



Investigating The Molecular Mechanisms of N-HELQ Interaction With POLD3 and RPA1

A thesis submitted to the University of Nottingham for the
Degree of Master of Research

Alexander Roberts, BSc (Hons)

February 2026

Abstract

Human cells are at risk of DNA damage from a multitude of sources throughout their cell cycle, including during DNA replication. This can lead to DNA double-strand break (DSB) lesions in the genome, from which replication stress can arise, resulting in further DSBs and consequently genomic instability, which can in turn lead to genetic diseases, primarily cancer. Human cells have several DNA damage response pathways to counteract this, including homologous recombination, which comprises multiple sub-pathways. Of these, break-induced replication has the highest mutagenic potential. The DNA helicase HELQ acts to limit resultant mutagenic DNA synthesis through interaction of its N-terminal region N-HELQ with the single-stranded DNA binding protein RPA subunit RPA1 and DNA polymerase Pol δ subunit POLD3. This is of interest to the Bolt lab, with precise molecular mechanisms of these protein-protein interactions (PPI) currently unknown. The work described in this thesis aimed to further the understanding of these mechanisms through identifying POLD3 and RPA1 amino acid residue sites necessary for the PPI.

In order to attempt this, the Bolt lab's existing qualitative bimolecular fluorescence complementation (BiFC) assay time-course method was developed into an endpoint method to provide additional quantitative data. This was followed by an attempt to PE3b prime edit human U2OS osteosarcoma cells to investigate whether previous findings that HELQ^{D142G/F143G} is unable to displace ssDNA-bound RPA *in vitro* could be reproduced *in vivo*. Here it is shown using the developed BiFC assay method that N-HELQ interacts with multiple POLD3 intrinsically disordered protein regions (IDPRs) but is unable to interact with RPA1, hence favouring a native cell investigative approach. The subsequent prime editing was unsuccessful. Further research on this would need to utilise a far-red fluorescent protein in BiFC assays to allow quantitative time-course data, in conjunction with identifying POLD3 IDPR residues of interest in POLD3-N-HELQ PPI and hence generating mutant proteins for use in BiFC assays. Meanwhile, further prime editing would likely need to use an improved system such as PE7 to increase success probability. Together, these approaches can contribute to ultimately understanding the molecular mechanisms of HELQ interaction with Pol δ and RPA.

Acknowledgments

First and foremost, I would like to thank my supervisor Professor Ed Bolt for giving me a place in his lab, and all of his advice, guidance, and help throughout the project. My time in the Bolt lab has allowed me to learn, develop, and hone critical skills and knowledge needed as a scientist.

I would also like to extend my thanks to the members of the Bolt Lab, past and present. PhD students Olivia Downs and Anna Lou-Hing taught me the fundamental microbiology and DNA work methods I needed for my MRes, with particular thanks to Olivia for her help in the prime editing side of things. Additionally, thanks go to postdoc Doctor Tom Killelea, the King of BiFC, whose advice and expertise were instrumental in developing the method, and also to Alice Nottingham for her help. Other thanks go to my fellow MRes students who I had the pleasure of working with this year – Anise Younis, Emily Haskins, and Zuzanna Hartlieb – and to this year's Bolt lab undergraduates. On a wider scale, I would like to express my gratitude to all of university staff who have helped in many ways this year.

Finally, I would like to thank my family for everything they do for me and keeping me going throughout this year and all the challenges it brought.

Abbreviations

AP Site – Apurinic or apyrimidinic site

ATP – Adenosine triphosphate

BER – Base excision repair

BiFC – Bimolecular fluorescence complementation

BIR – Break induced replication

bp – Base pairs

BSI – Biological sequence information

CCDS – Consensus coding sequence

cDNA – Complementary DNA

CFU – Colony forming units

CMG – Cdc45-MCM-GINS

CRISPR – Clustered regularly interspaced short palindromic repeats

CRL – Cullin-RING ubiquitin E3 ligase

crRNA – CRISPR RNA

CTD – C-terminal domain

Cryo-EM – Cryogenic electron microscopy

DBD – DNA-binding domain

dH₂O – Distilled water

DNA – Deoxyribonucleic acid

DMEM – Dulbecco's Modified Eagles Media

DNA-PK_{CS} – DNA-dependent protein kinase catalytic subunit

dNTP – Deoxyribonucleoside triphosphate

DSB – Double-stranded break

DSBR – Double-strand break repair

DUE – DNA unwinding element

EMSA – Electrophoretic mobility shift assays

FBS – Foetal bovine serum

FP – Fluorescent protein

GAN – GINS-associated nuclease

GFP – Green fluorescent protein

GIN – Go-Ichi-Ni-San

HDR – Homology-directed repair

HR – Homologous recombination

IDPR – Intrinsically disordered protein region

Indel – Insertion and deletion

MCM – Minichromosome maintenance

MMR – Mismatch repair

MMEJ – Microhomology-mediated end joining

mRNA – Messenger RNA

NCBI – National Center for Biotechnology Information

NEB – New England Biolabs

NER – Nucleotide excision repair

NHEJ – Non-homologous end joining

NLS – Nuclear localisation signal

NTD – N-terminal domain

NTP – Nucleoside triphosphate

ORB – Origin recognition box

ORC – Origin recognition complex

PBS – Phosphate-buffered saline

PCNA – Proliferating cell nuclear antigen

PCR – Polymerase chain reaction

pegRNA – Prime editing guide RNA

PIP – PCNA-interacting protein

PPI – Protein-protein interaction

PRI – POLD3-interacting region

PRM – Peptide recognition model

RFC – Replication factor C

RNA – Ribonucleic acid

RPA – Replication protein A

SEC – Size-exclusion chromatography

SCF – Skp1-Cullin1-F-box

SD – Standard deviation

SDW – Sterile distilled water

SDM – Site-directed mutagenesis

SF1-6 – Superfamilies 1-6

sgRNA – Single guide RNA

SPR – Surface plasmon resonance

SSDA – Synthesis-dependent strand annealing

SSA – Single-strand annealing

SSB – Single-stranded binding

ssDNA – Single stranded DNA

TC – Tissue culture treated

tracrRNA – Trans-activating CRISPR RNA

TS – Template switching

VBNC – Viable but non-culturable

WHD – Winged-helix domain

WT – Wild type

YFP – Yellow fluorescent protein

Table of Contents

Abstract.....	2
Acknowledgments	3
Abbreviations.....	4
Chapter 1: Introduction	11
1.1 Molecular Biology and Proteins	11
1.1.1 The Sequence Hypothesis and Central Dogma of Molecular Biology	11
1.1.2 Protein Structure and Function	14
1.1.3 Protein-Protein Interaction	17
1.2 DNA Replication	20
1.2.1 The Cell Cycle and DNA Replication Origins.....	20
1.2.2 Initiation and the Replisome	27
1.2.3 Elongation and Termination	32
1.2.4 Replication Stress	37
1.3 DNA Damage and Repair	39
1.3.1 DNA Damage Response Pathways	39
1.3.2 Homologous Recombination	43
1.4 HELQ	47
1.4.1 Discovery, Gene, and Nomenclature	47
1.4.2 Classification	48
1.4.3 Structure and Function.....	51
1.5 Project Aims	58
Chapter 2: Materials and Methods	59
2.1 Materials	59
2.1.1 Bacterial Strains	59
2.1.2 Plasmids	59
2.1.3 Primers	63
2.1.4 Antibiotics	65
2.1.5 Commercial Enzymes	65
2.2 Laboratory Recipes.....	68
2.2.1 Bacterial Growth Media Compositions.....	68
2.2.2 Gel Solutions Compositions.....	69
2.2.3 General Buffers Compositions	70
2.2.4 General Reagents Compositions.....	71
2.3 General Microbiology Protocols	71
2.3.1 Production of Competent <i>E. coli</i> Cells	71

2.3.2 Agar Plate Preparation	72
2.3.3 Transformation of Competent <i>E. coli</i> With Purified Plasmid DNA	72
2.3.4 Growing Overnight Cultures of Transformed <i>E. coli</i>	72
2.3.5 Producing Glycerol Stocks of Transformed <i>E. coli</i>	73
2.4 General Protocols Concerning DNA.....	73
2.4.1 Miniprep Extraction and Purification of Plasmid DNA From Transformed DH5 α <i>E. coli</i> Overnight Cultures	73
2.4.2 Analytical Restriction Digestion of Purified Plasmid DNA	73
2.4.3 Agarose Gel Electrophoresis of DNA Samples	73
2.4.4 Extraction of DNA From Electrophoresed Agarose Gels	74
2.4.5 Polymerase Chain Reaction Amplification From Purified Plasmid DNA	74
2.4.6 Purification of Product DNA From PCR Reaction Mixture.....	75
2.4.7 Site-Directed Mutagenesis By PCR Amplification From Purified Plasmid DNA	76
2.4.8 Restriction Cloning To Generate Recombinant Plasmids	76
2.5 General Protein Assay Protocols	77
2.5.1 Pilot Protein Overexpression	77
2.5.2 SDS-PAGE Analysis of Protein Overexpression Samples	78
2.6 Bimolecular Fluorescence Complementation Assays	79
2.6.1 Endpoint BiFC Assay Using Cotransformed BL21-AI <i>E. coli</i>	79
2.6.2 Test For BL21-AI <i>E. coli</i> Viable But Non-Culturable State	81
2.6.3 Test of BL21-AI <i>E. coli</i> Growth Media Autofluorescence and Effect on Growth Rate	83
2.7 Human Cell Culturing	84
2.7.1 Human Cell Line	84
2.7.2 Human Cells Growth Medium and Supplements	85
2.7.3 Routine Growth and Passaging of Human Cells.....	85
2.8 PE3b Prime Editing of Human Cells	86
2.8.1 Prime Editing Oligonucleotides	86
2.8.2 pegRNA Plasmid Construction By Golden Gate Assembly	87
2.8.3 sgRNA Plasmid Construction	88
2.8.4 Maxiprep Extraction and Purification of Prime Editing Plasmids DNA From Transformed DH5 α <i>E. coli</i> Overnight Cultures	89
2.8.5 Transfection By Lipofection of Human Cells With Prime Editing Plasmids	90
2.8.6 Human Cell Genomic Extraction and Analysis	91
2.9 Molecular Modelling	92
Chapter 3: Analysis of POLD3-N-HELQ Protein-Protein Interaction.....	93
3.1 Introduction	93
3.1.1 POLD3 Structure and Function	93
3.1.2 Bimolecular Fluorescence Complementation Assays	100

3.2 Analysis of POLD3 Quincunx Motif Mutation Impact on POLD3-N-HELQ Protein-Protein Interaction	103
3.2.1 Producing Plasmid Stocks of BiFC Assay Plasmids	103
3.2.2 BiFC Assay	106
3.3 Analysis of POLD3 Intrinsically Disordered Protein Regions and Putative Nuclear Localisation Signal Interaction With N-HELQ	115
3.3.1 BiFC Assay	115
3.3.2 Test For Viable But Non-Culturable State	122
3.4 Analysis of BL21-AI <i>E. coli</i> Growth Media Autofluorescence and Effect on Growth Rate	129
3.5 Summary	135
Chapter 4: Analysis of RPA1-N-HELQ Protein-Protein Interaction	138
4.1 Introduction	138
4.1.1 RPA1 Structure and Function.....	138
4.1.2 PE3b Prime Editing	146
4.2 Analysis of RPA1-N-HELQ Protein-Protein Interaction Using BiFC Assay.....	150
4.2.1 Restriction Cloning of NmVenus-RPA Subunit Plasmids and Pilot Protein Overexpression	150
4.2.2 BiFC Assay	155
4.3 PE3b Prime Editing of <i>HELQ</i> In U2OS Cells	160
4.3.1 PE3b Prime Editing Plasmid Construction	160
4.3.2 U2OS Cell Transfection and Genomic Analysis	162
4.4 Summary	166
Chapter 5: Discussion and Future Work	168
5.1 BiFC Assay Method Development	168
5.2 Characterising POLD3-N-HELQ Interaction Using BiFC Assays	169
5.3 Characterising RPA1-N-HELQ Interaction Using BiFC Assays.....	170
5.4 Attempting To Prime Edit U2OS Cells	171
Chapter 6: References.....	172
Chapter 7: Appendix	235

Chapter 1: Introduction

1.1 Molecular Biology and Proteins

1.1.1 The Sequence Hypothesis and Central Dogma of Molecular Biology

Molecular biology is an interdisciplinary field that aims to understand the biology of life in a cellular context on a molecular level (Cell Editorial Team, 2024; Li et al., 2013; Wrede, 2007) by studying the 'structure and function of molecules, [macromolecules, and their] systems associated with biological processes' (National Library of Medicine, 2019). These are a broad range of processes that primarily fall under the areas of: genetic information, mechanisms, and inheritance; biomolecule synthesis and interaction; and cell chemistry and interaction (Alberts et al., 2022). Due to this, molecular biology has significant overlap with the fields of genetics, biochemistry, biophysics, structural biology, and cell biology (Alberts et al., 2022; Li et al., 2013; National Library of Medicine, 2019; Tsegaye et al., 2021).

Fundamental to molecular biology is biological sequence information (BSI), including genetic information, which in all living organisms on Earth is encoded in deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) by trinucleotide codon sequences (Ao et al., 2022; Barbieri, 2016; Minchin et al., 2019; Saier, 2019). The flow of BSI is described in Francis Crick's sequence hypotheses (Crick, 1958). The sequence hypothesis states that the genetic information encoded in DNA or RNA determines the resultant amino acid sequences, and hence structure, function, and properties of synthesised proteins; this BSI is transferred residue-by-residue between the biological heteropolymer macromolecules DNA, RNA, and proteins (Ebeling et al., 1995; Krupyanskii, 2023; Ille et al., 2022; Zhang et al., 2016). To accomplish this, DNA is transcribed into a complementary messenger RNA (mRNA) by a DNA-dependent RNA polymerase, which in eukaryotes occurs in the cell nucleus; the mRNA is then transported to ribosomes, where it is translated into a protein (Mercadante et al., 2023).

Restrictions on the direction of this BSI flow are described in Crick's central dogma of molecular biology. The model is a framework that states BSI can be transferred from nucleic

acid to nucleic acid or protein, but not from protein to another protein or nucleic acid (Crick, 1958; Crick, 1970). This is illustrated in Figure 1.1 and Table 1.1.

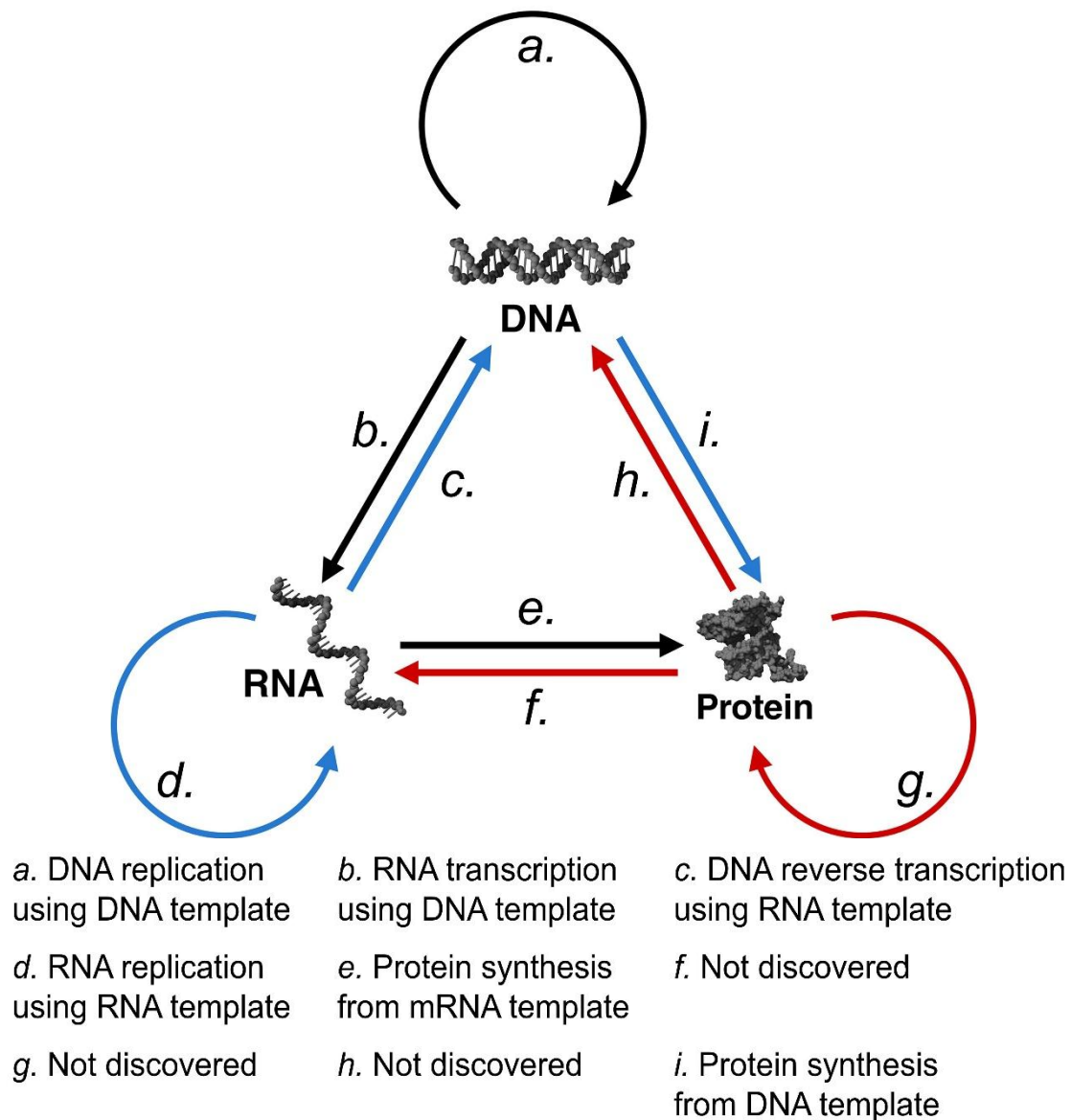


Figure 1.1 Directions of biological sequence information transfer (BSI) in the central dogma of molecular biology.

Arrows indicate all possible directions of BSI transfer within and between DNA, RNA, and protein. Black arrows indicate general transfer directions; blue arrows indicate special transfer directions; and red arrows indicate transfer directions for which evidence has not yet been shown. Figure adapted from Figure 1 in Ille et al. (2022).

Table 1.1 Details of biological sequence information transfer mechanisms within and between DNA, RNA, and protein in the central dogma of molecular biology.

Category	Directionality	Mechanism	Functional mediators
General transfers	DNA → DNA	DNA replication using DNA as a template (Lehman et al., 1958; Mercadante et al., 2023)	DNA-dependent DNA polymerase
	DNA → RNA	Transcription of ribosomal RNA, messenger RNA (mRNA), or transfer RNA using DNA as a template (Brenner et al., 1961; Gros et al., 1961; Mercadante et al., 2023)	DNA-dependent RNA polymerase
	RNA → protein	Protein synthesis using mRNA as a template during translation (Brenner et al., 1961; Gros et al., 1961; Mercadante et al., 2023)	Ribosome, tRNAs, factors, and enzymes
Special transfers	RNA → DNA	Reverse transcription of DNA using RNA as a template across all three domains of life and reverse-transcribing viruses (Ahlquist., 2006; Baltimore 1970; Mustafin et al., 2019; Temin et al., 1970)	RNA-dependent DNA polymerase (reverse transcriptase)
	RNA → RNA	RNA replication using RNA as a template in RNA-interfering eukaryotes and RNA viruses (Baltimore et al., 1963; Hua et al., 2021; Montgomery et al., 1998)	RNA-dependent RNA polymerase

	DNA → protein	<i>In vitro</i> protein synthesis in cell-free systems using DNA as a template (Chong, 2014; McCarthy et al., 1965)	Ribosome, tRNAs, factors, and enzymes
Unknown transfers	Protein → DNA	Not documented	Not applicable
	Protein → RNA	Not documented	Not applicable
	Protein → protein	Not documented	Not applicable

Table is modified from Table 1 in Ille et al. (2022).

Notably, the central dogma has been challenged through the existence of prions (Ille et al., 2022). Prions are transmissible pathogenic misfolded protein isoform aggregates which bind to the cellular isoform, autocatalytically changing its conformation to that of the prion, thus exponentially increasing cellular prion levels; the prions then spread cell-to-cell to continue accumulating (Prusiner, 1982; Prusiner, 1998; Prusiner et al., 1998; Sigurdson et al., 2019). It has been argued that this transfers BSI in a protein → protein direction, hence violating the central dogma (Daus, 2016; Koonin, 2012). Regardless, the central dogma remains valid and a ‘fundamental principle of molecular biology’ (Ille et al., 2022).

1.1.2 Protein Structure and Function

The central dogma has three key general BSI transfers, which are DNA replication, transcription, and translation; they ultimately result in a synthesised protein (Ille et al., 2022; Seo et al., 2023; Zagrovic et al., 2023). The protein’s properties, structure, and function are determined by its constituent amino acid residues (James et al., 2025). There are 22 canonical proteinogenic amino acids that are incorporated into proteins as residues linked by peptide bonds; their unique properties are determined by their side-chain R-group (Fichtner et al., 2017; Lopez et al., 2024). These amino acids are summarised in Figure 1.2. The residues can then be post-translationally modified, further altering their properties and in turn the protein’s properties, structure, and function (Lopez et al., 2024; Pan et al., 2022). There are a wide range of post-translational modifications, including ‘acetylation, methylation, phosphorylation, ubiquitination, glycosylation, SUMOylation, and redox’ (Zhong et al., 2023).

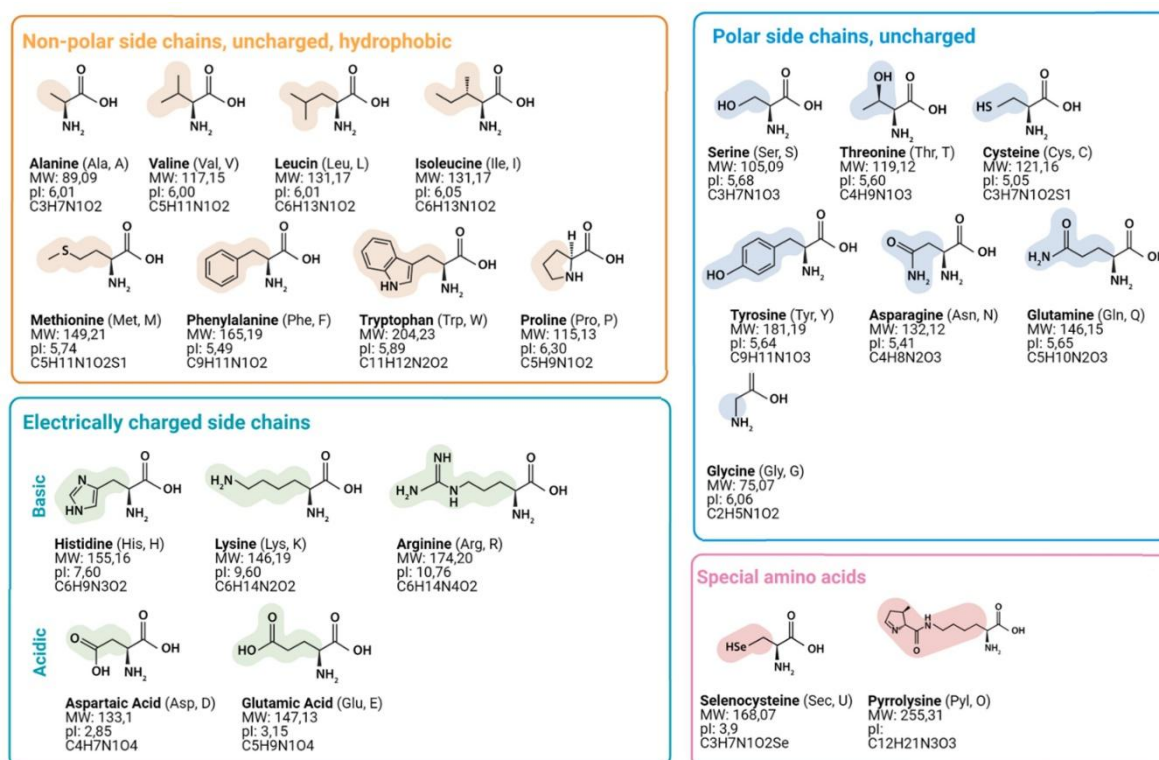


Figure 1.2 Summary of the canonical proteinogenic amino acids.

There are 22 proteinogenic amino acids encoded for by genetic information. They are divided up into 20 standard amino acids and two special amino acids (JPT, 2025a). Standard amino acids participate in protein synthesis, while 'selenocysteine occurs at the active site of several enzymes', and pyrrolysine is not found in humans. Their unique properties are determined by their side-chain R-groups (Lopez et al., 2024). Polar amino acids are hydrophilic, and either uncharged or charged (JPT, 2025b; Satoh et al., 2022; Zhou et al., 2018). Basic amino acids are positively charged, while acidic amino acids are negatively charged (JPT, 2025a). Amino acids are listed above with their full names, three-letter abbreviations, and single-letter abbreviations, below which are their molar weight (MW) in g/mol, isoelectric point (pI) in pH, and chemical formula. R-groups are highlighted in chemical structure diagrams. Figure is adapted from Amino Acid Chart by JPT (2025a).

The protein structure itself can be divided up into four levels: primary, secondary, tertiary, and quaternary. A protein's primary structure is its linear residue sequence, while its secondary structure is comprised of α -helix or β -pleated sheets, arising from structural folding due to

interactions between CO and NH functional groups in different residues forming hydrogen bonds (Rehman et al., 2024).

A protein's tertiary structure is its definitive three-dimensional configuration. It arises from further folding due to intermolecular interactions between R-groups and other R-groups or the peptide backbone (Rehman et al., 2024). These interactions are, in order of strongest to weakest: disulphide bonds, ionic interactions, hydrogen bonds, hydrophobic interactions, and Van der Waals interactions. Disulphide bonds are covalent; the others are non-covalent (Ferenczy et al., 2022; Stollar et al., 2020). For a protein to fold into its correct shape, hydrophobic interactions typically force hydrophobic residues to cluster together as a core (Rehman et al., 2024). Meanwhile, hydrogen bonds form between partially negative nitrogen or oxygen atoms or partially positive hydrogen atoms in a polar R-group and the opposite partial charge in another R-group or the peptide backbone (Rehman et al., 2024; Stollar et al., 2020). Hydrophobic interactions and hydrogen bonds are the main tertiary structure intermolecular interactions, with hydrophobic interactions contributing more to a protein's structural stability and hydrogen bonds contributing more to its mechanical resistance (Ferenczy et al., 2022; Pace et al., 2011). Van der Waals interactions occur 'when molecules with complementary shapes approach each other', including non-polar residues (Stollar et al., 2020). Ionic interactions occur between oppositely charged ions, such as basic and acidic R-groups. Disulphide bonds form between sulphydryl groups in two cystines (Rehman et al., 2024; Stollar et al., 2020).

A protein's tertiary structure has a distinct hierarchy: domains are distinct compact folding units which can be divided into subdomains; subdomains are in turn assembled from folds; folds are typically comprised of structural motifs; and motifs are short structural segments of generally 5-25 residues (Saito et al., 2007; Schaeffer et al., 2016; Stollar et al., 2020; Sun et al., 2004). Additionally, if two or more proteins form a functional structure, they become protomer subunits in an oligomer complex with a quaternary structure that uses the same types of intermolecular interactions as a tertiary structure (Carugo, 2014; Rehman et al., 2024). Tertiary and quaternary structures are fundamental to a protein's function, as they are

the state in which it interacts with other molecules (Morris et al., 2022; Rehman et al., 2024). Proteins have a vast array of functions, all of which require interacting with/binding to other molecules via a region in the protein termed a binding site (Morris et al., 2022; Ringe, 1995).

1.1.3 Protein-Protein Interaction

Protein-protein interactions (PPIs) occur as steps in biochemical pathways, which feed into a vast range of larger biological processes (La et al., 2013; Rao et al., 2014; Seychell et al., 2021; Wang et al., 2005). PPIs can be classified based on several factors: molecular docking mechanism, stability, lifetime, fold, oligomer type, and symmetry (Nooren et al., 2003; Peng et al., 2017). The molecular docking mechanism can be direct, where proteins physically contact each other, or indirect, where they physically contact an intermediate(s) and interact through it instead (Peng et al., 2017). The stability can be obligate or non-obligate, where an obligate PPI's 'protomers are not found as stable structures on their own *in vivo*' (Nooren et al., 2003). In regard to lifetime, a PPI can be permanent, where proteins form a stable long-term complex with a higher affinity disassociation constant of 1×10^{-9} M or lower, or transient, where they form an unstable short-term complex with a lower affinity disassociation constant of 1×10^{-6} M or higher (La et al., 2013; Peng et al., 2017). Obligate PPIs are always permanent; most non-obligate PPIs are transient (Peng et al., 2017). On the basis of fold, PPIs can be domain-domain, where an independently-folding globular peptide recognition module (PRM) of 50-150 residues binds to another PPI, creating a relatively large contact interface, or domain-peptide, where a PRM binds to a linear motif of around 10 residues, creating a relatively small contact interface (Peng et al., 2017; Schillinger et al., 2009; Stein et al., 2009). Domain-peptide PPI PRMs do not change conformation when binding to linear motifs, and so the PPIs are usually transient, with the PRM able to easily change interaction partners (Peng et al., 2017). In terms of oligomer type, a complex can have identical or non-identical protomers, being a homo- or heterooligomer, respectively, and two or more protomers, such as a dimer, trimer, or tetramer. Homodimers can further be divided based on symmetry, being either isologous, interacting using the same binding site location on each protomer, or heterologous, using binding site locations on each protomer (Nooren et al., 2003).

The PPI binding event occurs through the interaction of a binding site on one protein with a complementary partner binding site on another protein, acting in a binding site pair (Deng et al., 2022; Konc et al., 2022; Nooren et al., 2003). A monovalent PPI uses only one binding site pair; a multivalent PPI involves multiple binding sites pairs involving interaction between one protein and one or more other proteins (Deng et al., 2022; Mirica et al., 2012; Teilum et al., 2021). A multivalent PPI of several relatively weak binding site pair interactions can be stronger than a relatively strong monovalent PPI (Mammen et al., 1998; Teilum et al., 2021). A binding site has specificity for what partner binding site it interacts with; if it is monospecific, it can only bind to one distinct binding site, or if it multispecific, it can bind to multiple possible binding sites (Nooren et al., 2003; Teilum et al., 2021). IDPRs are often highly multispecific, or promiscuous (Schrieber et al., 2011). A binding site's specificity allows the protein to interact with functional protein partners while avoiding nonspecific PPIs, may be nonfunctional or pathogenic (Bahadur et al., 2008; Johnson et al., 2011). Its specificity arises from its complementarity (Nooren et al., 2003), which is determined by a broad range of shape-related properties, such as 'volume, depth, geometric features, and flexibility', and physicochemical properties, such as 'electrostatic potential, hydrophobicity, hydrogen bond potential, and aromaticity' (Scott et al., 2022). These properties are due to the proteins' tertiary or quaternary structure, in turn arising from the sequence and properties of its constituent residues (Brinda et al., 2005; James et al., 2025; Janin et al., 2008; Rehman et al., 2024; Talavera et al., 2011).

The binding event has three stages. First, there is a 'fast formation of a [nonspecific] encounter complex [by a pair of] free proteins in solution' (Marín-López et al., 2018). The nonspecific encounter complex has weakly associated conformations. This allows the second stage, where the two proteins search each other's surface to find complementary binding site pairs and reach the optimum binding geometry (Nooren et al., 2003; Schilder et al., 2013). This may involve interaction of one to three anchor residues in a binding site, usually in the smaller protein if they are different sizes, with the partner binding site to stabilise the nonspecific encounter complex into a transient encounter complex (Csermely et al., 2010; Rajamani et al., 2004). At this point the transient encounter complex is mainly stabilised 'by long-range

electrostatic interactions and desolvation forces' (Schilder et al., 2013). Finally, the transient encounter complex becomes a bound, specific complex, in which short-range intermolecular interactions between surface residues become important (Konc et al., 2022; Schilder et al., 2013). This may involve latch residues at the periphery of one or both binding sites changing into a final conformation to finalise specific complex stability (Csermely et al., 2010; Rajamani et al., 2004).

Any of the binding event stages may involve a variety of different conformational changes by one or both of the interacting proteins (Csermely et al., 2010). Additionally, proteins may have alternate conformational modes of binding with each other (Kundrotas et al., 2013). The intermolecular interaction forces between residues in the binding event are the same as those that determine a protein's tertiary and quaternary structure (Ferenczy et al., 2022; Schreiber et al., 2009; Seychell et al., 2021; Stollar et al., 2020; Westermarck et al., 2013). Several models have been proposed for the binding event process, including lock and key, induced fit, and conformational selection (Spasov, 2024). It should also be noted that a PPI can be cooperative, where binding of one protomer to another influences the complex's affinity for further binding of one of the protomers, or competitive, where two proteins compete to bind at the same binding site (Peacock et al., 2017; Tian et al., 2025).

The ultimate nature of a PPI is determined not just by the protein's themselves, but a number of external factors. These include cellular regulation, which is affected by signals, stimuli, and cell state, and the presence and concentration of other cellular macromolecules, (Ngounou Wetie et al., 2014; Speer et al., 2021). Thus, PPI is an extremely complex event driven by a multitude of factors (Sardiu et al., 2011).

1.2 DNA Replication

1.2.1 The Cell Cycle and DNA Replication Origins

DNA replication is arguably the most important of the three general BSI transfers (Ausiannikava et al., 2017, Ille et al., 2022; Seo et al., 2023; Zagrovic et al., 2023), possibly even ‘the most fundamental biological process’ (Ausiannikava et al., 2017). It can be defined as biological process by which a cell produces a new identical copy of DNA from an original DNA template (Chaudhry et al., 2023; Mercadante et al., 2023). This is a semiconservative process, as shown in Figure 1.3 (Mercadante et al., 2023; Meselson et al., 1958; Watson et al., 1953). DNA replication is ubiquitous across all three domains of life, which are bacteria, archaea, and eukaryotes, and has a fundamental role in DNA repair and ultimately allows for protein synthesis (Ausiannikava et al., 2017; Chaudhry et al., 2023; Ekundayo et al., 2019; Mercadante et al., 2023; O’Donell et al., 2013).

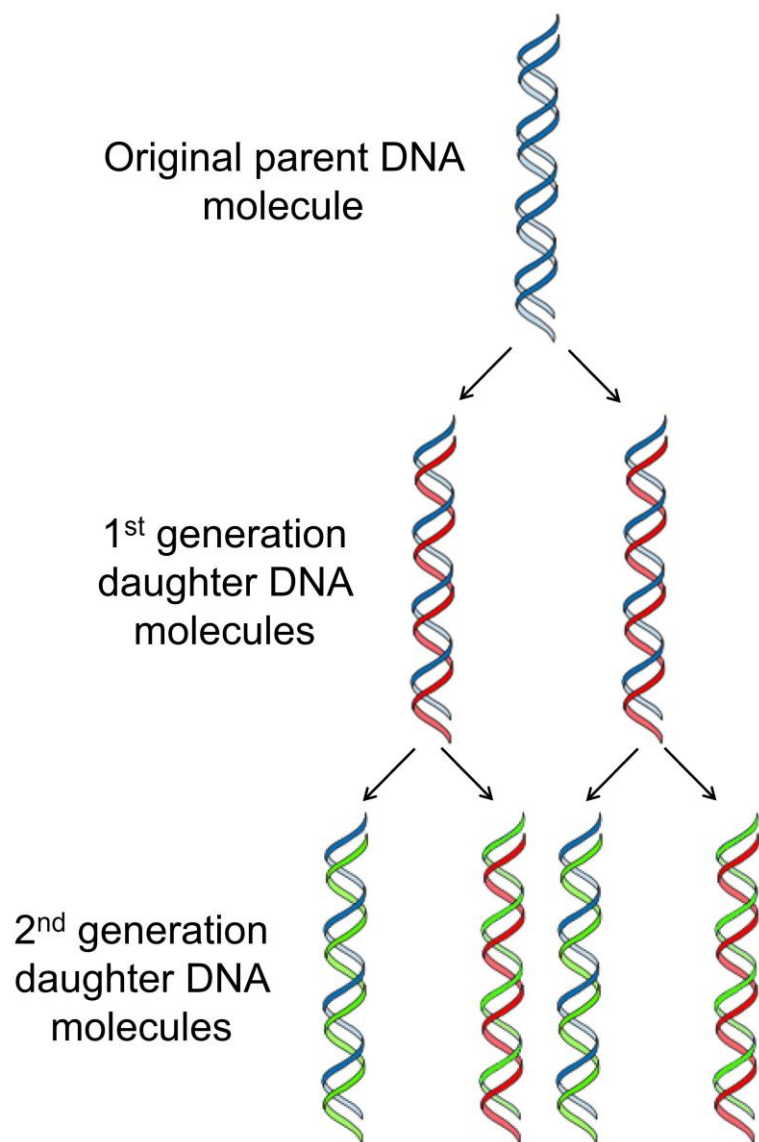


Figure 1.3 Semiconservative DNA replication process.

DNA is a semiconservative process, thus the original template parent double-stranded DNA (dsDNA) molecule splits so that both two resultant daughter dsDNA molecules each consist of one strand from the old parent molecule and one strand of newly synthesised DNA (Mercadante et al., 2023; Meselson et al., 1958; Watson et al., 1953). Figure is adapted from Figure 6 in Meselson et al. (1958) and Figure 7 in Wong (2009) chapter 'DNA'.

DNA replication's fundamental role lies in the cell cycle. The cell cycle is a complex process by which a cell is maintained, grows, and proliferates, culminating in the production of two new daughter cells from a single old parent cell by cell division (Schafer, 1998). DNA replication

functions as a prerequisite to cell division through replicating a cell's genome so that both daughter cells to have genetic information identical to each other and their parent cell (Cezanne et al., 2024; Chaudhry et al., 2023; Ekundayo et al., 2019; Mercadante et al., 2023; O'Donnell et al., 2013). While DNA replication's principals are common across the three domains of life, it differs in terms of cell cycle timing and molecular machinery (Ausiannikava et al., 2017; Cezanne et al., 2024). The variance in the cell cycle's phases and genome DNA replication timings across the domains of life is illustrated in Figure 1.4.

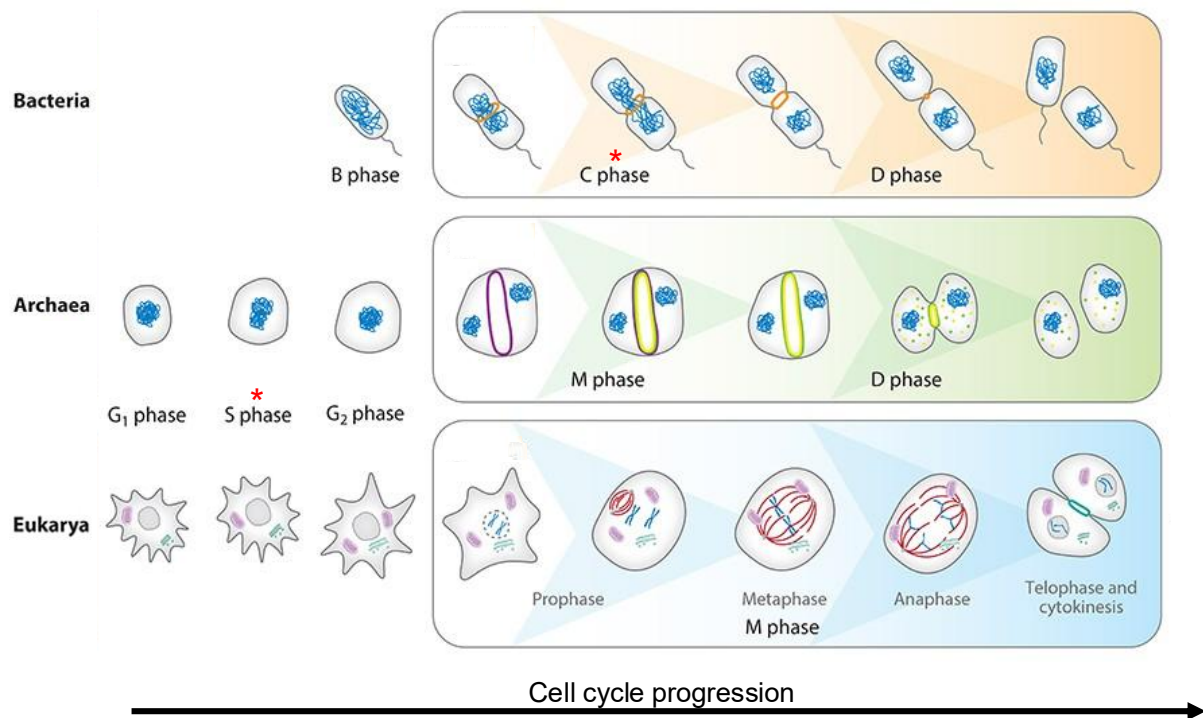


Figure 1.4 Overview of cell cycle phases and genome DNA replication timings across the three domains of life.

The bacterial cell cycle progresses through three distinct phases: B phase (period after birth), C phase (DNA replication), and D phase (cell division) (Wang et al., 2009). The superphylum TACK archaeal and eukaryotic cell cycles are similar and very different from the bacterial cell cycle, progressing through five and four distinct phases, respectively: G₁ phase (cellular growth), S phase (DNA replication), G₂ phase (cellular growth), M phase (chromosome segregation and cell division by mitosis in eukaryotes, divided into prophase, metaphase, anaphase, and telophase; chromosome segregation in archaea), and D phase (cell division in archaea) (Lindås et al., 2013; McIntosh, 2016). Other archaeal cell cycles are largely varied or unknown, and may possess less distinct phases, for example class Haloarchaea may lack a distinct S phase (Cezanne et al., 2024; Cui et al., 2021). Shaded boxes indicate genome segregation and cell division phases in the different cell cycles. Red asterisks indicate DNA replication phases. Figure is adapted from Figure 2 in Cezanne et al. (2024).

There are also variances in the number of DNA replication rounds initiated per cell cycle across the domains of life, related to control of the cell cycle (Cezanne et al., 2024). The eukaryotic cell cycle is highly regulated, with progression dependent on numerous cyclin-dependent kinases and their regulatory subunits, and satisfying checkpoints during G₁ (to prevent replication of damaged DNA), G₂ (to prevent mitotic segregation of defective chromosomes), and M phase (to prevent mitosis-related chromosome dysfunction) (Blagosklonny et al., 2013). In particular, the G₁ checkpoint, also known as the restriction point, determines if a cell proceeds with DNA replication or enters a quiescent G₀ phase, a reversible state of metabolically-active cell cycle arrest (Blagosklonny et al., 2013; Linke et al., 1996; Marescal et al., 2020; Wang, 2021). This results in sequential cell cycle phases with a single round of DNA replication per cell cycle (Cezanne et al., 2024; Romanowski et al., 1996; Sclafani et al., 2007). By contrast, the mechanisms of the cell cycle control across archaea are poorly understood and show a wide variety of mechanisms; this is true even in superphylum TACK order Sulfolobales archaea, which have been shown to possess distinct cell cycle stages similar to eukaryotes (Figure 1.4) (Ausiannikava et al., 2017; Cezanne et al., 2024). Coupled with the variation in known archaeal cell cycle stages (Figure 1.4), this may result in considerable variation in DNA replication rounds per cell cycle across archaea (Ausiannikava et al., 2017; Cezanne et al., 2024; Duggin et al., 2008; Hildenbrand et al., 2011).

Meanwhile, bacterial cell cycle control is unlike that of eukaryotes, as progression to a phase is not 'dependent on the completion of the previous phase'. Due to this, phases may overlap, resulting in multiple simultaneous rounds of DNA replication per cell cycle. (Cezanne et al., 2024). This is dependent bacterial cell growth speed. Slow growing bacterial cells, defined by 'mass doubling time > C + D period' time, initiate and undergo a single round of DNA replication per cell cycle. Conversely, fast growing bacterial cells, defined by 'mass doubling time $\leq$ C + D period time', initiate a single round of DNA replication per cell cycle, but can undergo multiple rounds simultaneously by initiating a new round before the old one has finished (Wang et al., 2009).

Rounds of DNA replication are initiated at genomic DNA sites known as replication origins, which is explained by the replicon hypothesis proposed by Jacob et al. (1963). DNA replication in ‘many organisms [uses] preferred genomic regions as origins’, likely to avoid DNA damage arising from DNA replication. To achieve this, *trans*-acting factors known as initiators, which are DNA replication initiator proteins, recognise and bind to *cis*-acting element specific DNA sequences known as replicators at these origins, after which they load DNA helicases onto the DNA, potentially acting alongside co-loader proteins, and the DNA helicase then drives recruitment of further DNA replication machinery (Ekundayo et al., 2019). The number of origins per chromosome where this occurs varies across the domains of life, as do the replication initiator proteins (Hu et al., 2023). This is detailed in Table 1.2.

Table 1.2 Details of replication origin number and DNA replication initiator proteins across the domains of life.

Domains of Life	Replication Origin Number	Chromosome Number	Genome Size	Replication Initiator Proteins
Bacteria	Single origin per chromosome	Single or multiple circular chromosomes	~0.6 to ~8.0 Mbp	DnaA
Archaea	Single or multiple origins per chromosome	Single circular chromosome (typically)	~0.5 to ~5.8 Mbp	Orc1/Cdc6
Eukaryotes	Multiple origins per chromosome	Multiple linear chromosomes	~10 to >100,000 Mbp	Orc1-5 and Cdc6

Table is modified from Table 2 in Hu et al. (2023).

While the replication initiator proteins vary considerably in their regulation and mechanisms, those listed in Table 1.2 are members of the ATPases Associated with various cellular Activities (AAA+) protein superfamily, save for the eukaryotic Orc 6 (Duderstadt et al., 2008; Duncker et al., 2009; Hu et al., 2023; Leonard et al., 2010). According to Snider et al. (2008), ‘the AAA+ superfamily is a ‘large and functionally diverse superfamily of NTPases that are characterised

by a conserved nucleotide-binding and catalytic' AAA+ module of 200-250 residues. AAA+ proteins have a wide variety of cellular functions, whose activity is regulated by adenosine triphosphate (ATP) binding and hydrolysis; the latter of which also provides the required energy (Dunn et al., 2023; Leonard et al., 2010; Snider et al., 2008).

Bacterial origins are termed *oriCs*, usually adjacent to the gene encoding the replication initiator protein DnaA (Leonard et al., 2013). These *oriCs* range from 250 base pairs (bp) to 2 kbp and contain three elements to control origin activity: 9 bp motifs called DnaA boxes, which vary in number and arrangement across bacteria; an AT-rich DNA unwinding element (DUE), and binding sites for replication initiation regulatory proteins (Ekundayo et al., 2019; Leonard et al., 2013). DnaA recognises and binds to DnaA boxes and is activated by ATP-binding (Ekundayo et al., 2019; Erzberger et al., 2006; Sekimizu et al., 1987). This results in a helical filament protein complex forming at the *oriC*; in *E. coli* this complex contains integration host factor, DiaA, and oligomeric ATP-DnaA (Katayama et al., 2017; Zawilak-Pawlik et al., 2017). The complex then unwinds the DUE into single-stranded DNA (ssDNA) molecules by melting (Rashid et al., 2025).

Archaeal replication origins share attributes of both bacteria and eukaryotes (Leonard et al., 2013). Archaeal origin positions and contents are similar to *oriCs*: they are usually adjacent to the gene encoding the replication initiator protein Cdc6/Orc1, a homologue eukaryotic Orc1 and Cdc6, and contain an AT-rich DUE flanked by one or more motifs called origin recognition boxes (ORBs or miniORBs depending on length), which vary in number, arrangement, and sequence (Ausiannikava et al., 2017; Cezanne et al., 2024; Ekundayo et al., 2019; Leonard et al., 2013). Cdc6/Orc1 recognises and binds to ORBs, although the number of Cdc6/Orc1 required to bind varies across archaeal species; the Cdc6/Orc1 is activated by ATP binding (Ausiannikava et al., 2017; Cezanne et al., 2024; Ekundayo et al., 2019; Samson et al., 2013). The ATP-Cdc6/Orc1 recruits minichromosome maintenance (MCM) helicases, which works alongside Cdc45 and Go-Ichi-Ni-San (GINS) GINS, which are respectively homologous to eukaryotic Cdc45 and GINS, in a Cdc45-MCM-GINS (CMG) complex to unwind the DUE(s) into

ssDNA molecules (Cezanne et al., 2024; Dewar et al., 2017; Makarova et al., 2012; Samson et al., 2016).

Eukaryotic origins are more complex than in bacteria or archaea. Rather than being signified by a specific sequence motif, their locations in eukaryotes other than *Saccharomyces cerevisiae* and *Saccharomycotina* species are determined by numerous factors: 'DNA accessibility, nucleotide sequence [content], nucleosome positioning, epigenetic features, DNA topology, certain DNA structural features [such as G-quadruplexes], as well as regulatory proteins, and transcriptional interference' (Ekundayo et al., 2019). During the cell cycle G1 phase, a heterohexameric origin recognition complex (ORC) comprised of Orc1-6 binds to an origin in an ATP-dependent manner (Bell et al., 1992; Diffley et al., 1994; Ekundayo et al., 2019; Hu et al., 2023; Xu et al., 2020; Yang, R. et al., 2024). The ORC recruits factor Cdc6, forming a ring-like structure and allowing for recruitment of the chromatin licensing and replication factor Cdt1 and a heterohexamer of replicative DNA helicase MCM complex comprised of Mcm2-7; this results in a pre-replication complex which 'licences' the origin (Ekundayo et al., 2019; Lattmann et al., 2022; Li et al., 2012; Parker et al., 2017; O'Donnell et al., 2013; Speck et al., 2005; Takara et al., 2011). MCM helicase activity remains inactive until S phase, when kinases direct chromatin association of numerous factors, primarily Cdc45 and GINS, which form a CMG complex with MCM (Aparicio et al., 2006; Bochman et al., 2009; Ilves et al., 2010; Langston et al., 2023; O'Donnell et al., 2013). This activates MCM helicase activity in the CMG at the licenced origin to unwind the DNA duplex into ssDNA molecules, while Cdc6 and Cdt1 are eliminated from the complex (Bochman et al., 2009; Ilves et al., 2010; Langston et al., 2023; O'Donnell et al., 2013).

1.2.2 Initiation and the Replisome

DNA replication can be divided into three stages: initiation, elongation, and termination (Song et al., 2023). During initiation, origin dsDNA is unwound into two ssDNA molecules (see Section 1.2.1), resulting in an open replication fork where replisomes are assembled (Ekundayo et al., 2019; Kaplan et al., 2010; Leman et al., 2013; O'Donnell et al., 2013; Yao et al., 2010). Replisomes are multiprotein complexes that carry out DNA replication. They include

six fundamental components that are conserved across all domains of life: a DNA helicase, a DNA single-stranded binding (SSB) protein, an RNA primase, a DNA polymerase, a DNA clamp loader, and a DNA clamp (Yao et al., 2010). These components are detailed in Table 1.3, and the *E. coli* and eukaryotic replisomes are illustrated in Figure 1.5.

Table 1.3 Details of replisome components across the three domains of life.

Component	<i>E. coli</i>	Eukaryotes	Archaea
DNA helicase	DnaB	MCM/CMG	MCM/CMG
Single-stranded binding protein	SSB	RPA	RPA
RNA primase	DnaG	Pol α -primase	Primase
DNA polymerase	Pol III	Pol $\alpha/\delta/\epsilon$	Pol
DNA clamp loader	γ/τ complex	RFC, Ctf18-RFC	RFC
DNA clamp	B	PCNA	PCNA
Other	?	Cdc45, GINS	GAN, GINS

E. coli replisome components are used as an example of a bacterial replisome as their replisome is currently the most understood in bacterial species and the *E. coli* species' status as a model organism (Beattie et al., 2015; Łazowski et al., 2024; Torres et al., 2025; Trojanowski et al., 2018). RecJ is not listed among *E. coli* other components as it is not considered part of the replisome, despite it being orthologous to Cdc45 and GINS-associated nuclease (GAN) (Makarova et al., 2012; Martínez-Carranza et al., 2024; Sanchez-Pulido et al., 2011; Torres et al., 2025). Table is modified from Table in Yao et al. (2010) (Greci et al., 2020; Martínez-Carranza et al., 2024; Yuan et al., 2024).

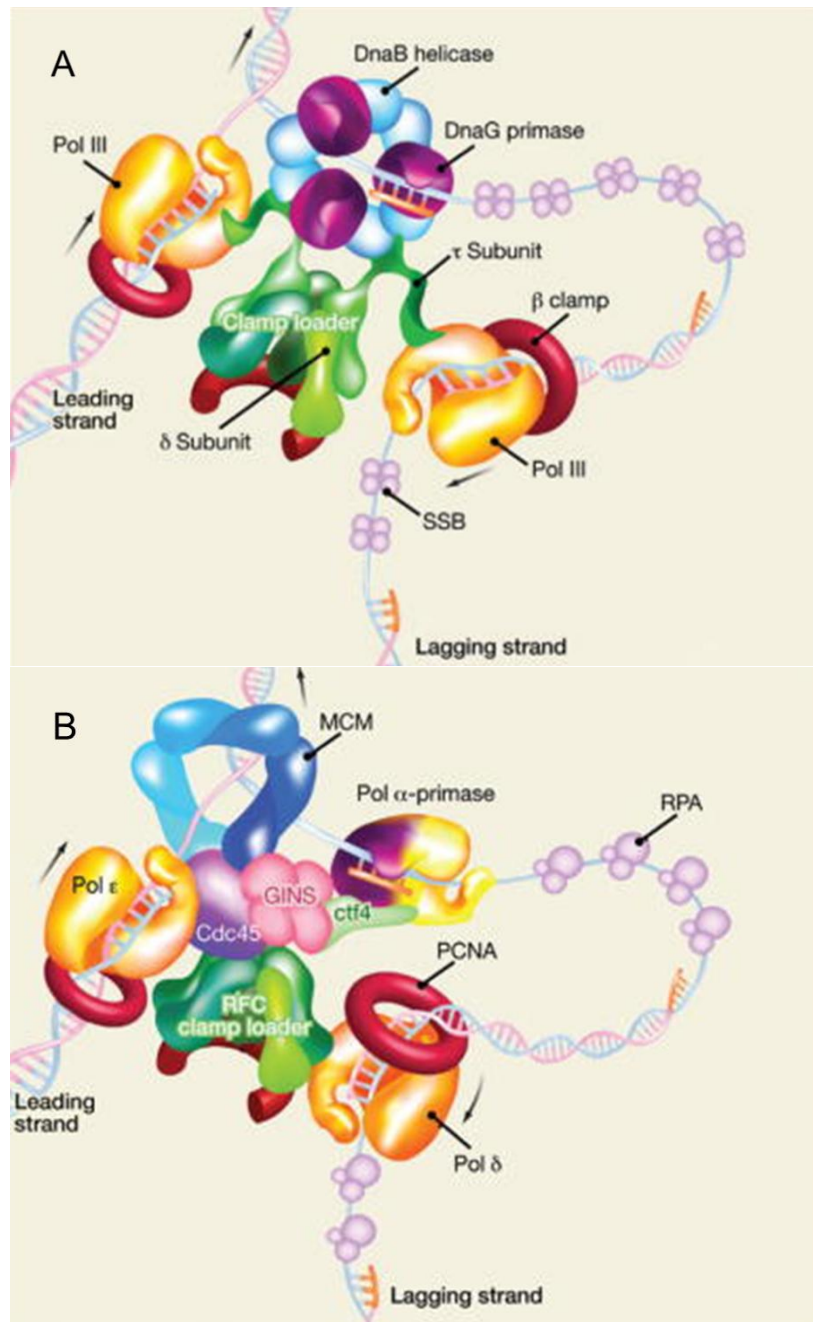


Figure 1.5 A. Components of the *E. coli* replisome.

Diagram of the *E. coli* replisome bound to chromosomal DNA during DNA syntheses following replication initiation. Key replisome components are annotated. Figure is adapted from Table panel A in Yao et al. (2010).

B. Components of the eukaryotic replisome.

Diagram of the eukaryotic replisome bound to chromosomal DNA during DNA syntheses following replication initiation. Key replisome components are annotated. Figure is adapted from Table panel B in Yao et al. (2010).

During initiation, the first replisome components assembled are the DNA helicase and SSB protein (Chaudry et al., 2023). In *E. coli*, once an *oriC* DUE has been unwound into ssDNA molecules, ATP-DnaA interacts with a DnaB-DnaC complex, loading the replicative DNA helicase DnaB onto the lagging strand, where it forms a homohexamer (Hayashi et al., 2020; LeBowitz et al., 1986; Itsathitphaisarn et al. 2012; Katayama et al., 2017; Keyamura et al., 2007; Marszalek et al., 1994; McMillan et al., 2024; San Martin et al., 1995). DnaB then translocates in a 5'-3' direction along the lagging strand, unwinding the duplex DNA into two ssDNA molecules using energy from ATP hydrolysis, resulting in two bidirectional replication forks (Kaguni, 2011; LeBowitz et al., 1986; Yao et al., 2010). SSB is recruited and binds onto the unwound ssDNA molecules; SSB has many roles in the replisome, including protecting the ssDNA from damage, preventing ssDNA reannealing (Meyer et al., 1990; Spenkelink et al., 2019), and '[removing] DNA secondary structures that would impede replication' (Yao et al., 2010). In eukaryotes, two CMGs in a head-to-head orientation translocate in the 3'-5' direction along the leading strand, unwinding the duplex DNA into two ssDNA molecules in an ATP-dependent manner, resulting in two bidirectional replication forks (Kose et al., 2019; Langston et al., 2023; Paker et al. 2017; Yao et al., 2010). The eukaryotic SSB protein is replication protein A (RPA), which has functional similarities to SSB (Brill et al., 1998; Paul et al., 2025; Yao et al., 2010). Archaeal replication initiation machinery is similar to that of eukaryotes, although archaeal CMG forms a homohexamer rather than a heterohexamer and 'most archaeal replicative helicases contain [GAN], which is orthologous to eukaryotic Cdc45 and bacterial RecJ' (Martínez-Carranza et al., 2024).

The remaining replisome components are then assembled (Chaudry et al., 2023). In *E. coli*, at each of the two replication forks, DnaB recruits the RNA primase DnaG to both the leading and lagging strands, where it synthesises primers of approximately 10-12 residues from the unwound ssDNA templates (Chang et al., 2000; Kitani et al., 1985; Lewis et al., 2016; O'Donnell et al., 2013; Radford et al., 2023; Spinks et al., 2021; Swart et al., 1993). SSB also interacts with DnaG to strengthen DnaG-primer association (Shereda et al., 2008). The next replisome component is DNA polymerase Pol III, a holoenzyme composed of 10 different proteins, with 17 subunits (Ozawa et al., 2013). Pol III contains a clamp loader comprised of subunits Pol III_γ,

Pol III δ , Pol III δ' , Pol III τ , Pol III χ , and Pol III ψ , which recognises the primed ssDNA, and so Pol III is loaded onto the leading and lagging strands (Fang et al., 2011; Simonetta et al., 2009; Yao et al., 2010). The clamp loader loads Pol III by opening the sliding DNA clamp, the Pol III β subunit, in an ATP-driven process; Pol III β functions in a head-to-tail ring-shaped homodimer that encircles and tethers PolIII to the ssDNA (Fang et al., 2011; Kong et al. 1992; Yao et al., 2010 Yao et al., 2012).

