al., 2000). This mimics how the human homologues of MutS function as heterodimers of several different proteins to recognise different substrates. A MutS dimer binds DNA with ATP bound to one of the monomers and hydrolysis by its ATPase domain causes a confirmational change that alters binding relative to the other monomer. This allows for MutS to open and close, acting as a sliding clamp along DNA, which aids in discrimination of mismatches and recruitment of the MutL

protein (Lamers et al., 2000). However, there are competing models for how MutS and its ATPase activity function and this is still an active area or research.

Upon its recruitment, MutL binds MutS, which further facilitates binding of MutH. MutL possesses ATPase activity and functions as a homodimer like MutL, but its function is not well characterised and remains controversial. MutL is essential for MMR, as inactivating mutations of its ATPase domain abolish all MMR (G. M. Li, 2008). The MutH protein possesses endonuclease activity and uses the hemi-methylation pattern of newly replicated DNA in *E. coli* to distinguish between template and daughter strands, cutting the unmethylated daughter strand, producing a nick (Modrich & Lahue, 1996). In humans, while homologues of MutH do not exist, the MutL homologue MutLα contains intrinsic endonuclease activity, which functions in its place (Kadyrov et al., 2006). UvrD then loads at the nick site and uses its helicase activity to displace the daughter strand. Several exonucleases then excise the strand depending on the break site relative to the mismatch (ExoI or ExoX are 3′→5′ exonuclease and ExoVII or RecJ are 5′→3′ exonuclease). DNA polymerase III (Pol δ in humans) then resynthesises a new strand in the resulting single-strand gap and DNA ligase (LIG1 in humans) seals the nick (Modrich & Lahue, 1996).

While MMR is a highly conserved process, it is understandably more complex in humans and there is much active research in this field. Repair is strand specific and bi-directional in humans too, but as hemi-methylation is not present the daughter strand is identified by nicking by MutLα in concert with PCNA, which is loaded asymmetrically by RFC allowing for preferential nascent strand identification (Kadyrov et al., 2006).

1.1.3.3 Double strand break repair

DNA double strand breaks (DSBs) are widely considered the most toxic of all DNA lesions as a single, unrepaired break can be enough to kill a cell. Due to the nature of the break, the repair of DSBs is more complicated and often error-prone, potentially leading to deletion of genetic material as well as chromosome

rearrangements. DSBs can be induced by ionising radiation, chemotherapeutic drugs, oxidative damage, mechanical chromosome breakage, replication fork dysfunction, or the deprotection of telomeres. However, they are also enzymatically induced in a regulated manner during V(D)J recombination and immunoglobulin class-switching (Khanna & Jackson, 2001).

There are two main pathways by which cells repair DSBs, non-homologous end joining (NHEJ) and homologous recombination (HR) each operating at different stages of the cell cycle and on slightly different substrates.

1.1.3.3.1 NHEJ

NHEJ is initiated by the heterodimer of Ku70-Ku80, which recognises DNA ends with one dimer binding a DNA end each (Britton et al., 2013). The Ku70-Ku80 complex acts as a scaffold to recruit other NHEJ proteins: DNA-dependent protein kinase catalytic subunit (DNA-PKs), LIG4 and scaffolding factors: XRCC4, XLF, PAXX. The recruitment of DNA-PKcs to DNA ends by the Ku70-Ku80 complex initiates the tethering of the DNA ends to form a long-range synapse. Recruitment of XLF, XRCC4 and LIG4, as well as binding of Ku80 to DNA-PKcs, activates its kinase activity and its autophosphorylation brings the DNA ends into close proximity. This positions the DNA ends to be ligated by LIG4 (T. G. W. Graham et al., 2016). The exact mechanism of how these proteins, and specifically the catalytic activity of DNA-PKcs, coordinate the two stages of DNA bridging is still a matter of investigation.

Many different DNA ends can be produced depending on the source of the DSB and it is these ends that is thought to render NHEJ error-prone, as contrary to popular discourse, it is usually accurate and efficient (Bétermier et al., 2014). If DNA ends are incompatible, then several proteins can process them to a state suitable for ligation. Pol λ and Pol μ fill in any overhands or gaps, the nuclease Artemis removes damaged nucleotides, tyrosyl-DNA phosphodiesterase 1 (TDP1) removes topoisomerase I adducts, TDP2 removes 5′ DNA-topoisomerase II

adducts, and aprataxin resolves abortive ligation intermediates (Stinson et al., 2020).

1.1.3.3.2 HR

HR is the other major pathway that repair DSBs and operates exclusively in the S and G2 phases of the cell cycle when sister chromatids are available for recombination. Due to the extensive sequence homology used in this repair pathway, it is largely error free.

The Ku70-Ku80 complex still recognises and identifies DNA ends in HR. The first step is end resection of the DNA ends by the MRN complex in concert with CtIP, where the MRE11 endonuclease that has 3′→5′ activity nicks far from the DSB and digests towards it, displacing Ku70-Ku80. Further resection is carried out by exonucleases such as EXO1 and the helicase BLM (Scully et al., 2019). The resulting ssDNA is rapidly coated with RPA to prevent annealing to other DNA. BRCA2 (or RAD52 in yeast) then displaces RPA from DNA to facilitate loading of RAD51 nucleoprotein filaments onto DNA. RAD51 is a recombinase that mediates a homology search for complementary DNA and ‘invades’ dsDNA by base pairing its bound ssDNA with that of a complementary strand, displacing another strand and forming a synapse called a displacement loop (D-loop). The BRCA1-BARD1 complex enhances the recombinase activity of RAD51 and is in fact indispensable for fully-functional HR. Pol δ (but sometimes translesion DNA polymerases) then extends the invading strand using the invaded donor DNA as a template, which is termed gene conversion (Zhao et al., 2017). The importance of BRCA1 and BRCA2 to this pathway is highlighted by the fact that germline mutations massively increase risk for several cancers (Sopik et al., 2015) and that they are synthetic lethal with PARP inhibitors, a standard cancer drug treatment (Faraoni & Graziani, 2018).

The heteroduplex DNA synapse can take several forms, which determine different HR pathways. For example, in meiosis, double Holliday junctions are formed, by invasion by template DNA of donor DNA on the homologous chromosomes, which

when resolved result in crossover of genetic information from one chromosome to the other. The major pathway in somatic cells is synthesis-dependent strand annealing, where no crossover happens due to the inhibition of Holliday junction formation by the displacement of the invading strand, which re-anneals to the opposite end of the original break where DNA synthesis can continue. (Chapman et al., 2012; Symington & Gautier, 2011)

1.1.3.4 Nucleotide excision repair

As the results presented in this thesis directly involves the mechanism of NER, this is expanded in much greater detail in a later section (see 1.3).

1.1.4 DNA damage checkpoint signalling

Progressing through the cell cycle with DNA damage present can be deleterious to genome integrity and to accurate DNA replication. Thus, a series of signalling pathways exist to promote DNA repair, and halt the cell-cycle, and promote other damage responses when damage is detected.

Transduction of these signalling pathways is regulated by three kinases, which can be considered the gatekeepers of the DNA damage response (DDR): ATM, ATR, and DNA-PK. They are members of the phosphoinositide 3-kinase related kinases (PIKKs) and thus serine/threonine kinases (targeting S/T-Q motifs), but are only known to phosphorylate protein substrates (Blackford & Jackson, 2017). Each of these kinases is tightly regulated by a protein co-factor that regulates their recruitment to DNA damage sites and prevents aberrant activation. Ku80 regulates DNA-PK (Singleton et al., 1999), NBS1 regulates ATM (Falck et al., 2005), and ATRIP regulates ATR (Zou & Elledge, 2003).

As outlined previously 1.1.3.3.1), DNA-PKcs is recruited to DSBs by Ku70-Ku80, which activates its kinase activity and promotes NHEJ repair of DSBs. Thus DNA-PK acts to promote repair of DSBs by NHEJ.

ATM phosphorylates a massive repertoire of proteins, many of which are kinases themselves, thus triggering a diverse signalling cascade in response to DSBs (Bakkenist & Kastan, 2003). It is recruited to DSBs and is activated by the MRN complex through interaction with the NBS1 subunit. There is also evidence to suggest that ATM can be activated by simple tethering to DNA ends, or ssDNA, as well as chromatin (Shiotani & Zou, 2009; Soutoglou & Misteli, 2008; You et al., 2007). A local substrate in the vicinity of DSBs is the histone H2AX, which is phosphorylated to form γH2AX, acting as a marker of DSBs and facilitating repair (Fernandez-Capetillo et al., 2004). One of the most well characterised substrates of ATM is the CHK2 kinase, which has a role in the G1 checkpoint — halting the cell cycle from progression into S phase. It does this through inhibition of MDM2, stabilising p53 and inducing a transcriptional program that includes upregulation of p21. p21 inhibits cyclin-CDK complexes, preventing progression of the cell cycle into S phase (Shaltiel et al., 2015).

ATR becomes activated during S phase and G2, signalling through the CHK1 kinase. In contrast to ATM, which is activated by DSBs, ATR is activated by ssDNA, specifically during replication stress where DNA polymerases become uncoupled from the CMG helicase (Byun et al., 2005). Replication stress can be caused by conflict with many of the DNA lesions discussed earlier (see 1.1.1) and is usually the mechanism by which DNA lesions become genomic instability and must therefore be tightly surveyed and regulated. ATR is recruited to ssDNA in an RPA-dependent fashion via its partner, ATRIP (You et al., 2002; Zou & Elledge, 2003). CHK1 is recruited and phosphorylated at multiple locations, relieving auto-inhibition of its C-terminal domain and then goes on to phosphorylate several nucleoplasmic and cytoplasmic substrates. (Oe et al., 2001). Important targets include the CDC25 phosphatases, which remove inhibitory phosphorylation marks on several CDKs. Thus, inhibiting CDC25 phosphatases arrests the cell cycle at the G1/S, intra-S, and G2/M (Shen & Huang, 2012). It is also through CDK inhibition that ATR-CHK1 signalling reduces replication origin firing and thus stabilises the replication fork, probably through maintaining the pool of essential factors that would be depleted in cases of excessive firing (Toledo et al., 2013).

Although the ATM and ATR pathways were considered to be independent, there is emerging evidence to suggest crosstalk between them (Weber & Ryan, 2015). The colossal number of proteins phosphorylated by ATM and ATR means they regulate several pathways that are too numerous to expand on in detail here.

1.2 The ubiquitin system

The ubiquitin system is a diverse signalling network that is critical to many cellular events. One of the earliest discoveries highlighting the importance of protein turnover as a way to regulate a cellular process was that of the ubiquitin-mediated degradation of cyclins involved in the cell cycle (Aaron Ciechanover et al., 1984; Finley et al., 1984; Glotzer et al., 1991). This was a revolutionary moment in the understanding of how cellular events are regulated. The importance of the ubiquitin system in regulating many fundamental biological processes have since been described, some of the earliest and most studied processes are transcription and the DNA damage response. Indeed, there is a wealth of knowledge of how ubiquitin modulates several DNA damage repair pathways, not least NER (Conaway et al., 2002; Harper & Elledge, 2007; Jackson & Durocher, 2013).

The ubiquitin system was initially discovered as a pathway to 'label' proteins for degradation by the covalent linking of a poly-ubiquitin chain (Hershko et al., 1980; K. D. Wilkinson et al., 1980). Research has since revealed a massively expanding capacity of the ubiquitin system to signal in a vast array of different ways. Through the covalent attachment of ubiquitin to target proteins, it can act as a signal much akin to phosphorylation (Hoege et al., 2002; Keith D. Wilkinson, 1999). Thus, ubiquitin functions as a diverse post-translation modification that controls many cellular processes via the more established proteasome-degradation pathway and non-canonical signalling.

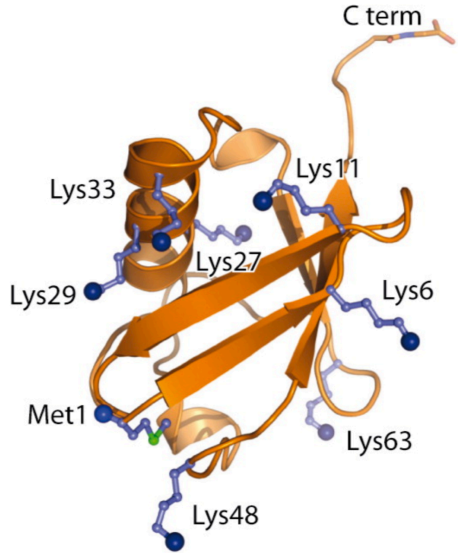


Figure 1-2 Structure of ubiquitin monomer
All lysines and Met1 side chains shown.
Adapted from (Komander & Rape, 2012)

1.2.1 The ubiquitin cascade

Ubiquitin is a small, 8.6 kDa protein of 76 amino acids, so named due to its ubiquitous presence in cells derived from all eukaryotes (Goldstein et al., 1975). Ubiquitin is usually conjugated to target proteins via its C-terminal glycine residue and the ϵ-amino group of a lysine residue in the target protein, termed ubiquitylation. While ubiquitylation of lysine residues is by far the most well studied, emerging evidence of ubiquitylation of serine, threonine and cysteine residues has been linked mostly to the endoplasmic-reticulum-associated protein degradation (ERAD) pathway. Ubiquitin's own seven internal lysines (K6, K11, K27, K29, K33, K48, K63) or its amino terminus are also used for conjugation of several ubiquitin monomers to form polypeptide chains (Yau & Rape, 2016). Homotypic ubiquitin chains are defined by the lysine residue linking them, as this is constant throughout the chain and defines the signal. For example, K48-linked chains are the most well studied and abundant chains which target proteins for degradation upon recognition by the 26S proteasome (Chau et al., 1989). K63-linked chains can signal proteins for degradation but have more frequently been found to act like scaffolds to bridge proteins, thus acting as signal transducers. In the repair of DSBs, for example, K63 ubiquitylation of histones recruits and enhances retention of 53BP1 and BRCA1 at DNA damage (Doil et al., 2009; Sobhian et al., 2007). The functions of other homotypic chains are much less well studied, but this is an ever-expanding area of research. Chain types are not limited to the homotypic type and mixed chains consisting of poly-ubiquitin linked by varying internal lysines have increasingly come under focus (Akutsu et al., 2016). Of course, with eight different linkage sites, oriented in different directions, and chains several ubiquitin moieties long, the structural repertoire becomes vast, underscoring the suitability of ubiquitin for signalling.

Ubiquitylation of proteins proceeds via a cascade of enzymes that catalyse the covalent linkage of free ubiquitin to a substrate lysine. The ubiquitylation system is a dynamic process that can be conceptualised in steps of writing, ubiquitylation; reading, recognition of and acting on the ubiquitin signal; and erasing, removal of

ubiquitin by de-ubiquitylating enzymes (DUBs), much like acetylation and methylation of histones, for example.

The ‘writing’ of ubiquitin to substrate proteins proceeds via three classes of enzymes: E1, E2, and E3. E1 are called activating enzymes as they interact with free ubiquitin to adenylate its C-terminus, which allows the formation of a thioester bond between the active site of the E1 enzyme and a cysteine residue of ubiquitin (A. Ciechanover et al., 1980; Haas et al., 1982; Hershko et al., 1980). The ubiquitin is then transferred to an E2 ubiquitin-conjugating enzyme, with ubiquitin again forming a thioester bond with a conserved cysteine in the active site (Scheffner et al., 1995). There are only 2 identified E1 enzymes in humans, 35 E2 enzymes, and over 1000 E3 ligases. Certain E2 enzymes will interact with only a subset of E3 ligases and E3 ligases too have defined substrates, thus specificity can be increasingly defined as ubiquitin moves down the cascade (Wijk & Timmers, 2010).

E3 ligases ultimately allow the transfer a thioester-bonded ubiquitin from an E2 enzyme to an isopeptide-bonded ϵ-amino group of lysine on target proteins. However, they do so via different mechanisms. Broadly, there are two families of E3 ligases, the Homologous to the E6AP Carboxyl Terminus (HECT) domain and Really Interesting New Gene (RING) domain family. The HECT E3 ligases transfer ubiquitin from an E2 enzyme to a cysteine residue in its own active site via a thioester bond and finally catalyse the isopeptide linkage of the ubiquitin to the lysine on the target protein. The last 60 amino acid residues of the HECT domain seem to specify its substrate, at least for *S. cerevisiae* Rsp5 (NEDD4 in humans) (H. C. Kim & Huibregtse, 2009). RING domain E3 ligases do not form a catalytic intermediate with ubiquitin, but instead define the substrate and act as a scaffold to ‘dock’ the E2-conjugated ubiquitin close to the target lysine where it is transferred (Passmore & Barford, 2004). The exact mechanism of how RING E3 ligases enhance the transfer of ubiquitin from E2 enzymes to lysine is not entirely clear, but it is likely that different mechanisms exist, depending on the E2 and RING E3’s involved (Dou et al., 2012; Koliopoulos et al., 2016; Plechanovov et al., 2012). Recently, a third family of E3 ligases have been defined called RING-in-between-RING (RBR) that possess both RING domains and an active-site cysteine like HECT E3’s (Dove & Klevit, 2017; Wenzel et al., 2011).

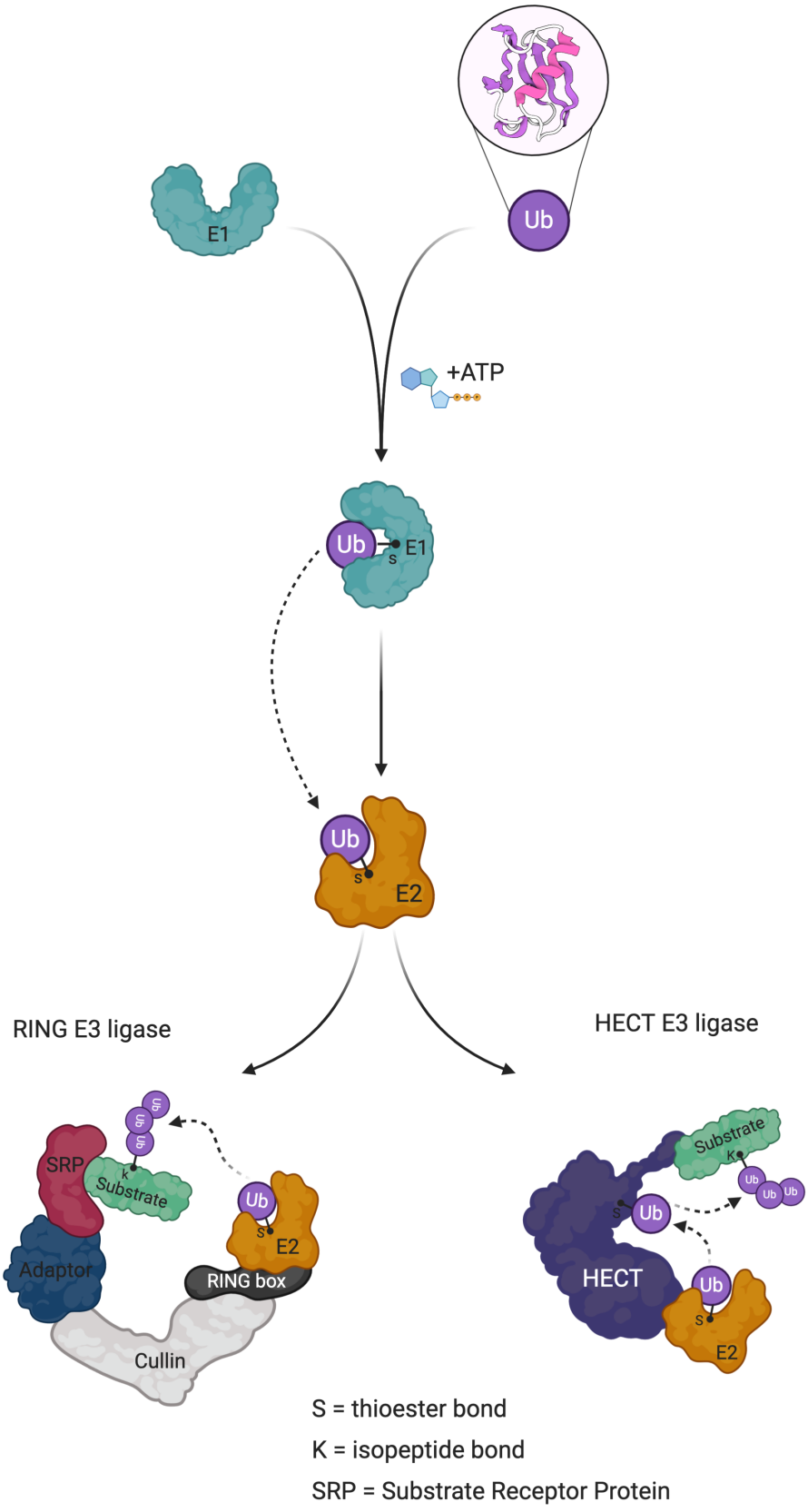


Figure 1-3 The ubiquitin cascade
Simplified overview of the ubiquitin cascade. See text for more details.

1.2.2 Ubiquitin proteasome pathway

The role of ubiquitin in the regulated degradation of proteins was the first, and is the most, studied observation of its function for which the Nobel Prize in Chemistry was awarded in 2004 (A. Ciechanover et al., 1980; Aaron Ciechanover, 2004; Etlinger & Goldberg, 1977; Hershko et al., 1980). The main mechanism of protein turnover is via the 26S proteasome, a behemoth protein factory of some 33 different proteins, totalling 2.5 MDa (Voges et al., 1999). Its structure comprises a 20S core particle (CP), which is found in all three domains of life, and in eukaryotes two 19S regulatory particles (RP) that cap either end of the CP. The RP is responsive for recognition and unfolding of substrates and presenting them to the CP which is cylindrically shaped and contains the proteolytic activity in its three central channels (Kunjappu & Hochstrasser, 2014).

It is now recognised that many different internal lysine linkages between ubiquitin can mark a protein for proteasomal degradation (except K63); however, the canonically recognised signal is the K48-linked chain (Kulathu & Komander, 2012). The ‘reader’ of degradative ubiquitin signals on proteins are a series of ubiquitin-binding domains (UBDs) within proteasome-shuttling proteins that transport target proteins to the proteasome for degradation. These proteins contain ubiquitin-like domains (UbLs) at their N-terminal and ubiquitin-associated domains (UBAs) at their C-terminal. The UBA domain recognises and binds polyubiquitin chains on target proteins, while the UbL domain interacts with the 19S RP complex of the proteasome to deliver the protein for degradation (Elsasser et al., 2002; Raasi et al., 2005). One of the best studied examples is the *S. cerevisiae* Rad23 protein (human homologues: HR23A and HR23B), which has two UBA domains through which it can interact with many proteins. Consequently, it is involved in many cellular processes from the cell cycle to NER, highlighting the importance of proteolytic degradation as a key regulator of distinct cellular events, not just in overall cellular homeostasis (Yokoi & Hanaoka, 2017). However, some studies have also implicated the 19S RP complex, which Rad23 interacts with, in proteolytic-independent functions in transcription and NER (Ferdous et al., 2002; Russell et al., 1999).

1.2.3 Ubiquitin as a PTM and signal transducer

While the initial focus of ubiquitylation research was on its role in targeting proteins for proteasomal degradation, it is now increasingly evident that ubiquitin's functions also extend to cell signalling. In particular, monoubiquitylation and K63-linked polyubiquitin chains have been best well characterised as PTMs involved in signal transduction. Of course, the existence of several different families of UBDs present on proteins involved in diverse cellular processes affirms ubiquitin as a way of modulating proteins functions. These UBDs can promote protein-protein interactions by binding other ubiquitylated proteins or modulate protein function by binding intramolecular ubiquitin moieties.

Examples of proteins containing functionally characterised UBDs include Vps9–a protein required in the yeast endocytic pathway–and the translesion synthesis DNA polymerase, Pol η .Vps9 was identified as containing a CUE (a type of UBD) domain that promotes its own monoubiquitylation by the Rsp5 E3 ligase and is important for modulating protein-interactions of Vps9 (Shih et al., 2003). Pol η contains a ubiquitin-binding ZnF (UBZ) UBD. Through its UBZ domain, pol η binds PCNA, which is monoubiquitylated upon DNA damage (Hoege et al., 2002; Kannouche & Lehmann, 2004). Inversely, pol η itself can be monoubiquitylated in its nuclear localisation signal (NLS), which inhibits its association with PCNA and hinders its aberrant activation. The authors also postulate that pol η might be autoinhibited by intramolecular binding of its UBZ with monoubiquitylation (Bienko et al., 2010).

K48-linked polyubiquitylation and proteasomal degradation of proteins can modulate pathways negatively by directly degrading enzymes, or positively by degrading an inhibitory protein. An example of the latter case is the control of the NF-κB transcription factor, which is kept inactive by being in a complex with IκBα, which masks NF-κB's NLS and sequesters it in the cytoplasm. Phosphorylation of IκBα by IKK brings about its polyubiquitylation and ultimate degradation, which allows for activation of NF-κB by disinhibition of its translocation to the nucleus (Karin & Ben-Neriah, 2000; Krappmann & Scheidereit, 2005). The activation of IKK kinase is itself bought about by polyubiquitylation of NEMO, an IKK regulatory

subunit, where the K63-linked chain serves as a docking surface to recruit activating proteins of IKK via their UBDs (Kanayama et al., 2004; H. Zhou et al., 2004). Here, the interdependence of ubiquitylation and phosphorylation, a well-established PTM in signal transduction, serves to highlight the importance of ubiquitylation as a key cellular signalling transducer. The single molecule ubiquitin is amenable to assembly in diverse topologies due to its many lysine linkages and can thus modulate signalling via sequestering of intramolecular UBDs, proteasomal degradation, and intermolecular recruitment of proteins.

Any signalling network needs 'erasers', as well as 'readers' and 'writers' if it is to precisely modulate the code. Indeed, deubiquitylating enzymes (DUBs) exist for precisely this function, approximately 100 of which have so far been identified in humans (Basar et al., 2021). There are two broad families, zinc-dependent JAMM metalloproteases and the papain-like cysteine proteases (Shrestha et al., 2014). Free ubiquitin itself is produced from polyubiquitin precursor proteins, by DUB-catalysed hydrolysis and the ubiquitin pool is maintained by hydrolysis of eupeptide and isopeptide of ubiquitin attached to other proteins. DUBs can cleave ubiquitin chains from the end (exo cleavage) or internally, between ubiquitin linkages (endo cleavage), as well as base cleavage of the Gly-Gly motif between ubiquitin and substrate proteins (or of free linear chains) (Mevissen et al., 2013; Reyes-Turcu et al., 2006). This activity is used by DUBs to selectively digest polyubiquitin chains on proteins to remove the signal deposited by ubiquitin. In the case of K48-linked chains this can, of course, serve to nullify degradation-signalling and stabilise target proteins and thus their activity in a pathway. K63-linked chain digestion could hinder protein recruitment and ablate signal transduction. This, of course, then raises the question of how DUBs themselves are regulated to stop unwanted erasing of ubiquitin signalling. Fascinatingly, many DUBs themselves contain UBDs, often ubiquitin-interacting motifs (UIMs) that impart substrate-linkage selectivity, or associate with adapter proteins that can either enhance or inhibit DUB activity (Komander et al., 2009). For instance, the ubiquitin-specific protease USP5 undergoes an allosteric conformational change upon binding to free ubiquitin that enhances its enzymatic activity (Reyes-Turcu et al., 2006).

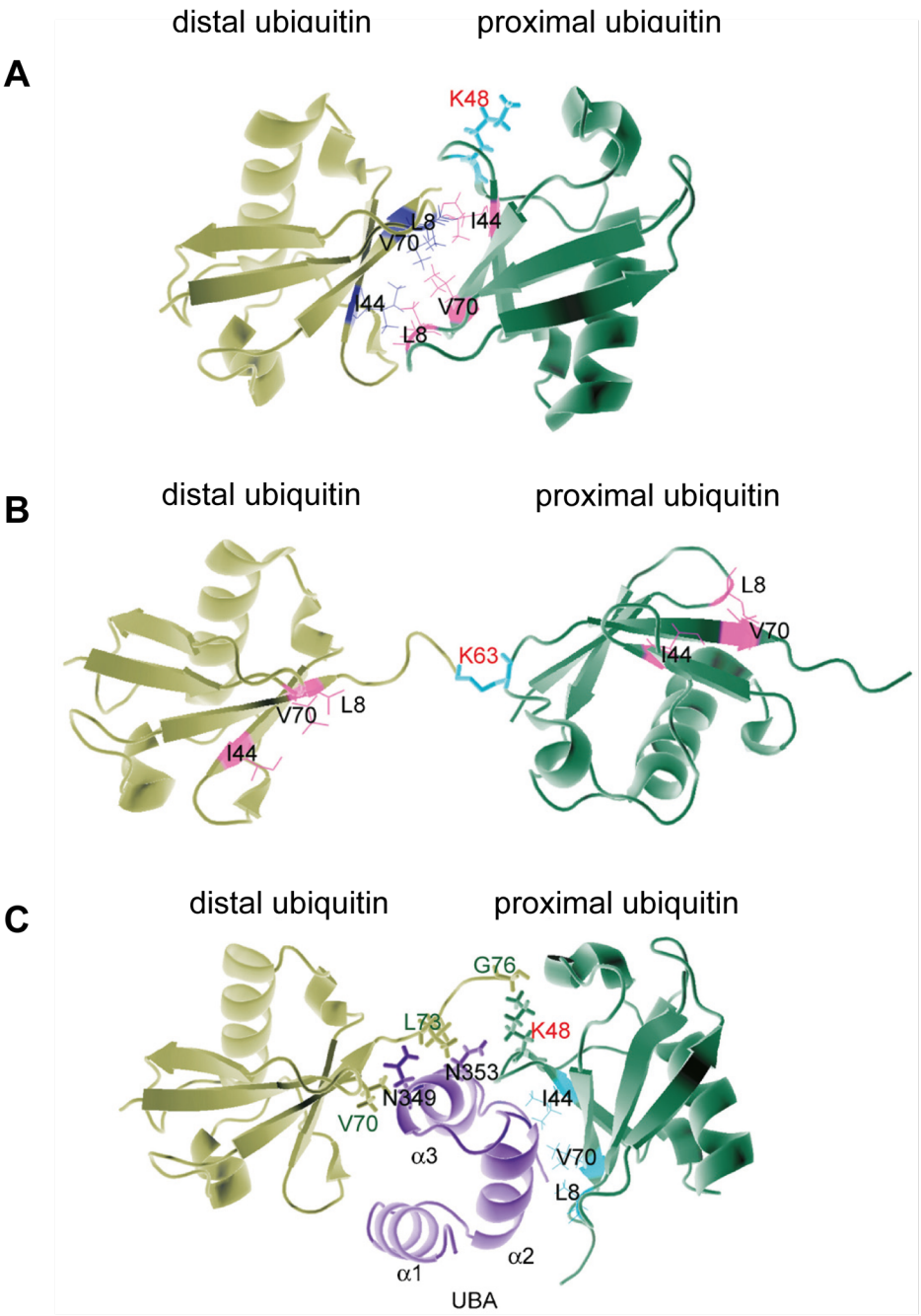


Figure 1-4 Structure of K48 and K63 ubiquitin linkages
A L48-linked di-ubiquitin mainly in a closed conformation. The side chains of hydrophobic patch residues (Leu8, Ile44 and Val70 in purple and pink) are buried between the two packed ubiquitin moieties.
B K63-linked di-ubiquitin adopts an open conformation, in which the hydrophobic patch on each ubiquitin monomer can independently interact with a ubiquitin-binding domain.
C Interaction of a Rad23 UBA domain with K48-linked di-ubiquitin. Note that the hydrophobic patch on the proximal ubiquitin molecule binds helix 2 of the UBA, whereas helix 3 of the UBA interacts with the C-terminal of the distal ubiquitin as well as the linker region.
Figure adapted from (W. Li & Ye, 2008).

1.2.4 Ubiquitin-like modifiers

While ubiquitin is a highly conserved protein among eukaryotes and its importance has been recognised since the 1980's, divergent ubiquitin-like proteins (UBLs) have more recently come under light as important PTMs in their own right. They all comprise a β-grasp fold that is characteristic of ubiquitin, hence 'ubiquitin-like'. Those discovered so far are SUMO, ISG15, NEDD8, HUB1 and ATG8 and their conjugation to proteins proceeds through an enzymatic cascade much akin to ubiquitin for which they have their own unique E1, E2, and E3 enzymes (Van Der Veen & Ploegh, 2012). They typically act as PTMs and signal transducers rather than targeting proteins for degradation. However, while they are not directly recognised by the proteasome, in several cases they can promote K48 polyubiquitylation of protein substrates to indirectly bring about their degradation. (Martin et al., 2009; Tatham et al., 2008).

SUMO proteins exist as four homologues in humans (SUMO 1–4), which are conjugated to lysines just like ubiquitin. SUMO2 and 3 are 97% identical and thus can't be differentiated by antibodies, so are usually referred to as SUMO2/3. SUMO1 is approximately 50% identical to SUMO2, and SUMO4 approximately 87% similar to SUMO2, although it is likely not normally expressed (Gareau & Lima, 2010). Therefore, sumoylation is mechanistically typically thought of as consisting of two modifications: either SUMO1 or SUMO2/3. Approximately 70% of SUMO substrates contain a consensus motif for sumoylation (ΨKx[DE]; Ψ denoting a large hydrophobic residue and K being the acceptor lysine) that is recognised by the single E2 enzyme, UBC9. SUMO interaction motifs (SIMs) are similar to UBDs in that they are domains in proteins that recognise and bind to SUMO chains, mediating signalling transduction (Song et al., 2004; K. A. Wilkinson & Henley, 2010).

NEDD8 is the UBL most similar to ubiquitin with 58% identity in primary sequence. Its importance is highlighted by the fact that NEDD8 knockouts are lethal in most eukaryotes except *S. cerevisiae* (Rabut & Peter, 2008; Watson et al., 2011). Like ubiquitin, it has its own conjugating enzyme cascade in humans: a single E1 (NAE),

two E2 enzymes (UBC12 and UBE2F) and several E3 enzymes. All of the NEDD8 E3 enzymes that have been identified so far also function as ubiquitin E3 ligases (Enchev et al., 2015). One example is the RBX1 protein, a subunit of Cullin-RING ligases (CRLs), the largest family of E3 ubiquitin ligases due to their multisubunit modularity. RBX1 can complex with the NEDD8 E2, UBC12 and facilitate neddylation of the cullin subunit of CRLs (D. T. Huang et al., 2009; Santonico, 2019). Neddylation of CRL's induces a confirmational change that 'activates' them, facilitating their ubiquitin E3 ligase activity (Duda et al., 2008; Lydeard et al., 2013). As RBX1 also binds ubiquitin E2s to promote ubiquitylation of protein substrates, it acts as both a NEDD8 and ubiquitin E3. Structural and mutational analysis of RBX1 revealed how it differentially regulates NEDD8~UBC12 and ubiquitin~UBCH5 so as to specifically neddylate and not ubiquitylate the cullin subunit (Scott et al., 2014). Thus, CRL's serve both as a NEDD8 E3 ligase, a NEDD8 substrate, and consequently a ubiquitin E3 ligase. A specific CRL, CRL4A is intimately involved in nucleotide excision repair and this is detailed further in the later section on NER.

1.3 Nucleotide excision repair

Nucleotide-excision repair (NER), like BER (1.1.3.1), is a DNA repair mechanism that relies on the enzymatic excision of damaged DNA and gap filling for repair. Crucially though, NER uses the same pair of enzymes to excise all DNA lesions by incision of the phosphodiester backbone. Therefore, rather than a plurality of enzymes—like glycosylases—that directly recognise specific lesions, NER relies on a common feature of all of its substrate lesions, the formation of a bulky distortion of the double helix, rather than the lesions itself. NER operates two distinct modes: global genome NER (GG-NER) that directly recognises distortions in the DNA helix; and transcription-coupled NER (TC-NER) that recognises RNAPII stalled by bulky lesions, thus operating in transcribed areas of the genome and influencing RNA synthesis. Only the initiating steps of lesion recognition differ between the two pathways, whereas they converge on common steps of incision and gap-filling.

Evidence was provided for thymine dimers (UV-induced CPDs, which are bulky lesions) stalling DNA synthesis in *E. coli* in 1963 by Setlow and colleagues, who

also hypothesised of a repair pathway for removal of these dimers by local DNA synthesis (Setlow et al., 1963). It was the excision of these dimers—which, importantly, were resistant to photolyase enzymes (see 1.1.3)—coinciding with the resumption of DNA synthesis, which was shown in a subsequent publication by the same group of researchers and provided the first formal proof of 'dark repair', as distinguished from 'light repair' (EPR) (Setlow & Carrier, 1964). A few months later, Hanawalt and colleagues revealed that 5-bromouracil (a thymine analogue) was incorporated into short segments of a single DNA strand immediately following UV-irradiation. They proposed that the excision of UV-produced thymine dimers, followed by short, non-conservative replication was a new mode of repair (Pettijohn & Hanawalt, 1964). These same results were quickly observed in mammalian cells too (Rasmussen & Painter, 1964). The mechanism remained unclear. Over the course of the next several decades, the mechanisms of NER from recognition to incision, excision, and gap-filling have been greatly studied and elucidated in *E. coli*, *S. cerevisiae*, and mammals (Prakash & Prakash, 2000; B. Van Houten, 1990; Wood, 1997).

1.3.1 Cockayne syndrome and Xeroderma pigmentosum

Cockayne syndrome (CS) and Xeroderma pigmentosum (XP) are both autosomal recessive diseases resulting from deleterious mutations in the genes essential to the two NER pathways. XP was the first disease of defective NER (and of repair in general) that was described, in that it was shown that cells from XP patients were deficient in excision of UV dimers and repair synthesis (Cleaver, 1968; Setlow et al., 1969). There is clinical heterogeneity among XP patients, but common phenotypes are extreme sensitivity to sunlight, pigment changes of the skin, an increased incidence of skin cancers, and a minority of patients suffer neurological abnormalities (Alan R. Lehmann et al., 2011). There are 8 complementation groups of XP: the genes XPA through XPG, which encode a variety of different protein products all required for the core NER reaction to repair bulky DNA lesions (Bowden et al., 2015). XPV is the 8th complementation gene; encoding Pol η, a low fidelity translesion polymerase involved in replicating DNA containing bulky DNA lesions (Johnson et al., 1999; Masutani et al., 1999). Consequently, XP patients

can possess genotypes that render both pathways of NER defective (those acting downstream of lesion recognition), while others are still competent for TC-NER (XPC and XPE mutations). XP patients are estimated to experience a 10,000-fold increased risk of non-melanoma skin cancer and a 2,000-fold increased risk of melanoma in just the first two decades of life. There is also an approximately 50-fold increase in various internal carcinomas, highlighting the role of NER in repairing damage from endogenous damaging agents as well as UV (Bradford et al., 2011). Of the 20–30% of patients who present with neurological degeneration, the majority of the mutations are in the XPD gene, which is part of the basal transcription factor TFIIH, signifying this may be due to additional effects on transcription as well as NER (Bradford et al., 2011).

Cockayne syndrome is a defect specifically related to the TC-NER pathway initiated by the CSA and CSB proteins, encoded by the *ERCC6* and *ERCC8* genes, respectively, hence 2 complementation groups exist. It was first described in two siblings who presented to Great Ormond Street hospital in London, UK (Cockayne, 1936). A seminal study reviewed 140 patients with CS that described many of the clinical manifestations of the disease, which presents as a progeroid syndrome (Nance & Berry, 1992). Clinical presentation and severity is extremely heterogenous amongst individuals, but common symptoms are: cachectic dwarfism, extreme sensitivity to sunlight, neurological impairments, hearing loss, and cardiovascular and renal complications. Intriguingly, in contrast to XP, CS patients do not display an increased incidence of cancer (Karikkineth et al., 2017). Cockayne syndrome was initially stratified into 3 categories based on the severity of the disease (Nance & Berry, 1992). Patients can often first present to clinicians at very different stages of the disease as it progresses at very different rates, with many patients dying in childhood, but with some progressing into adulthood. However, many clinicians believe patients exist on a spectrum and the original categories are not representative of different forms of disease (Laugel, 2013; B. T. Wilson et al., 2016). There is no correlation between underlying mutations of CSA or CSB and different CS categories. UV sensitivity syndrome (UVSS) and cerebro-oculo-facio-skeletal (COFS) are two other diseases that result from mutations in CSA, CSB, or UVSSA in the case of UVSS, and CSB, XPG, or XPD (as well as undetermined genotypes) in the case of COFS. They were originally classified as

distinct syndromes due to UVSSA patients having very mild CS symptoms and COFS patients having very severe CS symptoms. However, clinicians now believe they constitute the range of CS phenotypes, just at the extreme ends (Laugel, 2013). UVVS and COFS cells have also been shown to be deficient in TC-NER (J. M. Graham et al., 2001; Itoh et al., 1994; Nakazawa et al., 2012; Schwertman et al., 2012; Zhang et al., 2012)

1.3.2 Global genome NER and general NER mechanisms

1.3.2.1 Lesion recognition

NER repairs a vast array of DNA lesions that are recognised through their capacity to distort the double helical structure of DNA, creating a bulge, and are referred to as ‘bulky’ lesions. Probably the most commonly studied bulky lesions—which are also likely the most prevalent lesions repaired by NER—are the UV-induced CPD and 6–4PPs, with CPDs being the most prevalent. This is especially true in an environmental context as terrestrial sunlight is concentrated in longer UV wavelengths, which favour CPD formation (Besaratinia et al., 2011). As well as distortions in the DNA helix, some disruption of Watson–Crick base pairing must occur for lesions to be suitable substrates. CPDs cause less helical distortion than 6–4PPs and are therefore weaker substrates, but their inclusion opposite artificial mismatched bases increased their excision three to four fold, highlighting the property of bipartite discrimination of lesions in NER (Hess et al., 1997; Mu et al., 1997).

The protein that recognises NER-substrate lesions in the GG-NER pathway in humans is XPC, which co-purifies in complex with RAD23B and CETN2 (Araki et al., 2001; Masutani et al., 1994). It is the initiator of NER in the GG-NER pathway, where it is the first protein recruited to a substrate DNA lesion (Sugasawa et al., 1998). XPC is the only factor exclusive to GG-NER as lesion recognition is by a distinct mechanism in TC-NER, and this pathways thus operates normally in XPC deficient cells (Shivji et al., 1994; J Venema et al., 1991; Jaap Venema et al., 1990). The crystal structure of the yeast homologue of XPC, Rad4, bound to a CPD lesion revealed significant insights into the mechanism of its recognition of bulky

lesions. Rad4 binds the undamaged DNA strand via N-terminal TGD and BHD domains and inserts a β-hairpin domain into the DNA duplex, flipping out the two damaged bases, and binding and stabilising the undamaged bases opposite (Min & Pavletich, 2007). Rad4/XPC binds to the phosphate backbone of the undamaged DNA strand and probes the DNA in an ATP-independent manner. Thus, it can only adopt a confirmation where the β-hairpin intrudes the DNA duplex when the free energy is low enough from disrupted Watson–Crick base pairing to allow base flipping (Camenisch et al., 2009; Chen et al., 2015). A study using temperature-jump spectroscopy showed that Rad4 probes and interrogates DNA by a twisting mechanism at microsecond timescales and is stabilised on damaged DNA by greater contacts with the β-hairpins that allow it to flip out damaged bases on the millisecond timescale (Velmurugu et al., 2016).

However, the above studies were either done on artificial templates containing CPDs opposite mismatched bases (not physiological), or large, bulky lesions. The question then is how XPC can distinguish a genuine CPD that doesn't disrupt Watson–Crick base pairing as strongly (Jing et al., 1998; H. J. Park et al., 2002). The UV-DDB complex may be the answer to that question. It is composed of DDB1 and DDB2/XPE and has a higher affinity for CPD lesions than XPC (Chu & Chang, 1988; Fujiwara et al., 1999; Wittschieben et al., 2005). Unlike XPC, which binds to the flipped-out undamaged bases and doesn't contact the CPD bases, DDB2 inserts itself into the DNA duplex and flips the CPD bases into a binding pocket. It can also accommodate 6–4PPs in nucleosomes and abasic sites, indicating that it recognises common distortions in the DNA structure rather than lesions directly (Fischer et al., 2011; Osakabe et al., 2015; Scrima et al., 2008). Via DDB1, DDB2 also serves as a dedicated substrate receptor (DCAF) of the DDB1-CUL4A-RBX1 (CRL4A) RING E3 ubiquitin ligase complex (Groisman et al., 2003). As the substrate determinant of the CRL4A complex, DDB2 can polyubiquitylate itself, bringing about its own proteolytic degradation. $CRL4^{DDB2}$ also polyubiquitylates XPC upon UV-induced DNA damage, but this does not induce its degradation. Instead, it stabilises XPC at the site of the damage (Sugasawa et al., 2005). This occurs concurrently with dissociation of RAD23B, which initially stabilises it before damage binding (Bergink et al., 2012; Ng et al., 2003; Ortolan et al., 2004). XPC is also sumoylated, which further stabilises it at damage and recruits a SUMO-

targeted E3 ubiquitin ligase, RNF111, that polyubiquitylates XPC with K63-linked chains that ultimately brings about its degradation, assisted by the VCP segregase (Poulsen et al., 2013; Puumalainen et al., 2014; Van Cuijk et al., 2015; Q. E. Wang et al., 2005). The opposing fates bought about by the concurrent ubiquitylation of DDB2 and XPC at sites of damage by the same ligase is thought to constitute a 'handover' of the damage from UV-DDB to XPC to facilitate recruitment of other NER factors for downstream incision events. UV-DDB localises to sites of both CPDs and 6–4PPs in cells, but XPC only localises to 6–4PPs in its absence, serving to highlight the importance of UV-DDB for NER-mediated repair of CPDs, one of the most common DNA lesions (Fitch et al., 2003).

1.3.2.2 Lesion verification and DNA unwinding

While the above mechanisms are unique to the GG-NER pathway and TC-NER has its own unique mode of lesion recognition, the two pathways then converge, and the steps outlined below are common in both NER pathways.

XPC next recruits TFIIH to damage via a direct interaction, to open DNA and verify the lesion for GG-NER, while TFIIH is recruited via different mechanisms for TC-NER (Bernardes de Jesus et al., 2008; Yokoi et al., 2000). Further lesion verification may be needed by TFIIH due to XPC's general affinity for single-stranded DNA, usually caused by lesion-destabilisation of the helix, rather than the lesion directly, which could lead to aberrant activation of NER (Buterin et al., 2002; Camenisch et al., 2009). The opening of DNA around the lesion may facilitate further binding sites for downstream factors, such as the NER endonucleases. TFIIH is a large, multi-protein complex essential for both NER and as a basal transcription factor for transcription initiation by RNA polymerase II, but it adopts slightly different compositions and conformations in both processes, namely the CAK complex containing Cdk7, which is not required for NER, (Kokic et al., 2019; Svejstrup et al., 1995). The two subunits of central importance to NER are TFIIH's helicases, XPB and XPD, which have opposite directionality; XPB is 3′ → 5′, and XPD is a 5′→ 3′ DNA helicase (Schaeffer et al., 1993). While the helicase activity of XPB was initially thought to open DNA around the lesion, akin to melting of

promoters during initiation, it has been shown to, surprisingly, be dispensable for NER, although its ATPase activity is required for TFIIH binding (Coin et al., 2007). In contrast, the helicase activity of XPD is essential for NER and is activated by the p44 regulatory subunit upon DNA binding, which allows XPD to open the DNA surrounding the lesion (Coin et al., 1998; Oksenych et al., 2009). When the XPD helicase runs into a damage on the strand it is translocating along, its helicase activity is inhibited and TFIIH is stabilised at the damage site (Bennett Van Houten et al., 2016; Kuper et al., 2014; Mathieu et al., 2010; Sugasawa et al., 2009).

XPA is a non-enzymatic protein that is essential for NER, but its mechanism hasn't been completely clarified. However, it is known that it also binds to the damaged strand and facilitates TFIIH damage verification, while the ssDNA binding protein RPA binds the non-damaged strand (Jones & Wood, 1993; Krasikova et al., 2010, 2018), presumably to stabilise the denatured region for incision. XPA enhances the helicase activity of TFIIH, while it 'searches' for damage and then enhances TFIIH stalling at damage (C. L. Li et al., 2015). RPA and XPA may function to stabilise the DNA bubble opened by TFIIH and demarcate its 5′ edge allowing recruitment of downstream endonucleases for incision (Camenisch et al., 2006; Kokic et al., 2019; Matsunaga et al., 1996).

1.3.2.3 Dual incision and excision of lesion-DNA

There are two endonucleases that function in NER, XPF and XPG, which make cuts 5′ and 3′ of the lesion, respectively. XPG is the first nuclease recruited to damage through interaction with TFIIH, which localises it 3′ of the damage (De Laat, Appeldoorn, Jaspers, et al., 1998; De Laat, Appeldoorn, Sugasawa, et al., 1998). Recruitment of XPG finalises the formation of the pre-incision complex as initially it functions as a structural scaffold, lying dormant until the recruitment of XPF, upon which incisions are made. Indeed, nuclease mutants of XPG are still competent for XPF recruitment and cleavage 5′ (by XPF), but not 3′, of a 6–4PP lesion (Constantinou et al., 1999; Wakasugi et al., 1997). No incision, 5′ or 3′, takes place in cell extracts with catalytic dead XPF, indicating that it only serves as a nuclease and operates downstream of XPG (Staresincic et al., 2009; Tapias et al.,

2004). This study also showed using an *in vitro* system that XPF makes the first incision 5′ of the damaged stand, leaving a 3′-OH substrate that can facilitate repair synthesis of DNA even without XPG incision. Recruitment of replication machinery was seen in mutant XPG cells but not mutant XPF cells, again indicating that XPF incision occurs first (Staresincic et al., 2009). This is consistent with data that show XPG binds to the DNA bubble at the ss/dsDNA junction, but only cuts substrates with 3′ ssDNA overhangs (E. Evans, Fellows, et al., 1997; E. Evans, Moggs, et al., 1997; Hohl et al., 2003). The completion of dual incision allows for the release/excision of the ~30mer lesion-containing DNA strand with TFIIH still bound to it (Kemp et al., 2012).

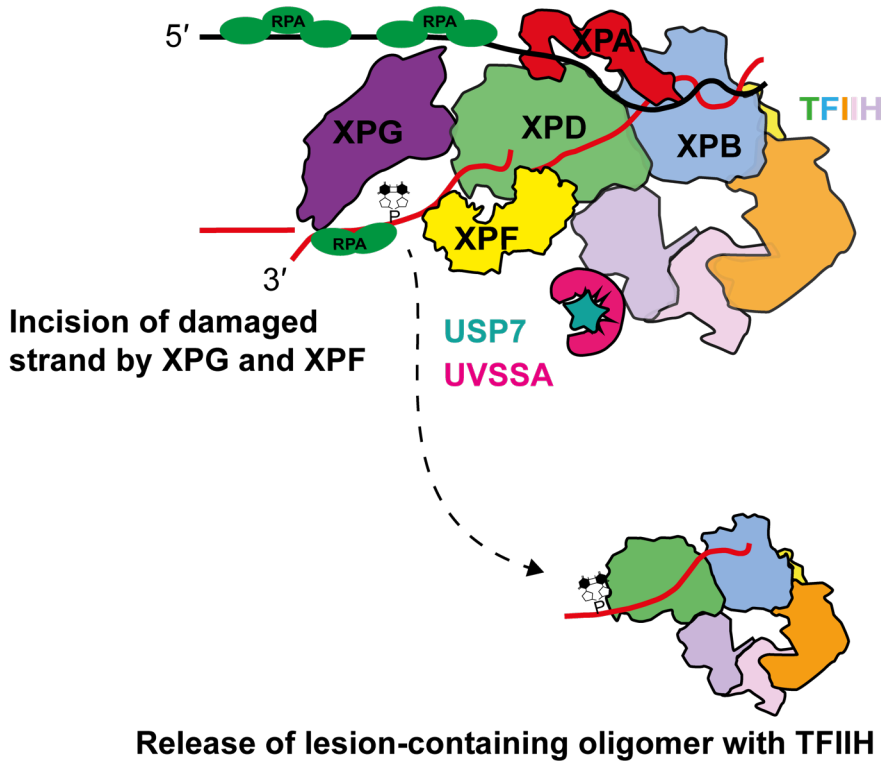


Figure 1-5 Schematic of damaged strand incision and removal
After NER is initiated by recognition of the lesion there is a 'handover' to TFIIH by XPC in GG-NER or UVSSA in TC-NER. TFIIH uses its helicase XPD to verify the lesion and position the endonucleases XPG and XPF for incision of the damaged strand 3′ and 3′ of the lesion, respectively. A ~25 nt oligomer is then excisied bound by TFIIH.

1.3.2.4 Gap-filling repair synthesis

Ultimately, the goal of any DNA repair mechanism is to return damaged DNA to its original form with the correct nucleotide sequence. In the case of NER, direct reversion of damaged bases isn't achieved like in BER, but rather a patch of DNA containing the offending lesion is excised, as discussed. Replacement of normal nucleotide sequence is then achieved by repair synthesis using the undamaged strand as a template.

The entire mammalian NER reaction has been reconstituted with purified proteins *in vitro,* which identified the minimally required replicative proteins as being Pol δ or Pol ϵ, PCNA, replication factor C (RFC), and DNA ligase I (Aboussekhra et al., 1995; Araújo et al., 2000; Mu et al., 1995; Wood et al., 1988). However, in cells this process is likely more complex, not least given there are at least 14 DNA polymerases in humans (Loeb & Monnat, 2008). It was shown that DNA ligase I was required in proliferating cells, while DNA ligase IIIα was employed in quiescent cells (Moser et al., 2007). The product of the XPV gene was shown to encode a translesion DNA polymerase, Pol η (Johnson et al., 1999; Masutani et al., 1999). Patients with mutations in XPV displayed phenotypes similar to patients with mutations in other XP proteins, so it was surprising to find that XP-V patient cell lines displayed normal NER, but had abnormal post-UV DNA synthesis (A. R. Lehmann et al., 1975). Depletion of the translesion polymerase, Pol κ, renders cells UV-sensitive and reduces, but doesn't ablate, NER (Ogi & Lehmann, 2006). It was further shown that Pol κ is responsible for up to 50% of repair synthesis in cells (Ogi et al., 2010). Translesion polymerases belong to the Y family of DNA polymerases and are of lower fidelity than canonical replicative polymerases due to their larger active site. But this also allows them to synthesise through lesions that would block the high-fidelity replicative polymerases (Loeb & Monnat, 2008). However, translesion polymerases (η and κ) should hypothetically not be involved in the NER reaction after a lesion is excised, as they are employed to synthesise through bulky lesions. This fits with the observation that cells lacking functional Pol η carry out normal NER. But knockdown of Pol κ reduced post-UV repair synthesis by 50% in cells, indicating a vital, but still poorly understood role for translesion polymerases in response to UV lesions. It could be that only 50% of UV lesions are

excised and repaired by the canonical NER pathway, while the rest may be replicated at low fidelity if not repaired in time, particularly in S-phase. This dependency for Pol κ may be exaggerated due to the high doses of UV used in the laboratory that don't reflect environmental exposure, so that the high number of lesions may saturate the canonical NER pathway. It could also be that translesion polymerases are better suited to the unique substrates posed by a TFIIH bound ~30nt bubble and single 5′ incision on one strand. In any case, a role for several polymerases of different fidelities in repair synthesis is clearly necessary in the response to UV damage.

1.3.3 Transcription-coupled NER

While the mechanisms described above account for lesion recognition in GG-NER and downstream steps are common to both pathways, the mechanism of lesion recognition in TC-NER is unique, more complex, and less well understood. While unscheduled DNA synthesis following UV and the release of thymine dimers from DNA—the discovery of NER—was observed in the 1960's (Pettijohn & Hanawalt, 1964; Setlow et al., 1963; Setlow & Carrier, 1964), it wasn't until over 20 years later that the first evidence of TC-NER was observed (Bohr et al., 1985; I. Mellon et al., 1986). In these studies, the researchers found that repair of CPDs was much more efficient inside a human gene than outside it. Further work used strand-specific probes against the two DNA strands, which showed that the rapid repair took place only in the transcribed strand and not in the non-transcribed strand in mammalian cells (Isabel Mellon et al., 1987). Repair was also enhanced by selective induction of transcription in a bacterial operon proving this phenomenon wasn't simply due to open chromatin (Isabel Mellon & Hanawalt, 1989). In yeast, RNAPII temperature-sensitive mutants and inhibition of transcription with α-amanitin abolished strand-specific repair and definitively proved the central role of RNAPII in the newly discovered TC-NER pathway (Christians & Hanawalt, 1992; Sweder & Hanawalt, 1992). Together, these studies showed that TC-NER is a highly conserved process, from prokaryotes through to higher eukaryotes. The biochemical interrogation of TC-NER was first studied using bacterial proteins, as this was the simplest and, at the time, the most established — bacterial NER having been

reconstituted *in vitro* in the 1980's (A. Sancar et al., 1981; Aziz Sancar & Rupp, 1983; Thomas et al., 1985; Yeung et al., 1983).

It was hypothesised that transcription complexes stalling at bulky lesions might constitute the initiation of TC-NER. Testing this hypothesis using *in vitro* systems of bacterial NER confirmed that RNAP is stalled by CPDs, but surprisingly the addition of RNAP to the reaction actually inhibited repair of the transcribed strand and favoured repair of the non-transcribed strand (C. P. Selby & Sancar, 1990). The clue to solving this problem was a missing factor, the mutant of which had been unknowingly identified decades earlier (Witkin, 1966). The mutation frequency decline (*mfd*) mutant strain of *E. coli* gives rise to increased mutations after UV, which were even shown to have strand bias, but neither the gene(s) nor the protein(s) responsible had been identified (Bockrath & Palmer, 1977; George & Witkin, 1974, 1975). Just a few months after their initial paper showing RNAP inhibited TC-NER, Selby and Sancar partially purified what they called transcription repair coupling factor (TRCF), which resolved the conundrum of RNAP-inhibition of repair and allowed reconstitution of bacterial TC-NER (Christopher P. Selby & Sancar, 1991). After coming across Witkin's papers they realised that her *mfd* mutants may have mutations in the gene encoding TRCF and proved that, indeed, only *mfd* cell extracts complemented with TRCF could carry out TC-NER *in vitro* (Christopher P. Selby et al., 1991).