In eukaryotes, Pol α is a dual-function RNA primase and DNA polymerase, composed of four subunits: Pri1, Pri2, Pol1, and Pol12 (PRIM1, PRIM2, POLA1, and POLA2 in humans, respectively) (Yuan et al., 2023). Pol α is recruited to both the leading and lagging strands at each of the two replication forks as a heterotetramer through interaction of Pol2 with CMG (Chilkova et al., 2007; Jones et al. 2023; Yuan et al., 2023). Using the unwound ssDNA templates, Pri1 synthesises an RNA primer, which Pol1 extends into an RNA-DNA hybrid primer of approximately 25-30 residues (Yuan et al., 2023). At the lagging strand's parent, replication factor C (RFC) interacts with the primed ssDNA, functioning as a DNA clamp loader to load proliferating cell nuclear antigen (PCNA) onto the DNA in an ATP-dependent process, where it forms a homotrimeric ring that encircles the ssDNA, and so acts as a sliding DNA clamp (Hingorani et al., 2002; Neuwald, 2005; Yao et al., 2012; Yuan et al., 2024). PCNA subsequently recruits DNA polymerase Pol δ , composed of three subunits (four in humans): Pol3p, Pol31p, and Pol32p (POLD1, POLD2, POLD3, and POLD4 in humans, respectively); PCNA is able to interact with all Pol δ subunits, including in humans (Khandagale et al., 2020; Li et al., 2009; Prindle et al., 2012; Mondol et al., 2019). This forms a holoenzyme composed of RFC, PCNA, and Pol δ (Schumacher et al., 2000). At the leading strand's parent, the DNA clamp loader is RFC variant Ctf18-RFC, which binds to the primed ssDNA in a complex with DNA polymerase Pol ϵ (Fujisawa, et al., 2017; Yuan et al., 2024). Pol ϵ is composed of four subunits: Pol2, Dpb2, Dpb3, and Dpb4 (POLE1, POLE2, POLE3, and POLE4 in humans, respectively) (He et al., 2024; Yuan et al., 2020). It works alongside Ctf18-RFC to loads PCNA onto the DNA in an ATP-dependent process (Fujisawa, et al., 2017; Yuan et al., 2024). Meanwhile, this latter part of the replisome assembly process in archaea is similar but simpler compared to that of eukaryotes (Greci et al., 2020; Makarova et al., 2013). Arguably the primary difference is that the DNA polymerase used varies across archaeal species, with all archaea except phylum

Crenarchaeota using at least one archaea-unique D-family DNA polymerase; additionally, the group(s) and number of B-family DNA polymerase(s) used vary across archaeal species (Greci et al., 2020; Feng et al., 2022; Makarova et al., 2014).

1.2.3 Elongation and Termination

DNA synthesis occurs during the elongation stage of DNA replication (Chaudry et al., 2023). In *E. coli*, at the leading strand Pol III extends the RNA primer by DNA synthesis and so replicates the parent strand in a 5'-3' direction (McHenry et al., 2011; Naufer et al., 2017). The Pol III catalytic core is comprised of Pol III α , Pol III ϵ , and Pol III θ , with Pol III α containing the polymerase active site itself (Naufer et al., 2017; Ozawa et al., 2013). As the DNA polymerase progresses along the bases of the template ssDNA molecule (Yao et al., 2010), it 'incorporates a deoxyribonucleotide triphosphate (dNTP) complementary to the template base and [catalyses] the formation of a phosphodiester bond between the 3'-OH of the primer strand and the α phosphorus of the dNTP' (Nakamura et al., 2013). This results in the release of a pyrophosphate group from the dNTP, and the pyrophosphate is then hydrolysed by a pyrophosphatase into two phosphate molecules; these two processes provide the energy for continued DNA synthesis (Burke et al., 2018; Kottur et al., 2018). As DnaB translocates and unwinds the duplex DNA, it drives the movement of the replication fork and associated replisome, with the continual DNA synthesis accelerating the unwinding (Kaguni, 2011; LeBowitz et al., 1986; Spinks et al., 2021; Stano et al., 2005; Yao et al., 2010). Due to DNA having an antiparallel structure, DNA synthesis in the same replisome occurs continuously at the leading strand, but not at the lagging strand (Naufer et al., 2017; Yao et al. 2010). Instead, replication of the lagging strand's parent occurs discontinuously in a 5'-3' direction, resulting in ssDNA molecules of 1-2 kb termed Okazaki fragments, hence requiring a new RNA primer for each fragment (Burgers, 2019; Makiela-Dzbenka et al., 2019; Sugino et al., 1972; Yao et al., 2010). Removal of RNA primers from Okazaki fragments is performed through a combination of RNase H1 and DNA polymerase Pol I, with Pol I synthesising DNA to fill in the gaps between fragments (Allen et al., 2011; Dohrmann et al., 2011; Miller et al., 1973; Wendel et al. 2021). The fragments are then joined by DNA ligase A into one continuous DNA strand (Kaguni, 2018).

DNA synthesis in eukaryotes utilises similar physiochemical mechanisms to that of *E. coli*, with some fundamental differences (O'Donnell et al., 2013; Serdyuk et al., 2025; Yao et al., 2010). Firstly, rather than using a single DNA polymerase to replicate both duplex parent strands, Pol ϵ , with catalytic subunit Pol2 (POLE1 in humans), replicates leading strand's parent and Pol δ , with catalytic subunit Pol3p (POLD1 in humans), replicates the lagging strand's parent (Chilkova et al., 2007; Fujisawa et al., 2017; Huang et al., 1999; Prindle et al., 2012; Stillman, 2008; Yuan et al., 2020; Zegerman, 2013). Additionally, eukaryotic Okazaki fragments are only 100-200 bases long, and their joining is more complex (Yao et al., 2010). Fragment primers are first cleaved by RNase H2, but 'it cannot cleave between the 3'-most ribonucleotide and the initial deoxynucleotide' (Balakrishnan et al., 2013). The remaining single-stranded primer flap structure is typically cleaved by flap endonucleases in a short flap or long flap pathway: in the short flap pathway, Fen1 and Pol δ remove the flap, while in the long flap pathway Fen1 and Dna2 remove it (Balakrishnan et al., 2011; Balakrishnan et al., 2013; Liu, B. et al., 2017). Pol δ synthesises DNA to fill in the gaps, and the fragments are joined by DNA ligase I (Balakrishnan et al., 2011; Howes et al., 2012). DNA synthesis in archaea resembles that of eukaryotes, although the exact proteins involved and their structures vary considerably across species (Ausiannikava et al., 2017; Cannone et al., 2015; Lin et al., 2022; Oyama et al., 2011).

DNA replication ends during the termination stage (Chaudry et al., 2023). Termination can be divided into five processes: resolution of topological stress, meeting of converging replication forks, replisome disassembly, final Okazaki fragment gap filling and joining, and decatenation. The first process, resolution of topological stress, relies on the formation of pre-catenanes, DNA structures that form due to topological stress (Dewar et al., 2017). During DNA replication, unwinding of the parental duplex can lead to positive supercoils forming ahead of the replication fork; accumulation of these DNA structures can impede or halt replication (Dewar et al., 2017; Schwartzman et al., 2019). Type I and type II DNA topoisomerases relax supercoils and thus dissipate them (Ashley et al., 2017; Vos et al., 2011; Zechiedrich et al., 1990). Alternatively, over-rotation of the replication fork can cause daughter strands to cross over, forming pre-catenanes behind replication forks (Cebrián et al., 2015; Dewar et al., 2017; Garnier et al., 2021; Postow et al., 2001). During termination, when two replication forks

converge, if there is less than 150 bp of parent DNA duplex between them (Vafabakhsh et al., 2012), supercoils cannot form between the replication forks and so topological stress resolution 'becomes absolutely dependent on the formation of pre-catenanes' (Dewar et al., 2017) (see Figure 1.6A).

During the second process, meeting of converging of replication forks, the replication forks merge together, and the replisomes seamlessly pass by each other due to being bound to opposite parental DNA strands (Dewar, 2024) (see Figure 1.6A). The third process, replisome disassembly, is an active, regulated process that occurs to prevent re-replication and interference with subsequent processes such as transcription (Dewar et al., 2017; Jones et al., 2024) (see Figure 1.6B). The fourth process, Okazaki fragment gap filling and joining, is required to complete lagging strand DNA replication, since when two converging replication forks meet, 'a single-stranded gap exists between the 3' end of the leading strand and the downstream Okazaki fragment of the opposing fork'. The gap is filled and the DNA ligated (Dewar et al., 2017) (see Figure 1.6B). The fifth and final process, decatenation, is necessary to resolve catenanes before chromosome segregation. Catenanes arise at the end of DNA replication after the last turn of the parental DNA duplex is replicated and the daughter strands are interwound; this also converts any unresolved pre-catenanes into catenanes (Bush et al., 2015; Dewar et al., 2017) (see Figure 1.6C).

In *E. coli*, DNA replication on their circular chromosomes terminates in termination a region opposite the *oriC* (Syeda et al., 2020). This is achieved through a Tus-*ter* trap termination system. Each side of the termination region is flanked by five 23 bp nonpalindromic *ter* sequences designated *terA* to *terJ*, and *ter* sequences are bound to by Tus protein (Coskun-Ari et al., 1994; Hidaka et al., 1998; Jameson et al., 2021). The structure of Tus means that has a permissive and nonpermissive orientation when bound to *ter* sequences depending on the direction of migrating replisome approach (Mulcair et al., 2006). Consequently, DnaB in migrating replisomes during DNA replication interacts with and displaces Tus from the first five *ter* sequences they encounter since the protein is in a permissive orientation at that direction of approach, allowing replisome progression. However, Tus at the next five *ter*

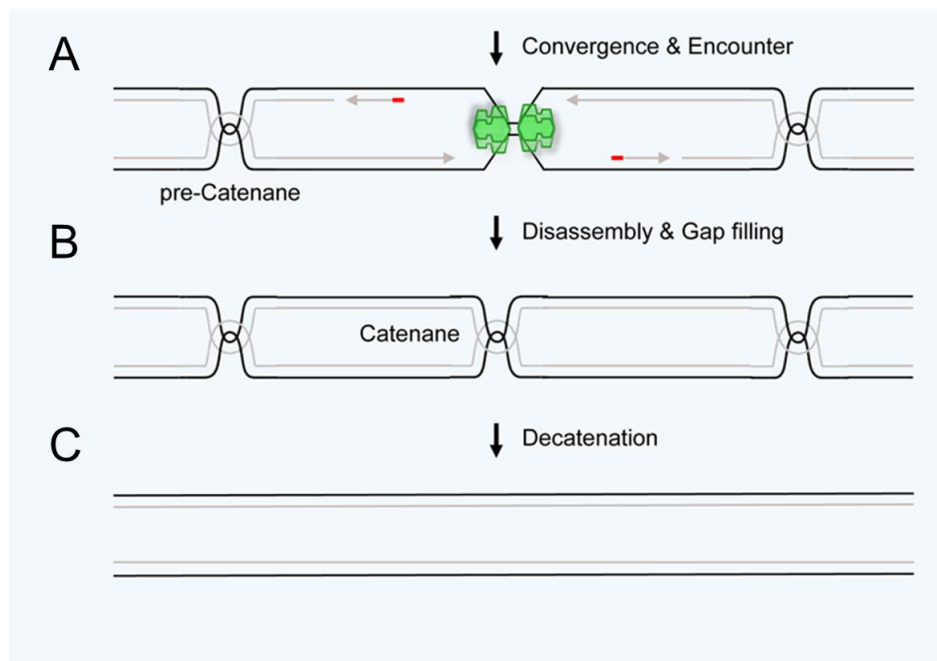


Figure 1.6 A. Generic illustration of resolution of topological stress and meeting of converging replication forks during DNA replication termination.

In the first termination process, topological stress is resolved through pre-catenane formation behind the replication forks when they are less than 150 bp apart ('convergence') (Dewar et al., 2012; Vafabakhsh et al., 2012). In the second process, converging forks meet ('encounter'), and replisomes seamlessly pass by each other (Dewar, 2024). DNA helicases are in green and RNA primers are in red. Figure is adapted from Figure 1E in Dewar et al. (2017).

B. Generic illustration of resolution of replisome disassembly and final Okazaki fragment gap filling and joining during DNA replication termination.

In the third process, the replisomes are actively disassembled from the two DNA duplexes (Dewar et al., 2017; Jones et al., 2024). In the fourth process, the final Okazaki fragment gap is filled with DNA, and the fragments is ligated to the 3' end of the lagging strand (Dewar et al., 2017). Figure is adapted from Figure 1F in Dewar et al. (2017).

C. Generic illustration of decatenation during DNA replication termination.

In the fifth and final process, a catenane arises at the end of DNA replication, and any unresolved turn pre-catenanes into catenanes. These are resolved before chromosome segregation (Bush et al., 2015; Dewar et al., 2017). Figure is adapted from Figure 1G in Dewar et al. (2017).

sequences is in a nonpermissive orientation from that direction of approach, and so DnaB cannot displace Tus and replisome progression is arrested, hence converging replication forks meet in the termination region (Dewar et al., 2017; Jameson et al., 2021; Mulcair et al., 2006; Mulugu et al., 2001). The five termination processes occur between the fifth and sixth *ter* sequences, *terC* and *terA* (Dewar et al., 2017). Positive supercoils are unwound by the type II topoisomerases DNA gyrase and Topo IV, while Topo IV unwinds pre-catenanes and is assisted by Topo III in unwinding pre-catenanes at the replication fork (Ashley et al., 2017; Lee et al., 2019). The converging replication forks meet and a 3' ssDNA flap is generated on each of the lagging strands (Dewar et al., 2017). Arrested replisomes remain DNA-bound at Tus-*ter* traps, as their arrest alone does not trigger replisome disassembly, although the exact mechanisms and regulation of *E. coli* replisome disassembly are currently unclear (Beattie et al., 2015; Dewar et al., 2017; Jones et al., 2024; Midgley-Smith et al., 2018; Moolman et al., 2016). The lagging strand flaps are removed, and the last Okazaki fragment is filled in and ligated to the upstream leading strand (Dewar et al., 2017). Finally, catenanes are resolved by Topo IV before chromosome segregation (Ashley et al., 2017; Dewar et al., 2017).

In eukaryotes, DNA replication terminates wherever two replication forks meet, rather than at a fixed location. This lack of specificity is primarily due to three factors: eukaryotic origins lack sequence-specificity, except in *S. cerevisiae*; eukaryotic origins fire at various times rather than simultaneously; and eukaryotic 'replication forks move at different speeds throughout the genome' (Jones et al., 2024). Positive supercoils are unwound by the topoisomerases Topo I and Topo II (TOP1 and TOP2 α in humans, respectively), while Topo II (TOP2 α in humans) unwinds pre-catenanes, and it has been hypothesized that Topo III (TOP3 α in humans) may also assist unwinding of pre-catenanes that include ssDNA (Bush et al., 2015; Lucas et al., 2001; McKie et al., 2021; Pommier et al., 2016; Pommier et al., 2022). The converging replication forks meet and a 3' ssDNA flap is generated on each of the lagging strands (Dewar et al., 2017). Replisome disassembly is driven by ubiquitination with lysine 48-linked ubiquitin chains on Mcm7 in CMG, promoted by Cullin-RING ubiquitin E3 ligase (CRL) CRL2^{LRR-1} in metazoans and the Skp1-Cullin1-F-box (SCF) protein complex SCF^{Dia2} in *S. cerevisiae* (Maric et al., 2014; Moreno et al., 2014; Moreno et al., 2020; Sonnevile et al., 2017; Xia, 2021). This

causes Mcm7 to pass ubiquitination threshold, so that the replisome is disassembled by p97-Ufd1-Npl4 (Moreno et al., 2020; Xia, 2021). Meanwhile, the lagging strand flaps are removed, and the last Okazaki fragment is filled in and ligated to the upstream leading strand (Dewar et al., 2017). Finally, catenanes are resolved by Top II (TOP2 α in humans) (Charbin et al., 2014; McKie et al., 2021; Rose et al. 1990). DNA replication termination in archaea has not been extensively investigated, although it has been shown that in *Sulfolobus acidocaldarius* and *Sulfolobus solfataricus*, which have three origins per chromosome, termination of fork pairs can occur asynchronously in nonspecific sites like in eukaryotes (Dalgaard et al., 2009; Goodall et al., 2023; Jones et al., 2024; Lundgren et al., 2004).

1.2.4 Replication Stress

DNA replication machinery normally replicates DNA with high fidelity; however, this efficacy can be threatened by DNA replication stress (Saxena et al., 2022). Replication stress can be defined as ‘inefficient DNA replication that causes DNA replication forks to progress slowly or stall’ (Burhans et al., 2007). This can be due to replisomes encountering endogenous or exogenous obstacles on chromosomes (Yoshida et al., 2021). Endogenous obstacles are arguably the most direct cause of replication stress, being physiological in nature and directly blocking the replisome; these can include non-B-DNA secondary structure formation and DNA lesions (Liu et al., 2024; Saxena et al., 2022; Yoshida et al., 2021). Non-B-DNA secondary structures contributing to replication stress include R-loops, hairpins, cruciforms, Z-DNA, triplex DNA, G-quadruplexes, and intercalated-motifs (Abou Assi et al., 2018; Herlihy et al. 2017; Kaushik Tiwari et al., 2016; Liu et al., 2024; Saxena et al., 2022; Williams et al., 2023; Zhao et al., 2010). DNA lesions contributing to replication stress include nicks, gaps, stretches of DNA single-stranded breaks, DNA double-stranded breaks (DSBs), interstrand crosslinks, adducts, and abasic sites (Saxena et al., 2022; Sarker et al., 2012; Tsegay et al., 2019; Vare et al., 2012; Zeman et al., 2014). DNA lesions arise due to DNA damage from various endogenous and exogenous sources (Alhmoud et al., 2020; Saxena et al., 2022). ‘Endogenous sources of DNA damage include hydrolysis, oxidation, alkylation, mismatch of DNA bases’ (Hakem, 2008), endogenous non-canonical nucleotides, and byproducts of cellular metabolism (Saxena et al., 2022). Exogenous sources of DNA damage include: ‘ionising radiation, ultraviolet radiation, various chemicals agents’ (Hakem, 2008) including mutagens, reactive oxygen species, and

exogenous non-canonical nucleotides (Saxena et al., 2022). Exogenous obstacles to replisome progression are environmental or artificial obstacles (Yoshida et al., 2021). These include alteration in cellular dNTP pools, increased origin firing, and clashes between replication and transcription complexes (Saxena et al., 2022; Yoshida et al., 2021).

Replication stress can have a number of potential consequences (Saxena et al., 2022). These can be localised to the replication fork, primarily replication fork collapse and DSBs (Deng et al., 2021; Patel et al., 2023; Saxena et al., 2022). These DSBs mean that replication stress consequences can extend to a cellular level by the accumulation of resultant mutations and chromosomal alterations including insertions, deletions, substitutions, duplications, and translocations (Grabarz et al., 2012; Saxena et al., 2022; Varga et al., 2005). These are partly induced by error-prone DNA repair pathways the cell employs to try and correct DSBs such as break-induced replication (BIR), non-homologous end joining (NHEJ), and microhomology-mediated end joining (MMEJ) (Saxena et al., 2022). The DSBs, mutations, and chromosomal alterations can ultimately culminate in genomic and epigenomic instability, disrupting genomic integrity and leading to genetic disease, primarily cancer (Tsegay et al., 2019; Zeman et al., 2014). This genomic instability can cause defects in cellular replication stress response proteins and pathways, enhancing genomic instability resulting from replication stress even further, particularly in cancer (Yoshida et al., 2021; Zeman et al., 2014). However, defects in replication stress response proteins can give rise to many diseases beyond just cancer; in humans these include Aicardi-Goutieres syndrome, Bloom syndrome, Fanconi anaemia, Nijmegen breakage syndrome-like disorder, Schimke immunoosseous dysplasia, and Spinocerebellar ataxia type 10 (Zeeman et al., 2014).

Cells have several mechanisms to counteract replication stress. On the replisome itself there are accessory helicases, which facilitate movement of replication forks past protein-bound DNA that would otherwise impede fork progression (Brüning et al., 2014). For example, in *E. coli* Rep translocates in a 3'-5' direction along the leading strand's template DNA and physically interacts with DnaB, while UvrD also translocates in a 3'-5' direction but does not directly interact with the replisome (Brüning et al., 2014; Matson, 1986; Yarranton et al., 1979). In

eukaryotes, accessory helicases translocate in a 5'-3' direction (Olson et al., 2024). *S. cerevisiae* possesses Rrm3 and Pif1, both members of the Pif1 DNA helicase family; metazoa also possess several additional DNA helicases, namely HELB, DDX11, FANCI and RTEL1 (Bochman et al., 2010; Olson et al., 2024). There are also DNA repair and tolerance pathways that activate in the event of a stressed or stalled replication fork to allow replication fork progression. These pathways include TLS, PrimPol-mediated repriming, template switching (TS), BIR, and other homologous recombination (HR) pathways (Saxena et al., 2022). However, even with these mechanisms, cells still encounter DNA damage, thus they possess several post-replication DNA pathways (Gao et al., 2017).

1.3 DNA Damage and Repair

1.3.1 DNA Damage Response Pathways

Across all domains of life, cells are at risk of DNA damage throughout the cell cycle, not just during DNA replication (Branzei et al. 2008; Chatterjee et al., 2017). This is due to continuous exposure to various endogenous and exogenous mutagens, which can cause DNA lesions, potentially giving rise to a number of different mutation types, as described in Section 1.2.4. As such, organisms possess DNA damage response pathways active throughout different cell cycle stages in order to repair the damage and hence maintain genome stability (Alhmoud et al., 2020; Branzei et al. 2008; Chatterjee et al., 2017). In eukaryotes there are principally five distinct DNA damage response pathways: mismatch repair (MMR), base excision repair (BER), NER, NHEJ, and HR (Chatterjee et al., 2017). Mutagens, their resulting DNA lesions, and the corresponding DNA damage response pathway are illustrated in Figure 1.7.

The MMR pathway is for correcting base-base mismatches, insertion/deletion mispairs (Li, 2008), and strand misalignments arising from DNA replication errors, HR, and direct DNA damage (Jiricny, 2013; Kunkel et al., 2005; Saxena et al., 2022; Schofield et al., 2003). MMR functions by excision of nascent DNA tracts containing mispairs and subsequent error-free resynthesis of the excised tract (Jiricny, 2013). MMR is widely conserved across bacteria and eukaryotes, but archaea tend to lack a canonical MMR pathway (Kunkel et al., 2013; White,

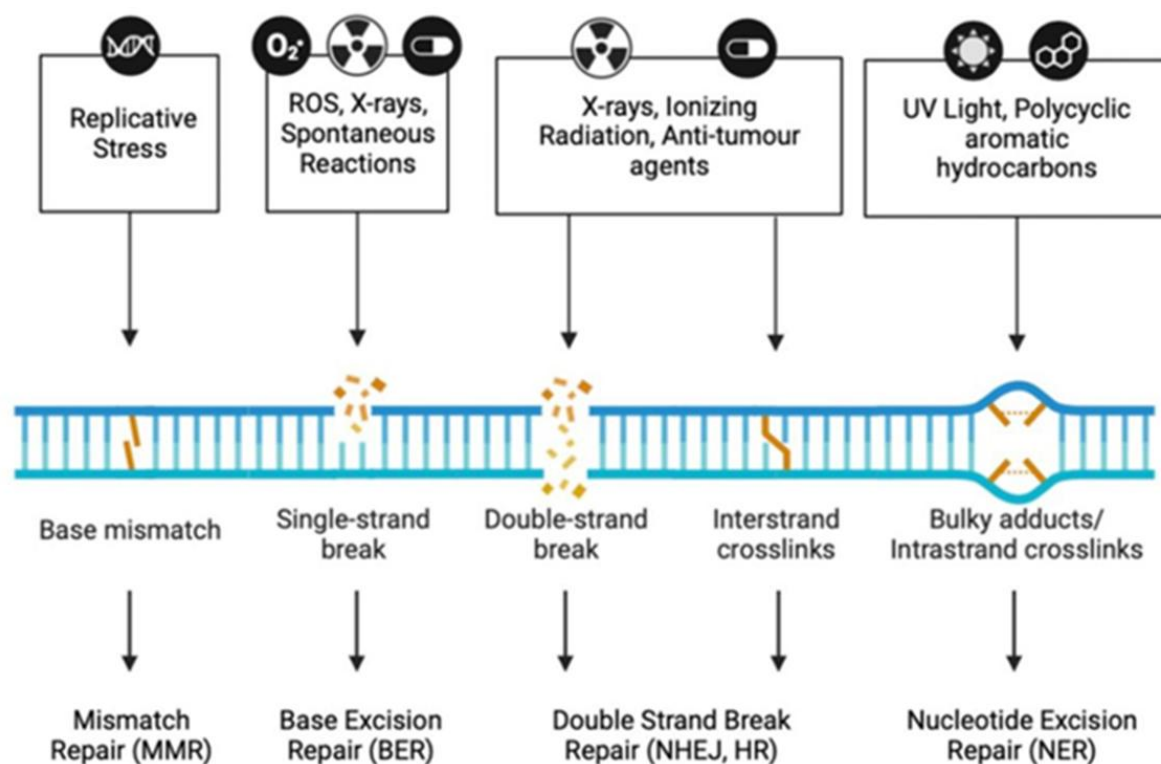


Figure 1.7 Sources of DNA damage, resultant DNA lesions, and corresponding DNA damage response pathway.

Across all domains of life, cells are continuously exposed to various endogenous and exogenous mutagens, which can cause DNA lesions. Therefore, organisms possess DNA damage response pathways active throughout different cell cycle stages to repair the damage and hence maintain genome stability (Alhmoud et al., 2020; Branzei et al. 2008; Chatterjee et al., 2017). In eukaryotes there are principally five distinct DNA damage response pathways: mismatch repair (MMR), base excision repair (BER), nucleotide excision repair (NER), non-homologous end-joining (NHEJ), and homologous end-joining (HR) (Chatterjee et al., 2017). Figure is Figure 1.3 in Lou-Hing (2024).

M.F. et al., 2018). In *E. coli*, mismatches are recognised in a methyl-directed system by MutS, which then works with MutL to direct the endonuclease MutH to nick the mismatch-containing segment (Hsieh et al., 2008). After the tract is excised, it is filled in by Pol III and ligated by DNA ligase (Fukui, 2010). Eukaryotes have a number of MutS and MutL homologues, primarily Msh2 and Msh6 (MSH2 and MSH6 in humans, respectively) in a MutS α complex for

MutS and Mlh1 and Pms2 (MLH1 and PMS2 in humans, respectively) in a MutL α complex for MutL, but none for MutH (Fukui, 2010; Hsieh et al., 2008; Reyes et al., 2015). Mismatches are recognised by MutS α by the presence of nicks, MutL α then nicks either side of the mismatched base (Haye et al., 2015; Hsieh et al., 2008), and the 'resulting DNA segment is excised by the exonuclease EXO1' (Hsieh et al., 2008). After the tract is excised, it is filled in by Pol δ and ligated by DNA ligase I (Fukui, 2010; Reyes et al., 2021).

The BER pathway is responsible for repairing non-bulky DNA lesions involving individual bases by removing and replacing damaged and non-canonical bases (Hindi et al. 2021; White, M.F. et al., 2018; Wozniak et al., 2022). It is a multi-stage process with two main pathways: a short-patch for repairing single nucleotide and long-patch pathway for repair of tracts of at least two nucleotides (Hindi et al., 2022; Robertson et al., 2009). BER relies on the activity of DNA glycosylases (Robertson et al., 2009; Wozniak et al., 2022). To initiate BER, a lesion is recognised by the DNA glycosylase appropriate for that specific lesion type (Robertson et al., 2009). There are two types of DNA glycosylases: monofunctional and bifunctional (Wozniak et al., 2022). Monofunctional glycosylases cleave the N-glycosidic bond, 'removing the damaged base removing the damaged base and creating an apurinic or apyrimidinic (AP) site' (Robertson et al., 2009). Then, the DNA backbone is cleaved by a DNA AP endonuclease or DNA AP lyase; bifunctional glycosylases also have AP lyase activity. This leaves a nick, which is filled in by DNA polymerase and then ligated by a DNA ligase (Robertson et al., 2009; Wozniak et al., 2022). This process is conserved across all domains of life, although in *E. coli* both the short-patch and long-patch pathways use the same set of proteins while in eukaryotes the pathways each use different sets (Krwawicz et al., 2007; White, M.F. et al., 2018). While many of the enzymes involved in BER are conserved across bacteria and eukaryotes, the exact enzymes used at each step vary considerably depending on the lesion type and location across bacteria, eukaryotes, and archaea (Hindi et al., 2021; Krokan et al., 2013; Robertson et al., 2009; White, M.F. et al., 2018; Wozniak et al., 2022).

The NER pathway is another base excision pathway, responsible for repairing bulky DNA lesions involving helix distortions (Batty et al., 2000; Rouillon et al., 2011). NER has two main

pathways: global genome NER, which can occur anywhere in the genome, and transcription-coupled NER, which occurs in the transcribed strand of active genes (Schärer, 2013). The core process is conserved across all domains of life, but the proteins involved vary considerably: *E. coli* primarily use a set of Uvr proteins; eukaryotes use a multitude of different proteins unrelated to the bacterial set; and across archaea there is a large diversity of NER proteins, ranging from the bacterial system to numerous eukaryotic NER system genes (White, M.F. et al., 2018; Wozniak et al., 2022). In brief, a bulky DNA lesion is first recognised by an appropriate specific factor patrolling the DNA, which in *E. coli* is an ATP-bound UvrA dimer; these factors are each able to recognise a wide variety of bulky lesions. A pre-incision complex then forms at the lesion, which in *E. coli* requires each UvrA monomer recruiting UvrB, which has both ATPase and 5'-3' DNA helicase activity. An endonuclease(s) is then recruited to cleave at both sites of the lesion; in *E. coli*, all of the complex dissociates from the DNA except one UvrB that binds UvrC, a dual endonuclease that makes the required cleavages. Finally, the DNA segment is excised by the activity of DNA helicase(s); in *E. coli*, UvrB and UvrC are displaced, and DNA helicase UvrD excises the fragment (Schärer, 2013; Wozniak et al., 2022). In prokaryotes the excised fragment is 12-13 nucleotides long, while in eukaryotes it is 24-32 nucleotides long (Sancar, 1996). Finally, the gap is then filled by DNA polymerase(s) and ligated by a DNA ligase(s); in *E. coli* these are Pol I and DNA ligase I, respectively, while in humans these are Pol δ , Pol ϵ , or Pol κ depending on the pathway and DNA ligase I or ligase III α -XRCC1, respectively (Moolenaar et al., 2000; Schärer, 2013; Spivak, 2015; Wozniak et al., 2022).

The NHEJ pathway is one of the two major DSB repair pathways in eukaryotes (Chang et al., 2017; Mao et al., 2008; Radhakrishnan et al., 2014). In fact, NHEJ is mainly present in eukaryotes; only a few bacteria possess a simplified NHEJ mechanism, instead using homology-directed mechanism alongside an alternative end joining mechanism to repair DSBs, and is rare in archaea, with only *Methanocella paludicola* having a complete NHEJ complex (Bartlett et al., 2013; Chayot et al., 2010; White, M.F. et al., 2018). In humans, NHEJ is the main, preferred DSB repair pathway, active during most stages of the cell cycle, predominantly G₁ phase (Chang et al., 2017; Mao et al., 2008). The first step in NHEJ is synapsis, when a DSB is recognised by the Ku heterodimer composed of KU70 and KU80, which binds to each of the DSB ends (Chang et al., 2017; Lieber et al., 2003; Zhao et al., 2017). The

next step is end-processing, the contents of which are variable depending on the type of DSB ends. Regardless, the DNA-dependent protein kinase catalytic subunit (DNA-PK_{CS}) and the endonuclease Artemis are recruited to the bound Ku; DNA-PK_{CS} is autophosphorylated and activates Artemis endonuclease activity, which resects any overhangs (Chang et al., 2017). The final step is ligation, when DNA gaps are filled in by Pol μ and Pol λ , and the ends are ligated by a complex of and DNA ligase IV (Lieber et al., 2003). Despite being the favoured DSB mechanism in humans, it is intrinsically error-prone, which can ultimately promote genome instability and hence genetic diseases such as cancer (Chang et al., 2017; Lieber et al., 2003; Mao et al., 2008; Sishc et al., 2017).

1.3.2 Homologous Recombination

The HR pathway is the second major DSB repair pathway in eukaryotes, and unlike NHEJ is universal across all domains of life (Chen, 2013; Clay et al., 2021; White, M.A. et al., 2018; Mao et al., 2008; White, M.F. et al., 2018). In addition to DSB repair, HR also functions in repair of other DNA lesions such as interstrand crosslinks and has shown evidence of repairing single-strand nicks and gaps in mammalian cells (Jasin et al., 2013; Li et al., 2008). In order to carry out this repair, HR requires a homologous template, and so in eukaryotes it is predominately active during late S and G₂ phase due to the presence of a sister chromatid (Chang et al., 2017; Zhao et al., 2017). This presence of a sister chromatid means HR also has a key role in counteracting DNA replication stress (Costes et al., 2012). This is because if a replication fork stalls due to encountering a DNA lesion, HR can facilitate lesion bypass to continue fork progression, but this generates tracts of ssDNA; HR can then fill in gaps using the sister chromatid as a template (Costes et al., 2012; Prado, 2014). Furthermore, HR can even recover replication forks that have collapsed due to DNA lesions (Costes et al., 2012; Saleh-Gohari et al., 2005). Given HR is involved in these processes, combined with its overall much higher fidelity than NHEJ, HR is hence crucial for maintaining genomic stability (Li et al., 2008; Mao et al., 2008; Moynahan, 2010; Zhao et al., 2017). Despite this, in humans NHEJ remains the preferred SB repair pathway even during S phase and G₂ phase (Beucher et al., 2009, Chang

et al., 2017), with the 'ratio of NHEJ to HR in wild-type (WT) mammalian somatic cells estimated to be 4:1' (Chang et al., 2017).

Even though HR has an overall much higher fidelity than NHEJ, it can still 'be up to 1000-fold [more] mutagenic than normal DNA replication' (Rodgers et al., 2016), with its mutagenic potential varying depending on the type of HR. This is because HR DSB repair can occur via three distinct sub-pathways, which are, in order of least to most mutagenic: double-strand break repair (DSBR) via double Holliday junction intermediates, synthesis-dependent strand annealing (SDSA), and BIR (Li et al., 2008; Rodgers et al., 2016). The first two sub-pathways repair two-ended DSBs, while BIR repairs one-ended DSBs (Zhao et al., 2024). Regardless of sub-pathway, HR can be divided into three stages: presynapsis, synapsis, and postsynapsis (Li et al., 2008). This is a complex process involving a multitude of proteins (Wright et al., 2018). During presynapsis, DSB ends are recognised in dsDNA, and resected by endonucleases in to produce 3'-OH-terminal ssDNA tails (Li et al., 2008; White, M.F. et al., 2018; Wright et al., 2018). The newly resected ends are initially bound to by a SSB protein (RPA in humans and most archaea and SSB in *E. coli*) to protect the ssDNA (Inoue et al., 2008; Wardell et al., 2017; White, M.F. et al., 2018; Wright et al., 2018). Recombinase mediator proteins then facilitate displacement of the SSB protein with a DNA recombinase (RAD51 in humans, RecA in *E. coli*, and RadA in archaea) which forms a nucleoprotein presynaptic filament with the ssDNA (Li et al., 2008; Liu et al., 2011; Seitz et al., 1998; White, M.F. et al., 2018; Wright et al., 2018). During synapsis, the filament invades another duplex DNA, and the recombinase catalyses homology search in the invaded duplex to find a sequence complementary to its bound ssDNA to act as a homologous template (Li et al., 2008; Wright et al., 2018). Upon finding the correct sequence, the 3' end of the invading strand intertwines with the strand containing the sequence, forming a displacement loop (D-loop) (Wright et al., 2018). It is during postsynapsis that the three distinct HR sub-pathways become apparent (Li et al., 2008; Wright et al., 2018).

During postsynapsis, the D-loop can be processed in a number of ways depending on the HR mechanism, each of which ultimately results in a crossover product or a non-crossover product that may have loss of heterozygosity (Li et al., 2008). SDSA is the preferred HR sub-

pathway in human somatic cells (Wright et al., 2018). After the D-loop has formed, a DNA polymerase (Pol δ with potential help from Pol ϵ in humans, Pol III in *E. coli*, and family-B and family-D polymerases in archaea) uses the homologous template to extend the D-loop (Hogrel et al., 2020; Wright et al., 2018). The D-loop then dissociates due to DNA helicase activity (Li et al., 2008; Wright et al., 2018), and the 'disengaged invading strand reanneals with the second end of the DSB', resulting in a non-crossover product. The synapsis and postsynapsis steps may need to be repeated multiple times to complete SDSA (Li et al., 2008). The main mutagenic potential risk in this HR sub-pathway is template-switching occurring (Rodgers et al., 2016). Meanwhile, BIR employs a similar mechanism to SDSA, but the D-loop migrates as a non-canonical replication fork that extends the invading strand by replicating the template chromosome's entire distal arm (Donnianni et al., 2019; Li et al., 2008). This means the second end of the DSB is never engaged, and so the non-crossover product has loss of heterozygosity (LOH), thus giving BIR the highest mutagenic potential among the HR sub-pathways (Li et al., 2008; Rodgers et al., 2016). In DSBR, the D-loop is extended but engages the second end of the DSB, forming a double Holliday junction (Li et al., 2008). The double Holliday junctions can be processed by a resolvase endonuclease (GEN1 in humans, RuvC in *E. coli*, and Hjc or Hje in archaea) into a crossover or non-crossover product (Kvaratskhelia et al., 2002; Li et al., 2008; Middleton et al., 2004; Shah Punatar et al., 2017; West, 2009). Alternatively, the double Holliday junctions can be dissolved, in which branch migration results in formation of a hemicatenane that is subsequently resolved by a topoisomerase (Li et al., 2008). In humans, this is carried out by the BRT complex composed of the DNA helicase BLM, topoisomerase III α , and scaffold proteins RMI1 and RMI2 in humans (Xue et al., 2013). Dissolution results in only noncrossover products (Li et al., 2008). While DSBR has the least mutagenic potential among the three HR sub-pathways, it still risks template switching like SDSA (Rodgers et al., 2016). The different HR sub-pathways and outcomes are illustrated in Figure 1.8.

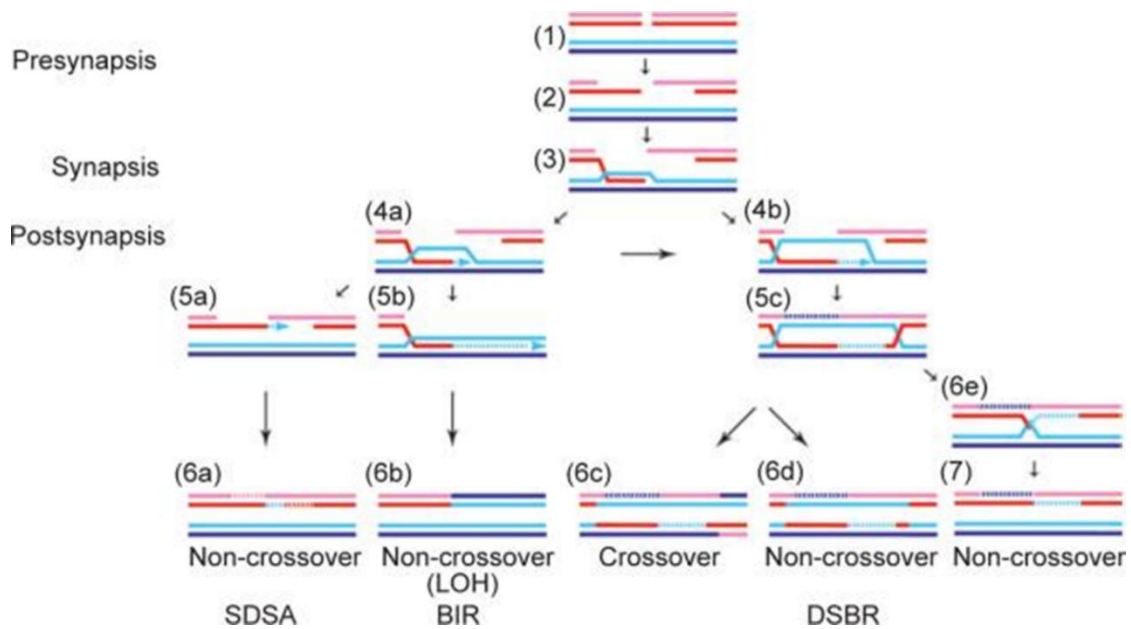


Figure 1.8 Homologous recombination sub-pathways and outcomes.

Homologous recombination (HR) DNA double-strand break (DSB) repair occurs in three stages: presynapsis, synapsis, and postsynapsis. During presynapsis, DSB ends are recognised (step 1), and subsequently resected to produce 3'-OH-terminal ssDNA tail. These are coated in single-stranded binding protein, which is displaced by a DNA recombinase, forming a presynaptic nucleoprotein filament (step 2). During synapsis, the filament invades another DNA duplex to find a complementary homologous template, which it intertwines with, resulting in a D-loop (step 3). Postsynapsis can use three different sub-pathways. In synthesis-dependent strand annealing (SDSA), the D-loop is extended by a DNA polymerase using the homologous template (step 4a), then dissociates from the template (step 5a), giving a non-crossover product (step 6a). In break-induced replication (BIR), the D-loop is converted into a replication fork that extends the invading strand by replicating the template chromosome's entire distal arm (step 5b), giving a crossover product with loss of heterozygosity (LOH) (step 6b). In double-strand break repair (DSBR), the second DSB end is engaged, forming double Holliday junctions (steps 4b and 5c). The double Holliday junctions can be processed by a resolvase endonuclease, generating crossover or non-crossover products (steps 6c and 6d). Alternatively, they can undergo branch migration (step 6e) before being dissolved by a topoisomerase complex, giving a non-crossover product (step 7). Figure and legend are adapted from Figure 1 in Li et al. (2008).

1.4 HELQ

1.4.1 Discovery, Gene, and Nomenclature

DNA helicases are a key type of enzyme involved in HR DSB repair (Huselid et al., 2020). DNA helicases are universal across all domains of life and are defined as motor proteins that couple nucleoside triphosphate (NTP) hydrolysis with unwinding of duplex DNA (Bosh et al., 2013; O'Donnell et al., 2018; Suhasini et al., 2010; Tuteja et al., 2004; Zhang, Y.M et al., 2024). To do this, they bind to ssDNA and/or dsDNA and utilise energy from NTP (typically ATP) hydrolysis via their NTPase activity to translocate along the DNA with a specific polarity (3'-5' or 5'-3') and break the hydrogen bonds between nucleobases in duplex DNA to unwind it (Bianco, 2022; Brosh, 2013; Kim et al., 2009; Tuteja et al., 2004). Hence, DNA helicases 'are involved in almost all aspect of nucleic acid metabolism' (Tuteja et al., 2006). During HR DSB repair, different DNA helicases are involved in a variety functions: promoting or supressing DSB end resection, displacing SSB proteins from resected ends, disrupting D-loops, signalling at replication forks, and migrating and/or dissolving Holliday junctions (Estep et al., 2018; Harmon et al., 1998; Huselid et al., 2020; Nishimura et al., 2012; Veaute et al., 2005). One such DNA helicase in humans is Helicase POLQ-like (HELQ) (He et al., 2023; Tang et al., 2023).

The discovery of human HELQ originates from the work of Boyd et al. (1976), who began mutation sensitivity screenings in *Drosophila melanogaster*. These screenings continued, and the *D. melanogaster mus308* gene was identified by Boyd et al. (1981), then cloned by Harris et al. (1996). Oshige et al. (1999) were able to first purify the protein product, later designated DNA polymerase Pol θ , which has an N-terminal helicase domain and a C-terminal DNA polymerase domain (Beagan et al., 2017; Pang et al. 2005). Sharief et al. (1999) identified the gene for human Pol θ , also known as POLQ, and the protein was first purified by Seki et al. (2003), who showed that the helicase domain is inactive. The N-terminal region human homologue with an active helicase domain was discovered by Marini et al. (2002). They isolated the human gene from complementary DNA (cDNA) expression libraries, subsequently overexpressed the protein in insect Sf9 cells, purified it, subjected it to DNA binding analysis, and designated it HEL308. The gene's cytogenetic location was identified as 4q21, specifically 4q21.23 (NCBI, 2025b). Due to its relationship with POLQ, HEL308 was later redesignated as

Helicase POLQ-like (HELQ) (Takata et al., 2013; Woodman et al., 2011a; Woodman et al., 2011b).

It was later discovered that HELQ has an archaeal homologue: Hel308. This was first separately discovered in two different archaeal species (Fujikane et al., 2005; Guy et al., 2005). Guy et al. (2005) expressed *Methanothermobacter thermautotrophicus* protein Hel308a in *E. coli* strain *dnaE486*, whose Pol III point mutation causes ‘replication fork stalling at semi-permissive or restrictive temperatures’; this resulted in a synthetically lethal phenotype. They then overexpressed the protein in BL21-AI *E. coli*, purified it, and subjected it to biochemical analyses. Fujikane et al. (2005) fractionated *Pyrococcus furiosus* cell extracts via chromatography, which they then screened for Holliday junction resolution activity, which protein Hjm showed. They then prepared a heat-stable protein library using *P. furiosus* genomic DNA and cloned the gene responsible from it, overexpressed it in BL21-CodonPlus(DE3)-RIL *E. coli*, purified it, and subject it to biochemical analyses. Fujikane et al. (2006) later carried out further biochemical and genetic analyses, which showed that Hjm was a Hel308a homologue. The archaeal homologue was subsequently designated Hel308 (Büttner et al., 2007; Richards et al., 2008; Woodman et al., 2007). While HELQ is conserved across eukaryotes and archaea, bacteria and fungi lack a HELQ orthologue (Richards et al., 2008).

1.4.2 Classification

In order to classify helicases, there are broadly six superfamilies: superfamilies 1-6 (SF1-6) (Singleton et al., 2007). Superfamily classification is based on factors such as core domain motif conservation, oligomeric state, and sequence homology (Singleton et al., 2007; Meir et al., 2021; Menon et al., 2025). SF1 and SF2 helicases are the largest superfamilies, and both possess 12 conserved domain motifs in two RecA-like domains at their core (Fairman-Williams et al., 2010; Singleton et al., 2002). Two of these motifs, Walker A and Walker B, are highly conserved across the two superfamilies, being located at the junction between the two RecA-like domains, where they function as an interface for ATP-binding and hydrolysis activity, thus providing the helicase core’s ATPase activity (Enemark, et al., 2008; Singleton et al., 2002;

Singleton et al., 2007). These cores are ‘often flanked by N- and C-terminal auxiliary domains or extensions, which have been shown to mediate substrate specificity or possess regulatory functions’ (Lang et al., 2024). SF3-6 helicases only possess one RecA-like domain per monomer (Byrd et al., 2012). In addition, SF1 and SF2 function as monomers or dimers, while SF3-6 function as single or double hexameric rings (Byrd et al., 2012; Enemark, et al., 2008; Gao et al., 2020; Singleton et al., 2007). The six superfamilies are summarised in Table 1.4.

Table 1.4 Summary of the six helicase superfamilies.

Superfamily	Oligomeric Structure	Main Substrate	Translocation Polarity	Other Key Structural Characteristics	Representative Examples
SF1	Monomer or dimer	DNA or RNA	3'-5' or 5'-3'	Two core Rec-A-like domains with 12 conserved motifs. Walker A and Walker B motifs provide ATPase activity (Enemark, et al., 2008; Fairman-Williams et al., 2010; Singleton et al., 2002; Singleton et al., 2007).	PcrA, UvrD, Pif1, and RecB
SF2	Monomer or dimer	DNA or RNA	3'-5' or 5'-3'	Similar to SF1, but also possess an N-terminal Q-motif before Walker A (Sarvari et al., 2026). Also possess a conserved DExH/D box motif block (Cencic et al., 2013).	Ded1, Rad54, RecG, and Hel308/HELQ

SF3	Hexameric ring	DNA or viral RNA (Singleton et al., 2007)	3'-5'	One core rec-A like domain of approximately 100 residues coordinated by five motifs (Byrd et al., 2012; Sarvari et al., 2026).	Papillomavirus E1, parvovirus Rep/NS1, and Simian virus 40 large T antigen (Sarvari et al., 2026)
SF4	Hexameric ring	DNA	5'-3'	One core Rec-A like domain and coordinated by six motifs (Byrd et al., 2012; Sarvari et al., 2026).	Bacteriophage T7 gp4 (Sarvari et al., 2026)
SF5	Hexameric ring	RNA	5'-3'	One core Rec-A like domain (Byrd et al., 2012).	Rho (Sarvari et al., 2026)
SF6	Hexameric ring	DNA	3'-5' or 5'-3'	One core Rec-A like domain (Byrd et al., 2012).	MCMs (Sarvari et al., 2026)

Table is adapted from Table 1 in Perera et al. (2022).

Within the superfamilies are families containing least three proteins and groups containing two proteins; families can in turn contain subfamilies (Fairman-Williams et al., 2010). These categories are classified based on typical features in three main factors: sequence, regarding sequence motifs and conservation; structure, regarding characteristic features and sequence motif conservation; and mechanisms, regarding nucleic acid substrate, translocation polarity, and biochemical activities (Bhattacharyya et al., 2014; Fairman-Williams et al., 2010; Singleton et al., 2007). SF2 is the largest superfamily, containing 11 families (Byrd et al., 2012; Fairman-Williams et al., 2010; Wyszomirski et al., 2012). SF2 families are classified together due to the conservation of a motif block consisting of DEAD, DEAH, and other variants, termed the

DExH/D box. They are then sub-divided based on how their conserved motifs relate to and couple their NTPase activity with substrate binding and unwinding (Cencic et al., 2013). These families are diverse in terms of their sequences, structures, and mechanisms (Byrd et al., 2012; Fang et al., 2023; Fairman-Williams et al., 2010), with roles in 'DNA transcription and repair, RNA metabolism, and chromosome rearrangement' (Lou-Hing, 2024). One small SF2 family is the Ski2-like family, consisting of large multidomain RNA helicases (Johnson et al., 2013), which are 'involved in RNA processing, splicing, and degradation pathways' (Tang et al., 2023), typically as parts of complexes (Johnson et al., 2013). Ski-2-like helicases are defined by their core, which like other SF2 families has up to 12 conserved motifs across the two RecA-like domains in their core (Batista et al., 2024; Fairman-Williams et al., 2010). These cores are ring-like, comprised of four domains in total, also containing a winged helix domain (WHD) and ratchet domain; this is their DExH/D box (Johnson et al., 2013). HELQ and Hel308 possess these domains and are thus members of the Ski2-like family (Batista et al., 2024; Tang et al., 2023).

1.4.3 Structure and Function

HELQ is comprised of 1,101 residues with a molecular mass of 124.5 kDa (Marini et al., 2002). HELQ can be divided into two main regions: a non-catalytic, intrinsically disordered N-terminal N-HELQ comprised of residues 1-240, and a catalytic C-terminal C-HELQ comprised of residues 241-1101 (Jenkins et al., 2021). It contains a POLD3-interacting region (PRI) (residues 1-76) and six domains, with N-HELQ containing the PRI and one domain, and the remaining five domains in C-HELQ (He et al., 2023; Jenkins et al., 2021; Zhao et al., 2024). Domain 1 (residues 128-237) is the domain in N-HELQ, being a 5' accessory four-helix bundle PWI-like domain (Jenkins et al., 2021). Domains 2 (residues 248-518) and 3 (residues 519-737) are the conserved RecA-like domains with their Walker A and Walker B motifs characteristic of a SF2 helicase (Byrd et al., 2012; Jenkins et al., 2021; Zhao et al., 2024). Domain 3 also contains a β -hairpin structure (Zhao et al., 2024). Domain 4 (residues 738-840) is the WHD, a winged-shaped spiral folding, while domain 5 (residues 841-1005) is the ratchet domain, a seven- or eight-helix bundle; together with domains 2 and 3 they make the DExH/D box helicase core characteristic of a Sk2-like family helicase (Batista et al., 2024; Johnson et al., 2013; Zhao et al., 2024). Finally, domain 6 (residues 1006-1101) is a 3' accessory helix-loop-helix (HLH)

structure (Zhao et al., 2024). The HELQ PRI and domains contain a number of functional residues; HELQ residue information is illustrated in Figure 1.9.

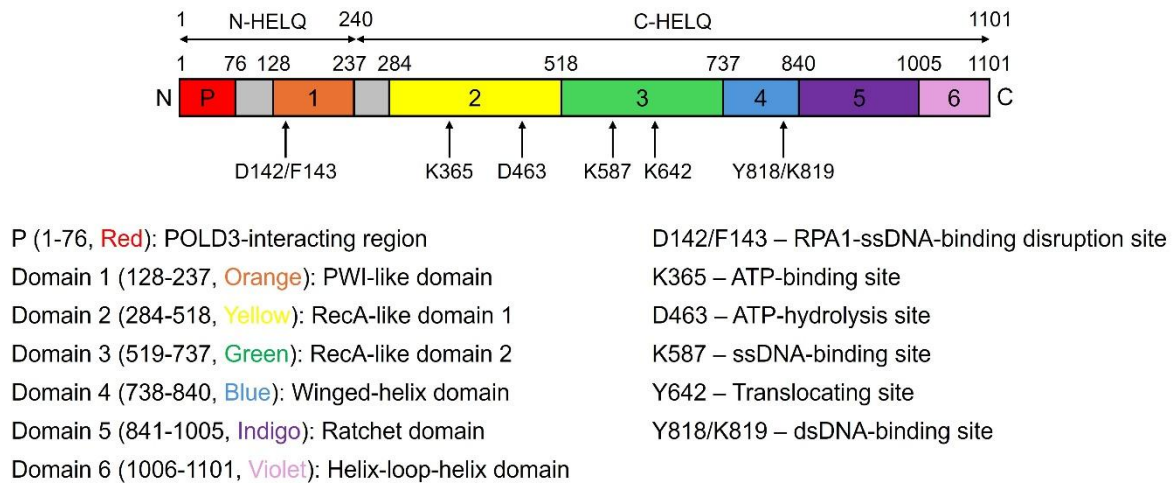


Figure 1.9 HELQ amino acid residue information.

HELQ contains two regions, N-HELQ and C-HELQ, with a POLD3-interacting region and six domains as denoted by letter/number and colour. POLD3-interacting region and domain residue ranges in addition to functional residue sites are indicated (He et al., 2023; Jenkins et al., 2021; Zhao et al., 2024). Figure is adapted from Figure 2A in Zhao et al. (2024).

The crystal structure of HELQ is currently unresolved (Zhao et al., 2024), but this is untrue of archaeal Hel308, which was first resolved in *Archaeoglobus fulgidus* by Büttner et al. (2007), and later in *Sulfolobus solfataricus* by Richards et al. (2008) and in *Pyrococcus furiosus* by Oyama et al. (2009). Despite this lack of resolution, bioinformatic analysis of a HELQ 3D structure predicted by the software AlphaFold 2 showed high similarity with *Archaeoglobus fulgidus* Hel308 structure (Zhao et al., 2023). This reflects a high sequence and structural conservation of HELQ orthologues across eukaryotes and archaea, although only metazoan HELQ proteins possess an N-HELQ region and PWI-like domain (Jenkins et al., 2021; Zhao et al., 2024). HELQ structure as predicted by the software AlphaFold 3 is shown in Figure 1.10 and illustrated with the PRI and domains in Figure 1.11.

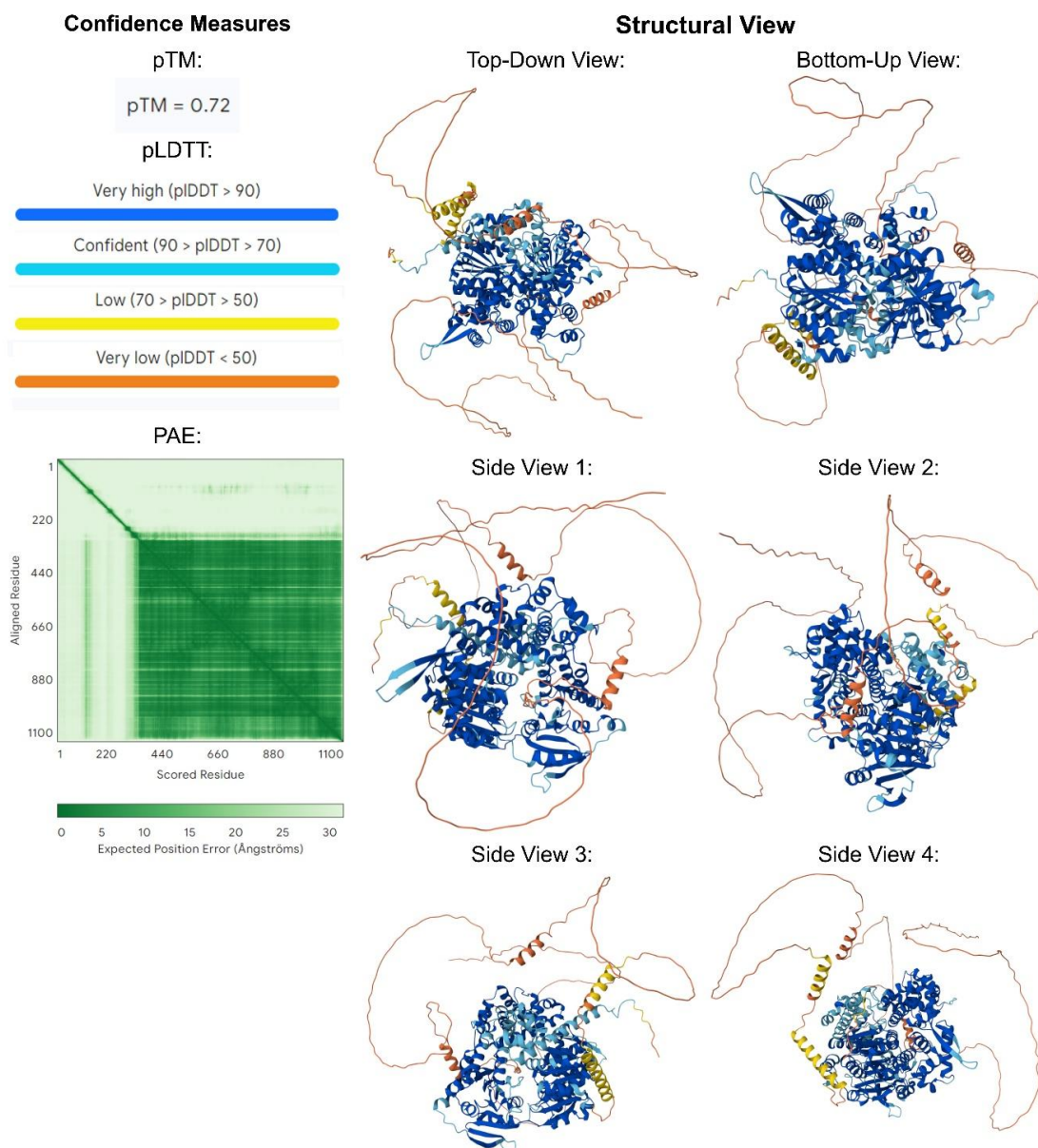


Figure 1.10 HELQ structure predicted by the software AlphaFold 3.

Structure is shown from multiple orientations. AlphaFold 3 uses multiple confidence measures (Abramson et al., 2024). Predicted Template Modelling (pTM) score measures overall structure prediction accuracy, ranging from 0-1 with threshold of 0.5; a higher score indicates greater accuracy (Wee et al., 2024). Predicted Local Distance Difference Test (pLDDT) score measures individual residue structure prediction confidence, ranging from 0-100; a higher score indicates greater confidence. pLDDT score is denoted by residue colour. Predicted Align Error (PAE) assesses overall structural prediction confidence, measured in angstroms. Lower PAE indicates greater prediction confidence (Pratt et al., 2025).

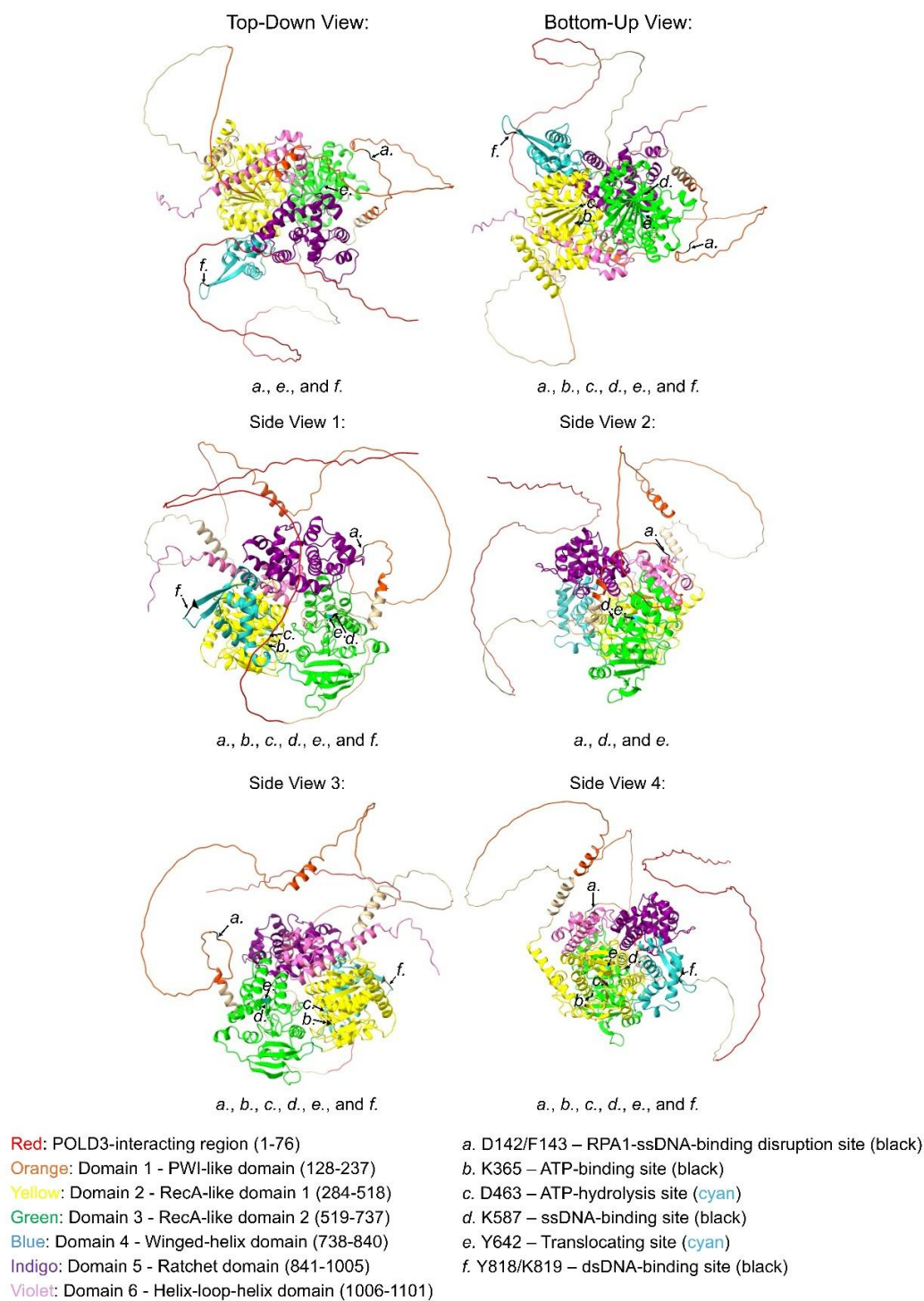


Figure 1.11. HELQ predicted structure organisation.

HELQ structure as predicted by AlphaFold 3 (Figure 1.10) is shown from multiple orientations. POLD3-interacting region and domains are indicated by colour. Functional residues are coloured black or cyan and indicated by letter where visible, with those indicated for each orientation listed underneath.

While Ski2-like family helicases are mainly RNA helicases, HELQ and Hel308 are DNA helicases (Byrd et al., 2012; Fang et al., 2023; Johnson et al., 2013; Zhao et al., 2024). As such, HELQ has four main biochemical functions for this: DNA binding, ATPase activity, DNA translocation, and DNA unwinding. HELQ achieves these through its structure, primarily through its C-HELQ region (He et al., 2023; Zhao et al., 2024). HELQ can bind to both ssDNA and dsDNA, with preference for ssDNA (Zhao et al., 2024). Its ssDNA-binding requires RecA-like domain 2 K587 and dsDNA-binding requires WHD Y818/K819 (Northall et al., 2017; Zhao et al., 2023). ssDNA-binding is required for its ATPase activity, which utilises RecA-like domain 1 K365 in ATP-binding and requires the domain's D463 for ATP-hydrolysis (Jenkins et al., 2021; Marini et al., 2002). This ATPase activity in turn provides energy for HELQ DNA translocation, which is 5'-3' (Jenkins et al., 2021; Marini et al., 2002). Its DNA translocation requires RecA-like domain 2 Y624 and the ratchet domain (Büttner et al., 2007; Jenkins et al., 2021). It has been shown that in 5'-3' translocation of archaeal Hel308, the entire helicase core is required, with the WHD assisting the ratchet domain, coordinated by the RecA-like domains (Lever et al., 2023). DNA translocation, alongside DNA binding and ATPase activity, feeds into DNA unwinding (Zhao et al., 2024). In archaeal Hel308, initial dsDNA strand separation is ATP-independent, being unwound by the RecA-like domain B β -hairpin; subsequent ATP-dependent DNA unwinding occurs tandemly with DNA translocation, with the HLH domain positioning dsDNA bound by the WHD be unwound by the ratchet domain (Büttner et al., 2007; Northall et al., 2017; Woodman et al., 2007). Thus, with regards to its helicase activity, HELQ can be summarised as a ssDNA-dependent ATPase and DNA helicase that translocates 3'-5' (Jenkins et al., 2021).

HELQ also possess a fifth main biochemical function that opposes its DNA helicase activity: DNA annealing. This utilises motif IVa in RecA-like domain B, consisting of residues 641-648; HELQ^{Y642A}, while inactivating DNA translocation, results in DNA hyper-annealing (Jenkins et al., 2021; Lou-Hing, 2024; Zhao et al., 2024). This function is ATP-dependent in the presence of RPA, otherwise ATP is unessential (Anand et al., 2022). Due to its main biochemical functions, including DNA annealing, HELQ is involved in multiple processes. It has a key role in

DNA replication, where its DNA helicase activity 'is crucial for unwinding the lagging strand of the replication fork' to continually restart lagging strand synthesis, but its main role is in maintaining genomic integrity, thus combating genomic instability and consequential genetic diseases such as cancer (Zhao et al., 2024). HELQ is fundamental to this as a caretaker gene, such that *HELQ*^{-/-} in human U2OS cells triggers cancer phenotypes, while HELQ overexpression is implicated in epithelial ovarian cancer, where it promotes resistance to DNA crosslinking-based treatments (Jenkins et al., 2021; Liu, D.N. et al., 2017; Long et al., 2018).

HELQ maintains genomic integrity through functioning in DNA DSB repair. It has been implicated to various degrees in multiple pathways: MMEJ, single-strand annealing (SSA), and HR, comprising three sub-pathways of DSBR, SDSA, and BIR. BIR repairs one-ended DSBs; the rest of these pathways repair two-ended DSBs (Zhao et al., 2024). Anand et al. (2022) showed that *HELQ*^{-/-} U2OS cells had impaired MMEJ and SSA, and suggested HELQ could be involved in initial DSB end resection. Previously, Kamp et al. (2021) showed that HELQ in *Caenorhabditis elegans* was essential for MME and SSA, and crucial for SDSA. Like SDSA, BIR shares the principal of D-loop extension rather than DSBR double Holliday junction formation, and a picture has been emerging of HELQ involvement in the HR sub-pathway (He et al., 2023; Li et al., 2008; Zhao et al., 2024). During BIR, including at stalled replication forks, RAD51 on the invading nucleoprotein filament stimulates ssDNA-bound HELQ DNA helicase activity (Anand et al., 2022; He et al., 2023; Zhao et al., 2024). As it translocates and unwinds the invaded duplex, it displaces RPA, allowing the filament to anneal to the necessary homologous template (Anand et al., 2022). HELQ is also involved during the later Pol δ DNA synthesis step, where HELQ acts to limit mutagenic BIR DNA synthesis. This role of HELQ is of interest to the Bolt lab, who posited a model for the process in He et al. (2023). This is summarised in Figure 1.12.

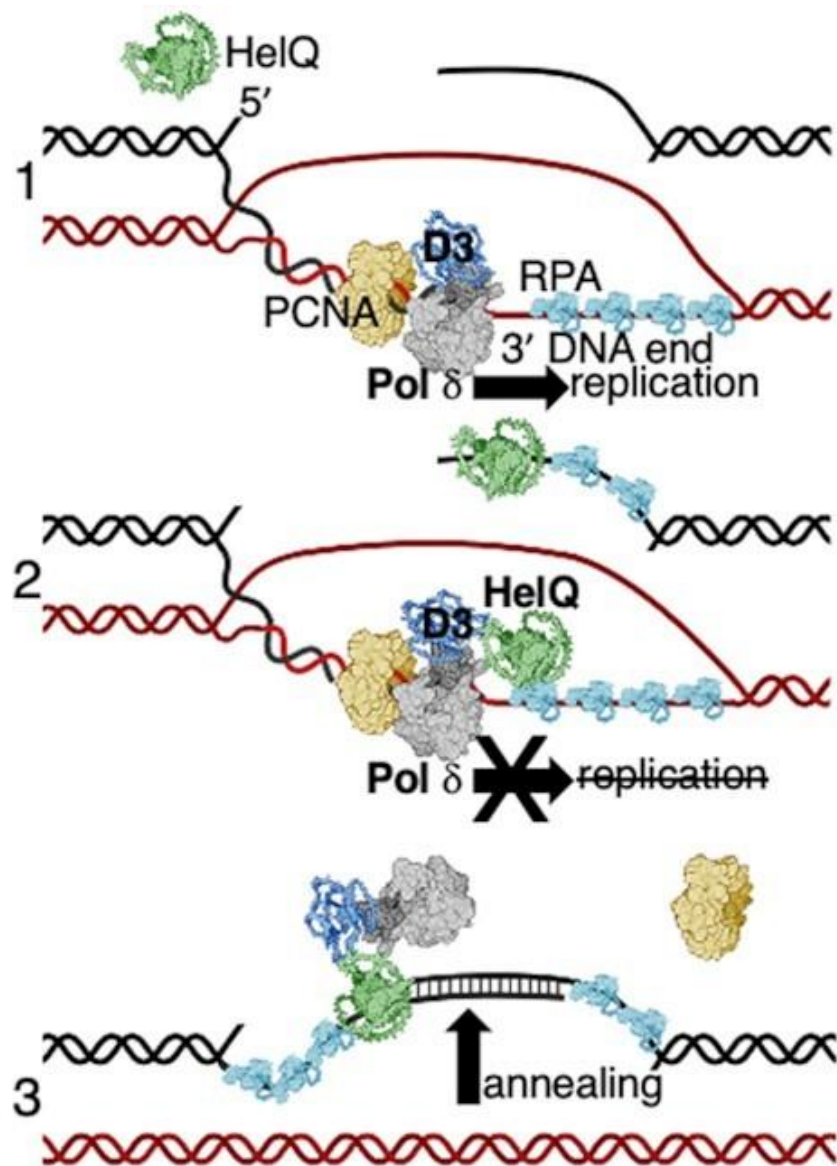


Figure 1.12 Bolt lab model of HELQ limiting mutagenic DNA synthesis during the homologous recombination sub-pathway break-induced replication (BIR).

1. During BIR, DNA polymerase Pol δ forms a holoenzyme with PCNA and DNA single-stranded binding protein RPA, pairing the 3' end of broken single-stranded DNA with an unbroken strand in an invaded duplex, using the latter strand as a template for DNA synthesis as the D-loop migrates in a non-canonical replication fork (Donnianni et al., 2019; He et al., 2023). 2. DNA helicase HELQ then interacts with RPA subunit RPA1, which orients HELQ towards the POLD; HELQ displaces the RPA and interacts with Pol δ DNA subunit POLD3, inhibiting the DNA synthesis. 3. PCNA dissociates from Pol δ, the latter of which stimulates HELQ to anneal the nascent DNA with its parent duplex (He et al., 2023; Jenkins et al., 2021). Figure is Figure 6 in He et al. (2023).

In developing this model, the Bolt lab has made progress in identifying the regions of HELQ, Pol δ , and RPA involved in their interaction. HELQ interacts with subunits within Pol δ and RPA (He et al., 2023; Jenkins et al., 2021). Pol δ is a heterotetramer comprised of four subunits (see Section 1.2.2); HELQ interacts with POLD3 (He et al., 2023; Zhou et al., 2012). RPA is a heterotetramer comprised of three subunits: RPA1, RPA2, and RPA3; HELQ interacts with RPA1 (Hass et al., 2012; Jenkins et al., 2021; Wold et al., 1997). HELQ interacts with POLD3 and RPA1 through its N-HELQ region, via its extreme N-terminal residues 1-76 for POLD3 and its PWI-like domain for RPA1, with PWI-like domain D142/F143 necessary to disrupt ssDNA-RPA interaction. However, the POLD3 and RPA1 residues necessary for interaction with N-HELQ are currently unknown (He et al., 2023; Jenkins et al., 2021).

1.5 Project Aims

The initial aim of this project was to investigate further the molecular mechanisms of POLD3-N-HELQ interaction. This utilised the Bolt lab's BL21-AI *E. coli* bimolecular fluorescence complementation (BiFC) assay system, which led to a second aim of developing the original time-course method into an endpoint method to provide quantitative data. Using this method, the level of interaction between POLD3 Quincunx motif mutants, and later POLD3 intrinsically disordered protein regions (IDPRs), with N-HELQ were assessed in order to locate potential sites of POLD3-N-HELQ interaction. This then gave rise to a third aim, which was to determine if the endpoint method could be used to show RPA1-N-HELQ interaction and hence locate potential interaction sites in RPA1. Following this, a fourth and final aim was to use an *in vivo* approach to attempt to assess full-length protein RPA-HELQ interaction via PE3b prime editing of U2OS cells.

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Bacterial Strains

Table 2.1 Details of *E. coli* strains used in this work.

Strain Name	Supplier	Features	Use
DH5 α	Invitrogen	<i>recA1</i> mutation to increase insert stability and <i>endA</i> mutation to increase DNA yield and quality (Thermo Fisher, 2025a). Genotype: <i>F⁻ ϕ80lacZΔM15 Δ(lacZYA-argF)U169 recA1 endA1 hsdR17(<i>r_K⁻</i>, <i>m_K⁺</i>) phoA supE44 λ⁻thi-1 gyrA96 relA1</i> (Thermo Fisher, 2025b).	Plasmid cloning
BL21-AI	Invitrogen	<i>lon</i> and <i>ompT</i> mutations to remove Lon and OmpT proteases, respectively, to reduce heterologous protein degradation. T7RNAP- <i>tetA</i> insertion into the <i>araB</i> locus encodes an arabinose-inducible T7 RNA polymerase and tetracycline for strain identity verification (Thermo Fisher, 2025d). Genotype: <i>F⁻ ompT hsdS_B (<i>r_B⁻</i>, <i>m_B⁻</i>) gal dcm araB::T7RNAP-tetA</i> (Thermo Fisher, 2025b).	Protein overexpression

2.1.2 Plasmids

Table 2.2 List of plasmids used in this work.

Plasmid	Description	Resistance	Source
pBAD-FLmVenus	DNA encoding full-length mVenus fluorescent protein cloned into pBAD/HisA vector via NheI and EcoRI restriction sites.	Ampicillin	Liu He (He et al., 2023)

pBAD-NmVenus	DNA encoding mVenus N-terminal residues 1-154 cloned into pBAD/HisA vector via NheI restriction site.	Ampicillin	Liu He (He et al., 2023)
pRSF-CmVenus	DNA encoding mVenus C-terminal residues 155-238 cloned into pRSF-1b vector via NcoI restriction site.	Kanamycin	Liu He (He et al., 2023)
pBAD-NmVenus-PoID3	DNA encoding human DNA polymerase POLD subunit POLD3 residues 2-466 cloned into pBAD-NmVenus vector via XhoI and HindIII restriction sites.	Ampicillin	Liu He (He et al., 2023)
pRSF-CmVenus-NHelQ	DNA encoding human DNA helicase HELQ N-terminal (N-HELQ) residues 2-240 cloned into pRSF-CmVenus vector via KpnI and XhoI restriction sites.	Kanamycin	Liu He (He et al., 2023)
pALH008	pBAD-NmVenus-POLD3 with the DNA encoding the POLD3 319-KKRRR-323 'Quincunx' motif mutated to encode GGGGG (pBAD-NmVenus-PoID3 ^{QC}).	Ampicillin	Anna Lou-Hing (Lou-Hing, 2024)
pOLD038	pBAD-NmVenus-POLD3 with DNA encoding the POLD3 319-KKRRR-323 Quincunx motif mutated to encode KKSRR (R321S) (pBAD-NmVenus-PoID3 ^{KKSRR}).	Ampicillin	Olivia Downs
pOLD039	pBAD-NmVenus-POLD3 with DNA encoding the POLD3 319-KKRRR-323 Quincunx motif mutated to encode KSRRR (K320S) (pBAD-NmVenus-PoID3 ^{KSRRR}).	Ampicillin	Olivia Downs
pOLD040	pBAD-NmVenus-POLD3 with DNA encoding the POLD3 319-KKRRR-323 Quincunx motif mutated to encode KKDRR (R321D) (pBAD-NmVenus-PoID3 ^{KKDRR}).	Ampicillin	Olivia Downs

pOLD041	pBAD-NmVenus-POLD3 with DNA encoding the POLD3 319-KKRRR-323 Quincunx motif mutated to encode AKKRA (K319A/R323A) (pBAD-NmVenus-PoID3 ^{AKKRA}).	Ampicillin	Olivia Downs
pOLD042	pBAD-NmVenus-POLD3 with DNA encoding 381-RKRKR-385 in the POLD3 putative NLS mutated to encode GGGGG (pBAD-NmVenus-PoID3 ^{PNLS}).	Ampicillin	Olivia Downs
pOLD001	DNA encoding POLD3 IDPR1 (residues 169-188) cloned into pBAD-NmVenus vector via XhoI and HindIII restriction sites (pBAD-NmVenus-PoID3 ^{IDPR1}).	Ampicillin	Olivia Downs
pOLD002	DNA encoding POLD3 IDPR2 (residues 199-232) cloned into pBAD-NmVenus vector via XhoI and HindIII restriction sites (pBAD-NmVenus-PoID3 ^{IDPR2}).	Ampicillin	Olivia Downs
pOLD003	DNA encoding POLD3 IDPR3 (residues 274-393) cloned into pBAD-NmVenus vector via XhoI and HindIII restriction sites (pBAD-NmVenus-PoID3 ^{IDPR3}).	Ampicillin	Olivia Downs
pOLD004	DNA encoding POLD3 IDPR4 (residues 406-466) cloned into pBAD-NmVenus vector via XhoI and HindIII restriction sites (pBAD-NmVenus-PoID3 ^{IDPR4}).	Ampicillin	Olivia Downs
p11d-tRPA(123)	Addgene plasmid #102613 that expresses the three subunits of the human ssDNA-binding protein RPA in <i>E. coli</i> (Addgene, 2025b).	Ampicillin	Marc Wold (Henricksen et al., 1994)

pAR001	DNA encoding human ssDNA-binding protein RPA subunit RPA1 residues 2-616 cloned into pBAD-NmVenus vector via KpnI and EcoRI restriction sites (pBAD-NmVenus-RPA1)	Ampicillin	This work
pAR002	DNA encoding human ssDNA-binding protein RPA subunit RPA2 residues 2-270 cloned into pBAD-NmVenus vector via XhoI and HindIII restriction sites (pBAD-NmVenus-RPA2).	Ampicillin	This work
SpyCas9 PE2*	Addgene plasmid #169850 that expresses an NLS-optimised Spy nickase Cas9 ^{H840A} reverse transcriptase M-MLVD ^{200N/T306K/W313F/T330P/L603W} fusion protein (nCas9-RT) in mammalian cells for prime editing (Addgene, 2025d).	Ampicillin	Scot Wolfe (Liu et al., 2021).
pU6-pegRNA-GG-acceptor	Addgene plasmid #132777 that contains that contains a cassette for cloning in a spacer via BsaI for expression of nCas9-RT pegRNA in prime editing (Addgene, 2025c).	Ampicillin	David Liu (Anzalone et al, 2019)
BPK1520	Addgene plasmid #65777 that contains a cassette for cloning in a spacer via BsmBI for expression of nCas9-RT sgRNA in prime editing (Addgene, 2025a).	Ampicillin	Keith Joung (Kleinstiver et al., 2015)
pAR003	pU6-pegRNA-GG-acceptor plasmid with DNA encoding pegRNA for <i>HELQ</i> D142A/F143A mutation.	Ampicillin	This work
pAR004	BPK1520 plasmid with DNA encoding sgRNA for <i>HELQ</i> D142A/F143A mutation.	Ampicillin	This work

pEGFP-C1	Vector that expresses the red-shifted GFPmut1 variant of wild-type GFP in mammalian cells (Addgene, 2002). The GFPmut1 can also be fused to the N-terminus of a partner protein (SnapGene, 2025).	Kanamycin	Clontech (SnapGene, 2025)
----------	---	-----------	---------------------------

2.1.3 Primers

EcoRI_RPA1_R was ordered from and synthesised by Eurofins Genomics. All other primers used in this work were ordered from and synthesised by Sigma-Aldrich. Primers were synthesised at 25 nM scale, desalt purified, delivered lyophilised dry, and resuspended in SDW to 100 µM stock concentration for storage at -20 °C.

Table 2.3 List of primers used in this work.

Primer Name	Template DNA	Use	5'-3' Sequence
KpnI_RPA1_v2_F	p11d-tRPA(123)	Forward primer for PCR amplification and Sanger sequencing of RPA1 DNA.	taatatggtaccatGTCGGCCA GCTGAGCG
EcoRI_RPA1_R	p11d-tRPA(123)	Reverse primer for PCR amplification and Sanger sequencing of RPA1 DNA.	tagccggaattcTCACATCAAT GCAC
XhoI_RPA2_v2_F	p11d-tRPA(123)	Forward primer for PCR amplification and Sanger sequencing of RPA2 DNA.	taatatctcgagTGGAAACAGT GGATTCG
HindIII_RPA2_v2_R	p11d-tRPA(123)	Reverse primer for PCR amplification and Sanger sequencing of RPA2 DNA.	cgtagcaagcttTTATTCTGCA TCTGTGG
XhoI_RPA3_v2_F	p11d-tRPA(123)	Forward primer for PCR amplification and Sanger sequencing of RPA3 DNA.	cgtaatctcgagGTGGACATG ATGGACTTG

HindIII_RPA3_v2_R	p11d-tRPA(123)	Reverse primer for PCR amplification and Sanger sequencing of RPA3 DNA.	gccgcgaagcttTCAATCATGT TGCACAATC
NheI_F	pBAD-NmVenus	Forward primer for PCR amplification and Sanger sequencing of NmVenus DNA.	atatGCTAGCATGGTGAGC AAGGGCGAG
NheI_R	pBAD-NmVenus	Reverse primer for PCR amplification and Sanger sequencing of NmVenus DNA.	atATGCTAGCGGCGGTGAT ATAGACGTTGTG
LKO.1 5'	pAR003	Forward primer for Sanger sequencing of DNA encoding prime editing pegRNA.	GACTATCATATGCTTACCG T
OS280	pAR004	Forward primer for Sanger sequencing of DNA encoding prime editing sgRNA.	CAGGGTTATTGTCTCATGA GCGG
HELQ_D142/F143_F	Extracted human U2OS cell genomic DNA	Forward primer for PCR amplification and Sanger sequencing of region surrounding DNA encoding HELQ D142/F143.	taataTGCAATGGCGCGATC TG
HELQ_D142/F143_R	Extracted human U2OS cell genomic DNA	Reverse primer for PCR amplification and Sanger sequencing of region surrounding DNA encoding HELQ D142/F143.	taatcTATTGGTGTCTGGGC C

2.1.4 Antibiotics

Table 2.4 List of antibiotics used in this work.

Antibiotic	Stock Concentration	Working Concentration
Ampicillin	100 mg/mL	100 µg/mL
Kanamycin	50 mg/mL	40 µg/mL

2.1.5 Commercial Enzymes

Table 2.5 List of commercial enzymes used in this work.

Enzyme	Supplier	Features
HindIII-HF	NEB	Restriction endonuclease that cuts sequence: 5'-A▼A G C T T-3' 3'-T T C G A▲A-5' (NEB, 2025f)
NheI-HF	NEB	Restriction endonuclease that cuts sequence: 5'-G▼C T A G C-3' 3'-C G A T C▲G'5' (NEB, 2025h)
AccI	NEB	Restriction endonuclease that cuts sequence: 5'-G▼T M K A C-3' 3'-C A K M T▲G-5' (NEB, 2025a)
BsaI-HFv2	NEB	Restriction endonuclease that cuts sequence: 5'-G G T C T C (N) ₁ ▼-3' 3'-C C A G A G (N) ₅ ▲-5' (NEB, 2025b)
PvuI-HF	NEB	Restriction endonuclease that cuts sequence: 5'-C G A T▼C G-3' 3'-G C▲T A G C-5' (NEB, 2025i)

BsmBI-v2	NEB	Restriction endonuclease that cuts sequence: 5'-C G T C T C (N) ₁ ▼-3' 3'-G C A G A G (N) ₅ ▲-5' (NEB, 2025c)
EcoRI-HF	NEB	Restriction endonuclease that cuts sequence: 5'-G▼A A T T C-3' 3'-C T T A A▲G-5' (NEB, 2025e)
KpnI-HF	NEB	Restriction endonuclease that cuts sequence: 5'-G G T A C▼C-3' 3'-C▲C A T G G-5' (NEB, 2025g)
XhoI	NEB	Restriction endonuclease that cuts sequence: 5'-C▼T C G A G-3' 3'-G A G C T▲C-5' (NEB, 2025n)
DpnI	NEB	Restriction endonuclease that cuts sequence: CH₃ 5'-G A▼T C-3' 3'-C T▲A G-5' CH₃ Only cuts when restriction site is methylated (NEB, 2025d).
NotI-HF	NEB	Restriction endonuclease that cuts sequence: 5'-G C▼G G C C G C-3' 3'-C G C C G G▲C G-5' (NEB, 2026b)
BamHI-HF	NEB	Restriction endonuclease that cuts sequence: 5'-G▼G A T C C-3'

		3'-C C T A G _▲ G-5' (NEB, 2026a)
Q5 high-fidelity DNA polymerase	NEB	Thermostable DNA polymerase that synthesises DNA in a 5'-3' direction with high-fidelity using unbound dNTPs, beginning from a template-bound primer. Also has 3'-5' exonuclease activity (Cooper, 2000; NEB, 2025j).
Quick CIP	NEB	Heat-labile calf intestinal alkaline phosphatase that 'nonspecifically catalyses the dephosphorylation of 5' and 3' ends of DNA and RNA phosphomonoesters' and hydrolyses dNTPs (NEB, 2025k).
T4 PNK	NEB	Polynucleotide kinase that phosphorylates 5' ends of DNA or RNA and dephosphorylates 3' ends (NEB, 2025m).
T4 DNA ligase	NEB	DNA ligase that catalyses the formation of a phosphodiester bond between complementary base pairs in dsDNA or dsRNA, being able to joins both cohesive and blunt ends. Also able to 'repair single stranded nicks in duplex DNA and some DNA/RNA hybrids' (NEB, 2025l).
RNAase A	Invitrogen	Nonprocessive bovine pancreatic endoribonuclease that cuts ssRNA (delCardayré et al., 1994), used as part of the Invitrogen PureLink Genomic DNA Mini Kit (Thermo Fisher, 2007).
Proteinase K	Invitrogen	Fungal nonspecific serine endoprotease from <i>Engyodontium album</i> (Kushnirov et al., 2020; Thermo Fisher, 2025c) that degrades proteins by hydrolysis (Ebeling et al., 1974), used as part of the Invitrogen PureLink Genomic DNA Mini Kit (Thermo Fisher, 2007).

2.2 Laboratory Recipes

2.2.1 Bacterial Growth Media Compositions

Table 2.6 List of recipes of bacterial growth media used in this work.

Growth Medium	Composition
Mu broth	10 g/L tryptone 10 g/L NaCl 5 g/L yeast extract 2 mM NaOH
Mu agar	200 mL Mu broth 3 g agar powder
2X M9 medium base salts	24.14 g/L $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$ 6 g/L KH_2PO_4 1 g/L NaCl 0.6 g/L MgSO_4 0.03 g/L $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$
Trace element solution	0.08 g/L ZnCl_2 0.4 g/L $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ 0.02 g/L $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ 0.02 g/L $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 0.02 g/L $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$ 0.02 g/L $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$
2X M9 minimal medium	50 mL 2X M9 medium base salts 0.01 g biotin 0.01 g thiamine hydrochloride 100 μL trace element solution 100 μL 100X MEM Vitamin Solution (Sigma-Aldrich)

	0.123 g (NH ₄)SO ₄ 1.86 mL 1 M NaOH 200 µL 60% (v/v) glycerol
Supplemented 2X M9 medium	50 mL 2X M9 minimal medium 0.5 g casamino acids
1X M9 medium base salts	12.07 g/L Na ₂ HPO ₄ ·12H ₂ O 3 g/L KH ₂ PO ₄ 0.5 g/L NaCl 0.3 g/L MgSO ₄ 0.015 g/L CaCl ₂ ·2H ₂ O
1X M9 minimal medium	50 mL 1X M9 medium base salts 0.01 g biotin 0.01 g thiamine hydrochloride 100 µL trace element solution 100 µL 100X MEM Vitamin Solution (Sigma-Aldrich) 0.123 g (NH ₄)SO ₄ 1.86 mL 1 M NaOH 200 µL 60% (v/v) glycerol

2.2.2 Gel Solutions Compositions

Table 2.7 List of recipes of gels used in this work.

Gel Type	Percentage	Composition
Agarose gel	1% (w/v)	1 g agarose powder 100 mL 1X TBE
	2% (w/v)	2 g agarose powder 100 mL 1X TBE
SDS separating gel	10% (w/v)	5.2 mL SDW

		3.3 mL 30% (w/v) acrylamide 1.3 mL 3 M Tris-HCl pH 8.8 100 µL 10% (w/v) SDS 100 µL 10% (w/v) APS 10 µL TEMED
SDS stacking gel	5% (w/v)	2.25 mL SDW 0.7 mL 30% (w/v) acrylamide 1.25 mL 0.5 M Tris-HCl pH 6.8 50 µL 10% (w/v) SDS 50 µL 10% (w/v) APS 10 µL TEMED

2.2.3 General Buffers Compositions

Table 2.8 List of recipes of buffers used in this work.

Buffer	Composition
10X TBE	1 M Tris 1 M boric acid 20 mM EDTA pH 8.0
4X SDS-PAGE loading dye	200 mM Tris-HCl pH 8 8% (w/v) SDS 0.4% (w/v) bromophenol blue 40% (v/v) glycerol
10X SDS-PAGE running buffer	250 mM Tris 1.92 M glycine 1% (w/v) SDS
10X annealing buffer	100 mM Tris-HCl pH 7.5 500 mM NaCl

2.2.4 General Reagents Compositions

Table 2.9 List of recipes of reagents used in this work.

Buffer	Composition
1X Coomassie blue stain solution	0.05% (w/v) Brilliant Blue R-250 10% (v/v) acetic acid 40 % (v/v) methanol
1X Coomassie destain solution	10% (v/v) acetic acid 20% (v/v) methanol

2.3 General Microbiology Protocols

2.3.1 Production of Competent *E. coli* Cells

Protocol was performed under sterile conditions. Glycerol stocks of competent DH5 α or BL21-AI *E. coli* were streaked onto Mu agar plates without antibiotics and incubated in a PIF120(200) Incubator (Carbolite) overnight at 37 °C (see Section 2.3.3). A single colony was picked and inoculated into 50 mL autoclave sterilised Mu broth in a baffled conical flask as an overnight starter culture, which was incubated in a Lab-Therm LT-W Shaking Incubator (Kühner Shaker) at 37 °C with 150 RPM shaking for 16-18 hours. Overnight starter culture was inoculated into 0.99 L Mu broth in a baffled conical flask for a 1:100 dilution, which was incubated in the Lab-Therm LT-W Shaking Incubator at 37 °C with 150 RPM shaking until OD₆₀₀ 0.5, measured with 1 mL culture samples in cuvettes on a DS-11+ Spectrophotometer (DeNovix). Cells were then pelleted in a Centrifuge 5430 R (Eppendorf) at 4 °C at 4000 RPM for 5 minutes, resuspended in 10 mL ice-cold 0.22 μ m filter-sterilised 0.1 M CaCl₂, and incubated on ice for 1 hour. Cells were pelleted again as before and resuspended in 35 mL of ice-cold 0.22 μ m filter-sterilised 0.1 M CaCl₂ and 15 mL autoclave-sterilised 80% (v/v) glycerol. 100-300 μ L cell aliquots of were flash-frozen on dry ice and stored at -80 °C.

2.3.2 Agar Plate Preparation

Autoclave-sterilised Mu agar was microwaved in an R-959(SL)M-AA Microwave Oven (Sharp) until fully melted, then incubated in a Drying Cabinet SS (LEEC) at 37 °C to maintain a liquid state if it was not used immediately. Once cooled, under sterile conditions, 25 mL liquid agar per plate to be prepared was transferred to a conical flask, any relevant antibiotic(s) were added for the correct working concentration(s), and the antibiotic(s) mixed in by shaking. Under sterile conditions, 25 mL liquid agar was poured evenly into a Petri dish for each plate and incubated at room temperature until solidified. Plates were then incubated in the Drying Cabinet SS at 37 °C until dried. Agar plates were stored at 4 °C if not used immediately.

2.3.3 Transformation of Competent *E. coli* With Purified Plasmid DNA

Protocol was performed under sterile conditions. 100 µL competent DH5α or BL21-AI *E. coli* were added to 1 µL miniprep-purified plasmid DNA, gently mixed by pipetting, and incubated on ice for 30 minutes. Mixture was heat shocked by incubating in a SUB14 Unstirred Water Bath (Grant Instruments) at 42 °C for 90 seconds then incubated on ice for 5 minutes. 900 µL autoclave-sterilised Mu broth was added to the mixture. Mixture was incubated in an Isotemp SWB 15 Shaking Water Bath (Fisher Scientific) at 37 °C with 150 RPM shaking for 1 hour. Cells were pelleted in a Heraeus Pico 17 Centrifuge (Thermo Scientific) at 17,000 *g* for 1 minute, 900 uL supernatant was removed, and the pellet resuspended in the remaining 100 uL Mu broth. Mixture was then spread evenly on an agar plate containing any relevant antibiotic(s), which was incubated in a PIF120(200) Incubator (Carbolite) t at 37 °C for 16-18 hours overnight and then stored at 4 °C.

2.3.4 Growing Overnight Cultures of Transformed *E. coli*

Under sterile conditions, a single colony was picked from an agar plate of transformed DH5α or BL21-AI *E. coli* and inoculated into 5 mL autoclave-sterilised Mu broth containing any relevant antibiotic(s) in a culture tube. Culture was incubated in a Rollordrum TC-7 (New Brunswick Scientific) tube rotator overnight at 37 °C at 25 RPM for 16-18 hours.

2.3.5 Producing Glycerol Stocks of Transformed *E. coli*

Under sterile conditions, 800 µL transformed BL21-AI *E. coli* overnight culture was added to 800 µL autoclave-sterilised 80% (v/v) glycerol in a cryovial and mixed via pipetting. Bacterial glycerol stock was then flash-frozen on dry ice and stored at -80 °C.

2.4 General Protocols Concerning DNA

2.4.1 Miniprep Extraction and Purification of Plasmid DNA From Transformed DH5α *E. coli* Overnight Cultures

Plasmid DNA was extracted and purified from transformed DH5α *E. coli* overnight cultures to make plasmid stocks using the Wizard Plus SV Minipreps DNA Purification System (Promega) via the centrifugation protocol using a Heraeus Pico 17 Centrifuge (Thermo Scientific), according to the manufacturer's instructions (https://www.promega.com/-/media/files/resources/protcards/wizard-plus-sv-minipreps-dna-purification-system-quick-protocol.pdf?rev=8293c52acfca4de495f1df90a7753d50&sc_lang=en) (Promega, 2009). Purified plasmid DNA concentration was measured in triplicate on a DS-11+ Spectrophotometer (DeNovix) and averaged. Purified plasmid DNA was stored at -20 °C.

2.4.2 Analytical Restriction Digestion of Purified Plasmid DNA

Analytical restriction digestion was carried out to verify the length of purified plasmid DNA. 20 µL reaction mixture was assembled containing 2 µL 10X rCutSmart buffer (NEB), 100 ng purified plasmid DNA, 0.5 µL of each restriction endonuclease, and SDW. Reaction mixture was incubated in a 1002 Water Bath (GFL) at 37 °C for 1 hour before being incubated on ice. Reaction mixture was then electrophoresed on an agarose gel (see Section 2.4.3).

2.4.3 Agarose Gel Electrophoresis of DNA Samples

Agarose gel electrophoresis was carried out to visualise the length of unlabelled DNA strands, such as analytical restriction digested plasmids or PCR products. Agarose powder for an

appropriate percentage (w/v) agarose gel was dissolved in 1X TBE by microwaving in an R-959(SL)M-AA Microwave Oven (Sharp). Once cooled, 0.2 µg/mL ethidium bromide was added, and the solution was mixed by shaking, poured evenly, and left to solidify. Gel was submerged in 1X TBE in a multiSUB Midi Horizontal Gel System (Fisher Scientific). 6X Purple Gel Loading Dye (NEB) was added to a DNA sample(s) and 14 µL of each sample mixture were loaded onto the set gel alongside 5 µL 1 kb DNA Ladder (NEB) or Quick-Load 1 kb Plus DNA Ladder (NEB), unless otherwise stated. Unless otherwise stated, gel was electrophoresed at 120 V for 1 hour, using a PowerPac Basic Supply System (Bio-Rad) or PowerEase 90 W Power Supply (Invitrogen). Gel was visualised using UV light and imaged in a U:Genius³ (Syngene) gel documentation system, unless gel extraction of DNA was required (see Section 2.4.4). Image was saved as a JPEG.

2.4.4 Extraction of DNA From Electrophoresed Agarose Gels

Gel extraction of DNA was carried out if an electrophoresed agarose gel could not be visualised under UV light. Electrophoresed agarose gel was visualised on a Dual LED Blue/White Light Transilluminator (Invitrogen) in a dark room and a gel fragment containing the desired DNA strand was extracted using a scalpel. Gel fragment was stored at -20 °C until DNA purification. DNA purification was carried out using the Wizard SV Gel and PCR Clean-Up System (Promega) via the centrifugation protocol using an AccuBlock Digital Dry Bath D1302 (Labnet International), Vortex-Genie 2 (Scientific Industries), and Heraeus Pico 17 Centrifuge (Thermo Scientific), according to the manufacturer's instructions (<https://www.promega.com/-/media/files/resources/protcards/wizard-sv-gel-and-pcr-clean-up-system-quick-protocol.pdf>) (Promega, 2024). Purified DNA concentration was measured on a DS-11+ Spectrophotometer (DeNovix). Purified DNA was stored at -20 °C.

2.4.5 Polymerase Chain Reaction Amplification From Purified Plasmid DNA

PCR amplification was carried out to amplify the DNA of a region of interest in a purified plasmid. Forward and reverse primers were designed to amplify the region of interest. Unless otherwise stated, 25 µL reaction mixture was assembled containing 5 µL 5X Q5 High GC

Enhancer (NEB), 5 μ L 5X Q5 reaction buffer (NEB), 50 ng template purified plasmid DNA, 1.25 μ L 10 μ M forward primer, 1.25 μ L 10 μ M reverse primer, 0.5 μ L 10 mM dNTPs, 0.25 μ L Q5 High-Fidelity DNA Polymerase (NEB), and SDW. Optimum primer annealing temperature was calculated using the online tool NEB Tm Calculator (<https://tmcalculator.neb.com/#!/main>). Reaction mixture was incubated in a Veriti 96-Well Thermal Cycler (Applied Biosystems) with the following thermocycling conditions:

Table 2.10 PCR amplification from purified plasmid DNA thermocycling conditions.

Cycle Step	Temperature (°C)	Time (seconds)	Number of Cycles
Initial denaturing	95	30	1
Denaturing	95	10	30
Annealing	Variable (dependent on primer Tm)	30	
Extension	72	30/product kb	
Final extension	72	120	1
Hold	16	Indefinite	1

PCR product DNA size was verified by electrophoresing 12 μ L sample containing 7 μ L SDW, 3 μ L PCR sample, and 2 μ L 6X Purple Gel Loading Dye (NEB) on a 1% (w/v) agarose gel alongside 1 kb DNA Ladder (NEB) (see Section 2.4.3), unless otherwise stated. Reaction mixture was then stored at -20 °C until PCR product DNA purification (see Section 2.4.6).

2.4.6 Purification of Product DNA From PCR Reaction Mixture

PCR product DNA purification was carried out using the Wizard SV Gel and PCR Clean-Up System (Promega) via the centrifugation protocol using a Heraeus Pico 17 Centrifuge (Thermo Scientific) (see Section 2.4.4). Purified DNA concentration was measured on a DS-11+ Spectrophotometer (DeNovix). Purified DNA was stored at -20 °C.

2.4.7 Site-Directed Mutagenesis By PCR Amplification From Purified Plasmid DNA

PCR amplification was carried out to amplify a mutagenised version of a purified plasmid in site-directed mutagenesis (SDM) (see Section 2.4.5). Forward and reverse primers were designed to amplify the mutagenic version using the online tool NEBase Changer (<https://nebasechanger.neb.com/>). Mutagenised PCR product DNA size was verified by electrophoresing 12 µL sample containing 7 µL SDW, 3 µL PCR sample, and 2 µL 6X Purple Gel Loading Dye (NEB) on a 1% (w/v) agarose gel alongside 1 kb DNA Ladder (NEB) (see Section 2.4.3). 10 µL KLD reaction was assembled containing 5 µL SDW, 1 µL PCR sample, 1 µL 10X T4 Ligase Reaction Buffer (NEB), 1 µL T4 PNK, 1 µL T4 DNA Ligase, and 1 µL DpnI. Reaction mixture was incubated at room temperature for 1 hour before transformation into 100 µL competent DH5α *E. coli* (see Section 2.3.3).

2.4.8 Restriction Cloning To Generate Recombinant Plasmids

Restriction cloning was performed to generate recombinant plasmids. Purified plasmid DNA and purified PCR product DNA were separately restriction digested in 20 µL reaction mixtures containing 2 µL 10X rCutSmart buffer (NEB), 1 µL of each restriction endonuclease, purified plasmid DNA or purified PCR product DNA, and SDW (see Section 2.4.2). Unless stated otherwise, reaction mixture was incubated at 37 °C for 1 hour. If both restriction endonucleases could be heat inactivated, reaction mixture was subsequently incubated in an AccuBlock Digital Dry Bath D1302 (Labnet International) at 80 °C for 20 minutes. If a reaction mixture contained a digested plasmid vector, the plasmid DNA was dephosphorylated of 5'-3' phosphates by addition of 0.8 µL SDW, 0.2 µL 10X rCutSmart buffer (NEB), and 1 µL Quick CIP followed by incubation in a 1002 Water Bath (GFL) at 37 °C for 10 minutes. Quick CIP was then heat inactivated by incubating the reaction mixture in the AccuBlock Digital Dry Bath D1302 at 80 °C for 2 minutes. Digested DNA was purified using the Wizard SV Gel and PCR Clean-Up System (Promega) via the centrifugation protocol using a Heraeus Pico 17 Centrifuge (Thermo Scientific) (see Section 2.4.4). Purified digested DNA concentration was measured on a DS-11+ Spectrophotometer (DeNovix). Purified digested DNA was stored at -20 °C.

Ligation reaction mixture composition was calculated using the online tool NEBioCalculator (<https://nebiocalculator.neb.com/#!/ligation>). 20 μ L reaction mixture was assembled containing 2 μ L 10X T4 DNA Ligase Reaction Buffer (NEB), 1 μ L T4 DNA Ligase (NEB), purified digested plasmid vector DNA, purified digested PCR product insert DNA, and SDW using the appropriate mass of insert DNA for a given mass of vector DNA and vector:insert ratio. Negative control reaction mixture lacking insert DNA was also assembled. Reaction mixtures were incubated at room temperature for 1 hour then each transformed into 100 μ L DH5 α *E. coli*, which were plated onto Mu agar plates containing appropriate antibiotic(s) (see Section 2.3.3). 5 mL overnight cultures were grown (see Section 2.2.4) and miniprepmed (see Section 2.4.1). Purified plasmid DNA was analytical restriction digested (see Section 2.4.2). Reaction mixtures were electrophoresed on an agarose gel (see Section 2.4.3). Successful recombinant plasmids were verified through whole plasmid sequencing, carried out by internally by Deep Seq (University of Nottingham), or Sanger sequencing of both the vector insertion region and of the insert region using the region's respective forward and reverse primers, carried out externally by GENEWIZ (Azenta Life Sciences).

2.5 General Protein Assay Protocols

2.5.1 Pilot Protein Overexpression

Pilot protein overexpression was performed to verify proteins of interest could be successfully overexpressed in BL21-AI *E. coli*. Protocol was performed under sterile conditions. Purified plasmid DNA with a region encoding a protein of interest, or a control plasmid, was transformed into 100 μ L competent BL21-AI *E. coli*, which were plated onto Mu agar plates containing 100 μ g/mL ampicillin an (see Section 2.3.3). 5 mL overnight cultures for each of the transformations were grown (see Section 2.3.4).

Mu broth was prewarmed in a PIF120(200) Incubator (Carbolite) overnight at 37 °C for 16-18 hours. Ampicillin was then added for 100 μ g/mL working concentration. 500 μ L of each overnight culture was pelleted in a Heraeus Pico 17 Centrifuge (Thermo Scientific) at 13,200 *g* for 1 minute and the pellet resuspended in 500 μ L prewarmed Mu broth containing 100

µg/mL ampicillin. For each transformation, resuspended cells were inoculated in duplicate into 4.875 mL prewarmed Mu broth containing 100 µg/mL ampicillin for a 1:40 dilution. 5 mL cultures were then incubated in an OLS Aqua Pro Shaking Water Bath (Grant Instruments) at 37 °C with 150 RPM shaking. Culture OD₆₀₀ was measured using a Spectronic 20+ (Thermo Electron Corporation). Once a culture reached OD₆₀₀ 0.5 ± 0.05 and was to overexpress protein, this was induced using 20% (w/v) L-arabinose and 1 M IPTG at respective 0.2% (w/v) and 0.5 mM working concentrations. Only one culture in each duplicate was induced for protein overexpression, the other acting as an uninduced negative control. After 3 hours of incubation from protein overexpression inducement, 1 mL culture was converted into an SDS-PAGE sample. 8 µL of each SDS-PAGE sample was electrophoresed alongside 4 µL Color Prestained Protein Standard, Broad Range (10-250 kDa) (NEB) in a 10% (w/v) SDS-PAGE gel at 120 V for 1 hour and 10 minutes (see Section 2.5.2).

2.5.2 SDS-PAGE Analysis of Protein Overexpression Samples

SDS-PAGE analysis was used to visualise the molecular weight of proteins following overexpression and verify successful overexpression. A given volume of protein-overexpressing transformed BL21-AI *E. coli* cells were pelleted in a Heraeus Pico 17 Centrifuge (Thermo Scientific) at 13,200 *g* for 1 minute and the pellet resuspended in 150 µL SDW. 50 µL 4X SDS loading dye was added, then 1 M DTT was added for a 5 mM working concentration. Resulting SDS-PAGE sample was incubated in an Analogue Dri-Block DB-2A (Techne) at 95 °C for 10 minutes. Sample was stored at -20 °C.

10% (w/v) SDS-PAGE gel was prepared using a Mini-PROTEAN II Electrophoresis Cell (Bio-Rad). Gel was comprised of a lower 10% (w/v) SDS separating gel and upper 5% (w/v) SDS stacking gel. Gel was submerged in 1X SDS-PAGE running buffer, and 8 µL of each sample was loaded into the gel alongside 4 µL Color Prestained Protein Standard, Broad Range (10-250 kDa) (NEB). Unless otherwise stated, gel was electrophoresed at 120 V for 1 hour, using a PowerPac Basic Supply System (Bio-Rad), PowerPac 300 (Bio-Rad), or Mini 300 V Plus Power Supply (Fisher Scientific). Gel was then stained in 1X Coomassie blue stain solution with rocking on a

Platform Rocker STR6 (Stuart Scientific) or Reax 3 (Heidolph) rocker for 1 hour, washed with water, and destained in 1X Coomassie destain solution with rocking on the Platform Rocker STR6 or Reax 3 for at least 18 hours. Gel was then illuminated using an A4 LED Light Box LC2004LED (Lightcraft) and imaged. Gel was stored submerged in dH₂O at room temperature.

2.6 Bimolecular Fluorescence Complementation Assays

2.6.1 Endpoint BiFC Assay Using Cotransformed BL21-AI *E. coli*

Endpoint BiFC assay was used to investigate the level of interaction between a given protein pair. Protocol was performed under sterile conditions. Pairs of purified plasmid DNA, each with a region encoding a control or test protein, including pBAD-NmVenus and pRSF-CmVenus or pBAD-NmVenus-PoD3 and pRSF-CmVenus-NH₂Q, were each cotransformed into 100 µL competent BL21-AI *E. coli*, which were plated onto Mu agar plates containing 100 µg/mL ampicillin and 40 µg/mL kanamycin (see Section 2.3.3). Cotransformations used 70 ng purified plasmid DNA for each of the plasmids in the pair. 5 mL overnight cultures for each of the competent BL21-AI *E. coli* cotransformations were grown (see Section 2.3.4), and if necessary bacterial glycerol stocks were produced (see Section 2.3.5).

Mu broth was prewarmed in a PIF120(200) Incubator (Carbolite) overnight at 37 °C for 16-18 hours. Ampicillin and kanamycin were then added for respective 100 µg/mL and 40 µg/mL working concentrations. For each cotransformation, the respective overnight culture was inoculated into 1.95 mL prewarmed Mu broth containing 100 µg/mL ampicillin and 40 µg/mL kanamycin in a clear flat-bottom non-tissue-culture treated (TC) 24-well plate for a 1:40 dilution in technical triplicate. Plate with lid was incubated in a FLUOstar Omega (BMG LABTECH) plate reader at 37 °C with 300 RPM double orbital shaking. Sample OD₆₀₀ was measured using the top optic every 10 minutes with 2 mL prewarmed Mu broth containing 100 µg/mL ampicillin and 40 µg/mL kanamycin as a blank. Once a sample reached OD₆₀₀ 0.5 ± 0.05, protein overexpression was induced using 20% (w/v) L-arabinose and 1 M IPTG at respective 0.2% (w/v) and 0.5 mM working concentrations. After 3 hours and 30 minutes of

incubation from protein overexpression inducement, sample was pelleted in a Heraeus Pico 17 Centrifuge (Thermo Scientific) at 1,000 *g* for 10 minutes, and the pellet gently resuspended in 1 mL 0.22 μm filter-sterilised 1X M9 minimal medium containing 100 $\mu\text{g}/\text{mL}$ ampicillin and 40 $\mu\text{g}/\text{mL}$ kanamycin. Resuspended samples were stored at 4 °C.

The next day, each cell sample was gently mixed by pipetting to resuspend the pellet. 400 μL of each sample was pelleted in the Heraeus Pico 17 Centrifuge at 12,000 *g* for 1 minute, 250 μL supernatant removed, and the pellet resuspended in the remaining 150 μL supernatant. Remaining 600 μL samples were stored at 4 °C. Resuspended 150 μL samples were transferred to a black flat-bottom non-TC 96-well plate in addition to a triplicate of 150 μL 1X M9 minimal medium containing 100 $\mu\text{g}/\text{mL}$ ampicillin and 40 $\mu\text{g}/\text{mL}$ kanamycin. Endpoint sample fluorescence intensity was measured using the FLUOstar Omega plate reader's top optic with excitation/emission wavelength filters 500/520 nm and using one of the 150 μL 1X M9 minimal medium triplicate as a blank. Gain adjustment was set by using 95% of the highest fluorescence intensity. Resuspended 150 μL samples and the 150 μL 1X M9 minimal medium triplicate were transferred to a flat-bottom clear non-TC 96-well plate. Endpoint sample OD₆₀₀ was measured using the FLUOstar Omega plate reader's top optic using one of the 150 μL 1X M9 minimal medium triplicate as a blank. Plate was then scanned using the built-in Cy2 filter, and imaged, in an Amersham Typhoon 5 Biomolecular Imager (Cytiva) with the plate in a Titre plate holder on a multi-stage. Images were saved as a greyscale TIFF and heatmap JPEG. Heatmap gradient was saved as a screenshot.

To verify control and test protein overexpression, within 48 hours either one or all sample(s) from each of the remaining 600 μL sample triplicates was gently mixed by pipetting to resuspend the pellet, then 500 μL of each of those resuspended samples was converted into an SDS-PAGE sample. 8 μL of each SDS-PAGE sample was electrophoresed alongside 4 μL Color Prestained Protein Standard, Broad Range (10-250 kDa) (NEB) in a 10% (w/v) SDS-PAGE gel at a given voltage and time length (see Section 2.5.2).

A given number of days after samples were first resuspended in 1 mL 1X M9 medium, a spot assay was performed to assess sample cell viability. Each sample was gently mixed by pipetting to resuspend the pellet, then serial dilutions were performed for each sample. 10 μ L sample was then diluted in 90 μ L autoclave-sterilised 1X M9 medium base salts as a 10^{-1} dilution, and each subsequent dilution diluted 10 μ L of the previous in 90 μ L 1X M9 medium base salts. 10 μ L of each serial dilution was spotted in triplicate onto Mu agar plates containing 100 μ g/mL ampicillin and 40 μ g/mL kanamycin at up to 25 spots per plate. Plates were air-dried for 15 minutes, incubated in the PIF120(200) Incubator overnight at 37 °C for 18 hours, and then stored at 4 °C. Colonies were counted using an SC6+ Colony Counter (Stuart) and averages were calculated and rounded to the nearest whole number. Spots with too much runoff or merging were not included in averages. Plates were illuminated using an A4 LED Light Box LC2004LED (Lightcraft) and imaged. Sample endpoint fluorescence intensity was separately normalised using endpoint OD₆₀₀ values and colony forming units (CFU)/mL, where colony count averages used to calculate CFU/mL were based on highest dilution factor among a sample's triplicates that had at least 2 valid spots, otherwise the highest dilution factor with at least one valid spot was used. Data was then plotted and analysed using two-tailed Welch's t-tests using the software GraphPad Prism 10 (<https://www.graphpad.com/>). Graphs were saved as JPEGs.

2.6.2 Test For BL21-AI *E. coli* Viable But Non-Culturable State

Test for viable but non-culturable state was carried out 13 days after POLD3 Quincunx motif mutants endpoint BiFC assay to determine if BL21-AI *E. coli* cell samples were viable and could be resuscitated. Protocol was performed under sterile conditions. 100 μ L competent BL21-AI *E. coli* were cotransformed with pBAD-NmVenus and pRSF-CmVenus or pBAD-NmVenus-POLD3 and pRSF-CmVenus-NHclQ, then plated onto Mu agar plates containing 100 μ g/mL ampicillin and 40 μ g/mL kanamycin (see Section 2.3.3). Cotransformations used 70 ng purified plasmid DNA for each of the plasmids in the pair. 5 mL overnight cultures for each of the competent BL21-AI *E. coli* cotransformations were grown (see Section 2.3.4).

Mu broth was prewarmed in a PIF120(200) Incubator (Carbolite) overnight at 37 °C for 16-18 hours. Ampicillin and kanamycin were then added for respective 100 µg/mL and 40 µg/mL working concentrations. 30.4 mL of Mu broth was removed, and 1.6 mL Tween 80 (Koch-Light Laboratories) was added to the removed Mu broth for 5% (v/v) concentration. This was then mixed using a Vortex-Genie 2 (Scientific Industries) and 0.22 µm filter-sterilised. Wells in a clear flat-bottom tissue-culture treated 48-well plate were filled with 990 µL or 1 mL Mu broth containing 100 µg/mL ampicillin and 40 µg/mL kanamycin with or without 5% (v/v) Tween 80, for 24 wells each. For each medium type, 10 µL pBAD-NmVenus + pRSF-CmVenus and pBAD-NmVenus-POLD3 + pRSF-CmVenus-NHelQ cotransformation overnight culture was inoculated in 990 µL medium for 1:100 dilutions in technical duplicate, and cell samples of BL21-AI *E. coli* expressing NmVenus + CmVenus, NmVenus-POLD3 + CmVenus-N-HELQ, and NmVenus-POLD3^{PNLS} stored in 1X M9 medium at 4 °C were gently mixed by pipetting and inoculated in technical duplicate. Plate with lid was incubated in a FLUOstar Omega (BMG LABTECH) plate reader at 37 °C with 200 RPM double orbital shaking for 48 hours. Sample OD₆₀₀ was measured using the top optic every 10 minutes with 1 mL respective medium as a blank. 1 mL samples were then pelleted in a Heraeus Pico 17 Centrifuge (Thermo Scientific) at 1,000 *g* for 10 minutes, and the pellet gently resuspended in 1 mL 0.22 µm filter-sterilised 1X M9 minimal medium containing 100 µg/mL ampicillin and 40 µg/mL kanamycin.

A spot assay was performed to assess sample cell viability. For each sample, 10 µL was diluted in 90 µL autoclave-sterilised 1X M9 medium base salts as a 10⁻¹ dilution, and each subsequent dilution diluted 10 µL of the previous in 90 µL 1X M9 medium base salts up to 10⁻¹⁰. 10 µL of each serial dilution was spotted in triplicate onto Mu agar plates containing 100 µg/mL ampicillin and 40 µg/mL kanamycin at up to 25 spots per plate. Plates were air-dried for 15 minutes, incubated in the PIF120(200) Incubator overnight at 37 °C for 18 hours, and then stored at 4 °C. Colonies were counted using an SC6+ Colony Counter (Stuart), and averages were calculated and rounded to the nearest whole number. Spots with too much runoff or merging were not included in averages. Colony count averages used to calculate CFU were based on highest dilution factor among a sample's triplicates that had at least 2 valid spots, otherwise the highest dilution factor with at least one valid spot was used. Data was then

plotted using the software GraphPad Prism 10 (<https://www.graphpad.com/>). Plates were illuminated using an A4 LED Light Box LC2004LED (Lightcraft) and imaged.

2.6.3 Test of BL21-AI *E. coli* Growth Media Autofluorescence and Effect on Growth Rate

Growth media test was performed to determine the time it took for BL21-AI *E. coli* cotransformed with pairs of BiFC assay plasmids to reach OD₆₀₀ 0.5. Protocol was performed under sterile conditions. 100 µL competent BL21-AI *E. coli* were cotransformed with pBAD-NmVenus and pRSF-CmVenus or pBAD-NmVenus-PoID3 and pRSF-CmVenus-NHclQ, then plated onto Mu agar plates containing 100 µg/mL ampicillin and 40 µg/mL kanamycin (see Section 2.3.3). Cotransformations used 70 ng purified plasmid DNA for each of the plasmids in the pair. 5 mL overnight cultures for each of the competent BL21-AI *E. coli* cotransformations were grown (see Section 2.3.4).

0.22 µm filter-sterilised 2X M9 minimal medium, and 2X M9 medium supplemented with 1% casamino acids (MP Biomedicals), containing 100 µg/mL ampicillin and 40 µg/mL kanamycin was prepared. Wells in a clear flat-bottom non-tissue culture treated (TC) 24-well plate were filled with 1.95 or 2 mL 2X M9 minimal medium, supplemented 2X M9 medium, or Mu broth containing 100 µg/mL ampicillin and 40 µg/mL kanamycin. Plate was scanned using the built-in Cy2, Cy5, and IRshort filters, and imaged, in an Amersham Typhoon 5 Biomolecular Imager (Cytiva) with the plate in a Titre plate holder on a multi-stage. Images were saved as a greyscale TIFFs and heatmap JPEGs. Heatmap gradient was saved as a screenshot. Media autofluorescence intensity was measured with a FLUOstar Omega (BMG LABTECH) plate reader at 37 °C, using the plate reader's top optic and emission/excitation filters of 500/520 nm and 590-50/670 nm, with gain set at 95% of highest fluorescence value for each filter pair (1473 and 3025, respectively). Media was then prewarmed by incubating the plate in the FLUOstar Omega plate reader at 37 °C.

Three 1 mL aliquots from each of the two overnight cultures were pelleted in a Heraeus Pico 17 Centrifuge (Thermo Scientific) at 1,000 *g* for 10 minutes, the supernatant discarded, and for each set of three aliquots, pellets were gently resuspended in 1 mL 2X M9 minimal medium, supplemented 2X M9 medium, or Mu broth containing 100 µg/mL ampicillin and 40 µg/mL kanamycin. Resuspended samples were incubated in an Isotemp SWB 15 Shaking Water Bath (Fisher Scientific) at 37 °C with 150 RPM shaking for 30 minutes to allow cells to adjust to the new medium. Samples were then inverted to resuspend any pellets. Samples were inoculated into 1.95 mL respective prewarmed medium in the 24-well plate for 1:40 dilutions in technical triplicate. Plate with lid was incubated in the FLUOstar Omega at 37 °C with double 300 RPM orbital shaking. Sample OD₆₀₀ was measured every 10 minutes for 16 hours overnight using the plate reader's top optic and 2 mL respective medium as a blank. Average OD₆₀₀ values were calculated using triplicate averages, and results were plotted using the software GraphPad Prism 10 (<https://www.graphpad.com>). Logarithmic phase growth rate was calculated using the formula:

$$\mu = \frac{\Delta \ln OD_{600}}{\Delta t}, \text{ where } t \text{ is time.}$$

Fluorescence data was analysed using two-tailed Welch's t-tests with two-stage step-up method with Q of 1% in GraphPad Prism 10 (Benjamini et al., 2006). Graphs were saved as JPEGs.