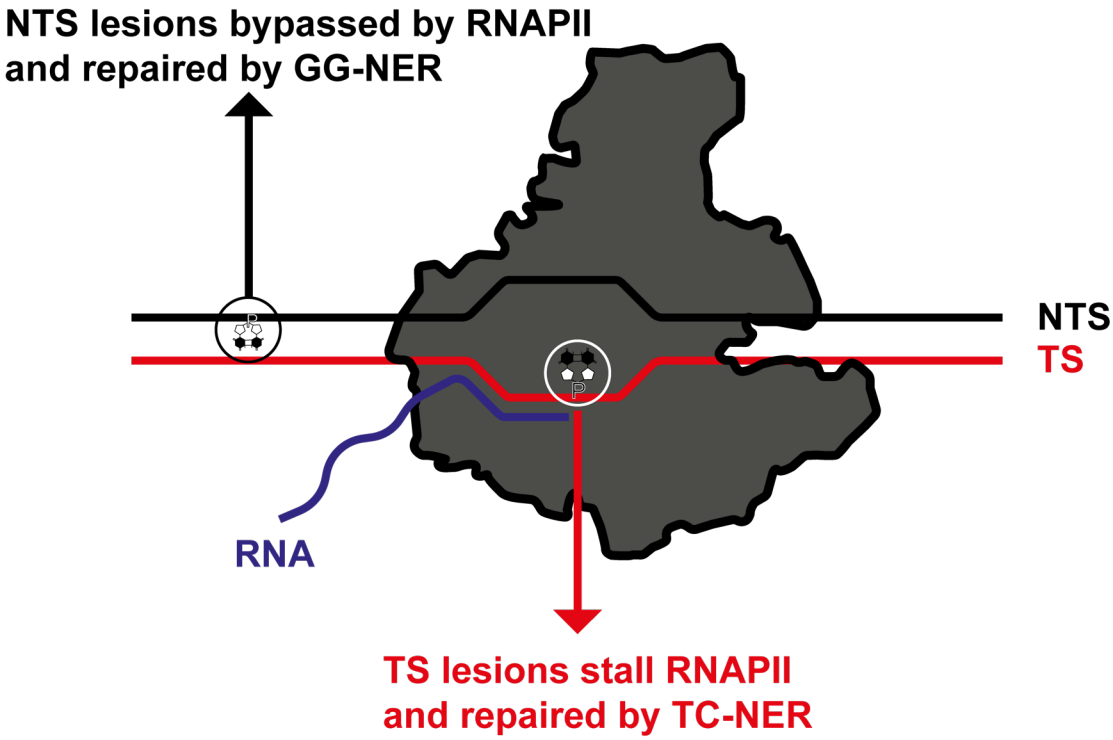


Figure 1-6 RNAPII stalling at lesions on the transribed strand triggers TC-NER
Schematic of RNAPII transcribing along a portion of DNA. Lesions on the non-transcribed strand (NTS) do not stall RNAPII as they do not enter its active site, whereas lesions on the transcribed-strand (TS) stall the polymerase as they enter the active site and block NTP polymerisation of the 5′ end of the RNA. This stalled RNAPII is recognised by mfd/Rad26/CSB to initiate TC-NER.

The Mfd/TRCF protein is an ATP-dependent DNA translocase that rescues arrested RNAP complexes by pushing RNAP forward and likely ejecting it from DNA (Howan et al., 2012; J. S. Park et al., 2002). Structural studies of Mfd/TRCF suggest it undergoes major conformational changes upon binding stalled RNAP and ATP hydrolysis, which allows Mfd to push against RNAP trapped at a damage and inserts itself into the RNAP clamp, prying it open and releasing RNAP (Shi et al., 2020). Such detail has not been uncovered for the functional homologues of Mfd: Rad26 in *S. cerevisiae* (Van Gool et al., 1994), and CSB in humans (C Troelstra et al., 1990; Christine Troelstra et al., 1992). Given the divergence of these proteins, the presence of extra domains, as well as more general complexity of the entire TC-NER process in eukaryotes (not least due to PTMs), an overview of the literature of these proteins deserves more detail, particularly given the experimental focus of this thesis.

1.3.4 Molecular mechanisms of Cockayne syndrome proteins in TC-NER

In contrast to prokaryotes, no complete reconstitution of human or yeast TC-NER has been accomplished with solely purified proteins in eukaryotes. However, the dual incision reaction has been reconstituted *in vitro* with RNAPII stalled at cisplatin damage on DNA templates, which was dependent on TFIIH, XPF, XPG, XPA, and RPA and was stimulated by CSB (Lainé & Egly, 2006). However, the recruitment of TFIIH to damage-stalled RNAPII was independent of CSB, in contrast to other studies *in vitro and in vivo* (Fousteri et al., 2006; Sarker et al., 2005; Dean Tantin, 1998; Tijsterman et al., 1997; Tu et al., 1997; van der Weegen et al., 2020). Furthermore, there was no requirement for CSA, UVSSA, or XAB2, which are essential for TC-NER *in vivo* (Henning et al., 1995; Nakatsu et al., 2000; Zhang et al., 2012). This highlights the extra complexity of human compared to bacterial TC-NER and points to additional, unidentified factors needed for the repair of RNAPII-stalled lesions. In fact, while CSB could correct the TC-NER defect in cells, and is recruited to transcribing RNAPII *in vitro,* it was unable to remove RNAPII stalled at a lesion from the DNA template (Christopher P. Selby & Sancar, 1997a; D Tantin et al., 1997).

The TC-NER defect in human Cockayne syndrome cells comprises two complementation groups, which display sensitivity to UV and a failure to recover RNA synthesis following shutdown after UV irradiation (Mayne & Lehmann, 1982; Schmickel et al., 1977; Van Hoffen et al., 1993; B. Venema et al., 1990). Interestingly, the *ERCC6* gene encoding the human functional homologue of Mfd, CSB, was actually cloned (C Troelstra et al., 1990) and CSB shown to complement the TC-NER defect in CS-B cells (Christine Troelstra et al., 1992) at around the same time as Mfd (Christopher P. Selby et al., 1991; Christopher P. Selby & Sancar, 1991). The *ERCC8* gene, encoding the CSA protein, was found to complement the second group of Cockayne syndrome cells a few years later (Henning et al., 1995).

CSB is a large 1493 amino acid, 168 kDa protein belonging to the Swi2/Snf2 family of DNA helicases/translocases and contains 7 helicase domains (Christine Troelstra et al., 1993). However, CSB possesses no helicase activity but instead

uses its ATPase activity to work as a 3′ → 5′ DNA translocase (Christopher P. Selby & Sancar, 1997a). It can remodel chromatin, at least *in vitro* (Cho et al., 2013; Citterio et al., 2000; J. Y. Lee et al., 2017). Early evidence from studies of CSB hinted that it might be involved in basal transcription in the absence of any DNA damage. Indeed, a fraction of CSB was shown to copurify with RNAPII and associate with RNAPII *in vitro* in the absence of DNA damage (D Tantin et al., 1997; Van Gool et al., 1997). Experiments *in vitro* also showed that CSB enhanced RNAPII elongation rate (Christopher P. Selby & Sancar, 1997b). Further still, extracts from Cockayne syndrome cells deficient in either CSA or CSB showed that they both displayed reduced levels of transcription that was rescued with transfection of the gene coding for the deficient protein (Dianov et al., 1997). Microarray analysis of cells lacking CSB have also shown a general dysregulation of gene expression (Newman et al., 2006; Y. Wang et al., 2014). Together, these studies provide a variety of evidence that supports the role of CSB and CSA in general transcription under normal conditions.

It seems possible that CSB functions in transcription and TC-NER through a unified mechanism explained by recent biochemical and structural work. *In vitro* transcription assays showed *S. cerevisiae* homologue of CSB, Rad26, uses its ATPase activity to resolve RNAPII stalled at a variety of obstructions. Rad26 was shown to be able to ‘push’ stalled RNAPII through A-tracts and DNA-binding polyamides, but not RNAPII stalled at CPDs, where a distinct conformation was adopted (J. Xu et al., 2017). Rad26 can also aid RNAPII in overcoming nucleosome barriers (J. Xu et al., 2020). It may be that Rad26/CSB ‘probes’ RNAPII during normal transcription at stalling barriers where its translocase activity can resolve RNAPII stalled at permissive barriers, but not RNAPII stalled at bulky lesions where it forms a stable complex and initiates TC-NER. A similar model was proposed for bacterial Mfd, which was found to translocate autonomously on DNA at a speed of 7 bp/s with a processivity of only 200 bp, too slow to catch up to an actively elongating RNAP, unless it is stalled, where it can help resolve the stalled RNAP to continue transcription (Le et al., 2018). The structure of the ATPase domain of Rad26 has been resolved in complex with RNAPII (J. Xu et al., 2017). This shows that Rad26 binds to the fork of the transcription bubble upstream of RNAPII and sits between the clamp (Rpb2 side) and stalk (Rpb4/7) regions,

creating an 80º bend in the DNA, something not seen in any of the several other RNAPII structures. This may in turn act as a binding site for the recruitment of downstream NER factors.

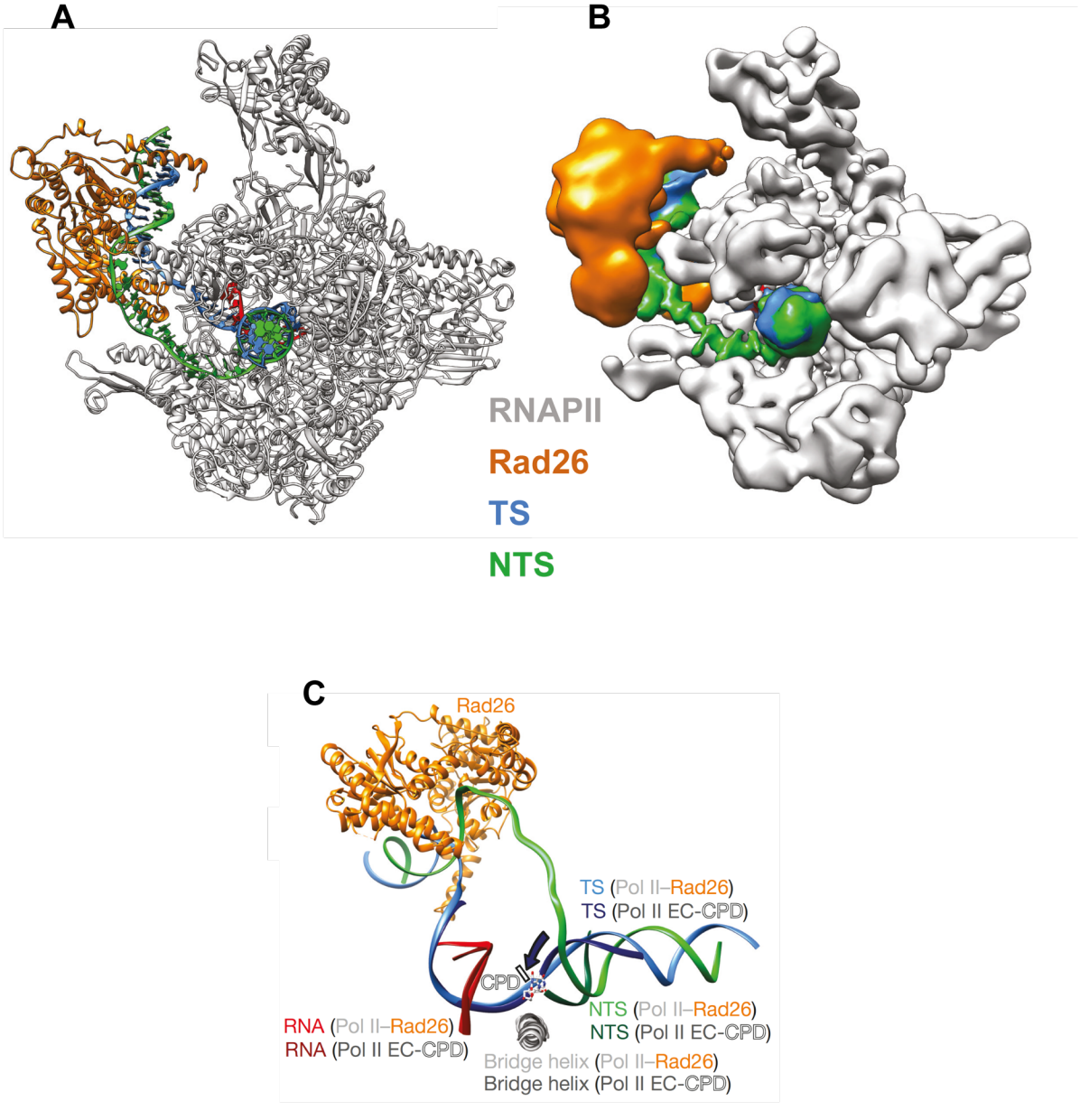


Figure 1-7 Rad26-RNAPII structure (adapted from J. Xu et al., 2017)
Rad26 binds to upstream transcription bubble and bends DNA 80º upwards. RNAPII transcription direction is out of the paper toward the reader.
A Atomic model of Rad26-RNAPII. **B** Cryo-EM density of Rad26-RNAPII.
C Overlay of Rad26-RNAPII and CPD-stalled RNAPII structures with RNAPII density removed.

While a small portion of CSB is associated with RNAPII under normal conditions, the presence of DNA damage dramatically stabilises this interaction (D Tantin et al., 1997; Van Den Boom et al., 2004; Van Gool et al., 1997). By immunoprecipitating RNAPII before and after UV irradiation in normal and Cockayne syndrome cells, one study was able to deduce the sequence of NER factors recruited to sites of damage (Fousteri et al., 2006). It showed that CSB was the first factor to be recruited to a stalled RNAPII and was essential for the recruitment of downstream repair factors: TFIIH, XPF, XPG, XPA, and RPA, which didn't happen in cells lacking CSA or CSB. Importantly, CSA was recruited to RNAPII in a UV-dependent and CSB-dependent manner. More recently, using a similar technique, the Luijsterburg group were able to map a region of the CSB C-terminal vital for interaction and recruitment of CSA (van der Weegen et al., 2020). They confirmed that CSB indeed initiates TC-NER by binding to lesion-stalled RNAPII and next recruits CSA, which is needed for UVSSA association, in turn the key factor for TFIIH recruitment. In some ways, CSB/CSA and UVSSA can thus be seen as the TC-NER counterparts of the damage-recognizers DDB1-DDB2 and XPC in GG-NER.

CSA is a 396 amino acid protein of 44 kDa. It has no intrinsic enzymatic activity, but contains 7 WD40 repeats which are motifs with β-propeller architecture and thought to facilitate protein interactions (H. X. Zhou & Wang, 2001). As well as interacting with CSB after UV, CSA also interacts with TFIIH and XAB2. XAB2 is a non-catalytic protein consisting mainly of 15 tetratricopeptide repeats and while an essential factor in the TC-NER pathway due to its interaction with XPA, not much more is known about it, except that it does not function in GG-NER (Kuraoka et al., 2008; Nakatsu et al., 2000). CSA interacts with the TFIIH subunit p44, which is also known to regulate XPD helicase activity (Coin et al., 1998; Henning et al., 1995). Although CSA has no catalytic activity itself, its action in TC-NER is presumed to derive from the fact that it resides in complex with the CRL4A E3 ubiquitin ligase as a DCAF through interaction with DDB1, essentially replacing DDB2 which only operates in GG-NER (see 1.3.2.1) (Groisman et al., 2003). Ubiquitylation by this $CLR4A^{CSA}$ complex is negatively regulated by the COP9 signalosome (CSN), which is a common partner of the cullin family of E3 ligases. CSN de-neddylates cullins, regulating (reducing) their ubiquitin ligase activity (also see 1.2.4) (Cope et al.,

2002; Lyapina et al., 2001; Yang et al., 2002). Surprisingly, Groisman and colleagues found that the CSN module also contained deubiquitylation activity *in vitro* and that neither $CLR4A^{CSA}$ nor $CLR4A^{DDB2}$ in their complex with CSN displayed any detectable ubiquitin ligase activity. However, as expected, CRL4A complexes lacking the CSN module had ubiquitin ligase activity. It seems then that CSN may somehow modulate protein modification during NER reactions by deneddylation of CRL4A and deubiquitylation of substrates. Strikingly, the two CRL4A complexes reacted oppositely in response to irradiation in that CSN dissociated from $CLR4A^{DDB2}$, but associated with $CLR4A^{CSA}$ after UV. Recent protein structures of CRL4A with the CSN module show that CRL4A cannot bind a substrate and CSN at the same time, which keeps substrate-unbound CRL4A molecules constitutively deactivated by CSN-mediated deneddylation (Cavadini et al., 2016; Fischer et al., 2011). Fischer and colleagues also demonstrated that $CLR4A^{CSA}$ can auto-ubiquitylate itself *in vitro*, which is stimulated by addition of CSB. $CLR4A^{CSA}$ was also shown to ubiquitylate CSB *in* vitro, mimicking the functional connection between $CLR4A^{DDB2}$ and XPC (Fischer et al., 2011). The findings that CSN associates with $CLR4A^{CSA}$ upon UV irradiation, thus inactivating it (Groisman et al., 2003); that CSB and CSA are both recruited to stalled RNAPII upon damage (Fousteri et al., 2006); and that CSB is ubiquitylated in response to UV and by $CLR4A^{CSA}$ are clearly inconsistent. CSA has also been shown to be necessary for the degradation of CSB following UV-irradiation within 4 hours (Groisman et al., 2006). However, other studies found CSB is not degraded within 6 hours following UV (Liebelt et al., 2020). Our laboratory has never been able to detect significant UV-induced degradation of CSB (unpublished data).

Protection of CSB from degradation may normally be ensured by UVSSA, which is stably associated with the USP7 DUB and stimulates TC-NER (Schwertman et al., 2012, 2013). UVSSA was found to be associated with RNAPII, TFIIH and CSB after UV (Nakazawa et al., 2012). It was shown that CSB was degraded more quickly after UV in cells with non-functional UVSSA, indicating that its binding partner, USP7, may stabilise CSB by deubiquitylating it (Nakazawa et al., 2012; Schwertman et al., 2012; Zhang et al., 2012). UVSSA also somehow increased RNAPII ubiquitylation after UV, but equally was essential for the reappearance of initiating RNAPII following its degradation (Nakazawa et al., 2012; Zhang et al.,

2012). This confusing web of interactions and interdependencies still require full elucidation in a mechanistically satisfying model.

The Svejstrup lab identified a UBD in the C-terminal of CSB that is essential for its role in TC-NER (Anindya et al., 2010). CSB lacking the C-terminal portion containing the UBD was able to facilitate assembly of repair factors around a stalled RNAPII, but ultimately couldn't support incision or damage removal. Replacement of CSB's UBD with a heterologous UBA from *S. cerevisiae* Rad23 rescued UV-sensitivity to CSB WT levels. The precise position of this UBD and its domain classification has since been clarified to be a winged-helix domain in structures that indeed confirm it can bind ubiquitin, in the first known interaction of a winged-helix domain and ubiquitin (Takahashi et al., 2019). However, the ubiquitylated protein substrate of this UBD is unknown. Two possibilities are that the UBD binds another ubiquitylated protein, or it could mediate intra-molecular interactions on CSB itself akin to how the leucine latch of CSB's N-terminal binds and regulates its ATPase activity (L. Wang et al., 2014). CSB's UBD provides a compelling connection of CSA (part of a ubiquitin 'writer' complex) to CSB (a ubiquitin 'reader') that appear to be intimately linked due to their similar, if not identical, phenotypes of Cockayne syndrome.

Interestingly, CSB has also been implicated in the repair of oxidative damage, otherwise usually attended to by glycosylases of the BER pathway (see 1.1.3.1). An increase in oxidative damage has thus been observed in cells and mice lacking CSB (de Waard et al., 2004; Kirkali et al., 2009; Menoni et al., 2012; Muftuoglu et al., 2009; Osterod et al., 2002; Tuo et al., 2003). CSB has also been found to interact with proteins involved in the BER pathway, such as the glycosylases OGG1 and NEIL1 (Muftuoglu et al., 2009; Tuo et al., 2002). In fact, a ubiquitylation site that renders CSB specifically sensitive to oxidative damage, but not UV irradiation, was identified and characterised by the Svejstrup laboratory (Ranes et al., 2016). The increased oxidative damage load resulting from inefficient repair is thought to be responsible for mitochondrial dysfunction in Cockayne syndrome cells, which is a popular theory for a mechanism of ageing (Scheibye-Knudsen et al., 2013; Sun et al., 2016).

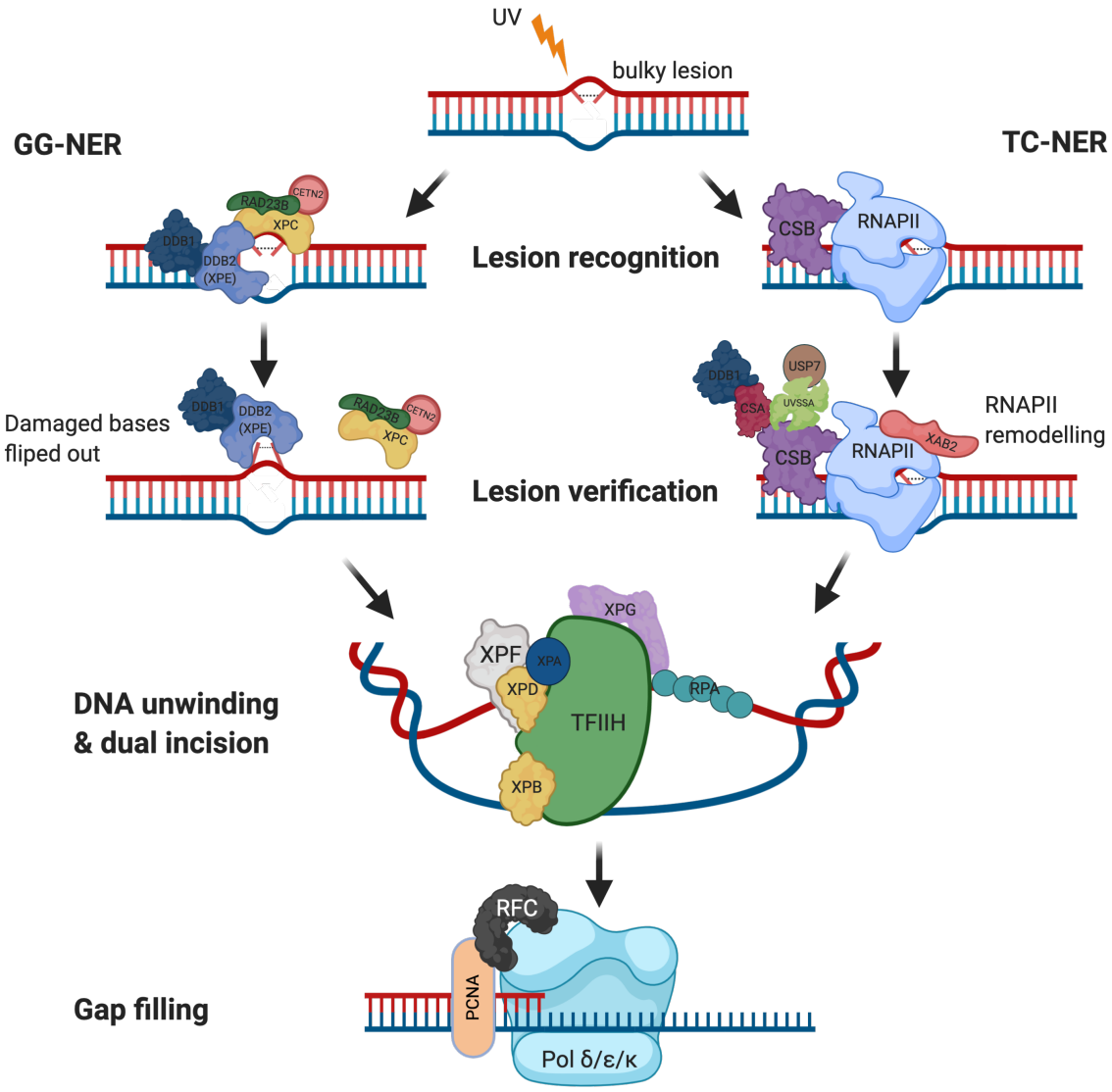


Figure 1-8 Schematic of NER reaction
Simplified cartoon overview of the two pathways of NER . See text for more details.

1.3.5 The fate of RNAPII stalled at damage

While RNAPII is the enzyme that synthesises RNA, transcription involves dozens of proteins that have evolved to assist RNAPII on its journey along a gene. As RNAPII doesn't possess ATP-dependent translocase activity, it translocates by Brownian motion that is favoured towards forward translocation by a ratchet mechanism, the stabilization of the forward-translocated state by binding of the correct next nucleotide to be incorporated, and by the association with elongation factors (Bar-Nahum et al., 2005; Svetlov & Nudler, 2013). This means that in certain circumstances RNAPII does backtrack, which creates a problem for continuation of transcription as the 3′ end of the RNA is now misaligned from the active site, blocking further nucleotide polymerisation. This is resolved by cleavage of the RNA end via RNAPII's intrinsic RNA cleavage activity, stimulated by TFIIS (Sigurdsson et al., 2010; Zatreanu et al., 2019). It can also be resolved by forward translocation of RNAPII to reposition the RNA 3′ end in the active site, which can be achieved for example, by the assistance provided by CSB (Christopher P. Selby & Sancar, 1997b; J. Xu et al., 2017). However, the stalling of RNAPII by obstacles in its path creates a more complicated problem. RNAPII is the single protein complex involved in expression of protein-coding genes, meaning that there are no translesion enzymes like there are for polymerases replicating the DNA (Waters et al., 2009). Irreversibly stalled RNAPIIs represent toxicity for the cell as they stall other RNAPIIs transcribing behind them and prevent any further transcription of that gene due to the unidirectionality of transcription (Saeki & Svejstrup, 2009).

RNAPII stalled at a CPD lesion shields the lesion inside its active site (Brueckner et al., 2007), which prevents lesion verification by TFIIH without remodelling of the stalled RNAPII (C. L. Li et al., 2015). Furthermore, excised oligonucleotides from TC-NER reactions are 24–32 nt in length, meaning incision happens at DNA nucleotides that would be occluded by a stalled RNAPII with a footprint of ~40 bp (J. C. Huang et al., 1992; Juch Chin Huang & Sancar, 1994; Svoboda et al., 1993). Therefore, damage-stalled RNAPII must be remodelled or removed to allow repair of lesions by NER, but the fate of RNAPII at lesions is debated. In bacterial NER, the UvrD helicase unwinds and removes the damaged DNA after incision, much

like TFIIH in humans (Aziz Sancar & Reardon, 2004). Recently, it has been shown that UvrD can also use its helicase activity to translocate RNAP backwards in TC-NER (Epshtein et al., 2014). However, Mfd (*E. coli* CSB) which initiates TC-NER by binding stalled RNAP, uses its translocase activity to move RNAP forwards and remove it from DNA (Fan et al., 2016; Shi et al., 2020). There may then be competing mechanisms for the removal or remodelling of RNAP stalled at damage. The importance of UvrD-mediated RNAP backtracking in TC-NER is disputed (Adebali, Chiou, et al., 2017; Adebali, Sancar, et al., 2017). In *S. cerevisiae,* Rad26 (CSB) initiates TC-NER and its mutation decreases repair of the transcribed strand, but surprisingly cells are not very UV sensitive (Van Gool et al., 1994). Cells lacking Rad26 and RPB9, a subunit of RNAPII, are much more UV-sensitive and TC-NER defective than the single mutants alone (S. Li & Smerdon, 2002). Deletion of Spt4 or C-terminal mutations in Spt5, subunits of the DSIF elongation factor, rescue the UV-sensitivity of Δ*rad26* cells (Ding et al., 2010; Jansen et al., 2000). Interestingly, both the Rad26-RNAPII and DSIF-RNAPII structures have been resolved, which show overlapping densities of Rad26 and DSIF on RNAPII, suggesting they cannot cooccur on the same RNAPII molecule (Bernecky et al., 2017; Ehara et al., 2017; J. Xu et al., 2017). Collectively, these data underscore the varying requirements of Rad26 for TC-NER in specific contexts and highlights the inhibitory effects of the RNAPII elongation machinery on TC-NER.

Aside from Rad26- or TFIIH-mediated removal/backtracking of RNAPII there is also evidence for the degradation of RNAPII in response to UV. This might facilitate the alternative, and slower, lesion recognition by the GG-NER pathway (Anindya et al., 2007; Nakazawa et al., 2020; Ratner et al., 1998; Reid & Svejstrup, 2004; Somesh et al., 2005, 2007; Tufegdžić Vidaković et al., 2020; M. D. Wilson et al., 2013; Woudstra et al., 2002). This process has been termed the ‘last resort pathway’ as it entails proteasomal degradation of the stalled RNAPII in such scenarios that TC-NER fails (M. D. Wilson et al., 2013; Woudstra et al., 2002). This is an established way of removing elongating RNAPIIs from DNA during several transcription stress events, such as head-on collisions (Hobson et al., 2012; Noe Gonzalez et al., 2021). In yeast, the polyubiquitylation of RNAPII for degradation is preceded by its monoubiquitylation by Rsp5 (NEDD4 in humans). Monoubiquitylated RNAPII is recognised by the Elongin-Cullin complex in concert with Def1, and a polyubiquitin

chain is then generated (Harreman et al., 2009; Somesh et al., 2005; Woudstra et al., 2002). The process is the same in humans, but with NEDD4 and CRL5$^{\text{Elongin A}}$ (Anindya et al., 2007; Yasukawa et al., 2008). More recently the Svejstrup laboratory and others identified a single lysine residue, K1268 on RPB1, the largest and catalytic subunit of RNAPII, that when mutated completely abolishes RNAPII ubiquitylation and degradation (Nakazawa et al., 2020; Tufegdžić Vidaković et al., 2020). Cells expressing endogenous RPB1 K1268R are sensitive to UV irradiation (but not as much as CSB KO cells) and exhibit a massive rewiring of gene expression upon UV irradiation, whereby short genes that are less likely to be damaged become overexpressed (Tufegdžić Vidaković et al., 2020). It might be then that RNAPII degradation also serves as a way to control cellular RNAPII levels to avert aberrant gene expression changes. RNAPII-stalling damage halts transcription of that gene meaning RNAPII occupancy of the gene is reduced. Long genes usually act as a 'sink' for RNAPII, but quickly become saturated with high levels of damage. This increases the local pool of RNAPII that subsequently become rerouted to short genes, which are by chance less likely to have damage. Amazingly, expression of K1268R RPB1 in the background of CSB KO allowed cells to recover transcription to the end of a 300 kb gene after UV, indicating they were TC-NER competent, even in the absence of CSB. This suggests that CSB's role might also be to protect RNAPII from degradation at the site of damage. This is consistent with more severe UV-induced degradation of RNAPII in yeast lacking Rad26 (Woudstra et al., 2002). This seemed to be contradicted by a study that showed reduced RNAPII ubiquitylation in Cockayne syndrome cells (Bregman et al., 1996). However, in the absence of CSA and CSB, cells shut down transcription more quickly after UV, meaning that less RNAPII runs into genes to meet the unrepaired DNA damage and present itself as a substrate for ubiquitylation (Anindya et al., 2007). The Ogi group found that ubiquitylation of RPB1 and UVSSA was necessary for recruitment of TFIIH to stalled RNAPII. A novel sequencing method also revealed that RPB1 ubiquitylation stimulated repair of the transcribed strand, highlighting a role for RPB1 ubiquitylation in facilitating TC-NER. Remarkably, a mouse expressing K1268R RPB1 displayed signs of Cockayne syndrome (Nakazawa et al., 2020).

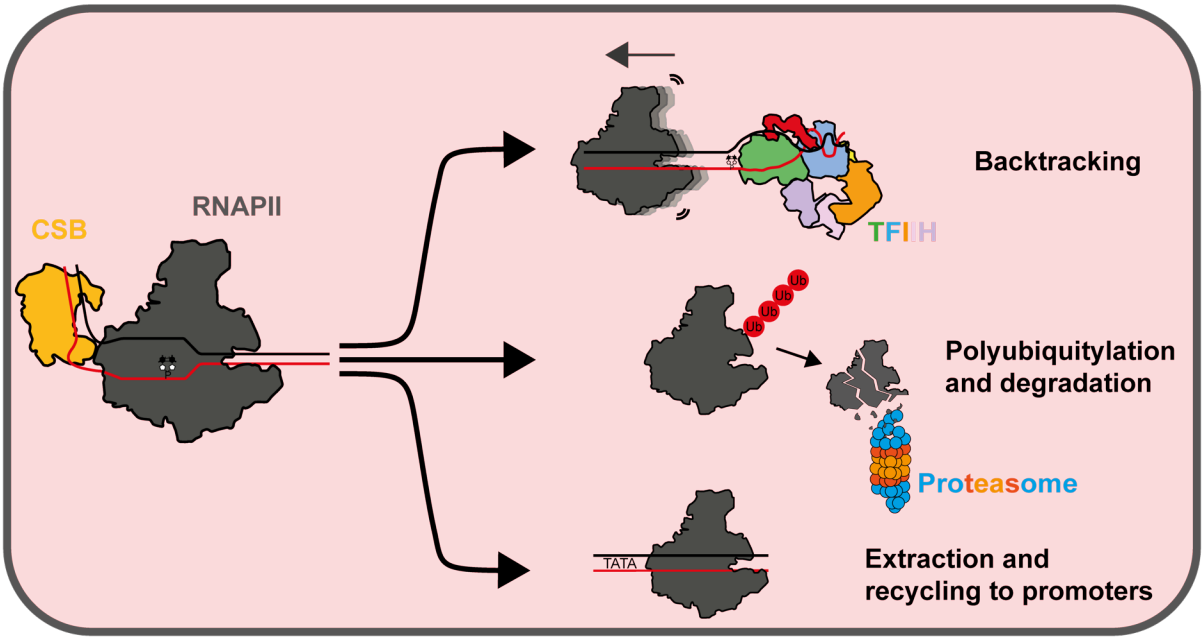


Figure 1-9 Schematic of the possible fates of RNAPII stalled at a lesion
Remodelling of a stalled RNAPII that conceals the lesion must be carried out to facilitate repair. There are several different possibilities and the importance of each's contribution to the TC-NER process is disputed. RNAPII may be backtracked away from the lesion (as is common during transcription stress), polyubiquitylated and degraded, or extracted from chromatin by another factor which allows it to be recycled to the promoter to continue 'scanning' for lesions if degradation of the RNAPII pool is not too severe. Resolution of RNAPII allows TFIIH to verify the lesion and for excision of a ~25 nt fragment of damaged DNA by 5′ and 3′ incisions of XPF and XPG nucleases, respectively.

1.4 Aims of this study

The cellular response to UV damage is clearly very complex and many questions remain about the fate of RNAPII at sites of damage and the role of CSB in facilitating its remodelling or removal. The vitally important role of ubiquitin in marking, signalling and regulating factors during lesion recognition in GG-NER is well established and understood. The role of ubiquitylation in TC-NER is well established too, but its consequences are less clear, and many questions remain unanswered. In this thesis I explore the interaction between CSB, DSIF, and RNAPII. I also characterise ubiquitylation sites on CSB and the cellular consequences of a CSB mutant that cannot be ubiquitylated in response to UV.

Chapter 2. Materials & Methods

2.1 General solutions

PBS (phosphate-buffered saline)

137 mM NaCl

2.7 mM KCl

10 mM Na_2HPO_4

1.8 mM KH_2PO_4

pH 7.4

Nuclease-free water

not DEPC treated (Thermo Fisher Scientific)

All assay-specific buffers are listed in the method description of the applicable assay

2.2 Molecular biology techniques

2.2.1 Plasmids

Plasmid name	Description	Source
pTRE3G-GFP-FLAG-CSB WT	CSB WT coding sequence with 5′ GFP tag and FLAG tag under the control of dox-inducible promoter	Previous lab member
pTRE3G-GFP-FLAG-CSB 5K>R	Same as CSB WT with 5 K>R mutations	Previous lab member
pTRE3G-GFP-FLAG-CSB 8K>R	Same as CSB WT with 8 K>R mutations	This study
pTRE3G-GFP-FLAG-CSBΔUBD	Same as CSB WT with terminal 273 amino acids deleted	Previous lab member
pDEST-FRT/TO-CSB WT	CSB coding sequence from equivalent pTRE3G plasmid cloned into vector with FRT sites and dox-inducible promoter	This study
pDEST-FRT/TO-CSB 5 K>R	CSB coding sequence from equivalent pTRE3G plasmid cloned into vector with FRT sites and dox-inducible promoter	This study
pDEST-FRT/TO-CSB 8 K>R	CSB coding sequence from equivalent pTRE3G plasmid cloned into vector with FRT sites and dox-inducible promoter	This study

pDEST-FRT/TO-CSB 19 K>R	CSB coding sequence from equivalent pTRE3G plasmid cloned into vector with FRT sites and dox-inducible promoter	This study
pDEST-FRT/TO-CSBΔUBD	CSB coding sequence from equivalent pTRE3G plasmid cloned into vector with FRT sites and dox-inducible promoter	This study
pOG44	Vector containing Flp recombinase for establishing Flp-In cell lines in combination with pDEST-FRT/TO vectors	(Thermo Fisher Scientific)
pFL-StrepII-CSB-FLAG-8xHIS	Plasmid for baculovirus generation. CSB WT with 5′ 2xStrepII tags and 3′ FLAG and 8xHIS tags	This study
pET-DUET-Sumo-SPT4-SPT5	Bacterial expression vector with DSIF. Spt4 is tagged 5′ with 10xHis, 7xArg, Sumo, 3C protease site in one cassette followed by Spt5 in separate cassette	Patrick Cramer laboratory
pGEX3-Dsk2	Bacterial expression plasmid containing GST-Dsk2	Previous lab member
pGs-21a-multiDsk	Bacterial expression plasmid containing GST-MultiDsk	(M. D. Wilson et al., 2012)
pSpCas9n(BB)-2A-GFP	Encodes GFP tagged Cas9 nickase mutant	(Ran, Hsu, Wright, et al., 2013)
pSpCas9WT(BB)-2A-GFP	Encodes GFP tagged Cas9	(Ran, Hsu, Wright, et al., 2013)

Table 2-1 All plasmids used in this study

2.2.2 PCR

All PCR reactions were performed with Q5 High-Fidelity DNA Polymerase 2X Master Mix (NEB) in 25 µl reactions according to the manufactures protocol. The standard protocol is outlined below.

PCR reaction

Component	Volume µl	Final Concentration
Q5 High-Fidelity 2X Master Mix	12.5	1X
10 µM Forward primer	1.25	0.5 µM
10 µM Reverse primer	1.25	0.5 µM
Template DNA	2	1–1000 pg
Nuclease-free Water	8	n/a

Cycling conditions

1.	Initial Denaturation	98°C for 30 seconds	
2.	Denaturation	98°C for 10 seconds	25–35 Cycles
3.	Annealing	50–72°C for 20 seconds	
4.	Extension	72°C for 30 seconds/kb	
5.	Final Extension	72°C for 2 minutes	

The annealing temperature was determined for each primer pair using NEB T_m calculator at www.tmcalculator.neb.com. If amplification failed, a T_m was determined empirically by using a gradient of 8 temperatures over

2.2.3 Site-directed mutagenesis

Mutagenesis of codons in DNA plasmids was performed using Q5 Site-Directed Mutagenesis Kit (NEB) according to manufactures protocols. Primer sequences containing the desired mutations were designed using NEB base changer at www.nebasechanger.neb.com. These primers were used in PCR reactions using the same conditions outlined above. After amplification, 1 µl of PCR product was incubated with 5 µl KLD reaction buffer, 1 µl 10X KLD enzyme mix, 3 µl nuclease-free water and incubated for 5 minutes at room temperature. 5 µl of this reaction was transformed into DH5α competent cells and plasmids were purified by mini preparation, as described above.

pTRE3G-GFP-FLAG-CSB 5K>R plasmid was previously generated and available in the laboratory. The additional 3 K>R mutations were made to generate pTRE3G-GFP-FLAG-CSB 8K>R using the primer sequences below.

Primer name	Sequence 5′→3′ (mutated codons in red)
CSB K774/783R Fwd	AGAGTCTACCAAAATTTCGTTGATTCCAGAGAAG
CSB K774/783R Rev	ATGCTGCTCATCTGTAAGACGGCAAAATAAG
CSB K1359/1360R Fwd	GGCATCATGAGAAGGGAGGGAAAAG
CSB K1359/1360R Rev	ATCCTGGCACTTCTCTGT

Table 2-2 Site-directed mutagenesis primers

2.2.4 CSB 19K>R generation

To generate a CSB 19>R coding sequence, a 1,750 bp dsDNA molecule with 17 K>R mutations encoded was synthesised and ordered from Genentech (sequence below). The sequence aligned to the C-terminal region of CSB's coding sequence from the codon encoding amino acid 918 up to 19 nucleotides 3′ of the stop codon. A PCR reaction was set up using 1 ng of pDEST-FRT/TO-CSB 8 K>R plasmid as template (to retain K774R and K783R present in CSB 8K>R) and the 1,750 bp dsDNA (CSB 17K>R) was used as a primer pair at a final concentration of 2 nM. The reaction was carried out as described above but with the annealing step removed, an extension step of 6 minutes, and 25 cycles of melting and annealing/extension. The resulting PCR product was subject to KLD reaction, transformed and miniprepped as described. The resulting plasmid was sequenced across the entire CSB coding sequence to confirm the presence of 19 K>R mutations and designated pDEST-FRT/TO-CSB 19 K>R.

CSB 17 K>R dsDNA forward strand sequence 5′→3′

GGCTTAGGTGTTAACCTGACGGGGGCAAACAGAGTTGTCATCTATGACCCAGACTGGAAC
CCAAGCACGGACACGCAGGCCCGGGAGCGAGCATGGAGAATAGGCCAGAAGAAGCAAGT
GACTGTGTACAGGCTCCTGACTGCGGGCACCATTGAAGAAAGGATCTACCACCGACAAAT
CTTCAAGCAGTTTTTGACAAATAGAGTGCTAAAAGACCCAAAACAAAGGCGGTTTTTCAGAT
CCAATGATCTCTATGAGCTATTTACTCTGACTAGTCCTGATGCATCCCAGAGCACTGAAACA
AGTGCAATTTTTGCAGGAACTGGATCAGATGTTCAGACACCCAAATGCCATCTAAAAAGAA
GGATTCAACCAGCCTTTGGAGCAGACCATGATGTTCCAAAACGCAAGAAGTTCCCTGCTTC
TAACATATCTGTAAATGATGCCACATCATCTGAAGAGAAATCTGAGGCTAAAGGAGCTGAA
GTAAATGCAGTAACTTCTAATCGAAGTGATCCTTTGAAAGATGACCCTCACATGAGTAGTAA
TGTAACTAGCAATGATAGGCTTGGAGAAGAGACAAATGCAGTATCTGGACCAGAAGAGTTG
TCAGTGATTAGTGGAAATGGGGAATGTTCAAATTCTTCAGGAACAGGCAAAACTTCTATGC
CATCTGGTGATGAAAGCATTGATGAAAGGTTAGGTCTTTCTTACAAAAGAGAAAGACCCAG
CCAGGCTCAAACAGAAGCTTTTTGGGAGAATAGACAAATGGAAAATAATTTTTATAAGCACA
AGTCAAGAACAAGACATCATAGTGTGGCAGAAGAAGAGACCCTGGAGAAACATCTGAGAC
CAAAGCAAAAGCCTAAGAACTCTAAGCATTGCAGAGACGCCAAGTTTGAAGGAACTCGAAT
TCCACACCTGGTGAAGAAAAGGCGTTACCAGAAGCAAGACAGTGAAAACAAGAGTGAGGC
CAAGGAACAGAGCAATGACGATTATGTTTTGGAAAGGCTTTTCAGAAGATCAGTTGGCGTG
CACAGTGTCATGAAGCACGATGCCATCATGGATGGAGCCAGCCCAGATTATGTACTGGTG
GAGGCAGAAGCCAACCGAGTGGCCCAGGATGCCCTGAGAGCACTGAGGCTCTCTCGTCA
GCGGTGTCTGGGAGCAGTGTCTGGTGTTCCCACCTGGACTGGCCACAGGGGGATTTCTG
GTGCACCAGCAGGAAAAAAGAGTAGATTTGGTAAGAAAAGGAATTCTAACTTCTCTGTGCA
GCATCCTTCATCAACATCTCCAACAGAGAAGTGCCAGGATGGCATCATGAGAAGGGAGGG
AAGAGATAATGTCCCTGAGCATTTTAGTGGAAGAGCAGAAGATGCAGACTCTTCATCCGGG
CCCCTCGCTTCCTCCTCACTCTTGGCTAGAATGAGAGCTAGAAACCACCTGATTCTGCCAG
AGCGTTTAGAAAGTGAAAGCGGGCACCTGCAGGAAGCTTCTGCCCTGCTGCCCACCACAG
AACACGATGACCTTCTGGTGGAGATGAGAAACTTCATCGCTTTCCAGGCCCACACTGATGG
CCAGGCCAGCACCAGGGAGATACTGCAGGAGTTTGAATCCAGGTTATCTGCATCACAGTC
TTGTGTCTTCCGAGAACTATTGAGAAATCTGTGCACTTTCCATAGAACTTCTGGTGGTGAAG
GAATTTGGAGACTCAGGCCAGAATACTGCTAAGGATCCAATGTAACTGTAT

2.2.5 Agarose gel electrophoresis

Agarose gels were prepared by dissolving UltraPure agarose (Thermo Fisher Scientific) in TBE to a final concentration between 0.8–2 %, depending on application, and the solution was heated in a microwave with frequent stirring until agarose was fully dissolved. GelRed nucleic acid gel stain (Biotium) was added to the solution to a final concentration of 1X from 10,000X stock. The solution was poured into a gel caster with well comb and left to cool and solidify. DNA was separated by submerging gels in TBE buffer and applying a constant voltage of 80 V. DNA gels were visualised on an Amersham Imager 600 (Cytiva).

TBE (Tris-Borate-EDTA)
89 mM Tris Base
89 mM Boric Acid
2 mM EDTA
pH 8.3

2.2.6 RNA extraction from human cells

RNA extraction for RT-qPCR was performed with the Qiagen RNeasy mini kit following the manufactures protocol. Cells were seeded in 6-well plates and harvested with RLT buffer in the dish at 70-80% confluence, which corresponded to 1–2x10^6 HEK 293 cells. On-column DNase digestion with the provided DNase was always performed.

2.2.7 Reverse transcription

Purified RNA was used to synthesise cDNA using random hexamers with the TaqMan reverse transcriptase kit (Applied Biosystems). Reactions were performed with 1 µg of RNA in 40 µl reactions and a control reaction with water instead of reverse transcriptase was always performed to control for DNA contamination. The following reagents and conditions were used:

Reagent	Volume µl
10X TaqMan RT Buffer	4
25 mM MgCl2	8.8
dNTPs	8
Random hexamers	2
RNAse Inhibitor	0.8
MultiScribe RTase (50 U/uL)	1
RNA	1 µg
Water	to 20 µl

cDNA synthesis reaction conditions

1. 25ºC for 10 minutes
2. 48ºC for 30 minutes
3. 95ºC for 5 minutes
4. 4 ºC hold

2.2.8 qPCR

Completed cDNA reactions (including those lacking absent transcriptase) were diluted 1:2 in water and 4 µl was used in 10 µl qPCR reactions in 384-well plates using Bio-Rad SYBR green reagents with the following conditions:

Reagent	Volume µl
Bio-Rad iTaq Universal SYBR Green Supermix 2X	5
Primer pair mix (5 µM each)	1
cDNA diluted 1:2	4

qPCR cycling conditions

1. Denaturation at 94°C for 15 seconds
2. Annealing at 58°C for 15 seconds
3. Extension at 72°C 20 seconds

Steps 1–3 were cycled 30 times

Amplification and imaging were performed on a CFX-384 Real-Time PCR System (Bio-Rad). Primers against GAPDH mRNA were used for normalisation and triplicates were always used. Analysis was performed in excel.

qPCR primers	Sequence 5′→3′
GAPDH mRNA Fwd	AGCCACATCGCTCAGACAC
GAPDH mRNA Rev	GCCCAATACGACCAAATCC
EXT1 nascent RNA end Fwd	TCAGGGTAAACAAGGGCAAC
EXT1 nascent RNA end Rev	CATCCTGGAGGATTGTTCGT
PUM1 nascent RNA end Fwd	AATCACTCGGCAGCCATAAG
PUM1 nascent RNA end Rev	CTGAAGCATTGAAGGTGGTG

Table 2-3 Primers for RRS qPCR of intron-exon junctions of nascent RNA transcripts

2.2.9 Nucleic acid quantification

DNA and RNA quantification were performed using UV spectrophotometry on a Nanodrop (Thermo Fisher Scientific). For low concentrations and more sensitive assays such as 4SU-seq, RNA was quantified using Qubit RNA BR and HS assays (Thermo Fisher Scientific).

2.2.10 DNA Sanger sequencing

The coding sequence of CSB was fully Sanger sequenced in all plasmids generated to confirm the presence of desired mutations. The samples were sequenced internally at the Francis Crick Institute by a dedicated facility using BigDye terminator cycle sequencing kit (Applied Biosystems).

Sequencing reaction

- 3 µl 5X sequencing buffer
- 200 ng DNA template
- 2 µl BigDye terminator
- 1 µl primer (10µM)
- Make to 20 µl with H_2O

PCR cycling conditions

1. 96ºC for 1 minute
2. 96ºC for 10 seconds
3. annealing temperature for 5 seconds
4. 60ºC for 4 minutes

Cycle steps 2–4 25 times

5. 60ºC for 8 minutes

All sequencing primers used in this study are outlined below.

Primer name	Sequence 5′→3′
CSB CDS #1	CCCACTCAAGTCAAACTCAGG
CSB CDS #2	GGATGACCTCACGTCATGTACGAC
CSB CDS #3	GACTATGAGCTGAAGCCTCTGC
CSB CDS #4	GGAGAAACTAATTCGAGATGTTGC
CSB CDS #5	TACCATAAATCCATACCTACTGCG
CSB CDS #6	TGCTGGACATACTTGAAGTATTCC
CSB CDS #7	TTCTATGCCATCTGGTGATGAAAG
CSB CDS #8	TTTAGTGGAAGAGCAGAAGATGC
CSB CDS #9	ATCACATGAAGCAGCACGAC
CSB CDS #10	TCATCTGAAGAGAAATCTGAGGC
StrepII_CSB_CO_Seq1_F	ATCGATTCGCGACCTACTCC
StrepII_CSB_CO_Seq2_F	CGGATTATTCATACCGTCCC
StrepII_CSB_CO_Seq3_F	TCCAAGCTGTGGAACCTTCT
StrepII_CSB_CO_Seq4_F	AGGAAGACGCTGAACCTGG
StrepII_CSB_CO_Seq5_F	TGAAGCTGGAAGACGACTCA
StrepII_CSB_CO_Seq6_F	ACTGCCACGGAATCCTGA
StrepII_CSB_CO_Seq7_F	ACGTGAAGATGTCCCTGAGC
StrepII_CSB_CO_Seq8_F	CCTACCTGAAGATGGACGGA
StrepII_CSB_CO_Seq9_F	CAGAGCACTGAAACCTCAGC
StrepII_CSB_CO_Seq10_F	CGAAAAGCTGGGACTGTCAT
StrepII_CSB_CO_Seq11_F	GTCGAGGCTGAAGCCAAC
StrepII_CSB_CO_Seq12_F	GAGGCTAGCGCTCTGCTG

CMV_Forward	CGCAAATGGGCGGTAGGCGTG
bGH Rev	TAGAAGGCACAGTCGAGG
M13 Forward	GTAAAACGACGGCCAG
M13 Reverse	CAGGAAACAGCTATGAC
pre-TRE tight F	AGGCGTATCACGAGGCCCTTTCGT
pre-TRE tight R	TATTACCGCCTTTGAGTGAGCTGA
CSB Exon2 SangSeq ng1/2 7/8	TCCAGGACCGACCGATACTC
CSB Exon2 SangSeq ng 3/4 5/6	TAGCTGGCTGCCTAAATGTCC
CSB Exon4 SangSeq	ACAAGTCCACCTTTCAGGTCA
CSB Exon5 SangSeq	GGCAGAGGCTTCAGCTCATA
CSB Exon2 SangSeq ngRNA9/10	GATGGGGAGGTGGAGGAGTA

Table 2-4 All sequencing primers used in this study

2.3 Bacterial techniques

2.3.1 Solutions

Lysogeny Broth (LB) (rich medium)

1% Bacto Tryptone (DIFCO)

0.5% Yeast Extract (DIFCO)

1% NaCl

pH adjusted to 7.0

SOC (heat-shock recovery medium)

2% Bacto Tryptone (DIFCO)

0.5% Yeast Extract (DIFCO)

10 mM NaCl

2.5 mM KCl

10 mM $MgCl_2$

10 mM $MgSO_4$

20 mM Glucose

pH adjusted to 7.0

2.3.2 Transformation of competent cells

Chemically competent DH5α E.coli cells (NEB) were used for growth of plasmid DNA. Frozen cells were thawed on ice and incubated with 100 ng of DNA or 5 µl of a ligation mix for 30 minutes, followed by 30 seconds heat-shock at 42ºC. The cells were then incubated on ice for 2 minutes followed by a 60-minute recovery in SOC medium at 37ºC. Cells were finally plated on LB-agar plates supplemented with 100 µg/ml ampicillin and incubated over night at 37ºC.

2.3.3 Extraction of plasmid DNA

GeneJET Plasmid Miniprep Kit (Thermo Scientific) were used to purify transformed plasmid DNA from *E.coli*. Typically, a single bacterial colony from LB plates was grown in 5 ml LB-Amp (or the appropriate antibiotic) medium overnight at 37°C. DNA plasmids were then purified from bacterial pellets according to manufacturer instructions.

2.3.4 Generation and extraction of Bacmid DNA

pFL plasmids were transformed into competent DH10Bac *E. coli* cells (Invitrogen) and plated on agar plates containing 50 µg/mL kanamycin, 7 µg/mL gentamicin, 10 µg/mL tetracycline, 100 µg/mL Bluo-gal, and 40 µg/mL IPTG and incubated at 37ºC for 24–48 hours until the appearance of blue/white colonies. White colonies were picked and grown up in LB containing the aforementioned antibiotics overnight. To purify Bacmid DNA the GeneJET miniprep kit mentioned above was used for resuspension, lysis and neutralisation steps, but DNA in the supernatant was precipitated with 0.7 volumes of isopropanol instead of binding to column. DNA was pelleted by centrifugation at 18,000 x *g* for 20 minutes and washed in 70% ethanol before air drying and resuspension in nuclease-free water.

2.3.5 Expression of Dsk2 and MultiDsk proteins

One Shot BL21(DE3) Star competent bacterial cells were transfected with pGEX3-Dsk2 or pGs-21a-multiDsk plasmid according to the manufacturer's instructions and plated on ampicillin selection plates. After overnight incubation at 37°C a colony was picked and inoculated into 20 ml of LB containing 100 µg/ml ampicillin and grown at 37°C overnight with shaking. The following day, 300 ml of LB-Amp was inoculated with 5 ml of the pre-inoculum and grown at 37°C with shaking at 200 rpm. Expression was induced with 1mM IPTG when bacterial growth reached OD_{600} of 0.8 for Dsk2 and 0.6 for MultiDsk. The induced culture was grown for 4 h at 30°C, with shaking at 200 rpm. Bacteria were then centrifuged, and the pellets were snap frozen in liquid nitrogen, and stored at −80°C until further processing.

2.3.6 Expression of DSIF protein

pET-DUET-Sumo-SPT4-SPT5 was transfected into BL21-CodonPlus(DE3)-RIL *E. coli* cells (Agilent) according to manufacturer's instructions. After outgrowth, cells were plated on agar plates with ampicillin and incubated overnight at 37ºC. Single colonies were grown in a 25 ml starter culture overnight at 37ºC. This starter culture was diluted into 500 ml and grown to OD_{600} 0.6. ZnCl2 was then added to 10 µM and expression induced with IPTG to 1 mM and cells were grown at 37ºC for 3 hours. Pellets were then harvested by centrifugation at 8,000 x *g*, snap frozen in liquid nitrogen and stored at -80ºC until further use.

2.4 Insect cell techniques

2.4.1 Insect cell lines

The insect cell line used for all experiments was Sf9, a clonal isolate of Sf21 *Spodoptera frugiperda* cells. These were obtained from Cell Services, The Francis Crick Institute.

2.4.2 Growth conditions

Insect cell lines were cultured in Gibco Sf-900 III SFM media (Thermo Fisher Scientific) unless stated otherwise. Cells were seeded at 0.5 x 10^6 cells/ml and grown until a confluency of 1 x 10^7 cells/ml when they were split. Cells were always grown in suspension with orbital rotation at 27ºC in a humidified incubator, unless otherwise stated. Cells were usually grown in conical flasks with filtered lids and filled to no more than 25% volume. For expression cells were grown in 2 L roller bottles.

2.4.3 Amplification of baculovirus

Sf9 cells were seeded at 5 x 10^5 cells/ml in 6-well dishes and allowed to attach by incubation at room temperature for 15 minutes. 2 µg bacmid DNA (2.3.4) was transfected into Sf9 cells using TransIT-LT1 transfection reagent (Mirus Bio) according to manufacturer's protocol. Transfected cells were left to grow and produce baculovirus in a humidified incubator at 27ºC without shaking. After 5 days, media containing shed baculovirus (P_1) was aspirated and saved.

To obtain higher baculovirus titres P_1 baculovirus media was diluted 20-fold into a fresh suspension of cells at 1 x 10^6 cells/ml, which were grown in suspension with shaking. Cells were monitored until they displayed approximately 20–30% viability, upon which the suspension was centrifuged to pellet cells and P_2 baculovirus-containing media was saved. Generation of high of viral titres was achieved by infecting Sf9 cells with P_2 as above but at 1:100 dilution into 50 ml cell suspension to generate P_3 which was used for expression.

2.4.4 Protein expression and harvesting

Optimal P_3 dilution and growth time for highest CSB expression was empirically determined using dilution-series and time-course experiments with western blot analysis. 20 ml of P_3 was added to 2 L of an Sf9 cell suspension at 2 x 10^6 and grown for 72 hours. The infected cells were harvested by centrifugation at 800 x *g* at 4ºC for 10 minutes. Media was removed and cell pellets were snap frozen in

liquid nitrogen and stored at -80ºC until further use. StrepII-CSB purification is outlined in section 2.8.2.