2.7 Human Cell Culturing

2.7.1 Human Cell Line

Human cell line used in this work was wild type (WT) U2OS, sourced from the lab of Claire Friel (University of Nottingham). Originally designated as the T2 cell line (Srimani et al., 2022), the U2OS cell line is a hypertriploid (Bayani et al., 2003) osteosarcoma cell line derived from a 'moderately differentiated osteogenic sarcoma of the tibia' of a 15-year-old female in 1964 (Pontén et al., 1967).

2.7.2 Human Cells Growth Medium and Supplements

Complete medium used in this work consisted of a stock serum- and glutamine-free DMEM (Gibco) supplemented with the following:

Table 2.11 List of DMEM supplements used in complete medium in this work.

Supplement	Stock Concentration	Working Concentration
Foetal bovine serum (FBS) (Sigma-Aldrich)	100%	10%
L-glutamine (Gibco)	200 mM	2 mM
Penicillin-streptomycin (Sigma-Aldrich)	10,000 units/mL penicillin, 10 mg/mL streptomycin	100 units/mL penicillin, 0.1 mg/mL streptomycin

Stock DMEM was stored at 4 °C. Supplements were aliquoted and stored at -20 °C. Stock DMEM was supplemented under sterile conditions, aliquoted into 50 mL, and stored at 4 °C.

2.7.3 Routine Growth and Passaging of Human Cells

Cells were cultured in complete medium in T75 flasks and incubated in an MCO-170AC CO₂ Incubator (PHCbi) at 37 °C and 5% CO₂. Cells were passaged for initial seeding and routinely upon reaching 80% confluency. Protocol was performed under sterile conditions. Medium was aspirated from cells and cells were washed twice in 1X phosphate-buffered saline (PBS) (Gibco). Cells were incubated within 2.5 mL 0.05% (w/v) trypsin-EDTA (Sigma-Aldrich) for 5 minutes, then thoroughly shaken. Trypsinisation was confirmed by viewing cells under an Invertoskop ID 03 (ZEISS) microscope. Trypsin activity was quenched by adding 7.5 mL medium. Cells were pelleted in a Heraeus Biofuge Primo Centrifuge (Thermo Scientific) at 350 *g* for 10 minute and then resuspended in 1 mL medium. Resuspended cells were transferred into a new T75 flask at a 1:12 split ratio and medium added to total 14 mL. 50% of cell medium was changed upon cells reaching 40% confluency.

2.8 PE3b Prime Editing of Human Cells

2.8.1 Prime Editing Oligonucleotides

To design oligonucleotides for the construction of human HELQ^{D142A/F143A} prime editing plasmids, WT human *HELQ* DNA sequence was sourced from the NCBI human *HELQ* CCDS (<https://www.ncbi.nlm.nih.gov/projects/CCDS/CcidsBrowse.cgi?REQUEST=ALLFIELDS&DATA=CCDS3603.1&ORGANISM=0&BUILDS=CURRENTBUILDS>) (NCBI, 2025a), and regions surrounding the sequence encoding HELQ^{D142/F143} were identified. This identified sequence and the sequence with desired prime editing modifications were used in the online tool pegFinder (<http://pegfinder.sidichenlab.org/>) (Chow et al., 2021). All prime editing oligonucleotides used in this work were ordered from and synthesised by Sigma-Aldrich at 0.02 µM scale, HPLC purified, delivered lyophilised dry, and resuspended in SDW to 100 µM stock concentration for storage at -20 °C.

Table 2.12 List of prime editing oligonucleotides used in this work.

Oligonucleotide Name	5'-3' Sequence	Use
HQPM_D142 F143_sgF	caccgTTCCGAGCAAAGATTTTCAG gttttaga	Forward oligonucleotide for annealing to HQPM_D142F143_sgr for component 1 in construction of pAR003.
HQPM_D142 F143_sgG	tagctctaaaacCTGAAAATCTTTGCTC GGAAC	Reverse oligonucleotide for annealing to HQPM_D142F143_sgf for component 1 in construction of pAR003.
HQPM_D142 F143_scaffF	GCTAGAAATAGCAAGTTAAAATAAG GCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCG	Forward oligonucleotide for annealing to HQPM_D142F143_scaffR for component 2 in construction of pAR003.

HQPM_D142 F143_scaffR	GCACCGACTCGGTGCCACTTTTTCA AGTTGATAACGGACTAGCCTTATTTT AACTTGCTATTTC	Reverse oligonucleotide for annealing to HQPM_D142F143_scaffF for component 2 in construction of pAR003.
HQPM_D142 F143_extensF	gtgcACAGCCGCCGCGCACTGAAAA TCTTTGCTCGGA	Forward oligonucleotide for annealing to HQPM_D142F143_extensR for coomponent 3 in construction of pAR003.
HQPM_D142F 143_extensR	aaaaTCCGAGCAAAGATTTTCAGT GGCGGCGGCTGT	Reverse oligonucleotide for annealing to HQPM_D142F143_extensF for component 3 in construction of pAR003.
HQPM_D142 F143_PE3_sgF	caccgTATAGCTCAAGTTGACGACC	Forward oligonucleotide for annealing to HQPM_D142F143_PE3_sgR for construction of pAR004.
HQPM_D142 F143_PE3_sgR	aaacGGTCGTCAACTTGAGCTATAc	Reverse oligonucleotide for annealing to HQPM_D142F143_PE3_sgF for construction of pAR004.

2.8.2 pegRNA Plasmid Construction By Golden Gate Assembly

SpyCas9 PE2*, pU6-pegRNA-GG-acceptor, and BPK1520 plasmids were each transformed into 100 μ L competent DH5 α *E. coli*, which were plated onto Mu agar plates containing 100 μ g/mL ampicillin (see Section 2.3.3). 5 mL overnight cultures were grown (see Section 2.3.4) and minipreped (see Section 2.4.1). Purified plasmid DNA was analytical restriction digested (see Section 2.4.2) then electrophoresed on a 1% (w/v) agarose gel (see Section 2.4.3).

Plasmid was constructed to encode pegRNA for *HELQ* D142A/F143A mutation according to methods described in Anzalone et al. (2019). 30 µL reaction mixture was assembled containing 3 µL 10X rCutSmart buffer (NEB), 1 µg purified pU6-pegRNA-GG-acceptor plasmid DNA, 1 µL BsaI-HFv2 (NEB), and SDW. Reaction mixture was incubated in a 1002 Water Bath (GFL) at 37 °C for 1 hour before being incubated on ice. Reaction mixture was electrophoresed in a 1% (w/v) agarose gel (see Section 2.4.3) and 2.2 kb DNA fragment was isolated by gel extraction (see Section 2.4.4). Appropriate oligonucleotides listed in Table 2.12 were annealed in 25 µL reaction mixtures containing 2.5 µL 10X annealing buffer, 1 µL of each appropriate 100 µM oligonucleotide, and SDW by incubating in a MiniAmp Thermal Cycler (Thermo Scientific) at 95 °C for 3 minutes followed by cooling at 0.1 °C/s to 22 °C. Resulting components 1 and 3 DNA were diluted to 1 µM by addition of SDW. Resulting component 2 DNA was phosphorylated and diluted to 1 µM in a 25 µL reaction mixture containing 2.5 µL 10X T4 DNA Ligase Reaction Buffer (NEB), 6.3 µL 4 µM component 2 DNA, 0.5 µL T4 PNK, and SDW that was incubated in the 1002 Water Bath at 37 °C for 1 hour before being incubated on ice. Golden Gate assembly 10 µL reaction mixture was assembled containing 1 µL 10X T4 DNA Ligase Reaction Buffer (NEB), 30 ng purified restriction digested pU6-pegRNA-GG-acceptor plasmid 2.2 kb fragment DNA, 1 µL of each of the three 1 µM components, 1 µL T4 DNA Ligase, and SDW. Reaction mixture was incubated at room temperature for 1 hour then transformed into 100 µL competent DH5α *E. coli*, which were plated onto Mu agar plates containing 100 µg/mL ampicillin (see Section 2.3.3). 5 mL overnight cultures were grown (see Section 2.3.4) and minipreped (see Section 2.4.1). Purified plasmid DNA was analytical restriction digested (see Section 2.4.2) then electrophoresed on a 1% (w/v) agarose gel (see Section 2.4.3). Purified plasmid DNA was stored at -20 °C. Successfully constructed pAR003 plasmid was verified through Sanger sequencing, carried out externally by GENEWIZ (Azenta Life Sciences using LOKO.1 5' primer).

2.8.3 sgRNA Plasmid Construction

Plasmid was constructed to encode sgRNA for *HELQ* D142A/F143A mutation according to methods described in Anzalone et al. (2019). HQPM_D142F143_PE3_sgF and

HQPM_D142F143_PE3_sgR oligonucleotides were diluted to 10 μ M in SDW. Phosphorylated oligonucleotide duplex DNA was generated in a 20 μ L reaction mixture containing 2 μ L 10X T4 DNA Ligase Reaction Buffer (NEB), 1 μ L of each 10 μ M HQPM_D142F143_PE3_sgF and HQPM_D142F143_PE3_sgR oligonucleotide, 0.5 μ L T4 PNK, and SDW. Reaction mixture was incubated in a MiniAmp Thermal Cycler (Thermo Scientific) at 37 $^{\circ}$ C for 1 hour, heated to 95 $^{\circ}$ C for 5 minutes, and cooled to 10 $^{\circ}$ C at 5 $^{\circ}$ C/min. BPK1520 plasmid was dephosphorylated in a 40 μ L reaction mixture containing 4 μ L 10X NEBuffer r3.1, 500 ng purified BPK1520 plasmid DNA, 1 μ L BsmBI-V2, 0.5 μ L Quick CIP, and SDW. Reaction mixture was incubated in an AccuBlock Digital Dry Bath D1302 (Labnet International) at 55 $^{\circ}$ C for 1 hour before being electrophoresed in a 1% (w/v) agarose gel (see Section 2.4.3). Plasmid DNA was then gel extracted (see Section 2.4.4). Plasmid DNA vector was ligated to the oligonucleotide duplex DNA insert in a 10 μ L reaction mixture containing 1 μ L 10X T4 DNA Ligase Reaction Buffer (NEB), 40 ng purified dephosphorylated digested plasmid DNA, 500 nM phosphorylated oligonucleotide duplex DNA, 0.5 μ L T4 DNA ligase, and SDW. Reaction mixture was incubated at room temperature for 1 hour before transformation into 100 μ L competent DH5 α *E. coli*, which were plated onto Mu agar plates containing 100 μ g/mL ampicillin (see Section 2.3.3). 5 mL overnight cultures were grown (see Section 2.3.4) and minipreped (see Section 2.4.1). Purified plasmid DNA was analytical restriction digested (see Section 2.4.2) then electrophoresed in a 1% (w/v) agarose gel (see Section 2.4.3). Purified plasmid DNA was stored at -20 $^{\circ}$ C. Successfully constructed pAR004 plasmid was verified through Sanger sequencing, carried out externally by GENEWIZ (Azenta Life Sciences) using OS280 primer.

2.8.4 Maxiprep Extraction and Purification of Prime Editing Plasmids DNA From Transformed DH5 α *E. coli* Overnight Cultures

Maxipreps were performed to reach desired masses of prime editing plasmids. Protocol was performed under sterile conditions. SpyCas9 PE2*, pAR003, and pAR004 plasmids were each transformed into 100 μ L competent BL21-AI *E. coli*, which were plated onto Mu agar plates containing 100 μ g/mL ampicillin (see Section 2.3.3). Individual colonies were picked from each plate and separately inoculated into 500 mL Mu broth containing 100 μ g/mL ampicillin in a baffled conical flask. Cultures were incubated in a New Brunswick Innova 44 Stackable

Incubator Shaker (Eppendorf) at 37 °C with 130 RPM shaking overnight for 18 hours. Plasmid DNA was extracted and purified from cultures using the ZymoPURE II Plasmid Maxiprep Kit (Zymo Research) via the centrifugation protocol and endotoxin removal step using an Avanti J-25 Centrifuge with a JLA-10.500 rotor (Beckman Coulter), Centrifuge 5430 R (Eppendorf), and Heraeus Pico 17 Centrifuge (Thermo Scientific), according to the manufacturer's instructions

(https://files.zymoresearch.com/protocols/d4202_d4203_zymopure_ii_plasmid_maxiprep.pdf) (Zymo Research, 2023). Purified plasmid DNA concentration was measured in triplicate on a DS-11+ Spectrophotometer (DeNovix) and averaged. Purified plasmid DNA was stored at -20 °C. Purified plasmid DNA was analytical restriction digested (see Section 2.4.2), then reaction mixtures were electrophoresed in a 1% (w/v) agarose gel (see Section 2.4.3). Purified plasmid DNA was stored at -20 °C.

2.8.5 Transfection By Lipofection of Human Cells With Prime Editing Plasmids

Cells were passaged (see Section 2.7.3), then resuspended cells were counted by mixing 10 µL cell suspension with 10 µL BioWhittaker Trypan Blue 0.4% 0.85% NaCl (Lonza Bioscience) and counting the number of cells in 10 µL mixture using a Neubauer Improved Haemocytometer Counting Chamber (Hawksley) under an Invertoskop ID 03 (ZEISS) microscope. Cell suspension was diluted such that 100,000 cells were seeded in 2 mL medium in each well of a clear flat-bottom non-TC 24-well plate. Plate was then incubated for 24 hours. Cells were imaged using a ZOE Fluorescent Cell Imager (Bio-Rad).

After incubation, transfection of cells by Lipofectamine 3000 Transfection Reagent (Invitrogen) was carried out according to the manufacturer's instructions (https://documents.thermofisher.com/TFS-Assets/LSG/manuals/lipofectamine3000_protocol.pdf) (Invitrogen, 2016) and methods described in Anzalone et al. (2019), although the latter used Lipofectamine 2000 Transfection Reagent (Invitrogen). Mixtures were assembled containing 0.75 µL or 1.5 µL Lipofectamine 3000 Reagent diluted in 25 µL supplement-free stock DMEM (Gibco). Separately, 52 µL master mixtures were assembled containing purified maxiprep pEGFP-C1 plasmid DNA, or

purified maxiprepmed SpyCas9 PE2*, pAR003, and/or pAR004 plasmid DNA individually or all three together, 2 μ L P3000 Reagent, and supplement-free stock DMEM. Final mixtures were assembled by mixing 25 μ L appropriate master mix with appropriate diluted Lipofectamine 3000 Reagent mixture via pipetting, such that final mixtures would contain 500 ng pEGFP-C1 plasmid DNA for a positive control, or 750 ng SpyCas9 PE2*, 250 ng pAR003, and/or 83 ng pAR004 plasmid individually or all three together, or no plasmid DNA. Final reaction mixtures were separately added to appropriate individual wells in the plate in duplicate. Plate was then incubated, cells were imaged using the ZOE Fluorescent Cell Imager before and at 26 hours, 46 hours, and 71 hours after transfection, in addition to green fluorescent channel images of the positive control at those time points. Images were saved as TIFs. At the latter two timepoints, genomic DNA of cells transfected with all three prime editing plasmids was extracted (see Section 2.8.6).

2.8.6 Human Cell Genomic Extraction and Analysis

Genomic DNA extraction of cells lipofected with all three prime editing plasmids using the PureLink Genomic DNA Mini Kit (Invitrogen) according to the manufacturer's instructions (https://documents.thermofisher.com/TFS-Assets/LSG/manuals/purelink_genomic_mini_man.pdf) (Invitrogen, 2007). Purified extracted genomic DNA concentration was measured in triplicate on a DS-11+ Spectrophotometer (DeNovix) and averaged. Extracted genomic DNA was stored at -20 °C. 50 ng genomic DNA from each extraction was then PCR amplified using HELQ_D142/F143_F and HELQ_D142/F143_R primers with annealing temperature of 65 °C and extension step time of 24 seconds to amplify the region of interest (see Section 2.4.5). PCR product DNA sizes were verified by electrophoresing 12 μ L sample containing 7 μ L SDW, 3 μ L PCR sample, and 2 μ L 6X Purple Gel Loading Dye (NEB) on a 2% (w/v) agarose gel alongside 100 bp DNA Ladder (NEB) at 120 V for 1 hour and 15 minutes (see Section 2.4.3).

12 μ L sample containing 0.5 μ L SDW, 8.5 μ L PCR sample, and 2 μ L 6X Purple Gel Loading Dye (NEB) for each PCR reaction mixture was then electrophoresed on a 2% (w/v) agarose gel alongside 1 kb DNA Ladder (NEB) at 120 V for 1 hour and 15 minutes (see Section 2.4.3).

Desired PCR product DNA was isolated by gel extraction (see Section 2.4.4). Purified PCR product DNA was Sanger sequenced internally by Deep Seq (University of Nottingham) using HELQ_D142/F143_F primer. Sanger sequencing trace ab1 files were converted to FASTQ format using the online tool Abi to Fastq Sequence Converter ([https://sequenceconversion.bugaco.com/converter/biology/sequences/abi to fastq.php](https://sequenceconversion.bugaco.com/converter/biology/sequences/abi%20to%20fastq.php)) for assessing prime editing success using the online tool CRISPResso2 (<https://crispresso2.pinellolab.org/submission>) (Clement et al., 2019).

2.9 Molecular Modelling

Protein residue sequences were obtained either from MANE Select transcript corresponding protein sequences on the online database Ensembl (<https://www.ensembl.org/index.html>) (Dyer et al., 2025) or plasmids modelled in the software SnapGene (<https://www.snapgene.com/>). Protein structural and interaction predication was carried out using the model AlphaFold 3, accessed via the website AlphaFold Server (<https://alphafoldserver.com/welcome>) (Abramson et al., 2024). Predicted structures were visualised using the software ChimeraX (<https://www.rbvi.ucsf.edu/chimerax/>) (Pettersen et al., 2024).

Chapter 3: Analysis of POLD3-N-HELQ Protein-Protein Interaction

3.1 Introduction

3.1.1 POLD3 Structure and Function

As described previously, DNA polymerase Pol δ functions to synthesise DNA with high-fidelity with 5'-3' directionality, primarily for DNA replication of the lagging strand's parental template strand (see Section 1.1.4) and extending D-loops during SDSA and BIR (see Section 1.2.2) (He et al., 2023; Lujan et al., 2016; Schmitt et al., 2009; Sneed et al., 2013). Pol δ is a heterotetramer complex, comprised of the subunits POLD1, POLD2, POLD3, and POLD4 (Khandagale et al., 2020; Zhou et al., 2012). These are summarised in Table 3.1. The apo-form Pol δ cryogenic electron microscopy (cryo-EM) structure was resolved by Shin et al. (2025) and is shown in Figure 3.1.

Table 3.1 Summary of human DNA polymerase Pol δ complex subunits.

Pol δ Subunit	Cytogenetic Location	Number of Amino Acid Residues	Molecular Mass (kDa)	Known Interacting Pol δ Subunits	Main Function Within Pol δ
POLD1	19q13.33 (NCBI, 2025c)	1,107 (Chung et al., 1991).	124 (Yang et al., 1992),	POLD2, POLD3*, and POLD4 (Podust et al., 1995; Podust et al., 2002)	Catalytic subunit that synthesises DNA with 3'-5' proofreading exonuclease activity for high-fidelity (Weedon et al., 2013).

POLD2	7p13 (NCBI, 2025d)	469 (Zhang et al., 1995)	51.3 (Zhang et al., 1995)	POLD1, POLD3, POLD4 (Podust et al., 1995; Podust et al., 2002)	Accessory subunit that acts as a scaffold to maintain complex stability and function (Cong et al., 2022).
POLD3	11q13.4 (NCBI, 2025e)	466 (Ensembl, 2025a)	51.4 (Ducoux et al., 2001; Hughes et al., 1999)	POLD1* and POLD2 (Podust et al., 1995; Podust et al., 2002)	Accessory subunit that regulates complex DNA synthesis (Alli et al., 2024).
POLD4	11q13.2 (NCBI, 2025f)	107 (Liu et al., 2000)	12.4 (Liu et al., 2000)	POLD1 and POLD2 (Podust et al., 1995; Podust et al., 2002)	Accessory subunit that regulates complex enzymatic activity by enhancing DNA synthesis but reducing proofreading activity (Lee et al., 2017; Xu et al., 2025).

* POLD1-POLD3 interaction has been described as questionable (Alli et al., 2024).

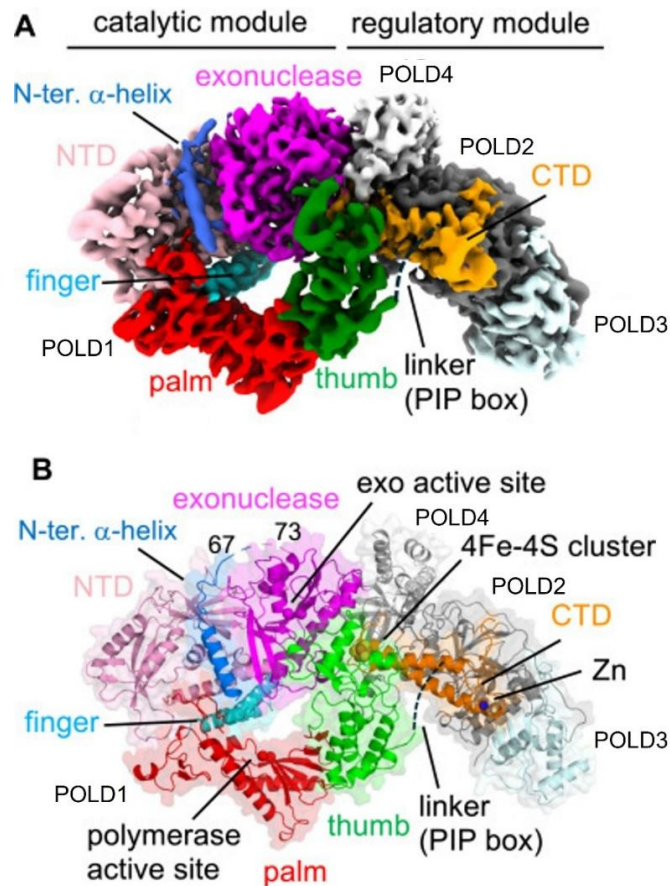


Figure 3.1 A. Apo-form human DNA polymerase Pol δ cryogenic electron microscopy structure with opaque surface.

Pol δ is comprised of a catalytic module and a regulatory module. The catalytic module is the catalytic POLD1 subunit, which has a N-terminal domain (NTD) (violet), exonuclease domain (indigo), polymerase domain, and C-terminal domain (CTD) (yellow). The NTD contains a an α -helix (blue). The polymerase domain contains a thumb subdomain (green), finger subdomain (cyan), and palm subdomain (red). The catalytic module is connected to the regulatory module by a PCNA-interacting protein (PIP) box linker. The regulatory module is comprised of accessory subunits POLD2 (dark grey), POLD3 (white), and POLD4 (light grey). Figure is adapted from Figure 2A in Shin et al. (2025).

B. Apo-form human DNA polymerase Pol δ cryogenic electron microscopy structure with cartoon and transparent surface.

POLD1 exonuclease and polymerase active sites are indicated. The POLD1 CTD contains a CysA motif with a Zn^{2+} ion and a CysB motif with a 4Fe-4S cluster. Figure is adapted from Figure 2B in Shin et al. (2025).

The POLD3 subunit is primarily involved in maintaining genomic integrity by preventing genomic instability and preserving replication forks under stress, via functioning as a controller of replicative DNA repair processes such as BIR (Alli et al., 2024; He et al., 2023; Tumini et al., 2016). POLD3 features an N-terminal WHD (residues 1-106), while the remainder of the protein is intrinsically disordered. The WHD 'is the only part of POLD3 that [has been] structurally resolved', hence its structure is mostly unresolved (Alli et al., 2024). Among the intrinsically disordered region are four designated IDPRs: IDPR1 (residues 169-188), IDPR2 (169-188), IDPR3 (274-393), and IDPR4 (residues 406-466).

POLD3 protein structure contains a number of motifs that enable its function during BIR (Alli et al., 2024). POLD3 first reaches the nucleus due to the presence of an IDPR3 NLS (residues 319-KKRRRIKL-326), which is recognised by an importin protein that transports POLD3 into the nucleus (Lu et al., 2021; Pohler et al., 2005). IDPR3 also contains a putative second NLS (PNLS) (residues 373-GENKRKRKR-385) (Alli et al., 2024). Like the other Pol δ subunits, POLD3 can bind to the PCNA DNA sliding clamp, which it achieves through its IDPR4 PIP box (residues 456-QVSITGFF-463); POLD3 is the subunit that binds to PCNA during BIR (Drosopoulos et al., 2020; Ensembl, 2025b; Gonzalez-Magaña et al., 2019; Khandagale et al., 2020; Zheng et al., 2020). Once bound to PCNA, POLD3 can assemble as part of Pol δ via its WHD β 3 sheet (residues 73-SCHKVAVV-80) for interaction with POLD2 (Baranovskiy et al., 2008; Drosopoulos et al., 2020; Li et al., 2009; Zheng et al., 2020). As part of the Pol δ complex, POLD3 prevents hyperactive DNA synthesis, maintains processivity, and facilitates bypass of lesions such as abasic sites, although the mechanisms of how it exerts this control are currently unknown (Alli et al., 2024). Additionally, during HR, POLD3 can interact with REV1 through its REV1 interaction region (residues 237-NFFGKAA-243); REV1 functions to protect nascent DNA strands at stalled replication forks (Pustovalova et al., 2016; Yang et al., 2015). POLD3 may also contain a DNA polymerase interacting motif (residues 390-KTYLDGEGCIV-400) for interaction with DNA polymerase Pol α , but such interaction has not been shown (Alli et al., 2024; Ensembl, 2025b). POLD3 residue information is illustrated in Figure 3.2, while its structure as predicted by AlphaFold 3 is shown in Figure 3.3, and illustrated with the WHD and IDPRs in Figure 3.4.

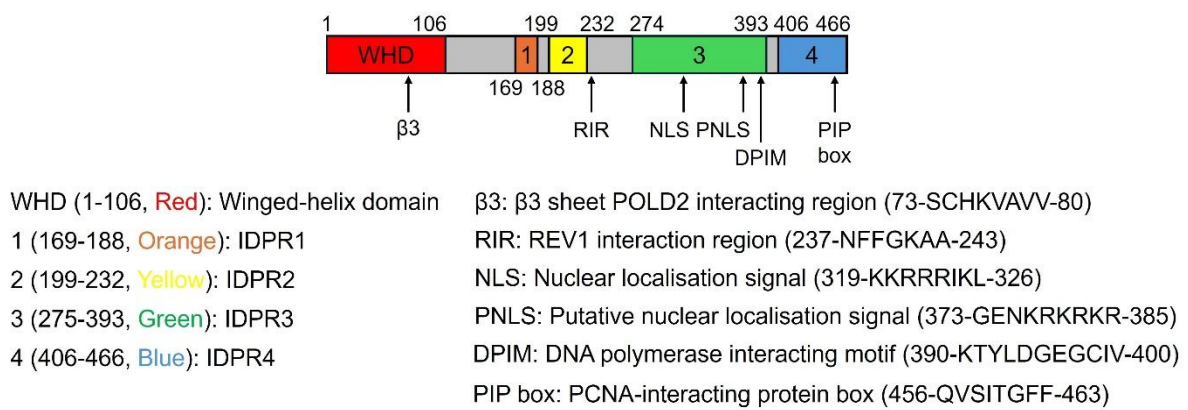


Figure 3.2 POLD3 amino acid residue information.

POLD3 contains an N-terminal winged helix domain (WHD) and four designated intrinsically disordered protein regions (IDPRs), as denoted by letters/number and colour. WHD and IDPR residue ranges in addition to functional motifs are indicated (Alli et al., 2024; Baranovskiy et al., 2008; Ensembl, 2025b; Gonzalez-Magaña et al., 2019; Pohler et al., 2005; Pustovalova et al., 2016).

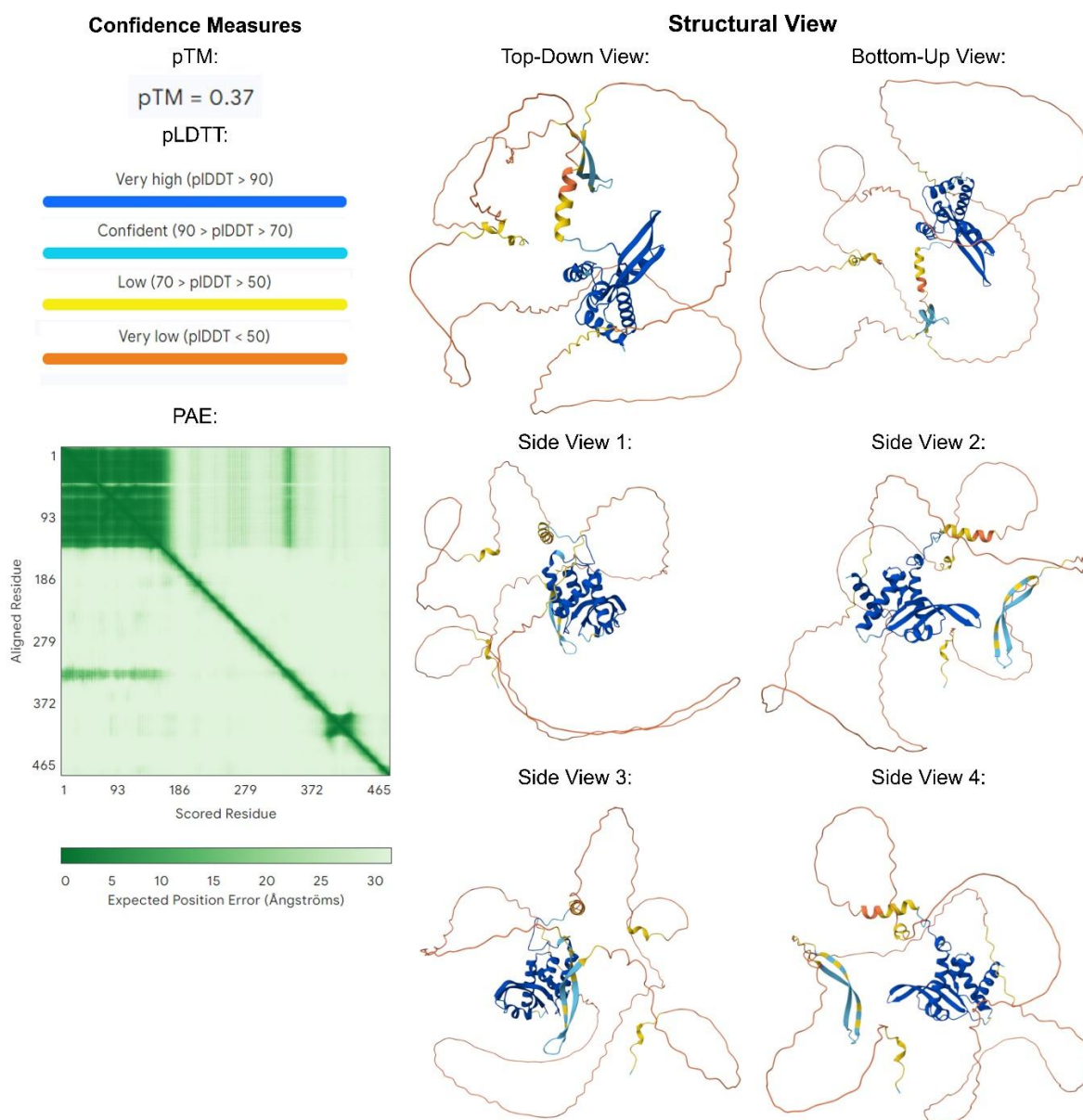
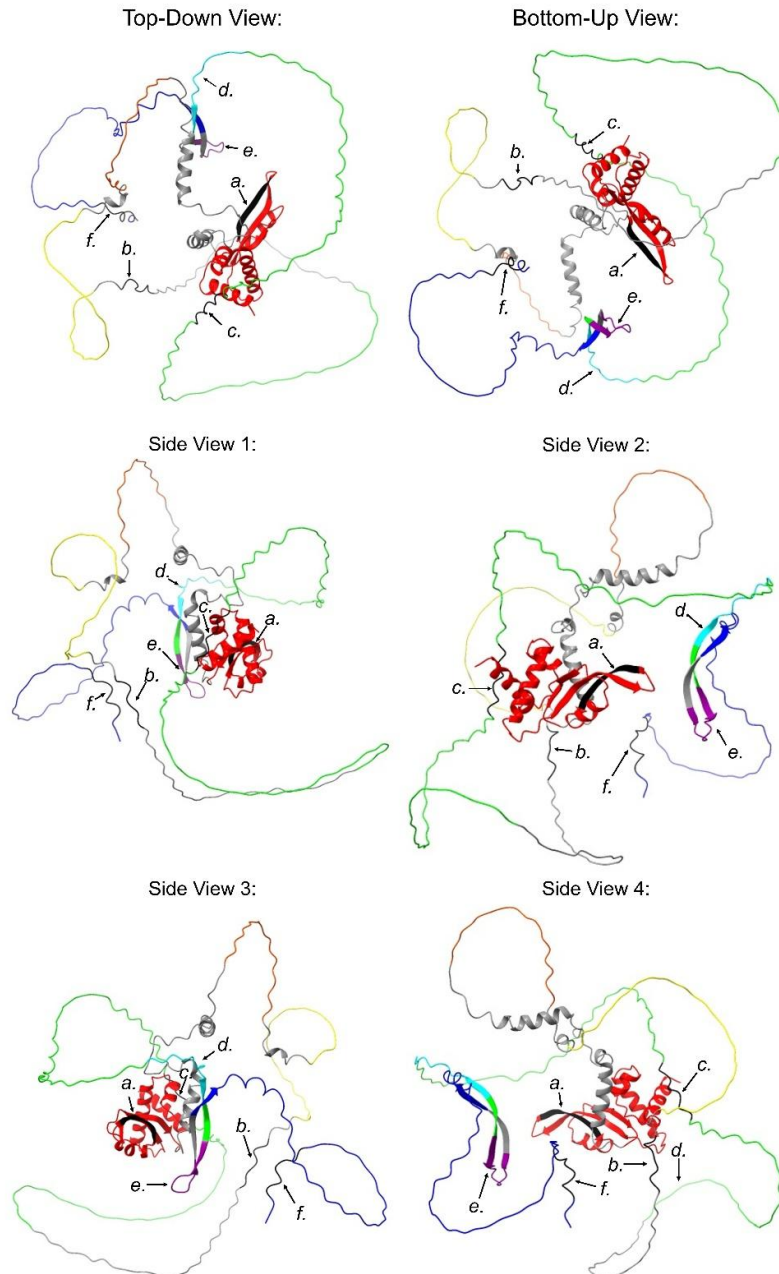


Figure 3.3 POLD3 structure predicted by the software AlphaFold 3.

Structure is shown from multiple orientations. AlphaFold 3 uses multiple confidence measures (Abramson et al., 2024). Predicted Template Modelling (pTM) score measures overall structure prediction accuracy, ranging from 0-1 with threshold of 0.5; a higher score indicates greater accuracy (Wee et al., 2024). Predicted Local Distance Difference Test (pLDDT) score measures individual residue structure prediction confidence, ranging from 0-100; a higher score indicates greater confidence. pLDDT score is denoted by residue colour. Predicted Align Error (PAE) assesses overall structural prediction confidence, measured in angstroms. Lower PAE indicates greater prediction confidence (Pratt et al., 2025).



Red: Winged-helix domain (1-106)
 Orange: IDPR1 (169-188)
 Yellow: IDPR2 (199-232)
 Green: IDPR3 (275-393)
 Blue: IDPR4 (406-466)

a. β 3 sheet POLD2 interacting region (73-SCHKVAVV-80, black)
 b. REV1 interaction region (237-NFFGKAA-243, black)
 c. Nuclear localisation signal (319-KKRRRIKL-326, black)
 d. Putative nuclear localisation signal (373-GENKRRKR-385, cyan)
 e. DNA polymerase interacting motif (390-KTYLDGEGCIV-400, purple)
 f. PCNA-interacting protein box (456-QVSITGFF-463, black)

Figure 3.4 POLD3 predicted structure organisation.

POLD3 structure as predicted by AlphaFold 3 (Figure 3.3) is shown from multiple orientations. Winged-helix domain (WHD) and protein intrinsically disordered regions (IDPRs) are indicated by colour. Functional motifs are coloured black, cyan, or purple and indicated by letter.

3.1.2 Bimolecular Fluorescence Complementation Assays

In previous work by the Bolt lab, He et al. (2023) attempted several methods to detect stable POLD3-N-HELQ interaction, including size-exclusion chromatography (SEC), affinity protein pull-downs, surface plasmon resonance (SPR), and electrophoretic mobility shift assays (EMSAs). These were unsuccessful, likely due to POLD3 and N-HELQ intrinsic disorder resulting in transient interactions (He et al., 2023; Wohl et al., 2025). Therefore, He et al. (2023) used a biomolecular fluorescence complementation (BiFC) assay approach. A BiFC assay is a non-invasive fluorescence-based technique used to detect PPI between protein pairs, traditionally using an *in vivo* method in living cells (Miller et al., 2015; Ohad et al., 2007). In principle, a cell coexpresses two fusion proteins, each being a protein of interest fused by a linker region to either a N-terminal or C-terminal non-fluorescent fragment of a fluorescent protein (FP). If the two proteins of interest interact, the non-fluorescent fragments reconstitute the full-length FP as a complex. If the reconstituted FP is excited at a specific radiation wavelength and consequently emits a specific desired higher wavelength, this indicates successful PPI (Dos Santos Rodrigues et al., 2023; Miller et al., 2015). This principal is illustrated in Figure 3.5. BiFC FP complex formation is often irreversible *in vivo*, hence a BiFC assay can detect even transient interactions with high sensitivity, and thus successfully showed POLD3-N-HELQ interaction (He et al., 2023; Miller et al., 2015; Uversky, 2020).

The BiFC concept was first developed by Ghosh et al. (2000), using green fluorescent protein (GFP) in BL21(DE3) *E. coli* cells, and the first BiFC assay was performed by Hu et al. (2002), using yellow fluorescent protein (YFP) in COS-1, NIH3T3, and HeLa mammalian cells. Since then, BiFC assay methods have been developed using a range of cell lines, FPs, and applications (Kodama et al., 2012; Miller et al., 2015; Ren et al., 2024). The method currently employed by the Bolt lab from He et al. (2023) is a time-course approach using mVenus in BL21-AI *E. coli* cells. mVenus is a YFP variant with emission/excitation maximum wavelengths of 515/527 nm (FPbase, 2025; Nagai et al., 2002). It was split into N-terminal fragment NmVenus (residues 1-154) and C-terminal fragment CmVenus (residues 155-238), fused to

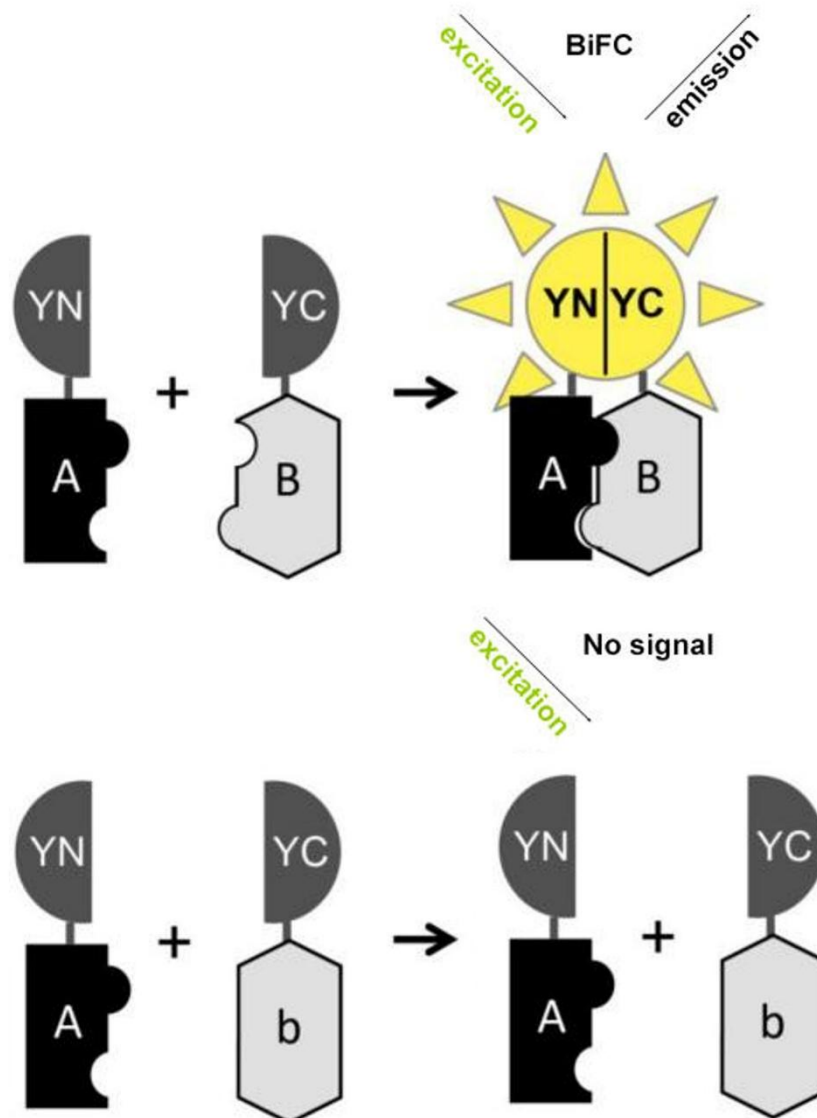


Figure 3.5 Principal of bimolecular fluorescence complementation (BiFC) assays.

Top: A cell coexpresses two fusion proteins. If proteins A and B interact, their respective fused N-terminal and C-terminal non-fluorescent fragments YN and YC reconstitute a fluorescent protein complex, in this case yellow fluorescent protein (YFP). If YFP is excited, it emits yellow fluorescence, indicating successful protein-protein interaction (PPI). Bottom: If proteins A and B do not interact, no YFP is formed. Consequently, there will be no yellow fluorescence, indicating no PPI. Figure is adapted from Figures 1 in Miller et al. (2015) and 1A in Ohad et al. (2007).

POLD3 and N-HELQ, respectively (He et al., 2023). These fragments were chosen because if PPI occurs in a BiFC assay using *E. coli* at 37 °C under aerobic conditions, the fragments can

rapidly reconstitute into a full-length FP (Dillard et al., 2018; He et al., 2023; Miller et al., 2015; Nagai et al., 2002; Tuttle et al., 2021). POLD3 and N-HELQ in the fusion proteins lacked their first residues encoded by start codons, to prevent over-expression of non-fusion proteins, which would in turn reduce fusion protein over-expression by competing with it for metabolic resources (Kim et al., 2020).

To summarise briefly the He et al. (2023) method, firstly overnight cultures of competent BL21-AI *E. coli* (co)transformed with one or a pair of plasmid(s) encoding BiFC assay proteins were set up, including NmVenus-POLD3 + CmVenus-N-HELQ and necessary controls. 100 μ L of each respective overnight culture was inoculated in 5 mL Mu broth and incubated at 37°C with shaking. Protein overexpression was induced at OD₆₀₀ 0.5 using L-arabinose and IPTG at 0.05% (w/v) and 1 mM respective working concentrations. At regular intervals over 3 hours, 100 μ L of cells from each culture were pelleted, resuspended in 150 μ L of ice-cold PBS, and transferred to a clear flat-bottom non-TC 96-well plate. Finally, the plate was scanned using the built-in Cy2 filter, and imaged, in an Amersham Typhoon 5 Biomolecular Imager (Cytiva).

In addition to identifying interacting proteins, BiFC assays can also be used to identify residues necessary for these PPIs by comparing fluorescence complementation efficacy between WT and mutant fusion proteins. If mutation of one or more specific residues of an interacting protein of interest alters the complementation efficacy and hence relative fluorescence level, it indicates that the mutation affects interaction between the protein pair, and thus the mutated residues may be necessary for the PPI (Kerppola, 2006a). This is because a PPI's nature fundamentally depends on the sequence and properties of the involved proteins' constituent residues (Stark, 2023). Hence, it was decided to continue the POLD3-N-HELQ BiFC assay work begun in He et al. (2023) in order to identify POLD3 residues necessary for the PPI, giving the first aim of this project.

3.2 Analysis of POLD3 Quincunx Motif Mutation Impact on POLD3-N-HELQ Protein-Protein Interaction

3.2.1 Producing Plasmid Stocks of BiFC Assay Plasmids

The POLD3-N-HELQ BiFC assay work was originally continued by Lou-Hing (2024), who used the He et al. (2023) method to investigate the effect on the PPI of mutating the polar, charged Quincunx motif (residues 319-KKRRR-323) in the NLS to polar, uncharged glycines (319-GGGGG-323) (JPT, 2025a; JPT, 2025b; Satoh et al., 2022). The preliminary data showed full PPI inhibition, so it was first decided to attempt to reproduce this result. BiFC assay plasmid stocks were produced by transforming competent DH5α *E. coli* with pBAD-FLmVenus, pBAD-NmVenus, pRSF-CmVenus, pBAD-NmVenus-PoID3, pALH008 (pBAD-NmVenus-PoID3^{QC}), or pRSF-CmVenus-NHelQ (see Section 2.3.3.), then overnight cultures were grown (see Section 2.3.4) and minipreped (see Section 2.4.1). The resulting plasmid DNA was analytically restriction digested (see Section 2.4.2) and electrophoresed on a 1% (w/v) agarose gel (see Section 2.4.3) (Figure 3.6). When digested using *AccI* and *HindIII*-HF, the pBAD-NmVenus-PoID3^{QC} plasmids bands showed unexpected sizes of approximately 5500 and 2250 bp compared to the expected sizes of 3513 and 2429 bp, and so new pBAD-NmVenus-PoID3^{QC} stocks were produced from different original stocks using the same method as before. These were then analytical restriction digested and electrophoresed on another 1% (w/v) agarose gel as before, this time using *PvuI*-HF and *HindIII*-HF (Figure 3.7). The old stocks now showed unexpected size bands of approximately 8500 and 700 bp compared to the expected sizes of 5000 and 942 bp, but the new stocks were the expected lengths, having bands of the same size as the digested pBAD-NmVenus-PoID3 positive control. Subsequently, attempts to reproduce the BiFC assay result from Lou-Hing (2024) were unsuccessful (data not shown).

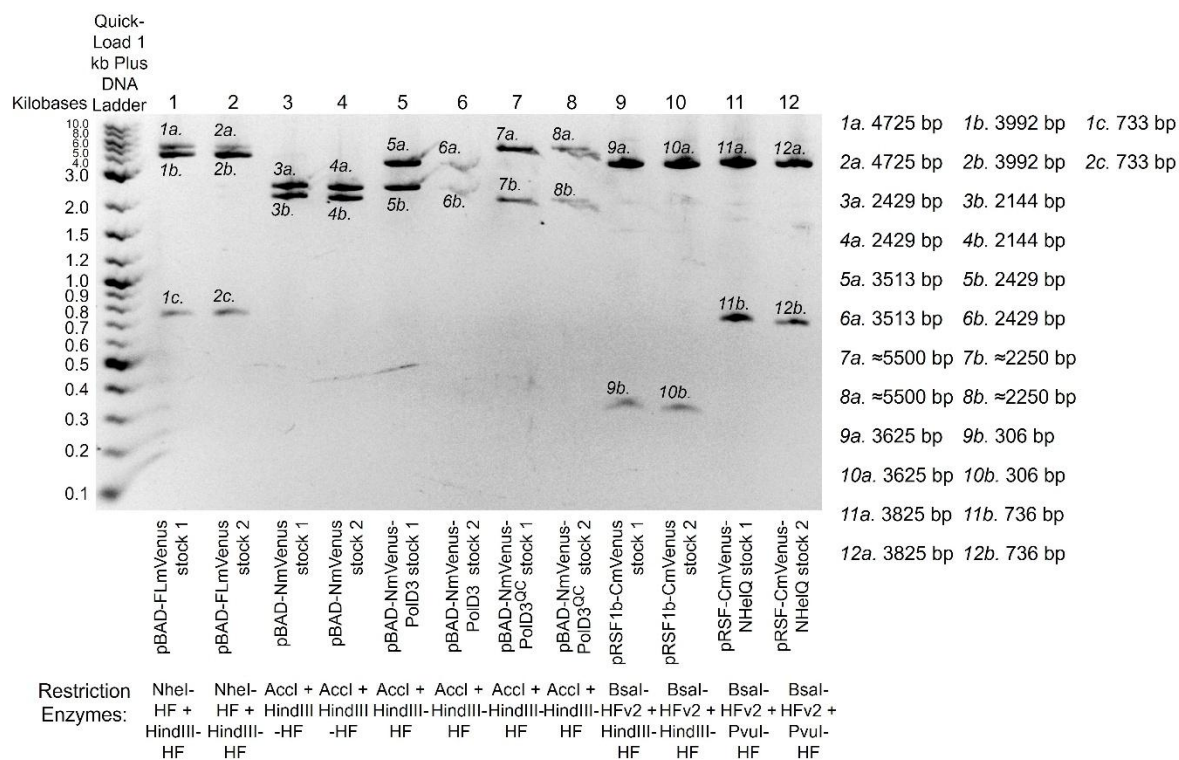


Figure 3.6 Agarose gel of electrophoresed analytical restriction digests of BiFC assay plasmid stocks verifying plasmid lengths.

Double analytical restriction digests were electrophoresed on a 1% (w/v) agarose gel alongside Quick-Load 1 kb Plus DNA Ladder (NEB). Image was inverted and brightness/contrast adjusted using the software ImageJ. Ladder DNA lengths in kilobases are shown to the left of the gel. Lane numbers are shown above the gel. Digested plasmids and restriction enzymes used in analytical digests are shown below the corresponding lane. Plasmid fragment lengths in base pairs (bp) are shown to the right of gel, corresponding to lane number and letter. DNA fragments given as approximations (≈) indicate bands of unexpected length, otherwise they are of expected length.

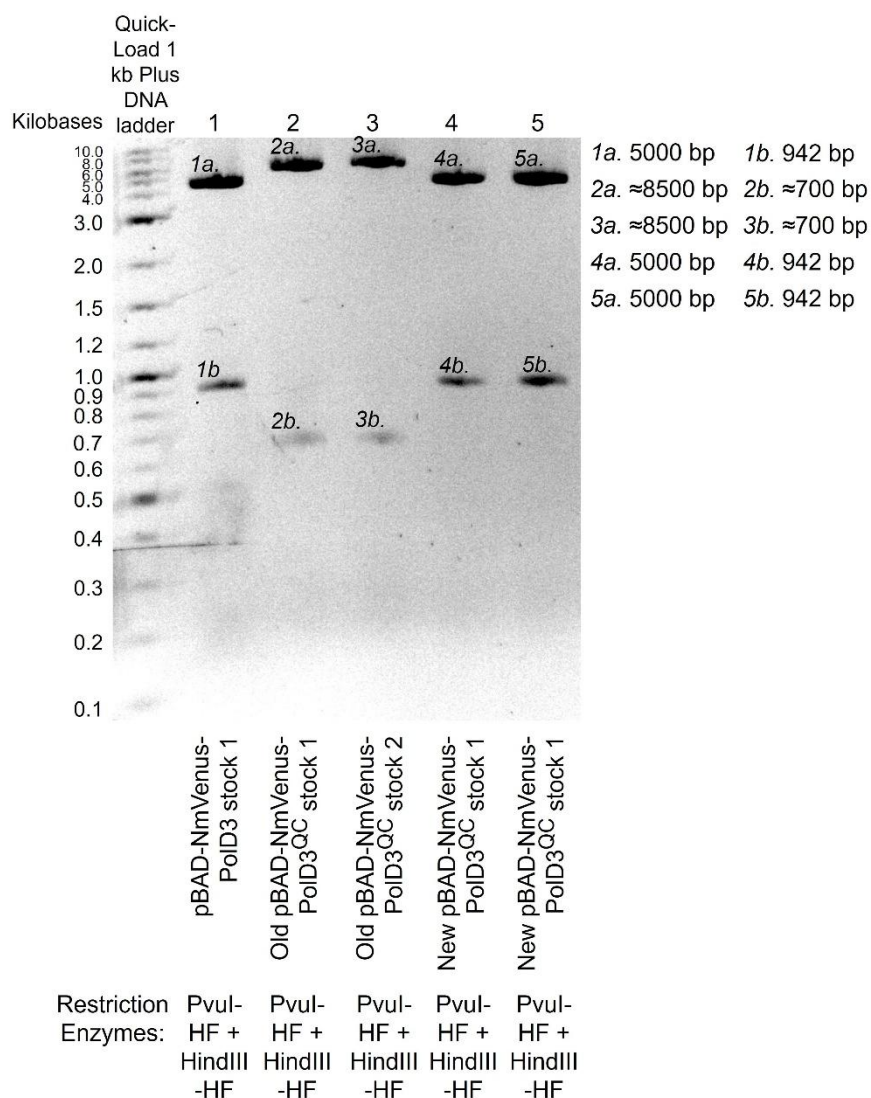


Figure 3.7 Agarose gel of electrophoresed analytical restriction digests of old and new pBAD-NmVenus-PoID3^{QC} plasmid stocks compared to pBAD-NmVenus-PoID3, verifying plasmid lengths.

Double analytical restriction digests were electrophoresed on a 1% (w/v) agarose gel alongside Quick-Load 1 kb Plus DNA Ladder (NEB). Image was inverted and brightness/contrast adjusted using the software ImageJ. Ladder DNA lengths in kilobases are shown to the left of the gel. Lane numbers are shown above the gel. Digested plasmids and restriction enzymes used in analytical digests are shown below the corresponding lane. Plasmid fragment lengths in base pairs (bp) are shown to the right of gel, corresponding to lane number and letter. DNA fragments given as approximations (≈) indicate bands of unexpected length, otherwise they are of expected length.

3.2.2 BiFC Assay

Due to being unable to reproduce the result from Lou-Hing (2024), it was decided to turn to a currently unpublished method developed by the Bolt lab's then-postdoc Doctor Tom Killelea that uses the same mVenus plasmids system. Unlike the time-course He et al. (2023) method, it is an endpoint method that incorporates a quantitative fluorescence scan using a FLUOstar Omega (BMG LABTECH) plate reader in addition to the qualitative Amersham Typhoon 5 Biomolecular Imager (Cytiva). By combining these quantitative and qualitative approaches, the qualitative Amersham Typhoon scan still provides a highly valid contextual summary, but its weakness of inherent subjectivity in terms of proving PPI can be covered by the quantitative FLUOstar Omega results, which allow in-depth numerical analysis; together, they can hence provide a more reliable and factual result than if using only one of them (Horstman et al., 2014; Pyo et al., 2023; Verhoef et al., 1997).

Several potential developments to the original Killelea method were identified, giving the second aim of this project. Firstly, an uninoculated Mu broth growth medium blank was added for greater BL21-AI *E. coli* OD₆₀₀ measurement accuracy (Atolia et al., 2020). Secondly, cell samples were stored overnight to allow greater FP reconstitution and maturation before endpoint measurements, thereby increasing the possibility of detecting successful PPI (Kerppola, 2006a; Kerppola, 2009). The cells were stored at 4 °C in 1X M9 minimal medium to prevent cell growth and proliferation, hence reducing stochastic variability while allowing for continued cell viability (Adiciptaningrum et al., 2015; Dağ et al., 2025; Paliy et al., 2007; Shaw et al., 1971; Thomas et al., 2018; Wei et al., 2018; Wolffe, 1995; Zhang et al., 2025). Thirdly, Amersham Typhoon scans were carried out in heatmap, not just greyscale, for greater fluorescence perception, and also contained a 1X M9 minimal medium triplicate to show that there is medium autofluorescence is negligible (Bhakta et al., 2024; Qutbi, 2024). Finally, OD₆₀₀ is sensitive to nonviable cells and cellular debris, which can reduce accuracy if using endpoint OD₆₀₀ measurements for fluorescence per viable cell normalisation (Beal et al., 2020; Costa et al., 2021). Therefore, a spot assay was included to estimate the number of viable cells in a sample because they are insensitive to most factors that can confound OD₆₀₀

measurements in terms of viability (Bhuyan et al., 2023; Costa et al., 2021; Sahu et al., 2023). Meanwhile, controls were kept the same. pBAD-NmVenus + pRSF-CmVenus was used as the positive control, since He et al. (2023) and their unpublished data showed that cotransformations with a non-fluorescent fragment and fusion protein risk false-positive results, which is not seen when using only the non-fluorescent fragments. pBAD-NmVenus-PoID3 + pRSF-CmVenus-NHelQ was used as the positive control due to having showed successful PPI and its mutants being the investigative subject of the BiFC assay. pBAD-FLmVenus was not included as a positive control because BiFC assay reconstituted FPs typically have fluorescence around 0.1-10% of the full-length protein, and so the full-length protein could adversely affect FLUOstar Omega fluorescence gain adjustment (Kerppola, 2009). Altogether, this led to the method described in Section 2.6.1.

Bolt lab PhD student Olivia Downs generated pBAD-NmVenus-PoID3 plasmids with Quincunx motif single or double point mutations by SDM (see Section 2.4.7) to determine if the mutations affect POLD3-N-HELQ PPI. The plasmids are detailed in Table 3.2.

Table 3.2 Details of pBAD-NmVenus-PoID3 Quincunx motif mutant plasmids generated by SDM.

Plasmid	Mutation(s)	Quincunx Motif Amino Acid Residue Sequence	Mutation Effect on Residues
pOLD038 (pBAD-NmVenus-PoID3 ^{KKSRR})	R321S	KKSRR	Polar positive charge to polar uncharged in middle of tract.*
pOLD039 (pBAD-NmVenus-PoID3 ^{KSRRR})	K320S	KSRRR	Polar positive charge to polar uncharged in middle of tract.*
pOLD040 (pBAD-NmVenus-PoID3 ^{KKDRR})	R321D	KKDRR	Polar positive charge to polar negative

			charge in middle of tract.*
pOLD041 (pBAD-NmVenus-PoID3 ^{AKRRA})	K319A/R323A	AKRRA	Polar positive charge to nonpolar uncharged at bookends to tract.*

* (JPT, 2025a; JPT, 2025b; Satoh et al., 2022; Zhou et al., 2018)

The BiFC assay was carried out as described in Section 2.6.1, using the following cotransformations: pBAD-NmVenus + pRSF-CmVenus, pBAD-NmVenus-PoID3 + pRSF-CmVenus-NHelQ, pBAD-NmVenus-PoID3^{QC} + pRSF-CmVenus-NHelQ, pBAD-NmVenus-PoID3^{KSRRR} + pRSF-CmVenus-NHelQ, pBAD-NmVenus-PoID3^{KKSRR} + pRSF-CmVenus-NHelQ, pBAD-NmVenus-PoID3^{KKDRR} + pRSF-CmVenus-NHelQ, and pBAD-NmVenus-PoID3^{AKRRA} + pRSF-CmVenus-NHelQ. Glycerol stocks were also produced for each cotransformation (see Section 2.3.5). In accordance with the original Killelea method, SDS-PAGE analysis to verify protein overexpression only used one sample from each triplicate (Figure 3.8). The samples were electrophoresed on a 10% (w/v) SDS-PAGE gel at 120 V for 1 hour and 10 minutes. NmVenus, NmVenus-POLD3, and NmVenus-POLD3 Quincunx motif mutants were visible in the appropriate lanes, but CmVenus and CmVenus-N-HELQ were not. A Western blot could not be performed for the latter two due to time constraints.