2.5 Mammalian cell culture techniques

2.5.1 Cell lines

Cell line	Description	Source	Additional antibiotics
MRC5va	SV40-transformed lung fibroblast cell-line with normal diploid karyotype	Svejstrup laboratory	none
CS1ANsv	SV40-transformed fibroblast cell-line derived from cockayne syndrome patient with *ERCC6* (CSB) inactivating mutations	Svejstrup laboratory	none
HEK293 Flp-In T-REx	HEK 293 cells containing single FRT Flp-In site under control of T-REx dox-inducible promoter	Thermo Fisher Scientific	100 µg/ml Zeocin 15 µg/ml blasticidin
U2OS Flp-In T-REx	U2OS cells containing FRT Flp-In site(s) under control of T-REx dox-inducible promoter	Steve West laboratory	100 µg/ml Zeocin 15 µg/ml blasticidin
CS1ANsv - TETon- CSB WT #5	CS1ANsv with doxycycline-inducible expression of GFP-FLAG-CSB WT	Svejstrup laboratory	0.2 µg/ml Puromycin
CS1ANsv - TETon- CSB WT #5A	Reselected single clone of CS1ANsv - TETon- CSB WT #5	Svejstrup laboratory	0.2 µg/ml Puromycin
CS1ANsv - TETon- CSB∆UBD	CS1ANsv with doxycycline-inducible expression of GFP-FLAG-CSB∆UBD (lacking terminal 273 amino acids)	Svejstrup laboratory	0.2 µg/ml Puromycin
CS1ANsv - TETon- CSB 5K>R	CS1ANsv with doxycycline-inducible expression of GFP-FLAG-CSB 5K>R	This study	0.2 µg/ml Puromycin
CS1ANsv - TETon- CSB 8K>R	CS1ANsv with doxycycline-inducible expression of GFP-FLAG-CSB 8K>R	This study	0.2 µg/ml Puromycin
HEK293 Flp-In T-REx CSB KO	HEK293 Flp-In T-REx cell line with CRISPR-Cas9 targeted indel mutation of exon 2 of *ERCC6* gene (CSB) rendering it inactive.	This study	100 µg/ml Hygromycin 15 µg/ml blasticidin
HEK293 Flp-In T-REx CSBWT	HEK293 Flp-In T-REx CSB KO with dox-inducible GFP-FLAG-CSB WT	This study	100 µg/ml Hygromycin 15 µg/ml blasticidin
HEK293 Flp-In T-REx CSB 5K>R	HEK293 Flp-In T-REx CSB KO with dox-inducible GFP-FLAG-CSB 5K>R	This study	100 µg/ml Hygromycin 15 µg/ml blasticidin

HEK293 Flp-In T-REx 8K>R	HEK293 Flp-In T-REx CSB KO with dox-inducible GFP-FLAG-CSB WT 8K>R	This study	100 µg/ml Hygromycin 15 µg/ml blasticidin
HEK293 Flp-In T-REx 19K>R	HEK293 Flp-In T-REx CSB KO with dox-inducible GFP-FLAG-CSB 19K>R	This study	100 µg/ml Hygromycin 15 µg/ml blasticidin
HEK293 Flp-In T-REx CSB∆UBD	HEK293 Flp-In T-REx CSB KO with dox-inducible GFP-FLAG-CSB∆UBD (lacking terminal 273 amino acids)	This study	100 µg/ml Hygromycin 15 µg/ml blasticidin
HEK293 Flp-In T-REx CSB USP2 WT	HEK293 Flp-In T-REx CSB KO with dox-inducible GFP-FLAG-CSB USP2 WT	This study	100 µg/ml Hygromycin 15 µg/ml blasticidin
HEK293 Flp-In T-REx CSB USP2 C276A	HEK293 Flp-In T-REx CSB KO with dox-inducible GFP-FLAG-CSB USP2 C276A	This study	100 µg/ml Hygromycin 15 µg/ml blasticidin
U2OS Flp-In T-REx CSB KO	U2OS Flp-In T-REx cell line with CRISPR-Cas9 targeted indel mutation of exon 5 of *ERCC6* gene (CSB) rendering it inactive.	This study	100 µg/ml Hygromycin 15 µg/ml blasticidin

Table 2-5 Mammalian cell lines used in this study

2.5.2 Growth conditions

Cells were usually cultured in DMEM as adherent monolayers in a 37°C humidified incubator with 5% CO_2 according to standard protocols. Cultures were split upon reaching approximately 80% confluency by dissociation from the plate by brief incubation with a trypsin/EDTA solution (Thermo Fisher Scientific) and a proportion of the cells were re-plated (usually 1:5 split every 2–3 days for HEK 293 cells). To freeze cells, 1 x 10^7 cells were resuspended in 1 ml freezing medium and transferred into cryogenic vials, and frozen using Nalgene Mr Frosty (Merck) suspended in isopropanol, which achieves a rate of cooling of 1°C/minute. Induction of gene expression was performed by addition of doxycycline to media at a final concentration of typically 1 ng/ml for 24–48 hours. Doxycycline was added to all cell lines even those absent a dox-inducible transgene.

Dulbecco's Modified Eagle's Medium (DMEM)

DMEM was obtained from Gibco (Thermo Fisher Scientific) and supplemented with 10% (v/v) tetracycline-free foetal bovine serum (Biosera) and penicillin (20

U/ml)/streptomycin (100 µg/mL), supplied as a 100X stock solution (Merck). Additional antibiotics were added as necessary and as outlined in Table **2-5**.

Freezing Medium
DMEM
10% (v/v) tetracycline-free foetal bovine serum
10% (v/v) DMSO

2.5.3 Lipid based transfection methods

Lipofectamine 3000 (Thermo Fisher Scientific) was used for the transfection of miniprepped DNA plasmids following manufacturer's protocol. Typically, 900 ng plasmid DNA and 100 ng pOG44 was transfected into a 6-well plate of 80% confluent cells (1 x 10^6) and 2 ml media without antibiotics and left to grow for 48 hours.

2.5.4 Establishment of stable cell lines

48–72 hours after transfection, cells were trypsinised and seeded into a three 15 cm dishes of increasing dilutions and left to attach overnight. The following day DMEM with the appropriate selection antibiotic (Table **2-5**) was added and replenished every 3–4 days. Cells were left to succumb to antibiotic-mediated death and monitored for the outgrowth antibiotic-resistant colonies of clonal cells. Once colonies had reached a size to be visible by eye, well separated colonies were isolated using cloning cylinders (Merck), trypsinised and each colony transferred to 24-well plates and expanded. About 12 colonies were isolated and grown for each cell line. To test for presence and expression of the transgene each cell population was seeded with and without doxycycline for 24 hours, harvested and whole cell extracts were prepared and analysed by western blotting (2.7). 2–3 different clonal cell populations were frozen and stored for future use per genotype. Saved cells were those with no leaky expression when grown without doxycycline and good induction when grown with doxycycline, ideally with similar expression levels between chosen populations.

2.6 CRISPR-Cas9 mediated genome editing

2.6.1 gRNA design

guide RNA (gRNA) sequences were designed at www.biorender.com against *ERCC6*. gRNAs were designed for use with the WT Cas9 and nickase Cas9 (requiring two gRNA's) proteins. Several gRNAs were chosen with consideration for the best combination of on-target and off-target score. DNA oligonucleotides encoding the gRNAs were ordered and synthesised from Integrated DNA Technologies. gRNAs were tested empirically in MRC5 cells for the best indel induction rate. The best gRNA for WT Cas9 and two gRNA's for nickase Cas9 were chosen based on the highest proportion of a single frameshift indel and taken forward for use in HEK 293 and U2OS cells.

2.6.2 Cloning of gRNA into Cas9 vector and transfection into cells

CRISPR-Cas9-nickase-mediated genome editing of HEK293 and U2OS Flp-In T-REx cells was performed as previously described (Ran, Hsu, Wright, et al., 2013). The oligonucleotides encoding gRNAs for targeting exon 2 (nickase Cas9) or exon 5 (WT Cas9) of *ERCC6* are listed below. Briefly, the forward and reverse strand oligonucleotides were annealed and ligated into pSpCas9WT(BB)-2A-GFP or pSpCas9n(BB)-2A-GFP linearized with BbsI, and plasmids were sequenced after cloning and transformation. To generate knockouts, cells were co-transfected with the two pSpCas9n(BB)-2A-GFP plasmids containing nickase-gRNA pairs 1 and 2 using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's instructions. 48 hours after transfection, high GFP-positive cells were sorted clonally by FACS into 96-well plates and grown until colonies were obtained. Clones were tested for the presence of CSB by western blotting and clones with complete absence of CSB were saved.

Primer	Sequence
ERCC6 sgRNA exon 5 Fwd	CACCGTCTGAGTATTTCCCCACAG
ERCC6 sgRNA exon 5 Rev	aaacCTGTGGGGAAATACTCAGAC
ERCC6 ngRNA exon2 #1 Fwd	CACCGCGTGGAGAAGGAGTATCGGT
ERCC6 ngRNA exon2 #1 Rev	aaacACCGATACTCCTTCTCCACGC
ERCC6 ngRNA exon2 #2 Fwd	CACCGCTCCACGTCAACGAGCTGGG
ERCC6 ngRNA exon2 #2 Rev	aaacCCCAGCTCGTTGACGTGGAGC

Table 2-6 DNA oligonucleotides encoding gRNA's cloned into Cas9 plasmids

2.6.3 TIDE analysis of genome editing of CSB KO cells

All cells that were absent for CSB protein by western blotting were further tested for the indel mutations created by Tracking of Indels by Decomposition (TIDE) analysis (Brinkman et al., 2014). Genomic DNA was extracted from cells and was amplified by PCR using primers aligning approximately 400 bp either side of the theoretical Cas9 cut site (gRNA alignment locus). A sequencing primer aligning approximately 200 bp upstream of the theoretical cut site was used for Sanger sequencing. The resulting chromatogram was uploaded to www.shinyapps.datacurators.nl/tide which provided information on the indels present. The cells with the highest proportion of a single frameshift indel were saved.

Primer name	Sequence
CSB Exon2 gDNA PCR Fwd	TGCTAACTTGGAAAACAGCAGC
CSB Exon2 gDNA PCR Rev	AGAAGATGGGCTGCACTCAC
CSB Exon5 gDNA PCR Fwd	ACCCTCGGAAAGTTTCATGCTA
CSB Exon5 gDNA PCR Rev	CAATCTCCTGCACTGGCACT
CSB Exon2 SangerSeq	TCCAGGACCGACCGATACTC
CSB Exon5 SangerSeq	GGCAGAGGCTTCAGCTCATA

Table 2-7 PCR and Sanger sequencing primers for checking Cas9-mediated genome editing

2.6.4 UV colony formation assay

A large population of cells was seeded into media containing 1 ng/ml doxycycline to induce CSB expression. After 24 hours induced cells were counted and seeded in duplicate at 2000 and 4000 cells per well of a 6-well plate. The next day cells were either mock-treated or UV-treated by removal of media and exposed to 4 J/m^2 UV-

C 254 nm irradiation using a custom conveyor belt machine. Doxycycline containing media was readded and cells were left to grow for 10–14 days. The experiment was stopped by removal of media, washing cells once with PBS and fixing the cells with 4% formaldehyde in PBS for 10 minutes at room temperature. Cells were washed two times with Milli-Q water and colonies were left to air dry. Colonies were stained *in situ* with 0.1% crystal violet solution for 30 minutes at room temperature. Excess crystal violet was removed by extensive washes with Milli-Q water. Colonies were left to air-dry and imaged with an Epson scanner.

2.7 Protein techniques

2.7.1 Whole-cell extracts

Cells were harvested by trypsinisation and pelleted by centrifugation. The cell pellet was washed with ice-cold PBS and pelleted again. Cell pellets were either snap frozen in liquid nitrogen and stored at -80ºC or processed immediately. Cell pellets were resuspended in ice-cold whole-cell extract buffer in 1.5 ml tubes. Typically, a 6-well plate of approximately 1 x 10^6 cells was lysed in 100–150 µl buffer, which was scaled up or down proportionally. Further lysis and shearing of genomic DNA (necessary to release chromatin-bound proteins) was facilitated by sonication of lysates in a 4ºC actively cooled water bath sonicator (Diagenode Bioruptor Plus). Lysates from 1 x 10^6 HEK 293 cells were sonicated on high for 7 cycles of 30 seconds on, 30 seconds off. The sonicated whole-cell extract (WCE) was clarified by centrifugation at 18, 000 x *g* for 20 minutes at 4ºC. The clarified supernatant was transferred to a fresh tube and snap frozen in liquid nitrogen and stored at -80ºC long-term. Typically, 30–40 µg of WCE was used for Western blot analysis.

Whole-cell extract buffer
50 mM Tris pH 7.5
500 mM NaCl
0.5% NP-40
5 mM EDTA
1 mM DTT
1X cOmplete protease inhibitor tablet (Merck)

2.7.2 Cell fractionation

Cell fractionation for soluble and chromatin extracts was performed using NE-PER Nuclear and Cytoplasmic Extraction Kit (Thermo Fisher Scientific) according to manufactures protocol with the following conditions. An 80% confluent 15cm dish was used per condition, which equates to approximately 2 x 10^7 cells. Cells were harvested by trypsinisation, pelleted, and after washing in PBS snap frozen in liquid nitrogen and stored at -80ºC until processing. Thawed pellets were resuspended in 1.2 ml of ice-cold CER I reagent, vigorously vortexed and incubated for 10 minutes on ice. 66 µl of ice-cold CER II reagent was added to the suspension, vortexed briefly, and incubated in ice for 1 minute. The suspension was centrifuged at 4ºC for 5 minutes at 16,000 x *g* and the supernatant (soluble extract) transferred to a fresh tube. 600 µl of ice-cold NER reagent was added to the insoluble pellet and resuspended. 250 U of BaseMuncher (Expedeon) and $MgCl_2$ to 1 mM final was added to digest DNA and fully release chromatin bound proteins and incubated on ice for 45 minutes with regular vortexing. Cell fractions were analysed by Western blotting.

2.7.3 Protein lysate quantification

Protein concentration of WCE and most cell lysates was determined by Pierce 660 nm protein assay (Thermo Fisher Scientific). 4–10 µl of WCE and 10 µl of protein standards were made up to 160 µl with Pierce 660 reagent in 96-well plates. Protein standards consisted of BSA in water at concentrations of 2, 1.5, 1, 0.75, 0.5, 0.25 mg/ml. The plate was shaken to mix and incubated for 5 minutes before reading the absorbance on a plate reader at 660 nm. Analysis of the data was done with excel and a standard curve was plotted with the absorbance of the protein standards. Protein concentration of samples was determined by interpolation of the standard curve using the samples corresponding absorbance.

For samples with buffer components not compatible with the Pierce 660 assay (e.g., purified proteins) the Qubit fluorometer protein assay (Thermo Fisher Scientific) was used according to manufacturer’s instructions.

2.7.4 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins in lysates were routinely separated by gel electrophoresis using precast 4–12% Bis-Tris gradient gels (Bio-Rad) with TGX running buffer (Bio-Rad). Precision Plus protein marker (Bio-Rad) was run on all gels to determine migration patterns of different sized protein markers from 10–250 kDa. To properly resolve high molecular weight proteins such as CSB and RPB1, 3–8% Tris-Acetate precast gels were used with the respective XT Tricine running buffer (Bio-Rad). Depending on the downstream application following SDS-PAGE, gels were either directly stained with InstantBlue Coomassie stain (Abcam) or further processed for western blot analysis.

Laemmli SDS-PAGE Loading Buffer
30% (v/v) glycerol
0.28 M Tris pH 6.8
1% (w/v) SDS
0.5M DTT
Bromophenol blue

TGX running buffer (Bio-Rad)
5 mM Tris pH 8.3
192 mM glycine
0.1% SDS

XT Tricine running buffer (Bio-Rad)
Proprietary formulation, which was used with 3–8% Criterion XT Tris-Acetate precast gels from Bio-Rad

2.7.5 Western blot analysis

After protein separation by SDS-PAGE, proteins were transferred from the gel to a nitrocellulose membrane (Amersham Protran Premium 0.2 NC) by wet electro-blotting. The wet protein transfer was set up by placing the membrane on top of the gel sandwiched between two pieces of Whatman paper (Cytiva, Thermo Fisher Scientific) and sponges. The sandwich was placed in a Bio-Rad Criterion Blotter filled with transfer buffer and an ice block and transferred for 30 minutes at

constant 100 V in a 4ºC cold room. The membrane was stained with Ponceau S solution (Merck) to stain all proteins and check transfer efficiency and loading evenness. The membrane was de-stained by rinsing in PBS for several minutes and incubated in blocking solution (5% milk in PBST) for 1 hour at room temperature. The membrane was then sliced, and the relevant portions placed into antibody solutions in 5% BSA with 0.02% Na-azide, which were incubated at 4ºC overnight with rotation.

For enhanced chemiluminescent (ECL) visualisation, membranes were incubated with HRP-conjugated secondary antibody diluted in blocking solution for 1 hour at room temperature, washed in PBST and incubated with an ECL reagent (Thermo Fisher Scientific SuperSignal West Pico or Dura) for 2 minutes. The chemiluminescent signal was detected by exposure of the membrane to Amersham Hyperfilm ECL (GE Healthcare). For near-infrared fluorescence visualisation, membranes were incubated in the dark with fluorescent dye-coupled secondary antibody diluted in blocking solution for 1 hour at room temperature. The membrane was washed extensively with PBST and imaged on the LI-COR Odyssey quantitative fluorescence imaging system.

Transfer buffer
25 mM Tris Base
192 mM glycine
10% (v/v) Methanol

PBS-Tween
137 mM NaCl
2.7 mM KCl
10 mM Na_2HPO_4
2 mM NaH_2PO_4
0.01% (v/v) Tween 20

Antibody	Antigen	Supplier
4H8	RPB1 phosphorylated-CTD	Abcam ab5408
8WG16	RPB1 unphosphorylated CTD	In-house
D8L4Y	RPB1 phosphorylated and phosphorylated CTD	Cell Signaling Technology #14958
StrepII tag	StrepII affinity tag	Abcam ab76949
SUPT5H	SPT5	Bethyl A300-869A
diGly	di-glycine ubiquitin motif	Cell Signaling Technology #5562
GFP	GFP	Abcam ab290
Vinculin	Vinculin	Sigma V9131
P4D1	Ubiquitin	Cell Signaling Technology #3936
CSA	CSA (ERCC8)	Abcam ab137033
CSB	CSB (ERCC6)	Bethyl A301-345A
Tubulin	Tubulin	In-house (Tat-1)
Histone H3	Histone H3	Abcam ab1971
Lamin A	Lamin A	Abcam ab133256

Table 2-8 Antibodies used in this study

2.8 Protein affinity chromatography

2.8.1 Extraction and purification of RNAPII from bovine calf thymus

Bovine RNAPII was purified as previously described with modifications (Hu et al., 2006). All steps were completed in a cold room at 4°C unless otherwise stated. Bovine calf thymus was homogenized for 3 min in buffer A using a 2 L blender (Waring). The homogenised material was centrifuged and the supernatant filtered through two layers of Bioprep nylon filter cloth. A 10% solution of polyethyleneimine, pH 7.8 was added to a final concentration of 0.05% and the material was stirred for 30 min, then centrifuged for 30 min at 12,000 x *g*. The resulting pellets were re-dissolved in buffer B. After centrifugation, the supernatant was loaded on a 120 ml Fast Flow Q Sepharose column (Cytiva), equilibrated in buffer B, by using a peristaltic pump at 5 ml/min. The column was washed with three column volumes of buffer B, followed by step-elution with buffer C. Eluates from the Fast Flow Q Sepharose column were analysed by western blotting for RPB1. Fractions with the highest yield of RPB1 were pooled and further purified using a 5 mL gravity flow column of 8WG16 (αRPB1 CTD) antibody-coupled Sepharose. The input was loaded overnight using a peristaltic pump at 0.1 ml/min. After application of the input material, the antibody column was washed with ten column volumes of buffer C, sealed, and allowed to equilibrate to room temperature (20–25 °C) for 15 min. RNAPII was eluted in batch using elution buffer, collecting 4 10 ml fractions at room temperature. The RNAPII-containing fractions were dialyzed against dialysis buffer overnight. A second elution method was attempted with RPB1 CTD peptides via competitive binding to the 8WG16 column. CTD peptides consisting of 4 heptad repeats (Biotin-eahx-$(YSPTSPS)_4$-COOH) were dissolved in PBS to a final concentration of 1 mg/ml. 5 ml of CTD elution buffer was applied to the column for 1 hour at RT with rotation before collection as eluate. This was repeated at 37 ºC for 1 hour and a final elution was done overnight at RT.

Buffer A
50 mM Tris pH 7.9
10 µM $ZnCl_2$
10% glycerol
cOmplete protease inhibitors (Merck)
PhosSTOP (Merck)

Buffer B
50 mM Tris pH 7.9
10 µM $ZnCl_2$
10% glycerol
150 mM $(NH4)_2SO_4$
cOmplete protease inhibitors (Merck)
PhosSTOP (Merck)

Buffer C
50 mM Tris pH 7.9
10 µM $ZnCl_2$
500 mM $(NH4)_2SO_4$
cOmplete protease inhibitors (Merck)
PhosSTOP (Merck)

Elution buffer
50 mM Tris pH7.8
10 µM $ZnCl_2$
500 mM $(NH4)_2SO_4$
40% 1,2 propanediol

Dialysis buffer
50 mM Tris pH 7.8
10 µM $ZnCl_2$
150 mM $(NH4)_2SO_4$)
5 mM DTT,
10% glycerol

2.8.2 Purification of StrepII-CSB WT from insect cells

Cell pellets described in 2.4.4 were thawed quickly in hands and resuspended in binding buffer at 1/10th of the volume of media they derived from (~ 2 x 107 cells /ml binding buffer). Suspension was lysed by dounce homogenisation with 10 strokes. Further lysis was performed with a probe-tip sonicator using the blunt probe (Branson Digital Sonifier 250) at 50% amplitude for 10 cycles of 15 sec on, 1 minute off. The lysate was clarified by ultracentrifugation at 125,000 x g for 1 hour at 4ºC. The clarified lysate was loaded onto an Äkta Explorer FPLC (Cytiva) using 1 ml StrepTrap HP columns (Cytiva). All applications of buffer to the column were conducted with a flow rate of 1 ml/minute. The column was equilibrated with 5 CV binding buffer before apply sample lysate containing StrepII-CSB to column. The column was washed with 10 CV wash buffer and re-equilibrated with 3 CV binding buffer or until conductance stabilised. Elution was done with elution buffer containing 2.5 mM desthiobiotin and 0.5 ml fractions were collected. A sample of fractions decided by UV absorbance were analysed by SDS-PAGE with InstantBlue Coomassie (Expedeon) staining of gels.

Binding buffer
100 mM Tris pH 8
150 mM NaCl
1 mM EDTA
10% glycerol
cOmplete protease inhibitors

Wash buffer
100 mM Tris pH 8
1 M NaCl
1 mM EDTA
0.01% NP-40
10% glycerol
cOmplete protease inhibitors

Elution buffer
100 mM Tris pH 8
150 mM NaCl
1 mM EDTA
2.5 mM desthiobiotin
10% glycerol
cOmplete protease inhibitors

2.8.3 Purification of His-tagged DSIF from *E. coli*

Cell pellets described in section 2.3.6 were thawed quickly in hands and then stored on ice. Cells were lysed in lysis buffer (5-10 volumes of pellet) for 5 minutes. Suspension was further lysed by sonication with probe-tip sonicator (Branson Digital Sonifier 250) at 20% amplitude with 8 cycles of 15 sec on, 30 sec off. Supernatant was clarified by centrifugation at 8,000 x *g* for 10 min. Ni-NTA beads (Qiagen) were pre-equilibrated with several volumes of wash buffer. An equal volume of Ni-NTA beads to the original cell pellet after induction was used for binding. Supernatant extract was loaded in batch on to Ni-NTA beads and incubated at 4ºC for 2–4 hours with rotation. Flowthrough was allowed to drain and the column was washed with 10 CV wash buffer before elution with 8 0.5 CV of elution buffer. All fractions were analysed by SDS-PAGE followed by Coomassie staining of the gel to check for the purity and concentration of fractions. The most concentrated fractions were dialysed against dialysis buffer overnight at 4ºC with stirring.

Lysis buffer
50 mM HEPES-KCl, pH 7.5
500 mM NaCl
10 µM ZnCl2
150 mM imidazole, pH 8.0
1 mM DTT
10% (v/v) glycerol
0.1 mM EDTA
0.1 mg/mL lysosyme
cOmplete protease inhibitor tablets (Merck)

Wash buffer
50 mM HEPES-KCl, pH 7.5
500 mM NaCl
10 µM ZnCl2
150 mM imidazole, pH 8.0
1 mM DTT

Elution buffer
50 mM HEPES-KCl, pH 7.5
500 mM NaCl
10 µM ZnCl2
400 mM imidazole, pH 8.0
1 mM DTT
10% (v/v) glycerol

Dialysis buffer
50 mM HEPES-KCl, pH 7.5
300 mM NaCl
10 µM ZnCl2
50 mM imidazole, pH 8.0
1 mM DTT
10% (v/v) glycerol

2.8.4 Purification of GST-Dsk2 and GST-MultiDsk proteins

The pellets described in 2.3.5 were thawed quickly at room temperature and transferred to ice.

Dsk2 cells were resuspended in 100 ml cold PBSA containing protease inhibitors. The suspension was sonicated with a tip probe sonicator (Branson Digital Sonifier 250) at 33% output, with 15 s ON, 30 s OFF pulses, for a total ON pulse duration of 10 min. Triton X-100 was added to the suspension to a final concentration of 0.5% and mixed gently, and lysate was then incubated on ice for 30 min and centrifuged at 18,000 x *g* for 10 min at 4°C to remove debris. Supernatant was added to the prewashed (2 washes in PBSA) glutathione sepharose beads. 1 ml of bead slurry (0.5 ml of packed beads) was used per 30 ml of cleared lysate. DTT was added to a final concentration of 2 mM and suspensions were incubated in the cold room for at least 4 hours or overnight with rotation. Then beads were pelleted at 500 x *g* for 5 min at 4°C and washed twice with ice-cold PBSA containing 0.1% Triton X-100 and protease inhibitors. Washed beads were resuspended to a 25% slurry in PBSA with 0.02% sodium azide and stored at 4°C.

MultiDsk cells were resuspended in 100 ml STE buffer (below) and incubated on ice for 15 minutes. Cells were further lysed by tip-probe sonication at 25% amplitude, 15 s ON, 30 s OFF pulses off for a total ON time of 1 minute and subsequently the cell extract was centrifuged at 18,000 x *g* for 30 minutes at 4ºC. The supernatant was collected and the final Triton-X concentration was raised to 3% and extract was incubated with 1 ml of pre-equilibrated glutathione sepharose bead slurry (0.5 ml of packed beads) per 30 ml of cleared lysate 1 ml of cleared lysate and incubated for 4 hours at 4ºC with rotation. The beads were pelleted at 500 x *g* for 5 min at 4°C and washed twice with ice-cold PBSA containing 0.1% Triton X-100 and protease inhibitors. Washed beads were resuspended to a 25% slurry in PBSA with 0.02% sodium azide and stored at 4°C.

STE buffer

10 mM Tris pH 8
100 mM NaCl
1 mM EDTA
2 mM DTT
1.5% Sarkosyl
0.1 mg/ml lysozyme
Protease inhibitors

2.9 *In vitro* transcription elongation complex assay

Transcription elongation complexes were set up as described (Saeki & Svejstrup, 2009) with modifications. RNAPII purified by Hannah Williams was used and purified as described in section 2.8.1.TECs were formed on a DNA transcription template using oligonucleotides listed in Table 2-9. Reactions were set up by incubating the transcribed strand (TS long/short) oligo with an RNA oligo template and allowing them to anneal. RNAPII was then and allowed to associate with the template before the biotin-labelled non-transcribed strand oligonucleotide (NTS long/short) was added. These biotin-tagged TEC complexes were incubated with streptavidin-coated magnetic Dynabeads (Thermo Fisher Scientific) and allowed to bind. Beads were washed with TXN wash buffer to remove any non-incorporated components and resuspended in TXN buffer. Finally, NTP's were added to start the transcription reaction (ATP and GTP for long oligos, or ATP only for short oligos) and incubated for 5 minutes in order to allow transcription until the G-rich region (or C-rich region with short oligos). Additional protein components were either added with RNAPII when forming TECs or added to the reaction after transcription had started so they could complex with already elongating RNAPII. TECs were captured using a DynaMag-2 Magnet (Thermo Fisher Scientific) and the supernatant aspirated. Laemmli buffer was added to each fraction and heated at 95ºC for 5 minutes before loading on a gel to perform SDS-PAGE and western blotting analysis.

TXN buffer
20 mM Tris pH 7.9
10 mM KCL
5 mM DTT
20 μM $ZnCl_2$
7 mM $MgCl_2$

TEC wash buffer
50 mM Tris pH7.9
500 mM NaCl
10 μM $ZnCl_2$
0.05% NP-40
10% glycerol

Oligo-nucleotide	Sequence
TS long	5′-GAGTTGGTTATGGTAGGTGAGTGTGTGATTGTGTGTTAGTGTGGTGTCGCTTGGGTTCTCTTTTCGCCTTGGGGGCTCCTCCTCCCTCCCTCTTTCCTGATGGCTGTTTGTTTCCTATAGCGTAGGCCTTAGACAATTGCGCATTCAGAC-3′
NTS long	5′-GTCTGAATGCGCAATTGTCTAAGGCCTACGCTATAGGAAACAAACAGCCATCAGGAAAGAGGGAGGGAGGAGGAGCCCCCAAGGCGAAAAGAGAACCCAAGCGACACTTCATTAACACACAATCACACACTCACCTACCATAACCAACTC-3′
RNA short	5′-UUUUUACAGCCAUC-3′
TS short	5′-GGCTCCTCCTCCCTCCCTCTTTCCTGATGGCTGTTTTTTTCCTATAGCGTAGGC-3′
NTS short	5′-GCCTACGCTATAGGAAACCTTGAAGTTAATCCAAAGAGGGAGGGAGGAGGAGCC-3′
RNA long	5′-ACCAGAACUACUUUUUACAGCCAUC-3′

Table 2-9 Oligonucleotides used for TEC assay. Red regions highlighted in TS is the induced stalling site of RNAPII when complementary NTP omitted. Regions highlighted in blue are annealing sites of TS and RNA.

2.10 5-ethynyl uridine assay to measure nascent RNA synthesis

5-EU incorporation assays were carried out as described previously (Williamson et al., 2017). MRC5 cells CS1AN cells expressing different CSB mutants were incubated for 48 hr with 1 ng/ml doxycycline. 2000 cells were seeded per well in triplicate into 96-well plates UV irradiated using a custom-built conveyor belt machine and allowed to recover for designated times. After 2- or 20-hours recovery, fresh media containing 0.75 mM 5-ethynyl uridine (5EU) was added and cells were incubated for 1 hour. 5EU-containing media was removed, and cells were fixed in PBS-buffered formaldehyde (3.7%) for 45 min at room temperature, washed once with PBS, followed by permeabilization with 0.5% TX-100 diluted in PBS for 30 min. Permeabilized cells were washed once with PBS, and then Alexa Fluor 647 Azide fluorophores were covalently attached to 5EU-labelled RNA by click reaction (100 mM Tris pH 8.5, 4 mM Cu_2SO_4, 10 µM Alexa azide 647, 100 mM ascorbic acid) for 1 hour at room temperature. Cells were washed 3 times with 100 mM Tris, pH 7.5 and stained with 4′,6-diamidino-2-phenylindole dihydrochloride (DAPI) at a final concentration of 1 µg/ml. Cells were washed once with PBS. Automated image acquisition of 6 fields per well was performed (Cellomics Array Scan VTI, Thermo Fisher Scientific) using a 10X objective. Image analysis was performed using HCS Studio 2.0. Cell nuclei were masked using the DAPI staining. The average intensity of Alexa Fluor 647-conjugated 5EU-labeled RNA was measured for each nucleus in at least 3 separate wells and plotted.

2.11 Incucyte imaging of live-cell growth

HEK293 cells (CSB WT, CSB 8R, CSB 19R, CSB KO, CSA KO) were seeded with 1 ng/ml doxycycline for 48 hours to induce gene expression. Cells were trypsinised and seeded in triplicate at 5,000–20,000 cells per well into poly-lysine coated 96-well plates and left to attach overnight. Cell media was removed, and cells were mock-treated or UV-treated with 20 J/m^2 254 nm UV-C irradiation using a custom conveyor belt machine and doxycycline media replaced. Growth was monitored and recorded by imaging of 4 fields of view of each well with 10x objective every 3 hours using Incucyte S3 system (Essen BioScience). A mask of the plate surface

area covered by cells in the field of view was created and applied using Incucyte image analysis software (Essen BioScience) and recorded for every image (4 per well, every 3 hours).

2.12 Ubiquitylated protein enrichment with Dsk2 resin

This was performed as described previously (Anindya et al., 2007; Tufegdzic Vidakovic et al., 2019). Dsk2 beads were prepared as described in 2.8.4. Cells were grown in 15 cm dishes with 1 ng/ml doxycycline for 48 hours before relevant UV treatments. After relevant recovery time after UV irradiation, cells were collected by scraping, centrifugation at 700 x g, rinsing in PBS and snap freezing of cell pellets in liquid nitrogen. Cell pellets were thawed and lysed in 1 ml TENT buffer with 250 U BaseMuncher (Expedeon) on ice for 45 min. Lysates were clarified by centrifugation at 18,000 x *g* for 20 minutes at 4ºC.

Dsk2 beads were pre-washed in TENT buffer. 50 µl of beads were used per sample of 1 mg lysate. Samples were normalised to 1 mg in 500 µl and added to beads and rotated for 2 hours at RT. The beads were then washed twice with 1 ml of TENT buffer before elution by heating at 95ºC in 50 µl Laemmli buffer for 5 minutes. 25 µl of Laemmli buffer containing ubiquitylated proteins were submitted to SDS-PAGE and western blot analysis 2.7.4).

TENT buffer
50 mM Tris-HCl pH 7.4
150 mM NaCl
2 mM EDTA
1% Triton X-100
cOmplete protease inhibitors (Merck)
PhosSTOP (Merck)
5 mM N-ethylmaleimide

2.13 Deep sequencing

2.13.1 4SU-labelled RNA generation and purification

The protocol was followed exactly as described in (Gregersen et al., 2020) after cell treatments. 2 x 10^6 cells of HEK293 cells (CSB WT, CSB 19R, CSB KO) were seeded in duplicate in 10 cm dishes and CSB transgenes induced for 48 hours with 1 ng/ml doxycycline. Next the media was removed, and cells were mock-treated or UV-treated with 20 J/m^2 254 nm UV-C irradiation using a custom conveyor belt machine and the media was readded. After the specified recovery time (untreated, 1 hour, 5 hours, 24 hours), 4-thiouridine (4SU) was added directly to the tissue culture media to a final concentration of 1mM for 15 minutes. The reaction was stopped by removing the media and lysis of cells directly in the dish by addition of 1 ml of TRIzol reagent (Thermo Fisher Scientific).

Cells in TRIzol solution were transferred to tubes and 200 µl UltraPure Phenol/Chloroform/Isoamyl alcohol 25:24:1 (Thermo Fisher) was added and mixed by vigorous shaking. After centrifugation at 12,000 x *g* for 15 min at 4ºC, the upper aqueous phase was transferred to a Phase-Lock-Gel tube (Qiagen). An equal volume of chloroform:isoamyl alcohol 24:1 was added and centrifuged for 5 minutes at 12,000 x *g*, 4°C. The upper aqueous phase was transferred to a new tube and 1.1 volume of isopropanol was added and incubated at room temperature for 20 min. The mixture was centrifuged at 7,500 x *g* for 5 minutes at 4ºC. The RNA pellet was washed in 85% ethanol, centrifuged at 7,500 x *g* for 5 min and all ethanol was removed and pellet air dried. The pellet was resuspended in 100 µl RNase free water and the RNA concentration was measured using Qubit BR RNA assay.
150 µg of human RNA was mixed with 1.5 µg yeast spike-in RNA in a total volume of 100 µl. In order to fragment the RNA, 20µl of 1 M NaOH was added and incubated for 20 minutes on ice. Fragmentation was rapidly slowed by addition of 80 µl 1 M Tris pH 6.8 and the reaction was fully quenched by running the sample twice through RNase-free p30 Bio-Rad spin columns in Tris pH 7.4. 3 µl BB buffer (2.5 µl 1 M Tris pH 7.4 and 0.5 µl EDTA) was added to fragmented RNA. Fragmented RNA was biotinylated by adding 50 µl 0.1 mg/ml MTSEA biotin-XX

linker dissolved in dimethylformamide and incubating the reaction for 30 min at RT in the dark. The biotinylated RNA was purified from excess biotin by using Phase-Lock-Gel tubes (Qiagen). An equal volume of UltraPure Phenol/Chloroform/Isoamyl alcohol 25:24:1 (Thermo Fisher) was added. After centrifugation at 12,000 x *g* for 5 minutes at 4°C, the upper aqueous phase was transferred to a Phase-Lock-Gel tube. RNA was precipitated by adding 1/10th volume of 5 M NaCl and 1.1 volume of isopropanol and incubated at room temperature for 10 minutes. The mixture was centrifuged at 20,000 x *g* for 20 minutes at 4°C. The RNA pellet was washed in 500 µl 85% ethanol and centrifuged at 20,000 x *g* for 5 minutes. All ethanol was removed, and the pellet was air dried before resuspending in 50 µl RNase free water.

Confirmation of 4SU incorporation was measured by dropping 5 µl of diluted RNA (2:5) onto a N+ membrane, UV crosslinked twice at 2000 µJ and blocked for 20 min at RT in blocking buffer (10% SDS, 1 mM EDTA in PBS). Membrane was probed with 1:50,000 dilution of 1 mg/ml streptavidin-horseradish peroxidase antibody (Pierce) in blocking solution for 15 minutes. The membrane was washed six times in PBS containing decreasing concentrations of SDS (10%, 1% and 0.1%, applied twice each). The signal of biotin-bound HRP was visualized by ECL detection 2.7.5).

Purification of biotinylated 4SU-labelled RNA from unlabelled RNA was performed using µMACS Streptavidin MicroBeads (Miltenyi Biotech) according to manufacturer's instructions. Briefly, samples were denatured at 65°C for 10 minutes followed by rapid cooling on ice for 5 minutes. 200 µL µMAC beads were added to each sample, mixed and left to incubated while rotating at RT for 15 minutes. Separation was carried out on magnetic columns provided and washed with 500 µL 55ºC wash buffer (100 mM Tris-HCl, pH 7.4, 10 mM EDTA, 1 M NaCl, 0.1% Tween-20). 4SU-labelled RNA was eluted by the addition of 100 µl freshly prepared 100 mM DTT followed by a second elution 5 minutes later. RNA was concentrated using the RNeasy MinElute Spin columns (Qiagen). RNA was eluted in 15 µL RNase free water and 1 µL was quantified by Qubit HS RNA assay.

2.13.2 Sequencing library preparation and deep sequencing

RNA samples were analysed for fragmentation and quality on Agilent TapeStation 4200. Samples were normalised and cDNA sequencing libraries were prepared using the KAPA RNA Hyper Prep kit with the following modifications:

- An initial incubation with the 2x fragment, prime and elute buffer of 1 minute at 65ºC
- During the 1st post-ligation cleanup, a 0.95X bead-based cleanup was used
- During the 2nd post-ligation cleanup, a 1X bead-based cleanup was used

Libraries were sequenced on an Illumina HiSeq 4000 as single-end 75 bp reads. Approximately 50 million reads were obtained per sample for 24 samples.

2.13.3 Bioinformatic analysis

All bioinformatic analysis of sequencing data was performed by Richard Mitter, Bioinformatics and Biostatistics Facility, The Francis Crick Institute.

Reads were aligned against the Homo sapiens GRCh38 and Saccharomyces cerevisiae sacCer3 genome builds using STAR v2.5.2a (Dobin et al., 2013) with Ensembl release 86 transcript annotations. For the purposes of visualization, genome alignment BAM files were created at the single sample level and also merged across biological replicates, sorted and indexed using Picard v2.1.1 (broadinstitute.github.io/picard). The yeast spike-in was used to account for differences in library size between samples. A yeast gene-level counts matrix was generated using the the SummarizeOverlaps function (mode = ”Union,” ignore.strand = FALSE) from the Bioconductor package GenomicAlignments (Lawrence et al., 2013). The matrix was passed to DESeq2′s estimateSizeFactors function (Love et al., 2014) to generate sample scale factors. BigWig files were generated by converting BAM files to bedGraph format using BEDtools’ genomeCoverageBed function (Quinlan & Hall, 2010) separately for each strand. Yeast scale factors were applied to normalize for differences in library size using the “-scale” argument. bedGraph files were in turn converted to bigWig format using the bedGraphToBigWig function from the KentTools package

(github.com/ucscGenomeBrowser/kent). An additional set of BigWig files was created where each treatment was normalised to the corresponding untreated control sample. This was conducted using deepTools v2.5.3's bigwigCompare function(Ramírez et al., 2014). A pseudocount of 1 was added to the original yeast scaled treatment and control bigwigs prior to calculation of a read-depth ratio. Average read depth profiles of the TSS region (-5kb:+100kb) were created for each bigwig from protein-coding genes from standard chromosomes (n= 19,919) using deepTools v2.5.3 with a bin size of 100 (Ramírez et al., 2014). Traveling ratios were calculated as in (Rahl et al., 2010). Briefly, each gene was divided into i) a promoter-proximal bin −30 bp to +300 bp around its TSS and ii) a gene body bin to the TTS. The traveling ratio is the ratio of read density in the promoter-proximal bin to that in the gene body. Genes shorter than 1kb and any that didn't have at least 1 mapped read in the promoter and gene-body consistently across all samples were considered to not be expressed and filtered out.

Chapter 3. Results I – Reconstituted transcription as an assay for DSIF and CSB interactions on RNAPII

3.1 Introduction

The molecular mechanism of how CSB functions in TC-NER is unknown. What is known is that CSB binds to RNAPII when it is stalled at DNA lesions such as CPD's (Fousteri et al., 2006; D Tantin et al., 1997; Van Den Boom et al., 2004; van der Weegen et al., 2020; J. Xu et al., 2017). As RNAPII stalls in such a way that conceals the lesion inside its active site, it must be removed or remodelled in such a way that allows access to the lesion and surrounding DNA by downstream repair factors, such as the endonucleases XPF and XPG. Based on its translocase activity as well as its direct interaction with RNAPII, CSB has been speculated to be the putative factor for this role.

Interestingly, genetic data indicate RNAPII elongation factors can inhibit the yeast CSB homolog, Rad26, in rescuing cells from UV sensitivity. A genetic screen in *S. cerevisiae* initially identified Spt4 (a subunit of the transcription elongation factor DSIF) depleted cells, concurrent with Rad26 depletion, as being less UV sensitive than Rad26-deficient cells alone (Jansen et al., 2000). Structural studies of elongating RNAPII bound to DSIF (Bernecky et al., 2016, 2017; Klein et al., 2011; Martinez-Rucobo et al., 2011) and RNAPII bound to Rad26, (W. Wang et al., 2018; J. Xu et al., 2017) reveal overlapping binding sites on RNAPII. Due to the high homology of DSIF, RNAPII, and the ATPase domain of Rad26, from yeast to human, it seemed likely that this structural overlap was conserved in human cells.

We therefore hypothesised that the role of CSB might be to remove or 'strip' RNAPII of the elongation factor DSIF that locks it into a forward translocating position. Removal of DSIF might then facilitate remodelling of RNAPII by way of backtracking or removal from DNA to expose the lesion for downstream repair factors.

To answer this question, an *in vitro* reconstituted transcription elongation assay was employed. In this system, a biotin-tagged DNA template is bound to

magnetic streptavidin beads to which purified RNAPII and an RNA template are added to form transcription elongation complexes (TECs). The original protocol for this assay employed radioactively-labelled nucleotides to label RNA during transcription and RNA extracted from these reactions was analysed by separation on a polyacrylamide gel and imaging by autoradiography. Based on the different RNA products produced under different conditions, conclusions about the transcription activity could be inferred. Because transcription reactions take place on magnetic beads that can be easily isolated and seperated from the reaction buffer, it was conceived as possible to analyse the reaction buffer for proteins that may have been removed from the DNA template (affixed to the beads) during the reaction. I therefore attempted to adapt this assay to assess the ability of CSB to remove DSIF from transcribing RNAPII by analysing the proteins of the bead and reaction buffer fractions by SDS-PAGE and Western blot.

3.2 Assay set-up

The *in vitro* transcription elongation assay has been employed in the lab previously to test the effect of different proteins and conditions using a yeast transcription system (Sigurdsson et al., 2010). Previously, RNA transcripts that were produced over the course of the reaction were isolated and compared using denaturing PAGE to qualitatively assess the effect of different conditions on RNA production.

As RNA production wouldn't yield useful information about the protein interactions with RNAPII, I decided to isolate protein from the bead fraction containing the transcription elongation complexes (TECs) and the supernatant that would contain proteins displaced during the reaction (see Figure 3-1).

To take advantage of this assay, several proteins needed to be purified. RNAPII, CSB and DSIF purification are outlined below and described in section 2.8 Protein affinity chromatography.

.

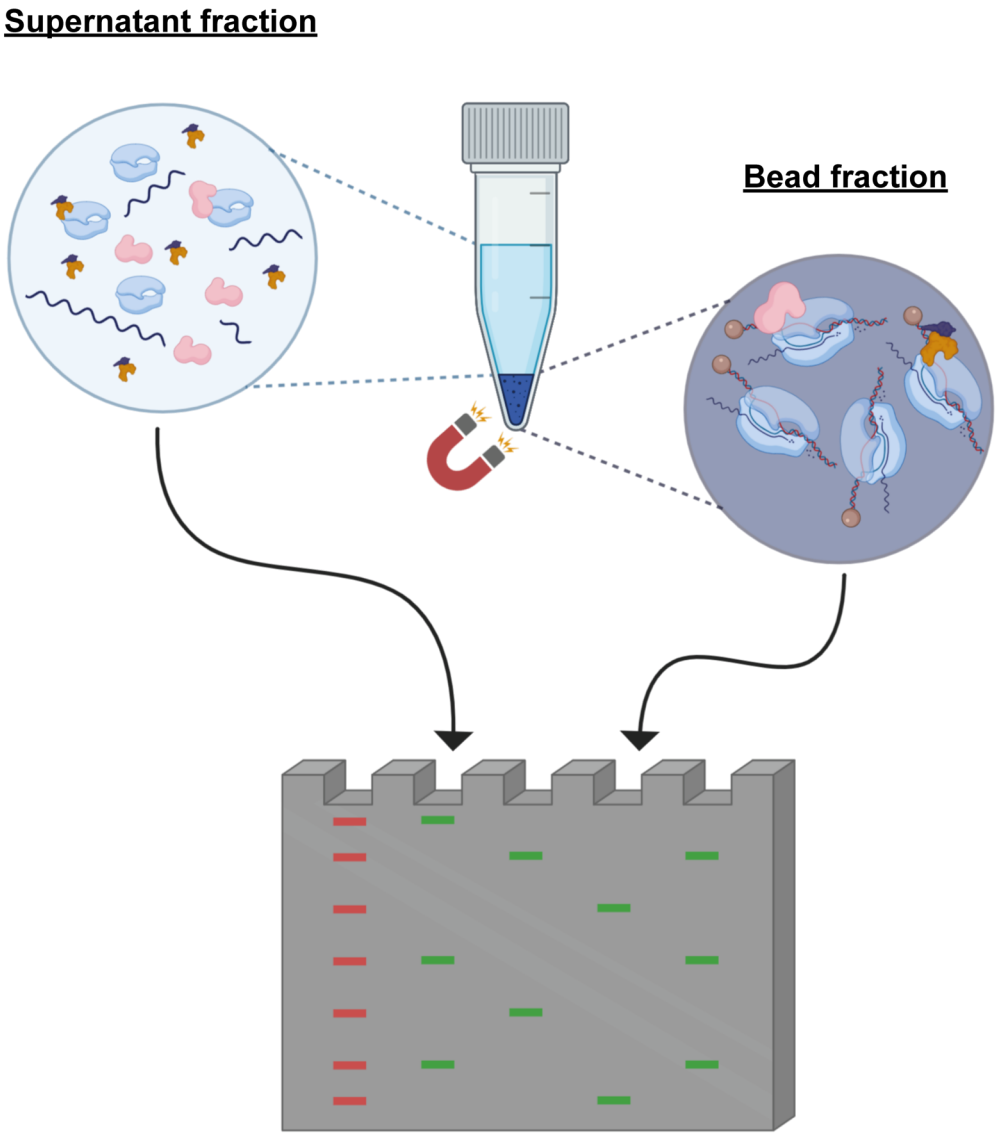


Figure 3-1 Cartoon model of TEC reaction set-up and analysis. RNAPII (blue) is engaged in fully formed TECs immobilised on magnetic beads (brown spheres) in bead fraction. Binding of CSB (pink) and DSIF (gold and blue) is representative of established binding location on RNAPII. Unbound, released, or actively dissociated protein and RNA exist in the supernatant fraction. These fractions can be separated by a magnet and each analysed by SDS-PAGE and western blotting.

3.3 RNAPII purification

The proteins for this assay needed to be purified in large quantities to facilitate optimisation of the TEC assay for RNAPII factor displacement.

Firstly, I attempted to purify RNAPII from bovine calf thymus. This is a long-standing protocol that is well established for isolating RNAPII complexes competent for *in vitro* transcription without the need for additional initiating factors.

A detailed protocol is outlined in section 2.8.1. Briefly, calf thymus polymerase was enriched using 150 mM ammonium sulphate precipitation and the lysates loaded onto a Q Sepharose anion exchange column. The most concentrated eluates were pooled (see Figure 3-2A and 2B) and loaded onto an agarose column crosslinked with 8WG16, a monoclonal antibody raised against the C-terminal domain (CTD) of RPB1, the main catalytic and largest subunit of RNAPII. RNAPII elution was attempted with propylene glycol or high concentrations of synthetic CTD peptide. Unfortunately, elution of RNAPII from the antibody column proved difficult (see Figure 3-2B). Elution's were attempted at 25ºC, 37ºC and overnight incubation with agitation of the beads with elution buffer, but none resulted in sufficient yields that were deemed necessary for productive TEC assays i.e., Coomassie stainable amounts. These problems are well known with mammalian RNAPII and are thought to be due to the highly repetitive antigen (52 repeats in the CTD) binding very strongly to the antibody compared to the more amenable yeast RNAPII (only 26 CTD repeats).

Given the difficulty in purifying sufficient quantities, the RNAPII used in subsequent assays was a preparation that was purified several years earlier by Barbara Svejstrup and was kindly gifted. This preparation was separated by SDS-PAGE and stained with Coomassie, which showed the presence of the RPB1 and RPB2 (with other subunits staining much more weakly), with no obvious signs of degradation (see Figure 3-2C). It was previously successfully used in TEC assays by Hannah Williams and shown to sufficiently produce full-length RNA products (data not shown).

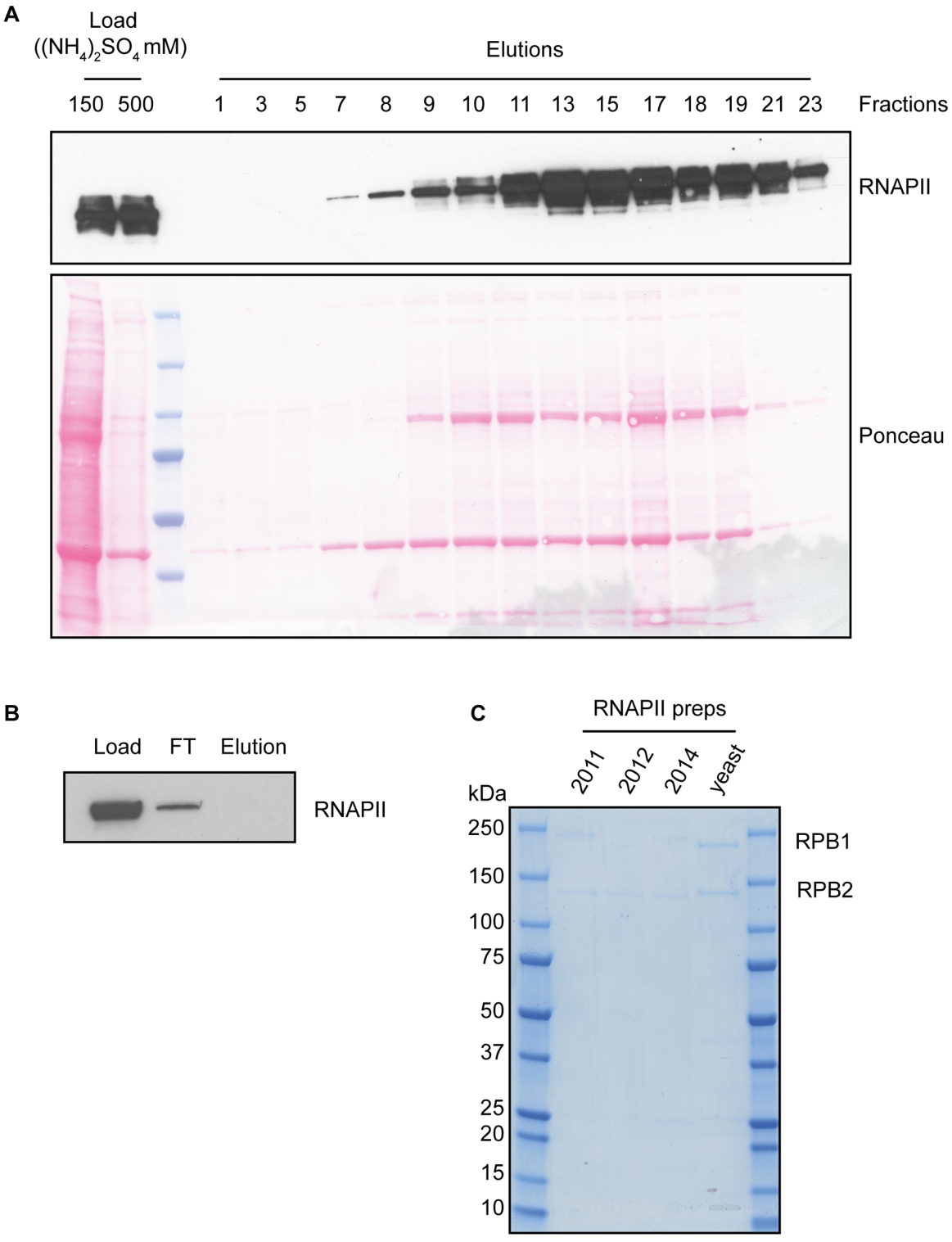


Figure 3-2 RNAPII expression and purification. A Western blot of RNAPII extraction using 8WG16 antibody against RPB1 CTD. 500 mM back-extract shows a higher extraction of RNAPII relative to total protein as seen in ponceau stain. **B** Western blot with 8WG16 antibody. Pooling of eluates 8–23 from **A** is labelled as load. FT is flowthrough. Elution with propylene glycol at 25ºC for 2 hours. **C** Coomassie stain of different calf thymus (year labelled) and yeast RNAPII purifications.

3.4 DSIF purification

A plasmid containing the coding sequence of human DSIF was kindly gifted from Patrick Cramer's laboratory. A modified pETduet vector contained in one cassette: SPT4 coding sequence (SUPT4H1) tagged with 10 histidine residues and a Sumo tag at the C-terminal separated by a 3C protease site to aid purification of the full complex. The other cassette contained SPT5 coding sequence (SUPT5H) only. The use of two cassettes is thought to support equimolar expression of both subunits for optimal complex formation. Transformation and expression 2.3.6) were carried out in BL21 bacterial cells and purification carried out on a Ni-NTA column as described in section 2.8.3.

While this preparation yielded sufficient yields of highly purified DSIF, SPT5–the largest subunit–was subject to degradation, producing a distinct lower-weight product (see Figure 3-3 band labelled *). Consultation with Carrie Bernecky, who had previously optimised this purification protocol in the Cramer laboratory, revealed that this was a common problem and she had determined the sequence of the degradation product by Edman degradation sequencing. This revealed a truncation of the first 140 amino acid residues of the acidic N-terminus domain, which is disordered and doesn't resolve in structures of DSIF and RNAPII, suggesting it doesn't contact RNAPII and may not be important for transcription (Bernecky et al., 2017). Other functional genetic studies in yeast have shown that a mutant of the yeast homologue of SPT5 lacking the acidic N-terminal domain has no defects in promoting *in vitro* transcription of the HIV-1 LTR (Ivanov et al., 2000). I therefore decided that for preliminary experiments it would be acceptable to use this preparation for the assays.

Cleavage of the histidine tag with a 3C protease was successful, as seen by the disappearance of the 30 kDa His-SUMO-SPT4 band and appearance of 13 and 16 kDa bands, corresponding to SPT4 and the His-SUMO tag, respectively. The 3C protease appears as 24 kDa band (see Figure 3-3B). As the protease is His-tagged it should be possible to remove it from the protein sample by simply running the eluate over a Ni-NTA column. Although the protease did bind the Ni-NTA column, so too did most of the non-tagged DSIF complex. Increasing concentrations of imidazole in the binding buffer did not increase DSIF in the

flowthrough while retaining the protease on the column. This could be caused by a natural affinity of DSIF for the Ni-NTA column, although I believe it is more likely caused by the SUMO protease binding to DSIF non-specifically.

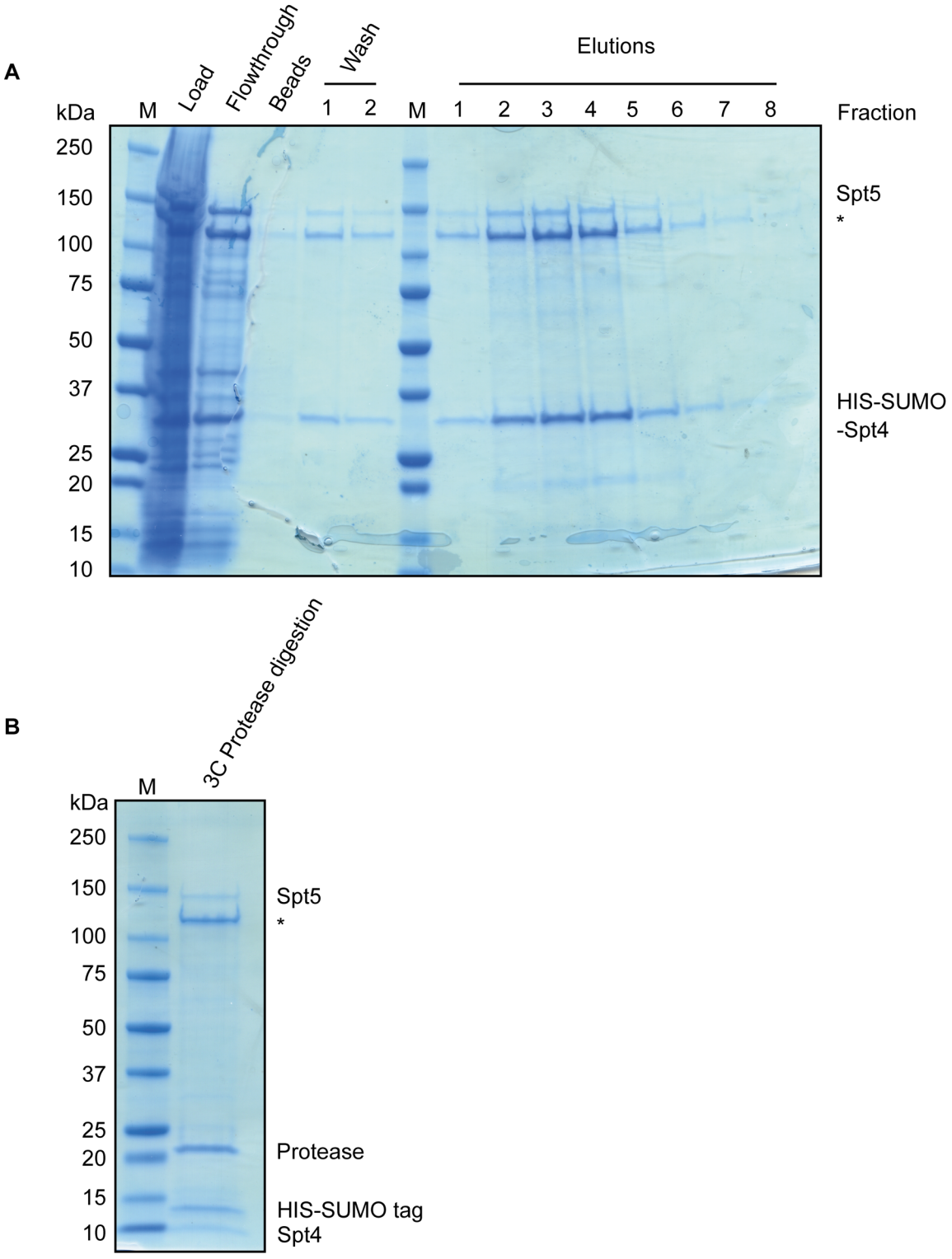


Figure 3-3 DSIF expression and purification. A Coomassie stained gel of fractions of DSIF purification. Full-length SPT5 is labelled such and * is N-terminal truncation. **B** Elution from Ni-NTA column after cleavage of HIS-SUMO tag from SPT4 showing continued presence of HIS-SUMO tag and HIS-tagged protease with DSIF, should not bind column after successful tag cleavage.