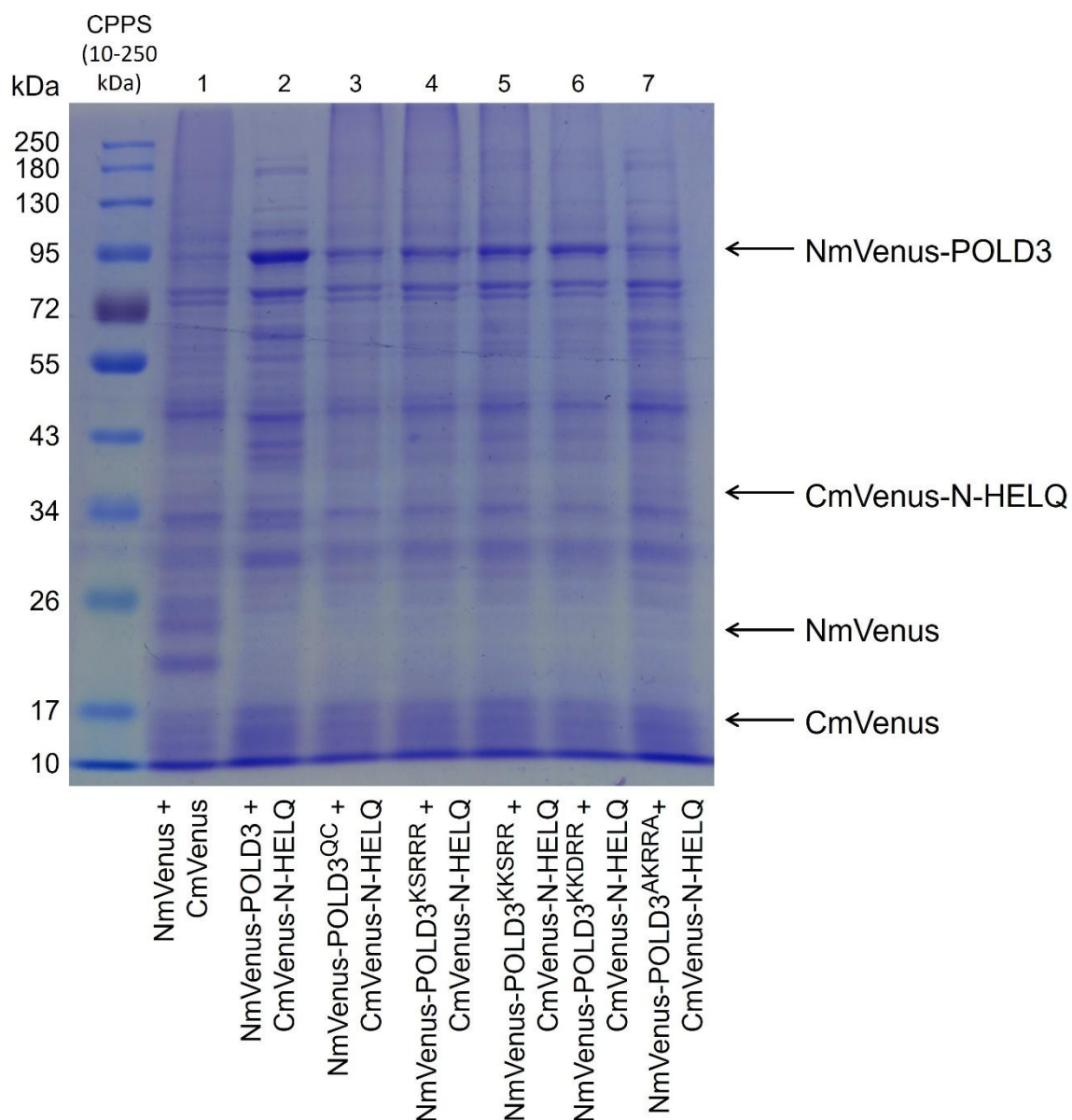


Figure 3.8 SDS-PAGE gel of POLD3 Quincunx motif mutants BiFC assay SDS-PAGE samples to verify protein overexpression.

One SDS-PAGE sample from each triplicate was electrophoresed on a 10% (w/v) SDS-PAGE gel at 120 V for 1 hour and 10 minutes alongside Color Prestained Protein Standard (CPPS), Broad Range (10-250 kDa) (NEB). CPPS molecular masses in kDa are shown to the left of the gel. Lane numbers are shown above the gel. The proteins overexpressed are shown below the corresponding lane. Expected overexpressed protein positions shown to the right of the gel. NmVenus-POLD3 Quincunx motif mutants are at the same position as NmVenus-POLD3.

A spot assay to determine sample cell viability was initially performed five days after cell samples were resuspended in 1X M9 medium using 10^{-1} - 10^{-8} serial dilutions, but only samples with plasmids encoding NmVenus + CmVenus grew. A spot assay was performed again two days later using 10^{-1} - 10^{-5} serial dilutions, with additional 10^{-6} - 10^{-8} serial dilutions for samples with plasmids encoding NmVenus + CmVenus (Figure 3.9). Corresponding sample average CFU/mL are listed in Table 3.3.

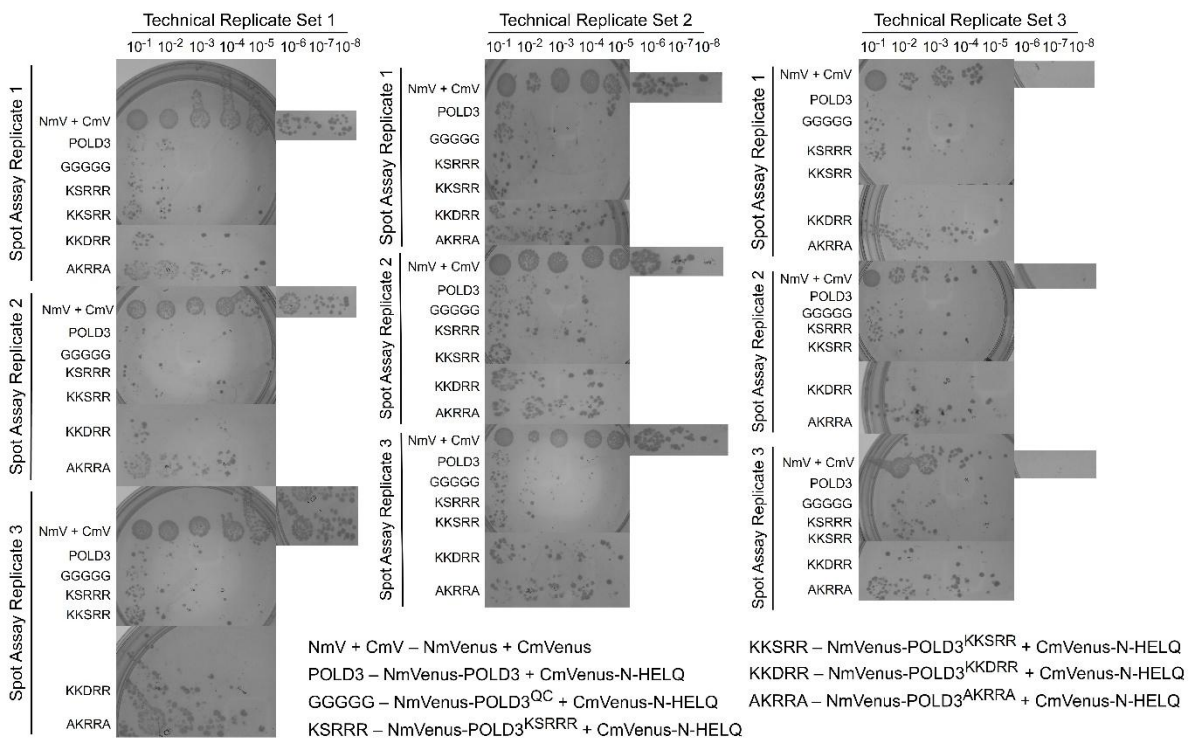


Figure 3.9 Spot assay of POLD3 Quincunx motif mutants BiFC assay BL21-AI *E. coli* cell samples to assess cell viability.

For each 1 mL sample replicate set, 10 μ L of each sample was diluted in 90 μ L of 1X M9 base salts for a 10^{-1} dilution, and this was serially diluted to 10^{-5} with an additional 10^{-6} - 10^{-8} dilutions for cells expressing NmVenus + CmVenus, using 10 μ L of each previous dilution. 10 μ L of each serial dilution was spotted in triplicate on a Mu agar plate containing 100 μ g/mL ampicillin + 40 μ g/mL kanamycin. The proteins that were overexpressed in each sample are abbreviated to the left of the corresponding row, with the key at the bottom.

Table 3.3 Average CFU/mL from POLD3 Quincunx motif mutants BiFC assay spot assay colony counts.

Proteins Overexpressed in BL21-AI <i>E. coli</i>	Technical Replicate	CFU/mL
NmVenus + CmVenus	1	1.7×10^{11}
	2	1.5×10^{10}
	3	2×10^7
NmVenus-POLD3 + CmVenus-N-HELQ	1	2×10^6
	2	4×10^6
	3	6×10^3
NmVenus-POLD3 ^{QC} + CmVenus-N-HELQ	1	1×10^4
	2	2×10^6
	3	2×10^7
NmVenus-POLD3 ^{KSRRR} + CmVenus-N-HELQ	1	3×10^5
	2	4×10^6
	3	3×10^7
NmVenus-POLD3 ^{KKSRR} + CmVenus-N-HELQ	1	2×10^7
	2	2×10^6
	3	3×10^3
NmVenus-POLD3 ^{KKDRR} + CmVenus-N-HELQ	1	3×10^5
	2	6×10^7
	3	4×10^7
NmVenus-POLD3 ^{AKRRA} + CmVenus-N-HELQ	1	3×10^7
	2	2.1×10^7

	3	5×10^7
--	---	-----------------

Colonies were counted using an SC6+ Colony Counter (Stuart) and rounded to the nearest whole number. Colony count averages used to calculate CFU/mL were based on highest dilution factor among a sample's triplicates that had at least 2 valid spots, otherwise the highest dilution factor with at least one valid spot was used.

Amersham Typhoon scans are shown in Figure 3.10. The 1X M9 minimal medium containing 100 µg/mL ampicillin + 40 µg/mL kanamycin showed negligible autofluorescence. Minimal fluorescence was observed in the NmVenus + CmVenus negative control. The NmVenus-POLD3 + CmVenus-N-HELQ positive control and all of the NmVenus-POLD3 Quincunx motif mutant + CmVenus-N-HELQ test pairs showed fluorescence, thus indicating PPI and that none of the POLD3 Quincunx motif mutations inhibited the interaction. Compared to the NmVenus-POLD3 + CmVenus-N-HELQ positive control technical triplicates, NmVenus-POLD3^{KKSRR} + CmVenus-N-HELQ and NmVenus-POLD3^{KKDRR} + CmVenus-N-HELQ both showed a consistently lower fluorescence, while NmVenus-POLD3^{AKRRA} + CmVenus-N-HELQ showed a consistently higher fluorescence.

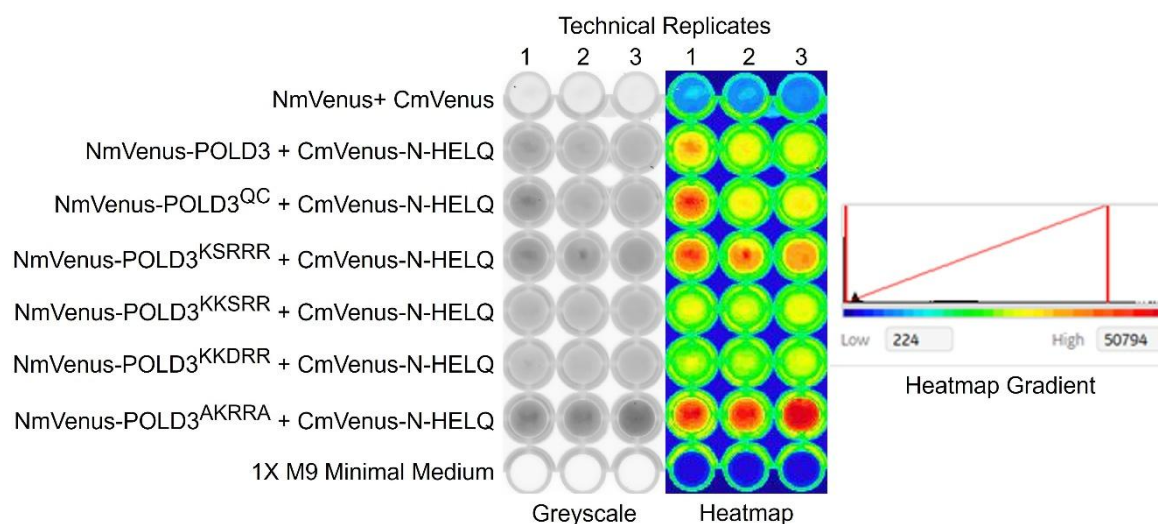


Figure 3.10 Amersham Typhoon 5 Biomolecular Imager (Cytiva) scan of POLD3 Quincunx motif mutants BiFC assay BL21-AI *E. coli* cell sample fluorescence to assess protein-protein interaction (PPI) between protein pairs.

Cells were suspended in 1X M9 minimal medium containing 100 µg/mL ampicillin + 40 µg/mL kanamycin. 400 µL of each sample was pelleted, 250 µL was removed, and cells were resuspended in remaining 150 µL, which was transferred to a clear flat-bottom non-tissue cultured 96-well plate. The plate was scanned using the built-in Cy2 filter. The scan was saved as both greyscale and heatmap, indicated below the scans. Overexpressed protein pairs are indicated to the left of the scans. Technical replicate set number is given above the scans. Heatmap gradient of colours corresponding to fluorescence intensity is shown on the right of the scans. Darker colour in greyscale indicates greater fluorescence intensity and greater fluorescence intensity indicates greater PPI.

The FLUOstar Omega fluorescence scan used a gain of 2048 and was normalised against OD₆₀₀ and CFU/mL; outliers of greater than 1 standard deviation (SD) were also removed from the CFU/mL normalisation (Figure 3.11). Two-tailed Welch's t-tests were performed using GraphPad Prism 10 to determine if sample triplicate mean fluorescence showed a significant difference compared to NmVenus + CmVenus, NmVenus-POLD3 + CmVenus-N-HELQ, and NmVenus-POLD3^{QC} + CmVenus-N-HELQ triplicate mean fluorescence. Full t-test statistics are listed in Appendix Table 7.2. When normalised by OD₆₀₀ and excluding repeat analyses,

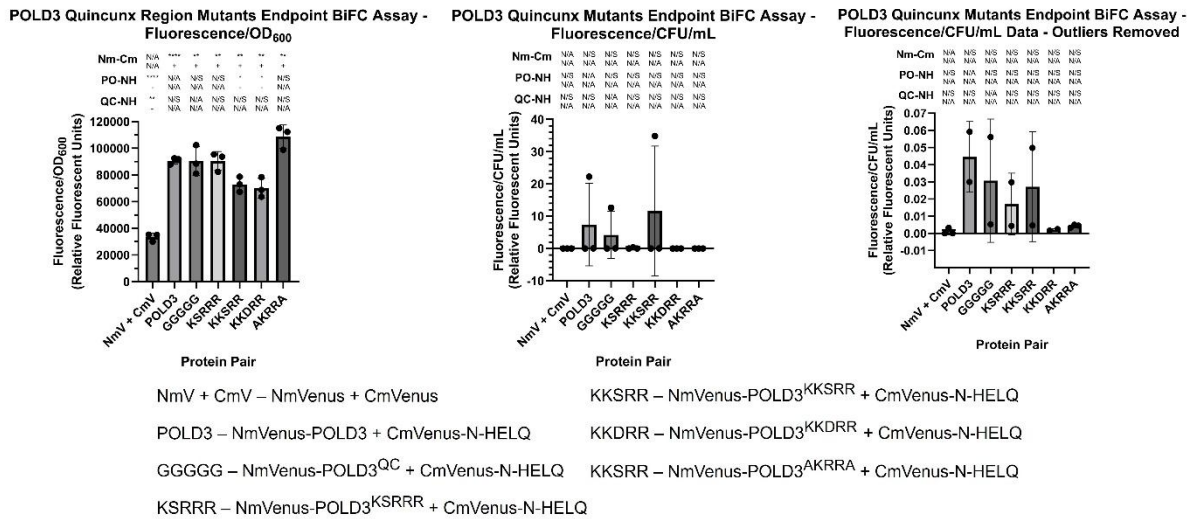


Figure 3.11 Normalised fluorescence values from FLUOstar Omega (BMG LABTECH) scan of POLD3 Quincunx motif mutants BiFC assay BL21-AI *E. coli* cell samples.

A gain of 2048 was used. Fluorescence was normalised by OD₆₀₀ (left) and colony forming units (CFU)/mL (middle). For fluorescence normalised by CFU/mL, results beyond 1 standard deviation were considered outliers and so removed (right). The column heights show mean and the error bars show $\pm$ one standard deviation. Two-tailed Welch's t-tests were performed using GraphPad Prism 10 to determine if triplicate means showed a significant difference compared to NmVenus + CmVenus (Nm-Cm), NmVenus-POLD3 + CmVenus-N-HELQ (PO-NH), and NmVenus-POLD3^{QC} + CmVenus-N-HELQ (QC-NH) fluorescence. Significance level is denoted by asterisks over the column: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$, N/S = 'no significance', and N/A = 'not applicable'. + indicates that that column protein pair has a higher mean than the comparative protein pair; - indicates it is lower.

compared to NmVenus + CmVenus fluorescence ($n = 3$, $M = 33441$, $SD = 2944$), all results were significantly higher, with NmVenus-POLD3 + CmVenus-N-HELQ ($n = 3$, $M = 90688$, $SD = 2464$) showing the largest degree of difference ($t(3.880) = 25.83$, $p < 0.0001$) and all NmVenus-POLD3 Quincunx motif mutant + CmVenus-N-HELQ pairs showing the same degree of difference ($p < 0.01$). Compared to NmVenus-POLD3 + CmVenus-N-HELQ fluorescence, NmVenus-POLD3^{KSRRR} + CmVenus-N-HELQ fluorescence ($n = 3$, $M = 72993$, $SD = 5616$) and

NmVenus-POLD3^{KKDRR} + CmVenus-N-HELQ fluorescence (n = 3, M = 70252, SD = 7475) results were significantly lower ($t(2.743) = 4.997$, $p = 0.0190$; $t(2.430) = 4.497$, $p = 0.0317$) while all other results showed no significant difference. Compared to NmVenus-POLD3^{QC} + CmVenus-N-HELQ fluorescence (n = 3, M = 90576, SD = 10813), no results showed a significant difference. Furthermore, the pattern of the protein pair triplicates' relative fluorescence levels matches that observed in the Amersham Typhoon scans. When normalised by CFU/mL, even with outliers removed, no result showed a significant difference.

3.3 Analysis of POLD3 Intrinsically Disordered Protein Regions and Putative Nuclear Localisation Signal Interaction With N-HELQ

3.3.1 BiFC Assay

Since the NmVenus-POLD3 Quincunx motif mutant + CmVenus-N-HELQ pairs still showed fluorescence indicating PPI, it was decided to investigate whether mutating a 381-RKRKR-385 motif in the PNLS to glycines (381-GGGGG-385) would affect POLD3-N-HELQ PPI (Alli et al., 2024). Like the mutation in pBAD-NmVenus-PoID3^{QC}, this would mutate a polar, charged residue tract to polar, uncharged glycines (JPT, 2025a; JPT, 2025b; Satoh et al., 2022). The NmVenus-POLD3 mutant for this was encoded for by Olivia Downs' plasmid pOLD042 (pBAD-NmVenus-PoID3^{PNLS}), which was generated by SDM like the POLD3 Quincunx motif mutant plasmids (see Section 2.4.7). Meanwhile, fellow Bolt lab MRes student Emily Haskins used the He et al. (2023) time-course BiFC assay method to show preliminary data indicating that POLD3 IDPR1 does not interact with N-HELQ, while IDPRs 2, 3, and 4 do, with IDPR3 showing the strongest PPI. IDPR3 contains both the confirmed NLS motif (residues 319-KKRRR-323) and the PNLS motif (381-RKRKR-385) (see Figure 3.2). Therefore, this next investigation stage was expanded to include the IDPRs to attempt to reproduce the result. The plasmids encoding NmVenus-POLD3 IDPR fusion proteins were generated by Downs using the method described in He et al. (2023), via PCR amplifying plasmid DNA encoding POLD3 IDPRs (see Section 2.4.5) and performing restriction cloning into pBAD-NmVenus via XhoI and HindIII restriction sites (see Section 2.4.8). The plasmids are detailed in Table 3.3.

Table 3.3 Details of pBAD-NmVenus-PoID3 IDPR plasmids generated by restriction cloning.

Plasmid	POLD3 IDPR	POLD Amino Acid Residues
pOLD001 (pBAD-NmVenus-PoID3 ^{IDPR1})	IDPR1	169-188
pOLD002 (pBAD-NmVenus-PoID3 ^{IDPR2})	IDPR2	199-232
pOLD003 (pBAD-NmVenus-PoID3 ^{IDPR3})	IDPR3	274-393
pOLD004 (pBAD-NmVenus-PoID3 ^{IDPR4})	IDPR4	406-466

The BiFC assay was carried out as described in Section 2.6.1, using the following cotransformations: pBAD-NmVenus + pRSF-CmVenus, pBAD-NmVenus-PoID3 + pRSF-CmVenus-NHelQ, pBAD-NmVenus-PoID3^{PNLS} + pRSF-CmVenus-NHelQ, pBAD-NmVenus-PoID3^{IDPR1} + pRSF-CmVenus-NHelQ, pBAD-NmVenus-PoID3^{IDPR2} + pRSF-CmVenus-NHelQ, pBAD-NmVenus-PoID3^{IDPR3} + pRSF-CmVenus-NHelQ, and pBAD-NmVenus-PoID3^{IDPR4} + pRSF-CmVenus-NHelQ. For the first three cotransformations, glycerol stocks produced from overnight cultures for the POLD3 Quincunx motif mutants BiFC assay were used rather than new cotransformations. This time, SDS-PAGE analysis used all samples (Figure 3.12). The samples were electrophoresed on 10% (w/v) SDS-PAGE gels at 120 V for 1 hour and 10 minutes. NmVenus, NmVenus-POLD3, and NmVenus-POLD3 Quincunx motif mutants were visible in the appropriate lanes, but CmVenus was nonvisible, and CmVenus-N-HELQ was nonvisible when co-overexpressed with a protein the size of NmVenus-POLD3, in this case NmVenus-POLD3^{PNLS}. A Western blot could not be performed for the nonvisible proteins due to time constraints.

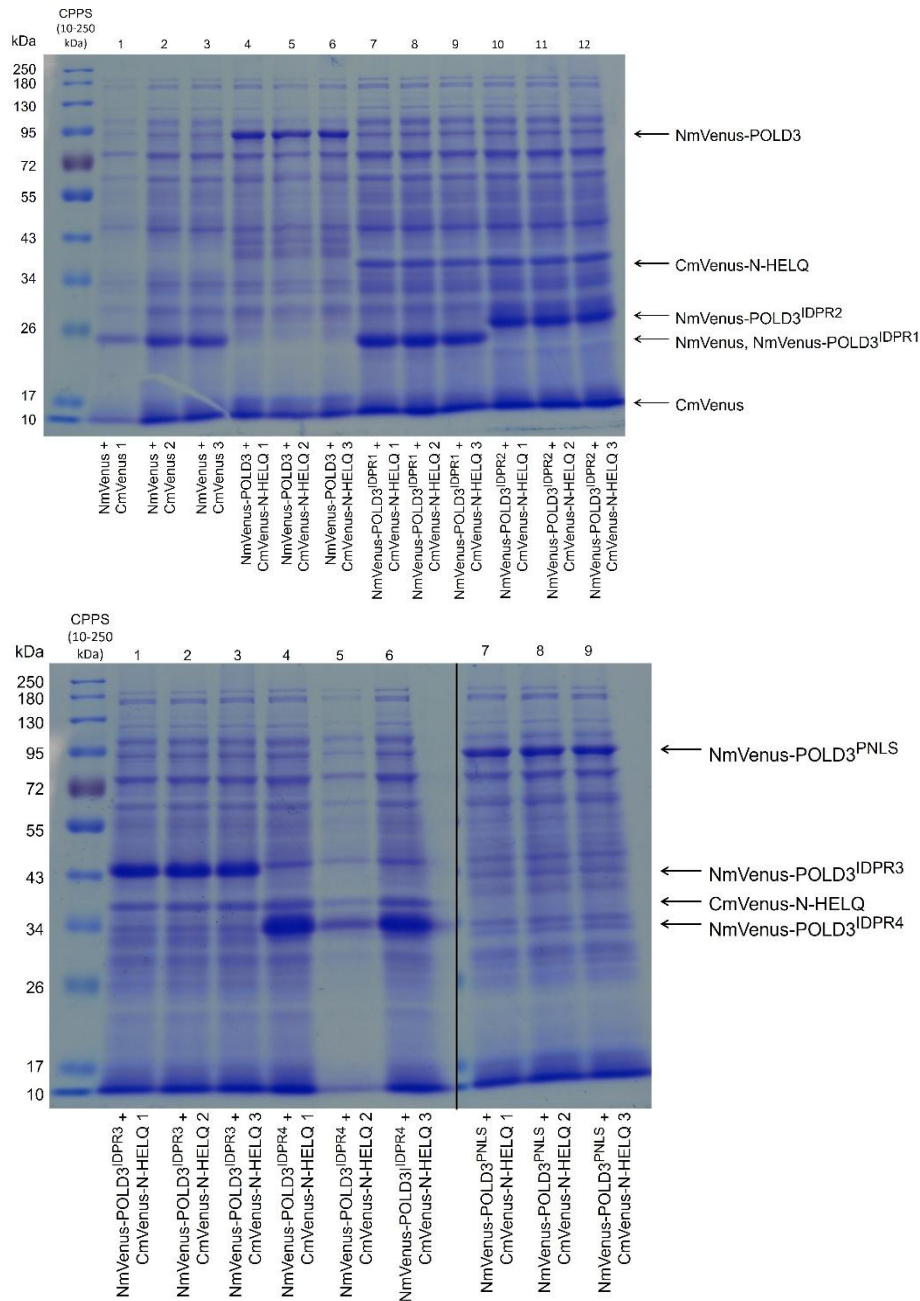


Figure 3.12 SDS-PAGE gels of POLD3 IDPRs and putative NLS BiFC assay SDS-PAGE samples to verify protein overexpression.

SDS-PAGE samples were electrophoresed on 10% (w/v) SDS-PAGE gels at 120 V for 1 hour and 10 minutes alongside Color Prestained Protein Standard (CPPS), Broad Range (10-250 kDa) (NEB). CPPS molecular masses in kDa are shown to the left of the gels. Lane numbers are shown above the gels. The proteins overexpressed are shown below the corresponding lane. Expected overexpressed protein positions shown to the right of gel.

A spot assay was performed eight days after cell samples were resuspended in 1X M9 medium, using 10^{-1} - 10^{-10} serial dilutions (Figure 3.13). No growth was observed from any of the cell samples with plasmids encoding NmVenus + CmVenus, NmVenus-POLD3 + CmVenus-N-HELQ, and NmVenus-PoID3^{PNLS} + CmVenus-HELQ. The average CFU/mL from samples that showed growth are listed in Table 3.4.

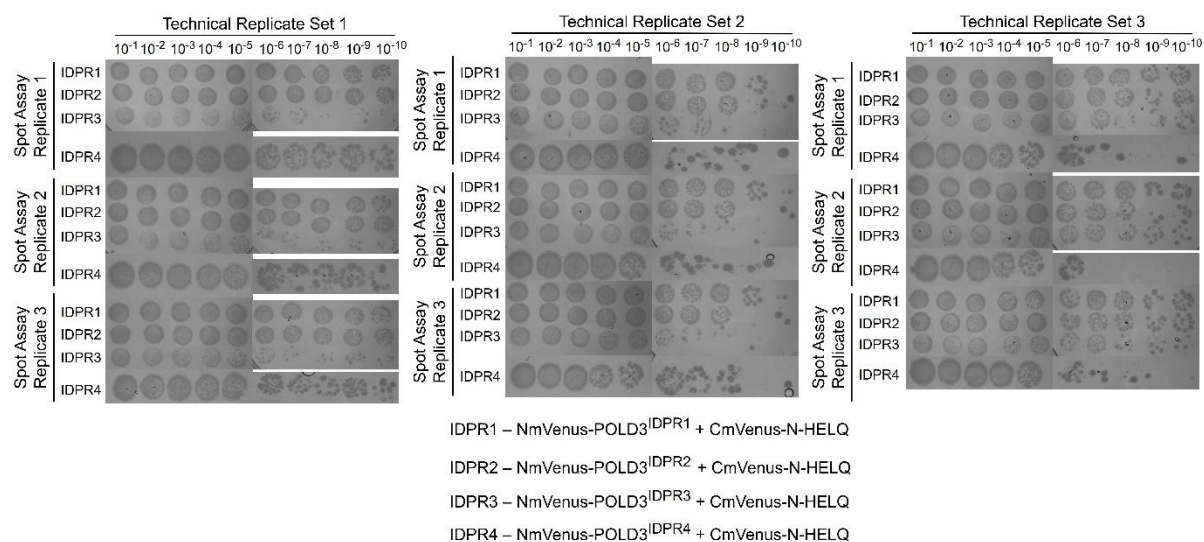


Figure 3.13 Spot assay of POLD3 IDPRs and putative NLS BiFC assay BL21-AI *E. coli* cell samples to assess cell viability.

For each 1 mL sample replicate set, 10 μ L of each sample was diluted in 90 μ L of 1X M9 base salts for 10^{-1} dilution, and this was serially diluted, using 10 μ L of each previous dilution. 10 μ L of each serial dilution was spotted in triplicate on a Mu agar plate containing 100 μ g/mL ampicillin + 40 μ g/mL kanamycin. The proteins that were overexpressed in each sample are abbreviated to the left of the corresponding row, with the key at the bottom

Table 3.4 Average CFU/mL from POLD3 IRPDRs and putative NLS BiFC assay spot assay colony counts.

Proteins Overexpressed in BL21-AI <i>E. coli</i>	Technical Replicate	CFU/mL
NmVenus-POLD3 ^{IDPR1} + CmVenus-N-HELQ	1	3.5×10^{13}
	2	3×10^{12}
	3	3×10^{13}

NmVenus-POLD3 ^{IDPR2} + CmVenus-N-HELQ	1	5.7x10 ¹³
	2	2x10 ¹²
	3	2x10 ¹³
NmVenus-POLD3 ^{IDPR3} + CmVenus-N-HELQ	1	6x10 ¹²
	2	2x10 ¹¹
	3	1.4x10 ¹³
NmVenus-POLD3 ^{IDPR4} + CmVenus-N-HELQ	1	1.6x10 ¹³
	2	1x10 ¹²
	3	2x10 ¹²

Colonies were counted using an SC6+ Colony Counter (Stuart) and rounded to the nearest whole number. Colony count averages used to calculate CFU/mL were based on highest dilution factor among a sample's triplicates that had at least 2 valid spots, otherwise the highest dilution factor with at least one valid spot was used.

Amersham Typhoon scans are shown in Figure 3.14. The 1X M9 minimal medium containing 100 µg/mL ampicillin + 40 µg/mL kanamycin showed negligible autofluorescence. Minimal fluorescence was observed in the NmVenus + CmVenus negative control and NmVenus-POLD3^{IDPR1} + CmVenus-N-HELQ. The NmVenus-POLD3 + CmVenus-N-HELQ positive control, NmVenus-POLD3^{IDPR2} + CmVenus-N-HELQ, NmVenus-POLD3^{IDPR3} + CmVenus-N-HELQ, and NmVenus-POLD3^{IDPR4} + CmVenus-N-HELQ showed fluorescence, thus indicating PPI, with NmVenus-POLD3^{IDPR3} + CmVenus-N-HELQ consistently showing the highest fluorescence across its technical triplicates. This pattern, excluding NmVenus-POLD3^{PNLS} + CmVenus-HELQ, reflects the fluorescence levels observed at the final timepoint in the Amersham Typhoon scan in Haskins' preliminary data.

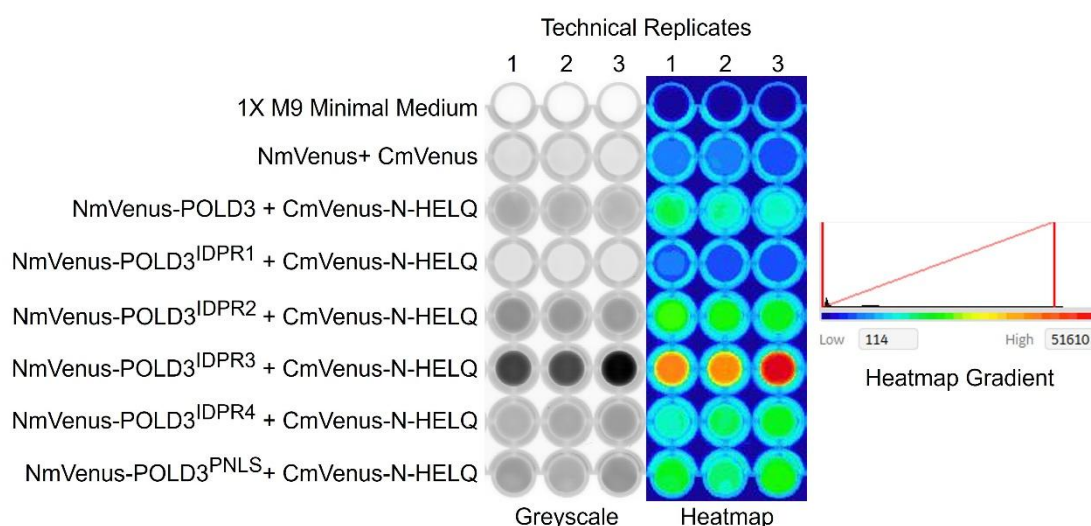


Figure 3.14 Amersham Typhoon 5 Biomolecular Imager (Cytiva) scan of POLD3 IDPRs and putative NLS BiFC assay BL21-AI *E. coli* cell sample fluorescence to assess protein-protein interaction (PPI) between protein pairs.

Cells were suspended in 1X M9 minimal medium containing 100 µg/mL ampicillin + 40 µg/mL kanamycin. 400 µL of each sample was pelleted, 250 µL was removed, and cells were resuspended in remaining 150 µL, which was transferred to a clear flat-bottom non-tissue cultured 96-well plate. The plate was scanned using the built-in Cy2 filter. The scan was saved as both greyscale and heatmap, indicated below the scans. Overexpressed protein pairs are indicated to the left of the scans. Technical replicate set number is given above the scans Heatmap gradient of colours corresponding to fluorescence intensity is shown on the right of the scans. Darker colour in greyscale indicates greater fluorescence intensity and greater fluorescence intensity indicates greater PPI.

The FLUOstar Omega fluorescence scan used a gain of 1829. Due to the lack of colonies for samples that used cells from glycerol stocks rather than cotransformations, cells were normalised using OD₆₀₀ but not CFU/mL (Figure 3.15). Two-tailed Welch's t-tests were performed using GraphPad Prism 10 to determine if sample triplicate mean fluorescence showed a significant different compared to NmVenus + CmVenus and NmVenus-POLD3 + CmVenus-N-HELQ triplicate mean fluorescence. Full t-test statistics are listed in Appendix Table 7.2. NmVenus + CmVenus fluorescence (n = 3, M = 29969, SD = 9449) compared to

**POLD3 IDPRs and PNLS Endpoint BiFC Assay -
Fluorescence/OD₆₀₀**

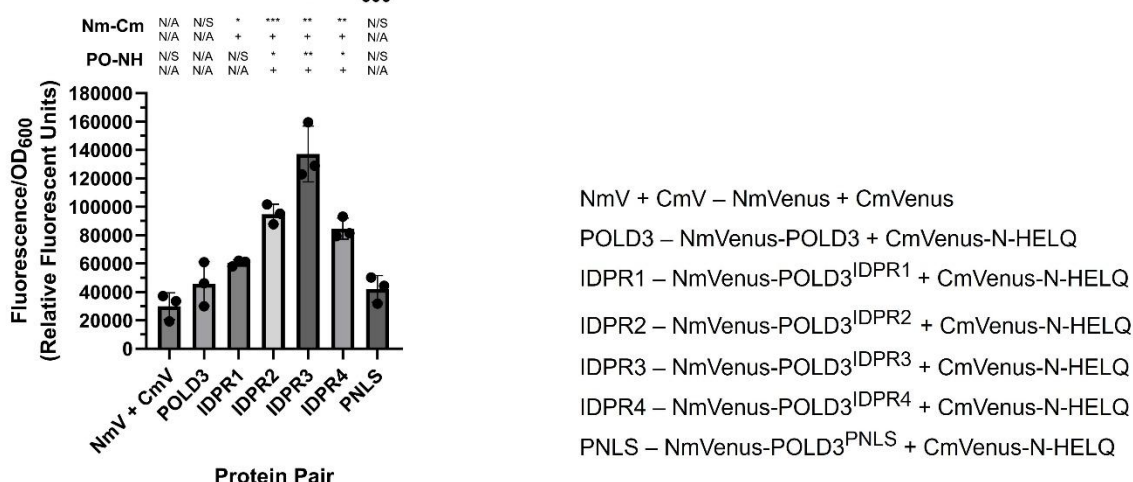


Figure 3.15 Normalised fluorescence values from FLUOstar Omega (BMG LABTECH) scan of POLD3 IDPRs and putative NLS BiFC assay BL21-AI *E. coli* cell samples.

A gain of 1829 was used. Fluorescence was normalised by OD₆₀₀. The column heights show mean and the error bars show $\pm$ one standard deviation. Two-tailed Welch's t-tests were performed using GraphPad Prism 10 to determine if triplicate means showed a significant difference compared to NmVenus + CmVenus (Nm-Cm) and NmVenus-POLD3 + CmVenus-N-HELQ (PO-NH). Significance level is denoted by asterisks over the column: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$, N/S = 'no significance', and N/A = 'not applicable'. + indicates that that column protein pair has a higher mean than the comparative protein pair; - indicates it is lower.

NmVenus-POLD3 + CmVenus-N-HELQ fluorescence ($n = 3$, $M = 45688$, $SD = 15566$) showed no significant difference ($t(3.298) = 1.495$, $p = 0.2237$). NmVenus-POLD3^{PNLS} + CmVenus-N-HELQ fluorescence ($n = 3$, $M = 42106$, $SD = 9438$) also showed no significant difference to either of them ($t(4.000) = 1.574$, $p = 0.1906$; $t(3.295) = 0.3409$, $p = 0.7538$). NmVenus-POLD3^{IDPR1} + CmVenus-N-HELQ fluorescence ($n = 3$, $M = 60450$, $SD = 2144$) was significantly higher than NmVenus + CmVenus fluorescence ($t(2.205) = 5.449$, $p = 0.0258$) but not NmVenus-POLD3 + CmVenus-N-HELQ fluorescence ($t(2.076) = 1.627$, $p = 0.2407$). To varying degrees, NmVenus-POLD3^{IDPR2} + CmVenus-N-HELQ fluorescence ($n = 3$, $M = 94799$, $SD = 7013$), NmVenus-

POLD3^{IDPR3} + CmVenus-N-HELQ fluorescence (n = 3, M = 137137, SD = 19670), and NmVenus-POLD3^{IDP4} + CmVenus-N-HELQ fluorescence (n = 3, M = 84541, and SD = 7416) was significantly higher than NmVenus + CmVenus ($t(3.691) = 9.543$, $p = 0.0010$; $t(2.876) = 8.506$, $p = 0.0040$; $t(3.786) = 7.869$, $p = 0.0018$) and to a lesser extent NmVenus-POLD3 + CmVenus-N-HELQ fluorescence ($t(2.780) = 4.982$, $p = 0.0186$; $t(3.799) = 6.315$, $p = 0.0038$; and $t(2.863) = 3.903$, $p = 0.0325$). NmVenus-POLD3^{IDPR3} + CmVenus-N-HELQ had the highest mean fluorescence. Like in the case of the POLD3 Quincunx motif mutants BiFC assay, the pattern of the protein pair triplicates' relative fluorescence levels matches that observed in the Amersham Typhoon scans.

3.3.2 Test For Viable But Non-Culturable State

Due to the spot assay showing no growth from cell samples that were expressing NmVenus + CmVenus, NmVenus-POLD3 + CmVenus-NHELQ, and NmVenus-POLD3^{PNLS} + CmVenus-NHELQ, it was hypothesised that these cells were in the viable but non-culturable (VBNC) state rather than dead. First observed by Xu et al., (1982), many bacterial species can enter a VBNC dormant survival state when exposed to various stressful environmental conditions and so remain viable but do not grow on routine growth media, remaining undetectable in spot assays (Pienaar et al., 2016; Ramamurthy et al., 2014; Yang, D. et al., 2024). The VBNC state has been observed in several pathogenic and non-pathogenic *E. coli* strains (Pienaar et al., 2016). In this case, it could be induced by the stress of being stored at a suboptimum temperature of 4 °C in nutrient-poor 1X M9 minimal medium, exacerbated by pre-stress caused by the cells being cultured from a glycerol stock compared to the other cell samples that were cultured from newly cotransformed cells (Dağ et al., 2025; Sleight et al., 2006; Thean et al., 2021; Wei et al., 2018). VBNC *E. coli* can be resuscitated back to a culturable state (Yang, D. et al., 2024). Wei et al. (2018) showed that O157:H7 *E. coli* could be resuscitated by incubation in LB medium containing 5% Tween 80 at 37 °C with shaking, with a higher number of colonies observed after plating at 48 hours of incubation compared to 24 hours. Therefore, this was used as the basis for a method to test if the cells were VBNC and so could be resuscitated to grow in a spot assay.

The test was carried out as described in Section 2.6.2. The pBAD-NmVenus + pRSF-CmVenus and pBAD-NmVenus-POLD3 + pRSF-CmVenus-NHelQ cotransformation overnight culture-inoculated duplicate samples functioned as positive controls to ensure cells could grow in Mu broth containing 100 µg/mL ampicillin and 40 µg/mL kanamycin with or without 5% (v/v) Tween 80. The 1 mL respective medium duplicates function as negative controls in addition to one of each serving as a blank for the samples in that respective medium. The growth curves for these samples and 1 mL respective medium OD₆₀₀ data are illustrated in Figure 3.16. No growth was observed in the negative control uninoculated media, while all positive control samples generated growth curves. This confirmed that BL21-AI *E. coli* cotransformed with plasmids encoding either only the non-fluorescent fragments or fusion proteins would be able to grow in Mu broth containing 100 µg/mL ampicillin and 40 µg/mL kanamycin with or without 5% (v/v) Tween 80.

The growth curves for the samples that were inoculated using the BiFC assay cell samples are shown in Figure 3.17. The NmVenus+POLD3^{PNLS} + CmVenus-N-HELQ sample 3 in Mu broth with 5% (v/v) Tween 80 replicate 2 was characteristic of a stationary phase, followed by a death phase, followed by a growth advantage in stationary phase phenotype (Chib et al., 2017; Gonzalez et al., 2023; Zambrano et al., 1993). The remaining VBNC test samples that showed growth curves all exhibited a resuscitation lag phase followed by a resuscitation phase characteristic of growth from VBNC cells (Yang, D. et al., 2024).

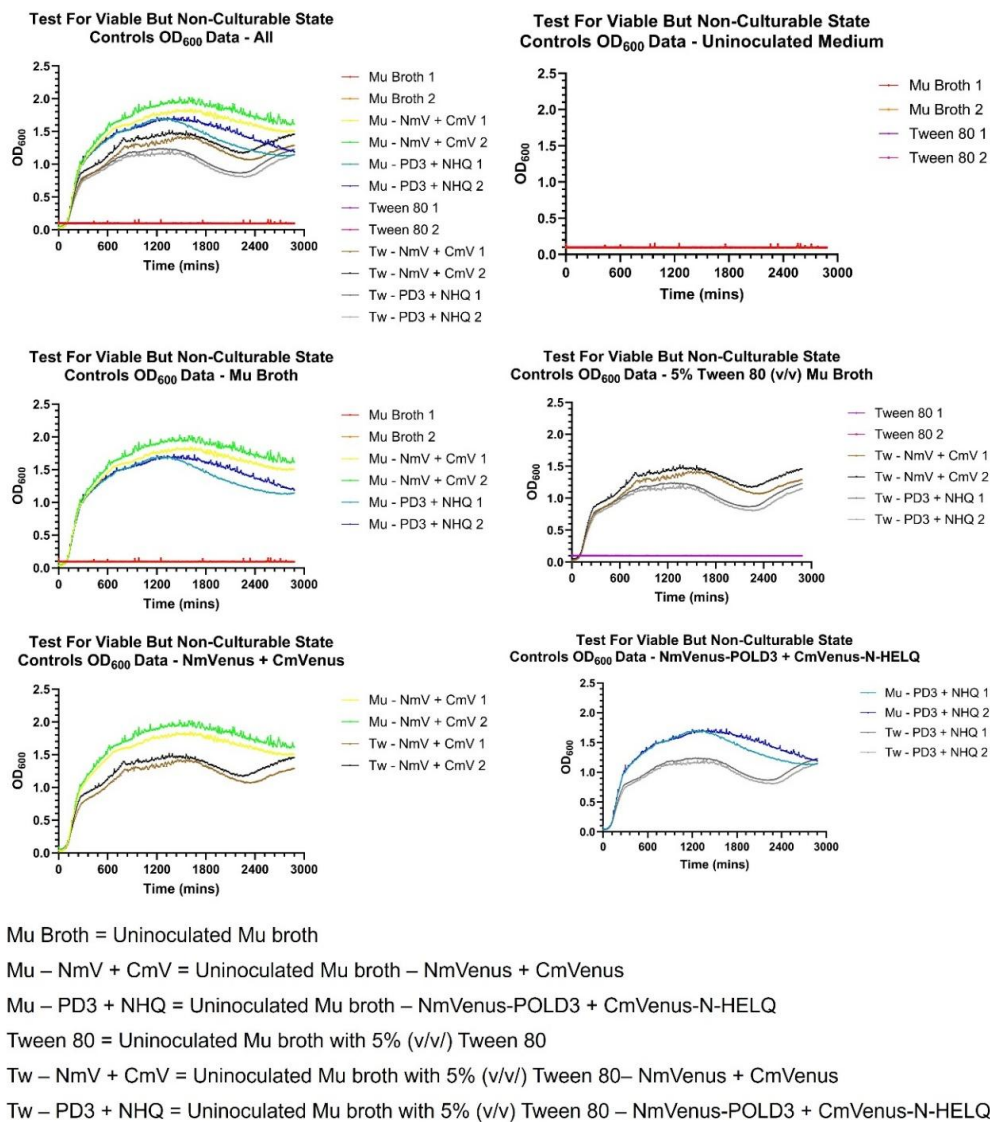


Figure 3.16 OD₆₀₀ data of control samples in POLD3 IDPRs BiFC assay test for BL21-AI *E. coli* viable but non-culturable state.

The test used Mu broth containing 100 µg/mL ampicillin and 40 µg/mL kanamycin with or without 5% (v/v) Tween 80. Each had 1 mL uninoculated medium as a negative control, and the 1 mL positive controls used cells inoculated at a 1:100 dilution from an overnight culture of cells expressing either NmVenus + CmVenus or NmVenus+POLD3 + CmVenus-N-HELQ. Each control was performed in duplicate. OD₆₀₀ was measured using a FLUOstar Omega (BMG LABTECH) plate reader. The control corresponding to each colour and replicate number is abbreviated next to the respective graphs, with the abbreviations listed in full below the graphs.

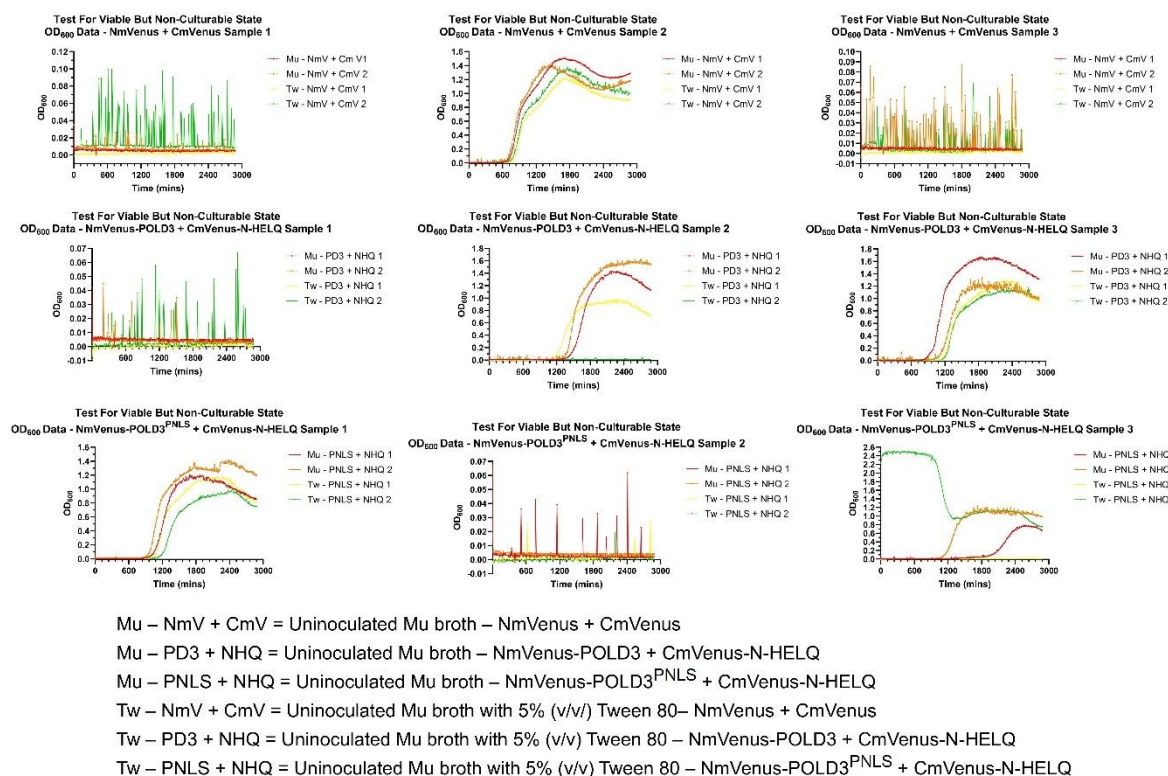


Figure 3.17 OD₆₀₀ data of test cell samples from POLD3 IDPRs BiFC assay test for BL21-AI *E.coli* viable but non-culturable (VBNC) state.

Each POLD3 IDPRS BiFC assay sample of cells expressing NmVenus + CmVenus, NmVenus+POLD3 + CmVenus-N-HELQ, or NmVenus+POLD3^{PNLS} + CmVenus-N-HELQ was inoculated for a 1:100 dilution in duplicate in Mu broth containing 100 µg/mL ampicillin and 40 µg/mL kanamycin with and without 5% (v/v) Tween 80. OD₆₀₀ was measured using a FLUOstar Omega (BMG LABTECH) plate reader. The VBNC test sample corresponding to each colour is abbreviated with replicate numbers next to the respective graphs, with the abbreviations listed in full below the graphs.

A spot assay of each of the VBNC test samples of cells expressing either NmVenus + CmVenus, NmVenus+POLD3 + CmVenus-N-HELQ, or NmVenus-POLD3^{PNLS} + CmVenus-N-HELQ, minus positive controls, is shown in Figure 3.18. The NmVenus+POLD3^{PNLS} + CmVenus-N-HELQ sample 3 in Mu broth without Tween 80 replicate 1 sample showed no colony growth despite generating a growth curve, while the NmVenus+POLD3 + CmVenus-N-HELQ sample 2 in Mu broth with 5% (v/v) Tween 80 replicate 2 sample showed colony growth despite not generating a growth curve. The remaining samples that generated growth curves all showed colony growth. Sample CFU/mL are shown in Table 3.5. Ultimately, NmVenus + CmVenus sample 2 replicates 1 and 2, NmVenus+POLD3 + CmVenus-N-HELQ sample 2 replicates 1 and 2 and sample 3 replicates 1 and 2, NmVenus-POLD3^{PNLS} + CmVenus-N-HELQ sample 1 replicates 1 and 2, and sample 3 replicate 2 showed colony growth. Five replicates had higher CFU/mL from Mu broth without Tween 80, two of which had the same log power as the CFU/mL from Mu broth with 5% (v/v) Tween 80, and four replicates had higher CFU/mL in Mu broth with 5% (v/v) Tween 80. Of these, all NmVenus + CmVenus replicates that showed colony growth had higher CFU/mL in Mu broth without 5% (v/v) Tween 80, while this was split evenly between Mu broth with and without 5% (v/v) Tween 80 for NmVenus+POLD3 + CmVenus-N-HELQ, and mostly Mu broth without 5% (v/v) Tween 80 for NmVenus+POLD3^{PNLS} + CmVenus-N-HELQ.

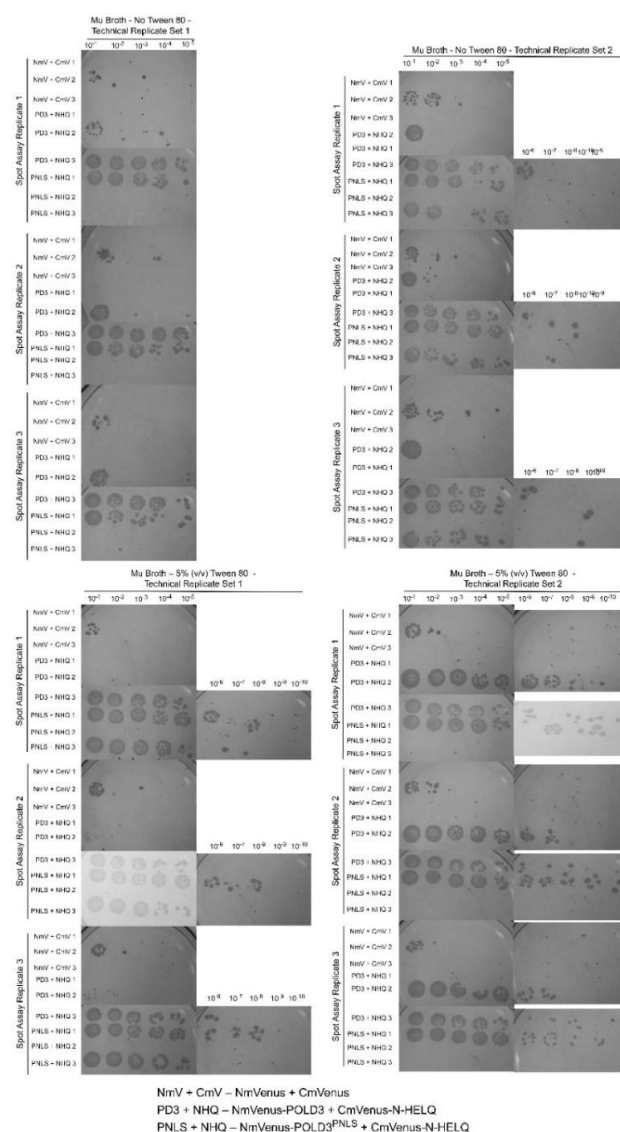


Figure 3.18 Spot assay from test for viable but non-culturable state of POLD3 IDPRs BiFC assay BL21-AI *E. coli* cell samples to verify resuscitation.

For each 1 mL cell sample of cells expressing either NmVenus + CmVenus, NmVenus+POLD3 + CmVenus-N-HELQ, or NmVenus-POLD3^{PNLS} + CmVenus-N-HELQ, minus positive controls, 10 μ L was diluted in 90 μ L of 1X M9 base salts for 10⁻¹ dilution, and this was serially diluted to 10⁻¹⁰, using 10 μ L of each previous dilution. 10 μ L of each serial dilution was spotted in triplicate on a Mu agar plate containing 100 μ g/mL ampicillin + 40 μ g/mL kanamycin. The proteins that were overexpressed in each sample are abbreviated to the left of the corresponding row, with the key at the bottom. 10⁻⁶-10⁻¹⁰ plates that showed no growth are omitted.

Table 3.5 Details of cell samples CFU/mL from spot assay colony counts in test for viable but non-culturable state of POLD3 IDPRs BiFC assay BL21-AI *E. coli* cell samples.

Proteins Coexpressed	Sample Number	Replicate	Mu Broth Without Tween 80	CFU/mL	Mu Broth With 5% (v/v) Tween 80	CFU/mL	Mu or Tw
NmVenus + CmVenus	1	1	-	N/A	-	N/A	N/A
		2	-	N/A	-	N/A	N/A
	2	1	+	2×10^6	+	4×10^4	Mu
		2	+	2×10^6	+	1×10^4	Mu
	3	1	-	N/A	-	N/A	N/A
		2	-	N/A	-	N/A	N/A
NmVenus-POLD3 + CmVenus-N-HELQ	1	1	-	N/A	-	N/A	N/A
		2	-	N/A	-	N/A	N/A
	2	1	+	3×10^7	+	3×10^{10}	Tw
		2	+	9.1×10^4	+	4×10^4	Mu*
	3	1	+	1.3×10^8	+	1×10^{12}	Tw
		2	+	9×10^7	+	8×10^7	Mu
NmVenus-POLD3 ^{PNLS} + CmVenus-N-HELQ	1	1	+	4×10^7	+	2×10^{11}	Tw
		2	+	1×10^9	+	1×10^{11}	Tw
	2	1	-	N/A	-	N/A	N/A
		2	-	N/A	-	N/A	N/A
	3	1	-	N/A	-	N/A	N/A

		2	+	7×10^7	-	5×10^7	Mu*
--	--	---	---	-----------------	---	-----------------	-----

A - indicates that there were not colonies observed for that replicate in the spot assay, and + indicates that there were colonies observed. Colonies were counted using an SC6+ Colony Counter (Stuart) and rounded to the nearest whole number. Colony count averages used to calculate CFU/mL were based on highest dilution factor among a sample's triplicates that had at least 2 valid spots, otherwise the highest dilution factor with at least one valid spot was used. 'Mu or Tw' column lists whether the replicate in Mu broth with ('Tw') or without ('Mu') 5% (v/v) Tween 80 for a respective sample has the higher CFU/mL, and * indicates that both have the same log power.

3.4 Analysis of BL21-AI *E. coli* Growth Media Autofluorescence and Effect on Growth Rate

Following the POLD3 IDPRs and PNLS BiFC assay, it was decided to attempt to develop the method further by progressing from an endpoint method to a time-course method. This is because an endpoint method only provides a single temporal snapshot, while a time-course method records data in a continuous temporal manner, allowing for observation of BiFC assay FP reconstitution and resultant fluorescence intensity change over time rather than at a single timepoint (Gujral et al., 2020; Kerppola, 2009; Weinreb et al., 2018). This would hence combine the respective time-course and quantitative strengths of the original He et al. (2023) method and the modified Killelea method, further improving the method (Horstman et al., 2014; Kang et al., 2019; Pyo et al., 2023; Verhoef et al., 1997).

However, the Mu broth exhibits a yellow autofluorescence, primarily due to its yeast extract containing flavins such as riboflavin (Jiménez-Nava et al., 2023; Masuda et al., 2010; Smith et al., 2024). This means that any reconstituted mVenus yellow fluorescence would likely be drowned out by medium autofluorescence if fluorescence measurements were taken while the *E. coli* were growing in Mu broth, especially since the reconstituted FP would likely fluoresce at 10% or less compared to the full-length protein (Bhakta et al., 2024; Kerppola, 2009). To avoid this, cells could instead be grown in M9 minimal medium, which has a far

lower autofluorescence due to its lack of rich nutrients (Bhakta et al., 2024). Furthermore, using 2X rather than 1X M9 minimal medium could increase fusion protein over-expression by up to two-fold, allowing for greater FP reconstitution, thus increasing relative fluorescent levels and hence making low-level PPIs easier to detect (Azatian et al., 2019; Kerppola, 2009). An issue, though, is that *E. coli* grows far slower in M9 minimal medium compared to nutrient-rich medium like Mu broth (Azatian et al., 2019; Elbing et al., 2019; Paudel et al., 2025). *E. coli* growth rate in M9 minimal medium can be improved by adding casamino acids, but this increases autofluorescence, even if lower than that of a nutrient-rich medium (Bhakta et al., 2024). Another possible solution to medium autofluorescence is changing the FP rather than the growth medium. Smith et al. (2024) showed that the autofluorescence spectra of LB medium, which like Mu broth contains tryptone and yeast extract, had extremely low red autofluorescence (Austin et al., 2021). The Bolt lab's FLUOstar Omega possesses excitation/emission wavelength filters of 590-50/670 nm, which could be applied to a BiFC assay using a far-red FP (Miller et al., 2015; Shang et al., 2024).