3.5 CSB purification

CSB has proven in the past to be a difficult protein to purify due to low levels of expression in human cells and the intrinsic instability of the protein leading to degradation (Svejstrup lab, unpublished observations; Christopher P. Selby & Sancar, 1997a, 1997b). The baculovirus expression system is a well-established method that generally supports significant overexpression of recombinant protein. This is accomplished through the ease of growing large amounts of *Spodoptera frugiperda* cells (SF9/SF21) as they are grown in suspension culture and shed baculovirus into the growth medium that can be easily harvested for infection of large amounts of cells for expression.

The StrepII tag is an engineered version of streptavidin that exploits the high affinity and specificity of the streptavidin-biotin interaction for affinity chromatography protein purification. Due to its small size of only eight amino acids, it also seemed less likely to interfere with the biochemical properties of CSB. Two StrepII tags separated by a flexible glycine linker (GGGSGGGSGGSA) were fused to the N-terminal of CSB, upstream of a 3C protease cleavage site.

This construct was cloned into a pFL plasmid (Fitzgerald et al., 2006) which was transformed into DH10Bac competent cells where recombination of the CSB gene from pFL into the bacmid takes place. Bacmid DNA was isolated as described in 2.3.4 and PCR amplification over the recombination site was performed to confirm the insertion of CSB DNA, which appears as a 5000 kb fragment (see Figure 3-4A). Bacmid DNA was transfected into SF9 cells to generate baculovirus and the media was harvested after a few days (P_1). Baculovirus titres were concentrated by infecting SF9 cells with P_1 titre, growing cells and harvesting the media and repeating until P_3 was obtained (2.4.3).

To optimise for maximal expression of CSB, different concentrations of P_3 were used to infect cells, which were harvested at different timepoints post-infection and protein was extracted and analysed by SDS-PAGE and Western blot. This showed a maximal expression of CSB at 72 hours post-infection with a 1:100 dilution of P_3, which was used for subsequent large-scale expression (see

section 2.4.4 and Figure 3-4B). After large-scale expression, cells were harvested and lysed in CSB binding buffer with sonication. The clarified lysate was loaded on to a StrepTrap column using an ÄKTA Pure FPLC machine. Washes with 1 M NaCl showed that binding of StrepII-CSB to the column was very specific (fractions 10 and 11 are still mostly flowthrough rather than wash as the conductance, a proxy for NaCl concentration, doesn't reach a peak until fraction 12 (see Figure 3-4C). StrepII-CSB was eluted from the column with 2.5 mM desthiobiotin. Coomassie staining of the eluate separated on a polyacrylamide gel showed the fast majority of protein in the eluate was full length CSB protein with some minor contaminating bands or degradation products (see Figure 3-4E).

To determine the nature of the faster migrating bands, the eluate was run and separated by SDS-PAGE. Western blotting with an antibody raised against the StrepII tag revealed an identical pattern to that seen in the Coomassie-stained gel, indicating that these bands were CSB degradation products with an intact N-terminal StrepII tag (see Figure 3-4D). As these degradation products constituted a very small amount of the eluate and there were no other contaminating proteins, this preparation was deemed pure enough to be used in subsequent experiments.

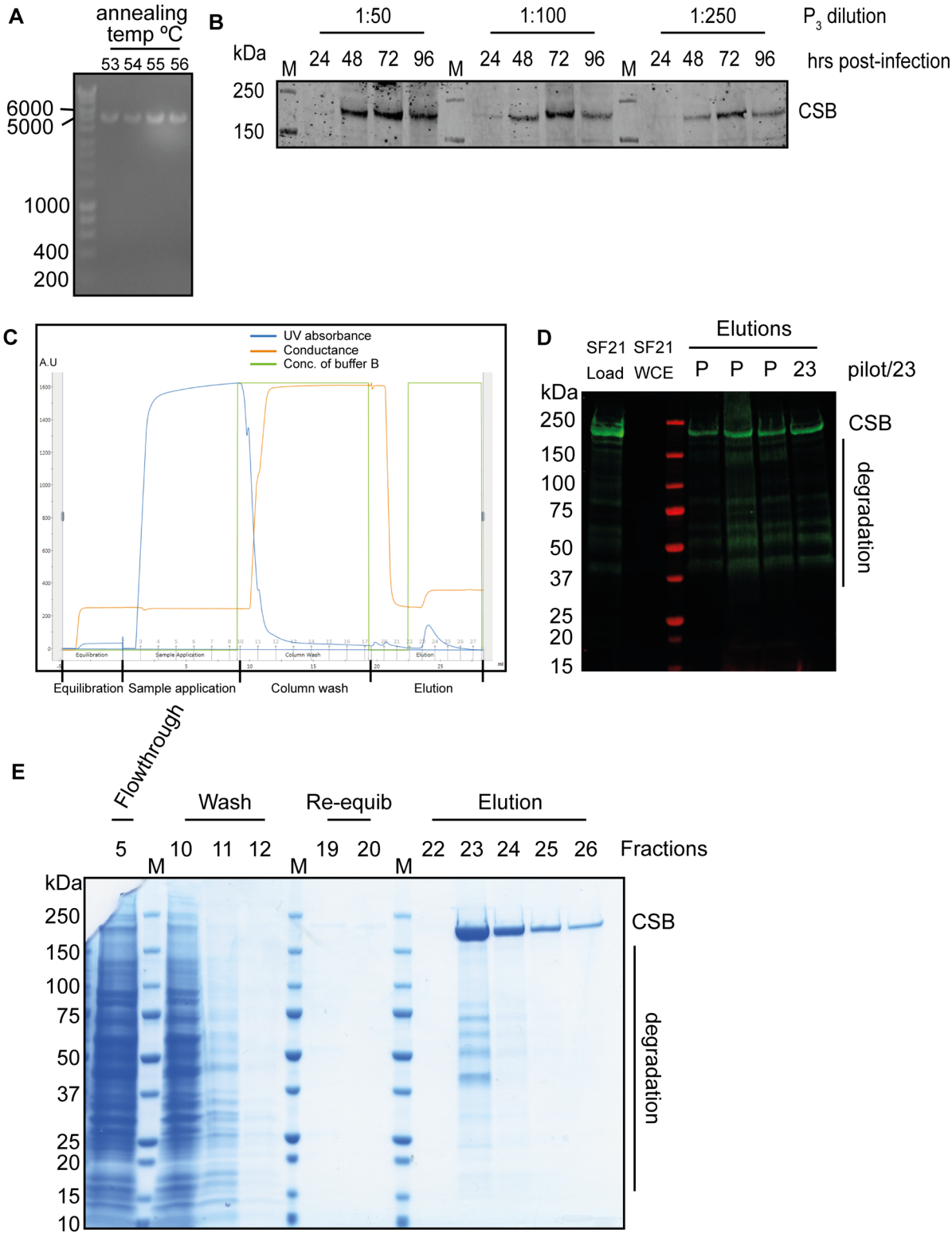


Figure 3-4 CSB baculovirus generation, expression and purification. A PCR of bacmid DNA showing amplicon corresponding to size of CSB coding sequence. **B** Western blot of CSB showing expression at different timepoints following infection of different dilutions of P_3 baculovirus. Lanes marked 'M' are marker. **C** HPLC graph showing UV absorbance, conductance and concentration of wash buffer (B) over volume. Numbers in grey are fractions. **D** Western blot using LI-COR system with StrepII antibody of eluates from pilot purification preparations and fraction 23 as shown in **E**. Lane 2 is whole cell extract of uninfected SF21 cells showing no background signal from antibody. **E** Coomassie stained SDS-PAGE gel of 10 µl of different fractions.

3.6 TEC assay optimisation

TEC complexes were set up as described 2.9, adapted from (Saeki & Svejstrup, 2009) and represented in Figure 3-5. Briefly, a 150 nt single-stranded DNA molecule was incubated with a 14 nt RNA template so that they anneal and form the transcription template to which RNAPII will bind. RNAPII purified from calf thymus was added to associate with the template before a biotin-labelled non-transcribed single-stranded DNA molecule was added. These biotin-tagged TEC complexes were added to streptavidin-coated magnetic beads and allowed to bind. Beads were washed with transcription wash buffer to remove any non-incorporated components and resuspended in transcription buffer. Finally NTP's were added to start the transcription reaction. Transcription to certain positions can be controlled through the omittance of certain NTP's. A G-stop (a guanosine-rich region that acts as a stall site due to lack of CTP) was included in the transcribed strand, 23 nucleotides downstream of the 3′ end of the RNA template. TEC assays can be started by 'walking' the polymerase up to the G-stop by omitting cytosine (CTP) and this was used to mimic a stalled polymerase, such as one stalled at a bulky lesion. Additional protein components were either added with RNAPII when forming TECs or added to the reaction after transcription had started so they could complex with already elongating RNAPII

RNAPII will stall without incorportation of an NTP for RNA polymerisation (Noe Gonzalez et al., 2021) and a stalled RNAPII is recognised and engaged by CSB regardless of the stalling signal (CPD or non-incorporation) and whether or not it triggers TC-NER (J. Xu et al., 2017). Hence, it was deemed appropriate that an impassable stalling signal that would trigger CSB binding, would be sufficient to study the ability of CSB to displace DSIF from RNAPII.

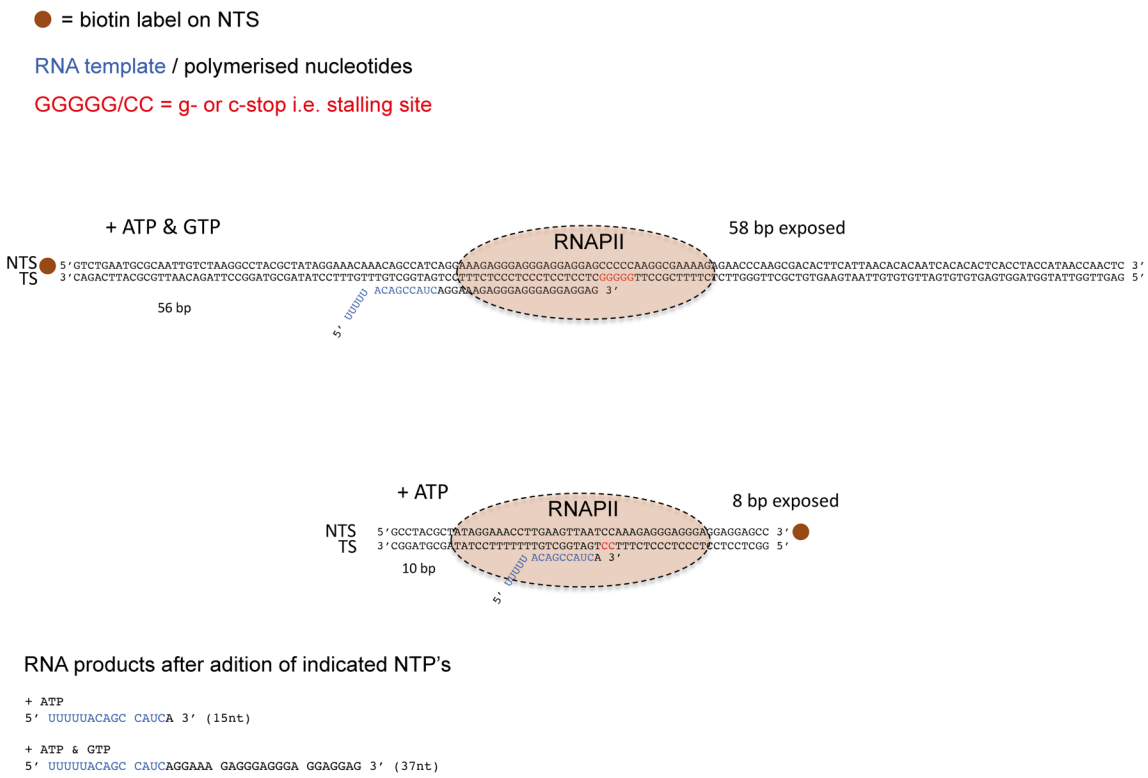


Figure 3-5 A schematic of fully assembled TEC complexes

RNAPII molecules were bound to a biotinylated DNA template with an RNA template and complexed with magnetic beads. Transcription was initiated with the addition of NTP's, the composition of which dictates to where on the DNA template RNAPII transcribes to before stalling. The standard assay was set up with 150 nt DNA oligonucleotides (top) and the addition of ATP and GTP that allowed transcription of 23 nucleotides up to a G-stop. Shorter DNA and longer RNA templates were also tested. The RNA template is highlighted in blue with black letters indicating polymerised nucleotides after initiation of transcription. Highlighted in red are the stalling signals on the transcribed strand. RNAPII is drawn to scale to show approximately the length of DNA it conceals.

3.6.1 Initial experiments

For initial experiments, TECs were assembled as described, added to streptavidin beads, and washed thoroughly. Transcription was started with the addition of NTPs less CTP and DSIF was added at the same time so that it would bind to an elongating polymerase. Reactions were incubated for 5 minutes at room temperature to allow transcription up to the G-stop. Beads were then immobilised and the reaction buffer, containing unbound DSIF, removed. Reactions were resuspended in transcription buffer, to which CSB was added and reactions incubated at room temperature for 5 minutes. Beads were then immobilised, and reaction buffer removed. Both fractions were boiled with Laemmli buffer and separated by SDS-PAGE and the fractions analysed for SPT5 by western blotting, with the relative amounts of SPT5 in each fraction assumed to represent DSIF as a whole (see Figure 3-1).

This revealed that slightly more DSIF occupied the bead fraction when polymerase was present, indicating a slight specificity of DSIF binding to fully formed TEC complexes. However, the large amount of binding of SPT5 to the bead fraction in the absence of polymerase suggests a lot of nonspecific binding, either to DNA or the beads directly (compare lanes 1 and 3 Figure 3-6A). It also revealed that the addition of CSB appeared to displace DSIF from the bead fraction as seen from a reduction of SPT5 on the beads following addition of CSB (compare lanes 1 with 2 and 3 with 4 Figure 3-6A) and an increase in SPT5 in the supernatant fraction (compare lanes 5 with 6 and 7 with 8 Figure 3-6A). However, this displacement was just as strong in the absence of polymerase as in its presence, indicating that the majority of DSIF was being displaced from DNA or the streptavidin beads and not necessarily just from stalled polymerase. This also raised the possibility that lots of free polymerase not forming TECs was binding non-specifically to the streptavidin beads.

3.6.2 Testing specificity of streptavidin Dynabeads

The assay described above had been developed in the lab for yeast proteins and was optimised using M280 streptavidin beads. Thermo Fisher Scientific have since developed many other streptavidin magnetic beads that differ in size and hydrophobicity. I tested these to see which beads best bound TEC complexes in the most specific manner. TECs were assembled as described previously or RNAPII was added alone to beads in transcription buffer. After washes, the beads were analysed by Western blot for the amount of RNAPII bound. Surprisingly, the M280 beads that were used previously, showed very weak binding of RNAPII (see Figure 3-6B lanes 1 and 5). While the binding of TECs to M280 beads was too limited to be detected by Western blot of RNAPII, it is possible that enough TECs bind to produce sufficient amounts of RNA that can be seen by much more sensitive radioautography, as had previously been the readout for this assay. While most of the beads showed non-specific binding of RNAPII, there was an increase in binding of RNAPII when it was incorporated into TECs before binding to C1 Dynabeads, indicating a higher specificity of these beads for biotin-tagged TECs over free RNAPII (compare lanes 4 with 8 Figure 3-6B). Therefore, going forward experiments were conducted with C1 streptavidin Dynabeads.

3.6.3 Testing of different length oligonucleotides

The apparent nonspecific binding of DSIF to the bead fraction could be mediated through direct binding to the beads themselves, or to excess upstream and downstream DNA not occupied by RNAPII. I therefore decided to shorten the DNA oligonucleotides to try and improve the specificity of DSIF binding to RNAPII. The original oligonucleotides were 150 bp, but footprinting studies have shown RNAPII occupies approximately 40 bp (Rice et al., 1993), so oligos were shortened to 54 bp to leave minimal DNA exposed (see Figure 3-5).

It has also been shown that the length of RNA protruding from polymerase is important for binding of DSIF (Crickard et al., 2016), so longer RNA oligonucleotide templates were tested to see if this increased specificity. However, there did not seem to be any effect of the longer RNA on binding of DSIF to the bead fraction as

there was just as much SPT5 bound in the absence of RNAPII as in its presence (see Figure 3-6C).

I next decided to see if I could improve specific binding of DSIF to RNAPII by adding it at different stages of the reaction. I either added it to already formed TECs at the same time as nucleotides to initiate transcription, or when incubating RNAPII with RNA (before beads were added). While there was significantly less SPT5 in the bead fraction when adding DSIF before fully formed TECs, this was likely because it then undergoes more extensive washing after beads are added, but it's still possible there is proportionally more specific binding to RNAPII. There is evidence of a slight increase in specificity when DSIF is incubated with RNAPII before fully forming TECs, as seen by the higher amount of SPT5 in the bead fraction in the presence of RNAPII compared to its absence, but this is a slight difference (compare lanes 5 with 6 Figure 3-6C).

3.6.4 Testing background binding of DSIF to streptavidin beads

The different streptavidin beads were tested for specific binding of fully formed TECs over RNAPII and I assumed this would reduce nonspecific binding of all proteins. However, I had not formally tested this. To assess the level of nonspecific binding of DSIF to streptavidin beads or DNA, TEC reactions were set up as usual, with or without RNAPII (lanes 1, 2, 3, and 4 Figure 3-6D). DSIF was also added to beads alone, in the absence of any oligonucleotides (lanes 5 and 6 Figure 3-6D). In addition, both my purification of DSIF and a protein from a preparation gifted from the Cramer lab were compared to test if the N-terminal truncation of SPT5 affected binding (Figure 3-6D).

This revealed that there was significant nonspecific binding of DSIF to the streptavidin beads in the absence of any RNAPII or oligonucleotides, indicating that almost all the SPT5 in the bead fractions I had been analysing was due to nonspecific binding to streptavidin beads. Both full-length and N-terminal truncated SPT5 were identical in their degree of binding.

The nonspecific interactions of proteins with components of the assay described here were never tested before because this would not necessarily interfere with the production of RNA from properly formed TECs, thus not introducing noise into the transcription (radioautography) readout. However, binding of DSIF to anything other than RNAPII in a TEC causes a significant noise issue when the readout is now the entire contents of the bead fraction. Due to these results, I decided to discontinue this project as this assay was not suitable to reliably discern protein interactions. Instead, the Dong Wang laboratory at UCSD have over the last year or more been attempting these experiments with their yeast protein setup. So far, no results have been produced.

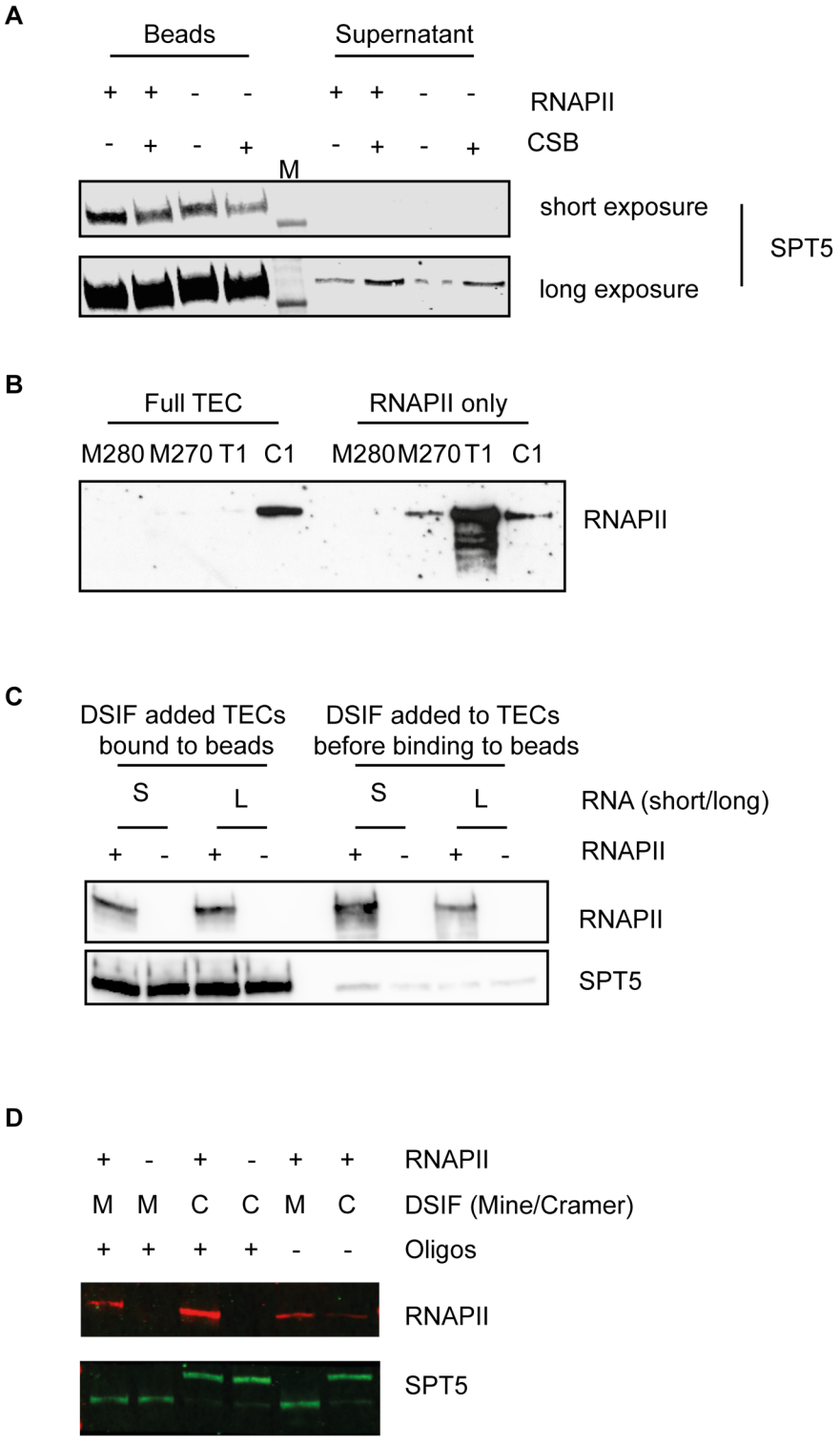


Figure 3-6 TEC assay optimisation shows that this assay is not suitable for investigating protein interactions. A Western blot of pilot TEC experiment showing that CSB mildly displaces SPT5 bound in the bead fraction. **B** Optimisation of TEC assay by testing of different streptavidin Dynabeads for specific incorporation of RNAPII via biotinylated oligos (Full TEC), and nonspecific binding by incubation of beads and RNAPII only. **C** Testing of different length RNA templates on SPT5 binding which shows no discernible difference. Testing specific binding of DSIF by adding complex at different points in TEC assembly. **D** DSIF non-specifically binds to beads. DSIF is present in bead fraction independent of RNAPII or any oligonucleotides.

Chapter 4. Results II – Establishing a cell system to study CSB ubiquitylation mutants

4.1 Introduction

It is known that the CUL4A E3 ubiquitin ligase uses several different cofactors, called DCAFs, to ubiquitylate many different target proteins during NER (J. Lee & Zhou, 2007). In the GG-NER pathway, UV-DDB recognises UV-lesions and is auto-ubiquitylated as one of its subunits, DDB1, is an integral component of the CUL4A E3 ligase (Matsuda et al., 2005). This results in its degradation and the recruitment of XPC, the main lesion recognition protein. XPC is ubiquitylated after lesion recognition by CUL4A, which seems to stabilise its association with DNA *in vitro* (Sugasawa et al., 2005). It has also been shown that another E3 ligase, RNF111, ubiquitylates XPC following UV irradiation, which targets it for degradation (Van Cuijk et al., 2015). In the TC-NER pathway, RNAPII, the de facto sensor of lesions in the transcribed strand, is known to become ubiquitylated and degraded following UV-irradiation (Anindya et al., 2007; Harreman et al., 2009; Tufegdžić Vidaković et al., 2020). Moreover, CSB contains a functionally important ubiquitin-binding domain (Anindya et al., 2010). Regulation by ubiquitin is clearly an important and abundant signalling mechanism in NER in general. The research in this area is more established for GG-NER, but apart from the existence of ubiquitylation of TC-NER factors, its mechanism is much less well understood.

CSA is another DCAF of CUL4A and an essential protein in TC-NER. Loss-of-function mutations in CSA, like in CSB, are responsible for causing Cockayne syndrome. CSB is one of the most important initiating factors in the TC-NER pathway and it has shown that CSA is essential for the ubiquitylation of CSB (Groisman et al., 2006), intimately linking the function of both proteins in TC-NER. But prior studies are lacking details on the nature of the functional interplay between CSA and CSB. Previous work in our lab has identified a ubiquitylated lysine residue that when mutated renders cells sensitive exclusively to oxidative damage, but not UV (Ranes et al., 2016). This demonstrated that CSB ubiquitylation of K991R is important for its role in BER of oxidative lesions, but not

TC-NER. The Groisman study clearly makes the case that CSA-dependent CSB ubiquitylation is important for TC-NER of UV lesions. It is likely then, like for XPC and RNAPII, there are specific ubiquitylation events that regulate CSB's mechanism of action in TC-NER.

In order to better understand the interplay between CSA and CSB, a former postdoc in the lab, Stefan Boeing, used a proteomics-based approach to map CSA-dependent ubiquitin sites by comparing CSB-ubiquitylation in WT cells and cells lacking CSA. In the following chapter I will present data that maps CSA-dependent ubiquitylation sites on CSB and their mutagenesis to abrogate CSB ubiquitylation. A cell system using the common model CS1AN patient cell line was used to attempt to create a system in which to study CSB ubiquitylation mutants. The CS1AN cell line is a fibroblast cell line derived from a patient suffering from Cockayne syndrome and immortalised with SV40 T antigen. This cell line has a normal karyotype (46,XX) and carries heterozygous mutations in CSB with one allele encoding a premature truncation (K377X) and the second allele encoding a frameshift mutation (R857X).

4.2 Identification of CSB ubiquitylated residues

To identify the lysine residues ubiquitylated on CSB, a mass spectrometry approach was employed. This work was carried out by Stefan Boeing, prior to my arrival. The ubiquitylation of lysines on proteins results in an isopeptide bond between the lysine and the C-terminal glycine of ubiquitin. Trypsin digestion, which cleaves on the C-terminal side of lysine or glycine residues, leaves a distinct K-G-G (diGly) remnant motif at the site of ubiquitylated lysines. This motif can be enriched by immunoprecipitation with diGly branch-specific antibodies (Cell Signalling Technology) to allow for mass spectrometry identification of ubiquitylated lysines. A stable isotope labelling with amino acids in cell culture (SILAC) approach (Boeing et al., 2016) was employed to compare HEK293 cells overexpressing CSB (as endogenous CSB is lowly abundant) with a CRISPR-derived CSA knockout in the same background. All cells were UV irradiated and incubated for 3 hours with MG132, a proteasome inhibitor, to preserve all ubiquitin. Therefore, the dataset

identified ubiquitylated residues that were CSA-dependent and present after UV irradiation, but not necessarily UV-dependent as both arms were subject to UV, precluding specific enrichment of UV-dependent residues.

Five residues on CSB were identified: K783, K1295, K1363, K1392, K1457, which were predominantly located in the C-terminal (see Figure 4-1). Mutation of lysine to the structurally conserved arginine residue is a common technique to inhibit ubiquitylation. Due to there being several lysine residues identified, and for it to be common for several mutations to be needed to ablate ubiquitylation (Danielsen et al., 2011), I decided to also mutate any lysine residues within 10 residues adjacent to those identified. Although one of the lysines identified was within the ATPase domain region and therefore might potentially affect catalytic activity, closer inspection showed that this lysine was not situated within either of the two conserved lobes and was therefore considered unlikely to contribute to catalytic activity; it was therefore mutated as well. Two constructs were made, where one contained the identified 5 lysines mutated to arginine (5R) and one with 3 additional mutations (8R).

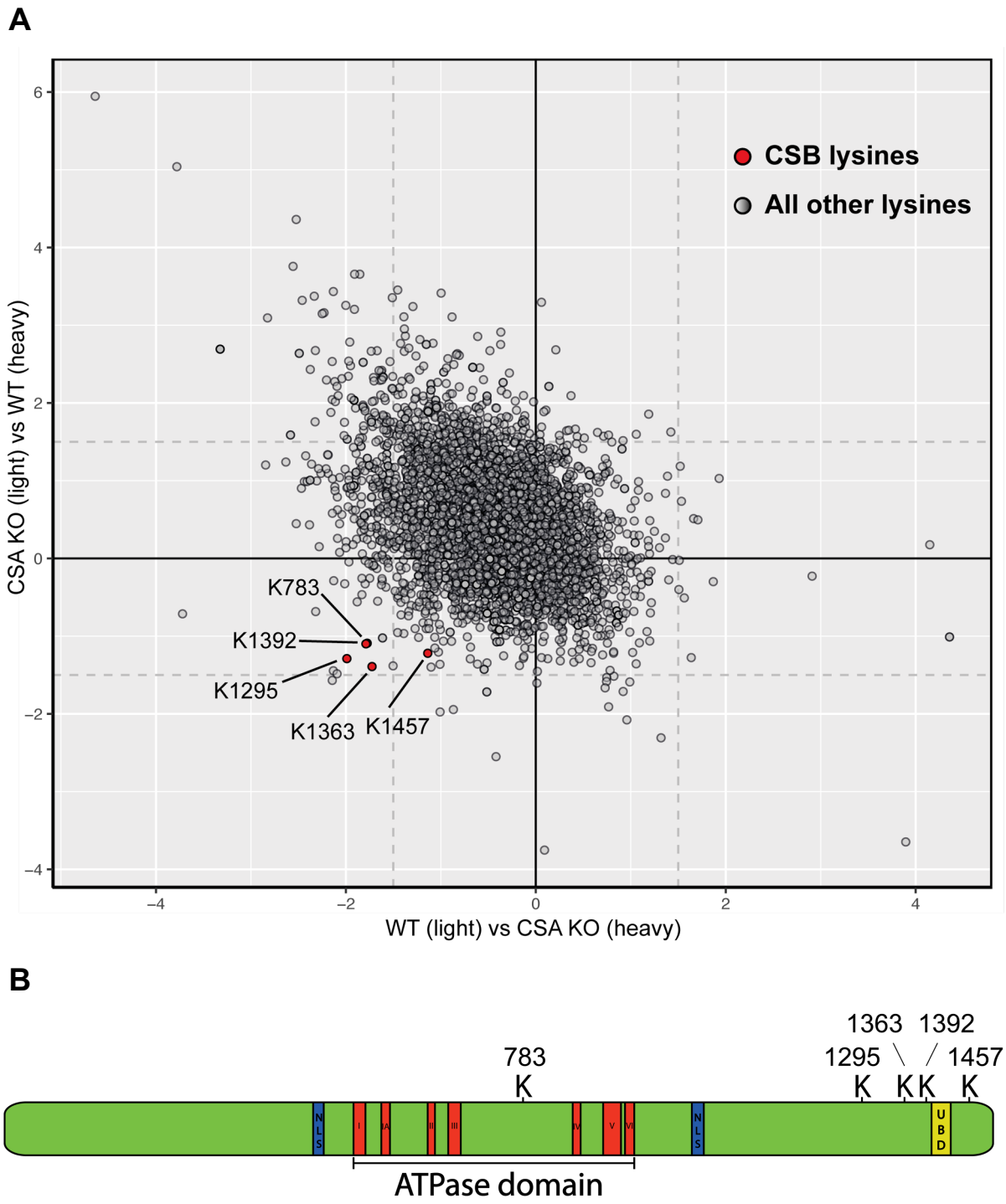


Figure 4-1 diGly mass spectrometry reveals CSA-dependent ubiquitylated residues on CSB. A SILAC mass spectrometry approach was employed to test diGly immunoprecipitated lysates after UV irradiation comparing CSA KO and parental cells. Forward and reverse arms of the experiment are plotted on opposing axis using iBAQ values.
Experiment performed by Stefan Boeing.

4.3 CSB mutant transgenes and cell system

The CS1AN cell line is a Cockayne syndrome patient-derived cell line harbouring heterozygous truncations of CSB (p.Lys337Ter and p.Arg857Ter), which results in a loss of CSB function (A. R. Lehmann, 1982; Mayne & Lehmann, 1982). This cell line has been used extensively in the literature and previously in our lab too, so naturally it was the cell line of choice to create a system to study the CSB ubiquitylation mutants. The CSB coding sequence, N-terminally tagged with GFP and FLAG, was cloned into pTRE3G plasmid containing a doxycycline-inducible promoter. The two ubiquitylation mutants (5R and 8R mentioned above) were made by mutagenizing (2.2.3) the CSB WT plasmid and confirmed by sequencing. Cell lines expressing CSB were made by lipid-based transfection of plasmids followed by selection with puromycin (see 2.5.4), yielding colonies deriving from an individual cell, which were expanded and tested for expression by Western blot (see

Figure 4-2B). CS1AN cells expressing WT CSB and CSB lacking the UBD, CSBΔUBD (deletion of the most C-terminal 273 amino acids) were already available in the lab (Anindya et al., 2010).

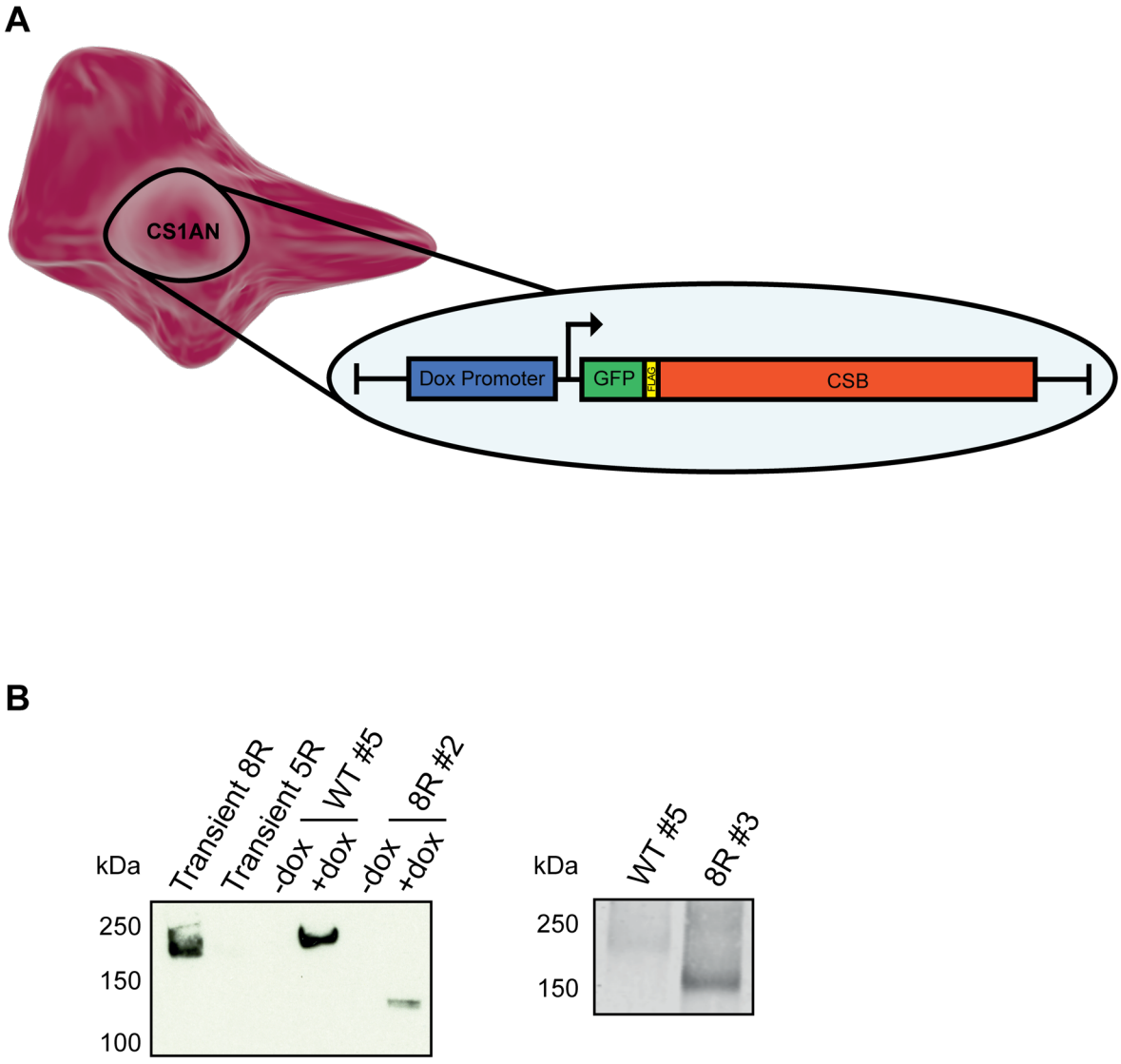


Figure 4-2 CS1AN cell line system shows integration of truncated transgenes.
A Schematic of transgene transfected into cells. **B** Western blots showing different sizes of proteins produced from random integration of GFP/FLAG-tagged WT or 8R CSB transgene in CS1AN cells blotting with GFP antibody. Insertion of a truncated CSB construct is shown and discussed further in text.

4.4 CS1AN + CSB WT fail to rescue transcription shutdown

It is well established that cells respond to some genotoxic insults, such as UV-irradiation, by an almost complete shutdown of global transcription that then recovers over 8–24 hours (Mayne & Lehmann, 1982). In cells deficient in TC-NER, there is a failure to recover RNA synthesis (RRS) and this is a classical assay to test primary patient cells in the diagnosis of Cockayne syndrome.

If the ubiquitylation of CSB is important for its function in TC-NER then a failure of the CSB ubiquitylation mutants to recover RNA synthesis would be proof of this. CSB 8R cells were initially the only ubiquitylation mutant tested, as 5R could later be tested if a phenotype was found in the more expansive mutant.

CS1AN cells expressing CSB WT, different CSB mutants, or no transgene were cultured in doxycycline for 24 hours to induce expression of the transgene. Cells were seeded in triplicate, then either left untreated or UV irradiated at 20 J/m^2. After 2 and 24 hours, cells were labelled with 5-ethynyl uridine (5EU) for 1 hour, to measure nascent transcription by incorporation into newly transcribed RNA. Cells were then fixed, permeabilised, stained with DAPI, and Alexa Fluor 647 Azide was conjugated to 5EU-labelled RNA by click chemistry. This allowed for the visualisation of the levels of nascent RNA transcripts in the cell by fluorescence microscopy. The average intensity of 5EU was measured for each nucleus and averaged over the triplicates for each condition (detailed protocol in section 2.10).

This revealed a shutdown of transcription 2 hours after UV in all cells (see Figure 4-3A), with expression recovering to different extents at 24 hours in the different cell lines (see Figure 4-3B). As expected, CS1AN and CSB∆UBD cells failed to recover RNA synthesis at 24 hours, whereas MRC5 cells (a similar fibroblast cell line expressing endogenous CSB) fully recovered RNA synthesis. CSB 8R cells failed to recover RNA synthesis at 24 hours, indicating a possible defect of this mutant in TC-NER. Surprisingly however, CSB WT cells recovered RNA synthesis only to moderate levels at 24 hours, but not to the levels of transcription in untreated cells, as I had expected. In conclusion, I was not able to reliably infer if

the phenotype of CSB 8R was due to the inability of CSB to be ubiquitylated as the vital control cell line, CSB WT, did not produce a positive result.

Discussions with Michael Ranes, who developed the CS1AN cell line expressing WT CSB, revealed that he had noticed that these cell lines, while initially displaying WT phenotype, started to revert to a CS1AN phenotype after being cultured for some time. He believed the outgrowth of CS1AN cells lacking expression of WT CSB may have been responsible for this, and he therefore often re-selected cells with puromycin. This would explain the semi-rescued phenotype I observed in this cell line. When establishing stable cell lines in CS1AN, I often observed expression of truncated CSB protein after transfection of CSB 8R (see Figure 4-2B). Owing to the large size of CSB, protein products that appear to be the same size as full-length protein by western blotting might contain small truncations that would not necessarily resolve differently from full-length CSB at the top of the gel. This could mislead one into believing the full CSB coding sequence had been integrated into the genome when small truncations may be present. For these reasons, I deemed this cell system to not be robust enough to draw reliable conclusions about CSB phenotypes and decided to take another approach.

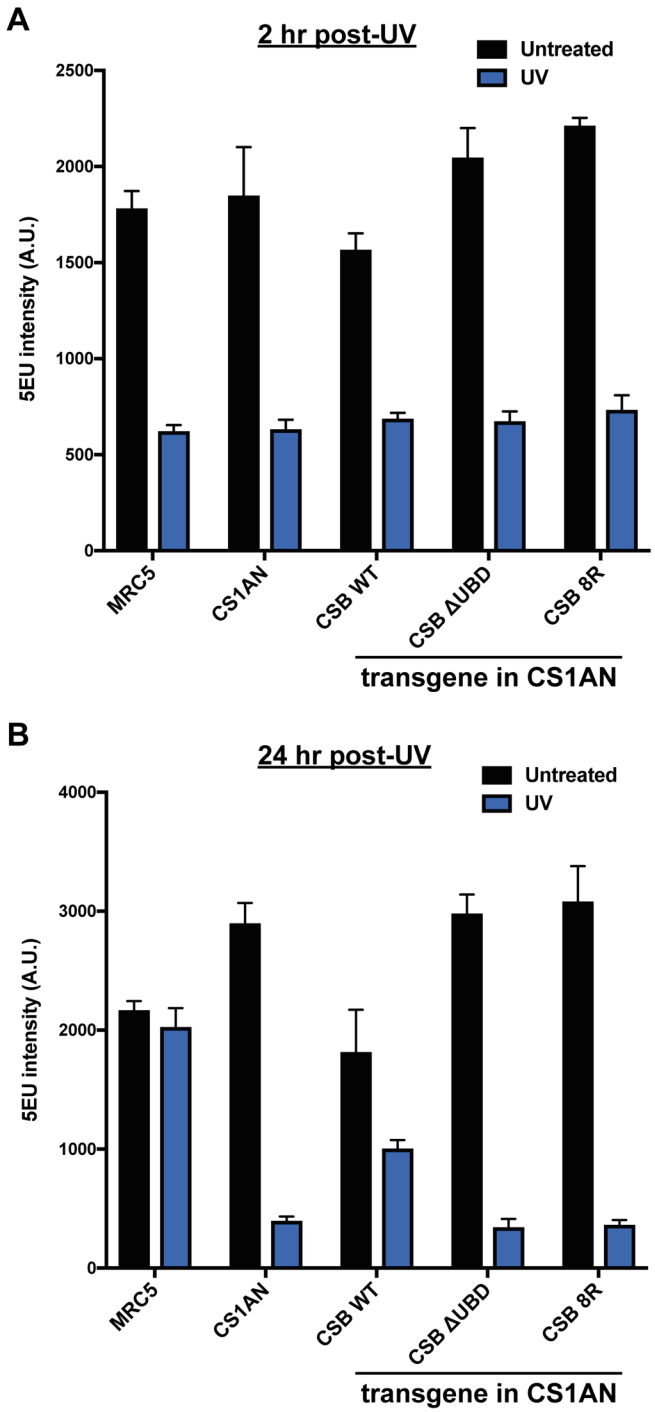


Figure 4-3 CS1AN cells expressing CSB WT fail to recover RNA synthesis following UV. Alexa Fluor 647 fluorescence intensity was measured for every nucleus of all wells of a sample and averaged to give a fluorescence per cell value. Untreated data are shown in black alongside post-UV data in blue. Error bars represent standard deviation from triplicates. **A** 2 hours after UV irradiation. **B** 24-hours after UV irradiation.

4.5 CRISPR-Cas9 CSB knockout

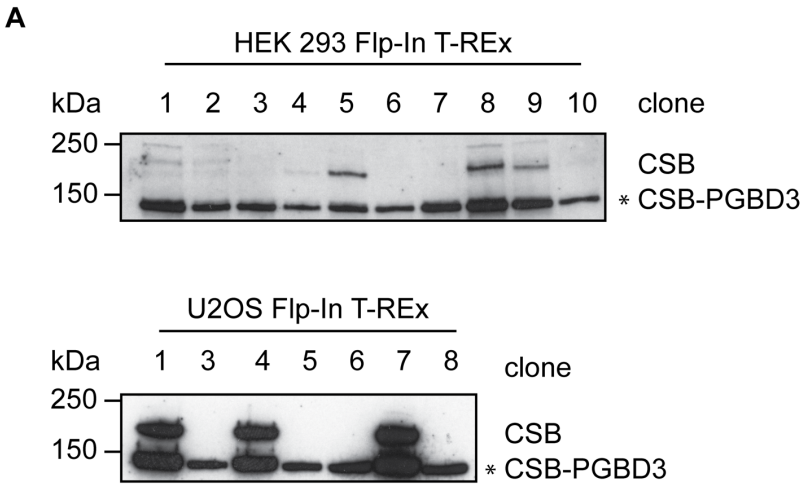
Historically, patient cell lines have been used as a classic system in order to study a proteins function in the background of mutations known to inactivate the endogenous gene/protein *in vivo*. However, with the recent advances in genomic engineering methods, it has become possible to knockout the endogenous gene being studied and express transgenes. I decided to use versions of the well-established cell lines, HEK293 and U2OS, that have been used routinely in the laboratory and have been engineered to contain a single Flp-FRT recombination site (see

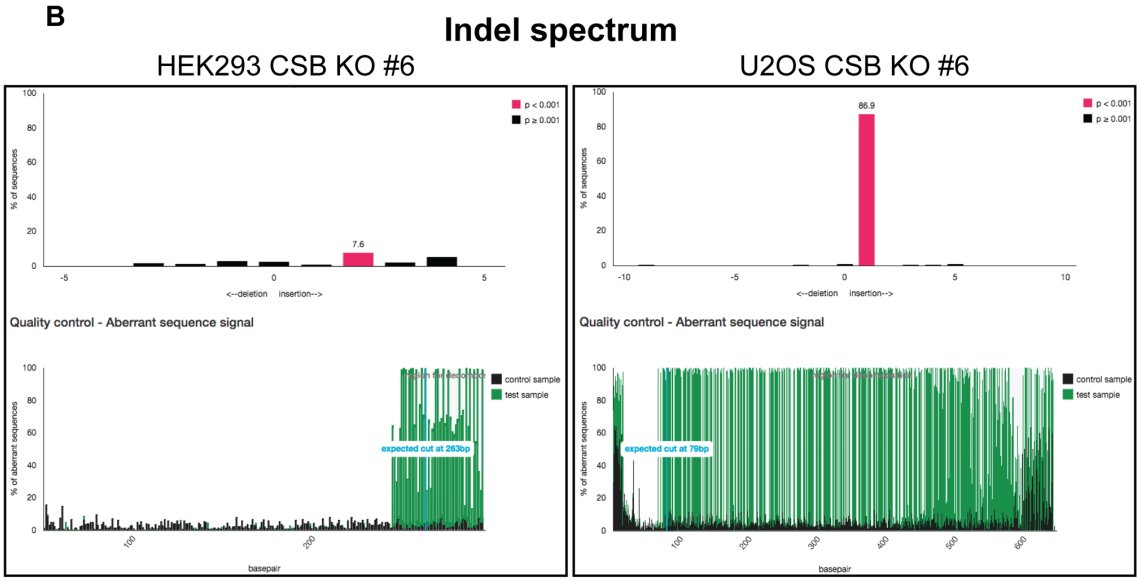
Figure 4-5A). Using these cell lines as a starting point for a genetic knock out and insertion of the full-length 'rescuing' transgene by recombination of FRT sites 5' and 3' of the coding sequence would overcome the problems observed in the CS1AN cells, such as truncated transgene insertion and expression (O'Gorman et al., 1991). HEK293 and U2OS cells were used as they were the cell lines available with FRT and T-REx sites integrated into the genome and thus ameanable to this approach. Also, they have properties that make them suitable for different experiments. HEK293 cells are small and fast growing that makes them suitable for growing up large numbers of cells quickly for experiments that might require large numbers of cells (Dsk2 pulldowns, IP's etc.). There are also a large number of genomic datasets in the laboratory performed in HEK293 derivative cell lines that would be a rich resource with which to compare genomic data I gather from any derivatives I make of these cell lines. However, the poor adherace to tissue culture plates of HEK293 cells make them difficult to use for assays that require media changes, multiple washes or other manipulations (e.g. 5-EU and clonogenic assays). For this reason, the U2OS cell line, which is very adherent was also used to make CSB knockouts. HEK293 cells are an embryonic epithelial kidney-derived hypotriploid cell line with a modal chromosome number of 64, occurring in 30% of cells and 4.2% of cells having higher ploidy. U2OS cells are an osteosarcoma-derived cell line with highly altered chromosomes, displaying mostly hypertriploid counts and several chromosomal rearranegments.

The CRIPSR-Cas9 system (Jinek et al., 2012) was employed for inducing mutations in the endogenous *ERCC6* (*CSB*) gene. Two approaches were taken using the WT Cas9 protein and an engineered 'nickase' version (Ran, Hsu, Lin, et al., 2013). The latter aims to reduce off-target effects by using a mutated version of Cas9 that only makes an incision in one DNA strand rather than a double strand break (DSB). A pair of sgRNAs complementary to opposite strands but offset of the target site increase fidelity as two off-target incision events would have to occur in close proximity to induce a DSB and the error-prone NHEJ repair pathway. Individual nicks would predominantly be repaired by the high-fidelity base excision repair pathway (BER) (Beard et al., 2019).

Several candidate gRNAs were first tested in MRC5 cells for cutting efficiency using TIDE analysis on the pool of transfected cells (Brinkman et al., 2014). The best gRNA of the WT and 'nickase' Cas9 was determined by the highest rates of out-of-frame indels and taken forward for use in HEK293 and U2OS cells. A plasmid containing GFP-tagged Cas9 and gRNA was transfected to allow for flow cytometry sorting of Cas9-positive single cells. Single cells were expanded and screened by Western blot for the presence of CSB (see Figure 4-4A). Clones that lacked detectable CSB were further screened by genomic DNA sequencing and analysed using TIDE to confirm homozygote out-of-frame indels (see Figure 4-4B). TIDE quantifies indels produced by Cas9 by analysing the Sanger sequencing chromatogram over the proposed cutsite. The HEK293 CSB KO cell line shows several indels. This could be due to different indels at each of the several alleles in the polyploid cell line as well as several cuts recurring after repair of the initial cut. It's unlikely, but plausible that Cas9 is diluted but endures in daughter cells after mitosis and continues editing in progeny giving rise to heterogeneously edited cell populations. It is also likely that single-cell sorting was not accurate and a several cells were sorted into a single well and the TIDE analysis represents a population of edited cells. In HEK293 cells, the nickase Cas9 using two guide RNA's produced more viable colonies with CSB knocked out than Cas9 WT with a single guide RNA. In U2OS cells, the WT Cas9 with single guide RNA produced more viable cells with CSB knocked out. Hence, cells that are used in all experiments going forward were produced with the nickase Cas9 for HEK293 cells and WT Cas9 for U2OS cells. CSB knockout clones were tested to confirm that they exhibited the

characteristic Cockayne syndrome phenotype of UV sensitivity by colony formation assay (see Figure 4-4C). Indeed, in contrast to parental cells which survived UV-irradiation, CSB KO cells failed to recover from such treatment.





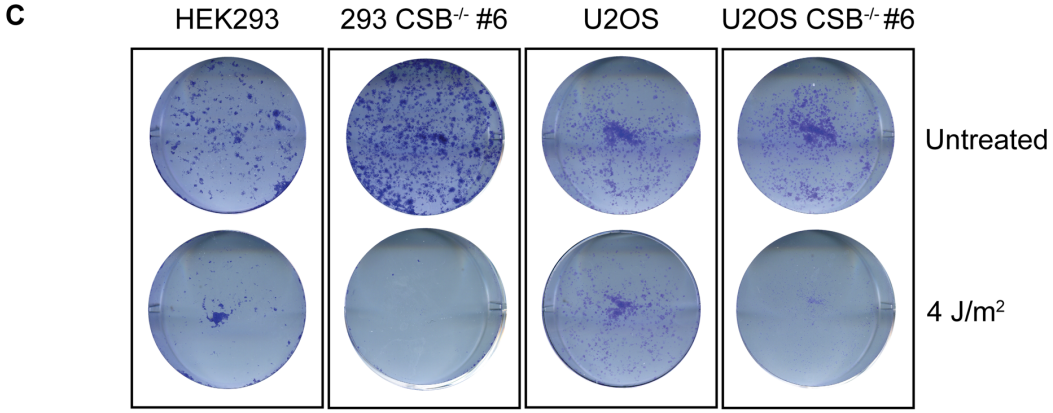


Figure 4-4 CRISPR knockout of CSB in HEK293 and U2OS Flp-In T-REx cell lines. A Western blot analysis showing complete lack of endogenous CSB protein in some clones. PGBD3, CSB piggyback protein (Newman et al., 2008). **B** Representative TIDE analysis of indels from two clones. **C** Clonogenics assay and crystal violet staining shows CSB KO cells are much more UV-sensitive than parental cells, as expected.

4.6 Flp-In cell line generation

The use of the Flp-In system requires unique plasmids containing recombination FRT sites. Hence, CSB WT and mutant coding sequences were cloned into pDEST-FRT/TO vectors, using the Gateway cloning system. HEK293 and U2OS cells were each transfected with a plasmid encoding the Flp recombinase (pOG44) and a pDEST-FRT/TO vector encoding either CSB WT, 5R, 8R, or CSBΔUBD (

Figure 4-5A) to allow for a single site integration into the FRT site. Cells were selected for the transgene with hygromycin, and positive clones were picked and expanded. Cells were tested for expression of CSB with and without doxycycline, which showed dose and time-dependent expression in HEK293 cells that was equivalent to endogenous levels of CSB (

Figure 4-5B). However, expression in U2OS cells was uneven between different mutant cell lines and repeated experiments showed sporadic expression that couldn't be titrated with different doses of doxycycline (

Figure 4-5C). Analysis of DNA damage markers in the U2OS cell lines showed inconsistent and irreproducible results (data not shown), which is not surprising as this is an osteosarcoma cell line and inactivation of DNA damage resonses and repair pathways is a common hallmark of tumours. I deemed it innapropriate to be studying a DNA damage repair pathway in a cell line that likely has mutated and altered responses to DNA damage that might not reflect physiological responses of most somatic tissues.

As the Flp-In FRT recombination site was introduced into the parental U2OS cells by another Crick laboratory rather than a company provider, it is possible that

several FRT loci were introduced. Recombination of the transgene into these cells could this lead to different copy numbers of transgenes in different cell lines leading to aberrant expression. Indeed, HEK293 Flp-In cells were developed by Thermo Fisher Scientific and Southern blot analysis was performed to confirm only a single integration of an FRT site. Hence, all derivative HEK293 cell lines carrying mutant CSB transgenes would integrate at a single locus in all cell lines, giving a higher likelihood of equal expression. For these reasons I decided not to use the generated U2OS cells going forward. The U2OS cell line was chosen as it is very adherent and amenable to washes, media changes and other manipulations without being susceptible to detachment, which is required for many of the assays I had planned to carry out (clonogenics, 5-EU RRS etc.). Contrarily, HEK293 cells, while considered adherent cell lines, are very susceptible to detachment with washes and media changes. Therefore, I would need to develop alternative assays to measure cell survival and RRS in response to UV irradiation in the HEK293 CSB cells.

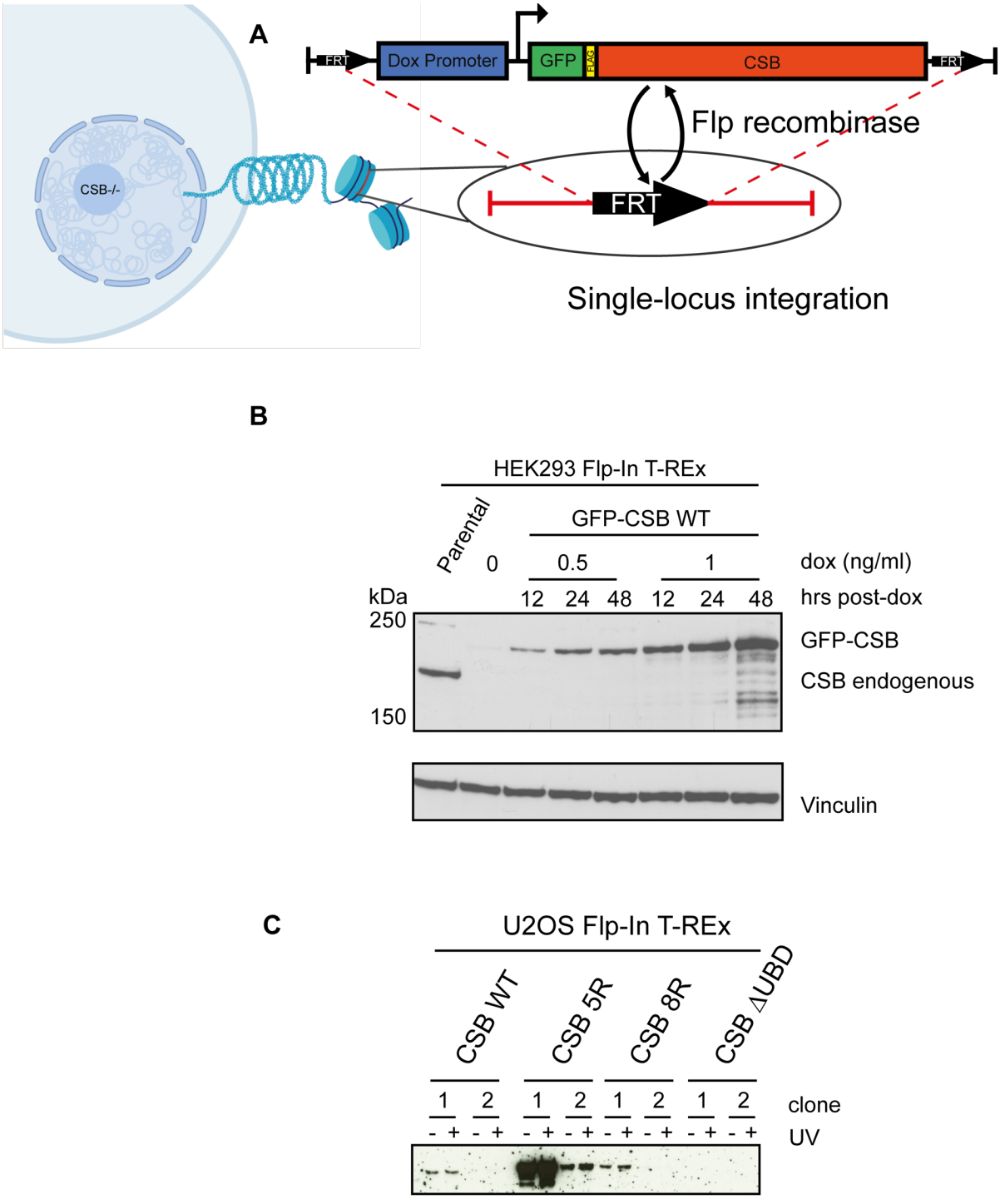


Figure 4-5 Generation of CSB expressing cell lines using Flp-In system in CSB KO cells
A Schematic of CSB transgene and its recombination into the genome via a single loci FRT site. **B** Western blot showing dose-dependent doxycycline expression of exogenous CSB to endogenous levels in HEK293 cells. Blotted with anti-CSB antibody. **C** Western blot showing very variable expression of exogenous CSB in several clones and genotypes at a single dose of doxycycline. Blotted with anti-CSB antibody.

Chapter 5. Results III – Functional characterisation of CSB ubiquitin mutants

5.1 Introduction

As mentioned previously, it is well established that regulation by ubiquitin plays an important role in NER, specifically among the initiating factors of the NER pathways, which is well characterised for XPC and UV-DDB of GG-NER. In TC-NER, it has been shown that CSA was necessary for CSB's ubiquitylation (Groisman et al., 2006). Previous work in the lab, had identified a ubiquitin binding domain (UBD) in the C-terminal region of CSB (Anindya et al., 2010), as essential for TC-NER. However, beyond these studies nothing is known about the mechanistic details of CSB's ubiquitylation and its role in TC-NER. In the previous chapter I presented results of a proteomic screen that identified the CSA-dependent ubiquitylation marks on CSB after UV.

The substrate of CSB's UBD remains elusive. Interestingly, regulatory domains in the C-terminal of chromatin remodelers that define their substrate are commonplace. For example, bromodomains exist in the C-terminal of the SWI/SNF family of chromatin remodelers, which target them to acetylated histone H3 (Clapier et al., 2017). The Swi2/Snf2 (to which CSB is most closely related) complex can be acetylated and such self-acetylation competes with histones for the C-terminal bromodomain (Dutta et al., 2014; J. H. Kim et al., 2010). Thus, intramolecular interactions of C-terminal domains and PTM's are an established mode of regulating activity in this family of chromatin remodelers.

This, thus, presents several interesting hypotheses of how CSB ubiquitylation may govern its mechanism. The UBD of CSB may target it to other ubiquitylated proteins, such as RNAPII. There is also evidence that CSB functions as a homodimer (Christiansen et al., 2005) and regulation of CSB activity may thus be facilitated via interactions of the UBD of one CSB molecule and ubiquitin of the other subunit of the dimer. Alternatively, binding of the C-terminal UBD intra-molecularly to ubiquitin on the CSB surface elsewhere may result in a dramatic

change of conformation. In an analogous way to Swi2/Snf2, ubiquitylation of CSB might thus act to recruit its own UBD, regulating its catalytic activity. In any case, whatever the mechanism, evidence suggests CSA acts in response to UV to impart ubiquitin marks on CSB important for the conclusion of the TC-NER reaction.

In the following section, I present data that aims to characterise the phenotypic effect of several CSB ubiquitylation mutants on the cellular response to UV-induced DNA damage. I present data that identifies further ubiquitylation sites on CSB induced by treatment of cells with cisplatin, a DNA-damaging drug which also induces RNAPII-stalling DNA damage.

5.2 CSB 8R has normal turnover kinetics

Having established HEK293 knockout cell lines expressing CSB mutant isoforms, I first wanted to test if I had disturbed the half-life of CSB under normal conditions, in the absence of exogenous DNA damage, by the mutation of lysine to arginine. Any significant change in the half-life of CSB 8R in the absence of DNA damage would thus make it difficult to interpret any phenotype occurring after UV-induced DNA damage. Indeed, if changes protein stability occurred, this could result in insufficient availability of CSB protein to recognise stalled RNAPII rather than inhibition of ubiquitin signalling, mechanistically important for CSB to resolve stalled RNAPII.

Cells were grown with doxycycline for 24 hours to express CSB WT or CSB 8R to equal levels. Dox was then removed to stop expression of the transgene, and cells were harvested 24, 48 and 72 hours later to assay for the level of CSB protein by Western blot analysis (Figure 5-1). Levels of both CSB WT and 8R were even at the time of dox removal (0 hrs) and took about 48 hours to be completely removed, giving a half-life of approximately 24 hours. The proteins had very similar turnover kinetics. Thus, it was deemed that the 8R mutations have no effect on intrinsic protein stability and any other phenotype would not be due to, for example, degradative ubiquitylation marks present under normal conditions. Going forward, all phenotypic assays were assessing differences between CSB WT and 8R

following UV-induced DNA damage. How the measured half-life of 24 hours compares to the half-life of endogenous CSB is not known. I am not aware of any studies measuring the half-life of CSB by pulse-chase or translation inhibition experiments.

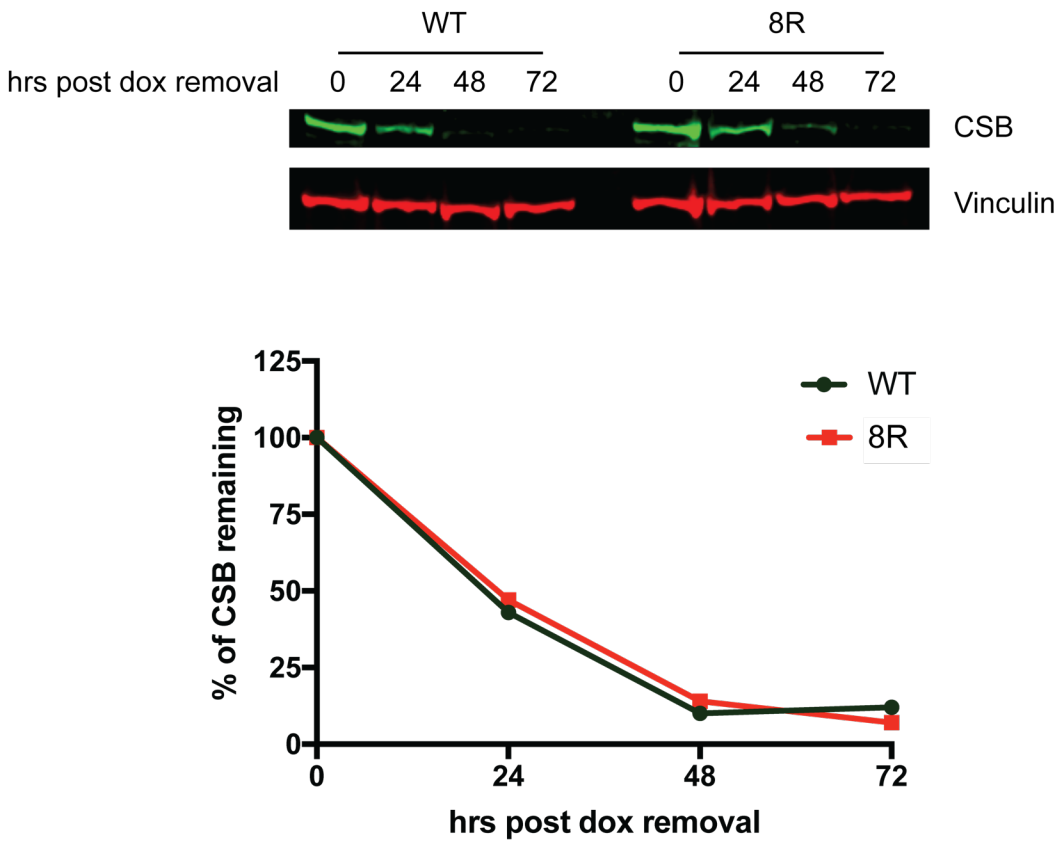


Figure 5-1 CSB 8R stability and turnover is normal
Western blot of CSB expression over time after doxycycline removal using CSB primary antibody and fluorescent secondary antibodies (Alexa Fluor 488). Quantification is normalised to expression at time of doxycycline wash out (0 hours) for each CSB genotype expressed as a transgene in HEK293 CSB KO cells. This experiment was repeated with a similar result.

5.3 CSB 8R restarts transcription following UV

Firstly, I decided to test if there was any effect of CSB K>R mutations on its role in resolving RNAPII-stalling DNA damage. One standard assay to address this question is to measure the recovery of RNA synthesis (RRS) following DNA damage. This had historically been done by labelling nascently transcribed RNA with nucleotide analogues, which can be visualised and measured as a proxy for the level of transcription occurring in the cells over the course of the labelling. This can be done by metabolic labelling of RNA with radioactive NTP's and autoradiography, or recently with 5EU labelling and subsequent click-chemistry with fluorophores (as in section 4.4).

These methods require extensive washing and manipulation of cells that is not suitable for easily detachable HEK293 cells. To remedy this, I used an adapted qPCR method using primers that would only amplify nascent RNA. Although qPCR can only amplify single loci and cannot measure global nascent RNA, amplifying at the 3′ end of long genes serves as a proxy for the full completion of repair in other genes of similar and shorter length (Tufegdžić Vidaković et al., 2020).To this end, two genes were chosen for this assay: EXT1 (one of the longest genes in the genome; 312 kb) and PUM1 (134 kb). These were chosen because they both represent long, highly expressed genes from HEK293 TT-seq datasets from the laboratory. Furthermore, UV irradiation of HEK293 cells have validated that both genes display a dramatic reduction of nascent transcription towards the 3' end of the gene body as measured by TT-seq (Tufegdžić Vidaković et al., 2020). Primers were designed to anneal to an intron exon junction towards the 3′ end of the gene, such that they were specific to nascent transcripts and wouldn't amplify mRNA (red arrows in Figure 5-2).

Cells were grown in doxycycline for 24 hours and either mock-treated or UV irradiated at 20 J/m^2 and then harvested at 0, 2, 24 hours post-UV by direct lysis in the dish. Total RNA was extracted using silica spin columns with DNase treatment to prevent amplification of genomic DNA. RNA was then quantified and normalised before reverse transcription with random hexamers. qPCR using BioRad SYBR

green technology was performed on cDNA in triplicate with primers aligning to GAPDH mRNA as well as the aforementioned nascent transcripts. Data shown is normalised to GAPDH mRNA levels for each timepoint and cell line individually and then every timepoint is normalised to the untreated condition for each cell line. As it is expected that nascent RNA levels will change between cell lines, timepoints and treatments, normalisation to a stable, long-lived, abundant RNA species that is unaffected by UV is needed; for this purpose, GAPDH mRNA was used.

A dramatic reduction in transcription in the 3′ end of EXT1 and PUM1 can be seen at 2 hours post-UV for all cell lines, as expected after inflicting transcription-blocking DNA damage and confirming the suitability of the assay to detect changes in transcription (see blue bars in Figure 5-2). Cell lines expressing CSB WT recovered transcription to near normal levels after 24 hours in both genes, while cell lines lacking CSB failed to recover RNA synthesis (see red bars), as expected. Unfortunately, cells expressing CSB 8R recovered RNA synthesis 24 hours post-UV to levels similar to WT. This indicates a normal phenotype of CSB 8R in response to transcription-blocking DNA damage.

Given this disappointing result, I decided to assay directly for the ubiquitylation status of CSB 8R in response to UV-induced DNA damage before performing further phenotypic experiments.

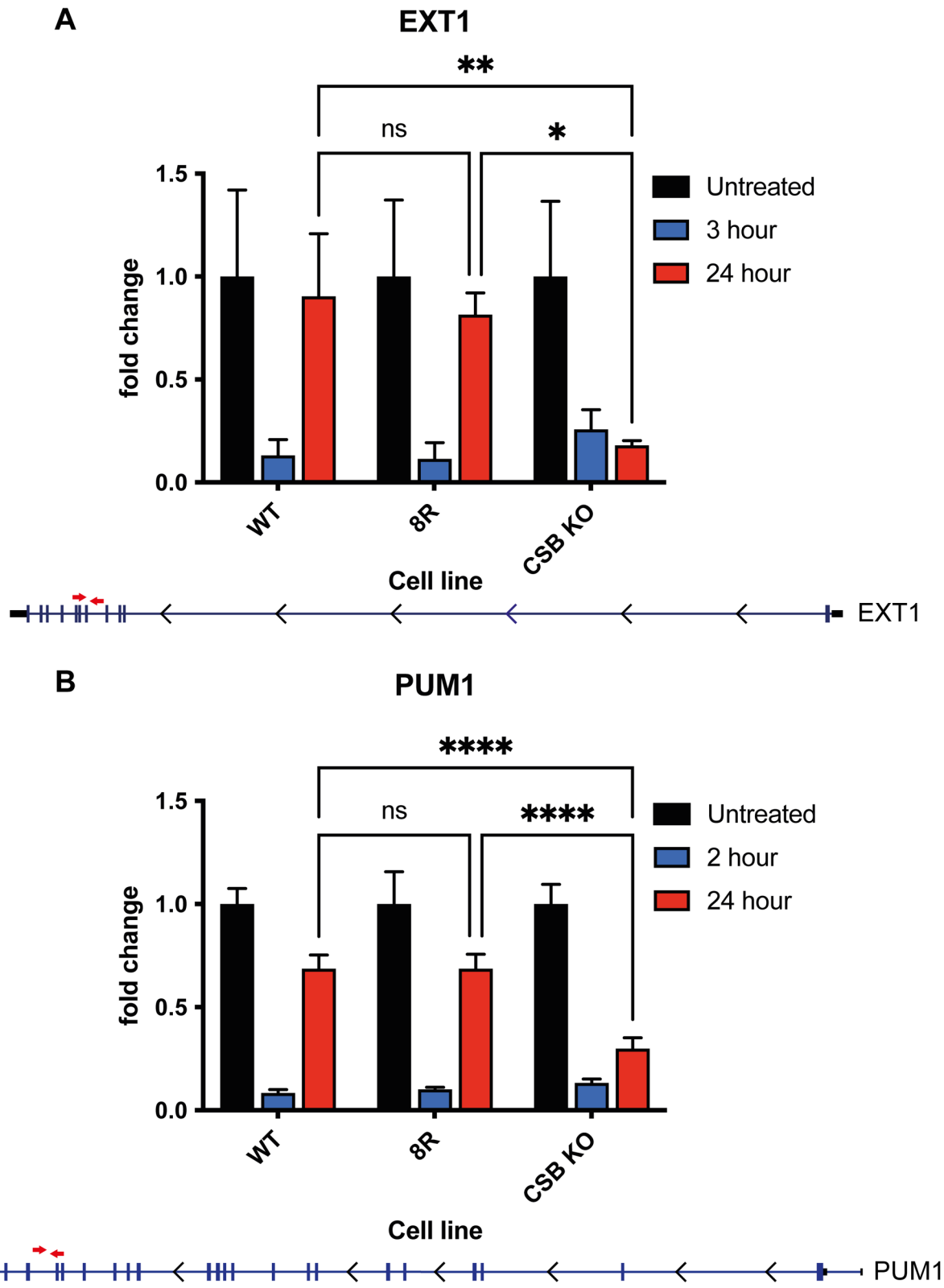


Figure 5-2 qPCR of intron-exon junction reveals normal RRS in CSB 8R.
qPCR experiments were performed with technical triplicates and data are represented as the mean of all replicates normalised to GAPDH mRNA and then the respective untreated condition for each genotype. Error bars represent standard deviation. For A statistics were done on means of technical triplicates from 3 experiments (3 biological replicates). Two-way ANOVA

multiple comparisons were done for 24-hour timepoints. Adjusted P values are represented by asterixs, where ns = P> 0.05; * = P <0.05; ** = P < 0.01; **** < 0.0001. **A** qPCR of EXT1 gene, approximately 300 kb from TSS. Data from three biological replicates of technical triplicates are shown. **B** qPCR of PUM1 gene, approximately 130 kb from TSS. Data are from technical triplicates.

5.4 CSB 8R is still ubiquitylated

In order to test whether the conservative K>R mutations of the lysine residues had indeed affected ubiquitylation, an enrichment for ubiquitylated proteins was performed, followed by CSB Western blotting. MutliDsk is a ubiquitin affinity resin that was developed in the laboratory (M. D. Wilson et al., 2012). It is a polymer of five ubiquitin-binding UBA domains from the yeast protein Dsk2, fused to GST. This allows for high affinity capture of ubiquitylated proteins from cell extracts.

HEK293 cells expressing CSB WT or CSB 8R were compared to parental HEK293 cells as a control to ensure similar levels of ubiquitylation were achieved on GFP and FLAG tagged CSB constructs. Cells were grown in doxycycline for 24 hours to achieve endogenous levels of transgene expression. Cells were either UV irradiated at 30 J/m^2 or mock treated. After 30 minutes incubation, cells were harvested and lysed. Cell lysates were incubated with equal amounts of MultiDsk beads and rotated in a cold room for several hours. The supernatant was removed and after several washes, the beads were boiled in Laemmli buffer to dissociate all bound proteins.

Western blot analysis of the input (pre-binding) and supernatant lysate (post-binding) with an antibody raised against ubiquitin, show a dramatic decrease in ubiquitylated proteins of all sizes in the lysate confirming that the beads effectively captured (depleted) the ubiquitylated species (see Figure 5-3A, compare lanes labelled ‘I’ and ‘S’). Treatment of lysates with the purified, recombinant catalytic domain of USP2, a potent deubiquitinating enzyme (DUB), prior to MultiDsk enrichment, confirmed the specificity of this antibody for ubiquitylated species and the efficiency of pull-out on the MultiDsk beads (see Figure 5-3A).

Western blot analysis of the bound material, with a CSB antibody, showed that CSB is captured in untreated and UV conditions. The appearance of a smear of higher molecular weight species after UV irradiation reveals a large increase in ubiquitylation of CSB in response to UV irradiation. While this confirmed that the exogenous CSB WT is indeed ubiquitylated to similar levels as endogenous CSB (see Figure 5-3B, compare lanes 1 with 5 and 2 with 7), there was, disappointingly, also WT levels of ubiquitylation of CSB 8R (see Figure 5-3B, compare lanes 3 with 8).

Note that in untreated conditions, although CSB is pulled down with MultiDsk beads, there is no higher molecular weight smear, indicating very little, if any, multi- or poly-ubiquitylation. There is a faint 'ladder' of larger molecular weight species above the 'main' CSB band, but the majority of species are composed of a single band with a distinct lack of high-molecular weight 'smear'. It is possible that the main band could represent a monoubiquitylated form of CSB, hence the absence of a smear. However, running untreated input lysates (lanes marked * in Figure 5-3B) alongside UV-treated MultiDsk-captured lysates shows no size difference between the main CSB bands, which would likely be seen if it was monoubiquitylated CSB. On the other hand, a shift of 8 kDa in the 200 kDa portion of the gel may not separate enough to give the resolution needed to distinguish between non- and mono-ubiquitylated CSB. USP2 treated lysates also show some CSB pulled down with MultiDsk, which would presumably be non-ubiquitylated and this migrates at the same size as CSB main band. Although binding is much reduced after USP2 treatment, it indicates some non-specific interaction of CSB with the beads remains. CSB may also bind to the beads via physiologically relevant interactions with other ubiquitylated proteins like RNAPII. CSB is recruited to RNAPII after UV and even though RNAPII's ubiquitylation is not responsible for this interaction (Tufegdžić Vidaković et al., 2020) it could still facilitate non-ubiquitylated CSB being enriched with Dsk2 beads.

In conclusion, it seems that even though some ubiquitylated CSB species exist in untreated conditions, the majority of CSB is not ubiquitylated in the absence of exogenous DNA damage. Whether these ubiquitin marks are important for CSB function or simply part of normal protein turnover is not known. There is a

significant increase in CSB ubiquitylation following UV irradiation, indicating a probable functional role for CSB ubiquitylation in removing UV-induced DNA damage. Surprisingly, the mutation of 8 lysine residues, 5 of which were previously identified as ubiquitylated *in vivo*, did not have any discernible difference on CSB ubiquitylation levels. This could be because these residues are not important for CSB ubiquitylation following UV irradiation. Alternatively, it is possible that other lysine residues in close structural proximity (but not necessarily close in primary sequence) could compensate for the 8R mutations and still permit modification, as often occurs with ubiquitylated proteins. Finally, there may also be several lysine residues not identified in the proteomics screen that can be still be ubiquitylated.

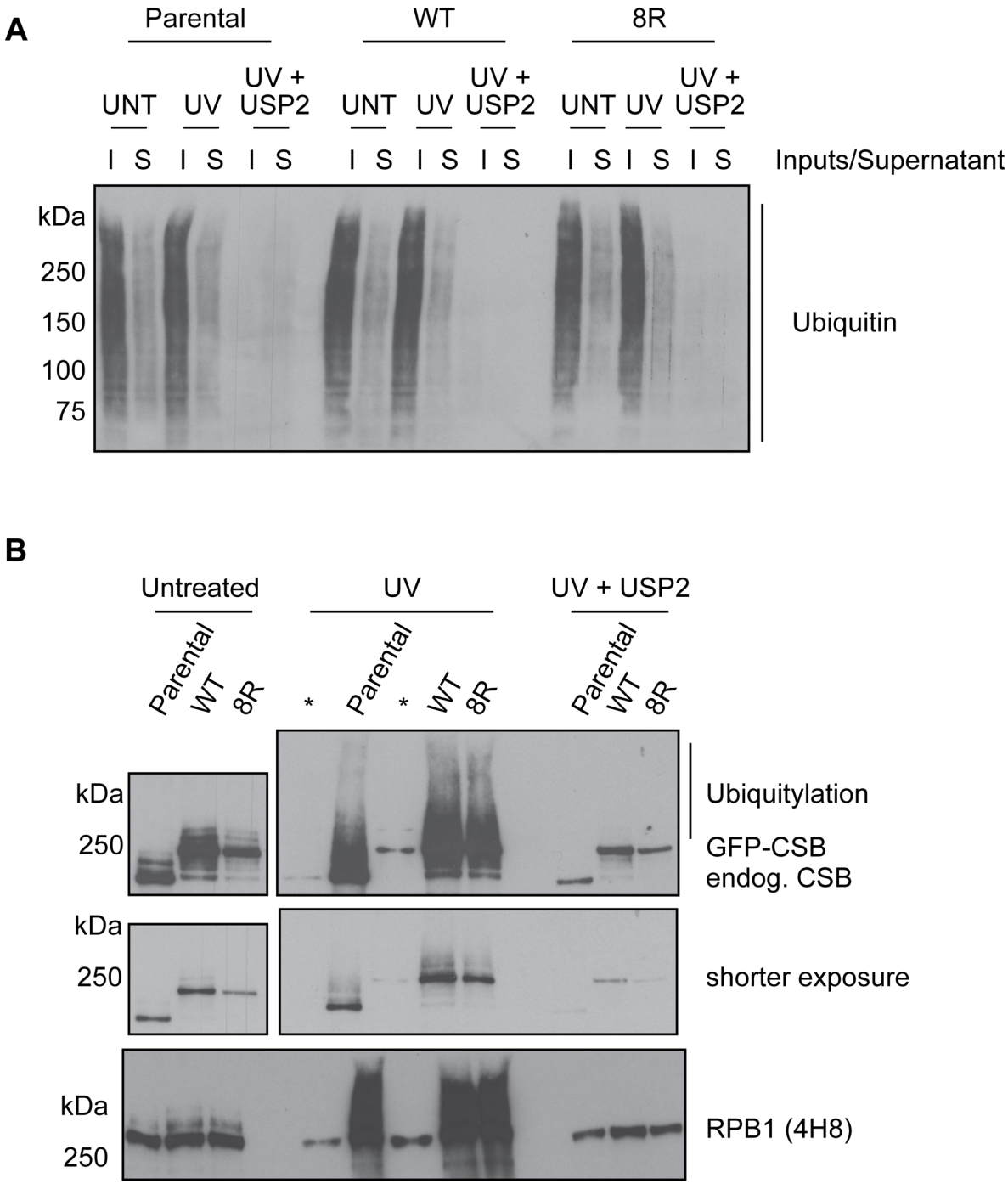


Figure 5-3 MultiDsk enrichment for ubiquitylated proteins reveals CSB 8R is still ubiquitylated. A Western blot analysis shows MultiDsk pulldown efficiently depletes ubiquitylated proteins from cell lysates. **B** MultiDsk eluates show CSB is ubiquitylated after UV irradiation and extent of ubiquitylation in transgene CSB is the same as endogenous CSB. CSB 8R is ubiquitylated to same extent as CSB WT. Blots were stripped and re-probed with 4H8 antibody that recognises the phosphorylated RPB1 subunit of RNAPII. This acts as a control for loading and ubiquitylation after UV irradiation.

5.5 Identification of additional CSB ubiquitylation sites

Stefan Boeing, who performed the original mass spectrometry screen that identified the 5 CSB lysine residues described above, had in the meantime also performed a new diGly mass spectrometry experiment. This experiment was performed as part of a collaboration with another lab and used cisplatin as the DNA damaging agent instead of UV irradiation. Cisplatin induces both intra-, and to a lesser extent, inter-strand DNA crosslinks, the former of which are RNAPII-stalling lesions and hence are repaired by TC-NER (Damsma et al., 2007; D. Wang & Lippard, 2005). Hence, any ubiquitylated residues identified in this screen were assumed to potentially be generally important for CSB's role in TC-NER.

The experiment was set up similarly to the mass spectrometry screen described in section 4.2. A SILAC methodology and enrichment with a diGly immunoprecipitation was used comparing 50 μM cisplatin treatment vs untreated in HEK293 cells overexpressing CSB. Interestingly, this revealed a total of 24 ubiquitylated CSB residues, which included all the 5 previously identified lysine residues and the 3 close-proximity residues mutated in CSB 8R. These were mostly clustered in the C-terminal region of CSB, with some in the ATPase region and some in the N-terminal region (Table 5-1).

All relevant lysine residues	CisPt Mass Spec	CSB 19R	CSB 8R	CSB 5R
258				
345				
606				
607				
650				
663				
725				
729				
751				
759				
774				
783				
971				
997				
1150				
1172				
1184				
1186				
1254				
1257				
1258				
1295				
1359				
1360				
1363				
1392				
1457				
1487				
1489				

Table 5-1 Mapped CSB ubiquitylated lysine residues and those mutated in different constructs

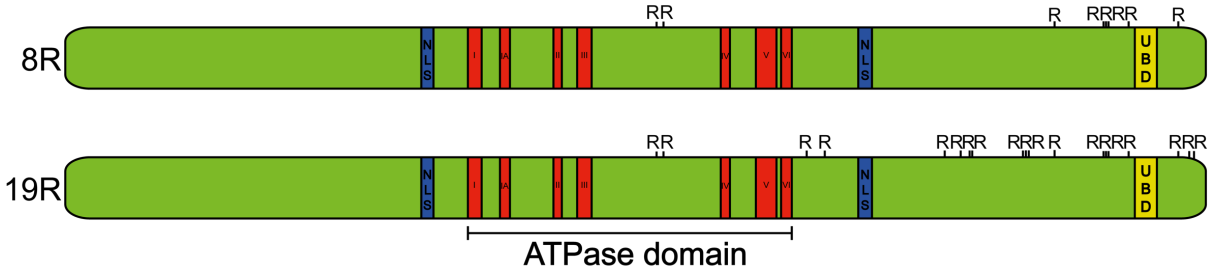


Figure 5-4 Schematic of CSB with relative positions of mutated ubiquitylated residues

5.6 Mutagenesis and cloning of 19R

Due to the identification of several more CSB ubiquitylation sites, I decided to investigate if further mutagenesis could ablate CSB ubiquitylation. The likelihood of missense mutations within the ATPase region inactivating the catalytic activity of CSB meant that no new mutations in this region were made. However, K774R and K783R from CSB 8R were maintained as it was already proven that these mutations did not confer a UV-sensitive phenotype but may accept ubiquitin in the presence of other arginine mutations. As most of the identified lysines were clustered in the C-terminal, which was the case in the UV mass spectrometry screen too, I decided to mutate all of these but not the two lysine residues in the N-terminal. Any lysines within three residues of those being mutated were also mutated to arginine and this was applied sequentially (K1184, 1257, 1359, 1360, 1489). This was to ensure these residues wouldn't serve to compensate as ubiquitin acceptors if the nearby mapped site was mutated.

Rather than doing several iterative rounds of site-directed mutagenesis, a DNA construct 2.2.4) was synthesised to replace the 3′ end (1,750 bp) of CSB 8R coding sequence thus preserving K774R and K783R and introducing a further 17 K>R mutations. An HpaI restriction site was introduced by silent mutation just upstream of the K971 codon in pDEST-FRT/TO-CSB8R plasmid to facilitate cloning of the DNA construct. Ultimately, traditional cloning methods using HpaI and BamHI restriction digest followed by ligation proved unsuccessful. As the synthetic DNA construct was double stranded, highly purified, and annealed with high affinity to the CSB 8R coding sequence, it was used as a primer pair in a two-step PCR reaction. This produced an amplified product, which was ligated to seal nicks, transformed into DH5-alpha E. coli and purified. Sanger sequencing revealed the product contained all of the intended 19 K>R mutations. Stable cell lines were produced by transfection of this plasmid (pDEST-FRT/TO-CSB19R) with the Flp recombinase (pOG44) as previously described (see section 0). Single cells were selected for with hygromycin and colonies grown from these cells were tested for the induction of CSB 19R in the absence and presence of doxycycline, in many cases showing levels of expression similar to CSB WT (see Figure 5-5).

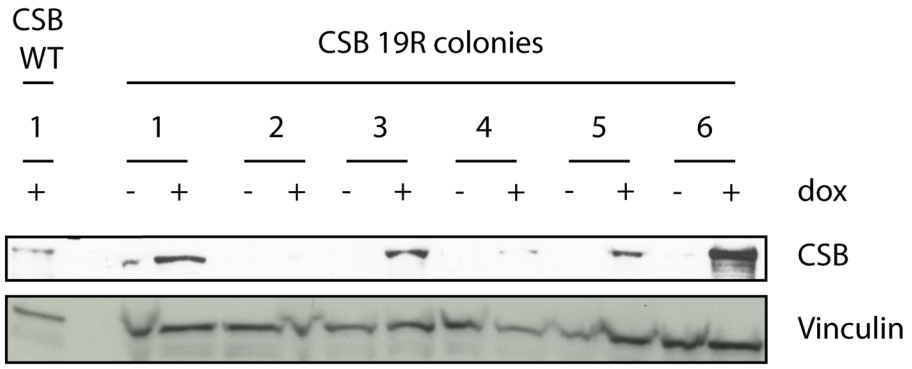


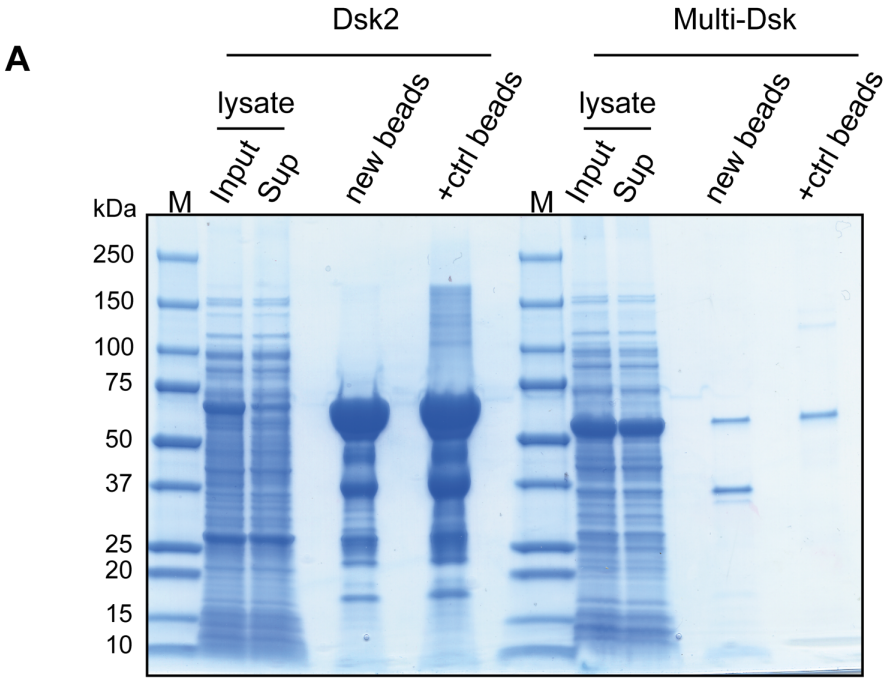
Figure 5-5 Induction of CSB 19R expression with doxycycline. Western blot analysis shows CSB 19R transgene expression is doxycycline dependent and of similar levels to CSB WT cell line. Blotted with anti-CSB antibody.

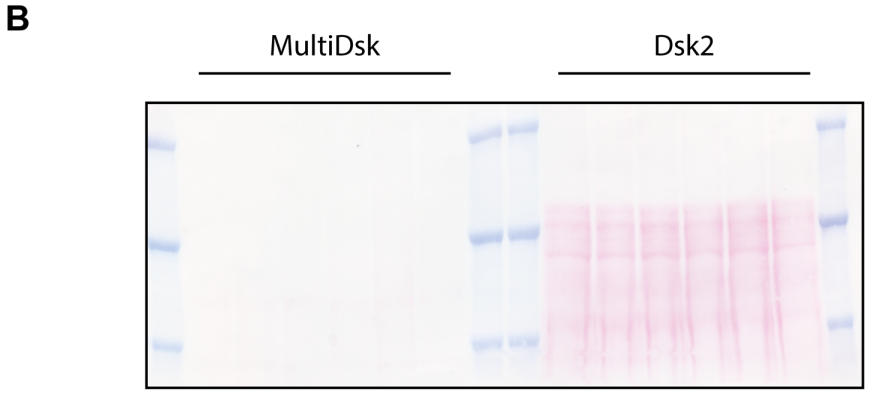
5.7 Purification of MultiDsk and Dsk2 resin

Previously, I had used the MultiDsk resin to evaluate the level of CSB ubiquitylation and had borrowed this reagent from a colleague. As I needed much more reagent for many experiments, I expressed and purified MultiDsk myself (see 2.8.4). Although in theory, MultiDsk should have high avidity for ubiquitin, as it is composed of 5 UBA domains from Dsk2 (M. D. Wilson et al., 2012), its expression and purification can be problematic and low yields lead to decreased efficiency in enriching ubiquitylated species. Therefore, I also expressed and purified full-length GST-tagged Dsk2, which has been shown to successfully enrich ubiquitylated proteins too (Anindya et al., 2007; Tufegdzic Vidakovic et al., 2019).

The problem with purifying MultiDsk is that it is expressed exclusively in inclusion bodies that need to be solubilised with sarkosyl, which also denatures the protein that then must be refolded with Triton X-100 (M. D. Wilson et al., 2012). This can lead to low yields of purified protein or denatured protein. Expression and purification of MultiDsk in E. coli worked well with acceptable yields (see Figure 5-6A). However, binding of GST-MultiDsk to glutathione beads was very inefficient (compare ‘Input’ with ‘Sup’ lane), suggesting that refolding of MultiDsk and the GST-tag may have been inefficient and inhibited binding to the glutathione beads. Although the amount of MultiDsk bound to beads was similar to the previous batch I had used (compare ‘new beads’ with ‘+ ctrl beads’), further testing showed that

Dsk2 beads pulled down much more protein – as seen by Ponceau staining of the eluates from an example experiment (see Figure 5-6C). Therefore, going forward Dsk2 beads were used to enrich ubiquitylated proteins from lysates.


A
Dsk2
Multi-Dsk
lysate
lysate
Input
Sup
new beads
+ctrl beads
Input
Sup
new beads
+ctrl beads
kDa
M
M
250
150
100
75
50
37
25
20
15
10


B
MultiDsk
Dsk2

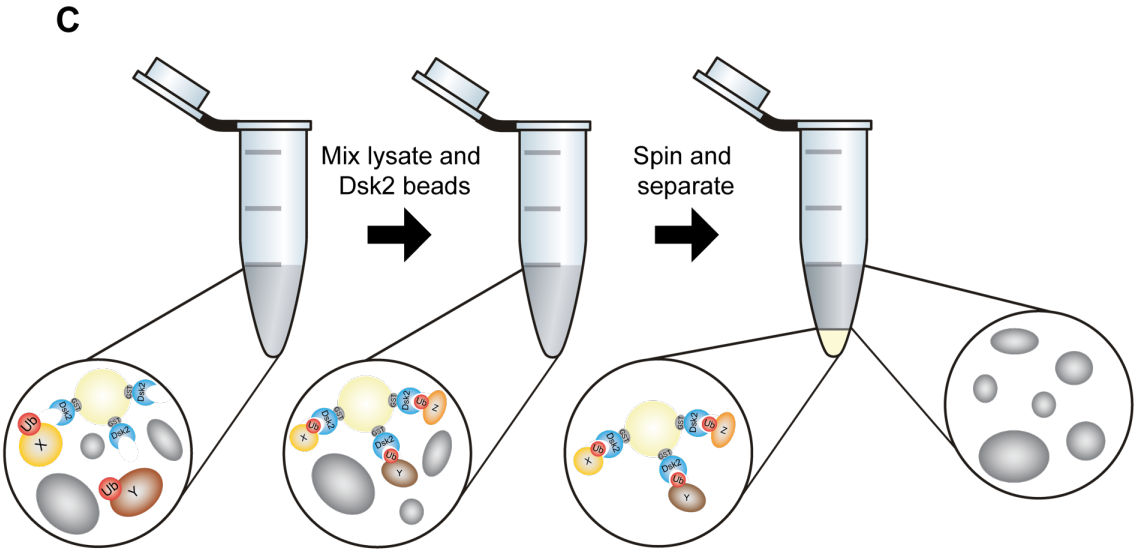

C
Mix lysate and
Dsk2 beads
Spin and
separate

Figure 5-6 Purification of MultiDsk and Dsk2 with overview of assay for enriching ubiquitylated proteins. A Coomassie stained gel showing expression and purification of GST-tagged MultiDsk and Dsk2. **B** Ponceau staining of elution's from MultiDsk and Dsk2 bead pulldowns reveals Dsk2 is much more capacity for ubiquitylated proteins per bead volume. **C** Schematic outlining protocol for ubiquitylated protein enrichment using Dsk2 beads.

5.8 CSB 19R is not ubiquitylated efficiently

Before attempting any phenotypic assays, I first characterised the ubiquitylation status of CSB 19R. Cells were grown with doxycycline for 24 hours to express the different CSB versions to equal levels. The experiment was performed similarly to the MultiDsk pulldown described earlier (5.4). Cells were either untreated or UV irradiated at 20 J/m^2 and lysed after a short recovery. Lysates were prepared and added to Dsk2 beads to enrich for ubiquitylated species. After washing, the contents of the beads were eluted and analysed by SDS-PAGE and Western blotting. As expected ,this showed an increase in ubiquitylation of CSB following UV irradiation in cells expressing CSB WT and CSB 8R (see

Figure 5-7A). By contrast, CSB 19R showed very little ubiquitylation after UV irradiation; certainly much less than CSB WT and 8R. Thus, it can be concluded then that the 19 K>R mutations in the C-terminal of CSB inhibit ubiquitylation that normally follows RNAPII-stalling DNA damage.

The same experiment was also performed in parental HEK293 cells and CSA KO HEK293 cells expressing endogenous CSB. Eluates from Dsk2 beads were analysed by Western blotting using antibodies against both CSB and RPB1. This revealed a slight inhibition in CSB ubiquitylation in CSA KO cells following UV irradiation (see

Figure 5-7B). Although the extent of this inhibition seems less than that of CSB 19R, it is hard to make comparisons between experiments due to differences in the sizes of endogenous and exogenous CSB, slight expression differences, and the variability in the quality of the Western blots between experiments. RPB1 is also ubiquitylated following UV irradiation, which is slightly inhibited in CSA KO cells as well. Previous results showed that CSA is not responsible for RPB1 ubiquitylation

(Anindya et al., 2007; Tufegdžić Vidaković et al., 2020) so this is likely an indirect effect.

There are several observations of note in these experiments that lead to several interpretations. While ubiquitylation of RNAPII in CSA KO cells is reduced, it is not completely ablated. This is consistent with proteomics data that show RPB1 is ubiquitylated in CSA KO cells (Tufegdžić Vidaković et al., 2020) and experiments that identify other E3 ubiquitin ligases that ubiquitylate RPB1 (Anindya et al., 2007; Harreman et al., 2009; Somesh et al., 2005, 2007; Yasukawa et al., 2008). CSB 19R also retains a very small amount of ubiquitylation as does CSB in CSA KO cells. The ubiquitin marks that remain could be distinct from the ones deposited by CSA and support a different function (like K991 in BER, for example). It could also be that proximal lysines can, to a small degree, act as ubiquitin acceptors in lieu of the cognate lysines that have been mutated to arginine. In any case these results validate that at least some of the many lysine residues identified in the diGly mass spectrometry screen are indeed ubiquitylated *in vivo* in response to UV-induced DNA damage. But there may still be ubiquitylated lysines unidentified. Using this assay to probe the ubiquitylation status of several other TC-NER factors in CSA KO cells would be interesting.

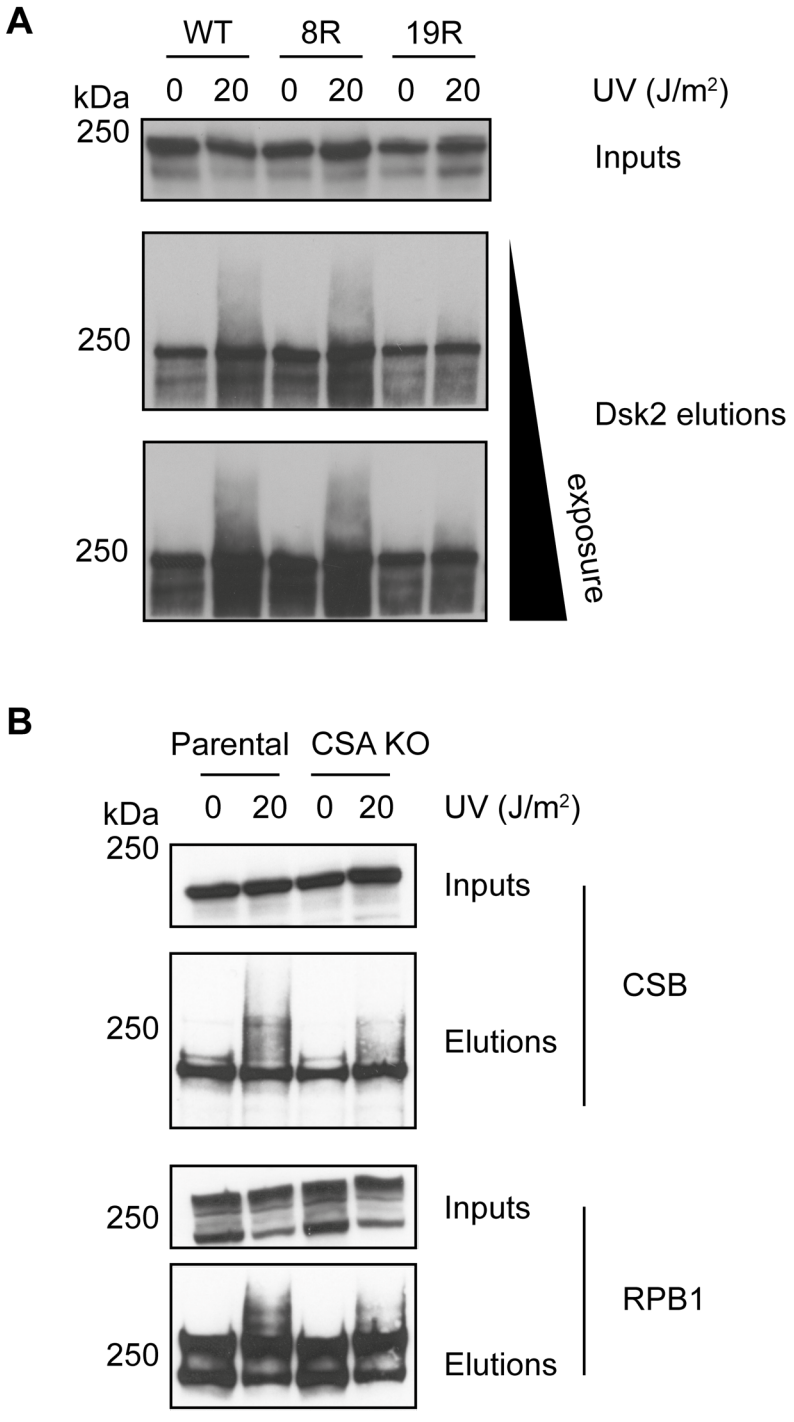


Figure 5-7 Dsk2 pulldown shows CSB 19R is less ubiquitylated than CSB WT in response to UV irradiation Dsk2 beads were used to enrich for ubiquitylated species in lysates prepared from cell lines expressing CSB WT, 8R, or 19R following mock or UV treatment. Elutes from these beads were analysed by SDS-PAGE and Western blotting with anti-CSB antibody. Inputs are sampled from the lysates taken before adding to Dsk2 beads showing equal expression of CSB in all cell lines.

5.9 CSB 19R is UV sensitive

The observation that CSB 19R shows almost no ubiquitylation following UV irradiation led us to ask if this conferred a UV sensitivity phenotype to cells. This would indicate that ubiquitylation of CSB is functionally consequential for its role in TC-NER. As described above, measuring cell survival following RNAPII-stalling DNA damage provides a rudimentary but effective way to broadly assay for a defect in TC-NER. Traditionally this has been done using the clonogenics assay (Selzer et al., 2002), but this is not suitable with HEK293 cells. Another approach is to use live-cell imaging technology that also offers several other advantages, such as being quicker, more scalable and offering close to real-time readout of cell growth. To this end, the IncuCyte system was employed which allows live-cell imaging of 96-well plates, while inside an incubator.

CSB KO, CSA KO, CSB WT, CSB 8R and CSB 19R HEK293 cell lines were grown in doxycycline for 24 hours to allow equal expression of CSB transgenes, where present. Each cell line was then seeded in triplicate for each condition into a 96-well plate. The following day, after cells had attached to the plate, they were either left untreated or subject to doses of 5, 10, or 20 J/m^2 of UV irradiation. The plate of cells was immediately transferred to the IncuCyte in a standard tissue-culture incubator where each well was imaged in 4 different locations every 3 hours for several days until they reached confluency. The graphs in Figure 5-8 show the surface area of the plate covered by cells as they grow over time.

In untreated conditions, all cell lines grew at very similar rates. Slight differences in growth can be explained by small differences in initial seeding density that always occurs due to natural variability in cell counting, dilution of cell solutions and pipetting errors of small volumes. In UV treated wells, the differences in cell confluency at 0 hours was exacerbated due to the removal and replacement of media which detaches some of the cells. Reattachment of cells at imaging sites in a non-even distribution skews results. Several replicates help to average out these biases. In UV-treated conditions there were differences in the growth response of the different cell lines. As expected, CSB KO cells were most sensitive to UV,

showing growth inhibition at as little as 5 J/m^2 and complete growth inhibition at 10 and 20 J/m^2. This resulted in eventual cell death as observed by the rounding up and detachment of cells in microscopy images at later time points (data not shown).

CSB WT expressing cell lines, as expected, were most resistant to UV-induced growth inhibition and only the high dose of 20 J/m^2 induced a lag in recovery before cell growth resumed, albeit at a slower rate than at the lower doses of UV irradiation (compare the slope gradient at log phase growth). CSB 8R expressing cell lines had very similar growth recovery profiles to CSB WT at 5 and 10 J/m^2. While a slower growth recovery at 20 J/m^2 seems apparent in CSB 8R cells, this is most likely an artefact of the lower initial cell density. The lag in recovery and the gradient of the slope during log phase growth is very similar to CSB WT cells and very different to the other cell lines showing UV-sensitivity.

Gratifyingly, CSB 19R and CSA KO cell lines showed very similar growth recovery profiles following UV irradiation, and while much more UV sensitive than CSB WT cells, they were not quite as sensitive as CSB KO cells. At 5 J/m^2 there was almost no difference in growth recovery in these cell lines compared to CSB WT, in contrast to CSB KO cells. However, at 10 and 20 J/m^2 there was a significant lag before the recovery of growth in these cell lines, but they both eventually continued to grow again, in contrast to CSB KO cells. To the best of my knowledge, a direct comparison between CSB and CSA KO cells in the same background has not previously been performed. Instead, it has been presumed that they are genetical equivalents, because mutation of either gene causes Cockayne syndrome, and because CSA is required to modify CSB (Groisman et al., 2006; Laugel, 2013).

Most importantly, these results indicate that CSB 19R confers UV-sensitivity to cells, observed as a failure to recover cell growth relative to CSB WT cells. The degree of UV-sensitivity is not as severe as CSB KO but is similar to CSA KO. This could suggest that inhibition of CSB ubiquitylation only partially affects its role in the DNA damage response, and CSB 19R may represent a separation of function mutation. It also provides further evidence that CSA is part of the E3 ligase complex that ubiquitylates CSB: mutating the ubiquitylation sites or the ubiquitin ligase results in the same level of damage sensitivity.

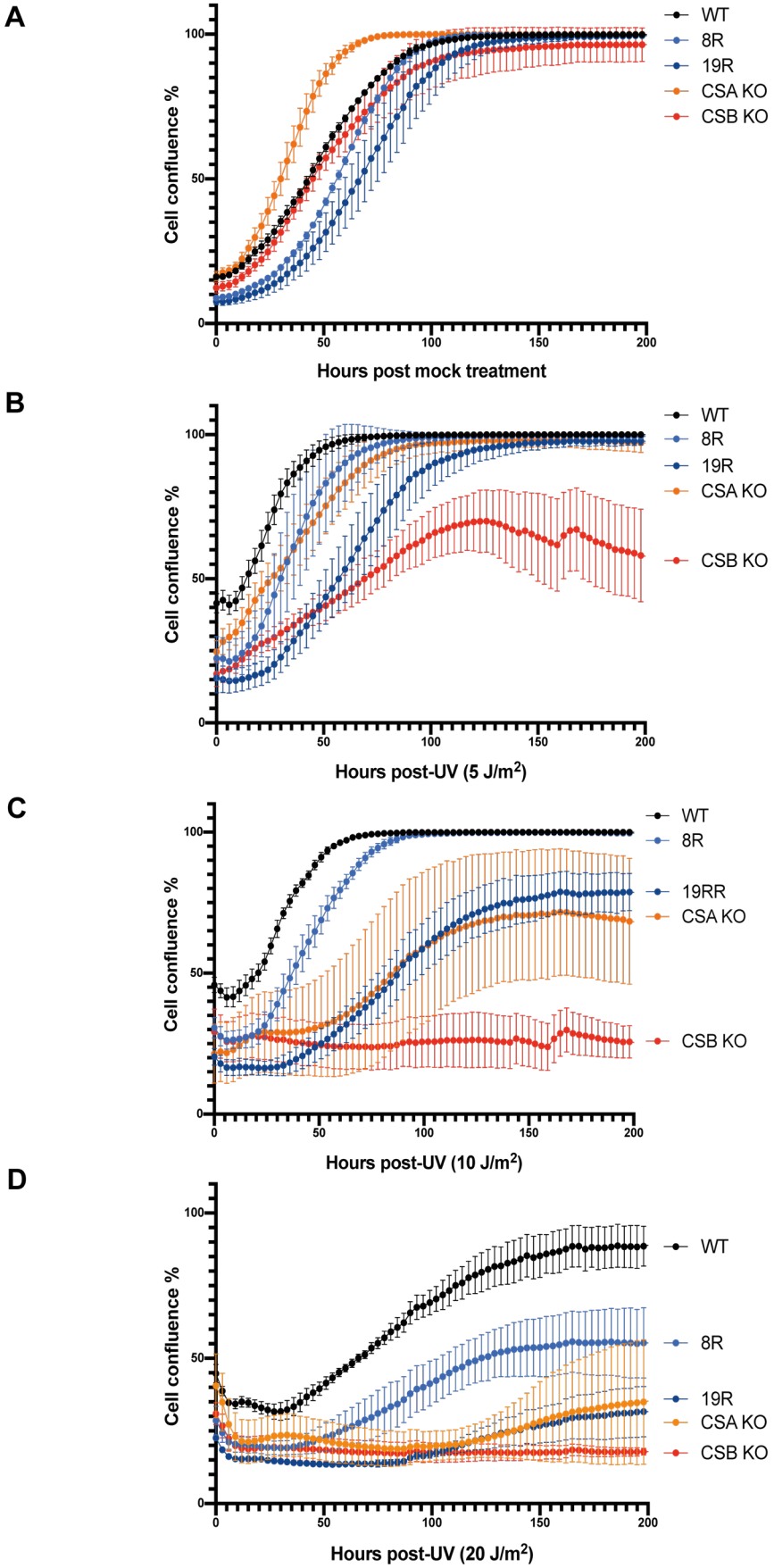

A
Cell confluence %
Hours post mock treatment
WT
8R
19R
CSA KO
CSB KO
B
Cell confluence %
Hours post-UV (5 J/m2)
WT
8R
19R
CSA KO
CSB KO
C
Cell confluence %
Hours post-UV (10 J/m2)
WT
8R
19RR
CSA KO
CSB KO
D
Cell confluence %
Hours post-UV (20 J/m2)
WT
8R
19R
CSA KO
CSB KO

Figure 5-8 IncuCyte growth curves show CSB 19R is UV-sensitive to a similar extent as CSA KO cells but less so than CSB KO cells.
Cells were seeded and grown in 96-well plates in an IncuCyte inside a cell culture incubator and imaged every 3 hours using phase contrast microscopy. Analysis was done using IncuCyte software that identifies all cells in an image field and calculates the surface area they occupy. Each sample was done in triplicate wells and images were taken in triplicate for every well. Data is shown as the mean of all images and wells for that timepoint. Error bars represent standard deviation. Data is representative of **A–D** represent increasing UV doses of untreated and 5, 10, 20 J/m^2.

5.10 CSB 19R is recruited to chromatin normally

It is well established that CSB is recruited to chromatin upon DNA damage, which is widely interpreted as CSB binding to stalled RNAPII (Lake et al., 2010; Van Den Boom et al., 2004). An inhibition of CSB 19R recruitment to chromatin following UV-irradiation would indicate that CSB ubiquitylation is important for recognition of, or stable association with, lesion-stalled RNAPII.

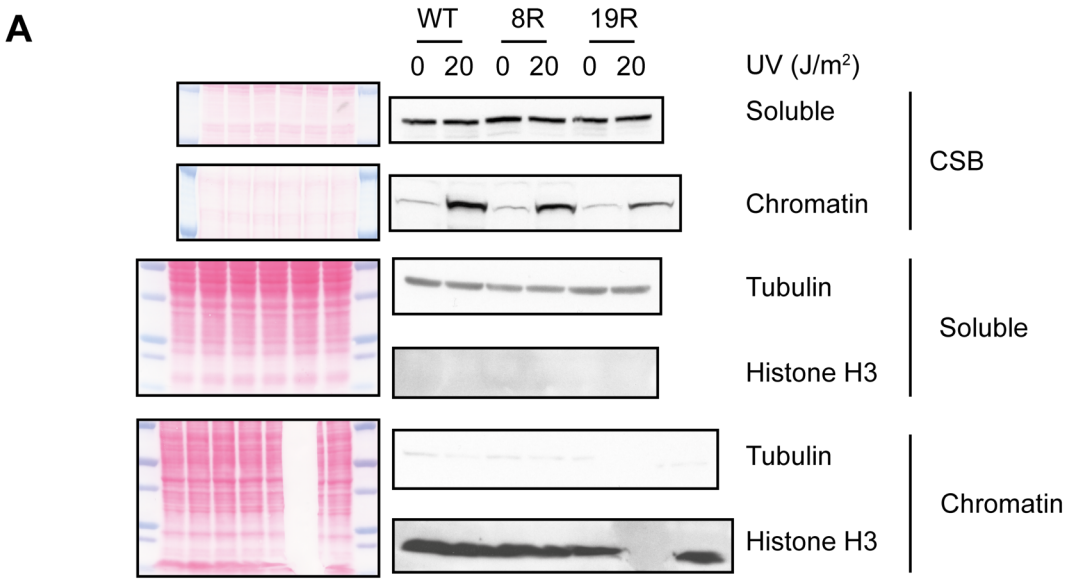
CSB WT, 8R and 19R cells were again cultured with doxycycline for 24 hours to express CSB transgenes to equal levels. Cells were then either left untreated or UV irradiated at 20 J/m^2 and left to recover for 1 hour. Cells were then harvested and subject to a fractionation procedure to isolate different cell compartments. The 'cytoplasmic' fraction was first collected by partially lysing cells in a hypotonic buffer along with Dounce homogenisation. After centrifugation and collection of the soluble supernatant, the resulting cell pellet was subject to lysis with a 500 mM NaCl buffer and benzonase digestion of DNA and RNA to fully release chromatin-bound proteins, resulting in the 'chromatin' enriched fraction.

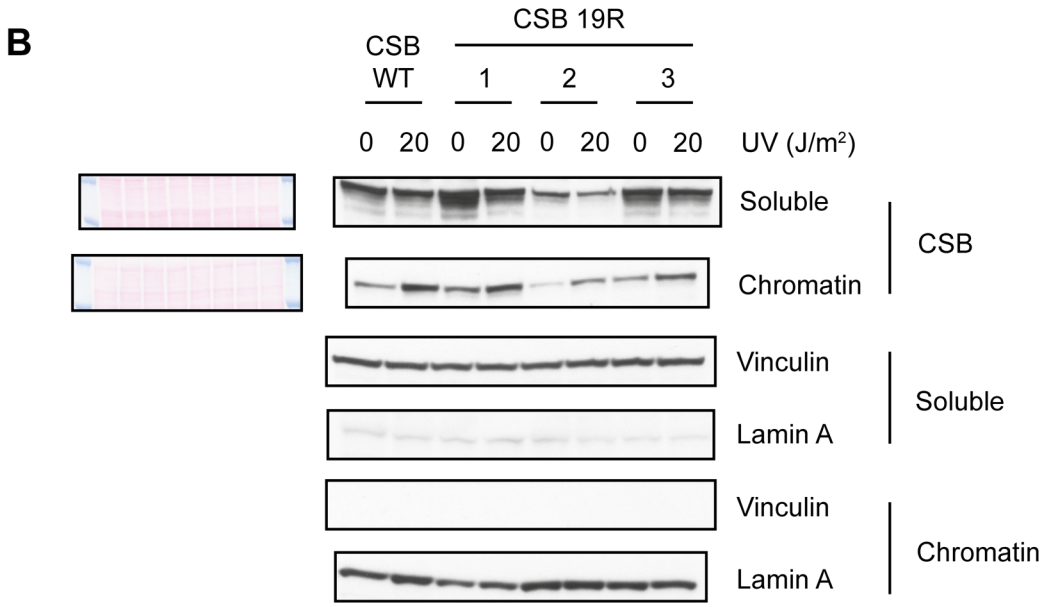
The successful fractionation of the different cell compartments was confirmed by Western blotting. As expected ,the soluble fraction contained tubulin, but not histone H3, while the chromatin fraction contained histone H3, but not tubulin (see Figure 5-9, bottom panels).

This assay revealed that CSB WT mainly resides within the soluble fraction under normal conditions, with only a small fraction residing on chromatin. CSB in the soluble fraction is almost certainly all from the nucleoplasm as CSB contains two nuclear localisation domains and immunofluorescence studies show CSB residing wholly in the nucleus (Iyama et al., 2018). Equal amounts of lysate by mass were

loaded on the gels in Figure 5-9; however, the yield of soluble lysate is approximately twice that relative to chromatin lysate for a given number of cells. Therefore, the proportion of cytoplasm to chromatin is not equimolar in the gels below and loading is biased toward chromatin. After UV irradiation, CSB WT binding to chromatin is dramatically increased. This also holds true for CSB 8R and CSB 19R. The degree of recruitment to chromatin of CSB 19R following UV irradiation seems to be slightly less than CSB WT, but there is still a dramatic increase of binding to chromatin compared to untreated conditions. Repetitions of this experiment produced variable results in the level of recruitment to chromatin. However, testing three differently cell line clones expressing CSB 19R showed there was a consistent increase in chromatin binding following UV, even if expression levels were slightly different (see Figure 5-9B).

I also performed this assay in parental and CSA KO cells that both express endogenous CSB (see Figure 5-9C). CSB is less ubiquitylated in CSA KO cells following UV irradiation. However, CSB is still recruited to chromatin after UV in these cells to the same extent as in parental cell lines. This indicates that CSB ubiquitylation is not required for UV-dependent association with chromatin and probably also not for the interaction with lesion-stalled RNAPII. It also suggests that CSB ubiquitylation is a downstream event of CSB's association with lesion-stalled RNAPII and may be required for the resolution of stalled RNAPII.