As such, 2X M9 minimal medium with or without a 1% casamino acids (w/v) supplement and Mu broth, all 100 µg/mL ampicillin + 40 µg/mL kanamycin, were analysed to determine their autofluorescence and their effect on the growth rate of BL21-AI *E. coli* cotransformed with plasmid pairs encoding NmVenus + CmVenus and NmVenus-POLD3 + CmVenus-HELQ, as described in Section 2.6.3. Media autofluorescence was measured using the FLUOstar Omega's 500/520 nm and 590-50/670 filters, with gain adjustments set against 95% of the highest fluorescence values when using the two filter pairs, giving 1473 and 3025, respectively (Figure 3.19). Sample autofluorescence means were analysed using two-tailed Welch's t-tests. Full t-test statistics are listed in Appendix Table 7.3. At gain 1473, 2X M9 minimal medium, when using 500/520 nm ($n = 6$, $M = 199832$, $SD = 3167$) and 590-50/670 nm filters ($n = 6$, $M = 1280$, $SD = 37.17$) compared to Mu broth ($n = 6$, $M = 232429$, $SD = 6043$; $n = 6$, $M = 1474$, $SD = 28.58$), showed a significantly lower autofluorescence ($t(7.554) = 11.70$, $p = 0.000004$; $t(9.381) = 10.14$, $p = 0.000002$). Meanwhile, supplemented 2X M9 medium, when using 500/520 nm ($n = 6$, $M = 211206$, $SD = 4866$) and 590-50/670 nm filters ($n = 6$, $M = 1378$, $SD = 100.0$) compared to Mu broth, showed a significantly lower autofluorescence only in the

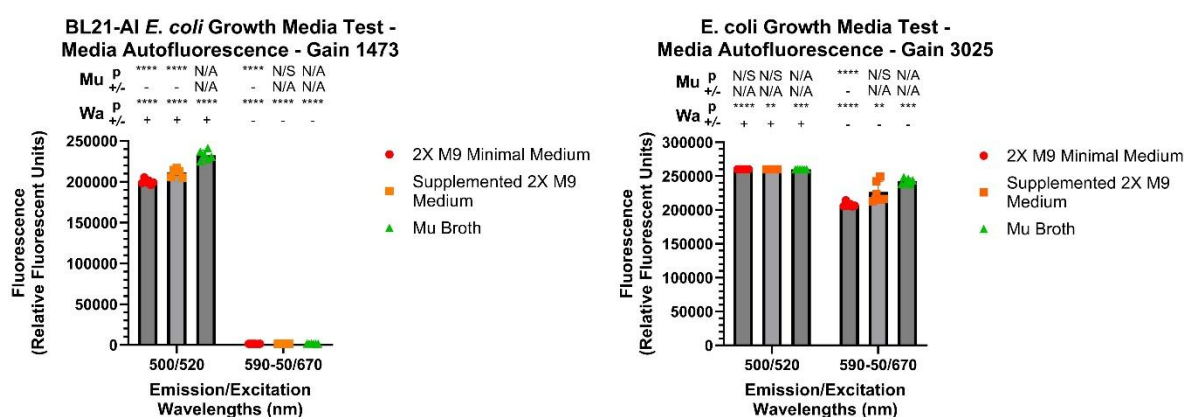


Figure 3.19 Autofluorescence of uninoculated growth media from BL21-AI *E. coli* growth media test.

Supplemented 2X M9 medium contains 1% (w/v) casamino acids. All three media contained 100 µg/mL ampicillin + 40 µg/mL kanamycin. For each medium, six 1.95 mL samples were used. Autofluorescence was measured in a clear non-tissue culture treated 24-well plate using a FLUOstar Omega (BMG LABTECH) plate reader. Gain adjustment was set at 95% of the highest fluorescence value for excitation/emission filters 500/520 nm and 590-50/670 nm, giving 1473 and 3025, respectively. Fluorescence values are indicated by colour and symbol for the different media, shown to the right. The column heights represent mean fluorescence values. Error bars represent standard deviations. Two-tailed Welch's t-tests with a two-stage step-up method were performed using GraphPad Prism 10 to determine if a medium's fluorescence mean showed a significant difference compared to Mu broth's fluorescence mean using the same filter pair and gain, indicated by 'Mu' (Benjamini et al., 2006). Two-tailed Welch's t-tests without a two-stage step-up method were performed to determine if a medium's fluorescence mean showed a significant difference when measured using the other filter pair and same gain, indicated by 'Wa'. Significance level is denoted by asterisks over the column in rows marked as 'p': ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$, N/S = 'no significance', and N/A = 'not applicable'. In the rows marked '+/-', + indicates that that column medium has a higher mean than the comparative medium mean; - indicates it is lower.

former case ($t(9.565) = 6.701$, $p = 0.000066$; $t(5.811) = 2.253$, $p = 0.066600$). When comparing respective media between filter pairs, at gain 1473 all media showed a significantly lower autofluorescence at the same degree when using 590-50/670 nm filters compared to 500/520 nm ($t(5.001) = 153.6$, $p < 0.0001$; $t(5.004) = 105.6$, $p < 0.0001$; $t(5.000) = 93.62$, $p < 0.0001$). At gain 3025 the media were again all significantly higher, but to varying degrees ($t(5.000) = 36.49$, $p < 0.0001$; $t(5.000) = 5.263$, $p = 0.0033$; $t(5.000) = 9.865$, $p = 0.0002$).

Media autofluorescence was visualised by scanning and imaging using the Amersham Typhoon 5 Biomolecular Imager's (Figure 3.20). The media was scanned using three built-in filters: Cy2, currently used for mVenus yellow fluorescence; Cy5, used for far-red fluorescence; and IRshort, used for short wavelength near-infrared fluorescence (Cytiva, 2020; He et al., 2023; Ferri et al., 2000). Under all three filters, 2X M9 minimal medium showed negligible autofluorescence and supplemented 2X M9 medium showed minimal autofluorescence. Mu broth showed higher autofluorescence than the other media under all three filters, extremely high under the Cy2 filter but far lower under Cy5 and IRshort filters.

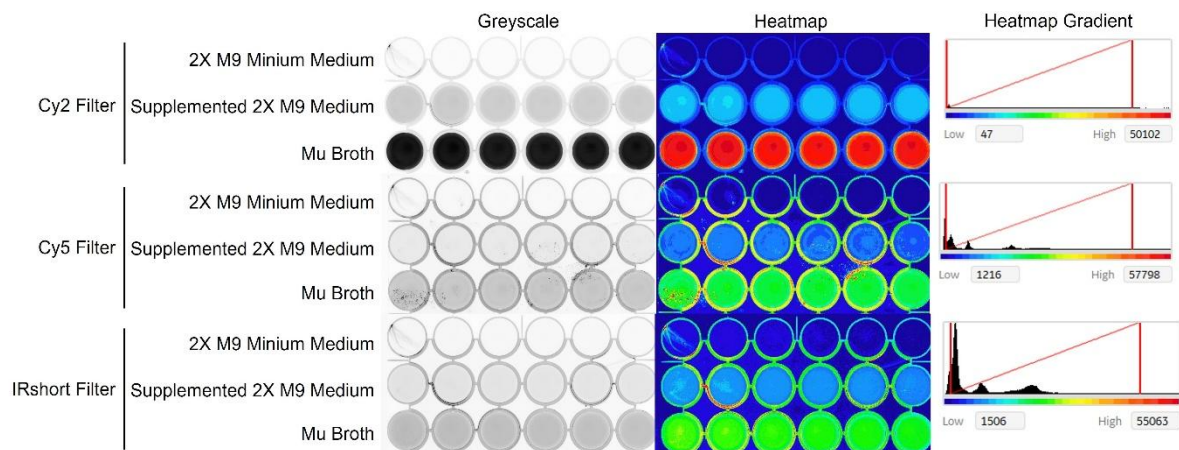


Figure 3.20 Amersham Typhoon 5 Biomolecular Imager (Cytiva) scan of uninoculated growth media autofluorescence from BL21-AI *E. coli* growth media test.

Supplemented 2X M9 medium contains 1% (w/v) casamino acids. All three media contained 100 µg/mL ampicillin + 40 µg/mL kanamycin. For each medium, six 1.95 mL samples were used in a clear non-tissue culture treated 24-well plate. The plate was scanned using the built-in Cy2, Cy5, and IRshort filter. The scans were as both greyscale and heatmap, indicated above the scans. The Cy2 filter greyscale scan was brightness/contrast adjusted using the software ImageJ. Darker areas on greyscale indicate greater autofluorescence. Heatmap colour gradient corresponding to autofluorescence level is shown on the right of the scans.

The media were then inoculated with BL21-AI *E. coli* cotransformed with plasmids encoding NmVenus + CmVenus or NmVenus-POLD3 + CmVenus-N-HELQ for 1:40 dilutions in technical triplicate and incubated in the FLUOstar Omega at 37 °C with double 300 RPM orbital shaking for 16 hours overnight, with sample OD₆₀₀ measured every 10 minutes (Figure 3.21). Logarithmic phase growth rate was calculated using the formula (Lindqvist et al., 2014):

$$\mu = \frac{\Delta \ln OD_{600}}{\Delta t}, \text{ where } t \text{ is time.}$$

For BL21-AI *E. coli* cotransformed with plasmids encoding NmVenus + CmVenus and NmVenus-POLD3 + CmVenus-N-HELQ, in 2X M9 minimal medium, supplemented 2X M9 medium, and Mu broth they respectively showed lag phase times of 140 minutes and 100 minutes, 10 and 0 minutes, and 10 minutes for both. Their exponential phase growth rates were respectively

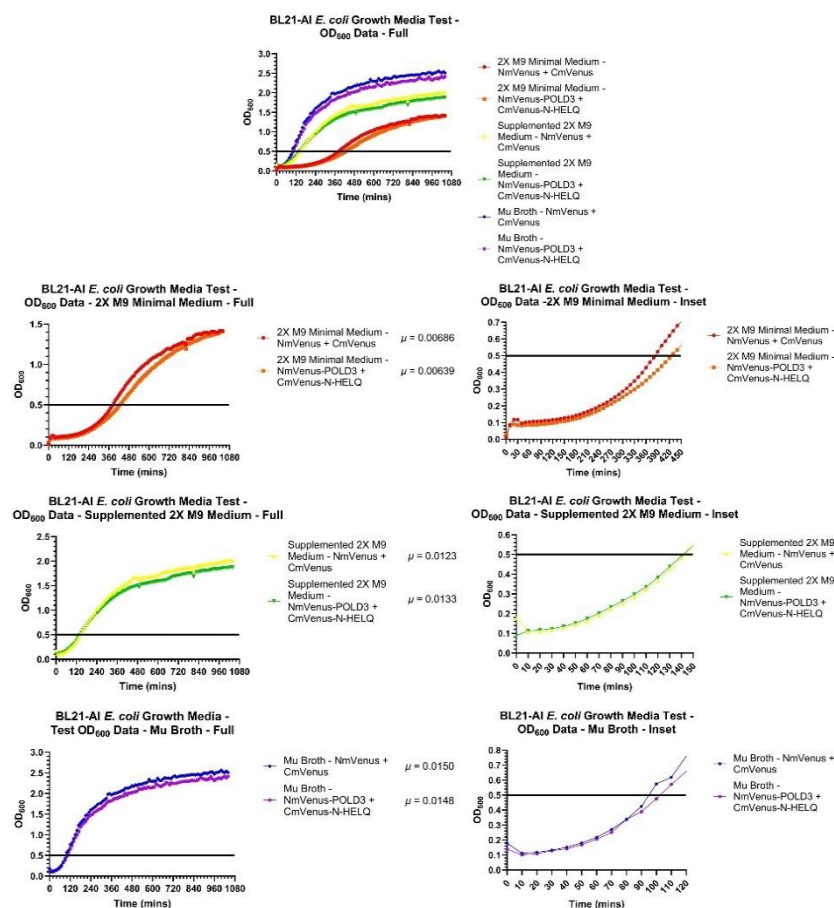


Figure 3.21 OD₆₀₀ data from BL21-AI *E. coli* growth media test.

Supplemented 2X M9 medium contains 1% (w/v) casamino acids. All three media contained 100 $\mu\text{g}/\text{mL}$ ampicillin + 40 $\mu\text{g}/\text{mL}$ kanamycin. For each medium, 1.95 mL was inoculated with 50 μL of BL21-AI *E. coli* cotransformed with plasmids encoding NmVenus + CmVenus or NmVenus-POLD3 + CmVenus-N-HELQ for 1:40 dilutions in technical triplicate in a clear non-tissue culture treated 24-well plate. The plate was incubated in a FLUOstar Omega (BMG LABTECH) plate reader at 37 °C with double 300 RPM orbital shaking for 16 hours overnight. Sample OD₆₀₀ was measured every 10 minutes with 2 mL of respective uninoculated media serving as blanks. Triplicate averages are plotted and the colour corresponding to each condition is indicated next to the graphs. The point at which growth curves cross the vertical line at on each graph indicates when they reached OD₆₀₀ 0.5, with this shown for each medium on the full growth curves on the left ('Full') and in greater detail in the graphs on the right ('Inset'). Exponential phase growth rates (μ) for each condition are rounded to three significant figures and listed to the right of the full growth curves.

0.00686 and 0.00639, 0.0123 and 0.0133, and 0.0150 and 0.0148. Cells reached OD₆₀₀ 0.5 after, respectively, 390 minutes and 430 minutes, 150 minutes for both, and 100 and 110 minutes. The three media in order of 2X M9 minimal medium, supplemented 2X M9 medium, and Mu broth thus showed a trend of shorter lag phase times, greater growth rates, and shorter times taken to reach OD₆₀₀ 0.5.

3.5 Summary

Previous Bolt lab work in Lou-Hing (2024) using the He et al. (2023) time-course BiFC assay method showed that mutating the POLD3 Quincunx motif (residues 319-KKRRR-323) in its IDPR3 NLS to glycines (319-GGGGG-323) fully inhibited POLD3-N-HELQ PPI (Alli et al., 2024). In order to attempt to reproduce this result, necessary BiFC assay plasmids stocks were produced. The initial pBAD-NmVenus-PoID3^{QC} plasmid stocks were of unexpected lengths when double analytical restriction digested and electrophoresed on a 1% (w/v) agarose gel, but further plasmids that originated from a different pBAD-NmVenus-PoID3^{QC} plasmid stock were the expected length. Subsequent attempts to reproduce the Lou-Hing (2024) result were unsuccessful, thus a new endpoint BiFC assay method was developed based on a currently unpublished method developed by the Bolt lab's then-postdoc Doctor Tom Killelea. Using this method, mutating the Quincunx motif to 319-GGGGG-323, 319-KSRRR-323, 319-KKSRR-323, 319-KKDRR-323, or 319-AKRRA-323 showed no inhibition of PPI. The pattern of relative fluorescence levels observed in the Amersham Typhoon biomolecular imager scan was also reflected in relative fluorescence values normalised using sample OD₆₀₀ when scanned using the FLUOstar Omega plate reader. However, when normalising the relative fluorescence values using CFU/mL, even with outliers removed, no result showed a significant difference compared to NmVenus + CmVenus fluorescence, NmVenus-POLD3 + CmVenus-N-HELQ fluorescence, or NmVenus-POLD3^{QC} + CmVenus-N-HELQ fluorescence in two-tailed Welch's t-tests.

It was decided to continue using the new BiFC assay method, this time to investigate whether mutating the POLD3 381-RKRKR-385 motif in its IDPR3 PNLs to glycines (381-GGGGG-385) would affect POLD3-N-HELQ PPI (Alli et al., 2024). Meanwhile, fellow Bolt lab MRes student Emily Haskins used the He et al. (2023) method to show preliminary data indicating that POLD3 IDPR1 does not interact with N-HELQ, while IDPRs 2, 3, and 4 do, with IDPR3 showing the strongest PPI. It was hence decided to incorporate the IDPRs into the assay. The pattern of relative fluorescence levels observed in the Amersham Typhoon biomolecular imager scan supported Haskins' data and was again reflected in relative fluorescence values normalised using sample OD600 when scanned using the FLUOstar Omega plate reader. In addition, NmVenus-POLD3^{PNLS} + CmVenus-N-HELQ showed fluorescence on par NmVenus-POLD3 + CmVenus-N-HELQ, indicating that this mutation does not inhibit PPI. No colonies grew in the spot assay from cell samples expressing NmVenus + CmVenus, NmVenus-POLD3 + CmVenus-N-HELQ, NmVenus-POLD3^{PNLS} + CmVenus-N-HELQ, and so relative fluorescence values could not be normalised using CFU/mL. It was hypothesised that cells in these samples were in the VBNC state rather than dead (Pienaar et al., 2016; Ramamurthy et al., 2014; Yang, D. et al., 2024). This was tested based on the method in Wei et al. (2018), who showed that O157:H7 *E. coli* could be resuscitated by incubation in LB medium containing 5% Tween 80 at 37 °C with shaking for 48 hours of incubation. Among the original BiFC assay cell samples of the technical triplicates of cells expressing NmVenus + CmVenus, NmVenus-POLD3 + CmVenus-N-HELQ, NmVenus-POLD3^{PNLS} + CmVenus-N-HELQ, at least one replicate showed colony growth in both its VBNC test duplicates, although no clear distinction could be made on whether 5% (v/v) Tween 80 resulted in greater colony growth.

It was then decided to attempt to develop the endpoint method into a time-course method that would provide relative fluorescence values over time. However, Mu broth has yellow autofluorescence, thus a far-red FP may be more suitable (Jiménez-Nava et al., 2023; Masuda et al., 2010; Miller et al., 2015; Shang et al., 2024; Smith et al., 2024). 2X M9 minimal medium with or without a 1% casamino acids (w/v) supplement and Mu broth, all containing 100 µg/mL ampicillin + 40 µg/mL kanamycin, were analysed to determine their autofluorescence and their effect on the growth rate of BL21-AI *E. coli* cotransformed with plasmid pairs

encoding NmVenus + CmVenus and NmVenus-POLD3 + CmVenus-HELQ. The three media, in the previously listed order, showed a trend of shorter lag phase times, greater growth rates, and shorter times taken to reach OD₆₀₀ 0.5. When using gain 1473 that was adjusted using excitation/emission 590-50/670 nm filters for far-red fluorescence rather than 500/520 nm for yellow fluorescence, 2X M9 minimal medium with and without 1% (w/v) casamino acids showed significantly less yellow autofluorescence than Mu broth, and all three media showed significantly less red autofluorescence than yellow autofluorescence.

To investigate further POLD3-N-HELQ PPI using BiFC assays, residues of interest would need to be identified in POLD3 IDPRs 2, 3, and 4, using software such as AlphaFold 3 in order for mutants to be generated (Abramson et al., 2024). Furthermore, three NmVenus-POLD3^{IDPR3} mutants would need to be generated, one each with the Quincunx motif or motif in the PNLS mutated to glycines and one with both mutations (Alli et al., 2024; Lou-Hing, 2024). This would allow investigating whether one or both of these regions contribute to the PPI. If the endpoint method in this work were to be used, it may be beneficial to use a far-red FP and FLUOstar Omega 590-50/670 nm filters, allowing for a time-course method. By identifying the residues involved in POLD-N-HELQ PPI, this could hence further the understanding of the molecular mechanisms behind the interaction and ultimately how HELQ limits mutagenic DNA synthesis during BIR (He et al., 2023).

Chapter 4: Analysis of RPA1-N-HELQ Protein-Protein Interaction

4.1 Introduction

4.1.1 RPA1 Structure and Function

As described previously, RPA is the main eukaryotic SSB and primarily functions to protect ssDNA from nucleolytic degradation, acting in a multitude of DNA metabolic activities, including in DNA replication as part of the replisome (see Section 1.2.2) and in DNA DSB repair HR pathways such as BIR (see Section 1.3.2) (Brill et al., 1998; Dueva et al. 2020; Inoue et al., 2008; Paul et al., 2025; Yao et al., 2010). RPA is a heterotrimer complex, comprised of the subunits RPA1, RPA2, and RPA3 (Dueva et al. 2020; Wold, 1997). These are summarised in Table 4.1. While RPA structure has been extensively studied, no apo-form cryo-EM or crystal structure is currently available (Fousek-Schuller et al., 2024). Therefore, RPA subunit residue information is shown in Figure 4.1 and its structure as predicted by AlphaFold 3 is shown in Figure 4.2 and annotated with domains in Figure 4.3.

Table 4.1 Summary of human ssDNA-binding protein RPA complex subunits.

RPA Subunit	Cytogenetic Location	Number of Amino Acid Residues	Molecular Mass (kDa)	Main Function Within RPA
RPA1	17p13.3 (Umbricht et al., 1994).	616 (Erdile et al., 1991)	70 (Erdile et al., 1991)	‘Responsible for [high-affinity ssDNA-binding] and many protein-protein interactions’ (Hass et al., 2012).
RPA2	1p35.3 (NCBI, 2025g)	270 (Erdile et al., 1990)	32 (Erdile et al., 1990)	Mediates RPA inter-subunit

				interactions and N-terminus phosphorylation may allow RPA dissociation from ssDNA (Dueva et al., 2020).
RPA3	7p21.3 (NCBI, 2025h)	121 (Umbricht et al., 1993)	14 (Umbricht et al., 1993)	Coordinates RPA1 and RPA2 interaction (Nasheuer et al., 2024).

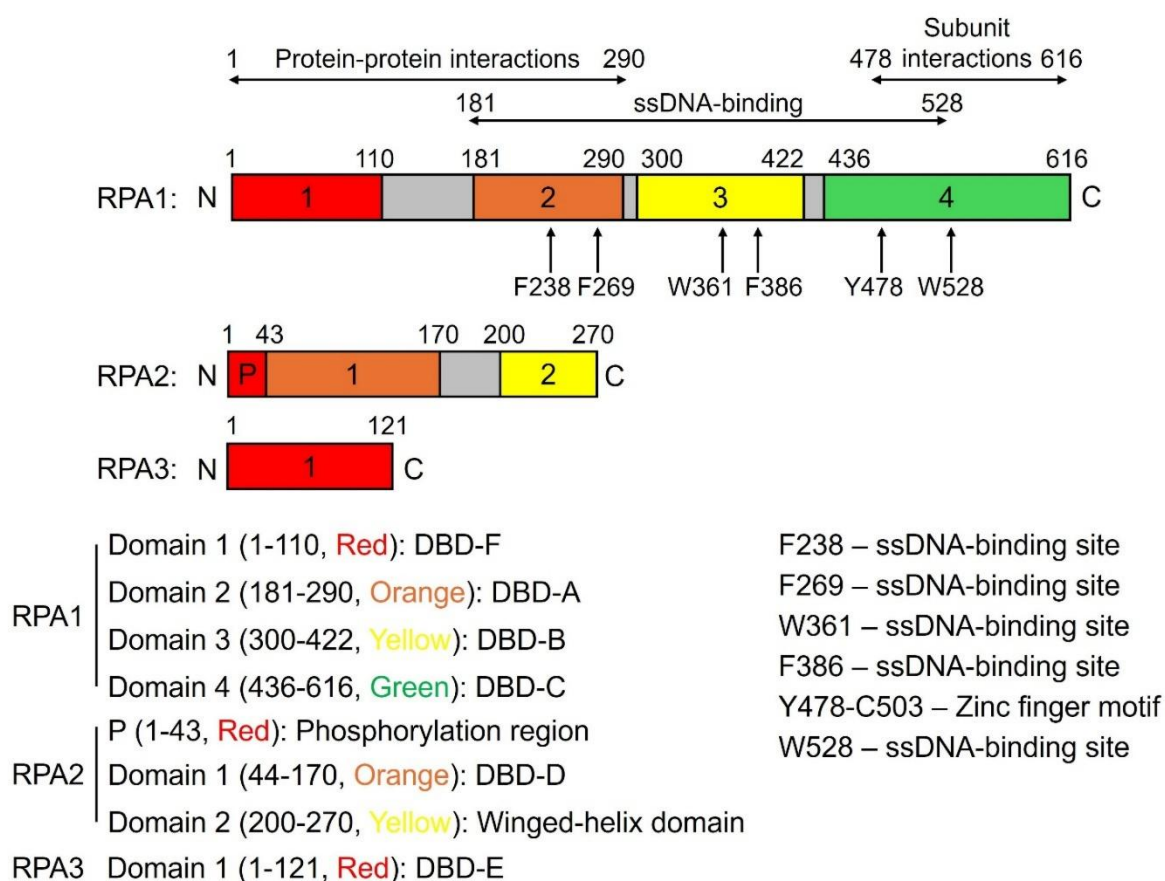


Figure 4.1 RPA subunits amino acid residue information.

RPA subunits mainly consist of DNA-binding domains (DBDs), while RPA2 also possesses an N-terminal phosphorylation region and C-terminal winged-helix domain (Nasheuer et al., 2024). These are denoted by letters/number and colour. RPA1 has three distinct overlapping regions for different functions: protein-protein interaction, single-stranded DNA (ssDNA)-binding, and RPA inter-subunit interactions (Fousek-Schuller et al., 2024). RPA1 ssDNA-binding functional aromatic residues and zinc finger motif are indicated (Cai et al., 2007; Ensembl, 2025c; Park et al., 1999; Walther et al., 1999). Figure is adapted from Figures 1A in Nasheuer et al. (2024) and Fousek-Schuller et al. (2024).

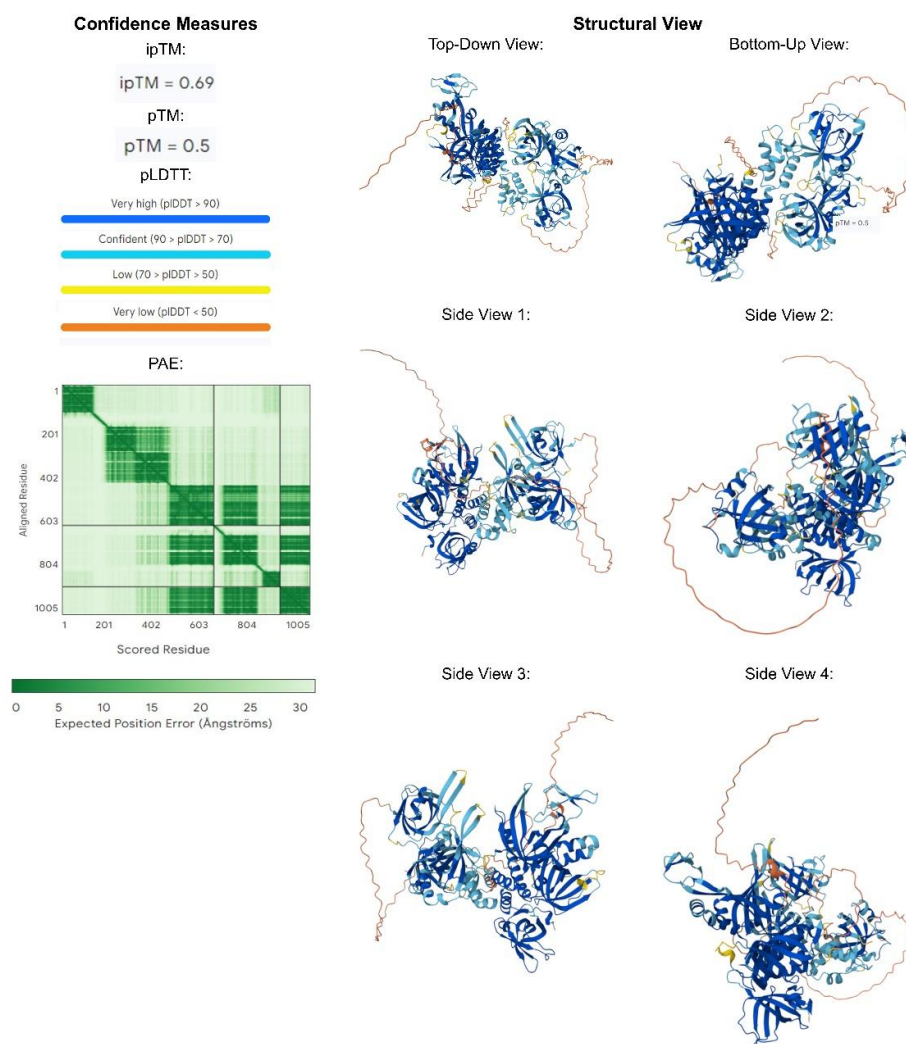


Figure 4.2 RPA structure predicted by the software AlphaFold 3.

Structure is shown from multiple orientations. AlphaFold 3 uses multiple confidence measures (Abramson et al., 2024). Interface Predicted Template Modelling (ipTM) score measures interaction interface prediction accuracy, ranging from 0-1 with threshold of 0.6; a higher score indicates greater accuracy. pTM score measures overall structure prediction accuracy, again ranging from 0-1 but with threshold of 0.5 (Wee et al., 2024). Predicted Local Distance Difference Test (pLDDT) score measures individual residue structure prediction confidence, ranging from 0-100; a higher score indicates greater confidence. pLDDT score is denoted by residue colour. Predicted Align Error (PAE) assesses overall structural prediction confidence, measured in angstroms. Lower PAE indicates greater prediction confidence (Pratt et al., 2025).

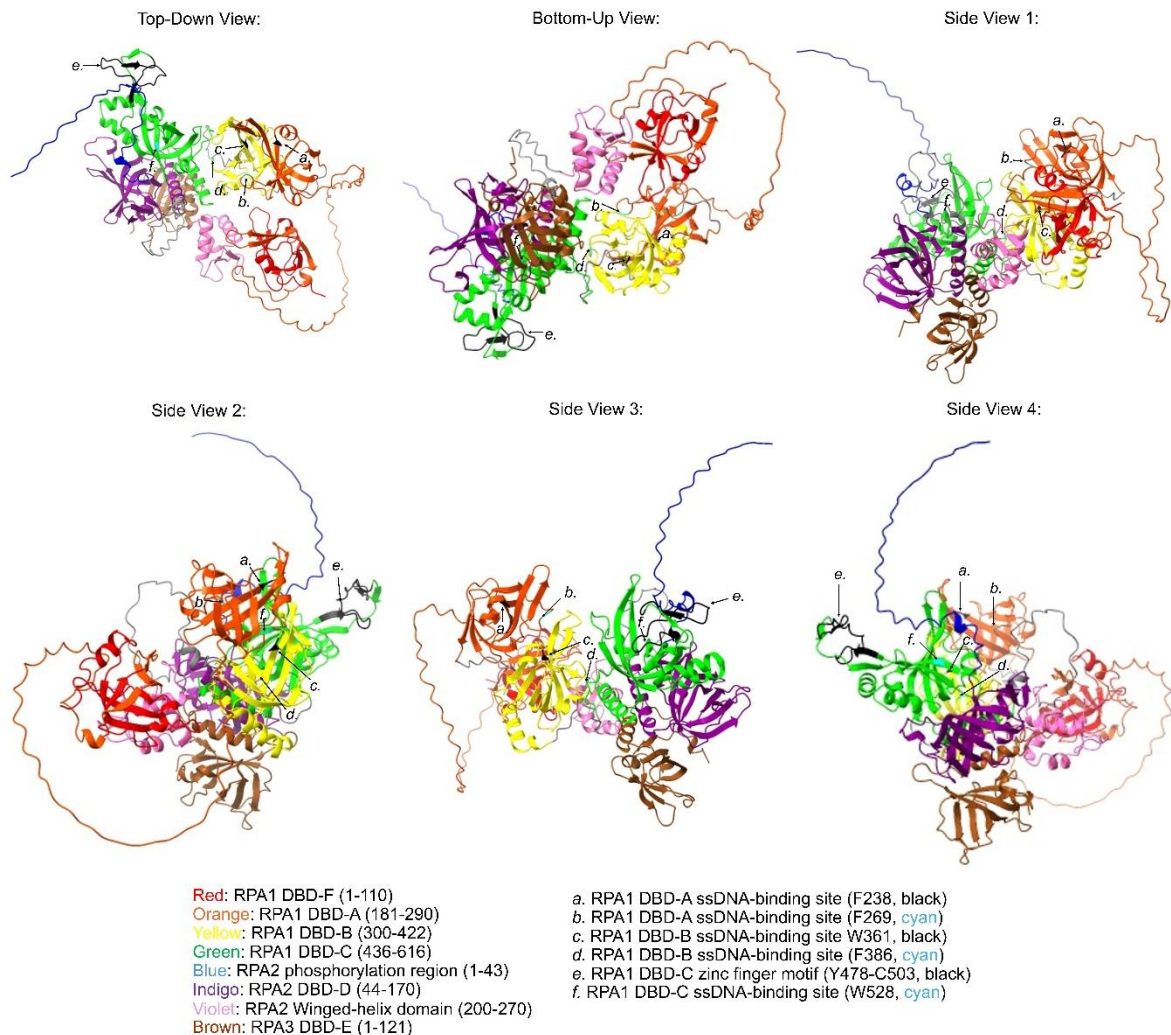


Figure 4.3 RPA predicted structure organisation.

RPA structure as predicted by AlphaFold 3 (Figure 4.2) is shown from multiple orientations. Subunit domains and regions are indicated by colour. Functional residues are coloured black or cyan and indicated by letter.

Like POLD3, the RPA1 subunit is primarily involved in maintaining genomic integrity by preventing genomic instability, in this case functioning both directly in protecting ssDNA against nucleolytic degradation and as a PPI hub in replicative DNA repair processes such as BIR (Alli et al., 2024; Dueva et al., 2020; He et al., 2023; Nasheuer et al., 2024). RPA1 is characterised by four DNA-binding domains (DBSs): DBD-F (residues 1-110), DBD-A (residues

181-290), DBD-B (residues 300-422), and DBD-C (residues 436-616) (Nasheuer et al., 2024). Each DBD contains an oligonucleotide/oligosaccharide-binding (OB)-fold (Haring et al., 2008), comprised of ‘twisted β sheets and α -helix capped ends’ (Nasheuer et al., 2024), while DBD-C also contains a zinc finger motif (residues 478-528) (Nasheuer et al., 2024; Park et al., 1999). It is this structure that enables RPA1 to function during BIR (Dueva et al., 2020).

During HR, RPA binds to a ssDNA intermediate, which during presynapsis includes newly resected 3' ssDNA tails at DSBs to protect against secondary structure formation and nucleolytic degradation (Dueva et al., 2020; Inoue et al., 2008; Wright et al., 2018). RPA1 enables RPA high-affinity ssDNA-binding, predominantly through the four OB-fold aromatic residues of DBD-A F238 and F269 and DBD-B W361 and F386 in the stacking between the ssDNA bases (Dueva et al., 2020; Haring et al., 2008; Hass et al., 2012; Nasheuer et al., 2024; Walther et al., 1999). This binding is also contributed to at a lesser extent by DBD-C through its OB-fold aromatic residue W528 and zinc finger motif (residues 478-503) (Cai et al., 2007; Dueva et al., 2020; Park et al., 1999). RPA1 is able to exist as part of RPA in this interaction through the RPA heterotrimerisation core, formed through hydrophobic interactions of OB-fold α -helices in DBD-C with RPA2 DBD-D and RPA3 DBD-E, with the DBD-C zinc finger motif providing structural stability (Dueva et al., 2020; Nasheuer et al., 2024). RPA is subsequently displaced by interaction with the DNA recombinase RAD51 through RPA1 DBD-A and the DBD-F-DBD-A linker region (Ma et al., 2017; Nasheuer et al., 2024; Stauffer et al., 2004; Wright et al., 2018). During the postsynapsis sub-pathway BIR, as the D-loop migrates as a non-canonical replication fork, RPA binds to the ssDNA template strand in the invaded duplex (Donnianni et al., 2019; He et al., 2023; Li et al., 2008). RPA1 may then interact with HELQ via the latter's N-HELQ PWI-like fold, displacing RPA (He et al., 2023). RPA1 structure as predicted by AlphaFold 3 is shown in Figure 4.4 and illustrated with its domains in Figure 4.5.

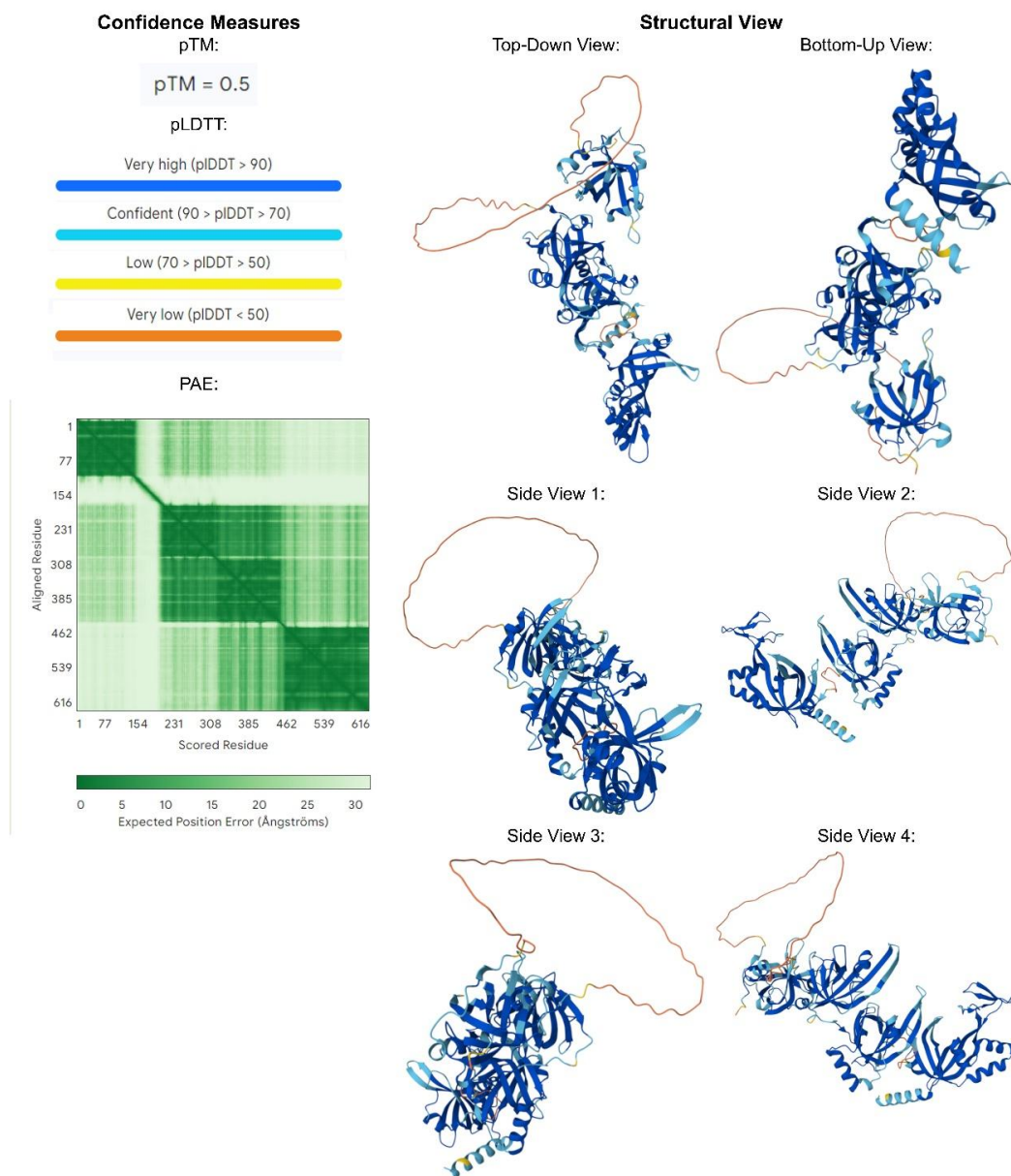


Figure 4.4 RPA1 structure predicted by the software AlphaFold 3.

Structure is shown from multiple orientations. AlphaFold 3 uses multiple confidence measures (Abramson et al., 2024). Predicted Template Modelling (pTM) score measures overall structure prediction accuracy, ranging from 0-1 with threshold of 0.5; a higher score indicates greater accuracy (Wee et al., 2024). Predicted Local Distance Difference Test (pLDDT) score measures individual residue structure prediction confidence, ranging from 0-100; a higher score indicates greater confidence. pLDDT score is denoted by residue colour. Predicted Align Error (PAE) assesses overall structural prediction confidence, measured in angstroms. Lower PAE indicates greater prediction confidence (Pratt et al., 2025).

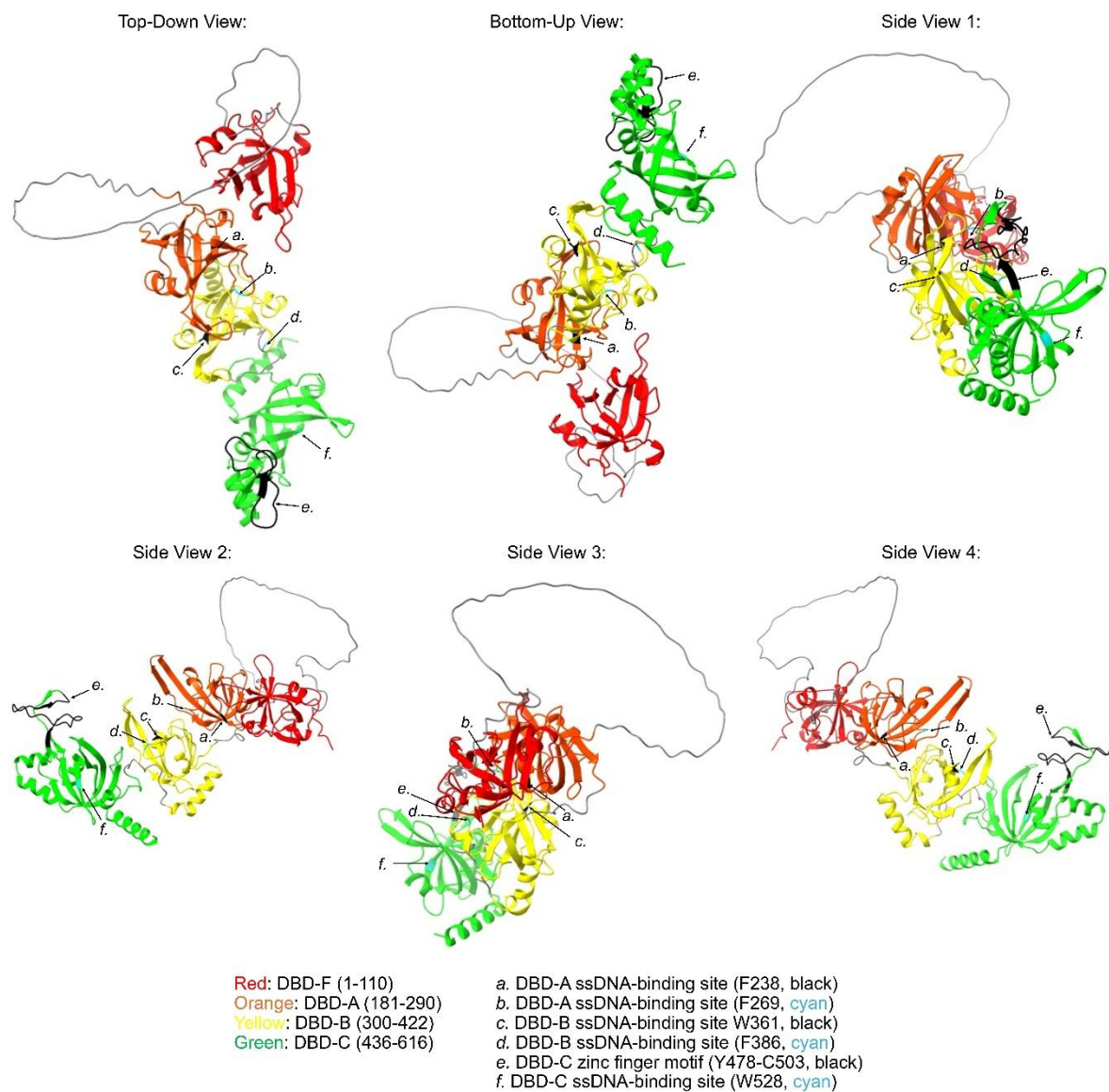


Figure 4.5 RPA1 predicted structure organisation.

RPA1 structure as predicted by AlphaFold 3 (Figure 4.4) is shown from multiple orientations. DNA-binding domains (DBDs) are indicated by colour. Functional motifs are coloured black or cyan and indicated by letter.

4.1.2 PE3b Prime Editing

Prime editing is a precision search-and-replace genome editing method that modifies the genomes of living cells, developed from the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 precision gene editing system (Anzalone et al., 2019; Chen et al., 2023). CRISPR are short, partially repeated DNA sequences found in the genomes of prokaryotes' (Asmamaw et al., 2021). CRISPR were first discovered in bacterial DNA, specifically in *E. coli* by Ishino et al. (1987), and were later discovered to also be present in archaeal DNA, first found in *Haloferax mediterranei* by Mojica et al. (1993). CRISPR were first hypothesised by Mojica et al. (2005), and later experimentally confirmed by Barrangou et al. (2007), to be a key element in prokaryotic adaptive immunity against viruses (Asmamaw et al., 2021). CRISPR are grouped into intergenic arrays, comprised of two to several hundred 23-50 bp short palindromic repeats separated by 17-86 bp viral DNA fragment spacers, (Gostimskaya, 2022; Shmakov et al., 2020). A CRISPR array typically exists in a CRISPR-*cas* locus, being associated with one or more adjacent *cas* genes which encode Cas proteins (Gostimskaya, 2022; Pinilla-Redondo et al., 2022; Shmakov et al., 2020). The CRISPR-Cas system immune response has three stages, mediated by its Cas protein(s): adaption, in which new spacers are incorporated into CRISPR arrays; processing, in which a long pre-CRISPR RNA (pre-crRNA) transcribed from a CRISPR array matures into multiple smaller crRNA (Karvelis et al., 2013; Liao et al., 2021; Shmakov et al., 2020); and interference, in which crRNA 'guides [Cas proteins] to bind and cleave complementary target DNA or RNA' (Shmakov et al., 2020).

Due to the number of Cas proteins, they are classified based on class and type, where Class 1 proteins, subdivided into Type I, III, and IV, are multi-subunit complexes and Class 2 proteins, subdivided into Type II, V, and VI, are single multidomain proteins, as shown in Figure 4.6 (Gostimskaya, 2022). Cas9 is a bacterial RNA-guided DNA endonuclease consisting of a single multidomain protein, classified as Class II Type 2 (Anders et al., 2014; Gostimskaya, 2022). During processing, trans-activating CRISPR RNA (tracrRNA) is required for pre-crRNA maturation into crRNAs that contain a single spacer flanked by repeats (Karvelis et al., 2013;

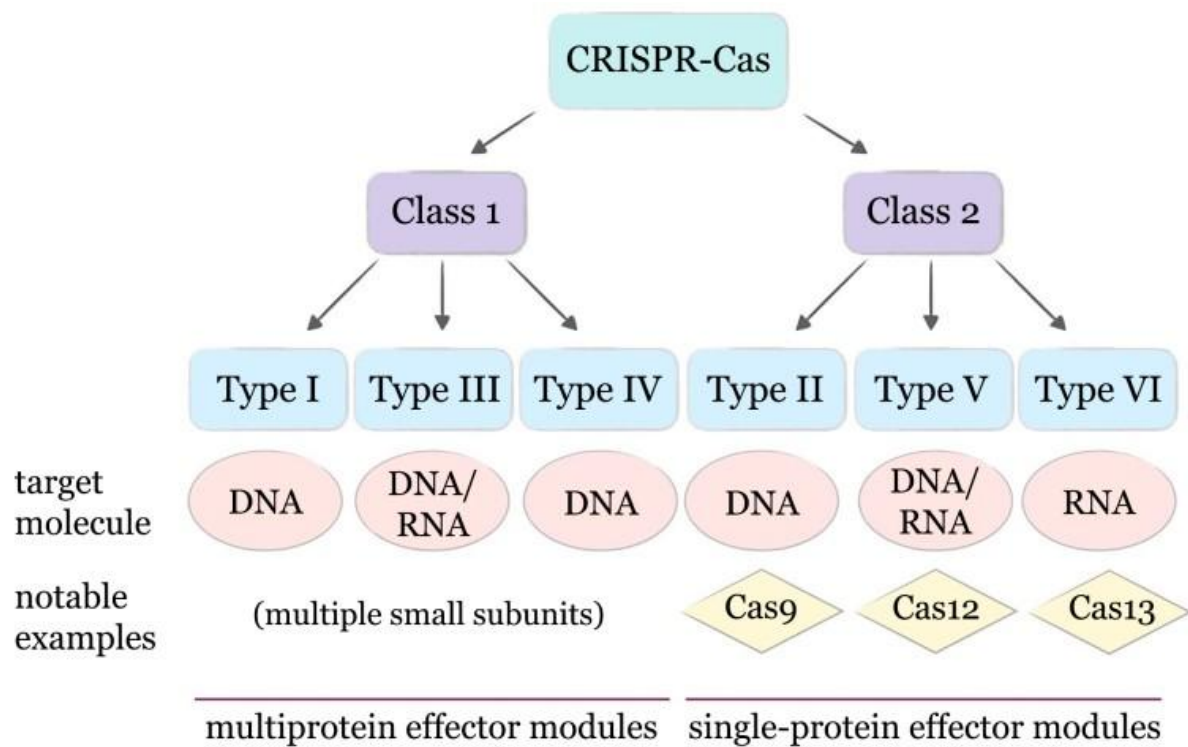


Figure 4.6 Classification of CRISPR-Cas systems based on Cas protein.

CRISPR-Cas systems are broadly divided into two Classes based whether the Cas protein is a multi-subunit complex or single protein and subdivided into Types based their target nucleic acid molecule(s). Figure is Figure 1 from Gostimskaya (2022).

Liao et al., 2021). During interference, Cas9 complexes with a crRNA and tracrRNA, with the crRNA guiding the Cas9 to a protospacer DNA sequence that is complementary to the crRNA spacer and flanked by a 2-5 bp protospacer adjacent motif; the Cas9 subsequently cleaves the DNA (NEB, 2014; Stamereilers et al., 2021).

The CRISPR-Cas9 system for precision gene editing was first used in eukaryotic cells in 2013 in five near-simultaneous publications and primarily consists of a Cas9 protein and a single guide RNA (sgRNA), which is a tracrRNA-crRNA chimera (Anders et al., 2014; Chen et al., 2023; Gostimskaya; 2022; Lu et al., 2022). The Cas9-sgRNA complex binds to the target DNA sequence and cleaves, generating a DSB. This is repaired by homology-directed repair (HDR) pathways such as HR, generally resulting in precise edits, or NHEJ, risking insertion and

deletion (indel) mutations in the edit, especially in cells outside of mitosis (Lu et al., 2022). If an exogenous donor DNA template is co-delivered alongside the Cas9-sgRNA, HDR can 'recombine the DNA template into the DSB site', being able to generate insertions and deletions (Chen et al., 2023). Furthermore, fusing a nuclease-inactive Cas9 with deaminases allows for base editing, performing all four transition mutations (C→T, G→A, A→G, and T→C) and two of the eight transversion mutations (C→G and C→A) without the risk of NHEJ generating unwanted mutations. However, this base editing is unable to perform all 12 base-to-base conversions and risks unwanted bystander products (Anzalone et al., 2019; Lu et al., 2022).

This led to the development of prime editing by Anzalone et al. (2019), capable of performing insertions, deletions, and all 12 base-to-base conversions without requiring a donor DNA template, nor generating DSBs or bystander products. They developed the system through multiple iterations, each with higher efficiency: PE1 (0.2-17% efficiency), PE2 (1.6-5.1-fold higher than PE1), and PE3 (3-fold higher than PE2), and PE3b (comparable to PE3) (Lu et al., 2022). PE3b also results in 13-fold average lower indels compared to PE3 (Anzalone et al., 2019). PE3b uses three components: a prime editing protein, a prime editing guide RNA (pegRNA), and sgRNA. The protein is a fusion of a nickase Cas9^{H840A} and reverse transcriptase M-MLV^{D200N/T306K/W313F/T330P/L603W} (nCas9-RT) and may also include optimised NLSs to further improve prime editing efficiency (Anzalone et al., 2019; Liu et al., 2021). In terms of mechanism, the pegRNA binds to the Cas9^{H840A}, which then binds to a target DNA sequence in a duplex complementary to the pegRNA 5' spacer sequence and nicks the non-target strand. The 3' ends of the nicked strand and pegRNA primer binding site subsequently anneal (Doman et al., 2022; Zhang, W. et al., 2024). The reverse transcriptase then uses the DNA-RNA duplex as a substrate, the non-target DNA 3' flap as a primer, and the pegRNA reverse transcription template to synthesise DNA with the desired edit (Doman et al., 2022). Through flap equilibration, the newly synthesised 3' flap is retained and the unmodified 5' flap is cleaved, while the sgRNA directs the Cas9^{H840A} to nick the target strand away from the non-target strand nick site (Anzalone et al., 2019; Lu et al., 2022). Finally, MMR results in the edit either being incorporated into the target strand or reverted back to the original sequence, with the

target strand nick making MMR favour editing (Doman et al., 2022). This process is illustrated in Figure 4.7.

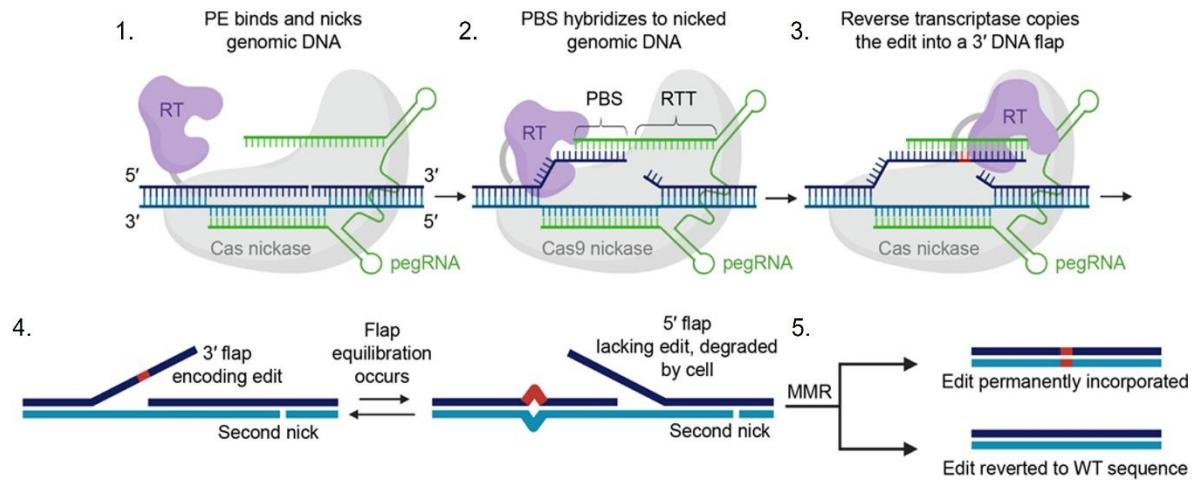


Figure 4.7 Mechanism of PE3b prime editing.

The PE3b system has five distinct stages. 1. The pegRNA (green) recruits a fusion protein via its Cas9^{H840A} nickase domain (grey) to a target DNA sequence in a duplex and the Cas9^{H840A} nicks the non-target strand. 2. The 3' ends of the nicked strand and pegRNA primer binding site (PBS) anneal (Doman et al., 2022; Zhang, W. et al., 2024). 3. The fusion protein M-MLV^{D200N/T306K/W313F/T330P/L603W} reverse transcriptase domain (purple) uses the DNA-RNA duplex as a substrate, the non-target DNA 3' flap as a primer, and the pegRNA reverse transcription template (RTT) to synthesise DNA with the desired edit (Doman et al., 2022). 4. The newly synthesised 3' flap is retained and the unmodified 5' flap is cleaved through flap equilibration, while the sgRNA directs the Cas9^{H840A} to nick the target strand away from the non-target strand nick site (Anzalone et al., 2019; Lu et al., 2022). 5. Mismatch repair (MMR) results in one of two outcomes: the edit is incorporated into the target strand (the target strand nick makes MMR favour this outcome) or reverted back to the original sequence (Doman et al., 2022). Figure is adapted from Figure 1 in Doman et al. (2022).

4.2 Analysis of RPA1-N-HELQ Protein-Protein Interaction Using BiFC Assay

4.2.1 Restriction Cloning of NmVenus-RPA Subunit Plasmids and Pilot Protein Overexpression

In previous work by the Bolt lab, Jenkins et al. (2021) used EMSAs and affinity pull-down assays to show RPA1-HELQ interaction but could not detect stable RPA-N-HELQ interaction using those methods or chemical crosslinking, microscale thermophoresis, and isothermal microcalorimetry. However, at that point the more sensitive He et al. (2023) mVenus time-course BiFC assay method had not yet been developed. Therefore, after the POLD3 IDPRs and PNLS BiFC assay, it was decided to investigate whether RPA1-N-HELQ interaction could be observed using the endpoint BiFC assay method, giving the third aim of this work.

p11d-tRPA(123) plasmid stocks were produced by transforming competent DH5 α *E. coli* with p11d-tRPA (see Section 2.3.3), then overnight cultures were grown (see Section 2.3.4) and minipreped (see Section 2.4.1). The resulting plasmid DNA was analytically restriction digested (see Section 2.4.2) and electrophoresed on a 1% (w/v) agarose gel (see Section 2.4.3) (Figure 4.8). DNA encoding RPA1, RPA2, and RPA3 was PCR amplified from p11d-tRPA (see Section 2.4.5). The RPA1 DNA PCR used a 30 μ L reaction mixture containing 1 μ L 10 mM dNTPs, with all other DNA and enzyme component volumes kept the same as a standard 25 μ L reaction; all other component volumes were scaled up for a 30 μ L reaction mixture. The desired RPA1 DNA PCR product subsequently required electrophoresis on a 1.2% agarose gel at 120 V for 1 hour (see Section 2.4.3) and gel extraction (see Section 2.4.4). PCR specific details are listed in Table 4.2. Electrophoresed agarose gels to verify PCR product size are shown in Figure 4.8.

Table 4.2 PCR amplification of DNA encoding RPA subunits from purified p11d-tRPA(123) plasmid DNA reaction details.

RPA Subunit	Forward Primer	Reverse Primer	Annealing Temperature (°C)	Extension Time (seconds)
RPA1	KpnI_RPA1_v2_F	EcoRI_RPA1_R	71	57
RPA2	XhoI_RPA2_v2_F	HindIII_RPA2_v2_R	68	25
RPA3	XhoI_RPA3_v2_F	HindIII_RPA3_v2_R	72	12

PCR product DNA was then purified (see Section 2.4.6). Restriction cloning was then performed to generate recombinant plasmids of pBAD-NmVenus plasmid vectors containing purified PCR product DNA inserts encoding RPA1, RPA2, or RPA3 (see Section 2.4.8). Restriction cloning specific details are listed in Table 4.3.

Table 4.3 Restriction cloning of pBAD-NmVenus plasmid vector and RPA DNA subunit inserts restriction digestion and ligation reaction details.

RPA Subunit	Restriction Enzymes	Plasmid Vector Restriction Digest Time (hours)	Plasmid Vector Mass (ng)	Vector:Insert Ratio
RPA1	KpnI-HF and EcoRI-HF	18*	50	3:1
RPA2	XhoI and HindIII-HF	1	24	3:1
RPA3	XhoI and HindIII-HF	1	24	3:1
			50	5:1
			30	7:1

* Plasmid vector restriction digestion reaction mixture was incubated overnight due restriction site proximity.

The resultant pAR001 (pBAD-NmVenus-RPA1) and pAR002 (pBAD-NmVenus-RPA2) were verified through analytical restriction digests that were electrophoresed on agarose gels (Figure 4.8) and subsequently whole plasmid sequencing and Sanger sequencing, respectively. All pBAD-NmVenus-RPA3 ligation reactions were unsuccessful.

To verify NmVenus-RPA1 and NmVenus-RPA2 could be successfully overexpressed in BL21-AI *E. coli*, pilot protein overexpression was performed using pBAD-NmVenus, pBAD-NmVenus-PolD3, pBAD-NmVenus-RPA1, and pBAD-NmVenusRPA2 (see Section 2.5.1), followed by SDS-PAGE analysis (see Section 2.5.2). pBAD-NmVenus-PolD3 functioned as a positive control, while pBAD-NmVenus and uninduced controls for each plasmid functioned as negative controls. The SDS-PAGE analysis showed that all induced BL21-AI *E. coli* cultures showed successful protein overexpression (Figure 4.9).

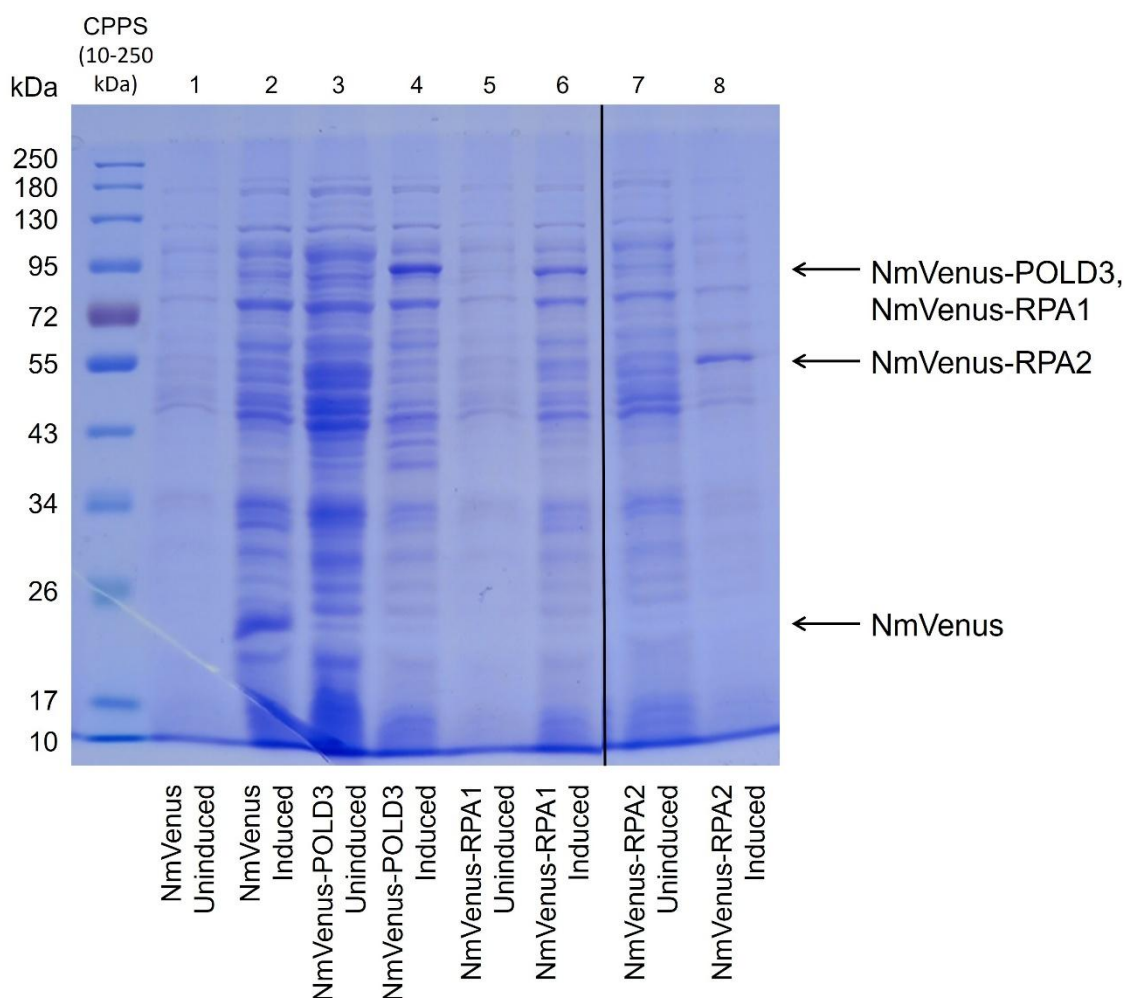


Figure 4.9 SDS-PAGE gel of NmVenus-RPA subunit proteins pilot overexpression SDS-PAGE samples to verify protein overexpression.

One SDS-PAGE sample from each BL21-AI *E. coli* culture was electrophoresed on a 10% (w/v) SDS-PAGE gel at 120 V for 34 minutes followed by 140 V for 36 minutes alongside Color Prestained Protein Standard (CPPS), Broad Range (10-250 kDa) (NEB). CPPS molecular masses in kDa are shown to the left of the gel. Lane numbers are shown above the gel. Whether a culture was uninduced or induced and the protein overexpressed are shown below the corresponding lane. Expected overexpressed protein positions shown to the right of the gel. Black line indicates truncated region of gel. All samples from induced cultures show successful protein overexpression not present in uninduced cultures.

4.2.2 BiFC Assay

The BiFC assay was carried out as described in Section 2.6.1, using the following cotransformations: pBAD-NmVenus + pRSF-CmVenus, pBAD-NmVenus-POLD3 + pRSF-CmVenus-NHelQ, pBAD-NmVenus-RPA1 + pRSF-CmVenus-NHelQ, and pBAD-NmVenus-RPA2 + pRSF-CmVenus-NHelQ. For the first two cotransformations, glycerol stocks produced from overnight cultures for the POLD3 Quincunx motif mutants BiFC assay were used rather than new cotransformations. SDS-PAGE analysis used all samples (Figure 4.10). The samples were electrophoresed on 10% (w/v) SDS-PAGE gels at 120 V for 1 hour and 10 minutes. NmVenus, NmVenus-POLD3, NmVenus-RPA1, and NmVenus-RPA2 were visible in the appropriate lanes, but CmVenus was nonvisible, and CmVenus-N-HELQ was nonvisible when co-overexpressed with NmVenus-POLD3 but not NmVenus-RPA1 and NmVenus-RPA2. A Western blot could not be performed for the nonvisible proteins due to time constraints.

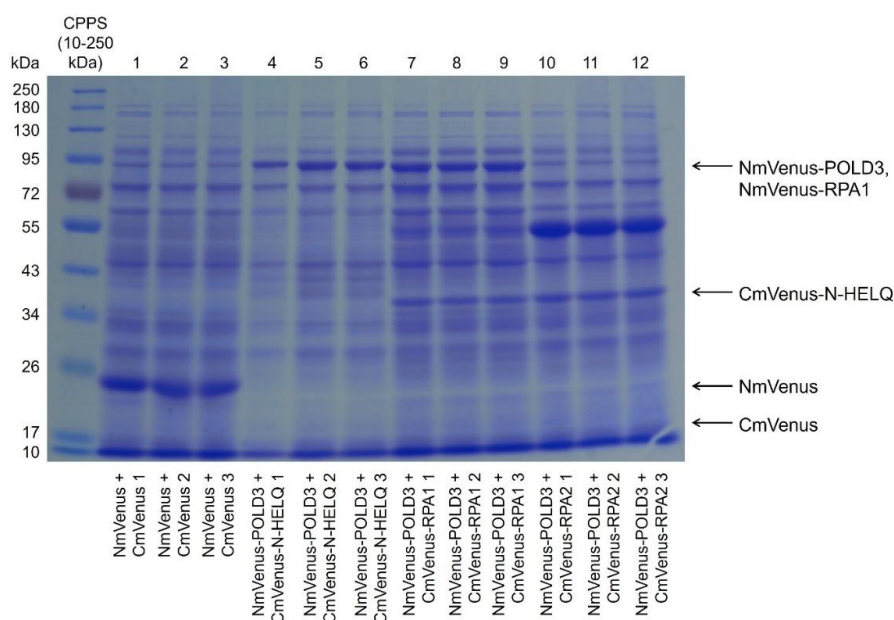


Figure 4.10 SDS-PAGE gel of RPA BiFC assay SDS-PAGE samples to verify protein overexpression.

SDS-PAGE samples were electrophoresed on 10% (w/v) SDS-PAGE gels at 120 V for 1 hour and 10 minutes alongside Color Prestained Protein Standard (CPPS), Broad Range (10-250 kDa) (NEB). CPPS molecular masses in kDa are shown to the left of the gels. Lane numbers are shown above the gels. The proteins overexpressed are shown below the corresponding lane. Expected overexpressed protein positions shown to the right of gel.

A spot assay was performed 10 days after cell samples were resuspended in 1X M9 medium, using 10^{-1} - 10^{-10} serial dilutions (Figure 4.11). Corresponding sample average CFU/mL are listed in Table 4.4.

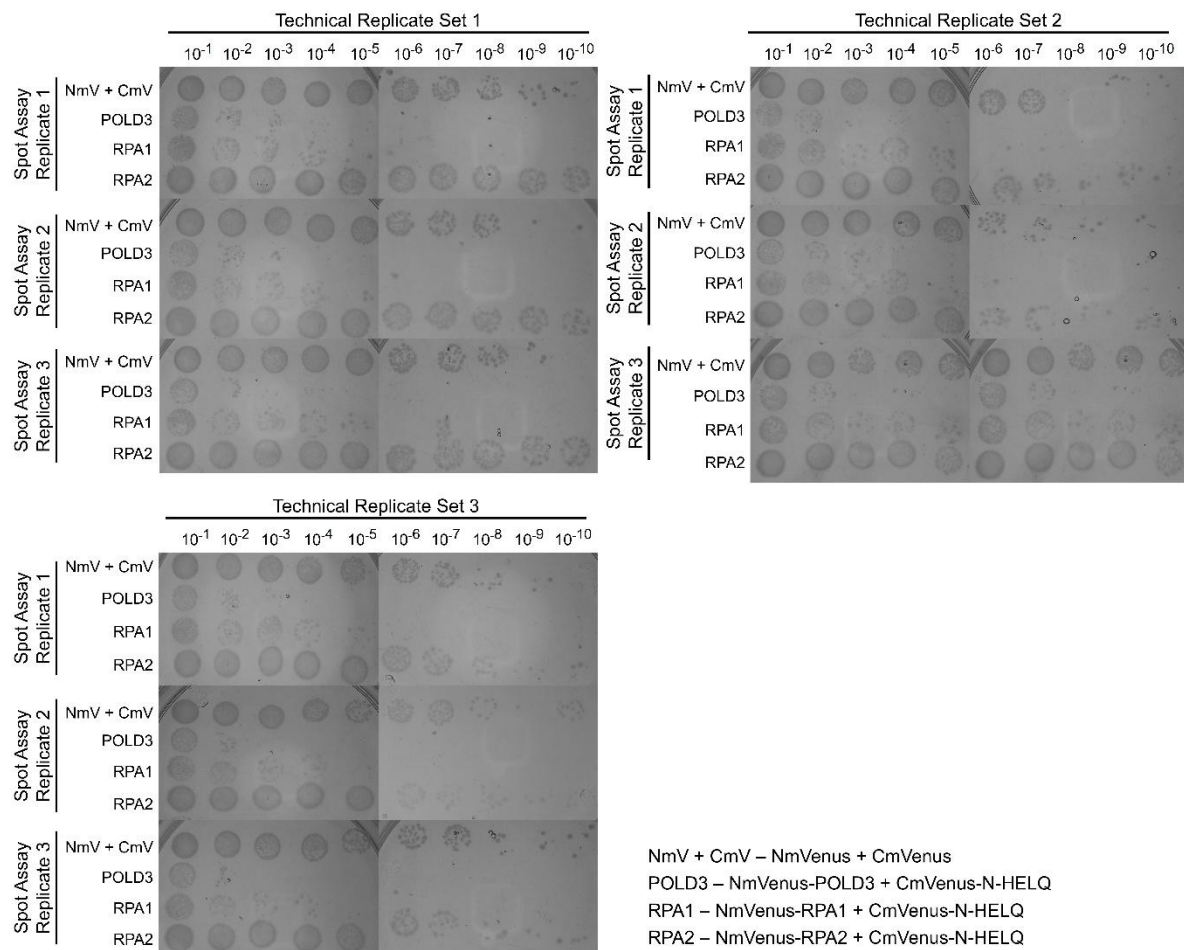


Figure 4.11 Spot assay of RPA1 and RPA2 BiFC assay BL21-AI *E. coli* cell samples to assess cell viability.

For each 1 mL sample replicate set, 10 μ L of each sample was diluted in 90 μ L of 1X M9 base salts for a 10^{-1} dilution, and this was serially diluted to 10^{-10} using 10 μ L of each previous dilution. 10 μ L of each serial dilution was spotted in triplicate on a Mu agar plate containing 100 μ g/mL ampicillin + 40 μ g/mL kanamycin. The proteins that were overexpressed in each sample are abbreviated to the left of the corresponding row, with the key at the bottom right.

Table 4.4 Average CFU/mL from RPA1 and RPA2 BiFC assay spot assay colony counts.

Proteins Overexpressed in BL21-AI <i>E. coli</i>	Technical Replicate	CFU/mL
NmVenus + CmVenus	1	3×10^{12}
	2	3×10^{12}
	3	9×10^{12}
NmVenus-POLD3 + CmVenus-N-HELQ	1	3×10^8
	2	6×10^8
	3	1×10^6
NmVenus-RPA1 + CmVenus-N-HELQ	1	3×10^8
	2	4×10^8
	3	6×10^8
NmVenus-RPA2 + CmVenus-N-HELQ	1	4×10^{13}
	2	3×10^{12}
	3	8×10^{10}

Colonies were counted using an SC6+ Colony Counter (Stuart) and rounded to the nearest whole number. Colony count averages used to calculate CFU/mL were based on highest dilution factor among a sample's triplicates that had at least 2 valid spots, otherwise the highest dilution factor with at least one valid spot was used.

Amersham Typhoon scans are shown in Figure 4.12. The 1X M9 minimal medium containing 100 µg/mL ampicillin + 40 µg/mL kanamycin showed negligible autofluorescence. Minimal fluorescence was observed in the NmVenus + CmVenus negative control. The NmVenus-POLD3 + CmVenus-N-HELQ positive control showed fluorescence greater than the negative control, while the NmVenus-RPA1 + CmVenus-N-HELQ and NmVenus-RPA2 + CmVenus-N-HELQ triplicates both showed a consistently lower fluorescence comparable to NmVenus + CmVenus, characteristic of no PPI.

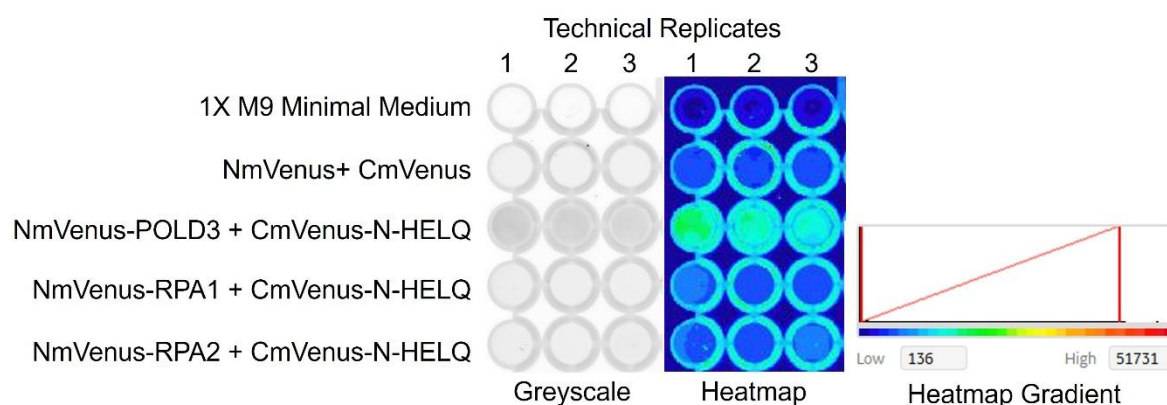


Figure 4.12 Amersham Typhoon 5 Biomolecular Imager (Cytiva) scan of RPA1 and RPA2 BiFC assay BL21-AI *E. coli* cell sample fluorescence to assess protein-protein interaction (PPI) between protein pairs.

Cells were suspended in 1X M9 minimal medium containing 100 µg/mL ampicillin + 40 µg/mL kanamycin. 400 µL of each sample was pelleted, 250 µL was removed, and cells were resuspended in remaining 150 µL, which was transferred to a clear flat-bottom non-tissue cultured 96-well plate. The plate was scanned using the built-in Cy2 filter. The scan was saved as both greyscale and heatmap, indicated below the scans. Overexpressed protein pairs are indicated to the left of the scans. Technical replicate set number is given above the scans. Heatmap gradient of colours corresponding to fluorescence intensity is shown on the right of the scans. Darker colour in greyscale indicates greater fluorescence intensity and greater fluorescence intensity indicates greater PPI.

The FLUOstar Omega fluorescence scan used a gain of 1828 and was normalised against OD₆₀₀ and CFU/mL; outliers of greater than 1 standard deviation (SD) were also removed from the CFU/mL normalisation (Figure 4.13). Two-tailed Welch's t-tests were performed using GraphPad Prism 10 to determine if sample triplicate mean fluorescence showed a significant difference compared to NmVenus + CmVenus, NmVenus-POLD3 + CmVenus-N-HELQ, and NmVenus-RPA1 + CmVenus-N-HELQ triplicate mean fluorescence. Full t-test statistics are listed in Appendix Table 7.4. When normalised by OD₆₀₀ and excluding repeat analyses, NmVenus + CmVenus fluorescence (n = 3, M = 25480, SD = 3489) and NmVenus-POLD3 + CmVenus-N-HELQ fluorescence (n = 3, M = 30866, SD = 7980) showed no significant difference

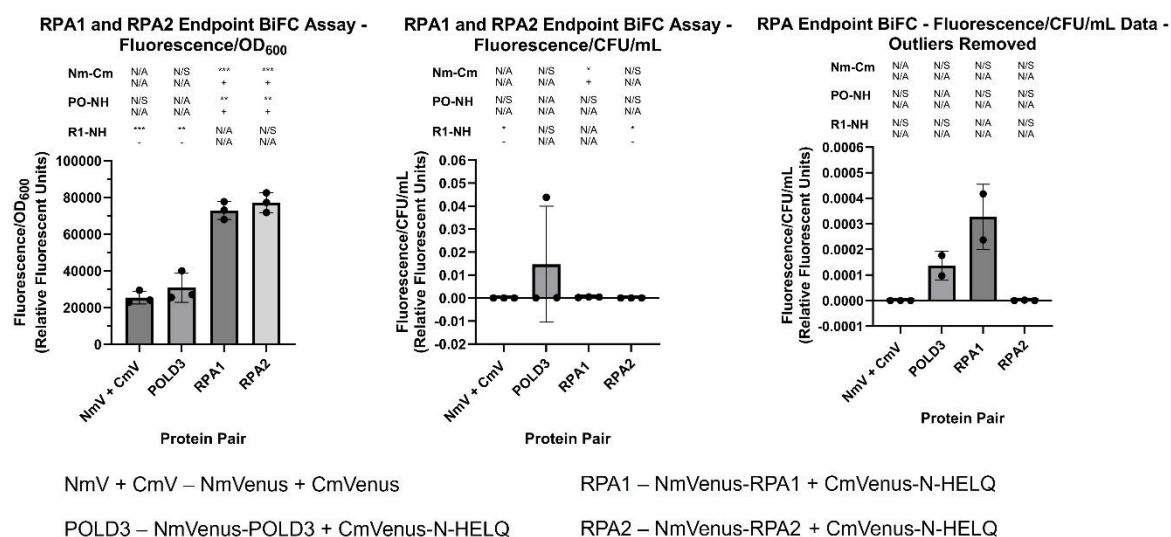


Figure 4.13 Normalised fluorescence values from FLUOstar Omega (BMG LABTECH) scan of RPA1 and RPA2 BiFC assay BL21-AI *E. coli* cell samples.

A gain of 1828 was used. Fluorescence was normalised by OD₆₀₀ (left) and colony forming units (CFU)/mL (middle). For fluorescence normalised by CFU/mL, results beyond 1 standard deviation were considered outliers and so removed (right). The column heights show mean and the error bars show $\pm$ one standard deviation. Two-tailed Welch's t-tests were performed using GraphPad Prism 10 to determine if triplicate means showed a significant difference compared to NmVenus + CmVenus (Nm-Cm), NmVenus-POLD3 + CmVenus-N-HELQ (PO-NH), and NmVenus-RPA1 + CmVenus-N-HELQ (R1-NH) fluorescence. Significance level is denoted by asterisks over the column: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$, N/S = 'no significance', and N/A = 'not applicable'. + indicates that that column protein pair has a higher mean than the comparative protein pair; - indicates it is lower.

compared to each other, but compared to NmVenus-RPA1 + CmVenus-N-HELQ fluorescence ($n = 3$, $M = 72927$, $SD = 4866$) both were significantly lower ($n = (t(3.626) = 13.72, p = 0.0003; t(3.307) = 7.794, p = 0.0031)$). This difference is also observed when both are compared to NmVenus-RPA2 + CmVenus-N-HELQ fluorescence ($n = 3$, $M = 77194$, $SD = 5389$) ($t(3.426) = 13.95, p = 0.0004; t(3.510) = 8.333, p = 0.002$). When normalising by CFU/mL, both NmVenus

+ CmVenus fluorescence ($n = 3$, $M = 1.177 \times 10^{-008}$, $SD = 4.661 \times 10^{-009}$) and NmVenus-RPA2 + CmVenus-N-HELQ fluorescence ($n = 3$, $M = 7.124 \times 10^{-007}$, $SD = 1.180 \times 10^{-006}$) are significantly lower compared to NmVenus-RPA1 + CmVenus-N-HELQ fluorescence ($n = 3$, $M = 0.0003897$, $SD = 0.0001410$) ($t(2.000) = 4.786$, $p = 0.0410$; $t(2.000) = 4.778$, $p = 0.0411$). When normalised by CFU/mL with outliers removed, no results showed a significant difference. Regardless of normalisation, the pattern of the protein pair triplicates' relative fluorescence levels measured using the FLUOstar Omega do not match that observed in the Amersham Typhoon scans.

4.3 PE3b Prime Editing of *HELQ* In U2OS Cells

4.3.1 PE3b Prime Editing Plasmid Construction

Following the BiFC assay to assess RPA1-N-HELQ interaction, it was considered that a native cell approach using RPA and HELQ may be more suitable. In previous work by the Bolt lab, Parkes (2024) successfully PE3b prime edited human U2OS cells to encode DD49^{D422A/D424A}. U2OS cells are commonly used to 'study DNA damage-induced cellular responses' in DNA repair research (Barta et al., 2023). Separately, Jenkins et al. (2021) found that N-HELQ PWI-like domain residues D142/F143 are necessary for HELQ to displace ssDNA-bound RPA. Therefore, it was decided to continue the prime editing work, now focusing on generating a *HELQ*^{D142A/F143A} U2OS cell line, thus giving the fourth aim of this work.

PE3b prime editing required generating stocks of three plasmids: SpyCas9 PE2* plasmid encoding the NLS-optimised nCas9-RT, a plasmid encoding pegRNA for *HELQ* D142A/F143A mutation (pAR003), and a plasmid constructed to encode sgRNA for *HELQ* D142A/F143A mutation (pAR004). Once the necessary oligonucleotides were designed and acquired (see Section 2.8.1), pAR003 and pAR004 were constructed according to methods described in Anzalone et al. (2019), via Golden Gate assembly (see Section 2.8.2) and restriction cloning (see Section 2.8.3), respectively, and verified through agarose gel electrophoresis and Sanger sequencing. Plasmids were subsequently maxipreped (see Section 2.8.4), again verified

through agarose gel electrophoresis. Electrophoresed agarose gels to verify plasmid sizes are shown in Figure 4.14.

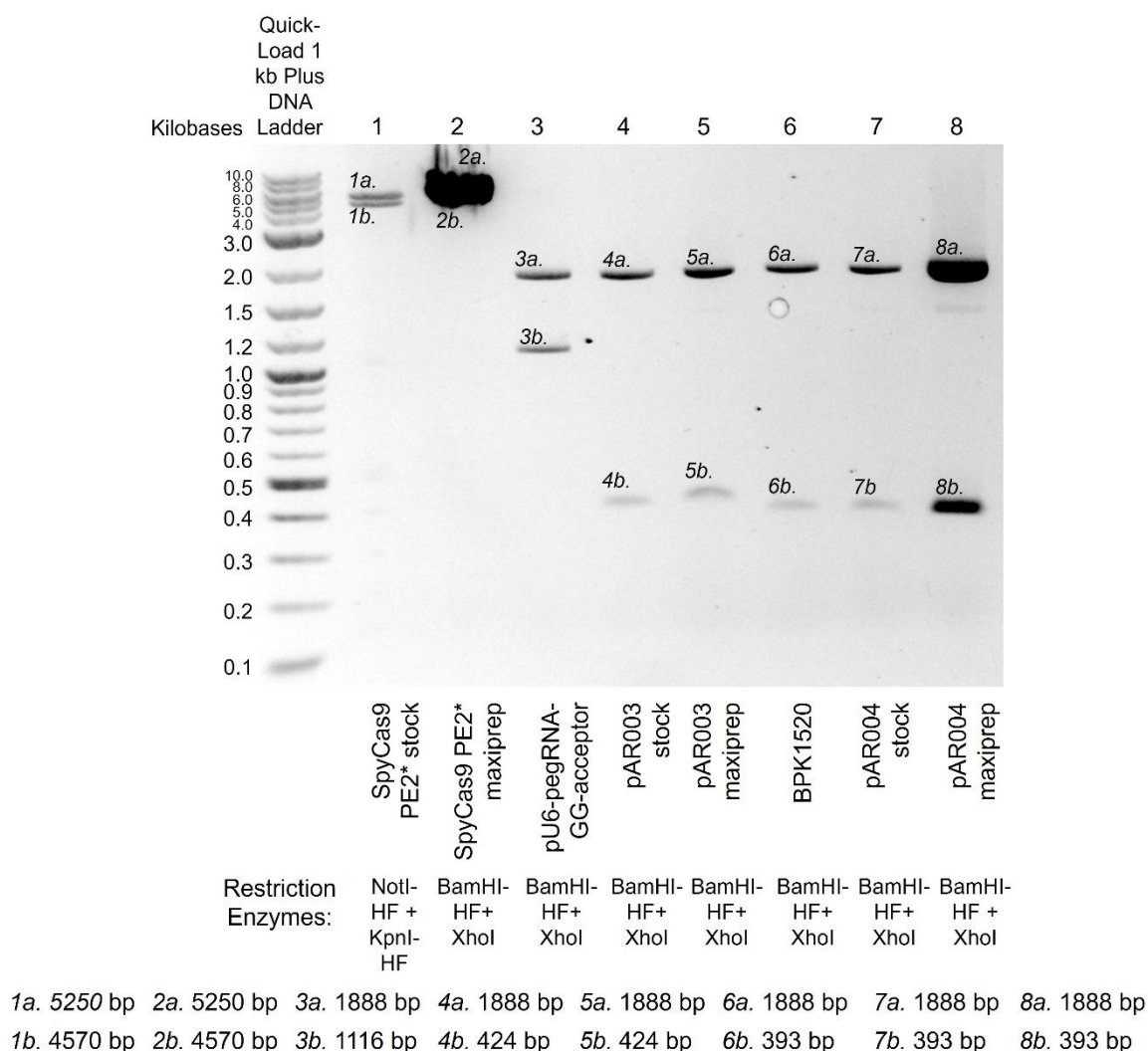


Figure 4.14 Agarose gel of electrophoresed analytical restriction digests of prime editing plasmid stocks and maxipreps verifying plasmid lengths.

Double analytical restriction digests were electrophoresed for on a 1% (w/v) agarose gel alongside Quick-Load 1 kb Plus DNA Ladder (NEB1) for hour at 120 V. Image was inverted and brightness/contrast adjusted using the software ImageJ. Ladder DNA lengths in kilobases are shown to the left of the gel. Lane numbers are shown are above the gel. Digested plasmids and restriction enzymes used in analytical digests are shown below the corresponding lane. Plasmid fragment lengths in base pairs (bp) are shown below the gel, corresponding to lane number and letter.

4.3.2 U2OS Cell Transfection and Genomic Analysis

Once the necessary plasmids were maxiprepmed, 100,000 cells were seeded in 2 mL medium in each well of a clear flat-bottom non-TC 24-well plate and incubated for 24 hours (see Section 2.8.5). The cells were then transfected by Lipofectamine 3000 Transfection Reagent (Invitrogen) as described in Anzalone et al. (2019). Cells were lipofected using either 0.75 μ L or 1.5 μ L Lipofectamine, with plasmid contents varying plasmid contents. Purified maxiprepmed pEGFP-C1 plasmid DNA alone served as a positive control to verify successful transfection; no plasmid DNA or purified maxiprepmed SpyCas9 PE2*, pAR003, or pAR004 plasmid DNA alone served as negative controls; and all three prime editing plasmids were used for the prime editing itself. Cells were imaged pre-transfection and at 26 hours, 46 hours, and 71 hours post-transfection using a ZOE Fluorescent Cell Imager (Bio-Rad). Cell images are shown in Figure 4.15. Cells transfected with pEGFP-C1 showed green fluorescence after 26 hours, indicating successful lipofection. After 71 hours, there were generally greater numbers of nonviable cells in samples transfected using 1.5 μ L Lipofectamine compared with 0.5 μ L Lipofectamine, especially apparent in negative control samples transfected with SpyCas9 PE2* or pAR003 alone and the test samples.

Genomic DNA was extracted from cell samples transfected with all three prime editing plasmids using the PureLink Genomic DNA Mini Kit (Invitrogen) from one sample in each duplicate after 46 hours and the second samples after 71 hours (see Section 2.8.6). The extracted genomic DNA was then PCR amplified and electrophoresed on an agarose gel to verify its length (Figure 4.16). The PCR products were then isolated by gel extraction from another agarose gel and Sanger sequenced. Prime editing success was assessed using the online tool CRISPResso2 (<https://crispresso2.pinellolab.org/submission>) (Clement et al., 2019). The prime editing was ultimately unsuccessful (Figure 4.17).

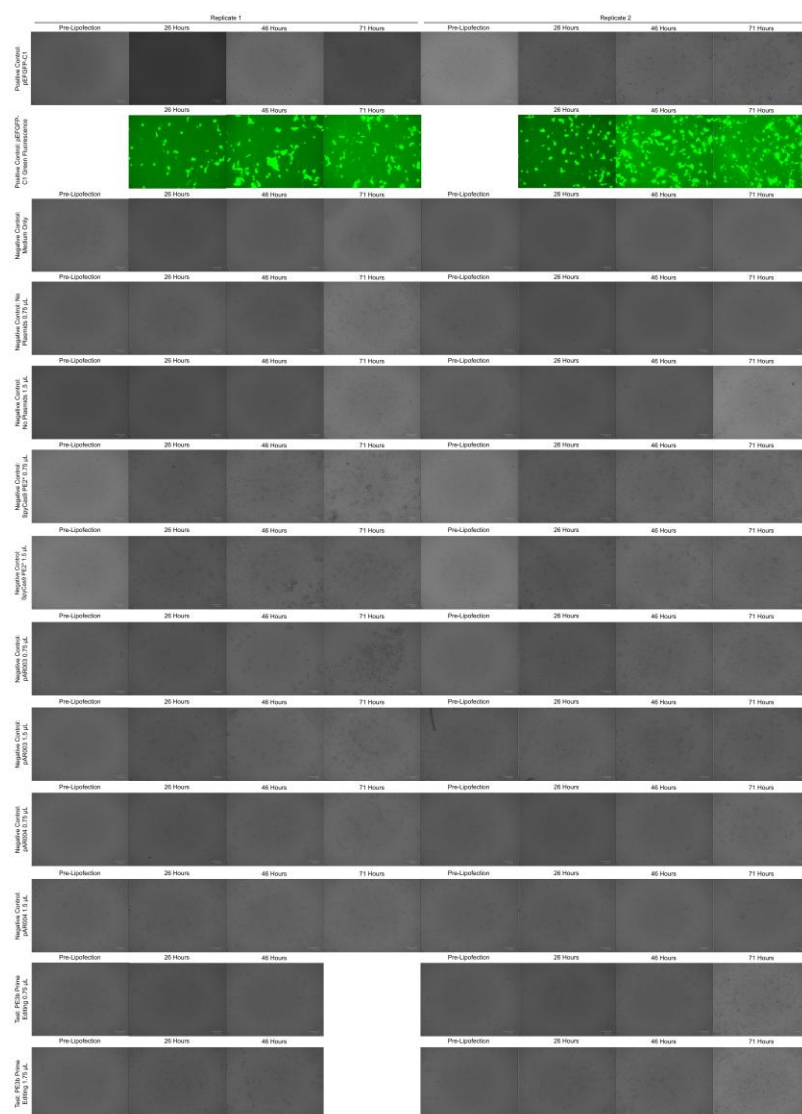


Figure 4.15 ZOE Fluorescent Cell Imager (Bio-Rad) images of human U2OS cells pre- and post-transfection by lipofection for PE3b prime editing to generate a *HELQ*^{D142A/F143A} mutant cell line.

Cells were imaged as TIFs pre-transfection and at 26 hours, 46 hours, and 71 hours post-transfection. The volume of Lipofectamine 3000 Transfection Reagent (Invitrogen) used, plasmids transfected, replicate number, and time point are shown above the respective images. Scale bars show 100 μ m. Cells were transfected with pEGFP-C1 encoding a green fluorescent protein as a positive control to verify successful lipofection. Negative controls were lipofected with one of three prime editing plasmids SpyCas9 PE2*, pAR003, and pAR004 alone or no plasmids. Test sample lipofections for used all three prime editing plasmids.

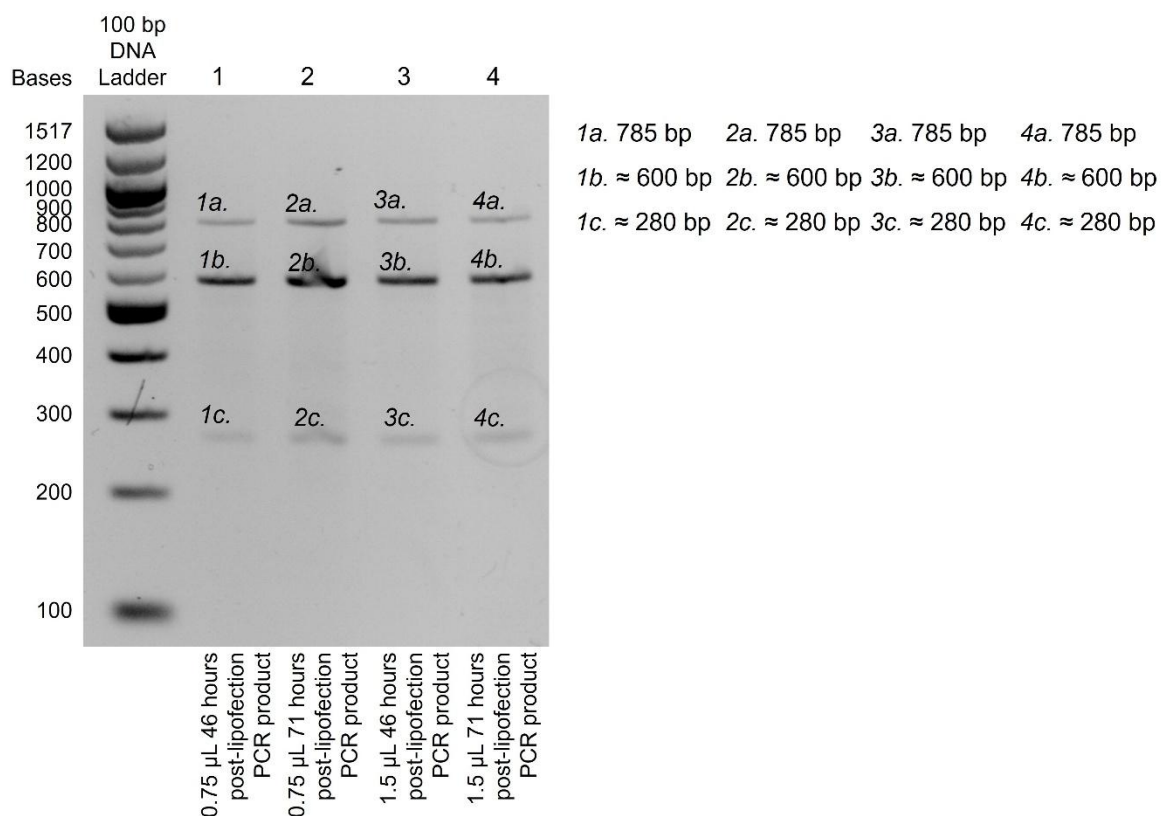


Figure 4.16 Agarose gel of electrophoresed purified PE3b prime editing U2OS cell genomic DNA PCR products.

PCR products were electrophoresed for on a 2% (w/v) agarose gel alongside 100 bp DNA Ladder (NEB1) for 1 hour and 15 minutes at 120 V. Image was inverted and brightness/contrast adjusted using the software ImageJ. Ladder DNA lengths in bases are shown to the left of the gel. Lane numbers are shown above the gel. The volume of Lipofectamine 3000 Transfection Reagent (Invitrogen) used in prime editing transfection and the timepoint of genomic DNA extraction for the respective genomic DNA PCR templates are shown below the corresponding lane. PCR product DNA lengths in base pairs (bp) are shown to the right of the gel, corresponding to lane number and letter. Lengths given as approximations ($\approx$) indicate bands of unexpected length, otherwise they are desired bands of expected length.

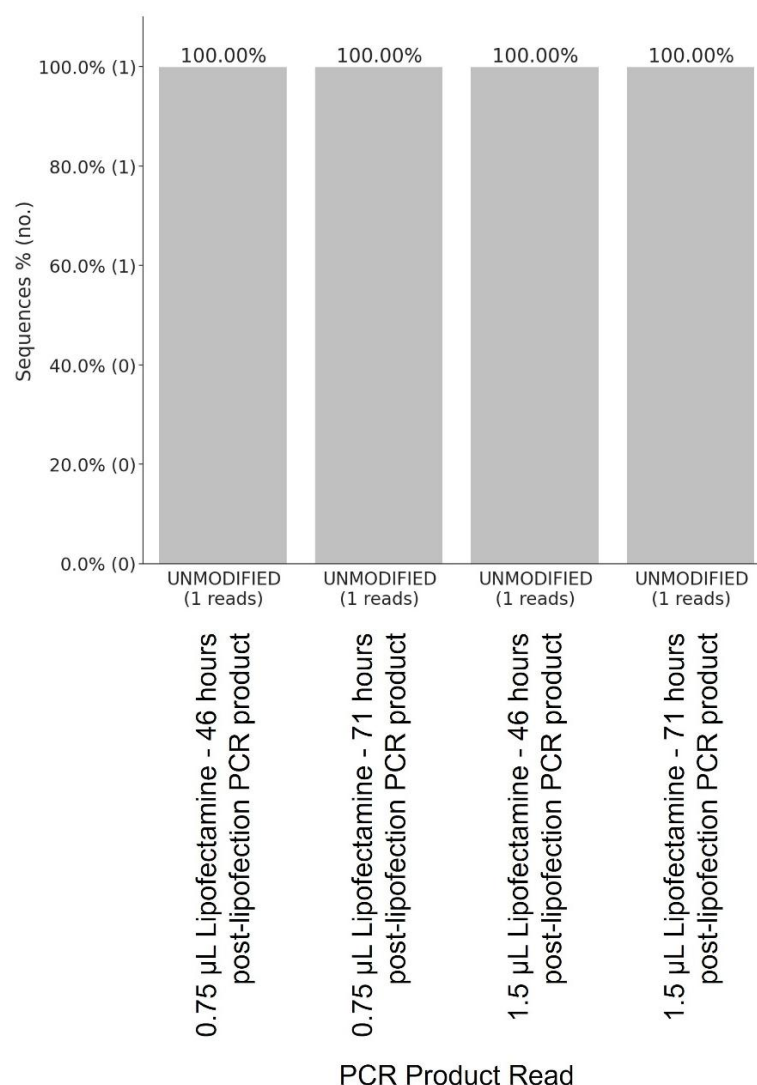


Figure 4.17 Merged CRISPResso2 allele assignment bar plot to assess prime editing success.

Prime editing single-end read sequencing analysis was performed using the online tool CRISPResso2 to compare PE3b prime editing genomic DNA PCR product FASTQ files with expected PCR amplicons containing successful prime edits (Clement et al., 2019). The DNA sequence 5'-TATAGCTCAAGTTGACGACC-3' from pAR004 that encodes the sgRNA was used to infer editing activity position. The PCR product DNA sequences being analysed are listed below the bar plot. The merged bar plot provides an overview on the number of sequence reads that are modified or unmodified sorted by alignment and editing frequency, showing the percentage of reads that are modified or unmodified. For a single read, a result of unmodified indicates that prime editing was unsuccessful, as is the case for the sequenced PCR products.

4.4 Summary

Previous Bolt lab work in Jenkins et al. (2021) used EMSAs and affinity pull-down assays show RPA1-HELQ interaction but could not detect stable RPA-N-HELQ interaction. Since at that point the He et al. (2023) mVenus time-course BiFC assay method had not yet been developed, this interaction was investigated using the endpoint BiFC assay method developed in this work. DNA encoding all three RPA subunits was successfully PCR amplified from p11d-tRPA(123), and subsequent restriction cloning was able to successfully generate pBAD-NmVenus-RPA1 (pAR001) and pBAD-NmVenus-RPA2 (pAR002), but not pBAD-NmVenus-RPA3. Pilot protein expression showed that both NmVenus-RPA1 and NmVenus-RPA2 successfully overexpress in BL21-AI *E. coli*. In a BiFC assay to assess PPI between NmVenus-RPA1 or NmVenus-RPA2 and CmVenus-N-HELQ, no interaction between either pair was detected from the relative fluorescence levels observed in an Amersham Typhoon biomolecular imager scan. However, the pattern observed in the Amersham Typhoon scan was not reflected in relative fluorescence values normalised using sample OD₆₀₀ or CFU/mL when scanned using the FLUOstar Omega plate reader.

Jenkins et al. (2021) had also found that N-HELQ PWI-like domain residues D142/F143 are necessary for HELQ to displace ssDNA-bound RPA, while Additional previous Bolt lab work by Parkes (2024) successfully PE3b prime edited human U2OS cells to encode DD49^{D422A/D424A}. Therefore, in order to pursue a native cell approach and prime edit a *HELQ*^{D142A/F143A} mutation in U2OS cells, plasmids were generated encode the necessary pegRNA (pAR003) and sgRNA (pAR004), via Golden Gate assembly and restriction cloning, respectively. These were maxiprepmed alongside SpyCas9 PE2* and transfected into U2OS cells via lipofection. Analysis of extracted genomic DNA PCR products via CRISPResso2 showed that the prime editing was unsuccessful. To investigate further RPA-HELQ interaction using prime editing and generate a U2OS *HELQ*^{D142A/F143A} cell line, a more efficient prime editing system would be needed to maximise the probability of successful edits, such as PE7 (Lee et al., 2025). Generating this cell line would hence allow investigating whether a *HELQ*^{D142A/F143A} mutation inhibits ssDNA-bound RPA displacement *in vivo*. As in the case of POLD3-N-HELQ interaction, this could hence

further the understanding of the molecular mechanisms behind the interaction and ultimately how HELQ limits mutagenic DNA synthesis during BIR (He et al., 2023).

Chapter 5: Discussion and Future Work

The aims of this project were fourfold. First, developing the original He et al. (2023) time-course method into an endpoint method to provide quantitative data. Second, to use this developed method to investigate further the molecular mechanisms of POLD3-N-HELQ interaction to locate potential sites of POLD3-N-HELQ interaction. Third, to determine if the endpoint method could be used to show RPA1-N-HELQ interaction and hence locate potential interaction sites in RPA1. Fourth and finally, to attempt to assess full-length protein RPA-HELQ interaction via PE3b prime editing of U2OS cells.

5.1 BiFC Assay Method Development

By developing upon the He et al. (2023) BiFC method and the Killelea method, the endpoint BiFC method used in this work combines their qualitative and quantitative strengths, enabling a more reliable and factual result than if using only one of the two (Horstman et al., 2014; Pyo et al., 2023; Verhoef et al., 1997). This method can hence be used to determine if any two proteins that can be overexpressed in BL21-AI *E. coli* can interact, and through the use of point mutations determine which specific residues are necessary for PPI, even in a transient interaction (Kerppola, 2006a; Kerppola, 2006b). However, a significant limitation of this method is that its endpoint nature of means it only provides a single temporal snapshot, while a time-course method like the original He et al. (2023) method allows observation of fluorescence intensity change over time (Gujral et al., 2020; Kerppola, 2009; Weinreb et al., 2018). Data gathered in Section 3.4 shows that when scanning the Mu broth used in this method using FLUOstar Omega filters 500/520 nm, its yellow relative autofluorescence is so high that it would likely drown out any reconstituted mVenus fluorescence (Figure 3.19), preventing observation of any fluorescence change over time (Bhakta et al., 2024; Jiménez-Nava et al., 2023; Kerppola, 2009; Masuda et al., 2010; Smith et al., 2024).

Conversely, the Mu broth has a far lower red autofluorescence when using 590-50/670 nm filters (Figure 3.20). Hence, to further develop this method into a time-course, a far-red FP would be needed (Miller et al., 2015; Shang et al., 2024). Based upon the filters, an ideal

choice of FP could be mNeptune: it has excitation/emission maxima 600/650 nm, no maturation issues at 37 °C, and has previously been used in human HeLa cell BiFC assays (Han et al., 2014; Miller et al., 2015). Alternatively, a later variant such as mCardinal could be used; this has 604/659 nm maxima and is 50% brighter than mNeptune, allowing it to overcome Mu broth autofluorescence even more than mNeptune (Chu et al., 2014; FPbase, 2026). It should also be noted that rather than normalising by OD600, OD700 measurements may be required instead due to the far-red FP potentially falsely increasing cell density measurements when using OD600 (Hecht et al., 2016). Additionally, Amersham Typhoon scans would need to use the built-in Cy5 or IRshort filter depending on the choice of FP (Figure 3.20). Hence, by further developing this method into an optimised time-course approach, it would allow for even more detailed investigations into the mechanisms of PPI.

5.2 Characterising POLD3-N-HELQ Interaction Using BiFC Assays

While previous Bolt lab work by Lou-Hing (2024) showed that by mutating the POLD3 NLS polar, charged Quincunx motif (residues 319-KKRRR-323) to polar, uncharged glycines (319-GGGGG-323) inhibited POLD3-N-HELQ interaction, data gathered in Section 3.2 shows that this is not the case (Figures 3.8-3.11, Table 3.3) (JPT, 2025a; JPT, 2025b; Satoh et al., 2022). In fact, rather than a single POLD3 residue tract being responsible for POLD3-N-HELQ interaction, data gathered in Section 3.3 indicates that POLD3 IDPR2, IDPR3, and IDPR4 all show interaction with N-HELQ, with the highest coming from IDPR3, hence supporting Haskins' preliminary data (Figures 3.12-3.15, Table 3.4). This therefore implies that POLD3 has multiple N-HELQ binding sites, and so is multivalent in its interaction to greatly increase its avidity for N-HELQ, as characteristic of a protein with multiple IDPRs (Deng et al., 2022; Fung et al., 2018; Konc et al., 2022; Mammen et al., 1998; Mirica et al., 2012; Nooren et al., 2003; Schrieber et al., 2011; Teilum et al., 2021). Furthermore, based upon He et al. (2023) requiring a BiFC assay to show POLD3-N-HELQ interaction after several less sensitive methods could not, this may also indicate that this is a transient, and therefore non-obligate, interaction, again as characteristic of intrinsically disordered proteins (La et al., 2013; Peng et al., 2017; Wright et al., 2015).

As mentioned in Section 3.5, to investigate further POLD3-N-HELQ PPI using BiFC assays, residues of interest would need to be identified in POLD3 IDPRs 2, 3, and 4, using software such as AlphaFold 3 in order for mutants to be generated (Abramson et al., 2024). Furthermore, three NmVenus-POLD3^{IDPR3} mutants would need to be generated, one each with the Quincunx motif or motif in the PNLs mutated to glycines and one with both mutations (Alli et al., 2024; Lou-Hing, 2024). This would allow investigating whether one or both of these regions contribute to the PPI. However, spot assays would need to be performed the next day after cell samples were resuspended in 1X M9 minimal medium due to the risk of entering the VBNC state, hence greatly reducing the accuracy of normalising relative fluorescence by CFU/mL (Figures 3.16-3.18, Table 3.5) (Dağ et al., 2025; Sleight et al., 2006; Thean et al., 2021; Wei et al., 2018; Yang, D. et al., 2024). Furthermore, to fully understand the molecular mechanisms of POLD3-N-HELQ interaction, a variety of additional atomic resolution, mass spectrometry, biophysical, and computational methods will be required, contributing to ultimately fully understanding how HELQ limits mutagenic DNA synthesis during BIR (He et al., 2023; Seychell et al., 2021).

5.3 Characterising RPA1-N-HELQ Interaction Using BiFC Assays

In previous work by the Bolt lab, Jenkins et al. (2021) were able to show RPA1-HELQ interaction but could not detect stable RPA-N-HELQ interaction. Therefore, in order to determine if RPA1-N-HELQ interaction could be observed using the BiFC assay method used in this work, DNA encoding all three RPA subunits was successfully PCR amplified from p11d-tRPA(123) and used in restriction cloning to generate pBAD-NmVenus-RPA1 (pAR001) and pBAD-NmVenus-RPA2 (pAR002) (Figure 4.8). Pilot protein expression showed that both NmVenus-RPA1 and NmVenus-RPA2 successfully overexpress in BL21-AI *E. coli* (Figure 4.9). In the BiFC assay, neither NmVenus-RPA1 nor NmVenus-RPA2 showed interaction with CmVenus-N-HELQ when scanned using the Amersham Typhoon (Figure 4.12), but did appear to show interaction when relative fluorescence was normalised by sample OD₆₀₀ and CFU/mL, thus the data is inconclusive. In addition, the NmVenus-POLD3 + CmVenus-N-HELQ positive control did not

show interaction when its relative fluorescence was normalised (Figure 4.12). This may be due to reduced protein yield from pre-stress to the BL21-AI *E. coli* caused by them being from a bacterial glycerol stock rather than a fresh cotransformation combined with the stress of being stored at 4 °C in 1X M9 minimal medium, thus in future BiFC assays using the method in this work should use fresh BL21-AI *E. coli* cotransformations for all protein pairs (Dağ et al., 2025; Sleight et al., 2006; Thean et al., 2021; Wei et al., 2018).

5.4 Attempting To Prime Edit U2OS Cells

In previous work by the Bolt lab, Parkes (2024) successfully PE3b prime edited human U2OS cells to encode DD49^{D422A/D424A}. Separately, Jenkins et al. (2021) found that N-HELQ PWI-like domain residues D142/F143 are necessary for HELQ to displace ssDNA-bound RPA. Therefore, a plasmid encoding pegRNA for *HELQ*^{D142A/F143A} mutation (pAR003), and a plasmid constructed to encode sgRNA for *HELQ*^{D142A/F143A} mutation (pAR004) were successfully generated and maxiprepped alongside SpyCas9 PE2* (Figure 4.14) for transfection by lipofection into U2OS cells to generate a U2OS *HELQ*^{D142A/F143A} mutant cell line (Figure 4.15). However, the prime editing was ultimately unsuccessful (Figures 4.16 and 4.17). Therefore, as mentioned in Section 4.4, in any future attempts to prime edit this cell line, a more efficient prime editing system would be needed to maximise the probability of successful edits, such as PE7 (Lee et al., 2025). This is currently the most efficient prime editing system due to fusing the La protein N-terminal domain (residues 1-194) to the nCas9-RT, thus stabilising the pegRNAs from degradation. Hence, by using this system, if a U2OS *HELQ*^{D142A/F143A} mutant cell line is generated, this can then be cultured in DNA-damaging mitomycin C. If these cells show a phenotype of reduced viability matching that of *HELQ* knockout U2OS cells generated by Cubbon (2022), it may indicate that N-HELQ PWI-like domain residues D142/F143 are indeed necessary for HELQ to displace ssDNA-bound RPA. Like in the case of investigating POLD3-N-HELQ interaction, this can thus contribute to ultimately fully understanding how HELQ limits mutagenic DNA synthesis during BIR (He et al., 2023; Jenkins et al., 2021; Seychell et al., 2021).

Chapter 6: References

Abou Assi, H., Garavís, M., González, C. et al. (2018), 'i-Motif DNA: structural features and significance to cell biology', *Nucleic Acids Research*, **46**(16), pp. 8038–8056. DOI: <https://doi.org/10.1093/nar/gky735>

Abramson, J., Adler, J., Dunger, J. et al. (2024), 'Accurate structure prediction of biomolecular interactions with AlphaFold 3', *Nature*, **630**(8016), pp. 493–500. DOI: <https://doi.org/10.1038/s41586-024-07487-w>

Addgene (2002), *pEGFP-C1 Vector Information*. Available at: https://media.addgene.org/data/plasmids/58/58467/58467-map_ifN0i9frczKJ.pdf.

Accessed: 24/8/25

Addgene (2025a), *BPK1520*. Available at: <https://www.addgene.org/65777/>. Accessed: 18/8/25

Addgene (2025b), *p11d-tRPA(123)*. Available at: <https://www.addgene.org/102613/>. Accessed: 28/8/25

Addgene (2025c), *pU6-pegRNA-GG-acceptor*. Available at: <https://www.addgene.org/132777/>. Accessed: 18/8/25

Addgene (2025d), *SpyCas9 PE2**. Available at: <https://www.addgene.org/169850/>. Accessed: 18/8/25

Adiciptaningrum, A., Osella, M., Moolman, M.C. et al. (2015), 'Stochasticity and homeostasis in the *E. coli* replication and division cycle', *Scientific Reports*, **5**, pp. 18261. DOI: <https://doi.org/10.1038/srep1826>

Ahlquist, P. (2006), 'Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses', *Nature Reviews. Microbiology*, **4**(5), pp. 371–382. DOI: <https://doi.org/10.1038/nrmicro1389>

Alberts, B., Heald, R., Johnson, A.D. et al. (2022), *Molecular Biology of the Cell*, 7th edition, International student edition. New York (NY): W.W. Norton & Company, Inc.