A
WT
8R
19R
0 20 0 20 0 20
UV (J/m²)
Soluble
Chromatin
CSB
Tubulin
Histone H3
Soluble
Tubulin
Histone H3
Chromatin


B
CSB 19R
CSB WT
1
2
3
0 20 0 20 0 20 0 20
UV (J/m²)
Soluble
Chromatin
CSB
Vinculin
Lamin A
Soluble
Vinculin
Lamin A
Chromatin

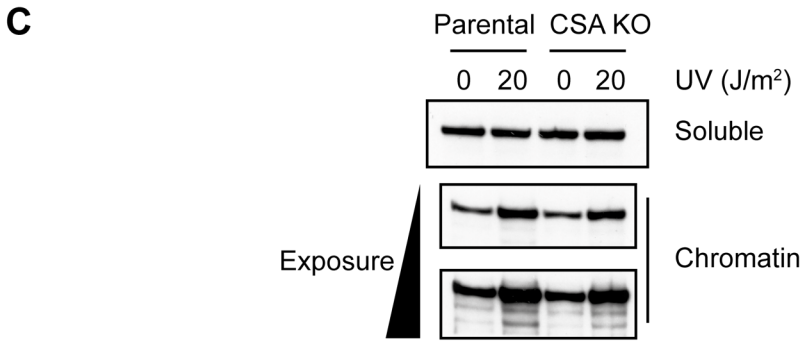

C
Parental
CSA KO
0 20 0 20
UV (J/m²)
Soluble
Exposure
Chromatin

Figure 5-9 Chromatin fractionation of cells shows CSB 19R still recruited to chromatin following UV irradiation A Western blot analysis of lysates after cell fractionation before and 1 hour after 20 J/m^2 of UV irradiation. Ponceau S-stained membranes on the left show equal loading. Tubulin and histone H3 blots show success of fractionation. **B** Same as in **A** except performed on 3 clones of CSB 19R cell lines compared to CSB WT cells. Vinculin and Lamin A blots show success of fractionation. **C** Same as in **A** except performed on parental HEK293 cells and CSA KO cells. All blots are a representative image from experiuments performed independently at least twice.

5.11 CSB 19R fails to restart transcription following UV

As described in the previous section (section 5.3), a qPCR method was adapted to measure nascent RNA production. This enables testing for the classic TC-NER defect of a failure to recover RNA synthesis following RNAPII-stalling during DNA damage. Given the interesting results showing that CSB 19R cells are UV-sensitive (section 5.9), it was important to phenotypically characterise this CSB mutant further using common model assays for TC-NER defects.

The experiment was carried out as described in section 5.9 using primers amplifying EXT1 nascent RNA, but with the additional cell line CSB 19R used. In one experiment, cells were also assayed at an additional, mid-recovery timepoint of 15 hours, which was used in the hope that it would reveal slower kinetics of CSB 19R repair if it was able to fully recover transcription at 24 hours, but had slower repair kinetics than CSB WT (Figure 5-10A).

The dramatic shutdown of transcription at 3 hours post-UV irradiation in all cell lines confirms that the experiment worked as expected. CSB WT and CSB 8R cell lines recovered transcription fully at 24 hours (Figure 5-10A and B). and their rate of recovery was similar, as seen at 15 hours (Figure 5-10A). CSB KO cells failed to recover transcription within 24 hours post-UV irradiation, which is consistent with these cells being deficient for TC-NER and being unable to remove RNAPII-stalling damage outside of the kinetically slower GG-NER pathway (Figure 5-10A, B, C). CSB 19R cells showed no recovery at 15 hours, but at 24 hours, had recovered slightly more than CSB KO, but not to the level of CSB WT and CSB 8R(Figure 5-10A, B, C). This indicates that CSB 19R retains some repair capacity, while not being as fully functional as CSB WT.

The level of transcription recovery of CSB 19R at 24 hours is slightly variable between the two independent experiments (Figure 5-10A and B), but always more than CSB KO cell and less than CSB WT cell and statistically significant when the biological replicates are analysed together (Figure 5-10C). I therefore decided that to get more robust and comprehensive data on the level of transcription restart after UV in CSB 19R cells, I would employ a genome-wide sequencing approach to measure nascent transcription.

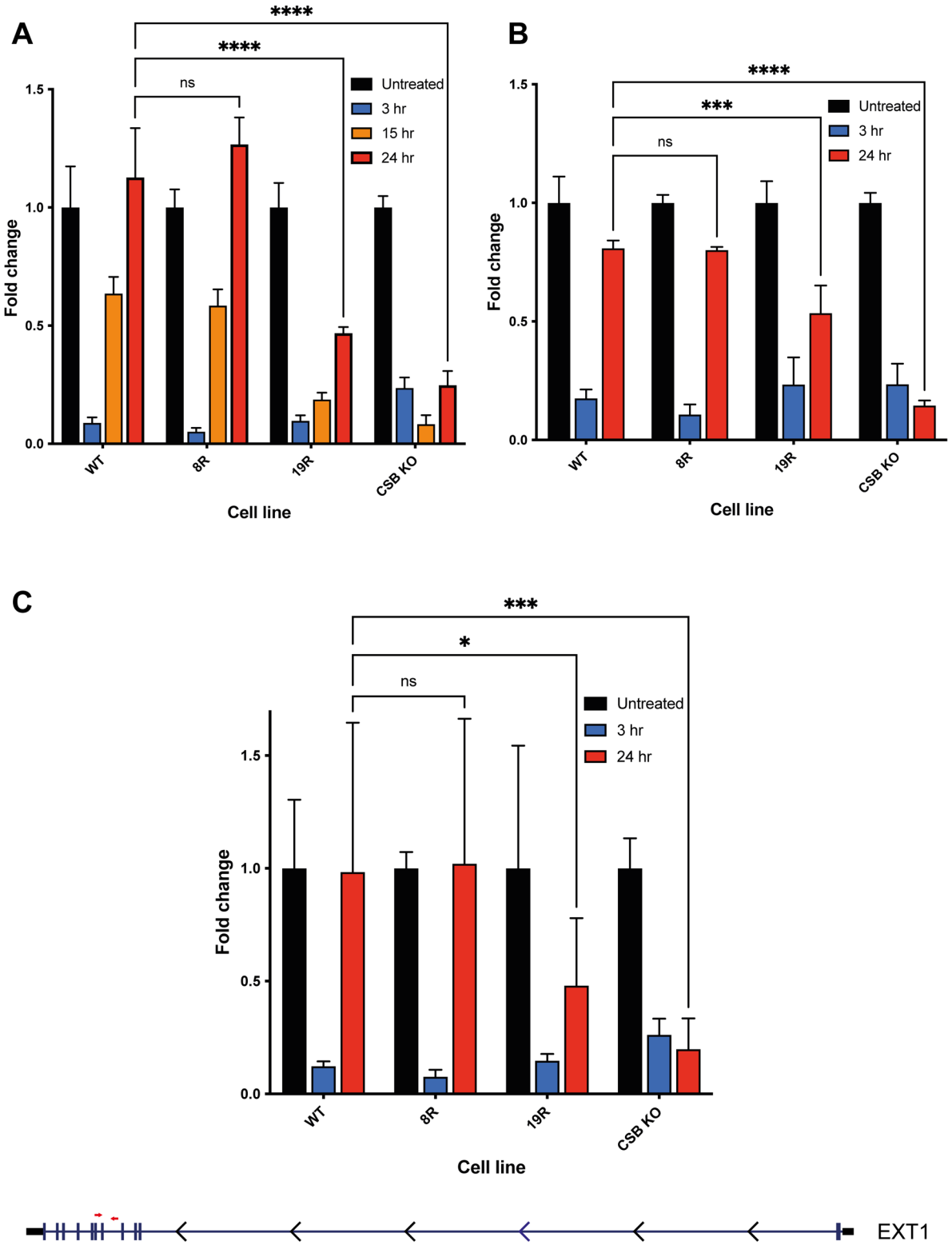

A
****
****
ns
Untreated
3 hr
15 hr
24 hr
Fold change
1.5
1.0
0.5
0.0
WT
8R
19R
CSB KO
Cell line
B
****
***
ns
Untreated
3 hr
24 hr
Fold change
1.5
1.0
0.5
0.0
WT
8R
19R
CSB KO
Cell line
C
***
*
ns
Untreated
3 hr
24 hr
Fold change
1.5
1.0
0.5
0.0
WT
8R
19R
CSB KO
Cell line
EXT1

Figure 5-10 RT-qPCR of nascent RNA shows CSB 19R fails to restart transcription following UV irradiation A An independent qPCR experiment of nascent EXT1 transcript at several time points after UV irradiation of 20 J/m^2 **B** An independent qPCR experiment of nascent EXT1 transcript at several time points after UV irradiation of 20 J/m^2. **C** Analysis of experiments A and B as biological replicates for samples for which there are 2 biological replicates. Each experiment (A and B) was performed with technical triplicates and data are represented as the mean of all replicates normalised to GAPDH mRNA and then the respective untreated condition for each genotype. Error bars represent standard deviation of the technical replicates. Primers target nascent transcript by amplifying an intron-exon junction approximately 300 kb downstream of the TSS of EXT1 gene. Two-way ANOVA multiple comparisons were done for 24-hour timepoints between CSB WT and all other cell lines. Adjusted P values are represented by asterixs, where ns = P> 0.05; * = P <0.05; *** = P < 0.001; **** < 0.0001.

5.12 CSB 19R has genome-wide transcription recovery defect

To get a snapshot of transcription across the entire genome, cells can be metabolically labelled with an analogue of uridine, 4-thiouridine (4SU), much like the 5EU labelling outlined in section 4.4. In this method, much shorter labelling times can be used due to very sensitive deep sequencing methods, which allows for a higher resolution 'snapshot' of transcription at specific timepoints. 4SU incorporates into RNA as it is transcribed by RNAPII and can be biotinylated to facilitate isolation of only 4SU-labelled nascent RNA from the pool of total RNA using streptavidin beads, which is then be used for DNA library preparation and deep sequenced (Gregersen et al., 2020).

CSB WT, 19R and CSB KO cells were cultured in doxycycline for 24 hours before being UV-irradiated at 20 J/m^2 or left untreated. Untreated cells were labelled by addition of 4SU to the media for 15 minutes, the reaction was stopped and cells lysed by addition of TRIzol. UV irradiated cells were incubated for 1, 5, and 24 hours and were labelled with 4SU for the final 15 minutes of incubation, then immediately lysed with TRIzol. Biological duplicates were done for each timepoint and cell line. Total RNA was purified and *S. cerevisiae* 4-thiouracil (4TU)-labelled RNA was spiked in relative to total amount of human RNA to enable normalisation. RNA was fragmented by controlled base hydrolysis to give a size distribution of approximately 25–500 nt. The quality and rate of fragmentation was analysed by TapeStation to confirm consistency before continuing. 4SU-labbeled RNA was biotinylated and purified using magnetic streptavidin-coated beads. Sequencing

libraries were prepared from purified 4SU-labelled RNA and sequenced using Illumina HiSeq 4000.

Bioinformatic analysis was performed by Richard Mitter. Replicates were merged, averaged and normalised relative to yeast RNA. Data is shown as metagene profiles normalised to the relevant untreated timepoint so as to give a ratio of transcription relative to 'normal' conditions. Genes were stratified by length as transcription defects might only become apparent several kb from the transcription start site (TSS).

As expected, 1 hour after UV, transcription in all cell lines is shifted toward the TSS as damage halts RNAPII progression – seen as a tall peak in the 5′ region of the gene (see Figure 5-11A) and transcription in the gene body falls below 1 in protein coding genes. This is completely in line with previous results obtained for TT-seq (Tufegdžić Vidaković et al., 2020) and confirms that the UV treatment and capture of nascent transcription has worked. At 5 hours post-UV, the peak over the TSS declines, which coincides with degradation of RNAPII and a decline in total transcription, but this is more pronounced in CSB KO cells. At 24 hours, CSB KO cells have much reduced transcription across all gene lengths, at any distance from the TSS. This is consistent with the lack of transcription restart in these cells. In contrast, there is only slightly less restart of transcription in CSB 19R cells than in WT cells, which is more prominent in >100 kb genes toward the end.

These data only show significant transcription activity up to 100 kb from the TSS, but if the trend seen for CSB 19R at 24 hours post-UV continues further into long genes, then at 300 kb from the TSS, transcription may well be significantly lower than CSB WT. This would agree with the qPCR data which was measured at approximately 300 kb from the TSS in the EXT1 gene. However, looking at the replicates independently, it can be seen that while the two CSB 19R replicates correlate almost perfectly, the CSB WT replicates have slightly different levels of transcription after UV exposure. While the CSB 19R profiles are quite different from one of the CSB WT replicates at 24 hours post-UV they are almost identical to the other replicate. This raises the possibility that the merged replicate data might be misleading and that CSB 19R can actually recover transcription after UV similarly

to CSB WT. While genome-wide data is generally more robust than qPCR, especially for drawing conclusions about genome-wide transcription changes, these data make it difficult to draw firm conclusions on the extent to which CSB 19R cells can restart transcription. However, both the TT-seq and qPCR data indicate some defect in the ability of CSB 19R to promote transcription recovery. Repeating this experiment with more replicates, or using multiple different clones of the same genotype, would reduce any variability and noise in the data arising during multi-step sample preparation.

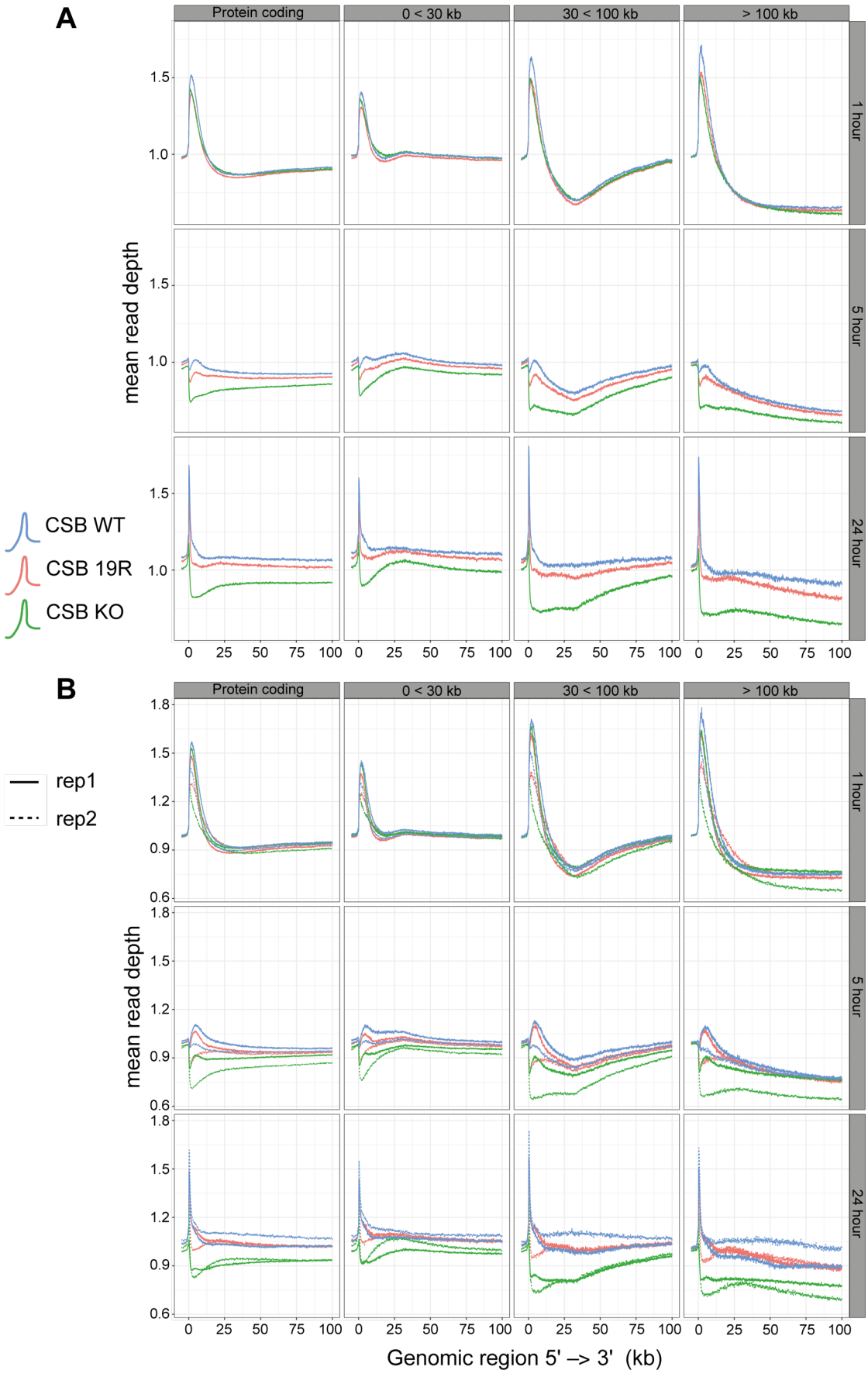

A
Protein coding
0 < 30 kb
30 < 100 kb
> 100 kb
1 hour
5 hour
24 hour
mean read depth
CSB WT
CSB 19R
CSB KO
B
rep1
rep2
Genomic region 5' –> 3' (kb)

Figure 5-11 CSB 19R RRS phenotype is less severe when assayed genome-wide compared to qPCR A Data shown are metagene profiles of all protein coding genes (n = 19,919) or groups of genes stratified by gene length (0<30 kb = 10,579, 30<100 kb = 5,783, >100 kb = 3,557) of merged duplicates. Profiles are all normalised to untreated conditions. **B** Data as in A but shown for individual replicates.

5.13 Constitutive deubiquitylated CSB by DUB fusion

Although CSB 19R shows a great reduction in ubiquitylation after UV, there still remains some ubiquitylation. Also, UV sensitivity of CSB 19R cells is not seen at low doses of UV. This may be due to remaining lysine residues that can act as acceptors of ubiquitin and compensate somewhat for the K>R mutations, resulting in a mild phenotype. There is, of course, also the possibility that such a high number of mutations confers structural alterations to CSB that interfere functions other than ubiquitylation, such as its catalytic activity. To investigate this, I am currently purifying recombinant CSB 19R to check it still maintains its ATPase activity in vitro, but this work remains incomplete.

Given the results with incomplete ablation of ubiquitylation, I also developed another mutant that I hoped would more efficiently prevent CSB ubiquitylation, while not introducing mutations that might confer other functional impediments. To this end, CSB WT was fused at the C-terminus with the catalytic domain of USP2, a ‘promiscuous’ DUB used recombinantly in Figure 5-3. Another 'control' construct was made, fusing CSB to USP2 but with the mutation C276A, which inactivates the catalytic activity of USP2. This is a much more robust control than was available for CSB 19R as it rules out the possibility of the USP2 protein interfering with CSB function in any other way other than its deubiquitinase catalytic activity.

Figure 5-12 CSB USP2 fusion construct

5.14 CSB-USP2 fusion protein is constitutively deubiquitylated

Once clones were established, I performed a Dsk2 pulldown to assess the level of ubiquitylation of CSB-USP2 before and after UV-irradiation. This assay was performed as described previously in 5.8.

The results show that CSB-USP2 WT is not ubiquitylated following UV irradiation, whereas CSB-USP2 C276A is still efficiently ubiquitylated (see Figure 5-13A). Note that CSB-USP2 C276A was slightly more expressed than CSB-USP2 WT, but not to the extent that it can account for the differences in ubiquitylation. The level of ubiquitylation of CSB-USP2 WT was even less than in CSA KO cells (see Figure 5-13B) indicating that USP2 removes CSA-independent ubiquitin marks from CSB as well. The same blots were then stripped and re-probed with an antibody against RPB1. This acts as a control for equal loading and the inputs indeed showed equal loading. It also revealed that RPB1 is slightly less ubiquitylated in CSB-USP2 WT cells to a similar extent as seen in CSA KO cells. This nicely supports the idea that CSB is dynamically associated with RNAPII upon UV-irradiation, but also makes it more difficult to make strong conclusions based on any phenotypic effect of expressing CSB-USP2 WT, as it might be acting indirectly, through de-ubiquitylation of CSB interactors, for example.

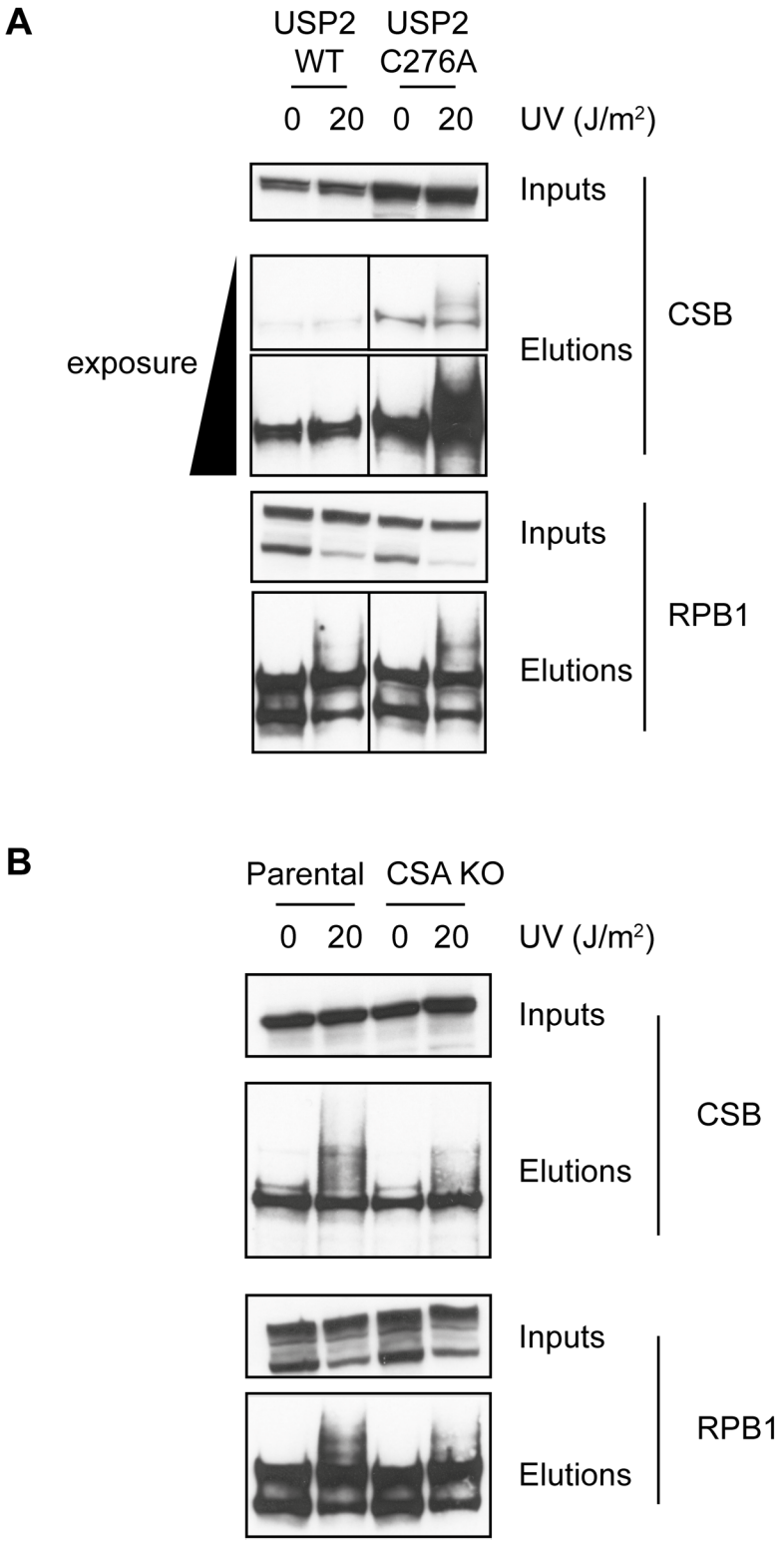


Figure 5-13 CSB-USP2 WT fusion protein is constitutively deubiquitylated
Western blot analysis of dsk2 pulldown of cell lysates before and after UV irradiation. Cell lines either express CSB fused to USP2 WT or a catalytic dead mutant, USP2 C276A.

Chapter 6. Discussion

6.1 Reconstituted transcription as an assay for DSIF and CSB interactions on RNAPII

The aim of Chapter 3 was to investigate the interactions of CSB and DSIF on RNAPII using a simplified *in vitro* transcription system. The fate of RNAPII stalled at a lesion remains one of the fundamental unknowns of the molecular mechanism of mammalian TC-NER. Structural studies of elongating RNAPII bound to DSIF (Bernecky et al., 2016, 2017; Klein et al., 2011; Martinez-Rucobo et al., 2011) and RNAPII bound to Rad26 (W. Wang et al., 2018; J. Xu et al., 2017) reveal overlapping binding sites on RNAPII. These studies suggest that CSB would remove DSIF upon binding stalled RNAPII. The TEC assays seemed like a perfect simplified system in which to definitively answer this question.

Obvious improvements to the protein purifications performed for the TEC assay could be made. While CSB purification was pure of any other contaminating proteins (Figure 3-4D), there was still a small number of degradative products that might necessitate further purification for publication if results were positive. FLAG and 8xHIS affinity tags were included in the C-terminus of CSB alongside the 2xStrepII tags on the N-terminus for this reason. Ultimately, further purification wasn't deemed necessary in this instance for the TEC assays. However, purifications attempted at a later date for testing CSB ATPase activity (data not shown) were two-step affinity purified, first via N-terminal StrepII tags followed by C-terminal HIS tags on a Ni-NTA column. Eluates from the StrepTrap column showed good yields of CSB but with some contaminants. These were loaded onto a Ni-NTA column and eluates were pure of contaminants and degradative products as determined by Coomassie staining, and Western blots with CSB antibody. Unfortunately, the yields after Ni-NTA clean-up were very low. This was likely due to the presence of EDTA and DTT in the StrepTrap elution buffer, which chelates and strips Ni^{2+} ions from the column. Reversing the order of columns used, using a FLAG column for clean-up, or omitting EDTA and DTT should help in acquiring large yields of extremely pure CSB for future studies such as ATPase assays.

DSIF purification, while extremely pure, suffered from degradation from the N-terminus of Spt5 resulting in significant amounts of degraded protein (Figure 3-3 DSIF expression and purification. Size-exclusion chromatography of DSIF eluates would allow for easy enrichment of the full-length product, while sacrificing overall yields. However, expression of DSIF in bacteria was good and the ease of growing large volumes of bacteria mean obtaining sufficient yields of full-length Spt5 should not be a problem.

Ultimately the TEC assays were unsuitable for studying protein interactions of RNAPII on DNA due to most of the RNAPII binding to the streptavidin magnetic beads (Figure 3-6D). The small amount of DSIF competed from the bead fraction into the supernatant in reactions with CSB was likely non-specific (Figure 3-6B). Fundamentally, this assay was designed for studying transcription and analysing RNA synthesised during the reaction as the readout. Using RNA production as a readout bypasses many of the issues experienced with non-specific binding of proteins and improper formation of TECs as these do not contribute to the RNA signal. This assay might have been better placed to ultimately test the ability of CSB to promote backtracking in the presence of TFIIH by purifying RNA from transcription reactions carried out in its presence, which would of course mean purifying the 10-subunit complex. Using DNase digestion is a simple method to footprint RNAPII with its bound protein partners to reveal RNAPII's position on DNA, which would highlight any backtracking (Rice et al., 1993). The ability of TFIIH to backtrack RNAPII has already been investigated *in vitro* using a different method, which proved that in conjunction with XPG, RNAPII can indeed be backtracked at transcription bubbles. Interestingly, CSB's ATPase activity was stimulated by XPG on bubble DNA substrates, but wasn't necessary for backtracking (Sarker et al., 2005). Similar experiments could be performed with addition of DSIF to see if it inhibits backtracking of RNAPII by locking it in an elongated state. DSIF-induced Inhibition of RNAPII backtracking would create a testable readout for the effects of CSB. If CSB does indeed remove DSIF from RNAPII, it may remove DSIF-induced inhibition of backtracking. Of course a negative result, doesn't prove that CSB cannot remove DSIF from RNAPII as it too may lock RNAPII in a forward translocating state and inhibit backtracking.

Using a similar TEC set-up as described in section 3.2, but instead extracting RNA and separating it on a gel as the readout may provide a more robust, if not less accurate, method to discern how CSB interacts with RNAPII. This approach was not taken initiailly as it does not directly test protein interactions and one must infer from the changes in RNA products what interactions are likely occurring. This also means that the reacitons must be carried out on a DNA template absent a G-stop or lesion as RNA production is the readout and no transcription occurs when RNAPII is irreversibly stalled. This of course means that while one can test the interactions between CSB, DSIF and RNAPII this way, it would be on a transcribing RNAPII, not a stalled RNAPII. TEC reactions with RNAPII alone will have a distinct banding pattern of RNA products with a main band of the full-lenth RNA and several shorter products representing stalling sites. Addition of DSIF or CSB to the reaction would probably enhance production of the full-length RNA and reduce smaller RNAs produced by RNAPII stalling. However, because CSB and DSIF enhance processive elongation by different mechanisms (Christopher P. Selby & Sancar, 1997b; Wada et al., 1998), the pattern of RNA products produced in reactions in which either is present might be distinct. Thus, carrying out TEC reactions with only CSB or DSIF present would allow one to know the pattern of RNA products produced when RNAPII is bound by either protein. Addition of both proteins to the reaction in different orders and assessment of the final RNA products would reveal a ‘CSB-‘ or ‘DISF-like’ RNA pattern, which would indicate whether CSB was able to displace DSIF or vice versa. Of course, the pattern of RNA products produced by CSB or DSIF may be too similar to differentiate in which case this would not be a suitable readout.

In most *in vitro* studies investigating CSB’s biochemical mechanism, CSA was not included in the reaction. While Cockayne syndrome is heterogenous in its phenotypic presentation, it is caused exclusively by mutations in CSA or CSB, with no significant correlation between phenotype and genotype (Laugel, 2013). Given the identical phenotypic consequences of CSA or CSB mutations, it is likely that they act within the same pathway and are dependent for each other’s full functionality. In support of this, it has been shown that $CRL4A^{CSA}$ ubiquitylates CSB directly *in vitr*o and is needed for CSB’s ubiquitylation *in vivo* (Fischer et al., 2011;

Groisman et al., 2006). It therefore seems prudent that further *in vitro* experiments trying to reconstitute CSB's role in TC-NER be carried out with the full $CRL4A^{CSA}$ complex. However, much of what we know about CSB comes from *in vitro* biochemical characterisation done without CSA's presence, so clearly it is not essential for its core catalytic activity. It might be the case that CSA can only modulate CSB activity in certain cellular contexts (e.g., lesions within chromatin) that aren't recapitulated in *in vitro* reactions, negating CSA's essentiality. This is analogous to the role of DDB2 (XPE) of the UV-DDB complex in GG-NER, which when mutated gives rise to Xeroderma pigmentosum. However, it is not required for reconstitution of GG-NER *in vitro*, but mildly stimulates the reaction of cell free extracts (Kulaksız et al., 2005; Wakasugi et al., 2001). This is likely because UV-DDB facilitates lesion recognition in chromatin, which is lacking from reconstitution reactions (Fei et al., 2011; Luijsterburg et al., 2012; Osakabe et al., 2015). Analogously, CSA may not be needed for CSB's biochemical mechanism in reconstituted reactions, but this merits further examination. Of course, the practice of rarely publishing negative results probably biases the observation of a dearth of CSA *in vitro* experiments.

6.2 Ubiquitylation site mapping on CSB

The mapping of ubiquitylated lysines proteome wide by diGly enrichment has been reported many times in the literature (Elia et al., 2015; Emanuele et al., 2011; W. Kim et al., 2011; Povlsen et al., 2012; Wagner et al., 2011; G. Xu et al., 2010). One of these studies even mapped ubiquitylation in the DNA damage response to UV irradiation and identified K783, K1295, K1392, and K1457 as being ubiquitylated on CSB (Elia et al., 2015). This is in agreement with the ubiquitylation sites identified by Stefan Boeing and used as a basis for this study (Boeing et al., 2016 and Figure 4-1). The reason a diGly mass spectrometry screen in response to UV was repeated was to compare CSA knockout and parental cells to further enrich for TC-NER specific ubiquitylation events. CSB was overexpressed in this experiment as it is very lowly expressed endogenously and the peptide counts were very low in the Elia et al., 2015 study. While Stefan's approach looked specifically for CSA-dependent ubiquitylation events with a CSA KO vs parental cell arm, all samples

were UV irradiated, precluding comparison with non-irradiated cells. While the SILAC approach doubles sample sizes by having a light and heavy sample for each arm, if feasible, I would have included a non-irradiated treatment arm to specifically enrich for CSA- and UV-dependent ubiquitylation events. The overexpression of CSB may favour non-physiological PTMs and other ways to enrich for lowly abundant CSB peptides could have been used, such as nuclear extracts or immunoprecipitation of endogenous/exogenous CSB or RNAPII. However, the identification of 5 lysine residues ubiquitylated on CSB after UV irradiation in a CSA-dependent manner, with 4 of those sites corroborated by another study, is indicative of the robustness of this experiment.

The initial diGly mass spectrometry screen identified 5 lysine residues that were ubiquitylated on CSB in a CSA dependent manner. As well as a construct with 5 K>R mutations, one with 3 additional lysines mutated in close proximity was made (8K>R). This strategy was taken due to the apparent promiscuous nature of E3 ligases, which is exemplified by poor conservation of ubiquitylated lysine residues in mammals and multiple ubiquitylation sites within proteins (Danielsen et al., 2011). For example, it has been demonstrated for Sic1, a substrate of the SCF E3 ligase, that ubiquitylation of any of 6 lysine residues in its N-terminus is sufficient for its degradation *in vivo* and mutation of all 6 is necessary to inhibit turnover (Petroski & Deshaies, 2003). Many other examples exist of proteins with multiple lysine substrate sites of a single E3 ligase (Blondel et al., 2000; Chi et al., 2001; Ganoth et al., 2001; Nakamura et al., 2000). CRLs do not possess intrinsic catalytic activity but complex with E2 enzymes, bringing them into the proximity with the substrate. An interesting observation of the CRL family (the largest E3 ligase family) is that their structure might underlie this substrate plasticity. The structure of all CRLs is similar, but that of $CRL4A^{CSA/DDB2}$ is most relevant here and has been well characterised (Cavadini et al., 2016; Duda et al., 2008; Fischer et al., 2011). The substrate receptor subunit (CSA/DDB2) is linked to the long flexible cullin scaffold via DDB1 at one end while the E2 binds the cullin arm via RBX1 at the other end. Activation of CRLs by neddylation of the cullin arm induces a more flexible and open conformation. This allows for structural flexibility that determines a ubiquitylation zone of 30–110 Å (Duda et al., 2008). CRLs have to adopt several confirmations as they must accommodate many different substrate receptor

subunits. For $CRL4A^{DDB2}$ the ubiquitylation target (XPC) is different from the recognised substrate (CPD DNA lesion) (Scrima et al., 2008). Mutation of cognate lysines then might simply shift ubiquitylation to nearby acceptor sites and therefore any lysines within 10 residues were also mutated to arginine. This of course isn't a fool proof strategy as residues far away in primary sequence may be proximal in tertiary structure. However, no structures of the C-terminal region of CSB or its species homologues exist due to its disordered nature. This might explain the WT phenotype of CSB 8R, which was still ubiquitylated after UV and not sensitive to UV irradiation (Figure 5-3, Figure 5-7, Figure 5-8Figure 5-10).

Fortunately, another diGly proteomics screen was performed by Stefan Boeing in collaboration with another laboratory using cisplatin to induce DNA damage, which stalls RNAPII and activates TC-NER (Damsma et al., 2007). This screen surprisingly turned up many more CSB ubiquitylation events than the previous screen (24 vs 5, Table 5-1). It subsequently became clear in further analysis that the lack of control events in the analysis such as FAND2 and FANCI ubiquitylation meant that the cisplatin treatment was likely unsuccessful (Garner & Smogorzewska, 2011). By this time however, I had already established the CSB 19R cells. Fortunately, CSB 19R was deficient in ubiquitylation following UV irradiation necessitating further phenotypic characterisation. The reason for the large number of CSB ubiquitylation sites identified in the cisplatin proteomics screen then was likely due to a combination of constitutive ubiquitylation sites, damage-induced events from infrequent engagement of CSB in repair of endogenous damage, and noise. Noise and false positive hits from proteome-wide screens is inevitable and common, but is usually overcome through enrichment with a treatment arm, but due to the comparison of fundamentally two untreated conditions these events would be over represented.

6.3 Cell system generation

Using Cockayne syndrome patient cell lines to study CSB has historically been widely employed by the scientific community for deduction of mechanisms of TC-NER (Citterio et al., 1998; A. R. Lehmann, 1982; Alan R. Lehmann et al., 1979;

Christine Troelstra et al., 1992). As such, a wealth of CS1AN-derived cell lines produced in the lab expressing CSB mutants useful as controls were immediately available to complement my research. Indeed, a ubiquitylation mutant of CSB defective in repair of oxidative lesions had successfully been characterised in this cell line in our lab (Ranes et al., 2016).

I had noticed during the screening process that some 8R CSB CS1AN cells I had generated expressed truncated forms of 8R CSB (Figure 4-2B). However, due to the large size of CSB (197 kDa with tags), small truncations wouldn't migrate differently from full length CSB to be able to discern if there were small truncations of an ostensibly full-lenth CSB. This might explain the strange phenotype seen of only partial RRS in WT CSB CS1AN cells 24 hours after UV, even though other fibroblast cell lines (MRC5) could fully restart transcription (Figure 4-3B). It might be that this was a mixed population of cells expressing some truncated CSB protein that is non-funcitonal but migrates with full-length CSB on a Western blot.

Commonly, a plasmid containing a transgene flanked by a promoter at the 5′ end and an antibiotic resistance gene at the 3′ end is transfected to generate cells stably expressing said gene. Presence of the gene product seen by western blot in antibiotic-selected cells can be sure to contain the full coding sequence as gene expression would specify the presence of a 5′ promoter and antibiotic resistance the presence of a 3′ antibiotic resistance gene. The system used to generate derivatives of CS1AN cell lines in this thesis, Tet-On 3G (Takarabio), uses a linear selection marker (puromycin resistance gene) on a separate DNA molecule to the plasmid containing the transgene. The manufacturer's stated reason for this is it reduces the chance that expression of the antibiotic resistance gene interferes with expression of the transgene. However, as the transgene isn't flanked at its 3′ end by an antibiotic resistant gene, puromycin resistant cells don't necessarily even indicate the presence of any transgene DNA, let alone the full CSB coding sequence. I attempted to amplify the integrated CSB gene from genomic DNA to try and ascertain if the full coding sequence was integrated, but this was unsuccessful. Hence, I decided to develop a cell system that could take advantage of a more reliable method of transgene genomic integration.

Flp recombinase-mediated genomic integration is highly specific and efficient (Buchholz et al., 1998; O'Gorman et al., 1991; Schlake & Bode, 1994; Senecoff et al., 1985). The plasmids for this system contain the antibiotic selection marker immediately downstream of the transgene, circumventing the problems with the linear selection marker. The plasmids also contain two recombination sites flanking the transgene that ensure integration of the full-length gene into the cell's genome. U2OS cells displayed very variable expression levels of transgene CSB and phenotypic results were also variable. The variable expression levels could be due to several FRT sites present in the genome leading to multiple integrations of CSB into differentially transcriptionally active regions. The HEK 293 cells were commercially produced and the presence of a single FRT site was confirmed by Southern blot, which likely explains the more consistent expression levels between the different derivative cell lines.

6.4 Functional characterisation of CSB ubiquitin mutants

6.4.1 Ubiquitylated CSB enrichment

Enrichment for ubiquitylated CSB was performed using GST-Dsk2 protein bound to beads in a pulldown analogous to immunoprecipitation. Dsk2 is optimal for enrichment of a wide variety of ubiquitin chains via its UBA domain which can bind monoubiquitin and polyubiquitin chains of linkages K48, 63, and the non-physiological 29/6 mixed chains which likely means it is non-selective toward chain linkage type (Hjerpe et al., 2009; Ohno et al., 2005; Raasi et al., 2005). Dsk2 enrichment then is a broad and sensitive technique to look for differences in ubiquitylation of proteins and indeed it was able to detect a significant decrease in lowly abundant CSB 19R ubiquitylation following UV (Figure 5-7A). Of course, this is a targeted strategy that doesn't look unbiasedly at multiple different PTMs that could be present on CSB and supressed by K>R mutations, such as SUMOylation and acetylation. Indeed, CSB has been shown to be sumoylated (Liebelt et al., 2020; Sin et al., 2016). However, these studies identify sumoylation of CSB at K32 and K205 of its N-terminus, which is in the opposite end of the protein relative to the K>R mutations used here. Sin et al., 2016, found that deletion of the last 30 amino acids of CSB ablated its sumoylation and rendered cells UV sensitive.

However, mutation of all the lysines in this region (K1457, K1487, and K1489, which were all mutated in CSB 19R) did not affect sumoylation of CSB or render cells UV sensitive. In Liebelt et al., 2020, the authors mutated K1359 and K1489 (both mutated in CSB 19R), but this didn't affect sumoylation. It is unlikely then that CSB 19R with all mutations in the C-terminus affects its sumoylation in the N-terminus. In any case, to test this directly an IP of CSB by GFP or FLAG and western blotting with a SUMO antibody would distinguish sumoylation levels between WT and 19R CSB.

Identification of ubiquitylated proteins after Dsk2 enrichment is done by Western blotting with an antibody against the protein of interest. While Dsk2 is specific for ubiquitin, it could conceivably pull down other PTM-proteins that are in complex with ubiquitylated ones. Additionally, ubiquitylated proteins can be simultaneously be modified in other ways such as sumoylated. Changes in higher-molecular weight species after Dsk2 enrichment might indicate changes in sumoylation rather than ubiquitylation. Another way to assess ubiquitylation of proteins is to transiently transfect His-tagged ubiquitin into cells and ubiquitylated proteins can then be enriched on Ni-NTA columns under denaturing conditions which will remove any interacting proteins and isolate only ubiquitylated species.

6.4.2 Measuring transcription restart defects

Short genes that are less likely to have damage and thus repaired quicker may give a false positive for RRS when the rest of the genome is still damaged and shutdown. On the other hand, measurements far from the promoter may be very sensitive to small reductions in repair capacity of cells. This is because of the increased number of lesions and how TC-NER proceeds, like transcription, unidirectionally from the promoter. Thus, RNAPII is released in waves from the promoter to 'scan' the genome resulting in sequential removal of lesions 5′ to 3′ (Lavigne et al., 2017; Liakos et al., 2020; Tufegdžić Vidaković et al., 2020). The dynamics of EXT1 RRS, which is a partial recovery at 15 hours, but fully restored at 24 hours, is slightly slower than with previous global transcription measurements (Mayne & Lehmann, 1982). This is likely a result of measurements taken some 290

kb from the promoter in conjunction with higher UV doses and thus higher damage loads. Ultimately, using the RT-qPCR assay to measure RRS showed CSB 19R transcription of EXT1 recovers slightly 24 hours after UV, but is mostly still shutdown, much like CSB and CSA KO cells.

The qPCR results merited further investigation of the RRS phenotype genome-wide. The use of 4SU-seq gives unparalleled resolution to examine nascent transcription over small time periods (Gregersen et al., 2020). While the 4SU-seq assay (Figure 5-11) confirmed that CSB 19R has an RRS defect after UV, it was not as pronounced as the defect of the EXT1 gene as measured by qPCR (Figure 5-10). However, stratification of genes by length showed that 24 hours after UV, CSB 19R cells had less transcription in >100 kb genes compared to CSB WT cells that declined with distance from the promoter. This is consistent with the qPCR data that shows a lack of recovery measured at 290 kb from the promoter. It's also consistent with the slightly better recovery of transcription of CSB 19R compared to CSB WT cells.

One of the difficulties with the 4SU-seq analysis was that there was a lot of variability in the basal transcription levels in untreated conditions, both between replicates and samples. I therefore decided to normalise in a similar way as to the qPCR analysis by taking a ratio of the later timepoints relative to untreated conditions. This nicely created uniform profiles between genotypes at 1-hour post-UV, which is expected at this time-point when damage is the main factor that constrains transcription to the 5′ of genes and CSB hasn't contributed much to repair. However, at 24 hours post-UV, transcription profiles between the two replicates of WT and CSB KO start to diverge, which is most obvious in >100 kb genes (Figure 5-11B). This could be due to the multiple enrichment and clean-up steps of biotinylated 4SU RNA that might contribute to inconsistency between replicates. The inclusion of more replicates would help to average out these inconsistencies.

6.4.3 Survival after UV

The 4SU-seq results for a mild RRS phenotype of CSB 19R are consistent with cell growth assays that show CSB 19R is sensitive to UV (Figure 5-8). At higher doses of UV, CSB KO cells show greater sensitivity than CSB 19R or CSA KO cells, which both show a similar sensitivity to UV with some cell growth post-UV but not to the extent of CSB WT cells. Survival of CSA KO cells, measured by clonogenic survival assays are variable, with some patient cell lines showing similar survival rates to CSB KO cells and others better survival (de Waard et al., 2004; Henning et al., 1995; van der Weegen et al., 2020). Results in the de Waard et al., 2004 study showed slightly different levels of survival to UV between CSA and CSB KO cells derived from mice depending on whether they were keratinocytes, ES or MEF cells, which suggests that the UV response might be different in different cell types. However, to the best of my knowledge, no studies have directly compared the UV sensitivity of CSA and CSB KO cells in isogenic cell lines. Further studies that directly measure cell viability such as ATP luciferase assays and apoptotic markers would be needed to confirm this result. But the extremely similar growth profiles of CSB 19R and CSA KO cells underscores the mechanistic connection of ubiquitylation of CSB by CSA in response to UV damage.

It has been shown that $CRL4A^{CSA}$ ubiquitylates CSB *in vitro* (Fischer et al., 2011) and it would be interesting to test if purified CSB 19R is unable to be ubiquitylated by CSA. This would further underscore the connection between CSB ubiquitylation by CSA.

6.4.4 Direct measurement of TC-NER

While RRS and UV-sensitivity have been phenotypes frequently assessed to clinically diagnose cockayne syndrome patients by cellular assays, they don't directly measure TC-NER. Thus, they serve as useful and tractable proxies. The exclusive measurement of repair of damage only in the transcribed strand is difficult unless one removes the GG-NER pathway completely. However, in normal cells (or, say CSB 19R cells), the vast majority of CPD removal and unscheduled DNA synthesis (gap-filling) is contributed by the GG-NER pathway, swamping any

contributions from TC-NER. However, to make firm conclusions about the effect of CSB mutants on repair of DNA damage, this repair must be measured directly. One well established method is to take advantage of the T4 endonuclease V that recognises and digests CPD-containing DNA. Using strand-specific oligonucleotide probes against a reporter gene in conjunction with T4 endonuclease V treatment and Southern blotting, one can measure repair of the transcribed strand *in vivo*. Damaged DNA will be digested by T4 endonuclease V, precluding probes binding, and the Southern blot will be blank. Over time, as repair preferentially occurs in the transcribed strand (due to faster kinetics of TC-NER) digestion will be inhibited and probes against the transcribed strand will start to be visible on the Southern blot while probes against the non-transcribed strand will not (as it is not repaired and still digested by T4endoV) (Gaillard et al., 2015).

There have since been sequencing methods that employ directional library preparations to be able to distinguish the transcribed and non-transcribed strands. The presence of CPDs precludes DNAs amplification by DNA polymerases resulting in damaged-DNA reads being sequenced. Over time, the increasing frequency of once absent sequences indicates their repair (Nakazawa et al., 2020). Assays such as these that directly measure TC-NER by CPD removal in the transcribed strand are ultimately necessary to be able to directly conclude that CSB 19R affects TC-NER and isn't responsible for UV-sensitivity and transcription defects via other mechanisms.

6.4.5 Recruitment to chromatin

CSB recruitment to chromatin upon DNA damage is a well-established phenomenon and is a consequence of an increase in the abundance of its substrate, lesion-stalled RNAPII. However, CSB recruitment to chromatin is not inhibited in CSA KO cells, indicating that CSA acts downstream of the initial recognition of stalled RNAPII (Figure 5-9C) (Fousteri et al., 2006; van der Weegen et al., 2020). CSB 19R was recruited to damaged chromatin at WT levels (Figure 5-9A), indicating that its ubiquitylation is not necessary for initial recognition and further supporting the idea that its ubiquitylation by CSA is a downstream signalling

event that further licenses repair. This further supports the idea that CSB 19R is deficient in CSA-dependent ubiquitylation. Assays that use buffers for subcellular fractionation are prone to contamination between fractions. It is known from immunofluorescence studies that CSB is mostly nucleolar under normal conditions (Iyama et al., 2018; Iyama & Wilson, 2016). The appearance of CSB in the soluble fraction in Figure 5-9 indicates the presence of cytoplasm and/or nucleoplasm in this fraction and the stabilisation of CSB with RNAPII after UV results in its increased abundance in the (insoluble) chromatin fraction, which is similar to other studies (Lake et al., 2010; Van Den Boom et al., 2004).

6.5 Other CSB ubiquitylation mutants

The mutation of 19 lysines in one protein is quite an extreme mutational load. However, the mutation of lysine to arginine is conservative and likely has minimal impact on structure. Nevertheless, to draw conclusions specifically about CSB's ubiquitylation it must be proven that CSB 19R retains its catalytic activity. In support of this, CSB 19R is able to associate with UV-damaged chromatin (Figure 5-9A) but ATPase dead mutants of CSB (K538R) cannot (see Lake et al., 2010). Attempts were made to purify CSB 19R from insect cells for *in vitro* ATPase assays, but the yields were low and subsequent analysis for ATPase activity was inconclusive, but this remains a priority. To investigate CSB ubiquitylation without the abundance of mutations that could impose other structural constraints affecting CSB activity besides ubiquitylation, a different construct was made (Figure 5-12). This involved the fusion of the USP2 DUB catalytic domain to the C-terminal of CSB (CSB-USP2 WT) that constitutively deubiquitylates it, while a catalytic dead USP2 fusion (CSB-USP2 C276A) did not (Figure 5-13A). Recapitulating the CSB 19R phenotypes with CSB-USP2 WT cells together with a WT phenotype of CSB-USP2 C276A would provide evidence against the hypothesis that CSB 19R phenotype could be due to other structural consequences of the numerous mutations. It could be that (de)ubiquitylation of CSB affects its ATPase activity, but this is hard to prove in the CSB 19R mutant where a lack of ATPase activity could be due to gross structural changes from a large number of mutations. Testing the ATPase activity of purified CSB-USP2 WT protein could help to untangle this. If it too had deficiencies in ATPase activity this would bolster the hypothesis that ubiquitylation is necessary for CSB's ATPase activity.

However, the presence of an artificial DUB within the TC-NER 'repairsome' could deubiquitylate other binding partners of CSB, confounding results. Indeed, RNAPII ubiquitylation is slightly reduced after UV in CSB-USP2 WT cells compared to CSB-USP2 C276A cells (Figure 5-13B). However, this is also the case in CSA KO cells even though CSA is not thought to be the UV-dependent E3 ubiquitin ligase of RNAPII (Anindya et al., 2007; Tufegdžić Vidaković et al., 2020). It could be that RNAPII ubiquitylation is a downstream step of CSB ubiquitylation and recruitment.

In which case, reduction of RNAPII ubiquitylation in CSB ubiquitylation mutants would be a normal phenotype. Untangling these phenomena will be difficult but is vital for a clear understanding of the fate of RNAPII at lesions. Further experiments to characterise the phenotypes of CSB-USP2 WT and CSB-USP2 CA using the qPCR and cell growth assays in response to UV is vital for confirming the phenotype of CSB 19R is due to the defect of its ubiquitylation. Also performing Dsk2 pulldowns in CSB-USP2 cells to assess RNAPII ubiquitylation will further strengthen the hypothesis that CSB ubiquitylation is important for TC-NER and the response to UV-induced DNA damage.

CSB degradation in response to UV was touted as a response to CSA-induced ubiquitylation (Groisman et al., 2006). However, these results have not been able to replicated by several members of our laboratory, or in the literature, and other studies have found CSB not to be degraded in response to UV (Lake et al., 2010; Liebelt et al., 2020). In Lake et al., 2010, chromatin fractionation of cells after UV showed an initial recruitment of CSB to the chromatin fraction that disappeared over time, but this cooccurred with a reappearance of CSB in the soluble fraction, suggesting its localisation is altered but not degraded. It's likely that in the Groisman study, chromatin was not fully solubilised and the disappearance of CSB from the lysates reflected its relocation from the soluble fraction to chromatin. My attempts at looking at CSB 19R degradation in whole cell extracts yielded variable results between the different clones that necessitates further investigation.

6.6 Cockayne syndrome proteins in general transcription

The role of CSB in transcription is well established (Balajee et al., 1997; Dianov et al., 1997; Newman et al., 2006; Proietti-De-Santis et al., 2006; Christopher P. Selby & Sancar, 1997b; Van Den Boom et al., 2004; Y. Wang et al., 2014, 2016). It is therefore not surprising that Cockayne syndrome patients also display neurodegenerative disease, which is commonly the result of transcription defects (Liu et al., 2017). The role of CSA and CSB ubiquitylation in this process is less clear. While, studies from extracts of CSA patient cell lines show they support reduced transcription *in vitro* (Dianov et al., 1997), no transcriptomics analysis of

CSA KO cells has been done like it has for CSB KO cells (Kyng et al., 2003; Newman et al., 2006; Ranes et al., 2016; Y. Wang et al., 2014). My isogenic cell line system was ideally suited for the this and I performed a total RNA-seq experiment to try and characterise the gene expression profiles of CSB WT, CSB 19R, CSB KO and CSA KO cells in the absence of any induced DNA damage. After inducing expression of transgenes with doxycycline for 48 hours, I isolated RNA for deep sequencing. Although there were great differences in gene expression between cells, unfortunately this was also true for the two different clonal populations of the same genotype. Hence, it was difficult to interpret the results and find any conclusive differences between the different genotypes. Due to the UV induced ubiquitylation of CSB, it could be that CSB 19R has normal transcription properties in the absence of DNA damage. However, the transcriptomic profiles of isogenic CSB KO and CSB WT cells was not any more different than between the two CSB WT clonal populations. This is surprising given results in other studies. Had I sequenced only one clone of each genotype; I may have been led to draw false conclusions from the differences between them. Given that these cells are isogenic, this should have limited variability between the clones. As all cell lines are cultured with doxycycline-free media they essentially exist as CSB KO cells until transgene induction. It may be that 48 hours was not enough to establish stable gene expression profiles as a result of CSB expression. This is supported by the fact that CSA KO cells, which have no dox-inducible transgene and are genetically different from all the other cell lines, were the most obviously different and clustered separately in principal component analysis. Repeating this experiment with induction of CSB transgenes for much longer timepoints could lead to a more definitive dataset where the clonal cells have similar gene expression profiles. Gene expression profiles of CSB-USP2 cell lines should be similar to CSB 19R cell lines and should be included to further corroborate the hypothesis that CSB ubiquitylation, or lack of it, is affecting gene expression under normal conditions.

6.7 Cockayne syndrome

How the Cockayne syndrome proteins cause the complex, severe and heterogenous phenotypes of the disease is much debated. Originally thought to be simply a consequence of a lack of the TC-NER pathway, this is not sufficient to explain the many phenotypes. Indeed, Xeroderma pigmentosum patients who are deficient in GG-NER repair of the whole genome and subsequently have higher unrepaired damage and mutational loads experience different, and in a lot of cases, milder symptoms. One of the striking differences is the lack of neurodegeneration in the majority of XP patients that is a hallmark of CS. Interestingly, some cases of XP-CS exist where patients display symptoms of both diseases. There a very few cases of XP-CS, but they are caused by mutations in either XPG, XPD, XPB, or XPF (Natale & Raquer, 2017). These are proteins that are involved in both NER and transcription with XPB and XPD serving as the helicase subunits of TFIIH. While XPG and XPF are endonucleases, XPG is known to interact with TFIIH during transcription and XPF is involved in chromatin looping via CTCF (Georges et al., 2019; Ito et al., 2007; Le May et al., 2012). Their absence during NER may also stabilise and sequester TFIIH at DNA damage that can't be repaired, hindering its participation in basal transcription. XPA cells do not display the severe RRS phenotype of CS cells, despite being ultimately deficient in both GG-NER and TC-NER (Vélez-Cruz et al., 2013). This points to a role for the Cockayne syndrome proteins in restarting transcription separate from repair of DNA damage. In light of the results from Tufegdžić Vidaković et al., 2020 that showed a non-degradable form of RNAPII in a CSB KO background permits restart of transcription, it points to a role for CSB in protecting RNAPII from massive degradation following UV. This would explain the milder phenotypes of XP patients who usually don't suffer from neurodegeneration, progeria, and stunted growth. Clearly, Cockayne syndrome is a disease not just of defective DNA repair, but also defective transcription.

By controlling RNAPII levels artificially, for example by using a quick-acting degron system, one might be able to recapitulate the phenotypes of Cockayne syndrome cells in the absence of DNA damage. If RNAPII levels can be degraded quickly in the absence of DNA damage, this would lead to transcription shutdown. If CSB's

role is to recover RNAPII levels then in cells with CSB, RNAPII levels should recover and transcription should resume. However cells with low levels of RNAPII and without CSB might never recover normal RNAPII levels and transcription won't restart desite the absence of DNA damage. Recapitulating the gene expression changes after UV in CSB KO cells might also be achieved by maintaining reduced levels of RNAPII. This would also strengthen the hypothesis that CSB's role is to maintain the free pool of RNAPII and protect it from excessive degredation from which it may not recover after UV.

The results in this thesis go some way to untangling the opaque molecular mechanism of CSB and how it's loss of function is responsible for such a complex and devastating disease. I hope these results can inspire hypotheses and experiments that will contribute to demystifying these mechanisms further.

Reference List

Aboussekhra, A., Biggerstaff, M., Shivji, M. K. K., Vilpo, J. A., Moncollin, V., Podust, V. N., Protić, M., Hübscher, U., Egly, J. M., & Wood, R. D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell*, *80*(6), 859–868. https://doi.org/10.1016/0092-8674(95)90289-9

Adebali, O., Chiou, Y. Y., Hu, J., Sancar, A., & Selby, C. P. (2017). Genome-wide transcription-coupled repair in Escherichia coli is mediated by the Mfd translocase. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(11), E2116–E2125. https://doi.org/10.1073/pnas.1700230114

Adebali, O., Sancar, A., & Selby, C. P. (2017). Mfd translocase is necessary and sufficient for Transcription-coupled repair in Escherichia coli. *Journal of Biological Chemistry*, *292*(45), 18386–18391. https://doi.org/10.1074/jbc.C117.818807

Ahel, I., Rass, U., El-Khamisy, S. F., Katyal, S., Clements, P. M., McKinnon, P. J., Caldecott, K. W., & West, S. C. (2006). The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates. *Nature*, *443*(7112), 713–716. https://doi.org/10.1038/nature05164

Akutsu, M., Dikic, I., & Bremm, A. (2016). Ubiquitin chain diversity at a glance. *Journal of Cell Science*, *129*(5), 875–880. https://doi.org/10.1242/jcs.183954

Anindya, R., Aygün, O., & Svejstrup, J. Q. (2007). Damage-Induced Ubiquitylation of Human RNA Polymerase II by the Ubiquitin Ligase Nedd4, but Not Cockayne Syndrome Proteins or BRCA1. *Molecular Cell*, *28*(3), 386–397. https://doi.org/10.1016/j.molcel.2007.10.008

Anindya, R., Mari, P. O., Kristensen, U., Kool, H., Giglia-Mari, G., Mullenders, L. H., Fousteri, M., Vermeulen, W., Egly, J. M., & Svejstrup, J. Q. (2010). A Ubiquitin-Binding Domain in Cockayne Syndrome B Required for Transcription-Coupled Nucleotide Excision Repair. *Molecular Cell*, *38*(5), 637–648. https://doi.org/10.1016/j.molcel.2010.04.017

Araki, M., Masutani, C., Takemura, M., Uchida, A., Sugasawa, K., Kondoh, J., Ohkuma, Y., & Hanaoka, F. (2001). Centrosome Protein Centrin 2/Caltractin 1 Is Part of the Xeroderma Pigmentosum Group C Complex That Initiates Global Genome Nucleotide Excision Repair. *Journal of Biological Chemistry*, *276*(22), 18665–18672. https://doi.org/10.1074/jbc.M100855200

Araújo, S. J., Tirode, F., Coin, F., Pospiech, H., Syväoja, J. E., Stucki, M., Hübscher, U., Egly, J. M., & Wood, R. D. (2000). Nucleotide excision repair of DNA with recombinant human proteins: Definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. I: *Genes and Development* (Vol. 14, Number 3). Cold Spring Harbor Laboratory Press. https://doi.org/10.5167/uzh-845

Bakkenist, C. J., & Kastan, M. B. (2003). DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature*, *421*(6922), 499–506. https://doi.org/10.1038/nature01368

Balajee, A. S., May, A., Dianov, G. L., Friedberg, E. C., & Bohr, V. A. (1997). Reduced RNA polymerase II transcription in intact and permeabilized Cockayne syndrome group B cells. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(9), 4306–4311. https://doi.org/10.1073/pnas.94.9.4306

Bar-Nahum, G., Epshtein, V., Ruckenstein, A. E., Rafikov, R., Mustaev, A., & Nudler, E. (2005). A ratchet mechanism of transcription elongation and its control. *Cell*, *120*(2), 183–193. https://doi.org/10.1016/j.cell.2004.11.045

Basar, M. A., Beck, D. B., & Werner, A. (2021). Deubiquitylases in developmental ubiquitin signaling and congenital diseases. *Cell Death and Differentiation*, *28*(2), 538–556. https://doi.org/10.1038/s41418-020-00697-5

Beard, W. A., Horton, J. K., Prasad, R., & Wilson, S. H. (2019). Eukaryotic base excision repair: New approaches shine light on mechanism. *Annual Review of Biochemistry*, *88*(1), 137–162. https://doi.org/10.1146/annurev-biochem-013118-111315

Bergink, S., Toussaint, W., Luijsterburg, M. S., Dinant, C., Alekseev, S., Hoeijmakers, J. H. J., Dantuma, N. P., Houtsmuller, A. B., & Vermeulen, W. (2012). Recognition of DNA damage by XPC coincides with disruption of the XPC-RAD23 complex. *Journal of Cell Biology*, *196*(6), 681–688. https://doi.org/10.1083/jcb.201107050

Bernardes de Jesus, B. M., Bjørås, M., Coin, F., & Egly, J. M. (2008). Dissection of the Molecular Defects Caused by Pathogenic Mutations in the DNA Repair Factor XPC. *Molecular and Cellular Biology*, *28*(23), 7225–7235. https://doi.org/10.1128/mcb.00781-08

Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J. M., & Cramer, P. (2016). Structure of transcribing mammalian RNA polymerase II. *Nature*, *529*(7587), 551–554. https://doi.org/10.1038/nature16482

Bernecky, C., Plitzko, J. M., & Cramer, P. (2017). Structure of a transcribing RNA polymerase II-DSIF complex reveals a multidentate DNA-RNA clamp. *Nature Structural and Molecular Biology*, *24*(10), 809–815. https://doi.org/10.1038/nsmb.3465

Besaratinia, A., Yoon, J., Schroeder, C., Bradforth, S. E., Cockburn, M., & Pfeifer, G. P. (2011). Wavelength dependence of ultraviolet radiation-induced DNA damage as determined by laser irradiation suggests that cyclobutane pyrimidine dimers are the principal DNA lesions produced by terrestrial sunlight. *The FASEB Journal*, *25*(9), 3079–3091. https://doi.org/10.1096/fj.11-187336

Bétermier, M., Bertrand, P., & Lopez, B. S. (2014). Is Non-Homologous End-Joining Really an Inherently Error-Prone Process? *PLoS Genetics*, *10*(1), e1004086. https://doi.org/10.1371/journal.pgen.1004086

Bhagwat, A. S., Hao, W., Townes, J. P., Lee, H., Tang, H., & Foster, P. L. (2016). Strand-biased cytosine deamination at the replication fork causes cytosine to thymine mutations in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(8), 2176–2181. https://doi.org/10.1073/pnas.1522325113

Bienko, M., Green, C. M., Sabbioneda, S., Crosetto, N., Matic, I., Hibbert, R. G., Begovic, T., Niimi, A., Mann, M., Lehmann, A. R., & Dikic, I. (2010). Regulation of Translesion Synthesis DNA Polymerase η by Monoubiquitination. *Molecular Cell*, *37*(3), 396–407. https://doi.org/10.1016/j.molcel.2009.12.039

Blackford, A. N., & Jackson, S. P. (2017). ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Molecular Cell*, *66*(6), 801–817. https://doi.org/10.1016/j.molcel.2017.05.015

Blondel, M., Galan, J. M., Chi, Y., Lafourcade, C., Longaretti, C., Deshaies, R. J., & Peter, M. (2000). Nuclear-specific degradation of Far1 is controlled by the localization of the F-box protein Cdc4. *EMBO Journal*, *19*(22), 6085–6097. https://doi.org/10.1093/emboj/19.22.6085

Bockrath, R. C., & Palmer, J. E. (1977). Differential repair of premutational UV-lesions at tRNA genes in E. coli. *MGG Molecular & General Genetics*, *156*(2), 133–140. https://doi.org/10.1007/BF00283485

Boeing, S., Williamson, L., Encheva, V., Gori, I., Saunders, R. E., Instrell, R., Aygün, O., Rodriguez-Martinez, M., Weems, J. C., Kelly, G. P., Conaway, J. W., Conaway, R. C., Stewart, A., Howell, M., Snijders, A. P., & Svejstrup, J. Q. (2016). Multiomic Analysis of the UV-Induced DNA Damage Response. *Cell Reports*, *15*(7), 1597–1610. https://doi.org/10.1016/j.celrep.2016.04.047

Bohr, V. A., Smith, C. A., Okumoto, D. S., & Hanawalt, P. C. (1985). DNA repair in an active gene: Removal of pyrimidine dimers from the DHFR gene of CHO cells is

much more efficient than in the genome overall. *Cell*, *40*(2), 359–369. https://doi.org/10.1016/0092-8674(85)90150-3

Bowden, N. A., Beveridge, N. J., Ashton, K. A., Baines, K. J., & Scott, R. J. (2015). Understanding xeroderma pigmentosum complementation groups using gene expression profiling after UV-light exposure. *International Journal of Molecular Sciences*, *16*(7), 15985–15996. https://doi.org/10.3390/ijms160715985

Bradford, P. T., Goldstein, A. M., Tamura, D., Khan, S. G., Ueda, T., Boyle, J., Oh, K. S., Imoto, K., Inui, H., Moriwaki, S. I., Emmert, S., Pike, K. M., Raziuddin, A., Plona, T. M., DiGiovanna, J. J., Tucker, M. A., & Kraemer, K. H. (2011). Cancer and neurologic degeneration in xeroderma pigmentosum: Long term follow-up characterises the role of DNA repair. *Journal of Medical Genetics*, *48*(3), 168–176. https://doi.org/10.1136/jmg.2010.083022

Bregman, D. B., Halaban, R., Van Gool, A. J., Henning, K. A., Friedberg, E. C., & Warren, S. L. (1996). UV-induced ubiquitination of RNA polymerase II: A novel modification deficient in Cockayne syndrome cells. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(21), 11586–11590. https://doi.org/10.1073/pnas.93.21.11586

Breiling, A., & Lyko, F. (2015). Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. I: *Epigenetics and Chromatin* (Vol. 8, Number 1, p. 24). BioMed Central Ltd. https://doi.org/10.1186/s13072-015-0016-6

Brem, R., Guven, M., & Karran, P. (2017). Oxidatively-generated damage to DNA and proteins mediated by photosensitized UVA. I: *Free Radical Biology and Medicine* (Vol. 107, pp. 101–109). Elsevier Inc. https://doi.org/10.1016/j.freeradbiomed.2016.10.488

Brinkman, E. K., Chen, T., Amendola, M., & Van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Research*, *42*(22), e168–e168. https://doi.org/10.1093/nar/gku936

Britton, S., Coates, J., & Jackson, S. P. (2013). A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. *Journal of Cell Biology*, *202*(3), 579–595. https://doi.org/10.1083/jcb.201303073

Brueckner, F., Hennecke, U., Carell, T., & Cramer, P. (2007). CPD damage recognition by transcribing RNA polymerase II. *Science*, *315*(5813), 859–862. https://doi.org/10.1126/science.1135400

Buchholz, F., Angrand, P. O., & Stewart, A. F. (1998). Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nature Biotechnology*, *16*(7), 657–662. https://doi.org/10.1038/nbt0798-657

Burgos-Barragan, G., Wit, N., Meiser, J., Dingler, F. A., Pietzke, M., Mulderrig, L., Pontel, L. B., Rosado, I. V., Brewer, T. F., Cordell, R. L., Monks, P. S., Chang, C. J., Vazquez, A., & Patel, K. J. (2017). Mammals divert endogenous genotoxic formaldehyde into one-carbon metabolism. *Nature*, *548*(7669), 549–554. https://doi.org/10.1038/nature23481

Buterin, T., Hess, M. T., Gunz, D., Geacintov, N. E., Mullenders, L. H., & Naegeli, H. (2002). Trapping of DNA nucleotide excision repair factors by nonrepairable carcinogen adducts. I: *Cancer Research* (Vol. 62, Number 15). https://doi.org/10.5167/uzh-14526

Byun, T. S., Pacek, M., Yee, M. C., Walter, J. C., & Cimprich, K. A. (2005). Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes and Development*, *19*(9), 1040–1052. https://doi.org/10.1101/gad.1301205

Caldecott, K. W. (2008). Single-strand break repair and genetic disease. I: *Nature Reviews Genetics* (Vol. 9, Number 8, pp. 619–631). Nature Publishing Group. https://doi.org/10.1038/nrg2380

Camenisch, U., Dip, R., Schumacher, S. B., Schuler, B., & Naegeli, H. (2006).

Recognition of helical kinks by xeroderma pigmentosum group a protein triggers DNA excision repair. *Nature Structural and Molecular Biology*, *13*(3), 278–284. https://doi.org/10.1038/nsmb1061

Camenisch, U., Träutlein, D., Clement, F. C., Fei, J., Leitenstorfer, A., Ferrando-May, E., & Naegeli, H. (2009). Two-stage dynamic DNA quality check by xeroderma pigmentosum group C protein. *EMBO Journal*, *28*(16), 2387–2399. https://doi.org/10.1038/emboj.2009.187

Cavadini, S., Fischer, E. S., Bunker, R. D., Potenza, A., Lingaraju, G. M., Goldie, K. N., Mohamed, W. I., Faty, M., Petzold, G., Beckwith, R. E. J., Tichkule, R. B., Hassiepen, U., Abdulrahman, W., Pantelic, R. S., Matsumoto, S., Sugasawa, K., Stahlberg, H., & Thomä, N. H. (2016). Cullin-RING ubiquitin E3 ligase regulation by the COP9 signalosome. *Nature*, *531*(7596), 598–603. https://doi.org/10.1038/nature17416

Chapman, J. R., Taylor, M. R. G., & Boulton, S. J. (2012). Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. I: *Molecular Cell* (Vol. 47, Number 4, pp. 497–510). Mol Cell. https://doi.org/10.1016/j.molcel.2012.07.029

Chatterjee, N., & Walker, G. C. (2017). Mechanisms of DNA damage, repair, and mutagenesis. I: *Environmental and Molecular Mutagenesis* (Vol. 58, Number 5, pp. 235–263). John Wiley and Sons Inc. https://doi.org/10.1002/em.22087

Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K., & Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science*, *243*(4898), 1576–1583. https://doi.org/10.1126/science.2538923

Chen, X., Velmurugu, Y., Zheng, G., Park, B., Shim, Y., Kim, Y., Liu, L., Van Houten, B., He, C., Ansari, A., & Min, J. H. (2015). Kinetic gating mechanism of DNA damage recognition by Rad4/XPC. *Nature Communications*, *6*(1), 5849. https://doi.org/10.1038/ncomms6849

Chi, Y., Huddleston, M. J., Zhang, X., Young, R. A., Annan, R. S., Carr, S. A., & Deshaies, R. J. (2001). Negative regulation of Gcn4 and Msn2 transcription factors by Srb10 cyclin-dependent kinase. *Genes and Development*, *15*(9), 1078–1092. https://doi.org/10.1101/gad.867501

Cho, I., Tsai, P. F., Lake, R. J., Basheer, A., & Fan, H. Y. (2013). ATP-Dependent Chromatin Remodeling by Cockayne Syndrome Protein B and NAP1-Like Histone Chaperones Is Required for Efficient Transcription-Coupled DNA Repair. *PLoS Genetics*, *9*(4), e1003407. https://doi.org/10.1371/journal.pgen.1003407

Christians, F. C., & Hanawalt, P. C. (1992). Inhibition of transcription and strand-specific DNA repair by α-amanitin in Chinese hamster ovary cells. *Mutation Research-DNA Repair*, *274*(2), 93–101. https://doi.org/10.1016/0921-8777(92)90056-9

Christiansen, M., Thorslund, T., Jochimsen, B., Bohr, V. A., & Stevnsner, T. (2005). The Cockayne syndrome group B protein is a functional dimer. *FEBS Journal*, *272*(17), 4306–4314. https://doi.org/10.1111/j.1742-4658.2005.04844.x

Chu, G., & Chang, E. (1988). Xeroderma pigmentosum group E cells lack a nuclear factor that binds to damaged DNA. *Science*, *242*(4878), 564–567. https://doi.org/10.1126/science.3175673

Ciechanover, A., Heller, H., Elias, S., Haas, A. L., & Hershko, A. (1980). ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proceedings of the National Academy of Sciences of the United States of America*, *77*(3), 1365–1368. https://doi.org/10.1073/pnas.77.3.1365

Ciechanover, Aaron. (2004). *Aaron Ciechanover – Nobel Lecture. NobelPrize.org. Nobel Media AB 2020. Mon. 14 Dec 2020. <https://www.nobelprize.org/prizes/chemistry/2004/ciechanover/lecture/>*.

https://www.nobelprize.org/prizes/chemistry/2004/ciechanover/lecture/
Ciechanover, Aaron, Finley, D., & Varshavsky, A. (1984). Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. *Cell*, *37*(1), 57–66. https://doi.org/10.1016/0092-8674(84)90300-3
Citterio, E., Rademakers, S., Van Der Horst, G. T. J., Van Gool, A. J., Hoeijmakers, J. H. J., & Vermeulen, W. (1998). Biochemical and biological characterization of wild-type and ATPase- deficient cockayne syndrome B repair protein. *Journal of Biological Chemistry*, *273*(19), 11844–11851. https://doi.org/10.1074/jbc.273.19.11844
Citterio, E., Van Den Boom, V., Schnitzler, G., Kanaar, R., Bonte, E., Kingston, R. E., Hoeijmakers, J. H. J., & Vermeulen, W. (2000). ATP-Dependent Chromatin Remodeling by the Cockayne Syndrome B DNA Repair-Transcription-Coupling Factor. *Molecular and Cellular Biology*, *20*(20), 7643–7653. https://doi.org/10.1128/mcb.20.20.7643-7653.2000
Clapier, C. R., Iwasa, J., Cairns, B. R., & Peterson, C. L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. I: *Nature Reviews Molecular Cell Biology* (Vol. 18, Number 7, pp. 407–422). Nature Publishing Group. https://doi.org/10.1038/nrm.2017.26
Cleaver, J. E. (1968). Defective repair replication of DNA in xeroderma pigmentosum. *Nature*, *218*(5142), 652–656. https://doi.org/10.1038/218652a0
Cockayne, E. A. (1936). Dwarfism with retinal atrophy and deafness. *Archives of Disease in Childhood*, *11*(61), 1–8. https://doi.org/10.1136/adc.11.61.1
Coin, F., Marinoni, J. C., Rodolfo, C., Fribourg, S., Pedrini, A. M., & Egly, J. M. (1998). Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nature Genetics*, *20*(2), 184–188. https://doi.org/10.1038/2491
Coin, F., Oksenych, V., & Egly, J. M. (2007). Distinct Roles for the XPB/p52 and XPD/p44 Subcomplexes of TFIIH in Damaged DNA Opening during Nucleotide Excision Repair. *Molecular Cell*, *26*(2), 245–256. https://doi.org/10.1016/j.molcel.2007.03.009
Conaway, R. C., Brower, C. S., & Conaway, J. W. (2002). Emerging roles of ubiquitin in transcription regulation. I: *Science* (Vol. 296, Number 5571, pp. 1254–1258). American Association for the Advancement of Science. https://doi.org/10.1126/science.1067466
Constantinou, A., Gunz, D., Evans, E., Lalle, P., Bates, P. A., Wood, R. D., & Clarkson, S. G. (1999). Conserved residues of human XPG protein important for nuclease activity and function in nucleotide excision repair. *Journal of Biological Chemistry*, *274*(9), 5637–5648. https://doi.org/10.1074/jbc.274.9.5637
Cooke, M. S., Evans, M. D., Dizdaroglu, M., & Lunec, J. (2003). Oxidative DNA damage: mechanisms, mutation, and disease. *The FASEB Journal*, *17*(10), 1195–1214. https://doi.org/10.1096/fj.02-0752rev
Cope, G. A., Suh, G. S. B., Aravind, L., Schwarz, S. E., Zipursky, S. L., Koonin, E. V., & Deshaies, R. J. (2002). Role of predicted metalloprotease motif of Jab1/Csn5 in cleavage of Nedd8 from Cul1. *Science*, *298*(5593), 608–611. https://doi.org/10.1126/science.1075901
Crickard, J. B., Fu, J., & Reese, J. C. (2016). Biochemical analysis of yeast suppressor of Ty 4/5 (Spt4/5) reveals the importance of nucleic acid interactions in the prevention of RNA polymerase II arrest. *Journal of Biological Chemistry*, *291*(19), 9853–9870. https://doi.org/10.1074/jbc.M116.716001
Damsma, G. E., Alt, A., Brueckner, F., Carell, T., & Cramer, P. (2007). Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nature Structural and Molecular Biology*, *14*(12), 1127–1133. https://doi.org/10.1038/nsmb1314
Danielsen, J. M. R., Sylvestersen, K. B., Bekker-Jensen, S., Szklarczyk, D., Poulsen, J.

W., Horn, H., Jensen, L. J., Mailand, N., & Nielsen, M. L. (2011). Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level. *Molecular and Cellular Proteomics*, *10*(3), 1–12. https://doi.org/10.1074/mcp.M110.003590

Davidson, E. A. (2006). Dna Repair and Mutagenesis, 2Nd Edition. I: *Shock* (Vol. 26, Number 2). https://doi.org/10.1097/01.shk.0000232588.61871.ff

Davies, R. J. H. (1995). Ultraviolet radiation damage in DNA. *Biochemical Society Transactions*, *23*(2), 407–418. https://doi.org/10.1042/bst0230407

De Bont, R., & van Larebeke, N. (2004). Endogenous DNA damage in humans: A review of quantitative data. I: *Mutagenesis* (Vol. 19, Number 3, pp. 169–185). Oxford Academic. https://doi.org/10.1093/mutage/geh025

De Laat, W. L., Appeldoorn, E., Jaspers, N. G. J., & Hoeijmakers, J. H. J. (1998). DNA structural elements required for ERCC1-XPF endonuclease activity. *Journal of Biological Chemistry*, *273*(14), 7835–7842. https://doi.org/10.1074/jbc.273.14.7835

De Laat, W. L., Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N. G. J., & Hoeijmakers, J. H. J. (1998). DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes and Development*, *12*(16), 2598–2609. https://doi.org/10.1101/gad.12.16.2598

de Waard, H., de Wit, J., Andressoo, J.-O., van Oostrom, C. T. M., Riis, B., Weimann, A., Poulsen, H. E., van Steeg, H., Hoeijmakers, J. H. J., & van der Horst, G. T. J. (2004). Different Effects of CSA and CSB Deficiency on Sensitivity to Oxidative DNA Damage. *Molecular and Cellular Biology*, *24*(18), 7941–7948. https://doi.org/10.1128/mcb.24.18.7941-7948.2004

Demple, B., & Harrison, L. (1994). Repair of oxidative damage to DNA: Enzymology and biology. I: *Annual Review of Biochemistry* (Vol. 63, pp. 915–948). Annual Reviews Inc. https://doi.org/10.1146/annurev.bi.63.070194.004411

Dianov, G. L., Houle, J. F., Iyer, N., Bohr, V. A., & Friedberg, E. C. (1997). Reduced RNA polymerase II transcription in extracts of Cockayne syndrome and xeroderma pigmentosum/Cockayne syndrome cells. *Nucleic Acids Research*, *25*(18), 3636–3642. https://doi.org/10.1093/nar/25.18.3636

Dianov, G. L., & Hübscher, U. (2013). Mammalian base excision repair: The forgotten archangel. I: *Nucleic Acids Research* (Vol. 41, Number 6, pp. 3483–3490). Oxford Academic. https://doi.org/10.1093/nar/gkt076

Ding, B., LeJeune, D., & Li, S. (2010). The C-terminal repeat domain of Spt5 plays an important role in suppression of Rad26-independent transcription coupled repair. *Journal of Biological Chemistry*, *285*(8), 5317–5326. https://doi.org/10.1074/jbc.M109.082818

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Doil, C., Mailand, N., Bekker-Jensen, S., Menard, P., Larsen, D. H., Pepperkok, R., Ellenberg, J., Panier, S., Durocher, D., Bartek, J., Lukas, J., & Lukas, C. (2009). RNF168 Binds and Amplifies Ubiquitin Conjugates on Damaged Chromosomes to Allow Accumulation of Repair Proteins. *Cell*, *136*(3), 435–446. https://doi.org/10.1016/j.cell.2008.12.041

Dou, H., Buetow, L., Sibbet, G. J., Cameron, K., & Huang, D. T. (2012). BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. *Nature Structural and Molecular Biology*, *19*(9), 876–883. https://doi.org/10.1038/nsmb.2379

Dove, K. K., & Klevit, R. E. (2017). RING-Between-RING E3 Ligases: Emerging Themes amid the Variations. I: *Journal of Molecular Biology* (Vol. 429, Number 22, pp. 3363–3375). Academic Press. https://doi.org/10.1016/j.jmb.2017.08.008

Duda, D. M., Borg, L. A., Scott, D. C., Hunt, H. W., Hammel, M., & Schulman, B. A. (2008). Structural Insights into NEDD8 Activation of Cullin-RING Ligases: Conformational Control of Conjugation. *Cell*, *134*(6), 995–1006. https://doi.org/10.1016/j.cell.2008.07.022

Dulbecco, R. (1949). Reactivation of ultra-violet-inactivated bacteriophage by visible light [2]. I: *Nature* (Vol. 163, Number 4155, pp. 949–950). Nature. https://doi.org/10.1038/163949b0

Dutta, A., Gogol, M., Kim, J. H., Smolle, M., Venkatesh, S., Gilmore, J., Florens, L., Washburn, M. P., & Workman, J. L. (2014). Swi/Snf dynamics on stress-responsive genes is governed by competitive bromodomain interactions. *Genes and Development*, *28*(20), 2314–2330. https://doi.org/10.1101/gad.243584.114

Ehara, H., Yokoyama, T., Shigematsu, H., Yokoyama, S., Shirouzu, M., & Sekine, S. I. (2017). Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science*, *357*(6354), 921–924. https://doi.org/10.1126/science.aan8552

Elia, A. E. H., Boardman, A. P., Wang, D. C., Huttlin, E. L., Everley, R. A., Dephoure, N., Zhou, C., Koren, I., Gygi, S. P., & Elledge, S. J. (2015). Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response. *Molecular Cell*, *59*(5), 867–881. https://doi.org/10.1016/j.molcel.2015.05.006

Elsasser, S., Gali, R. R., Schwikart, M., Larsen, C. N., Leggett, D. S., Müller, B., Feng, M. T., Tübing, F., Dittmar, G. A. G., & Finley, D. (2002). Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nature Cell Biology*, *4*(9), 725–730. https://doi.org/10.1038/ncb845

Emanuele, M. J., Elia, A. E. H., Xu, Q., Thoma, C. R., Izhar, L., Leng, Y., Guo, A., Chen, Y. N., Rush, J., Hsu, P. W. C., Yen, H. C. S., & Elledge, S. J. (2011). Global identification of modular cullin-RING ligase substrates. *Cell*, *147*(2), 459–474. https://doi.org/10.1016/j.cell.2011.09.019

Enchev, R. I., Schulman, B. A., & Peter, M. (2015). Protein neddylation: Beyond cullin-RING ligases. I: *Nature Reviews Molecular Cell Biology* (Vol. 16, Number 1, pp. 30–44). https://doi.org/10.1038/nrm3919

Epshtein, V., Kamarthapu, V., McGary, K., Svetlov, V., Ueberheide, B., Proshkin, S., Mironov, A., & Nudler, E. (2014). UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature*, *505*(7483), 372–377. https://doi.org/10.1038/nature12928

Etlinger, J. D., & Goldberg, A. L. (1977). A soluble ATP dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(1), 54–58. https://doi.org/10.1073/pnas.74.1.54

Evans, E., Fellows, J., Coffer, A., & Wood, R. D. (1997). Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *EMBO Journal*, *16*(3), 625–638. https://doi.org/10.1093/emboj/16.3.625

Evans, E., Moggs, J. G., Hwang, J. R., Egly, J. M., & Wood, R. D. (1997). Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO Journal*, *16*(21), 6559–6573. https://doi.org/10.1093/emboj/16.21.6559

Evans, M. D., Griffiths, H. R., & Lunec, J. (1997). Reactive Oxygen Species and their Cytotoxic Mechanisms. *Advances in Molecular and Cell Biology*, *20*(C), 25–73. https://doi.org/10.1016/S1569-2558(08)60271-4

Falck, J., Coates, J., & Jackson, S. P. (2005). Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature*, *434*(7033), 605–611. https://doi.org/10.1038/nature03442

Fan, J., Leroux-Coyau, M., Savery, N. J., & Strick, T. R. (2016). Reconstruction of

bacterial transcription-coupled repair at single-molecule resolution. *Nature*, *536*(7615), 234–237. https://doi.org/10.1038/nature19080

Faraoni, I., & Graziani, G. (2018). Role of BRCA mutations in cancer treatment with poly(ADP-ribose) polymerase (PARP) inhibitors. I: *Cancers* (Vol. 10, Number 12). MDPI AG. https://doi.org/10.3390/cancers10120487

Fei, J., Kaczmarek, N., Luch, A., Glas, A., Carell, T., & Naegeli, H. (2011). Regulation of nucleotide excision repair by UV-DDB: Prioritization of damage recognition to internucleosomal dna. *PLoS Biology*, *9*(10), e1001183. https://doi.org/10.1371/journal.pbio.1001183

Ferdous, A., Kodadek, T., & Johnston, S. A. (2002). A nonproteolytic function of the 19S regulatory subunit of the 26S proteasome is required for efficient activated transcription by human RNA polymerase II. *Biochemistry*, *41*(42), 12798–12805. https://doi.org/10.1021/bi020425t

Fernandez-Capetillo, O., Lee, A., Nussenzweig, M., & Nussenzweig, A. (2004). H2AX: The histone guardian of the genome. I: *DNA Repair* (Vol. 3, Numbers 8–9, pp. 959–967). https://doi.org/10.1016/j.dnarep.2004.03.024

Finley, D., Ciechanover, A., & Varshavsky, A. (1984). Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant ts85. *Cell*, *37*(1), 43–55. https://doi.org/10.1016/0092-8674(84)90299-X

Fischer, E. S., Scrima, A., Böhm, K., Matsumoto, S., Lingaraju, G. M., Faty, M., Yasuda, T., Cavadini, S., Wakasugi, M., Hanaoka, F., Iwai, S., Gut, H., Sugasawa, K., & Thomä, N. H. (2011). The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell*, *147*(5), 1024–1039. https://doi.org/10.1016/j.cell.2011.10.035

Fitch, M. E., Nakajima, S., Yasui, A., & Ford, J. M. (2003). In Vivo Recruitment of XPC to UV-induced Cyclobutane Pyrimidine Dimers by the DDB2 Gene Product. *Journal of Biological Chemistry*, *278*(47), 46906–46910. https://doi.org/10.1074/jbc.M307254200

Fitzgerald, D. J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T. J., & Berger, I. (2006). Protein complex expression by using multigene baculoviral vectors. *Nature Methods*, *3*(12), 1021–1032. https://doi.org/10.1038/nmeth983

Fousteri, M., Vermeulen, W., van Zeeland, A. A., & Mullenders, L. H. F. (2006). Cockayne Syndrome A and B Proteins Differentially Regulate Recruitment of Chromatin Remodeling and Repair Factors to Stalled RNA Polymerase II In Vivo. *Molecular Cell*, *23*(4), 471–482. https://doi.org/10.1016/j.molcel.2006.06.029

Friedberg, E. C. (2008). A brief history of the DNA repair field. *Cell Research*, *18*(1), 3–7. https://doi.org/10.1038/cr.2007.113

Fujiwara, Y., Masutani, C., Mizukoshi, T., Kondo, J., Hanaoka, F., & Iwai, S. (1999). Characterization of DNA recognition by the human UV-damaged DNA-binding protein. *Journal of Biological Chemistry*, *274*(28), 20027–20033. https://doi.org/10.1074/jbc.274.28.20027

Gaillard, H., Wellinger, R. E., & Aguilera, A. (2015). Methods to study transcription-coupled repair in chromatin. *Methods in Molecular Biology*, *1288*, 273–288. https://doi.org/10.1007/978-1-4939-2474-5_15

Ganoth, D., Bornstein, G., Ko, T. K., Larsen, B., Tyers, M., Pagano, M., & Hershko, A. (2001). The cell-cycle regulatory protein Cks1 is required for SCFSkp2-mediated ubiquitinylation of p27. *Nature Cell Biology*, *3*(3), 321–324. https://doi.org/10.1038/35060126

Gareau, J. R., & Lima, C. D. (2010). The SUMO pathway: Emerging mechanisms that shape specificity, conjugation and recognition. I: *Nature Reviews Molecular Cell Biology* (Vol. 11, Number 12, pp. 861–871). https://doi.org/10.1038/nrm3011

Garner, E., & Smogorzewska, A. (2011). Ubiquitylation and the Fanconi anemia pathway. I: *FEBS Letters* (Vol. 585, Number 18, pp. 2853–2860). NIH Public

Access. https://doi.org/10.1016/j.febslet.2011.04.078
George, D. L., & Witkin, E. M. (1974). Slow excision repair in an mfd mutant of Escherichia coli B/r. *MGG Molecular & General Genetics*, *133*(4), 283–291. https://doi.org/10.1007/BF00332704
George, D. L., & Witkin, E. M. (1975). Ultraviolet light-induced responses of an mfd mutant of escherichia coli B/r having a slow rate of dimer excision. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, *28*(3), 347–354. https://doi.org/10.1016/0027-5107(75)90229-8
Georges, A., Gopaul, D., Denby Wilkes, C., Giordanengo Aiach, N., Novikova, E., Barrault, M. B., Alibert, O., & Soutourina, J. (2019). Functional interplay between Mediator and RNA polymerase II in Rad2/XPG loading to the chromatin. *Nucleic Acids Research*, *47*(17), 8988–9004. https://doi.org/10.1093/nar/gkz598
Glotzer, M., Murray, A. W., & Kirschner, M. W. (1991). Cyclin is degraded by the ubiquitin pathway. *Nature*, *349*(6305), 132–138. https://doi.org/10.1038/349132a0
Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D. H., Niall, H. D., & Boyse, E. A. (1975). Isolation of a polypeptide that has lymphocyte differentiating properties and is probably represented universally in living cells. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(1), 11–15. https://doi.org/10.1073/pnas.72.1.11
Graham, J. M., Anyane-Yeboa, K., Raams, A., Appeldoorn, E., Kleijer, W. J., Garritsen, V. H., Busch, D., Edersheim, T. G., & Jaspers, N. G. J. (2001). Cerebro-oculo-facio-skeletal syndrome with a nucleotide excision-repair defect and a mutated XPD gene, with prenatal diagnosis in a triplet pregnancy. *American Journal of Human Genetics*, *69*(2), 291–300. https://doi.org/10.1086/321295
Graham, T. G. W., Walter, J. C., & Loparo, J. J. (2016). Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. *Molecular Cell*, *61*(6), 850–858. https://doi.org/10.1016/j.molcel.2016.02.010
Gregersen, L. H., Mitter, R., & Svejstrup, J. Q. (2020). Using TTchem-seq for profiling nascent transcription and measuring transcript elongation. *Nature Protocols*, *15*(2), 604–627. https://doi.org/10.1038/s41596-019-0262-3
Groisman, R., Kuraoka, I., Chevallier, O., Gaye, N., Magnaldo, T., Tanaka, K., Kisselev, A. F., Harel-Bellan, A., & Nakatani, Y. (2006). CSA-dependent degradation of CSB by the ubiquitin-proteasome pathway establishes a link between complementation factors of the Cockayne syndrome. *Genes and Development*, *20*(11), 1429–1434. https://doi.org/10.1101/gad.378206
Groisman, R., Polanowska, J., Kuraoka, I., Sawada, J. I., Saijo, M., Drapkin, R., Kisselev, A. F., Tanaka, K., & Nakatani, Y. (2003). The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell*, *113*(3), 357–367. https://doi.org/10.1016/S0092-8674(03)00316-7
Haas, A. L., Warms, J. V., Hershko, A., & Rose, I. A. (1982). Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *Journal of Biological Chemistry*, *257*(5), 2543–2548. https://doi.org/10.1016/s0021-9258(18)34958-5
Harper, J. W., & Elledge, S. J. (2007). The DNA Damage Response: Ten Years After. I: *Molecular Cell* (Vol. 28, Number 5, pp. 739–745). Cell Press. https://doi.org/10.1016/j.molcel.2007.11.015
Harreman, M., Taschner, M., Sigurdsson, S., Anindya, R., Reid, J., Somesh, B., Kong, S. E., Banks, C. A. S., Conaway, R. C., Conaway, J. W., & Svejstrup, J. Q. (2009). Distinct ubiquitin ligases act sequentially for RNA polymerase II polyubiquitylation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(49), 20705–20710. https://doi.org/10.1073/pnas.0907052106
Henner, W. D., Grunberg, S. M., & Haseltine, W. A. (1982). Sites and structure of γ radiation-induced DNA strand breaks. I: *Journal of Biological Chemistry* (Vol. 257,

Number 19). https://doi.org/10.1016/s0021-9258(18)33827-4
Henning, K. A., Li, L., Iyer, N., McDaniel, L. D., Reagan, M. S., Legerski, R., Schultz, R. A., Stefanini, M., Lehmann, A. R., Mayne, L. V., & Friedberg, E. C. (1995). The Cockayne syndrome group A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA polymerase II TFIIH. *Cell*, *82*(4), 555–564. https://doi.org/10.1016/0092-8674(95)90028-4
Hershko, A., Ciechanover, A., Heller, H., Haas, A. L., & Rose, I. A. (1980). Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proceedings of the National Academy of Sciences of the United States of America*, *77*(4), 1783–1786. https://doi.org/10.1073/pnas.77.4.1783
Hess, R. T., Schwitter, U., Petretta, M., Giese, B., & Naegeli, H. (1997). Bipartite substrate discrimination by human nucleotide excision repair. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(13), 6664–6669. https://doi.org/10.1073/pnas.94.13.6664
Hjerpe, R., Aillet, F., Lopitz-Otsoa, F., Lang, V., England, P., & Rodriguez, M. S. (2009). Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. *EMBO Reports*, *10*(11), 1250–1258. https://doi.org/10.1038/embor.2009.192
Hobson, D. J., Wei, W., Steinmetz, L. M., & Svejstrup, J. Q. (2012). RNA Polymerase II Collision Interrupts Convergent Transcription. *Molecular Cell*, *48*(3), 365–374. https://doi.org/10.1016/j.molcel.2012.08.027
Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G., & Jentsch, S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature*, *419*(6903), 135–141. https://doi.org/10.1038/nature00991
Hohl, M., Thorel, F., Clarkson, S. G., & Schärer, O. D. (2003). Structural determinants for substrate binding and catalysis by the structure-specific endonuclease XPG. *Journal of Biological Chemistry*, *278*(21), 19500–19508. https://doi.org/10.1074/jbc.M213155200
Houten, Bennett Van, Kuper, J., & Kisker, C. (2016). Role of XPD in cellular functions: To TFIIH and beyond. *DNA Repair*, *44*, 136–142. https://doi.org/10.1016/j.dnarep.2016.05.019
Howan, K., Smith, A. J., Westblade, L. F., Joly, N., Grange, W., Zorman, S., Darst, S. A., Savery, N. J., & Strick, T. R. (2012). Initiation of transcription-coupled repair characterized at single-molecule resolution. *Nature*, *490*(7420), 431–434. https://doi.org/10.1038/nature11430
Hu, X., Malik, S., Negroiu, C. C., Hubbard, K., Velalar, C. N., Hampton, B., Grosu, D., Catalano, J., Roeder, R. G., & Gnatt, A. (2006). A Mediator-responsive form of metozoan RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(25), 9506–9511. https://doi.org/10.1073/pnas.0603702103
Huang, D. T., Ayrault, O., Hunt, H. W., Taherbhoy, A. M., Duda, D. M., Scott, D. C., Borg, L. A., Neale, G., Murray, P. J., Roussel, M. F., & Schulman, B. A. (2009). E2-RING Expansion of the NEDD8 Cascade Confers Specificity to Cullin Modification. *Molecular Cell*, *33*(4), 483–495. https://doi.org/10.1016/j.molcel.2009.01.011
Huang, J. C., Svoboda, D. L., Reardon, J. T., & Sancar, A. (1992). Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. I: *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 89, Number 8). https://doi.org/10.1073/pnas.89.8.3664
Huang, Juch Chin, & Sancar, A. (1994). Determination of minimum substrate size for human excinuclease. *Journal of Biological Chemistry*, *269*(29), 19034–19040.

https://doi.org/10.1016/s0021-9258(17)32270-6
Ide, H., & Kotera, M. (2004). Human DNA glycosylases involved in the repair of oxidatively damaged DNA. I: *Biological and Pharmaceutical Bulletin* (Vol. 27, Number 4, pp. 480–485). Biol Pharm Bull. https://doi.org/10.1248/bpb.27.480
Ito, S., Kuraoka, I., Chymkowitch, P., Compe, E., Takedachi, A., Ishigami, C., Coin, F., Egly, J. M., & Tanaka, K. (2007). XPG Stabilizes TFIIH, Allowing Transactivation of Nuclear Receptors: Implications for Cockayne Syndrome in XP-G/CS Patients. *Molecular Cell*, *26*(2), 231–243. https://doi.org/10.1016/j.molcel.2007.03.013
Itoh, T., Ono, T., & Yamaizumi, M. (1994). A new UV-sensitive syndrome not belonging to any complementation groups of xeroderma pigmentosum or Cockayne syndrome: siblings showing biochemical characteristics of Cockayne syndrome without typical clinical manifestations. *Mutation Research-DNA Repair*, *314*(3), 233–248. https://doi.org/10.1016/0921-8777(94)90068-X
Ivanov, D., Kwak, Y. T., Guo, J., & Gaynor, R. B. (2000). Domains in the SPT5 Protein That Modulate Its Transcriptional Regulatory Properties. *Molecular and Cellular Biology*, *20*(9), 2970–2983. https://doi.org/10.1128/mcb.20.9.2970-2983.2000
Iyama, T., Okur, M. N., Golato, T., McNeill, D. R., Lu, H., Hamilton, R., Raja, A., Bohr, V. A., & Wilson, D. M. (2018). Regulation of the Intranuclear Distribution of the Cockayne Syndrome Proteins. *Scientific Reports*, *8*(1), 17490. https://doi.org/10.1038/s41598-018-36027-6
Iyama, T., & Wilson, D. M. (2016). Elements That Regulate the DNA Damage Response of Proteins Defective in Cockayne Syndrome. *Journal of Molecular Biology*, *428*(1), 62–78. https://doi.org/10.1016/j.jmb.2015.11.020
Jackson, S. P., & Durocher, D. (2013). Regulation of DNA Damage Responses by Ubiquitin and SUMO. I: *Molecular Cell* (Vol. 49, Number 5, pp. 795–807). Cell Press. https://doi.org/10.1016/j.molcel.2013.01.017
Jansen, L. E. T., Den Dulk, H., Brouns, R. M., De Ruijter, M., Brandsma, J. A., & Brouwer, J. (2000). Spt4 modulates Rad26 requirement in transcription-coupled nucleotide excision repair. *EMBO Journal*, *19*(23), 6498–6507. https://doi.org/10.1093/emboj/19.23.6498
Jiang, Y., Rabbi, M., Kim, M., Ke, C., Lee, W., Clark, R. L., Mieczkowski, P. A., & Marszalek, P. E. (2009). UVA generates pyrimidine dimers in DNA directly. *Biophysical Journal*, *96*(3), 1151–1158. https://doi.org/10.1016/j.bpj.2008.10.030
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829
Jing, Y., Kao, J. F. L., & Taylor, J. S. (1998). Thermodynamic and base-pairing studies of matched and mismatched DNA dodecamer duplexes containing cis-syn, (6-4) and Dewar photoproducts of TT. I: *Nucleic Acids Research* (Vol. 26, Number 16). https://doi.org/10.1093/nar/26.16.3845
Jinks-Robertson, S., & Bhagwat, A. S. (2014). Transcription-Associated mutagenesis. *Annual Review of Genetics*, *48*(1), 341–359. https://doi.org/10.1146/annurev-genet-120213-092015
Johnson, R. E., Kondratick, C. M., Prakash, S., & Prakash, L. (1999). hRAD30 mutations in the variant form of xeroderma pigmentosum. *Science*, *285*(5425), 263–265. https://doi.org/10.1126/science.285.5425.263
Jones, C. J., & Wood, R. D. (1993). Preferential Binding of the Xeroderma Pigmentosum Group A Complementing Protein to Damaged DNA. *Biochemistry*, *32*(45), 12096–12104. https://doi.org/10.1021/bi00096a021
Kadyrov, F. A., Dzantiev, L., Constantin, N., & Modrich, P. (2006). Endonucleolytic Function of MutLα in Human Mismatch Repair. *Cell*, *126*(2), 297–308. https://doi.org/10.1016/j.cell.2006.05.039
Kanayama, A., Seth, R. B., Sun, L., Ea, C. K., Hong, M., Shaito, A., Chiu, Y. H., Deng,

L., & Chen, Z. J. (2004). TAB2 and TAB3 activate the NF-κB pathway through binding to polyubiquitin chains. *Molecular Cell*, *15*(4), 535–548. https://doi.org/10.1016/j.molcel.2004.08.008

Kannouche, P. L., & Lehmann, A. R. (2004). Ubiquitination of PCNA and the polymerase switch in human cells. I: *Cell Cycle* (Vol. 3, Number 8, pp. 1009–1011). Taylor and Francis Inc. https://doi.org/10.4161/cc.3.8.1074

Karikkineth, A. C., Scheibye-Knudsen, M., Fivenson, E., Croteau, D. L., & Bohr, V. A. (2017). Cockayne syndrome: Clinical features, model systems and pathways. I: *Ageing Research Reviews* (Vol. 33, pp. 3–17). Elsevier Ireland Ltd. https://doi.org/10.1016/j.arr.2016.08.002

Karin, M., & Ben-Neriah, Y. (2000). Phosphorylation meets ubiquitination: The control of NF-κB activity. I: *Annual Review of Immunology* (Vol. 18, pp. 621–663). Annu Rev Immunol. https://doi.org/10.1146/annurev.immunol.18.1.621

KELNER, A. (1949). Effect of visible light on the recovery of Streptomyces griseus conidia. *Proceedings of the National Academy of Sciences of the United States Of*, *35*(2), 73–79. https://doi.org/10.1073/pnas.35.2.73

Kemp, M. G., Reardon, J. T., Lindsey-Boltz, L. A., & Sancar, A. (2012). Mechanism of release and fate of excised oligonucleotides during nucleotide excision repair. *Journal of Biological Chemistry*, *287*(27), 22889–22899. https://doi.org/10.1074/jbc.M112.374447

Khanna, K. K., & Jackson, S. P. (2001). DNA double-strand breaks: Signaling, repair and the cancer connection. *Nature Genetics*, *27*(3), 247–254. https://doi.org/10.1038/85798

Kim, H. C., & Huibregtse, J. M. (2009). Polyubiquitination by HECT E3s and the Determinants of Chain Type Specificity. *Molecular and Cellular Biology*, *29*(12), 3307–3318. https://doi.org/10.1128/mcb.00240-09

Kim, J. H., Saraf, A., Florens, L., Washburn, M., & Workman, J. L. (2010). Gcn5 regulates the dissociation of SWI/SNF from chromatin by acetylation of Swi2/Snf2. *Genes and Development*, *24*(24), 2766–2771. https://doi.org/10.1101/gad.1979710

Kim, W., Bennett, E. J., Huttlin, E. L., Guo, A., Li, J., Possemato, A., Sowa, M. E., Rad, R., Rush, J., Comb, M. J., Harper, J. W., & Gygi, S. P. (2011). Systematic and quantitative assessment of the ubiquitin-modified proteome. *Molecular Cell*, *44*(2), 325–340. https://doi.org/10.1016/j.molcel.2011.08.025

Kirkali, G., de Souza-Pinto, N. C., Jaruga, P., Bohr, V. A., & Dizdaroglu, M. (2009). Accumulation of (5′S)-8,5′-cyclo-2′-deoxyadenosine in organs of Cockayne syndrome complementation group B gene knockout mice. *DNA Repair*, *8*(2), 274–278. https://doi.org/10.1016/j.dnarep.2008.09.009

Klein, B. J., Bose, D., Baker, K. J., Yusoff, Z. M., Zhang, X., & Murakami, K. S. (2011). RNA polymerase and transcription elongation factor Spt4/5 complex structure. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(2), 546–550. https://doi.org/10.1073/pnas.1013828108

Klungland, A., & Lindahl, T. (1997). Second pathway for completion of human DNA base excision-repair: Reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO Journal*, *16*(11), 3341–3348. https://doi.org/10.1093/emboj/16.11.3341

Kokic, G., Chernev, A., Tegunov, D., Dienemann, C., Urlaub, H., & Cramer, P. (2019). Structural basis of TFIIH activation for nucleotide excision repair. *Nature Communications*, *10*(1), 1–9. https://doi.org/10.1038/s41467-019-10745-5

Koliopoulos, M. G., Esposito, D., Christodoulou, E., Taylor, I. A., & Rittinger, K. (2016). Functional role of TRIM E3 ligase oligomerization and regulation of catalytic activity . *The EMBO Journal*, *35*(11), 1204–1218. https://doi.org/10.15252/embj.201593741

Komander, D., Clague, M. J., & Urbé, S. (2009). Breaking the chains: Structure and function of the deubiquitinases. *Nature Reviews Molecular Cell Biology*, *10*(8), 550–563. https://doi.org/10.1038/nrm2731

Komander, D., & Rape, M. (2012). The ubiquitin code. *Annual Review of Biochemistry*, *81*, 203–229. https://doi.org/10.1146/annurev-biochem-060310-170328

Krappmann, D., & Scheidereit, C. (2005). A pervasive role of ubiquitin conjugation in activation and termination of IκB kinase pathways. I: *EMBO Reports* (Vol. 6, Number 4, pp. 321–326). European Molecular Biology Organization. https://doi.org/10.1038/sj.embor.7400380

Krasikova, Y. S., Rechkunova, N. I., Maltseva, E. A., & Lavrik, O. I. (2018). RPA and XPA interaction with DNA structures mimicking intermediates of the late stages in nucleotide excision repair. *PLoS ONE*, *13*(1), e0190782. https://doi.org/10.1371/journal.pone.0190782

Krasikova, Y. S., Rechkunova, N. I., Maltseva, E. A., Petruseva, I. O., & Lavrik, O. I. (2010). Localization of xeroderma pigmentosum group A protein and replication protein A on damaged DNA in nucleotide excision repair. *Nucleic Acids Research*, *38*(22), 8083–8094. https://doi.org/10.1093/nar/gkq649

Krokan, H. E., & Bjørås, M. (2013). Base excision repair. *Cold Spring Harbor Perspectives in Biology*, *5*(4), 1–22. https://doi.org/10.1101/cshperspect.a012583

Kubota, Y., Nash, R. A., Klungland, A., Schär, P., Barnes, D. E., & Lindahl, T. (1996). Reconstitution of DNA base excision-repair with purified human proteins: Interaction between DNA polymerase β and the XRCC1 protein. *EMBO Journal*, *15*(23), 6662–6670. https://doi.org/10.1002/j.1460-2075.1996.tb01056.x

Kulaksız, G., Reardon, J. T., & Sancar, A. (2005). Xeroderma Pigmentosum Complementation Group E Protein (XPE/DDB2): Purification of Various Complexes of XPE and Analyses of Their Damaged DNA Binding and Putative DNA Repair Properties. *Molecular and Cellular Biology*, *25*(22), 9784–9792. https://doi.org/10.1128/mcb.25.22.9784-9792.2005

Kulathu, Y., & Komander, D. (2012). Atypical ubiquitylation-the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages. I: *Nature Reviews Molecular Cell Biology* (Vol. 13, Number 8, pp. 508–523). Nature Publishing Group. https://doi.org/10.1038/nrm3394

Kunjappu, M. J., & Hochstrasser, M. (2014). Assembly of the 20S proteasome. I: *Biochimica et Biophysica Acta - Molecular Cell Research* (Vol. 1843, Number 1, pp. 2–12). Elsevier. https://doi.org/10.1016/j.bbamcr.2013.03.008

Kuper, J., Braun, C., Elias, A., Michels, G., Sauer, F., Schmitt, D. R., Poterszman, A., Egly, J. M., & Kisker, C. (2014). In TFIIH, XPD Helicase Is Exclusively Devoted to DNA Repair. *PLoS Biology*, *12*(9), e1001954. https://doi.org/10.1371/journal.pbio.1001954

Kuraoka, I., Ito, S., Wada, T., Hayashida, M., Lee, L., Saijo, M., Nakatsu, Y., Matsumoto, M., Matsunaga, T., Handa, H., Qin, J., Nakatani, Y., & Tanaka, K. (2008). Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair. *Journal of Biological Chemistry*, *283*(2), 940–950. https://doi.org/10.1074/jbc.M706647200

Kyng, K. J., May, A., Brosh, R. M., Cheng, W. H., Chen, C., Becker, K. G., & Bohr, V. A. (2003). The transcriptional response after oxidative stress is defective in Cockayne syndrome group B cells. *Oncogene*, *22*(8), 1135–1149. https://doi.org/10.1038/sj.onc.1206187

Lainé, J. P., & Egly, J. M. (2006). Initiation of DNA repair mediated by a stalled RNA polymerase IIO. *EMBO Journal*, *25*(2), 387–397. https://doi.org/10.1038/sj.emboj.7600933

Lake, R. J., Geyko, A., Hemashettar, G., Zhao, Y., & Fan, H. Y. (2010). UV-Induced Association of the CSB Remodeling Protein with Chromatin Requires ATP-

Dependent Relief of N-Terminal Autorepression. *Molecular Cell*, *37*(2), 235–246. https://doi.org/10.1016/j.molcel.2009.10.027

Lamers, M. H., Perrakis, A., Enzlin, J. H., Winterwerp, H. H. K., De Wind, N., & Sixma, T. K. (2000). The crystal structure of DNA mismatch repair protein MutS binding to a G·T mismatch. *Nature*, *407*(6805), 711–717. https://doi.org/10.1038/35037523

Laugel, V. (2013). Cockayne syndrome: The expanding clinical and mutational spectrum. *Mechanisms of Ageing and Development*, *134*(5–6), 161–170. https://doi.org/10.1016/j.mad.2013.02.006

Lavigne, M. D., Konstantopoulos, D., Ntakou-Zamplara, K. Z., Liakos, A., & Fousteri, M. (2017). Global unleashing of transcription elongation waves in response to genotoxic stress restricts somatic mutation rate. *Nature Communications*, *8*(1), 2076. https://doi.org/10.1038/s41467-017-02145-4

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T., & Carey, V. J. (2013). Software for Computing and Annotating Genomic Ranges. *PLoS Computational Biology*, *9*(8). https://doi.org/10.1371/journal.pcbi.1003118

Le May, N., Fradin, D., Iltis, I., Bougnères, P., & Egly, J. M. (2012). XPG and XPF Endonucleases Trigger Chromatin Looping and DNA Demethylation for Accurate Expression of Activated Genes. *Molecular Cell*, *47*(4), 622–632. https://doi.org/10.1016/j.molcel.2012.05.050

Le, T. T., Yang, Y., Tan, C., Suhanovsky, M. M., Fulbright, R. M., Inman, J. T., Li, M., Lee, J., Perelman, S., Roberts, J. W., Deaconescu, A. M., & Wang, M. D. (2018). Mfd Dynamically Regulates Transcription via a Release and Catch-Up Mechanism. *Cell*, *172*(1–2), 344-357.e15. https://doi.org/10.1016/j.cell.2017.11.017

Ledesma, F. C., El Khamisy, S. F., Zuma, M. C., Osborn, K., & Caldecott, K. W. (2009). A human 5′-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. *Nature*, *461*(7264), 674–678. https://doi.org/10.1038/nature08444

Lee, J. Y., Lake, R. J., Kirk, J., Bohr, V. A., Fan, H. Y., & Hohng, S. (2017). NAP1L1 accelerates activation and decreases pausing to enhance nucleosome remodeling by CSB. *Nucleic Acids Research*, *45*(8), 4696–4707. https://doi.org/10.1093/nar/gkx188

Lee, J., & Zhou, P. (2007). DCAFs, the Missing Link of the CUL4-DDB1 Ubiquitin Ligase. I: *Molecular Cell* (Vol. 26, Number 6, pp. 775–780). Cell Press. https://doi.org/10.1016/j.molcel.2007.06.001

Lehmann, A. R. (1982). Three complementation groups in Cockayne syndrome. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, *106*(2), 347–356. https://doi.org/10.1016/0027-5107(82)90115-4

Lehmann, A. R., Kirk-Bell, S., Arlett, C. F., Paterson, M. C., Lohman, P. H., de Weerd-Kastelein, E. A., & Bootsma, D. (1975). Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV irradiation. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(1), 219–223. https://doi.org/10.1073/pnas.72.1.219

Lehmann, Alan R., Kirk-Bell, S., & Mayne, L. (1979). Abnormal Kinetics of DNA Synthesis in Ultraviolet Light-irradiated Cells from Patients with Cockayne's Syndrome. *Cancer Research*, *39*(10), 4237–4241.

Lehmann, Alan R., McGibbon, D., & Stefanini, M. (2011). Xeroderma pigmentosum. I: *Orphanet Journal of Rare Diseases* (Vol. 6, Number 1, p. 70). BioMed Central. https://doi.org/10.1186/1750-1172-6-70

Li, C. L., Golebiowski, F. M., Onishi, Y., Samara, N. L., Sugasawa, K., & Yang, W. (2015). Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair. *Molecular Cell*, *59*(6), 1025–1034.

https://doi.org/10.1016/j.molcel.2015.08.012
Li, G. M. (2008). Mechanisms and functions of DNA mismatch repair. I: *Cell Research* (Vol. 18, Number 1, pp. 85–98). Nature Publishing Group. https://doi.org/10.1038/cr.2007.115
Li, S., & Smerdon, M. J. (2002). Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in Saccharomyces cerevisiae. *EMBO Journal*, *21*(21), 5921–5929. https://doi.org/10.1093/emboj/cdf589
Li, W., & Ye, Y. (2008). Polyubiquitin chains: Functions, structures, and mechanisms. I: *Cellular and Molecular Life Sciences* (Vol. 65, Number 15, pp. 2397–2406). https://doi.org/10.1007/s00018-008-8090-6
Liakos, A., Konstantopoulos, D., Lavigne, M. D., & Fousteri, M. (2020). Continuous transcription initiation guarantees robust repair of all transcribed genes and regulatory regions. *Nature Communications*, *11*(1), 1–16. https://doi.org/10.1038/s41467-020-14566-9
Liebelt, F., Schimmel, J., Verlaan-De Vries, M., Klemann, E., Van Royen, M. E., Van Der Weegen, Y., Luijsterburg, M. S., Mullenders, L. H., Pines, A., Vermeulen, W., & Vertegaal, A. C. O. (2020). Transcription-coupled nucleotide excision repair is coordinated by ubiquitin and SUMO in response to ultraviolet irradiation. *Nucleic Acids Research*, *48*(1), 231–248. https://doi.org/10.1093/nar/gkz977
Lindahl, T. (1993). Instability and decay of the primary structure of DNA. I: *Nature* (Vol. 362, Number 6422, pp. 709–715). Nature. https://doi.org/10.1038/362709a0
Lindahl, T., & Nyberg, B. (1972). Rate of Depurination of Native Deoxyribonucleic Acid. *Biochemistry*, *11*(19), 3610–3618. https://doi.org/10.1021/bi00769a018
Liu, E. Y., Cali, C. P., & Lee, E. B. (2017). RNA metabolism in neurodegenerative disease. I: *DMM Disease Models and Mechanisms* (Vol. 10, Number 5, pp. 509–518). Company of Biologists Ltd. https://doi.org/10.1242/dmm.028613
Loeb, L. A., & Monnat, R. J. (2008). DNA polymerases and human disease. I: *Nature Reviews Genetics* (Vol. 9, Number 8, pp. 594–604). Nature Publishing Group. https://doi.org/10.1038/nrg2345
Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12). https://doi.org/10.1186/s13059-014-0550-8
Luijsterburg, M. S., Lindh, M., Acs, K., Vrouwe, M. G., Pines, A., van Attikum, H., Mullenders, L. H., & Dantuma, N. P. (2012). DDB2 promotes chromatin decondensation at UV-induced DNA damage. *Journal of Cell Biology*, *197*(2), 267–281. https://doi.org/10.1083/jcb.201106074
Lyapina, S., Cope, G., Shevchenko, A., Serino, G., Tsuge, T., Zhou, C., Wolf, D. A., Wei, N., Shevchenko, A., & Deshaies, R. J. (2001). Promotion of NEDD8-CUL1 conjugate cleavage by COP9 signalosome. *Science*, *292*(5520), 1382–1385. https://doi.org/10.1126/science.1059780
Lydeard, J. R., Schulman, B. A., & Harper, J. W. (2013). Building and remodelling Cullin-RING E3 ubiquitin ligases. *EMBO Reports*, *14*(12), 1050–1061. https://doi.org/10.1038/embor.2013.173
Martin, N., Schwamborn, K., Schreiber, V., Werner, A., Guillier, C., Zhang, X. D., Bischof, O., Seeler, J. S., & Dejean, A. (2009). PARP-1 transcriptional activity is regulated by sumoylation upon heat shock. *EMBO Journal*, *28*(22), 3534–3548. https://doi.org/10.1038/emboj.2009.279
Martinez-Rucobo, F. W., Sainsbury, S., Cheung, A. C. M., & Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO Journal*, *30*(7), 1302–1310. https://doi.org/10.1038/emboj.2011.64
Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokol, M., Yuasa, M., Araki, M., Iwai, S., Takio, K., & Hanaoka, F. (1999). The XPV (xeroderma pigmentosum

variant) gene encodes human DNA polymerase η. *Nature*, *399*(6737), 700–704. https://doi.org/10.1038/21447

Masutani, C., Sugasawa, K., Yanagisawa, J., Sonoyama, T., Ui, M., Enomoto, T., Takio, K., Tanaka, K., Van Der Spek, P. J., Bootsma, D., Hoeijmakers, J. H. J., & Hanaoka, F. (1994). Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *EMBO Journal*, *13*(8), 1831–1843. https://doi.org/10.1002/j.1460-2075.1994.tb06452.x

Mathieu, N., Kaczmarek, N., & Naegeli, H. (2010). Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(41), 17545–17550. https://doi.org/10.1073/pnas.1004339107

Matsuda, N., Azuma, K., Saijo, M., Iemura, S. I., Hioki, Y., Natsume, T., Chiba, T., Tanaka, K., & Tanaka, K. (2005). DDB2, the xeroderma pigmentosum group E gene product, is directly ubiquitylated by Cullin 4A-based ubiquitin ligase complex. *DNA Repair*, *4*(5), 537–545. https://doi.org/10.1016/j.dnarep.2004.12.012

Matsumoto, Y., & Kim, K. (1995). Excision of deoxyribose phosphate residues by DNA polymerase β during DNA repair. *Science*, *269*(5224), 699–702. https://doi.org/10.1126/science.7624801

Matsunaga, T., Park, C. H., Bessho, T., Mu, D., & Sancar, A. (1996). Replication protein A confers structure-specific endonuclease activities to the XPF-ERCC1 and XPG subunits of human DNA repair excision nuclease. *Journal of Biological Chemistry*, *271*(19), 11047–11050. https://doi.org/10.1074/jbc.271.19.11047

Mayne, L., & Lehmann, A. R. (1982). Failure of RNA synthesis to recover after UV irradiation: An early defect in cells from individuals with Cockayne syndrome and xeroderma pigmentosum. *Mutation Research*, *96*(1), 140. https://doi.org/10.1016/0027-5107(82)90047-1

Mellon, I., Bohr, V. A., Smith, C. A., & Hanawalt, P. C. (1986). Preferential DNA repair of an active gene in human cells. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(23), 8878–8882. https://doi.org/10.1073/pnas.83.23.8878

Mellon, Isabel, & Hanawalt, P. C. (1989). Induction of the Escherichia coli lactose operon selectively increases repair of its transcribed DNA strand. *Nature*, *342*(6245), 95–98. https://doi.org/10.1038/342095a0

Mellon, Isabel, Spivak, G., & Hanawalt, P. C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell*, *51*(2), 241–249. https://doi.org/10.1016/0092-8674(87)90151-6

Menoni, H., Hoeijmakers, J. H. J., & Vermeulen, W. (2012). Nucleotide excision repair-initiating proteins bind to oxidative DNA lesions in vivo. *Journal of Cell Biology*, *199*(7), 1037–1046. https://doi.org/10.1083/jcb.201205149

Mevissen, T. E. T., Hospenthal, M. K., Geurink, P. P., Elliott, P. R., Akutsu, M., Arnaudo, N., Ekkebus, R., Kulathu, Y., Wauer, T., El Oualid, F., Freund, S. M. V., Ovaa, H., & Komander, D. (2013). OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. *Cell*, *154*(1), 169. https://doi.org/10.1016/j.cell.2013.05.046

Min, J. H., & Pavletich, N. P. (2007). Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature*, *449*(7162), 570–575. https://doi.org/10.1038/nature06155

Modrich, P., & Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annual Review of Biochemistry*, *65*(1), 101–133. https://doi.org/10.1146/annurev.bi.65.070196.000533

Moreno, N. C., De Souza, T. A., Garcia, C. C. M. H., Ruiz, N. Q., Corradi, C., Castro, L. P., Munford, V., Ienne, S., Alexandrov, L. B., & Menck, C. F. M. (2020). Whole-

exome sequencing reveals the impact of UVA light mutagenesis in xeroderma pigmentosum variant human cells. *Nucleic Acids Research*, *48*(4), 1941–1953. https://doi.org/10.1093/nar/gkz1182

Moser, J., Kool, H., Giakzidis, I., Caldecott, K., Mullenders, L. H. F., & Fousteri, M. I. (2007). Sealing of Chromosomal DNA Nicks during Nucleotide Excision Repair Requires XRCC1 and DNA Ligase IIIα in a Cell-Cycle-Specific Manner. *Molecular Cell*, *27*(2), 311–323. https://doi.org/10.1016/j.molcel.2007.06.014

Mouret, S., Philippe, C., Gracia-Chantegrel, J., Banyasz, A., Karpati, S., Markovitsi, D., & Douki, T. (2010). UVA-induced cyclobutane pyrimidine dimers in DNA: A direct photochemical mechanism? *Organic and Biomolecular Chemistry*, *8*(7), 1706–1711. https://doi.org/10.1039/b924712b

Mu, D., Park, C. H., Matsunaga, T., Hsu, D. S., Reardon, J. T., & Sancar, A. (1995). Reconstitution of human DNA repair excision nuclease in a highly defined system. *Journal of Biological Chemistry*, *270*(6), 2415–2418. https://doi.org/10.1074/jbc.270.6.2415

Mu, D., Tursun, M., Duckett, D. R., Drummond, J. T., Modrich, P., & Sancar, A. (1997). Recognition and repair of compound DNA lesions (base damage and mismatch) by human mismatch repair and excision repair systems. *Molecular and Cellular Biology*, *17*(2), 760–769. https://doi.org/10.1128/mcb.17.2.760

Muftuoglu, M., de Souza-Pinto, N. C., Dogan, A., Aamann, M., Stevnsner, T., Rybanska, I., Kirkali, G., Dizdaroglu, M., & Bohr, V. A. (2009). Cockayne syndrome group B protein stimulates repair of formamidopyrimidines by NEIL1 DNA glycosylase. *Journal of Biological Chemistry*, *284*(14), 9270–9279. https://doi.org/10.1074/jbc.M807006200

Nakamura, S., Roth, J. A., & Mukhopadhyay, T. (2000). Multiple Lysine Mutations in the C-Terminal Domain of p53 Interfere with MDM2-Dependent Protein Degradation and Ubiquitination. *Molecular and Cellular Biology*, *20*(24), 9391–9398. https://doi.org/10.1128/mcb.20.24.9391-9398.2000

Nakatsu, Y., Asahina, H., Citterio, E., Rademakers, S., Vermeulen, W., Kamiuchi, S., Yeo, J. P., Khaw, M. C., Saijo, M., Kodo, N., Matsuda, T., Hoeijmakers, J. H. J., & Tanaka, K. (2000). XAB2, a novel tetratricopeptide repeat protein involved in transcription-coupled DNA repair and transcription. *Journal of Biological Chemistry*, *275*(45), 34931–34937. https://doi.org/10.1074/jbc.M004936200

Nakazawa, Y., Hara, Y., Oka, Y., Komine, O., van den Heuvel, D., Guo, C., Daigaku, Y., Isono, M., He, Y., Shimada, M., Kato, K., Jia, N., Hashimoto, S., Kotani, Y., Miyoshi, Y., Tanaka, M., Sobue, A., Mitsutake, N., Suganami, T., … Ogi, T. (2020). Ubiquitination of DNA Damage-Stalled RNAPII Promotes Transcription-Coupled Repair. *Cell*, *180*(6), 1228-1244.e24. https://doi.org/10.1016/j.cell.2020.02.010

Nakazawa, Y., Sasaki, K., Mitsutake, N., Matsuse, M., Shimada, M., Nardo, T., Takahashi, Y., Ohyama, K., Ito, K., Mishima, H., Nomura, M., Kinoshita, A., Ono, S., Takenaka, K., Masuyama, R., Kudo, T., Slor, H., Utani, A., Tateishi, S., … Ogi, T. (2012). Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase IIo processing in transcription-coupled nucleotide-excision repair. *Nature Genetics*, *44*(5), 586–592. https://doi.org/10.1038/ng.2229

Nance, M. A., & Berry, S. A. (1992). Cockayne Syndrome: Review of 140 cases. *American Journal of Medical Genetics*, *42*(1), 68–84. https://doi.org/10.1002/ajmg.1320420115

Natale, V., & Raquer, H. (2017). Xeroderma pigmentosum-Cockayne syndrome complex. I: *Orphanet Journal of Rare Diseases* (Vol. 12, Number 1). BioMed Central Ltd. https://doi.org/10.1186/s13023-017-0616-2

Newman, J. C., Bailey, A. D., Fan, H. Y., Pavelitz, T., & Weiner, A. M. (2008). An abundant evolutionarily conserved CSB-PiggyBac fusion protein expressed in

cockayne syndrome. *PLoS Genetics*, *4*(3). https://doi.org/10.1371/journal.pgen.1000031

Newman, J. C., Bailey, A. D., & Weiner, A. M. (2006). Cockayne syndrome group B protein (CSB) plays a general role in chromatin maintenance and remodeling. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(25), 9613–9618. https://doi.org/10.1073/pnas.0510909103

Ng, J. M. Y., Vermeulen, W., Van der Horst, G. T. J., Bergink, S., Sugasawa, K., Vrieling, H., & Hoeijmakers, J. H. J. (2003). A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein. *Genes and Development*, *17*(13), 1630–1645. https://doi.org/10.1101/gad.260003

Noe Gonzalez, M., Blears, D., & Svejstrup, J. Q. (2021). Causes and consequences of RNA polymerase II stalling during transcript elongation. *Nature Reviews Molecular Cell Biology*, *22*(1), 3–21. https://doi.org/10.1038/s41580-020-00308-8

O'Gorman, S., Fox, D. T., & Wahl, G. M. (1991). Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science*, *251*(4999), 1351–1355. https://doi.org/10.1126/science.1900642

Odell, I. D., Wallace, S. S., & Pederson, D. S. (2013). Rules of engagement for base excision repair in chromatin. *Journal of Cellular Physiology*, *228*(2), 258–266. https://doi.org/10.1002/jcp.24134

Oe, T., Nakajo, N., Katsuragi, Y., Okazaki, K., & Sagata, N. (2001). Cytoplasmic occurrence of the Chk1/Cdc25 pathway and regulation of Chk1 in Xenopus oocytes. *Developmental Biology*, *229*(1), 250–261. https://doi.org/10.1006/dbio.2000.9968

Ogi, T., & Lehmann, A. R. (2006). The Y-family DNA polymerase κ (pol κ) functions in mammalian nucleotide-excision repair. *Nature Cell Biology*, *8*(6), 640–642. https://doi.org/10.1038/ncb1417

Ogi, T., Limsirichaikul, S., Overmeer, R. M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N. G., Mullenders, L. H. F., Yamashita, S., Fousteri, M. I., & Lehmann, A. R. (2010). Three DNA Polymerases, Recruited by Different Mechanisms, Carry Out NER Repair Synthesis in Human Cells. *Molecular Cell*, *37*(5), 714–727. https://doi.org/10.1016/j.molcel.2010.02.009

Ohno, A., Jee, J. G., Fujiwara, K., Tenno, T., Goda, N., Tochio, H., Kobayashi, H., Hiroaki, H., & Shirakawa, M. (2005). Structure of the UBA domain of Dsk2p in complex with ubiquitin: Molecular determinants for ubiquitin recognition. *Structure*, *13*(4), 521–532. https://doi.org/10.1016/j.str.2005.01.011

Oksenych, V., De Jesus, B. B., Zhovmer, A., Egly, J. M., & Coin, F. (2009). Molecular insights into the recruitment of TFIIH to sites of DNA damage. *EMBO Journal*, *28*(19), 2971–2980. https://doi.org/10.1038/emboj.2009.230

Ortolan, T. G., Chen, L., Tongaonkar, P., & Madura, K. (2004). Rad23 stabilizes Rad4 from degradation by the Ub/proteasome pathway. *Nucleic Acids Research*, *32*(22), 6490–6500. https://doi.org/10.1093/nar/gkh987

Osakabe, A., Tachiwana, H., Kagawa, W., Horikoshi, N., Matsumoto, S., Hasegawa, M., Matsumoto, N., Toga, T., Yamamoto, J., Hanaoka, F., Thomä, N. H., Sugasawa, K., Iwai, S., & Kurumizaka, H. (2015). Structural basis of pyrimidine-pyrimidone (6-4) photoproduct recognition by UV-DDB in the nucleosome. *Scientific Reports*, *5*, 16330. https://doi.org/10.1038/srep16330

Osterod, M., Larsen, E., Le Page, F., Hengstler, J. G., Van der Horst, G. T. J., Boiteux, S., Klungland, A., & Epe, B. (2002). A global DNA repair mechanism involving the Cockayne syndrome B (CSB) gene product can prevent the in vivo accumulation of endogenous oxidative DNA base damage. *Oncogene*, *21*(54), 8232–8239. https://doi.org/10.1038/sj.onc.1206027

Park, H. J., Zhang, K., Ren, Y., Nadji, S., Sinha, N., Taylor, J. S., & Kang, C. H. (2002).

Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(25), 15965–15970. https://doi.org/10.1073/pnas.242422699

Park, J. S., Marr, M. T., & Roberts, J. W. (2002). E. coli transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell*, *109*(6), 757–767. https://doi.org/10.1016/S0092-8674(02)00769-9

Passmore, L. A., & Barford, D. (2004). Getting into position: The catalytic mechanisms of protein of protein ubiquitylation. I: *Biochemical Journal* (Vol. 379, Number 3). https://doi.org/10.1042/BJ20040198

Petroski, M. D., & Deshaies, R. J. (2003). Context of multiubiquitin chain attachment influences the rate of Sic1 degradation. *Molecular Cell*, *11*(6), 1435–1444. https://doi.org/10.1016/S1097-2765(03)00221-1

Pettijohn, D., & Hanawalt, P. (1964). Evidence for repair-replication of ultraviolet damaged DNA in bacteria. *Journal of Molecular Biology*, *9*(2), 395–410. https://doi.org/10.1016/S0022-2836(64)80216-3

Plechanovov, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H., & Hay, R. T. (2012). Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature*, *489*(7414), 115–120. https://doi.org/10.1038/nature11376

Poulsen, S. L., Hansen, R. K., Wagner, S. A., van Cuijk, L., van Belle, G. J., Streicher, W., Wikström, M., Choudhary, C., Houtsmuller, A. B., Marteijn, J. A., Bekker-Jensen, S., & Mailand, N. (2013). RNF111/Arkadia is a SUMO-targeted ubiquitin ligase that facilitates the DNA damage response. *Journal of Cell Biology*, *201*(6), 787–807. https://doi.org/10.1083/jcb.201212075

Povlsen, L. K., Beli, P., Wagner, S. A., Poulsen, S. L., Sylvestersen, K. B., Poulsen, J. W., Nielsen, M. L., Bekker-Jensen, S., Mailand, N., & Choudhary, C. (2012). Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass. *Nature Cell Biology*, *14*(10), 1089–1098. https://doi.org/10.1038/ncb2579

Prakash, S., & Prakash, L. (2000). Nucleotide excision repair in yeast. I: *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* (Vol. 451, Numbers 1–2, pp. 13–24). Elsevier. https://doi.org/10.1016/S0027-5107(00)00037-3

Proietti-De-Santis, L., Drané, P., & Egly, J. M. (2006). Cockayne syndrome B protein regulates the transcriptional program after UV irradiation. *EMBO Journal*, *25*(9), 1915–1923. https://doi.org/10.1038/sj.emboj.7601071

Puumalainen, M. R., Lessel, D., Rüthemann, P., Kaczmarek, N., Bachmann, K., Ramadan, K., & Naegeli, H. (2014). Chromatin retention of DNA damage sensors DDB2 and XPC through loss of p97 segregase causes genotoxicity. *Nature Communications*, *5*, 3695. https://doi.org/10.1038/ncomms4695

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Raasi, S., Varadan, R., Fushman, D., & Pickart, C. M. (2005). Diverse polyubiquitin interaction properties of ubiquitin-associated domains. *Nature Structural and Molecular Biology*, *12*(8), 708–714. https://doi.org/10.1038/nsmb962

Rabut, G., & Peter, M. (2008). Function and regulation of protein neddylation. 'Protein modifications: beyond the usual suspects' review series. I: *EMBO Reports* (Vol. 9, Number 10, pp. 969–976). https://doi.org/10.1038/embor.2008.183

Rahl, P. B., Lin, C. Y., Seila, A. C., Flynn, R. A., McCuine, S., Burge, C. B., Sharp, P. A., & Young, R. A. (2010). C-Myc regulates transcriptional pause release. *Cell*, *141*(3), 432–445. https://doi.org/10.1016/j.cell.2010.03.030

Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A., & Manke, T. (2014). DeepTools: A

flexible platform for exploring deep-sequencing data. *Nucleic Acids Research*, *42*(W1). https://doi.org/10.1093/nar/gku365

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y., & Zhang, F. (2013). XDouble nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. *Cell*, *154*(6), 1380–1389. https://doi.org/10.1016/j.cell.2013.08.021

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143

Ranes, M., Boeing, S., Wang, Y., Wienholz, F., Menoni, H., Walker, J., Encheva, V., Chakravarty, P., Mari, P. O., Stewart, A., Giglia-Mari, G., Snijders, A. P., Vermeulen, W., & Svejstrup, J. Q. (2016). A ubiquitylation site in Cockayne syndrome B required for repair of oxidative DNA damage, but not for transcription-coupled nucleotide excision repair. *Nucleic Acids Research*, *44*(11), 5246–5255. https://doi.org/10.1093/nar/gkw216

Rasmussen, R. E., & Painter, R. B. (1964). Evidence for repair of ultra-violet damaged deoxyribonucleic acid in cultured mammalian cells. *Nature*, *203*(4952), 1360–1362. https://doi.org/10.1038/2031360a0

Ratner, J. N., Balasubramanian, B., Corden, J., Warren, S. L., & Bregman, D. B. (1998). Ultraviolet radiation-induced ubiquitination and proteasomal degradation of the large subunit of RNA polymerase II: Implications for transcription- coupled DNA repair. *Journal of Biological Chemistry*, *273*(9), 5184–5189. https://doi.org/10.1074/jbc.273.9.5184

Reid, J., & Svejstrup, J. Q. (2004). DNA damage-induced Def1-RNA polymerase II interaction and Def1 requirement for polymerase ubiquitylation in vitro. *Journal of Biological Chemistry*, *279*(29), 29875–29878. https://doi.org/10.1074/jbc.C400185200

Reyes-Turcu, F. E., Horton, J. R., Mullally, J. E., Heroux, A., Cheng, X., & Wilkinson, K. D. (2006). The Ubiquitin Binding Domain ZnF UBP Recognizes the C-Terminal Diglycine Motif of Unanchored Ubiquitin. *Cell*, *124*(6), 1197–1208. https://doi.org/10.1016/j.cell.2006.02.038

Rice, G. A., Chamberlin, M. J., & Kane, C. M. (1993). Contacts between mammalian RNA polymerase II and the template DNA in a ternary elongation complex. *Nucleic Acids Research*, *21*(1), 113–118. https://doi.org/10.1093/nar/21.1.113

Rochette, P. J., Therrien, J. P., Drouin, R., Perdiz, D., Bastien, N., Drobetsky, E. A., & Sage, E. (2003). UVA-induced cyclobutane pyrimidine dimers form predominantly at thymine-thymine dipyrimidines and correlate with the mutation spectrum in rodent cells. *Nucleic Acids Research*, *31*(11), 2786–2794. https://doi.org/10.1093/nar/gkg402

Russell, S. J., Reed, S. H., Huang, W., Friedberg, E. C., & Johnston, S. A. (1999). The 19S regulatory complex of the proteasome functions independently of proteolysis in nucleotide excision repair. *Molecular Cell*, *3*(6), 687–695. https://doi.org/10.1016/S1097-2765(01)80001-0

Saeki, H., & Svejstrup, J. Q. (2009). Stability, Flexibility, and Dynamic Interactions of Colliding RNA Polymerase II Elongation Complexes. *Molecular Cell*, *35*(2), 191–205. https://doi.org/10.1016/j.molcel.2009.06.009

Sancar, A., Kacinski, B. M., Mott, D. L., & Rupp, W. D. (1981). Identification of the uvrC gene product. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(9 II), 5450–5454. https://doi.org/10.1073/pnas.78.9.5450

Sancar, Aziz, & Reardon, J. T. (2004). Nucleotide excision repair in E. coli and man. *Advances in Protein Chemistry*, *69*, 43–71. https://doi.org/10.1016/S0065-3233(04)69002-4

Sancar, Aziz, & Rupp, W. D. (1983). A novel repair enzyme: UVRABC excision

nuclease of Escherichia coli cuts a DNA strand on both sides of the damaged region. *Cell*, *33*(1), 249–260. https://doi.org/10.1016/0092-8674(83)90354-9

Santonico, E. (2019). New Insights into the Mechanisms Underlying NEDD8 Structural and Functional Specificities. I: *Ubiquitin Proteasome System - Current Insights into Mechanism Cellular Regulation and Disease*. IntechOpen. https://doi.org/10.5772/intechopen.83426

Sarker, A. H., Tsutakawa, S. E., Kostek, S., Ng, C., Shin, D. S., Peris, M., Campeau, E., Tainer, J. A., Nogales, E., & Cooper, P. K. (2005). Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: Insights for transcription-coupled repair and Cockayne syndrome. *Molecular Cell*, *20*(2), 187–198. https://doi.org/10.1016/j.molcel.2005.09.022

Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J. H. J., Chambon, P., & Egly, J. M. (1993). DNA repair helicase: A component of BTF2 (TFIIH) basic transcription factor. *Science*, *260*(5104), 58–63. https://doi.org/10.1126/science.8465201

Scheffner, M., Nuber, U., & Huibregtse, J. M. (1995). Protein ubiquitination involving an E1–E2–E3 enzyme ubiquitin thioester cascade. *Nature*, *373*(6509), 81–83. https://doi.org/10.1038/373081a0

Scheibye-Knudsen, M., Croteau, D. L., & Bohr, V. A. (2013). Mitochondrial deficiency in Cockayne syndrome. *Mechanisms of Ageing and Development*, *134*(5–6), 275–283. https://doi.org/10.1016/j.mad.2013.02.007

Schlake, T., & Bode, J. (1994). Use of Mutated FLP Recognition Target (FRT) Sites for the Exchange of Expression Cassettes at Defined Chromosomal Loci. *Biochemistry*, *33*(43), 12746–12751. https://doi.org/10.1021/bi00209a003

Schmickel, R. D., Chu, E. H. Y., Trosko, J. E., & Chang, C. C. (1977). Cockayne syndrome: a cellular sensitivity to ultraviolet light. *Pediatrics*, *60*(2), 135–139.

Schwertman, P., Lagarou, A., Dekkers, D. H. W., Raams, A., Van Der Hoek, A. C., Laffeber, C., Hoeijmakers, J. H. J., Demmers, J. A. A., Fousteri, M., Vermeulen, W., & Marteijn, J. A. (2012). UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nature Genetics*, *44*(5), 598–602. https://doi.org/10.1038/ng.2230

Schwertman, P., Vermeulen, W., & Marteijn, J. A. (2013). UVSSA and USP7, a new couple in transcription-coupled DNA repair. I: *Chromosoma* (Vol. 122, Number 4, pp. 275–284). https://doi.org/10.1007/s00412-013-0420-2

Scott, D. C., Sviderskiy, V. O., Monda, J. K., Lydeard, J. R., Cho, S. E., Harper, J. W., & Schulman, B. A. (2014). Structure of a RING E3 trapped in action reveals ligation mechanism for the ubiquitin-like protein NEDD8. *Cell*, *157*(7), 1671–1684. https://doi.org/10.1016/j.cell.2014.04.037

Scrima, A., Koníčková, R., Czyzewski, B. K., Kawasaki, Y., Jeffrey, P. D., Groisman, R., Nakatani, Y., Iwai, S., Pavletich, N. P., & Thomä, N. H. (2008). Structural Basis of UV DNA-Damage Recognition by the DDB1-DDB2 Complex. *Cell*, *135*(7), 1213–1223. https://doi.org/10.1016/j.cell.2008.10.045

Scully, R., Panday, A., Elango, R., & Willis, N. A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. I: *Nature Reviews Molecular Cell Biology* (Vol. 20, Number 11, pp. 698–714). Nature Publishing Group. https://doi.org/10.1038/s41580-019-0152-0

Selby, C. P., & Sancar, A. (1990). Transcription preferentially inhibits nucleotide excision repair of the template DNA strand in vitro. *Journal of Biological Chemistry*, *265*(34), 21330–21336. https://doi.org/10.1016/s0021-9258(17)45364-6

Selby, Christopher P., & Sancar, A. (1991). Gene-and strand-specific repair in vitro: Partial purification of a transcription-repair coupling factor. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(18), 8232–

8236. https://doi.org/10.1073/pnas.88.18.8232
Selby, Christopher P., & Sancar, A. (1997a). Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. *Journal of Biological Chemistry*, *272*(3), 1885–1890. https://doi.org/10.1074/jbc.272.3.1885
Selby, Christopher P., & Sancar, A. (1997b). Cockayne syndrome group B protein enhances elongation by RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(21), 11205–11209. https://doi.org/10.1073/pnas.94.21.11205
Selby, Christopher P., Witkin, E. M., & Sancar, A. (1991). Escherichia coli mfd mutant deficient in 'mutation frequency decline' lacks strand-specific repair: In vitro complementation with purified coupling factor. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(24), 11574–11578. https://doi.org/10.1073/pnas.88.24.11574
Selzer, R. R., Nyaga, S., Tuo, J., May, A., Muftuoglu, M., Christiansen, M., Citterio, E., Brosh, R. M., & Bohr, V. A. (2002). Differential requirement for the ATPase domain of the Cockayne syndrome group B gene in the processing of UV-induced DNA damage and 8-oxoguanine lesions in human cells. I: *Nucleic Acids Research* (Vol. 30, Number 3, pp. 782–793). Oxford University Press. https://doi.org/10.1093/nar/30.3.782
Senecoff, J. F., Bruckner, R. C., & Cox, M. M. (1985). The FLP recombinase of the yeast 2-μm plasmid: Characterization of its recombination site. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(21), 7270–7274. https://doi.org/10.1073/pnas.82.21.7270
Setlow, R. B., & Carrier, W. L. (1964). the Disappearance of Thymine Dimers From Dna: an Error-Correcting. *Proceedings of the National Academy of Sciences of the United States Of*, *51*(2), 226–231. https://doi.org/10.1073/pnas.51.2.226
Setlow, R. B., Regan, J. D., German, J., & Carrier, W. L. (1969). Evidence that xeroderma pigmentosum cells do not perform the first step in the repair of ultraviolet damage to their DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *64*(3), 1035–1041. https://doi.org/10.1073/pnas.64.3.1035
Setlow, R. B., Swenson, P. A., & Carrier, W. L. (1963). Thymine dimers and inhibition of DNA synthesis by ultraviolet irradiation of cells. *Science*, *142*(3598), 1464–1466. https://doi.org/10.1126/science.142.3598.1464
Shaltiel, I. A., Krenning, L., Bruinsma, W., & Medema, R. H. (2015). The same, only different - DNA damage checkpoints and their reversal throughout the cell cycle. I: *Journal of Cell Science* (Vol. 128, Number 4, pp. 607–620). Company of Biologists Ltd. https://doi.org/10.1242/jcs.163766
Shen, T., & Huang, S. (2012). The Role of Cdc25A in the Regulation of Cell Proliferation and Apoptosis. *Anti-Cancer Agents in Medicinal Chemistry*, *12*(6), 631–639. https://doi.org/10.2174/187152012800617678
Shi, J., Wen, A., Zhao, M., Jin, S., You, L., Shi, Y., Dong, S., Hua, X., Zhang, Y., & Feng, Y. (2020). Structural basis of Mfd-dependent transcription termination. *Nucleic Acids Research*, *48*(20), 11762–11772. https://doi.org/10.1093/nar/gkaa904
Shibutani, S., Takeshita, M., & Grollman, A. P. (1991). Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature*, *349*(6308), 431–434. https://doi.org/10.1038/349431a0
Shih, S. C., Prag, G., Francis, S. A., Sutanto, M. A., Hurley, J. H., & Hicke, L. (2003). A ubiquitin-binding motif required for intramolecular monoubiquitylation, the CUE domain. *EMBO Journal*, *22*(6), 1273–1281. https://doi.org/10.1093/emboj/cdg140
Shiotani, B., & Zou, L. (2009). Single-Stranded DNA Orchestrates an ATM-to-ATR

Switch at DNA Breaks. *Molecular Cell*, *33*(5), 547–558. https://doi.org/10.1016/j.molcel.2009.01.024

Shivji, M. K. K., Eker, A. P. M., & Wood, R. D. (1994). DNA repair defect in xeroderma pigmentosum group C and complementing factor from HeLa cells. *Journal of Biological Chemistry*, *269*(36), 22749–22757. https://doi.org/10.1016/s0021-9258(17)31709-x

Shrestha, R. K., Ronau, J. A., Davies, C. W., Guenette, R. G., Strieter, E. R., Paul, L. N., & Das, C. (2014). Insights into the mechanism of deubiquitination by jamm deubiquitinases from cocrystal structures of the enzyme with the substrate and product. *Biochemistry*, *53*(19), 3199–3217. https://doi.org/10.1021/bi5003162

Sigurdsson, S., Dirac-Svejstrup, A. B., & Svejstrup, J. Q. (2010). Evidence that Transcript Cleavage Is Essential for RNA Polymerase II Transcription and Cell Viability. *Molecular Cell*, *38*(2), 202–210. https://doi.org/10.1016/j.molcel.2010.02.026

Sin, Y., Tanaka, K., & Saijo, M. (2016). The C-terminal region and SUMOylation of cockayne syndrome group B protein play critical roles in transcription-coupled nucleotide excision repair. *Journal of Biological Chemistry*, *291*(3), 1387–1397. https://doi.org/10.1074/jbc.M115.683235

Singleton, B. K., Torres-Arzayus, M. I., Rottinghaus, S. T., Taccioli, G. E., & Jeggo, P. A. (1999). The C Terminus of Ku80 Activates the DNA-Dependent Protein Kinase Catalytic Subunit. *Molecular and Cellular Biology*, *19*(5), 3267–3277. https://doi.org/10.1128/mcb.19.5.3267

Sobhian, B., Shao, G., Lilli, D. R., Culhane, A. C., Moreau, L. A., Xia, B., Livingston, D. M., & Greenberg, R. A. (2007). RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science*, *316*(5828), 1198–1202. https://doi.org/10.1126/science.1139516

Somesh, B. P., Reid, J., Liu, W. F., Søgaard, T. M. M., Erdjument-Bromage, H., Tempst, P., & Svejstrup, J. Q. (2005). Multiple mechanisms confining RNA polymerase II ubiquitylation to polymerases undergoing transcriptional arrest. *Cell*, *121*(6), 913–923. https://doi.org/10.1016/j.cell.2005.04.010

Somesh, B. P., Sigurdsson, S., Saeki, H., Erdjument-Bromage, H., Tempst, P., & Svejstrup, J. Q. (2007). Communication between Distant Sites in RNA Polymerase II through Ubiquitylation Factors and the Polymerase CTD. *Cell*, *129*(1), 57–68. https://doi.org/10.1016/j.cell.2007.01.046

Song, J., Durrin, L. K., Wilkinson, T. A., Krontiris, T. G., & Chen, Y. (2004). Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(40), 14373–14378. https://doi.org/10.1073/pnas.0403498101

Sopik, V., Phelan, C., Cybulski, C., & Narod, S. A. (2015). BRCA1 and BRCA2 mutations and the risk for colorectal cancer. *Clinical Genetics*, *87*(5), 411–418. https://doi.org/10.1111/cge.12497

Soutoglou, E., & Misteli, T. (2008). Activation of the cellular DNA damage response in the absence of DNA lesions. *Science*, *320*(5882), 1507–1510. https://doi.org/10.1126/science.1159051

Staresincic, L., Fagbemi, A. F., Enzlin, J. H., Gourdin, A. M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S. G., Vermeulen, W., & Schärer, O. D. (2009). Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO Journal*, *28*(8), 1111–1120. https://doi.org/10.1038/emboj.2009.49

Stinson, B. M., Moreno, A. T., Walter, J. C., & Loparo, J. J. (2020). A Mechanism to Minimize Errors during Non-homologous End Joining. *Molecular Cell*, *77*(5), 1080-1091.e8. https://doi.org/10.1016/j.molcel.2019.11.018

Sugasawa, K., Akagi, J. ichi, Nishi, R., Iwai, S., & Hanaoka, F. (2009). Two-Step

Recognition of DNA Damage for Mammalian Nucleotide Excision Repair: Directional Binding of the XPC Complex and DNA Strand Scanning. *Molecular Cell*, *36*(4), 642–653. https://doi.org/10.1016/j.molcel.2009.09.035

Sugasawa, K., Ng, J. M. Y., Masutani, C., Iwai, S., Van Der Spek, P. J., Eker, A. P. M., Hanaoka, F., Bootsma, D., & Hoeijmakers, J. H. J. (1998). Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Molecular Cell*, *2*(2), 223–232. https://doi.org/10.1016/S1097-2765(00)80132-X

Sugasawa, K., Okuda, Y., Saijo, M., Nishi, R., Matsuda, N., Chu, G., Mori, T., Iwai, S., Tanaka, K., Tanaka, K., & Hanaoka, F. (2005). UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell*, *121*(3), 387–400. https://doi.org/10.1016/j.cell.2005.02.035

Sun, N., Youle, R. J., & Finkel, T. (2016). The Mitochondrial Basis of Aging. I: *Molecular Cell* (Vol. 61, Number 5, pp. 654–666). Cell Press. https://doi.org/10.1016/j.molcel.2016.01.028

Svejstrup, J. Q., Wang, Z., Feave, W. J., Wu, X., Bushnell, D. A., Donahue, T. F., Friedberg, E. C., & Kornberg, R. D. (1995). Different forms of TFIIH for transcription and DNA repair: Holo-TFIIH and a nucleotide excision repairosome. *Cell*, *80*(1), 21–28. https://doi.org/10.1016/0092-8674(95)90447-6

Svetlov, V., & Nudler, E. (2013). Basic mechanism of transcription by RNA polymerase II. I: *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* (Vol. 1829, Number 1, pp. 20–28). Biochim Biophys Acta. https://doi.org/10.1016/j.bbagrm.2012.08.009

Svoboda, D. L., Taylor, J. S., Hearst, J. E., & Sancar, A. (1993). DNA repair by eukaryotic nucleotide excision nuclease. Removal of thymine dimer and psoralen monoadduct by HeLa cell-free extract and of thymine dimer by Xenopus laevis oocytes. *Journal of Biological Chemistry*, *268*(3), 1931–1936. https://doi.org/10.1016/s0021-9258(18)53943-0

Sweder, K. S., & Hanawalt, P. C. (1992). Preferential repair of cyclobutane pyrimidine dimers in the transcribed strand of a gene in yeast chromosomes and plasmids is dependent on transcription. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(22), 10696–10700. https://doi.org/10.1073/pnas.89.22.10696

Symington, L. S., & Gautier, J. (2011). Double-strand break end resection and repair pathway choice. *Annual Review of Genetics*, *45*, 247–271. https://doi.org/10.1146/annurev-genet-110410-132435

Takahashi, T. S., Sato, Y., Yamagata, A., Goto-Ito, S., Saijo, M., & Fukai, S. (2019). Structural basis of ubiquitin recognition by the winged-helix domain of Cockayne syndrome group B protein. *Nucleic Acids Research*, *47*(7), 3784–3794. https://doi.org/10.1093/nar/gkz081

Tantin, D, Kansal, A., & Carey, M. (1997). Recruitment of the putative transcription-repair coupling factor CSB/ERCC6 to RNA polymerase II elongation complexes. *Molecular and Cellular Biology*, *17*(12), 6803–6814. https://doi.org/10.1128/mcb.17.12.6803

Tantin, Dean. (1998). RNA polymerase II elongation complexes containing the cockayne syndrome group B protein interact with a molecular complex containing the transcription factor IIH components xeroderma pigmentosum B and p62. *Journal of Biological Chemistry*, *273*(43), 27794–27799. https://doi.org/10.1074/jbc.273.43.27794

Tapias, A., Auriol, J., Forget, D., Enzlin, J. H., Schärer, O. D., Coin, F., Coulombe, B., & Egly, J. M. (2004). Ordered Conformational Changes in Damaged DNA Induced by Nucleotide Excision Repair Factors. *Journal of Biological Chemistry*, *279*(18), 19074–19083. https://doi.org/10.1074/jbc.M312611200

Tatham, M. H., Geoffroy, M. C., Shen, L., Plechanovova, A., Hattersley, N., Jaffray, E. G., Palvimo, J. J., & Hay, R. T. (2008). RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nature Cell Biology*, *10*(5), 538–546. https://doi.org/10.1038/ncb1716

Thomas, D. C., Levy, M., & Sancar, A. (1985). Amplification and purification of UvrA, UvrB, and UvrC proteins of Escherichia coli. *Journal of Biological Chemistry*, *260*(17), 9875–9883. https://doi.org/10.1016/s0021-9258(17)39318-3

Tijsterman, M., Verhage, R. A., Van De Putte, P., Tasseron-De Jong, J. G., & Brouwer, J. (1997). Transitions in the coupling of transcription and nucleotide excision repair within RNA polymerase II-transcribed genes of Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(15), 8027–8032. https://doi.org/10.1073/pnas.94.15.8027

Toledo, L. I., Altmeyer, M., Rask, M. B., Lukas, C., Larsen, D. H., Povlsen, L. K., Bekker-Jensen, S., Mailand, N., Bartek, J., & Lukas, J. (2013). XATR prohibits replication catastrophe by preventing global exhaustion of RPA. *Cell*, *155*(5), 1088. https://doi.org/10.1016/j.cell.2013.10.043

Troelstra, C, Odijk, H., de Wit, J., Westerveld, A., Thompson, L. H., Bootsma, D., & Hoeijmakers, J. H. (1990). Molecular cloning of the human DNA excision repair gene ERCC-6. *Molecular and Cellular Biology*, *10*(11), 5806–5813. https://doi.org/10.1128/mcb.10.11.5806

Troelstra, Christine, Hesen, W., Bootsma, D., & Hoeijmakers, J. H. J. (1993). Structure and expression of the excision repair gene ERCC6, involved in the human disorder cockayne's syndrome group B. *Nucleic Acids Research*, *21*(3), 419–426. https://doi.org/10.1093/nar/21.3.419

Troelstra, Christine, van Gool, A., de Wit, J., Vermeulen, W., Bootsma, D., & Hoeijmakers, J. H. J. (1992). ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. *Cell*, *71*(6), 939–953. https://doi.org/10.1016/0092-8674(92)90390-X

Tu, Y., Bates, S., & Pfeifer, G. P. (1997). Sequence-specific and domain-specific DNA repair in xeroderma pigmentosum and Cockayne syndrome cells. *Journal of Biological Chemistry*, *272*(33), 20747–20755. https://doi.org/10.1074/jbc.272.33.20747

Tufegdzic Vidakovic, A., Harreman, M., Dirac-Svejstrup, A. B., Boeing, S., Roy, A., Encheva, V., Neumann, M., Wilson, M., Snijders, A. P., & Svejstrup, J. Q. (2019). Analysis of RNA polymerase II ubiquitylation and proteasomal degradation. *Methods*, *159–160*, 146–156. https://doi.org/10.1016/j.ymeth.2019.02.005

Tufegdžić Vidaković, A., Mitter, R., Kelly, G. P., Neumann, M., Harreman, M., Rodríguez-Martínez, M., Herlihy, A., Weems, J. C., Boeing, S., Encheva, V., Gaul, L., Milligan, L., Tollervey, D., Conaway, R. C., Conaway, J. W., Snijders, A. P., Stewart, A., & Svejstrup, J. Q. (2020). Regulation of the RNAPII Pool Is Integral to the DNA Damage Response. *Cell*, *180*(6), 1245-1261.e21. https://doi.org/10.1016/j.cell.2020.02.009

Tuo, J., Chen, C., Zeng, X., Christiansen, M., & Bohr, V. A. (2002). Functional crosstalk between hOgg1 and the helicase domain of Cockayne syndrome group B protein. *DNA Repair*, *1*(11), 913–927. https://doi.org/10.1016/S1568-7864(02)00116-7

Tuo, J., Jaruga, P., Rodriguez, H., Bohr, V. A., & Dizdaroglu, M. (2003). Primary fibroblasts of Cockayne syndrome patients are defective in cellular repair of 8-hydroxyguanine and 8-hydroxyadenine resulting from oxidative stress. *The FASEB Journal*, *17*(6), 668–674. https://doi.org/10.1096/fj.02-0851com

Van Cuijk, L., Van Belle, G. J., Turkyilmaz, Y., Poulsen, S. L., Janssens, R. C., Theil, A. F., Sabatella, M., Lans, H., Mailand, N., Houtsmuller, A. B., Vermeulen, W., & Marteijn, J. A. (2015). SUMO and ubiquitin-dependent XPC exchange drives nucleotide excision repair. *Nature Communications*, *6*(1), 1–10.

https://doi.org/10.1038/ncomms8499

Van Den Boom, V., Citterio, E., Hoogstraten, D., Zotter, A., Egly, J. M., Van Cappellen, W. A., Hoeijmakers, J. H. J., Houtsmuller, A. B., & Vermeulen, W. (2004). DNA damage stabilizes interaction of CSB with the transcription elongation machinery. *Journal of Cell Biology*, *166*(1), 27–36. https://doi.org/10.1083/jcb.200401056

Van Der Veen, A. G., & Ploegh, H. L. (2012). Ubiquitin-like proteins. *Annual Review of Biochemistry*, *81*(1), 323–357. https://doi.org/10.1146/annurev-biochem-093010-153308

van der Weegen, Y., Golan-Berman, H., Mevissen, T. E. T., Apelt, K., González-Prieto, R., Goedhart, J., Heilbrun, E. E., Vertegaal, A. C. O., van den Heuvel, D., Walter, J. C., Adar, S., & Luijsterburg, M. S. (2020). The cooperative action of CSB, CSA, and UVSSA target TFIIH to DNA damage-stalled RNA polymerase II. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15903-8

Van Gool, A. J., Citterio, E., Rademakers, S., Van Os, R., Vermeulen, W., Constantinou, A., Egly, J. M., Bootsma, D., & Hoeijmakers, J. H. J. (1997). The Cockayne syndrome B protein, involved in transcription-coupled DNA repair, resides in an RNA polymerase II-containing complex. *EMBO Journal*, *16*(19), 5955–5965. https://doi.org/10.1093/emboj/16.19.5955

Van Gool, A. J., Verhage, R., Swagemakers, S. M. A., Van De Putte, P., Brouwer, J., Troelstra, C., Bootsma, D., & Hoeijmakers, J. H. J. (1994). RAD26, the functional S.cerevisiae homolog of the cockayne syndrome B gene ERCC6. *EMBO Journal*, *13*(22), 5361–5369. https://doi.org/10.1002/j.1460-2075.1994.tb06871.x

Van Hoffen, A., Natarajan, A. T., Mayne, L. V., Zeeland, A. A. va., Mullenders, L. H. F., & Venema, J. (1993). Deficient repair of the transcribed strand of active genes in cockayne's syndrome cells. *Nucleic Acids Research*, *21*(25), 5890–5895. https://doi.org/10.1093/nar/21.25.5890

Van Houten, B. (1990). Nucleotide excision repair in Escherichia coli. I: *Microbiological Reviews* (Vol. 54, Number 1, pp. 18–51). https://doi.org/10.1128/mmbr.54.1.18-51.1990

Vélez-Cruz, R., Zadorin, A. S., Coin, F., & Egly, J. M. (2013). Sirt1 suppresses RNA synthesis after UV irradiation in combined xeroderma pigmentosum group D/Cockayne syndrome (XP-D/CS) cells. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(3). https://doi.org/10.1073/pnas.1213076110

Velmurugu, Y., Chen, X., Sevilla, P. S., Min, J. H., & Ansari, A. (2016). Twist-open mechanism of DNA damage recognition by the Rad4/XPC nucleotide excision repair complex. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(16), E2296–E2305. https://doi.org/10.1073/pnas.1514666113

Venema, B., Mullenders, L. H. F., Natarajan, A. T., Van Zeeland, A. A., & Mayne, L. V. (1990). The genetic defect in cockayne syndrome is associated with a defect in repair of UV-induced DNA damage in transcriptionally active DNA. I: *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 87, Number 12). https://doi.org/10.1073/pnas.87.12.4707

Venema, J, van Hoffen, A., Karcagi, V., Natarajan, A. T., van Zeeland, A. A., & Mullenders, L. H. (1991). Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active genes. *Molecular and Cellular Biology*, *11*(8), 4128–4134. https://doi.org/10.1128/mcb.11.8.4128

Venema, Jaap, van Hoffen, A., Natarajan, A. T., van Zeeland, A. A., & Mullenders, L. H. F. (1990). The residual repair capacity of xeroderma pigmentosum complementation group C fibroblasts is highly specific for transcriptionally active DNA. *Nucleic Acids Research*, *18*(3), 443–448.

https://doi.org/10.1093/nar/18.3.443
Vignard, J., Mirey, G., & Salles, B. (2013). Ionizing-radiation induced DNA double-strand breaks: A direct and indirect lighting up. I: *Radiotherapy and Oncology* (Vol. 108, Number 3, pp. 362–369). Radiother Oncol. https://doi.org/10.1016/j.radonc.2013.06.013
Voges, D., Zwickl, P., & Baumeister, W. (1999). The 26S proteasome: A molecular machine designed for controlled proteolysis. I: *Annual Review of Biochemistry* (Vol. 68, pp. 1015–1068). https://doi.org/10.1146/annurev.biochem.68.1.1015
Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G. A., Winston, F., Buratowski, S., & Handa, H. (1998). *DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs*. Genes and Development. https://doi.org/10.1101/gad.12.3.343
Wagner, S. A., Beli, P., Weinert, B. T., Nielsen, M. L., Cox, J., Mann, M., & Choudhary, C. (2011). A Proteome-wide, Quantitative Survey of In Vivo Ubiquitylation Sites Reveals Widespread Regulatory Roles. *Molecular & Cellular Proteomics*, *10*(10), M111.013284. https://doi.org/10.1074/mcp.m111.013284
Wakasugi, M., Reardon, J. T., & Sancar, A. (1997). The Non-catalytic function of XPG protein during dual incision in human nucleotide excision repair. *Journal of Biological Chemistry*, *272*(25), 16030–16034. https://doi.org/10.1074/jbc.272.25.16030
Wakasugi, M., Shimizu, M., Morioka, H., Linn, S., Nikaido, O., & Matsunaga, T. (2001). Damaged DNA-binding Protein DDB Stimulates the Excision of Cyclobutane Pyrimidine Dimers in Vitro in Concert with XPA and Replication Protein A. *Journal of Biological Chemistry*, *276*(18), 15434–15440. https://doi.org/10.1074/jbc.M011177200
Wang, D., & Lippard, S. J. (2005). Cellular processing of platinum anticancer drugs. I: *Nature Reviews Drug Discovery* (Vol. 4, Number 4, pp. 307–320). Nature Publishing Group. https://doi.org/10.1038/nrd1691
Wang, L., Limbo, O., Fei, J., Chen, L., Kim, B., Luo, J., Chong, J., Conaway, R. C., Conaway, J. W., Ranish, J. A., Kadonaga, J. T., Russell, P., & Wang, D. (2014). Regulation of the Rhp26ERCC6/CSB chromatin remodeler by a novel conserved leucine latch motif. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(52), 18566–18571. https://doi.org/10.1073/pnas.1420227112
Wang, Q. E., Zhu, Q., Wani, G., El-Mahdy, M. A., Li, J., & Wani, A. A. (2005). DNA repair factor XPC is modified by SUMO-1 and ubiquitin following UV irradiation. *Nucleic Acids Research*, *33*(13), 4023–4034. https://doi.org/10.1093/nar/gki684
Wang, W., Xu, J., Chong, J., & Wang, D. (2018). Structural basis of DNA lesion recognition for eukaryotic transcription-coupled nucleotide excision repair. I: *DNA Repair* (Vol. 71, pp. 43–55). Elsevier B.V. https://doi.org/10.1016/j.dnarep.2018.08.006
Wang, Y., Chakravarty, P., Ranes, M., Kelly, G., Brooks, P. J., Neilan, E., Stewart, A., Schiavo, G., & Svejstrup, J. Q. (2014). Dysregulation of gene expression as a cause of cockayne syndrome neurological disease. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(40), 14454–14459. https://doi.org/10.1073/pnas.1412569111
Wang, Y., Jones-Tabah, J., Chakravarty, P., Stewart, A., Muotri, A., Laposa, R. R., & Svejstrup, J. Q. (2016). Pharmacological Bypass of Cockayne Syndrome B Function in Neuronal Differentiation. *Cell Reports*, *14*(11), 2554–2561. https://doi.org/10.1016/j.celrep.2016.02.051
Waters, L. S., Minesinger, B. K., Wiltrout, M. E., D'Souza, S., Woodruff, R. V., & Walker, G. C. (2009). Eukaryotic Translesion Polymerases and Their Roles and

Regulation in DNA Damage Tolerance. *Microbiology and Molecular Biology Reviews*, *73*(1), 134–154. https://doi.org/10.1128/mmbr.00034-08

Watson, I. R., Irwin, M. S., & Ohh, M. (2011). NEDD8 Pathways in Cancer, Sine Quibus Non. I: *Cancer Cell* (Vol. 19, Number 2, pp. 168–176). https://doi.org/10.1016/j.ccr.2011.01.002

Weber, A. M., & Ryan, A. J. (2015). ATM and ATR as therapeutic targets in cancer. I: *Pharmacology and Therapeutics* (Vol. 149, pp. 124–138). Elsevier Inc. https://doi.org/10.1016/j.pharmthera.2014.12.001

Weinfeld, M., Mani, R. S., Abdou, I., Aceytuno, R. D., & Glover, J. N. M. (2011). Tidying up loose ends: The role of polynucleotide kinase/phosphatase in DNA strand break repair. I: *Trends in Biochemical Sciences* (Vol. 36, Number 5, pp. 262–271). Trends Biochem Sci. https://doi.org/10.1016/j.tibs.2011.01.006

Wenzel, D. M., Lissounov, A., Brzovic, P. S., & Klevit, R. E. (2011). UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature*, *474*(7349), 105–108. https://doi.org/10.1038/nature09966

Wijk, S. J. L., & Timmers, H. T. M. (2010). The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. *The FASEB Journal*, *24*(4), 981–993. https://doi.org/10.1096/fj.09-136259

Wilkinson, K. A., & Henley, J. M. (2010). Mechanisms, regulation and consequences of protein SUMOylation. I: *Biochemical Journal* (Vol. 428, Number 2, pp. 133–145). https://doi.org/10.1042/BJ20100158

Wilkinson, K. D., Urban, M. K., & Haas, A. L. (1980). Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *Journal of Biological Chemistry*, *255*(16), 7529–7532. https://doi.org/10.1016/s0021-9258(19)43857-x

Wilkinson, Keith D. (1999). Ubiquitin-dependent signaling: The role of ubiquitination in the response of cells to their environment. I: *Journal of Nutrition* (Vol. 129, Number 11, pp. 1933–1936). American Institute of Nutrition. https://doi.org/10.1093/jn/129.11.1933

Williamson, L., Saponaro, M., Boeing, S., East, P., Mitter, R., Kantidakis, T., Kelly, G. P., Lobley, A., Walker, J., Spencer-Dene, B., Howell, M., Stewart, A., & Svejstrup, J. Q. (2017). UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene. *Cell*, *168*(5), 843-855.e13. https://doi.org/10.1016/j.cell.2017.01.019

Wilson, B. T., Stark, Z., Sutton, R. E., Danda, S., Ekbote, A. V., Elsayed, S. M., Gibson, L., Goodship, J. A., Jackson, A. P., Keng, W. T., King, M. D., McCann, E., Motojima, T., Murray, J. E., Omata, T., Pilz, D., Pope, K., Sugita, K., White, S. M., & Wilson, I. J. (2016). The Cockayne Syndrome Natural History (CoSyNH) study: Clinical findings in 102 individuals and recommendations for care. *Genetics in Medicine*, *18*(5), 483–493. https://doi.org/10.1038/gim.2015.110

Wilson, M. D., Harreman, M., & Svejstrup, J. Q. (2013). Ubiquitylation and degradation of elongating RNA polymerase II: The last resort. I: *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* (Vol. 1829, Number 1, pp. 151–157). https://doi.org/10.1016/j.bbagrm.2012.08.002

Wilson, M. D., Saponaro, M., Leidl, M. A., & Svejstrup, J. Q. (2012). MultiDsk: A Ubiquitin-Specific Affinity Resin. *PLoS ONE*, *7*(10), 46398. https://doi.org/10.1371/journal.pone.0046398

Witkin, E. M. (1966). Radiation-induced mutations and their repair. *Science*, *152*(3727), 1345–1353. https://doi.org/10.1126/science.152.3727.1345

Wittschieben, B., Iwai, S., & Wood, R. D. (2005). DDB1-DDB2 (xeroderma pigmentosum group E) protein complex recognizes a cyclobutane pyrimidine dimer, mismatches, apurinic/apyrimidinic sites, and compound in DNA. *Journal of Biological Chemistry*, *280*(48), 39982–39989. https://doi.org/10.1074/jbc.M507854200

Wood, R. D. (1997). Nucleotide excision repair in mammalian cells. I: *Journal of Biological Chemistry* (Vol. 272, Number 38, pp. 23465–23468). American Society for Biochemistry and Molecular Biology. https://doi.org/10.1074/jbc.272.38.23465

Wood, R. D., Robins, P., & Lindahl, T. (1988). Complementation of the xeroderma pigmentosum DNA repair defect in cell-free extracts. *Cell*, *53*(1), 97–106. https://doi.org/10.1016/0092-8674(88)90491-6

Woudstra, E. C., Gilbert, C., Fellows, J., Jansen, L., Brouwer, J., Erdjument-Bromage, H., Tempst, P., & Svejstrup, J. Q. (2002). A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature*, *415*(6874), 929–933. https://doi.org/10.1038/415929a

Xu, G., Paige, J. S., & Jaffrey, S. R. (2010). Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nature Biotechnology*, *28*(8), 868–873. https://doi.org/10.1038/nbt.1654

Xu, J., Lahiri, I., Wang, W., Wier, A., Cianfrocco, M. A., Chong, J., Hare, A. A., Dervan, P. B., DiMaio, F., Leschziner, A. E., & Wang, D. (2017). Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature*, *551*(7682), 653–657. https://doi.org/10.1038/nature24658

Xu, J., Wang, W., Xu, L., Chen, J. Y., Chong, J., Oh, J., Leschziner, A. E., Fu, X. D., & Wang, D. (2020). Cockayne syndrome B protein acts as an ATP-dependent processivity factor that helps RNA polymerase II overcome nucleosome barriers. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(41), 25486–25493. https://doi.org/10.1073/pnas.2013379117

Yang, X., Lykke-Andersen, K., Tsuge, T., Xiao, D., Wang, X., Rodriguez-Suarez, R. J., Zhang, H., & Wei, N. (2002). The COP9 signalosome inhibits p27kip1 degradation and impedes G1-S phase progression via deneddylation of SCF Cul1. *Current Biology*, *12*(8), 667–672. https://doi.org/10.1016/S0960-9822(02)00791-1

Yasukawa, T., Kamura, T., Kitajima, S., Conaway, R. C., Conaway, J. W., & Aso, T. (2008). Mammalian Elongin A complex mediates DNA-damage-induced ubiquitylation and degradation of Rpb1. *EMBO Journal*, *27*(24), 3256–3266. https://doi.org/10.1038/emboj.2008.249

Yau, R., & Rape, M. (2016). The increasing complexity of the ubiquitin code. I: *Nature Cell Biology* (Vol. 18, Number 6, pp. 579–586). Nature Publishing Group. https://doi.org/10.1038/ncb3358

Yeung, A. T., Mattes, W. B., Oh, E. Y., & Grossman, L. (1983). Enzymatic properties of purified Escherichia coli uvrABC proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(20 I), 6157–6161. https://doi.org/10.1073/pnas.80.20.6157

Yokoi, M., & Hanaoka, F. (2017). Two mammalian homologs of yeast Rad23, HR23A and HR23B, as multifunctional proteins. I: *Gene* (Vol. 597, pp. 1–9). Elsevier B.V. https://doi.org/10.1016/j.gene.2016.10.027

Yokoi, M., Masutani, C., Maekawa, T., Sugasawa, K., Ohkuma, Y., & Hanaoka, F. (2000). The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA. *Journal of Biological Chemistry*, *275*(13), 9870–9875. https://doi.org/10.1074/jbc.275.13.9870

You, Z., Bailis, J. M., Johnson, S. A., Dilworth, S. M., & Hunter, T. (2007). Rapid activation of ATM on DNA flanking double-strand breaks. *Nature Cell Biology*, *9*(11), 1311–1318. https://doi.org/10.1038/ncb1651

You, Z., Kong, L., & Newport, J. (2002). The role of single-stranded DNA and polymerase α in establishing the ATR, Hus1 DNA replication checkpoint. *Journal of Biological Chemistry*, *277*(30), 27088–27093. https://doi.org/10.1074/jbc.M204120200

Zatreanu, D., Han, Z., Mitter, R., Tumini, E., Williams, H., Gregersen, L., Dirac-

Svejstrup, A. B., Roma, S., Stewart, A., Aguilera, A., & Svejstrup, J. Q. (2019). Elongation Factor TFIIS Prevents Transcription Stress and R-Loop Accumulation to Maintain Genome Stability. *Molecular Cell*, *76*(1), 57-69.e9. https://doi.org/10.1016/j.molcel.2019.07.037

Zhang, X., Horibata, K., Saijo, M., Ishigami, C., Ukai, A., Kanno, S. I., Tahara, H., Neilan, E. G., Honma, M., Nohmi, T., Yasui, A., & Tanaka, K. (2012). Mutations in UVSSA cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair. *Nature Genetics*, *44*(5), 593–597. https://doi.org/10.1038/ng.2228

Zhao, W., Steinfeld, J. B., Liang, F., Chen, X., Maranon, D. G., Jian Ma, C., Kwon, Y., Rao, T., Wang, W., Sheng, C., Song, X., Deng, Y., Jimenez-Sainz, J., Lu, L., Jensen, R. B., Xiong, Y., Kupfer, G. M., Wiese, C., Greene, E. C., & Sung, P. (2017). BRCA1-BARD1 promotes RAD51-mediated homologous DNA pairing. *Nature*, *550*(7676). https://doi.org/10.1038/nature24060

Zhou, H., Wertz, I., O'Rourke, K., Ultsch, M., Seshagir, S., Eby, M., Xiao, W., & Dixit, V. M. (2004). Bcl10 activates the NF-κB pathway through ubiquitination of NEMO. *Nature*, *427*(6970), 167–171. https://doi.org/10.1038/nature02273

Zhou, H. X., & Wang, G. (2001). Predicted structures of two proteins involved in human diseases. *Cell Biochemistry and Biophysics*, *35*(1), 35–47. https://doi.org/10.1385/CBB:35:1:35

Zou, L., & Elledge, S. J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science*, *300*(5625), 1542–1548. https://doi.org/10.1126/science.1083430