Alhmoud, J.F., Woolley, J.F., Al Moustafa, A.E. et al. (2020), 'DNA Damage/Repair Management in Cancers', *Cancers*, **12**(4), pp. 1050. DOI: <https://doi.org/10.3390/cancers12041050>

Allen, J.M., Simcha, D.M., Ericson, N.G. et al. (2011), 'Roles of DNA polymerase I in leading and lagging-strand replication defined by a high-resolution mutation footprint of ColE1 plasmid replication', *Nucleic Acids Research*, **39**(16), pp. 7020–7033. DOI: <https://doi.org/10.1093/nar/gkr157>

Alli, N., Lou-Hing, A., Bolt, E.L. et al. (2024), 'POLD3 as Controller of Replicative DNA Repair', *International Journal of Molecular Sciences*, **25**(22), 12417. DOI: <https://doi.org/10.3390/ijms252212417>

Anand, R., Buechelmaier, E., Belan, O. (2022), 'HELQ is a dual-function DSB repair enzyme modulated by RPA and RAD51', *Nature*, **601**(7892), pp. 268–273. DOI: <https://doi.org/10.1038/s41586-021-04261-0>

Anders, C., and Jinek, M. (2014), 'In vitro enzymology of Cas9', *Methods in Enzymology*, **546**, pp. 1–20. DOI: <https://doi.org/10.1016/B978-0-12-801185-0.00001-5>

Anzalone, A.V., Randolph, P.B., Davis, J.R. et al. (2019), 'Search-and-replace genome editing without double-strand breaks or donor DNA', *Nature*, **576**(7785), pp. 149–157. DOI: <https://doi.org/10.1038/s41586-019-1711-4>

Ao, C., Jiao, S., Wang, Y. et al. (2022). 'Biological Sequence Classification: A Review on Data and General Methods', *Research (Washington, D.C.)*, 2022, 0011. DOI: <https://doi.org/10.34133/research.0011>

Aparicio, T., Ibarra, A., and Méndez, J. (2006), 'Cdc45-MCM-GINS, a new power player for DNA replication', *Cell Division*, **1**, 18. DOI: <https://doi.org/10.1186/1747-1028-1-18>

Ashley, R.E., Dittmore, A., McPherson, S.A. (2017), 'Activities of gyrase and topoisomerase IV on positively supercoiled DNA', *Nucleic Acids Research*, **45**(16), pp. 9611–9624. DOI: <https://doi.org/10.1093/nar/gkx649>

Asmamaw, M., and Zawdie, B. (2021), 'Mechanism and Applications of CRISPR/Cas-9-Mediated Genome Editing', *Biologics: Targets & Therapy*, **15**, pp. 353–361. DOI: <https://doi.org/10.2147/BTT.S326422>

- Atolia, E., Cesar, S., Arjes, H.A et al. (2020), 'Environmental and Physiological Factors Affecting High-Throughput Measurements of Bacterial Growth', *mBio*, **11**(5), e01378-20. DOI: <https://doi.org/10.1128/mBio.01378-20>
- Ausiannikava, D., and Allers, T. (2017), 'Diversity of DNA Replication in the Archaea', *Genes*, **8**(2), pp. 56. DOI: <https://doi.org/10.3390/genes8020056>
- Austin, E., Geisler, A.N., Nguyen, J. et al. (2021), 'Visible light. Part I: Properties and cutaneous effects of visible light', *Journal of the American Academy of Dermatology*, **84**(5), pp. 1219–1231. DOI: <https://doi.org/10.1016/j.jaad.2021.02.048>
- Azatian, S.B., Kaur, N., and Latham, M.P. (2019), 'Increasing the buffering capacity of minimal media leads to higher protein yield', *Journal of Biomolecular NMR*, **73**(1-2), pp. 11–17. DOI: <https://doi.org/10.1007/s10858-018-00222-4>
- Bahadur, R.P., and Zacharias, M. (2008), 'The interface of protein-protein complexes: analysis of contacts and prediction of interactions', *Cellular and Molecular Life Sciences: CMLS*, **65**(7-8), pp. 1059–1072. DOI: <https://doi.org/10.1007/s00018-007-7451-x>
- Balakrishnan, L., and Bambara, R.A. (2011), 'Eukaryotic lagging strand DNA replication employs a multi-pathway mechanism that protects genome integrity', *The Journal of Biological Chemistry*, **286**(9), pp. 6865–6870. DOI: <https://doi.org/10.1074/jbc.R110.209502>
- Balakrishnan, L., and Bambara, R.A. (2013), 'Okazaki fragment metabolism', *Cold Spring Harbor Perspectives in Biology*, **5**(2), a010173. DOI: <https://doi.org/10.1101/cshperspect.a010173>
- Baltimore, D., Eggers, H.J., Franklin, R.M. et al. (1963), 'Poliovirus-induced RNA polymerase and the effects of virus-specific inhibitors on its production', *Proceedings of the National Academy of Sciences of the United States of America*, **49**(6), pp. 843–849. DOI: <https://doi.org/10.1073/pnas.49.6.843>
- Baltimore, D. (1970), 'RNA-dependent DNA polymerase in virions of RNA tumour viruses', *Nature*, **226**(5252), pp. 1209–1211. DOI: <https://doi.org/10.1038/2261209a0>

Baranovskiy, A.G., Babayeva, N.D., Liston, V.G. et al. (2008), 'X-ray structure of the complex of regulatory subunits of human DNA polymerase delta', *Cell Cycle (Georgetown, Tex.)*, **7**(19), pp. 3026–3036. DOI: <https://doi.org/10.4161/cc.7.19.6720>

Barbieri, M. (2016), 'What is information?' *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, **374**(2063), 20150060. DOI: <https://doi.org/10.1098/rsta.2015.0060>

Barrangou, R., Fremaux, C., Deveau, H. et al. (2007), 'CRISPR provides acquired resistance against viruses in prokaryotes', *Science (New York, N.Y.)*, **315**(5819), pp. 1709–1712. DOI: <https://doi.org/10.1126/science.1138140>

Barta, N., Ördög, N., Pantazi, V. et al. (2023), 'Identifying Suitable Reference Gene Candidates for Quantification of DNA Damage-Induced Cellular Responses in Human U2OS Cell Culture System', *Biomolecules*, **13**(10), pp. 1523. DOI: <https://doi.org/10.3390/biom13101523>

Bartlett, E.J., Brissett, N.C., and Doherty, A.J. (2013), 'Ribonucleolytic resection is required for repair of strand displaced nonhomologous end-joining intermediates', *Proceedings of the National Academy of Sciences of the United States of America*, **110**(22), pp. E1984–E1991. DOI: <https://doi.org/10.1073/pnas.1302616110>

Batista, M., Langendijk-Genevaux, P., Kwapisz, M. et al. (2024), 'Evolutionary and functional insights into the Ski2-like helicase family in Archaea: a comparison of *Thermococcales* ASH-Ski2 and Hel308 activities', *NAR Genomics and Bioinformatics*, **6**(1), lqae026. DOI: <https://doi.org/10.1093/nargab/lqae026>

Batty, D.P., and Wood, R.D. (2000), 'Damage recognition in nucleotide excision repair of DNA', *Gene*, **241**(2), pp. 193–204. DOI: [https://doi.org/10.1016/s0378-1119\(99\)00489-8](https://doi.org/10.1016/s0378-1119(99)00489-8)

Bayani, J., Zielenska, M., Pandita, A. et al. (2003), 'Spectral karyotyping identifies recurrent complex rearrangements of chromosomes 8, 17, and 20 in osteosarcomas', *Genes, Chromosomes & Cancer*, **36**(1), pp. 7–16. DOI: <https://doi.org/10.1002/gcc.10132>

Beagan, K., Armstrong, R.L., Witsell, A. et al. (2017), 'Drosophila DNA polymerase theta utilizes both helicase-like and polymerase domains during microhomology-mediated end joining and interstrand crosslink repair', *PLOS Genetics*, **13**(5), e1006813. DOI: <https://doi.org/10.1371/journal.pgen.1006813>

- Beal, J., Farny, N.G., Haddock-Angelli, T., et al. (2020), 'Robust estimation of bacterial cell count from optical density', *Communications Biology*, **3**(1), pp. 512. DOI: <https://doi.org/10.1038/s42003-020-01127-5>
- Beattie, T.R., and Reyes-Lamothe, R. (2015), 'A Replisome's journey through the bacterial chromosome', *Frontiers in Microbiology*, **6**, pp. 562. DOI: <https://doi.org/10.3389/fmicb.2015.00562>
- Bell, S.P., and Stillman, B. (1992), 'ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex', *Nature*, **357**(6374), pp. 128–134. DOI: <https://doi.org/10.1038/357128a0>
- Benjamini, Y., Krieger, A.M., and Yekutieli, D. (2006), 'Adaptive linear step-up procedures that control the false discovery rate', *Biometrika*, **93**(3), pp. 491–50. DOI: <https://doi.org/10.1093/biomet/93.3.491>
- Beucher, A., Birraux, J., Tchouandong, L. et al., (2009), 'ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2', *The EMBO Journal*, **28**(21), pp. 3413–3427. DOI: <https://doi.org/10.1038/emboj.2009.276>
- Bhakta, S., and Zong, D. (2024), *M9 Medium - Bennett Lab Wiki - Rice University Campus Wiki*. Available at: <https://wiki.rice.edu/confluence/display/BIODESIGN/M9+Medium>. Accessed: 6/12/25
- Bhattacharyya, B., and Keck, J.L. (2014), 'Grip it and rip it: structural mechanisms of DNA helicase substrate binding and unwinding', *Protein Science: a publication of the Protein Society*, **23**(11), pp. 1498–1507. DOI: <https://doi.org/10.1002/pro.2533>
- Bhuyan, S., Yadav, M., Giri, S.J. et al. (2023), 'Microliter spotting and micro-colony observation: A rapid and simple approach for counting bacterial colony forming units', *Journal of Microbiological Methods*, **207**, 106707. DOI: <https://doi.org/10.1016/j.mimet.2023.106707>
- Bianco, P. R. (2022), 'Insight into the biochemical mechanism of DNA helicases provided by bulk-phase and single-molecule assays', *Methods (San Diego, Calif.)*, **204**, pp. 348–360. DOI: <https://doi.org/10.1016/j.ymeth.2021.12.002>

Blagosklonny, M.V., and Pardee, A.B. (2013), 'The Restriction Point of the Cell Cycle', in: Madame Curie Bioscience Database [Internet]. Austin (TX): Landes Bioscience; 2000-2013. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK6318/>

Bochman, M.L., and Schwacha, A. (2009), 'The Mcm complex: unwinding the mechanism of a replicative helicase', *Microbiology and Molecular Biology Reviews*, **73**(4), pp. 652–683. DOI: <https://doi.org/10.1128/MMBR.00019-09>

Bochman, M. L., Sabouri, N., and Zakian, V.A. (2010), 'Unwinding the functions of the Pif1 family helicases', *DNA Repair*, **9**(3), pp. 237–249. <https://doi.org/10.1016/j.dnarep.2010.01.008>

Boyd, J.B., Golino, M D., Nguyen, T.D. et al. (1976), 'Isolation and characterization of X-linked mutants of *Drosophila melanogaster* which are sensitive to mutagens', *Genetics*, **84**(3), pp. 485–506. DOI: <https://doi.org/10.1093/genetics/84.3.485>

Boyd, J.B., Golino, M.D., Shaw, K.E. et al. (1981), 'Third-chromosome mutagen-sensitive mutants of *Drosophila melanogaster*', *Genetics*, **97**(3-4), pp. 607–623. DOI: <https://doi.org/10.1093/genetics/97.3-4.607>

Branzei, D., and Foiani, M. (2008), 'Regulation of DNA repair throughout the cell cycle', *Nature Reviews. Molecular Cell Biology*, **9**(4), pp. 297–308. DOI: <https://doi.org/10.1038/nrm2351>

Brinda, K.V., Surolia, A., and Vishveshwara, S. (2005), 'Insights into the quaternary association of proteins through structure graphs: a case study of lectins', *The Biochemical Journal*, **391**(Pt 1), pp. 1–15. DOI: <https://doi.org/10.1042/BJ20050434>

Brosh, R.M., Jr (2013), 'DNA helicases involved in DNA repair and their roles in cancer', *Nature Reviews. Cancer*, **13**(8), pp. 542–558. DOI: <https://doi.org/10.1038/nrc3560>

Brenner, S., Jacob, F., and Meselson, M. (1961), 'An unstable intermediate carrying information from genes to ribosomes for protein synthesis', *Nature*, **190**, pp. 576–581. DOI: <https://doi.org/10.1038/190576a0>

Brill, S.J., and Bastin-Shanower, S. (1998), 'Identification and characterization of the fourth single-stranded-DNA binding domain of replication protein A', *Molecular and Cellular Biology*, **18**(12), pp. 7225–7234. DOI: <https://doi.org/10.1128/MCB.18.12.7225>

- Brüning, J.G., Howard, J.L., and McGlynn, P. (2014), 'Accessory replicative helicases and the replication of protein-bound DNA', *Journal of Molecular Biology*, **426**(24), pp. 3917–3928. DOI: <https://doi.org/10.1016/j.jmb.2014.10.001>
- Burgers, P.M. (2019), 'Solution to the 50-year-old Okazaki-fragment problem', *Proceedings of the National Academy of Sciences of the United States of America*, **116**(9), pp. 3358–3360. DOI: <https://doi.org/10.1073/pnas.1900372116>
- Burhans, W.C., and Weinberger, M. (2007), 'DNA replication stress, genome instability and aging', *Nucleic Acids Research*, **35**(22), pp. 7545–7556. DOI: <https://doi.org/10.1093/nar/gkm1059>
- Burke, C.R., and Lupták, A. (2018), 'DNA synthesis from diphosphate substrates by DNA polymerases', *Proceedings of the National Academy of Sciences of the United States of America*, **115**(5), pp. 980–985. DOI: <https://doi.org/10.1073/pnas.1712193115>
- Bush, N.G., Evans-Roberts, K., and Maxwell, A. (2015), 'DNA Topoisomerases', *EcoSal Plus*, **6**(2), 10.1128/ecosalplus.ESP-0010-2014. DOI: <https://doi.org/10.1128/ecosalplus.ESP-0010-2014>
- Büttner, K., Nehring, S., and Hopfner, K.P. (2007), 'Structural basis for DNA duplex separation by a superfamily-2 helicase', *Nature Structural & Molecular Biology*, **14**(7), pp. 647–652. DOI: <https://doi.org/10.1038/nsmb1246>
- Byrd, A.K., and Raney, K.D. (2012), 'Superfamily 2 helicases', *Frontiers in Bioscience (Landmark edition)*, **17**(6), pp. 2070–2088. DOI: <https://doi.org/10.2741/4038>
- Cai, L., Roginskaya, M., Qu, Y. et al. (2007), 'Structural characterization of human RPA sequential binding to single-stranded DNA using ssDNA as a molecular ruler', *Biochemistry*, **46**(28), pp. 8226–8233. DOI: <https://doi.org/10.1021/bi7004976>
- Cannone, G., Xu, Y., Beattie, T.R. et al. (2015), 'The architecture of an Okazaki fragment-processing holoenzyme from the archaeon *Sulfolobus solfataricus*', *The Biochemical Journal*, **465**(2), pp. 239–245. DOI: <https://doi.org/10.1042/BJ20141120>
- Carugo, O. (2014), 'Protomers of protein hetero-oligomers tend to resemble each other more than expected', *SpringerPlus*, **3**, pp. 680. DOI: <https://doi.org/10.1186/2193-1801-3-680>

- Cebrián, J., Castán, A., Martínez, V. et al. (2015), 'Direct Evidence for the Formation of Precatenanes during DNA Replication', *The Journal of Biological Chemistry*, **290**(22), pp. 13725–13735. DOI: <https://doi.org/10.1074/jbc.M115.642272>
- Cell Editorial Team (2024), 'Molecular biology: The fundamental science fueling innovation', *Cell*, **187**(23), pp. 6415–6416. DOI: <https://doi.org/10.1016/j.cell.2024.10.043>
- Cencic, R., and Pelletier, J. (2013), 'Throwing a monkey wrench in the motor: targeting DExH/D box proteins with small molecule inhibitors', *Biochimica et Biophysica Acta*, **1829**(8), pp. 894–903. DOI: <https://doi.org/10.1016/j.bbagr.2013.01.008>
- Cezanne, A., Foo, S., Kuo, Y.W., and Baum, B. (2024), 'The Archaeal Cell Cycle', *Annual Review of Cell and Developmental Biology*, **40**(1), pp. 1–23. DOI: <https://doi.org/10.1146/annurev-cellbio-111822-120242>
- Chang, H.H.Y., Pannunzio, N.R., Adachi, N. et al. (2017), 'Non-homologous DNA end joining and alternative pathways to double-strand break repair', *Nature Reviews. Molecular Cell Biology*, **18**(8), pp. 495–506. DOI: <https://doi.org/10.1038/nrm.2017.48>
- Chang, P., and Mariani, K.J. (2000), 'Identification of a region of *Escherichia coli* DnaB required for functional interaction with DnaG at the replication fork', *The Journal of Biological Chemistry*, **275**(34), pp. 26187–26195. DOI: <https://doi.org/10.1074/jbc.M001800200>
- Charbin, A., Bouchoux, C., and Uhlmann, F. (2014), 'Condensin aids sister chromatid decatenation by topoisomerase II', *Nucleic Acids Research*, **42**(1), pp. 340–348. DOI: <https://doi.org/10.1093/nar/gkt882>
- Chatterjee, N., and Walker, G.C. (2017), 'Mechanisms of DNA damage, repair, and mutagenesis', *Environmental and Molecular Mutagenesis*, **58**(5), pp. 235–263. DOI: <https://doi.org/10.1002/em.22087>
- Chaudhry, R., and Khaddour, K. (2023), 'Biochemistry, DNA Replication', in: *StatPearls [Internet]*. Updated 1st May 2023. Treasure Island (FL): StatPearls Publishing. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK482125/>

Chayot, R., Montagne, B., Mazel, D. et al. (2010), 'An end-joining repair mechanism in *Escherichia coli*', *Proceedings of the National Academy of Sciences of the United States of America*, **107**(5), pp. 2141–2146. DOI: <https://doi.org/10.1073/pnas.0906355107>

Chen, P.J., and Liu, D.R. (2023), 'Prime editing for precise and highly versatile genome manipulation', *Nature Reviews. Genetics*, **24**(3), pp. 161–177. DOI: <https://doi.org/10.1038/s41576-022-00541-1>

Chen, X.J. (2013), 'Mechanism of homologous recombination and implications for aging-related deletions in mitochondrial DNA', *Microbiology and Molecular Biology Reviews: MMBR*, **77**(3), pp. 476–496. DOI: <https://doi.org/10.1128/MMBR.00007-13>

Chib, S., Ali, F., and Seshasayee, A.S.N. (2017), 'Genomewide Mutational Diversity in *Escherichia coli* Population Evolving in Prolonged Stationary Phase', *mSphere*, **2**(3), e00059-17. DOI: <https://doi.org/10.1128/mSphere.00059-17>

Chilkova, O., Stenlund, P., Isoz, I. et al. (2007). The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA', *Nucleic Acids Research*, **35**(19), pp. 6588–6597. DOI: <https://doi.org/10.1093/nar/gkm741>

Chow, R.D., Chen, J.S., Shen, J. et al. (2021), 'A web tool for the design of prime-editing guide RNAs', *Nature Biomedical Engineering*, **5**(2), pp. 190–194. DOI: <https://doi.org/10.1038/s41551-020-00622-8>

Chong, S. (2014), 'Overview of cell-free protein synthesis: historic landmarks, commercial systems, and expanding applications', *Current Protocols in Molecular Biology*, **108**, pp. 16.30.1–16.30.11. DOI: <https://doi.org/10.1002/0471142727.mb1630s108>

Chu, J., Haynes, R.D., Corbel, S.Y. et al. (2014), 'Non-invasive intravital imaging of cellular differentiation with a bright red-excitable fluorescent protein', *Nature Methods*, **11**(5), pp. 572–578. DOI: <https://doi.org/10.1038/nmeth.2888>

Chung, D.W., Zhang, J.A., Tan, C.K. et al. (1991), 'Primary structure of the catalytic subunit of human DNA polymerase delta and chromosomal location of the gene', *Proceedings of the National Academy of Sciences of the United States of America*, **88**(24), pp. 11197–11201. DOI: <https://doi.org/10.1073/pnas.88.24.11197>

Clay, D.E., and Fox, D.T. (2021), 'DNA Damage Responses during the Cell Cycle: Insights from Model Organisms and Beyond', *Genes*, **12**(12), pp. 1882. DOI: <https://doi.org/10.3390/genes12121882>

Clement, K., Rees, H., Canver, M.C. et al. (2019), 'CRISPResso2 provides accurate and rapid genome editing sequence analysis', *Nature Biotechnology*, **37**(3), pp. 224–226. DOI: <https://doi.org/10.1038/s41587-019-0032-3>

Cong, F., Long, J., Liu, J. et al. (2022), 'An integrative analysis revealing POLD2 as a tumor suppressive immune protein and prognostic biomarker in pan-cancer', *Frontiers in Genetics*, **13**, pp. 877468. DOI: <https://doi.org/10.3389/fgene.2022.877468>

Cooper, G.M. (2000), 'DNA Replication', in *The Cell: A Molecular Approach. 2nd edition*. Sunderland, Massachusetts: Sinauer Associates. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK9940/>

Coskun-Ari, F.F., Skokotas, A., Moe, G.R. et al. (1994), 'Biophysical characteristics of Tus, the replication arrest protein of *Escherichia coli*', *The Journal of Biological Chemistry*, **269**(6), pp. 4027–4034. Available at: [https://www.jbc.org/article/S0021-9258\(17\)41737-6/pdf](https://www.jbc.org/article/S0021-9258(17)41737-6/pdf)

Costa, P., Gomes, A.T.P.C., Braz, M. et al. (2021), 'Application of the Resazurin Cell Viability Assay to Monitor *Escherichia coli* and *Salmonella* Typhimurium Inactivation Mediated by Phages', *Antibiotics (Basel, Switzerland)*, **10**(8), pp. 974. DOI: <https://doi.org/10.3390/antibiotics10080974>

Costes, A., and Lambert, S.A. (2012), 'Homologous recombination as a replication fork escort: fork-protection and recovery', *Biomolecules*, **3**(1), pp. 39–71. DOI: <https://doi.org/10.3390/biom3010039>

Crick, F.H. (1958), 'On protein synthesis', *Symposia of the Society for Experimental Biology*, **12**, pp. 138–163.

Crick, F.H. (1970), 'Central dogma of molecular biology', *Nature*, **227**(5258), pp. 561–563. DOI: <https://doi.org/10.1038/227561a0>

- Csermely, P., Palotai, R., and Nussinov, R. (2010), 'Induced fit, conformational selection and independent dynamic segments: an extended view of binding events', *Trends in Biochemical Sciences*, **35**(10), pp. 539–546. DOI: <https://doi.org/10.1016/j.tibs.2010.04.009>
- Cubbon, A. (2022), 'Investigating a human DNA repair helicase that supports CRISPR editing', PhD Thesis, University of Nottingham. Available at: <https://eprints.nottingham.ac.uk/68641/>. Accessed: 31/1/26
- Cui, H.L., and Dyll-Smith, M.L. (2021), 'Cultivation of halophilic archaea (class *Halobacteria*) from thalassohaline and athalassohaline environments', *Marine Life Science & Technology*, **3**(2), pp. 243–251. DOI: <https://doi.org/10.1007/s42995-020-00087-3>
- Cytiva (2020), *Amersham Typhoon Biomolecular Imager*. Available at: <https://info.cytivalifesciences.com/rs/614-VFU-334/images/Amersham-Typhoon-Biomolecular-Imager-data-file-CY12743-10Jul20-DF.pdf>. Accessed: 16/12/25
- Dağ, C., Gocenler, O., Cakir N. et al. (2025), 'M9 Medium Composition Alters *E. coli* Metabolism During Recombinant Expression', *bioRxiv*, 2025.07.28.667263. DOI: doi: <https://doi.org/10.1101/2025.07.28.667263>
- Dalgaard, J.Z., Eydmann, T., Koulintchenko, M. et al., (2009), 'Random and site-specific replication termination', *Methods in Molecular Biology (Clifton, N.J.)*, **521**, pp. 35–53. DOI: https://doi.org/10.1007/978-1-60327-815-7_3
- Daus, M.L. (2016), 'Disease Transmission by Misfolded Prion-Protein Isoforms, Prion-Like Amyloids, Functional Amyloids and the Central Dogma', *Biology*, **5**(1), pp. 2. DOI: <https://doi.org/10.3390/biology5010002>
- delCardayré, S.B., and Raines, R.T. (1994), 'Structural determinants of enzymatic processivity. Biochemistry', **33**(20), pp. 6031–6037. DOI: <https://doi.org/10.1021/bi00186a001>
- Deng, L., Chen, C.L., Zhai, Y. et al. (2021), 'Editorial: DNA Replication Stress and Cell Fate', *Frontiers in Cell and Developmental Biology*, **9**, pp. 778486. DOI: <https://doi.org/10.3389/fcell.2021.778486>

Deng, Y., Efremov, A.K., and Yan, J. (2022), 'Modulating binding affinity, specificity, and configurations by multivalent interaction', *Biophysical Journal*, **121**(10), pp. 1868–1880. DOI: <https://doi.org/10.1016/j.bpj.2022.04.017>

Dewar, J.M., and Walter, J.C. (2017), 'Mechanisms of DNA replication termination', *Nature Reviews. Molecular Cell Biology*, **18**(8), pp. 507–516. DOI: <https://doi.org/10.1038/nrm.2017.42>

Dewar, J.M. (2024), 'DNA replication recruits a friend to overcome a challenging break-up', *The EMBO Journal*, **43**(18), pp. 3815–3817. DOI: <https://doi.org/10.1038/s44318-024-00204-3>

Diffley, J.F., Cocker, J.H., Dowell, S.J. et al. (1994), 'Two steps in the assembly of complexes at yeast replication origins *in vivo*', *Cell*, **78**(2), pp. 303–316. DOI: [https://doi.org/10.1016/0092-8674\(94\)90299-2](https://doi.org/10.1016/0092-8674(94)90299-2)

Dillard, K.E., Brown, M.W., Johnson, N.V. et al. (2018), 'Assembly and Translocation of a CRISPR-Cas Primed Acquisition Complex', *Cell*, **175**(4), pp. 934–946.e15. DOI: <https://doi.org/10.1016/j.cell.2018.09.039>

Donnianni, R.A., Zhou, Z.X., Lujan, S.A. et al. (2019), 'DNA Polymerase Delta Synthesizes Both Strands during Break-Induced Replication', *Molecular Cell*, **76**(3), pp. 371–381.e4. DOI: <https://doi.org/10.1016/j.molcel.2019.07.033>

Dohrmann, P.R., Manhart, C.M., Downey, C.D. et al. (2011), 'The rate of polymerase release upon filling the gap between Okazaki fragments is inadequate to support cycling during lagging strand synthesis', *Journal of Molecular Biology*, **414**(1), pp. 15–27. DOI: <https://doi.org/10.1016/j.jmb.2011.09.039>

Doman, J.L., Sousa, A.A., Randolph, P.B. et al. (2022), 'Designing and executing prime editing experiments in mammalian cells', *Nature Protocols*, **17**(11), pp. 2431–2468. DOI: <https://doi.org/10.1038/s41596-022-00724-4>

Dos Santos Rodrigues, F.H., Delgado, G.G., Santana da Costa, T. et al. (2023), 'Applications of fluorescence spectroscopy in protein conformational changes and intermolecular contacts', *BBA Advances*, **3**, 100091. DOI: <https://doi.org/10.1016/j.bbadv.2023.100091>

- Drosopoulos, W.C., Vierra, D.A., Kenworthy, C.A. et al. (2020), 'Dynamic Assembly and Disassembly of the Human DNA Polymerase δ Holoenzyme on the Genome *In Vivo*', *Cell Reports*, **30**(5), pp. 1329–1341.e5. DOI: <https://doi.org/10.1016/j.celrep.2019.12.101>
- Ducoux, M., Urbach, S., Baldacci, G. (2001), 'Mediation of proliferating cell nuclear antigen (PCNA)-dependent DNA replication through a conserved p21(Cip1)-like PCNA-binding motif present in the third subunit of human DNA polymerase delta', *The Journal of Biological Chemistry*, **276**(52), pp. 49258–49266. DOI: <https://doi.org/10.1074/jbc.M106990200>
- Duderstadt, K.E., and Berger, J.M. (2008), 'AAA+ ATPases in the initiation of DNA replication', *Critical Reviews in Biochemistry and Molecular Biology*, **43**(3), pp. 163–187. DOI: <https://doi.org/10.1080/10409230802058296>
- Dueva, R., and Iliakis, G. (2020), 'Replication protein A: a multifunctional protein with roles in DNA replication, repair and beyond', *NAR Cancer*, **2**(3), zcaa022. DOI: <https://doi.org/10.1093/narcan/zcaa022>
- Duggin, I.G., McCallum, S.A., and Bell, S.D. (2008), 'Chromosome replication dynamics in the archaeon *Sulfolobus acidocaldarius*', *Proceedings of the National Academy of Sciences of the United States of America*, **105**(43), pp. 16737–16742. DOI: <https://doi.org/10.1073/pnas.0806414105>
- Duncker, B.P., Chesnokov, I.N., and McConkey, B.J. (2009), 'The origin recognition complex protein family', *Genome Biology*, **10**(3), pp. 214. DOI: <https://doi.org/10.1186/gb-2009-10-3-214>
- Dunn, J., and Grider, M.H. (2023), 'Physiology, Adenosine Triphosphate', in *StatPearls [Internet]*. Updated 13th February. Treasure Island (FL): StatPearls Publishing. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK553175/>
- Dyer, S.C., Austine-Orimoloye, O., Azov, A.G. et al. (2025), 'Ensembl 2025', *Nucleic Acids Research*, **53**(D1), pp. D948–D957. DOI: <https://doi.org/10.1093/nar/gkae1071>
- Ebeling, M., and Nadler, W. (1995), 'On constructing folding heteropolymers', *Proceedings of the National Academy of Sciences of the United States of America*, **92**(19), pp. 8798–8802. DOI: <https://doi.org/10.1073/pnas.92.19.8798>

- Ebeling, W., Hennrich, N., Klockow, M. et al. (1974), 'Proteinase K from *Tritirachium album* Limber', *European Journal of Biochemistry*, **47**(1), pp. 91–97. DOI: <https://doi.org/10.1111/j.1432-1033.1974.tb03671.x>
- Ekundayo, B., and Bleichert, F. (2019), 'Origins of DNA replication', *PLOS Genetics*, **15**(9), e1008320. DOI: <https://doi.org/10.1371/journal.pgen.1008320>
- Elbing, K.L., and Brent, R. (2019), 'Recipes and Tools for Culture of *Escherichia coli*', *Current Protocols in Molecular Biology*, **125**(1), e83. DOI: <https://doi.org/10.1002/cpmb.83>
- Enemark, E.J., and Joshua-Tor, L. (2008), 'On helicases and other motor proteins', *Current Opinion in Structural Biology*, **18**(2), pp. 243–257. DOI: <https://doi.org/10.1016/j.sbi.2008.01.007>
- Ensembl (2025a), Transcript: ENST00000263681.7 POLD3-201. Available at: https://www.ensembl.org/Homo_sapiens/Transcript/Summary?db=core;g=ENSG000000077514;r=11:74493851-74669117;t=ENST00000263681. Accessed: 27/11/25
- Ensembl (2025b), Transcript: ENST00000263681.7 (POLD3-201) - Export Transcript Data. Available at: https://www.ensembl.org/Homo_sapiens/Export/Output/Transcript?db=core;flank3_display=0;flank5_display=0;g=ENSG000000077514;output=fasta;r=11:74493851-74669117;strand=feature;t=ENST00000263681;param=peptide;genomic=unmasked;format=HTML. Accessed: 28/11/25
- Ensembl (2025c), Transcript: ENST00000254719.10 (RPA1-201) - Export Transcript Data. Available at: https://www.ensembl.org/Homo_sapiens/Export/Output/Transcript?db=core;flank3_display=0;flank5_display=0;g=ENSG00000132383;output=fasta;r=17:1829702-1900082;strand=feature;t=ENST00000254719;param=peptide;genomic=unmasked;format=HTML. Accessed: 31/12/25
- Erdile, L. F., Wold, M.S., and Kelly, T. J. (1990), 'The primary structure of the 32-kDa subunit of human replication protein A', *The Journal of Biological Chemistry*, **265**(6), pp. 3177–3182. Available at: [https://www.jbc.org/article/S0021-9258\(19\)39750-9/pdf](https://www.jbc.org/article/S0021-9258(19)39750-9/pdf)

Erdile, L.F., Heyer, W.D., Kolodner, R. et al. (1991), 'Characterization of a cDNA encoding the 70-kDa single-stranded DNA-binding subunit of human replication protein A and the role of the protein in DNA replication', *The Journal of Biological Chemistry*, **266**(18), pp. 12090–12098. Available at: [https://www.jbc.org/article/S0021-9258\(18\)99069-1/pdf](https://www.jbc.org/article/S0021-9258(18)99069-1/pdf)

Erzberger, J.P., Mott, M.L., and Berger, J.M. (2006), 'Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling', *Nature Structural & Molecular Biology*, **13**(8), pp. 676–683. DOI: <https://doi.org/10.1038/nsmb1115>

Estep, K.N., and Brosh, R.M., Jr (2018), 'RecQ and Fe-S helicases have unique roles in DNA metabolism dictated by their unwinding directionality, substrate specificity, and protein interactions', *Biochemical Society Transactions*, **46**(1), pp. 77–95. DOI: <https://doi.org/10.1042/BST20170044>

Fairman-Williams, M.E., Guenther, U.P., and Jankowsky, E. (2010), 'SF1 and SF2 helicases: family matters', *Current Opinion in Structural Biology*, **20**(3), pp. 313–324. DOI: <https://doi.org/10.1016/j.sbi.2010.03.011>

Fang, J., Engen, J.R., and Beuning, P. J. (2011), '*Escherichia coli* processivity clamp β from DNA polymerase III is dynamic in solution', *Biochemistry*, **50**(26), pp. 5958–5968. DOI: <https://doi.org/10.1021/bi200580b>

Fang, T., Wang, X., and Huangfu, N. (2023), 'Superfamily II helicases: the potential therapeutic target for cardiovascular diseases', *Frontiers in Cardiovascular Medicine*, **10**, 1309491. DOI: <https://doi.org/10.3389/fcvm.2023.1309491>

Feng, X., Zhang, B., Gao, Z. et al. (2022), 'A Well-Conserved Archaeal B-Family Polymerase Functions as an Extender in Translesion Synthesis', *mBio*, **13**(1), e0265921. DOI: <https://doi.org/10.1128/mbio.02659-21>

Ferenczy, G.G., and Kellermayer, M. (2022), 'Contribution of hydrophobic interactions to protein mechanical stability', *Computational and Structural Biotechnology Journal*, **20**, pp. 1946–1956. DOI: <https://doi.org/10.1016/j.csbj.2022.04.025>

Ferri, G.L., Isola, J., Berger, P. et al. (2000), 'Direct eye visualization of Cy5 fluorescence for immunocytochemistry and in situ hybridization', *The Journal of Histochemistry and*

Cytochemistry: official journal of the Histochemistry Society, **48**(3), pp. 437–444. DOI: <https://doi.org/10.1177/002215540004800314>

Fichtner, M., Voigt, K., and Schuster, S. (2017), 'The tip and hidden part of the iceberg: Proteinogenic and non-proteinogenic aliphatic amino acids', *Biochimica et Biophysica Acta. General Subjects*, **1861**(1 Pt A), pp. 3258–3269. DOI: <https://doi.org/10.1016/j.bbagen.2016.08.008>

Fousek-Schuller, V.J., and Borgstahl, G.E.O. (2024), 'The Intriguing Mystery of RPA Phosphorylation in DNA Double-Strand Break Repair', *Genes*, **15**(2), pp. 167. DOI: <https://doi.org/10.3390/genes15020167>

FPbase (2025), *mVenus*. Available at: <https://www.fpbases.org/protein/mvenus/>. Accessed: 30/11/25

FPbase (2026), *mCardinal*. Available at: <https://www.fpbases.org/protein/mcardinal/>. Accessed: 31/1/26

Fukui, K. (2010), 'DNA mismatch repair in eukaryotes and bacteria', *Journal of Nucleic Acids*, **2010**, pp. 260512. DOI: <https://doi.org/10.4061/2010/26051>

Fujikane, R., Komori, K., Shinagawa, H. et al. (2005), 'Identification of a novel helicase activity unwinding branched DNAs from the hyperthermophilic archaeon, *Pyrococcus furiosus*', *The Journal of Biological Chemistry*, **280**(13), pp. 12351–12358. DOI: <https://doi.org/10.1074/jbc.M413417200>

Fujikane, R., Shinagawa, H., and Ishino, Y. (2006), 'The archaeal Hjm helicase has recQ-like functions, and may be involved in repair of stalled replication fork', *Genes to Cells: devoted to molecular & cellular mechanisms*, **11**(2), pp. 99–110. DOI: <https://doi.org/10.1111/j.1365-2443.2006.00925.x>

Fujisawa, R., Ohashi, E., Hirota, K. et al. (2017), 'Human CTF18-RFC clamp-loader complexed with non-synthesising DNA polymerase ϵ efficiently loads the PCNA sliding clamp', *Nucleic Acids Research*, **45**(8), pp. 4550–4563. DOI: <https://doi.org/10.1093/nar/gkx096>

- Fung, H.Y.J., Birol, M., and Rhoades, E. (2018), 'IDPs in macromolecular complexes: the roles of multivalent interactions in diverse assemblies', *Current Opinion in Structural Biology*, **49**, 36–43. DOI: <https://doi.org/10.1016/j.sbi.2017.12.007>
- Gao, Y., Mutter-Rottmayer, E., Zlatanou, A. et al. (2017), 'Mechanisms of Post-Replication DNA Repair', *Genes*, **8**(2), pp. 64. DOI: <https://doi.org/10.3390/genes8020064>
- Gao, Y., and Yang, W. (2020), 'Different mechanisms for translocation by monomeric and hexameric helicases', *Current Opinion in Structural Biology*, **61**, pp. 25–32. DOI: <https://doi.org/10.1016/j.sbi.2019.10.003>
- Garnier, F., Couturier, M., Débat, H. et al. (2021), 'Archaea: A Gold Mine for Topoisomerase Diversit', *Frontiers in Microbiology*, **12**, pp. 661411. DOI: <https://doi.org/10.3389/fmicb.2021.661411>
- Ghosh, I., Hamilton, A.D. and Regan, L. (2000), 'Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein', *Journal of the American Chemical Society*, **122**(23), 5658-5659. DOI: <https://doi.org/10.1021/ja994421w>
- Gonzalez, J.M., and Aranda, B. (2023), 'Microbial Growth under Limiting Conditions-Future Perspectives', *Microorganisms*, **11**(7), pp. 1641. DOI: <https://doi.org/10.3390/microorganisms11071641>
- Gonzalez-Magaña, A., Ibáñez de Opakua, A., Romano-Moreno, M. et al. (2019), 'The p12 subunit of human polymerase δ uses an atypical PIP box for molecular recognition of proliferating cell nuclear antigen (PCNA)', *The Journal of Biological Chemistry*, **294**(11), pp. 3947–3956. DOI: <https://doi.org/10.1074/jbc.RA118.006391>
- Goodall, D.J., Warecka, D., Hawkins, M. et al. (2023), 'Interplay between chromosomal architecture and termination of DNA replication in bacteria', *Frontiers in Microbiology*, **14**, 1180848. DOI: <https://doi.org/10.3389/fmicb.2023.1180848>
- Gostimskaya, I. (2022), 'CRISPR-Cas9: A History of Its Discovery and Ethical Considerations of Its Use in Genome Editing', *Biochemistry. Biokhimiia*, **87**(8), pp. 777–788. DOI: <https://doi.org/10.1134/S0006297922080090>

- Grabarz, A., Barascu, A., Guirouilh-Barbat, J. et al. (2012), 'Initiation of DNA double strand break repair: signaling and single-stranded resection dictate the choice between homologous recombination, non-homologous end-joining and alternative end-joining', *American Journal of Cancer Research*, **2**(3), pp. 249–268. Available at: <https://pmc.ncbi.nlm.nih.gov/articles/PMC3365807/>
- Greci, M.D., and Bell, S.D. (2020), 'Archaeal DNA Replication', *Annual Review of Microbiology*, **74**, pp. 65–80. DOI: <https://doi.org/10.1146/annurev-micro-020518-115443>
- Gros, F., Hiarr, H., Gilbert, W. et al. (1961), 'Unstable ribonucleic acid revealed by pulse labelling of *Escherichia coli*', *Nature*, **190**, pp. 581–585. DOI: <https://doi.org/10.1038/190581a0>
- Gujral, H., Kushwaha, A.K., and Khurana, S. (2020), 'Utilization of Time Series Tools in Life-sciences and Neuroscience', *Neuroscience Insights*, **15**, 2633105520963045. DOI: <https://doi.org/10.1177/2633105520963045>
- Guy, C.P., and Bolt, E. L. (2005), 'Archaeal Hel308 helicase targets replication forks *in vivo* and *in vitro* and unwinds lagging strands', *Nucleic Acids Research*, **33**(11), pp. 3678–3690. DOI: <https://doi.org/10.1093/nar/gki685>
- Hakem, R. (2008), 'DNA-damage repair; the good, the bad, and the ugly', *The EMBO Journal*, **27**(4), pp. 589–605. DOI: <https://doi.org/10.1038/emboj.2008.15>
- Han, Y., Wang, S., Zhang, Z. et al. (2014), '*In vivo* imaging of protein-protein and RNA-protein interactions using novel far-red fluorescence complementation systems', *Nucleic Acids Research*, **42**(13), e103. DOI: <https://doi.org/10.1093/nar/gku408>
- Haring, S.J., Mason, A.C., Binz, S.K. et al. (2008), 'Cellular functions of human RPA1', *The Journal of Biological Chemistry*, **283**(27), pp. 19095–19111. DOI: <https://doi.org/10.1074/jbc.M800881200>
- Harmon, F.G., and Kowalczykowski, S.C. (1998), 'RecQ helicase, in concert with RecA and SSB proteins, initiates and disrupts DNA recombination', *Genes & Development*, **12**(8), pp. 1134–1144. DOI: <https://doi.org/10.1101/gad.12.8.1134>

- Harris, P.V., Mazina, O.M., Leonhardt, E.A. et al. (1996), 'Molecular cloning of *Drosophila* *mus308*, a gene involved in DNA cross-link repair with homology to prokaryotic DNA polymerase I genes', *Molecular and Cellular Biology*, **16**(10), pp. 5764–5771. DOI: <https://doi.org/10.1128/MCB.16.10.5764>
- Hass, C.S., Lam, K., and Wold, M.S. (2012), 'Repair-specific functions of replication protein A', *The Journal of Biological Chemistry*, **287**(6), pp. 3908–3918. DOI: <https://doi.org/10.1074/jbc.M111.287441>
- Hayashi, C., Miyazaki, E., Ozaki, S. et al. (2020), 'DnaB helicase is recruited to the replication initiation complex via binding of DnaA domain I to the lateral surface of the DnaB N-terminal domain', *The Journal of Biological Chemistry*, **295**(32), pp. 11131–11143. DOI: <https://doi.org/10.1074/jbc.RA120.014235>
- Haye, J.E., and Gammie, A.E. (2015), 'The Eukaryotic Mismatch Recognition Complexes Track with the Replisome during DNA Synthesis', *PLOS Genetics*, **11**(12), e1005719. DOI: <https://doi.org/10.1371/journal.pgen.1005719>
- He, L., Lever, R., Cubbon, A. et al. (2023), 'Interaction of human HelQ with DNA polymerase delta halts DNA synthesis and stimulates DNA single-strand annealing', *Nucleic Acids Research*, **51**(4), pp. 1740–1749. DOI: <https://doi.org/10.1093/nar/gkad032>
- He, Q., Wang, F., Yao, N.Y. et al. (2024), 'Structures of the human leading strand Polε-PCNA holoenzyme', *Nature Communications*, **15**(1), pp. 7847. DOI: <https://doi.org/10.1038/s41467-024-52257-x>
- Hecht, A., Endy, D., Salit, M. et al. (2016), 'When Wavelengths Collide: Bias in Cell Abundance Measurements Due to Expressed Fluorescent Proteins', *ACS Synthetic Biology*, **5**(9), pp. 1024–1027. DOI: <https://doi.org/10.1021/acssynbio.6b00072>
- Henricksen, L.A., Umbricht, C.B., and Wold, M.S. (1994), 'Recombinant replication protein A: expression, complex formation, and functional characterization', *The Journal of Biological Chemistry*, **269**(15), pp. 11121–11132. Available at: [https://www.jbc.org/article/S0021-9258\(19\)78100-9/pdf](https://www.jbc.org/article/S0021-9258(19)78100-9/pdf)
- Herlihy, A.E., and de Bruin, R.A. (2017), 'The Role of the Transcriptional Response to DNA Replication Stress', *Genes*, **8**(3), pp. 92. DOI: <https://doi.org/10.3390/genes8030092>

Hidaka, M., Kobayashi, T., Takenaka, S. et al. (1989), 'Purification of a DNA replication terminus (ter) site-binding protein in *Escherichia coli* and identification of the structural gene', *The Journal of Biological Chemistry*, **264**(35), pp. 21031–21037.

Hildenbrand, C., Stock, T., Lange, C. et al. (2011), 'Genome copy numbers and gene conversion in methanogenic archaea', *Journal of Bacteriology*, **193**(3), pp. 734–743. DOI: <https://doi.org/10.1128/JB.01016-10>

Hindi, N.N., Elsakrmy, N., and Ramotar, D. (2021), 'The base excision repair process: comparison between higher and lower eukaryotes', *Cellular and Molecular Life Sciences: CMLS*, **78**(24), pp. 7943–7965. DOI: <https://doi.org/10.1007/s00018-021-03990-9>

Hingorani, M.M., and Coman, M.M. (2002), 'On the specificity of interaction between the *Saccharomyces cerevisiae* clamp loader replication factor C and primed DNA templates during DNA replication', *The Journal of Biological Chemistry*, **277**(49), pp. 47213–47224. DOI: <https://doi.org/10.1074/jbc.M206764200>

Hogrel, G., Lu, Y., Alexandre, N. et al. (2020), 'Role of RadA and DNA Polymerases in Recombination-Associated DNA Synthesis in Hyperthermophilic Archaea', *Biomolecules*, **10**(7), pp. 1045. DOI: <https://doi.org/10.3390/biom10071045>

Horstman, A., Tonaco, I. A., Boutilier, K. et al. (2014), 'A cautionary note on the use of split-YFP/BiFC in plant protein-protein interaction studies', *International Journal of Molecular Sciences*, **15**(6), pp. 9628–9643. DOI: <https://doi.org/10.3390/ijms15069628>

Howes, T.R., and Tomkinson, A.E. (2012), 'DNA ligase I, the replicative DNA ligase', *Sub-Cellular Biochemistry*, **62**, pp. 327–341. DOI: https://doi.org/10.1007/978-94-007-4572-8_17

Hsieh, P., and Yamane, K. (2008), 'DNA mismatch repair: molecular mechanism, cancer, and ageing', *Mechanisms of Ageing and Development*, **129**(7-8), pp. 391–407. DOI: <https://doi.org/10.1016/j.mad.2008.02.012>

Hu, C.D., Chinenov, Y., and Kerppola, T.K. (2002), 'Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation', *Molecular Cell*, **9**(4), pp. 789–798. DOI: [https://doi.org/10.1016/s1097-2765\(02\)00496-3](https://doi.org/10.1016/s1097-2765(02)00496-3)

Hu, Y., and Stillman, B. (2023), 'Origins of DNA replication in eukaryotes', *Molecular Cell*, **83**(3), pp. 352–372. DOI: <https://doi.org/10.1016/j.molcel.2022.12.024>

Hua, X., Berkowitz, N.D., Willmann, M.R. et al. (2021), 'Global Analysis of RNA-Dependent RNA Polymerase-Dependent Small RNAs Reveals New Substrates and Functions for These Proteins and SGS3 in *Arabidopsis*', *Non-Coding RNA*, **7**(2), pp. 28. DOI: <https://doi.org/10.3390/ncrna7020028>

Huang, D., Pospiech, H., Kesti, T. et al. (1999), 'Structural organization and splice variants of the POLE1 gene encoding the catalytic subunit of human DNA polymerase epsilon', *The Biochemical Journal*, **339**(Pt 3), pp. 657–665.

Hughes, P., Tratner, I., Ducoux, M. et al. (1999), 'Isolation and identification of the third subunit of mammalian DNA polymerase delta by PCNA-affinity chromatography of mouse FM3A cell extracts', *Nucleic Acids Research*, **27**(10), pp. 2108–2114. DOI: <https://doi.org/10.1093/nar/27.10.2108>

Huselid, E., and Bunting, S.F. (2020), 'The Regulation of Homologous Recombination by Helicases', *Genes*, **11**(5), pp. 498. DOI: <https://doi.org/10.3390/genes11050498>

Ille, A.M., Lamont, H., Mathews, M.B. (2022), 'The Central Dogma revisited: Insights from protein synthesis, CRISPR, and beyond', *Wiley Interdisciplinary Reviews, RNA*, **13**(5), e1718. <https://doi.org/10.1002/wrna.1718>

Ilves, I., Petojevic, T., Pesavento, J.J. et al. (2010), 'Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins', *Molecular Cell*, **37**(2), pp. 247–258. DOI: <https://doi.org/10.1016/j.molcel.2009.12.030>

Inoue, J., Honda, M., Ikawa, S. et al. (2008), 'The process of displacing the single-stranded DNA-binding protein from single-stranded DNA by RecO and RecR proteins', *Nucleic Acids Research*, **36**(1), pp. 94–109. DOI: <https://doi.org/10.1093/nar/gkm1004>

Invitrogen (2007), *PureLink Genomic DNA Mini Kit*. Available at: https://documents.thermofisher.com/TFS-Assets/LSG/manuals/purelink_genomic_mini_man.pdf. Accessed: 19/9/25

Invitrogen (2016), *Lipofectamine 3000 Regent Protocol*. Available at: https://documents.thermofisher.com/TFS-Assets/LSG/manuals/lipofectamine3000_protocol.pdf. Accessed: 19/9/25

Ishino, Y., Shinagawa, H., Makino, K. et al. (1987), 'Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product', *Journal of Bacteriology*, **169**(12), pp. 5429–5433. DOI: <https://doi.org/10.1128/jb.169.12.5429-5433.1987>

Itsathitphaisarn, O., Wing, R. A., Eliason, W.K. et al. (2012), 'The hexameric helicase DnaB adopts a nonplanar conformation during translocation', *Cell*, **151**(2), pp. 267–277. DOI: <https://doi.org/10.1016/j.cell.2012.09.014>

Jacob, F., Brenner, S., and Cuzin, F. (1963), 'On the regulation of DNA replication in bacteria', *Cold Spring Harbor Symposia on Quantitative Biology*, **28**, pp. 329–348. DOI:

James, J.E., and Lascoux, M. (2025), 'Amino Acid Properties, Substitution Rates, and the Nearly Neutral Theory', *Genome Biology and Evolution*, **17**(3), evaf025. DOI: <https://doi.org/10.1093/gbe/evaf025>

Jameson, K.H., Rudolph, C J., and Hawkins, M. (2021), 'Termination of DNA replication at Tus-ter barriers results in under-replication of template DNA', *The Journal of Biological Chemistry*, **297**(6), pp. 101409. DOI: <https://doi.org/10.1016/j.jbc.2021.101409>

Janin, J., Bahadur, R.P., and Chakrabarti, P. (2008), 'Protein-protein interaction and quaternary structure', *Quarterly Reviews of Biophysics*, **41**(2), pp. 133–180. DOI: <https://doi.org/10.1017/S0033583508004708>

Jasin, M., and Rothstein, R. (2013), 'Repair of strand breaks by homologous recombination', *Cold Spring Harbor Perspectives in Biology*, **5**(11), a012740. DOI: <https://doi.org/10.1101/cshperspect.a012740>

Jenkins, T., Northall, S.J., Ptchelkine, D. et al. (2021), 'The HelQ human DNA repair helicase utilizes a PWI-like domain for DNA loading through interaction with RPA, triggering DNA unwinding by the HelQ helicase core', *NAR Cancer*, **3**(1), zcaa043. DOI: <https://doi.org/10.1093/narcan/zcaa043>

- Jiménez-Nava, R.A., Zepeda-Vallejo, L.G., Santoyo-Tepole, F. et al. (2023), 'RP-HPLC Separation and ^1H NMR Identification of a Yellow Fluorescent Compound-Riboflavin (Vitamin B₂)-Produced by the Yeast *Hyphopichia wangnamkhiaoensis*', *Biomolecules*, **13**(9), pp. 1423. DOI: <https://doi.org/10.3390/biom13091423>
- Jiricny, J. (2013), 'Postreplicative mismatch repair', *Cold Spring Harbor Perspectives in Biology*, **5**(4), a012633. DOI: <https://doi.org/10.1101/cshperspect.a012633>
- Johnson, M.E., and Hummer, G. (2011), 'Nonspecific binding limits the number of proteins in a cell and shapes their interaction networks', *Proceedings of the National Academy of Sciences of the United States of America*, **108**(2), pp. 603–608. DOI: <https://doi.org/10.1073/pnas.1010954108>
- Johnson, S.J., and Jackson, R.N. (2013), 'Ski2-like RNA helicase structures: common themes and complex assemblies', *RNA Biology*, **10**(1), pp. 33–43. DOI: <https://doi.org/10.4161/rna.22101>
- Jones, M.L., Aria, V., Baris, Y. et al. (2023), 'How Pol α -primase is targeted to replisomes to prime eukaryotic DNA replication', *Molecular Cell*, **83**(16), pp. 2911–2924.e16. DOI: <https://doi.org/10.1016/j.molcel.2023.06.035>
- Jones, R.M., Reynolds-Winczura, A., and Gambus, A. (2024), 'A Decade of Discovery-Eukaryotic Replisome Disassembly at Replication Termination', *Biology*, **13**(4), pp. 233. DOI: <https://doi.org/10.3390/biology13040233>
- JPT (2025a), *All About Amino Acids*. Available at: <https://www.jpt.com/support-contact/resources/amino-acids/>. Accessed: 1/12/25
- JPT (2025b), *Polar and Nonpolar Amino Acids: Key Insights*. Available at: <https://www.jpt.com/blog/polar-nonpolar-amino-acids/>. Accessed: 4/12/25
- Kaguni, J.M. (2011), 'Replication initiation at the *Escherichia coli* chromosomal origin', *Current Opinion in Chemical Biology*, **15**(5), pp. 606–613. DOI: <https://doi.org/10.1016/j.cbpa.2011.07.016>

- Kaguni, J.M. (2018), 'The Macromolecular Machines that Duplicate the Escherichia coli Chromosome as Targets for Drug Discovery', *Antibiotics (Basel, Switzerland)*, **7**(1), pp. 23. DOI: <https://doi.org/10.3390/antibiotics7010023>
- Kamp, J.A., Lemmens, B.B.L.G., Romeijn, R J. et al. (2021), 'Helicase Q promotes homology-driven DNA double-strand break repair and prevents tandem duplications', *Nature Communications*, **12**(1), pp. 7126. DOI: <https://doi.org/10.1038/s41467-021-27408-z>
- Kang, X., Hajek, B., Wu, F. et al. (2019), 'Time series experimental design under one-shot sampling: The importance of condition diversity', *PLOS One*, **14**(10), e0224577. DOI: <https://doi.org/10.1371/journal.pone.0224577>
- Kaplan, D.L., and Bruck, I. (2010), 'Methods to study how replication fork helicases unwind DNA', *Methods in Molecular Biology (Clifton, N.J.)*, **587**, pp. 127–135. DOI: https://doi.org/10.1007/978-1-60327-355-8_9
- Karvelis, T., Gasiunas, G., Miksys, A. et al. (2013), 'crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*', *RNA Biology*, **10**(5), pp. 841–851. DOI: <https://doi.org/10.4161/rna.24203>
- Katayama, T., Kasho, K., and Kawakami, H. (2017), 'The DnaA Cycle in *Escherichia coli*: Activation, Function and Inactivation of the Initiator Protein', *Frontiers in Microbiology*, **8**, pp. 2496. DOI: <https://doi.org/10.3389/fmicb.2017.02496>
- Kaushik Tiwari, M., Adaku, N., Peart, N. et al. (2016), 'Triplex structures induce DNA double strand breaks via replication fork collapse in NER deficient cell', *Nucleic Acids Research*, **44**(16), pp. 7742–7754. DOI: <https://doi.org/10.1093/nar/gkw515>
- Kerppola, T.K. (2006a), 'Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells', *Nature Protocols*, **1**(3), pp. 1278–1286. DOI: <https://doi.org/10.1038/nprot.2006.201>
- Kerppola, T.K. (2006b), 'Visualization of molecular interactions by fluorescence complementation', *Nature Reviews. Molecular Cell Biology*, **7**(6), pp. 449–456. DOI: <https://doi.org/10.1038/nrm1929>

Kerppola, T.K. (2009), 'Visualization of molecular interactions using bimolecular fluorescence complementation analysis: characteristics of protein fragment complementation', *Chemical Society Reviews*, **38**(10), pp. 2876–2886. DOI: <https://doi.org/10.1039/b909638h>

Keyamura, K., Fujikawa, N., Ishida, T. et al. (2007), 'The interaction of DiaA and DnaA regulates the replication cycle in *E. coli* by directly promoting ATP DnaA-specific initiation complexes', *Genes & Development*, **21**(16), pp. 2083–2099. DOI: <https://doi.org/10.1101/gad.1561207>

Khandagale, P., Thakur, S., and Acharya, N. (2020), 'Identification of PCNA-interacting protein motifs in human DNA polymerase δ ', *Bioscience Reports*, **40**(4), BSR20200602. DOI: <https://doi.org/10.1042/BSR20200602>

Kim, J., Darlington, A., Salvador, M. et al. (2020), 'Trade-offs between gene expression, growth and phenotypic diversity in microbial populations', *Current Opinion in Biotechnology*, **62**, pp. 29–37. DOI: <https://doi.org/10.1016/j.copbio.2019.08.004>

Kim, J.H., and Seo, Y.S. (2009), 'In vitro assays for studying helicase activities', *Methods in Molecular Biology (Clifton, N.J.)*, **521**, pp. 361–379. DOI: https://doi.org/10.1007/978-1-60327-815-7_20

Kitani, T., Yoda, K., Ogawa, T. et al. (1985), 'Evidence that discontinuous DNA replication in *Escherichia coli* is primed by approximately 10 to 12 residues of RNA starting with a purine', *Journal of Molecular Biology*, **184**(1), pp. 45–52. DOI: [https://doi.org/10.1016/0022-2836\(85\)90042-7](https://doi.org/10.1016/0022-2836(85)90042-7)

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q. et al. (2015), 'Engineered CRISPR-Cas9 nucleases with altered PAM specificities', *Nature*, **523**(7561), pp. 481–485. DOI: <https://doi.org/10.1038/nature14592>

Kodama, Y., and Hu, C.D. (2012), 'Bimolecular fluorescence complementation (BiFC): a 5-year update and future perspectives', *BioTechniques*, **53**(5), pp. 285–298. <https://doi.org/10.2144/000113943>

Konc, J., and Janežič, D. (2022), 'Protein binding sites for drug design', *Biophysical Reviews*, **14**(6), pp. 1413–1421. DOI: <https://doi.org/10.1007/s12551-022-01028-3>

- Kong, X.P., Onrust, R., O'Donnell, M. et al. (1992), 'Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp', *Cell*, **69**(3), pp. 425–437. DOI: [https://doi.org/10.1016/0092-8674\(92\)90445-i](https://doi.org/10.1016/0092-8674(92)90445-i)
- Koonin, E.V. (2012), 'Does the central dogma still stand?', *Biology Direct*, **7**, pp. 27. DOI: <https://doi.org/10.1186/1745-6150-7-27>
- Kose, H.B., Larsen, N.B., Duxin, J.P. et al. (2019), 'Dynamics of the Eukaryotic Replicative Helicase at Lagging-Strand Protein Barriers Support the Steric Exclusion Model', *Cell Reports*, **26**(8), pp. 2113–2125.e6. DOI: <https://doi.org/10.1016/j.celrep.2019.01.086>
- Kottur, J., and Nair, D.T. (2018), 'Pyrophosphate hydrolysis is an intrinsic and critical step of the DNA synthesis reaction', *Nucleic Acids Research*, **46**(12), pp. 5875–5885. DOI: <https://doi.org/10.1093/nar/gky402>
- Krokan, H.E., and Bjørås, M. (2013), 'Base excision repair', *Cold Spring Harbor Perspectives in Biology*, **5**(4), a012583. DOI: <https://doi.org/10.1101/cshperspect.a012583>
- Krupyanskii, Y.F. (2023), 'Determination of DNA architecture of bacteria under various types of stress, methodological approaches, problems, and solutions', *Biophysical Reviews*, **15**(5), pp. 1035–1051. DOI: <https://doi.org/10.1007/s12551-023-01122-0>
- Krwawicz, J., Arczewska, K.D., Speina, E. et al. (2007), 'Bacterial DNA repair genes and their eukaryotic homologues: 1. Mutations in genes involved in base excision repair (BER) and DNA-end processors and their implication in mutagenesis and human disease', *Acta Biochimica Polonica*, **54**(3), pp. 413–434.
- Kundrotas, P.J., & Vakser, I.A. (2013), 'Protein-protein alternative binding modes do not overlap', *Protein Science: a publication of the Protein Society*, **22**(8), pp. 1141–1145. DOI: <https://doi.org/10.1002/pro.2295>
- Kunkel, T.A., and Erie, D.A. (2005), 'DNA mismatch repair', *Annual Review of Biochemistry*, **74**, pp. 681–710. DOI: <https://doi.org/10.1146/annurev.biochem.74.082803.133243>
- Kushnirov, V.V., Dergalev, A.A., and Alexandrov, A.I. (2020), 'Proteinase K resistant cores of prions and amyloids', *Prion*, **14**(1), pp. 11–19. DOI: <https://doi.org/10.1080/19336896.2019.1704612>

Kvaratskhelia, M., Wardleworth, B.N., Bond, C.S. et al. (2002), 'Holliday junction resolution is modulated by archaeal chromatin components *in vitro*', *The Journal of Biological Chemistry*, **277**(4), pp. 2992–2996. DOI: <https://doi.org/10.1074/jbc.M109496200>

La, D., Kong, M., Hoffman, W. et al. (2013), 'Predicting permanent and transient protein-protein interfaces', *Proteins*, **81**(5), pp. 805–818. DOI: <https://doi.org/10.1002/prot.24235>

Lang, N., Jagtap, P.K.A., and Hennig, J. (2024), 'Regulation and mechanisms of action of RNA helicases', *RNA Biology*, **21**(1), pp. 24–38. DOI: <https://doi.org/10.1080/15476286.2024.2415801>

Langston, L.D., Georgescu, R.E., and O'Donnell, M.E. (2023), 'Mechanism of eukaryotic origin unwinding is a dual helicase DNA shearing process', *Proceedings of the National Academy of Sciences of the United States of America*, **120**(52), e2316466120. DOI: <https://doi.org/10.1073/pnas.2316466120>

Lattmann, E., Deng, T., Walser, M. et al. (2022), 'A DNA replication-independent function of pre-replication complex genes during cell invasion in *C. elegans*', *PLOS Biology*, **20**(2), e3001317. DOI: <https://doi.org/10.1371/journal.pbio.3001317>

Łazowski, K., Woodgate, R., and Fijalkowska, I.J. (2024), '*Escherichia coli* DNA replication: the old model organism still holds many surprise', *FEMS Microbiology Reviews*, **48**(4), fuae018. DOI: <https://doi.org/10.1093/femsre/fuae018>

LeBowitz, J.H., and McMacken, R. (1986), 'The *Escherichia coli* dnaB replication protein is a DNA helicase', *The Journal of Biological Chemistry*, **261**(10), pp. 4738–4748.

Lee, C.M., Wang, G., Pertsinidis, A. et al. (2019), 'Topoisomerase III Acts at the Replication Fork To Remove Precatenanes', *Journal of Bacteriology*, **201**(7), e00563-18. DOI: <https://doi.org/10.1128/JB.00563-18>

Lee, J., Kweon, J., and Kim, Y. (2025), 'Emerging trends in prime editing for precision genome editing', *Experimental & Molecular Medicine*, **57**(7), pp. 1381–1391. DOI: <https://doi.org/10.1038/s12276-025-01463-8>

- Lee, M.Y.W.T., Wang, X., Zhang, S., et al. (2017), 'Regulation and Modulation of Human DNA Polymerase δ Activity and Function', *Genes*, **8**(7), pp. 190. DOI: <https://doi.org/10.3390/genes8070190>
- Lehman, I.R., Zimmerman, S.B., Adler, J. et al. (1958), 'ENZYMATIC SYNTHESIS OF DEOXYRIBONUCLEIC ACID. V. CHEMICAL COMPOSITION OF ENZYMATICALLY SYNTHESIZED DEOXYRIBONUCLEIC ACID', *Proceedings of the National Academy of Sciences of the United States of America*, **44**(12), pp. 1191–1196. DOI: <https://doi.org/10.1073/pnas.44.12.1191>
- Leman, A.R., and Noguchi, E. (2013), 'The replication fork: understanding the eukaryotic replication machinery and the challenges to genome duplication', *Genes*, **4**(1), pp. 1–32. DOI: <https://doi.org/10.3390/genes4010001>
- Leonard, A.C., and Grimwade, J.E. (2010), 'Regulating DnaA complex assembly: it is time to fill the gaps', *Current Opinion in Microbiology*, **13**(6), pp. 766–772. DOI: <https://doi.org/10.1016/j.mib.2010.10.001>
- Leonard, A.C., and Méchali, M. (2013), 'DNA replication origins', *Cold Spring Harbor Perspectives in Biology*, **5**(10), a010116. DOI: <https://doi.org/10.1101/cshperspect.a010116>
- Lever, R.J., Simmons, E., Gamble-Milner, R. et al. (2023), 'Archaeal Hel308 suppresses recombination through a catalytic switch that controls DNA annealing', *Nucleic Acids Research*, **51**(16), pp. 8563–8574. DOI: <https://doi.org/10.1093/nar/gkad572>
- Lewis, J.S., Jergic, S., and Dixon, N.E. (2016), 'The *E. coli* DNA Replication Fork', *The Enzymes*, **39**, pp. 31–88. DOI: <https://doi.org/10.1016/bs.enz.2016.04.001>
- Li, G.M. (2008), 'Mechanisms and functions of DNA mismatch repair', *Cell Research*, **18**(1), pp. 85–98. DOI: <https://doi.org/10.1038/cr.2007.115>
- Li, H., and Stillman, B. (2012), 'The origin recognition complex: a biochemical and structural view', *Sub-Cellular Biochemistry*, **62**, pp. 37–58. DOI: https://doi.org/10.1007/978-94-007-4572-8_3
- Li, X., and Heyer, W.D. (2008), 'Homologous recombination in DNA repair and DNA damage tolerance', *Cell Research*, **18**(1), pp. 99–113. DOI: <https://doi.org/10.1038/cr.2008.1>

- Li, X., Stith, C.M., Burgers, P.M. et al. (2009), 'PCNA is required for initiation of recombination-associated DNA synthesis by DNA polymerase delta', *Molecular Cell*, **36**(4), pp. 704–713. DOI: <https://doi.org/10.1016/j.molcel.2009.09.036>
- Li, Y., and Zhao, D. (2013), 'Basics of Molecular Biology'. *Molecular Imaging: Fundamentals and Applications*, pp. 541–601. DOI: https://doi.org/10.1007/978-3-642-34303-2_16
- Liao, C., and Beisel, C.L. (2021), 'The tracrRNA in CRISPR Biology and Technologies', *Annual Review of Genetics*, **55**, pp. 161–181. DOI: <https://doi.org/10.1146/annurev-genet-071719-022559>
- Lieber, M.R., Ma, Y., Pannicke, U. et al. (2003), 'Mechanism and regulation of human non-homologous DNA end-joining', *Nature Reviews. Molecular Cell Biology*, **4**(9), pp. 712–720. DOI: <https://doi.org/10.1038/nrm1202>
- Lin, T., Zhang, L., Jiang, D. et al. (2022), 'Biochemical characterization and mutational analysis of a novel flap endonuclease 1 from *Thermococcus barophilus* Ch5', *The International Journal of Biochemistry & Cell Biology*, **143**, pp. 106154. DOI: <https://doi.org/10.1016/j.biocel.2021.106154>
- Lindås, A.C., and Bernander, R. (2013), 'The cell cycle of archaea', *Nature Reviews. Microbiology*, **11**(9), pp. 627–638. DOI: <https://doi.org/10.1038/nrmicro3077>
- Lindqvist, R., and Barmark, G. (2014), 'Specific growth rate determines the sensitivity of *Escherichia coli* to lactic acid stress: implications for predictive microbiology', *BioMed Research International*, **2014**, pp. 471317. DOI: <https://doi.org/10.1155/2014/471317>
- Linke, S.P., Clarkin, K.C., Di Leonardo, A. et al. (1996), 'A reversible, p53-dependent G0/G1 cell cycle arrest induced by ribonucleotide depletion in the absence of detectable DNA damage', *Genes & Development*, **10**(8), pp. 934–947. DOI: <https://doi.org/10.1101/gad.10.8.934>
- Liu, B., Hu, J., Wang, J. et al. (2017), 'Direct Visualization of RNA-DNA Primer Removal from Okazaki Fragments Provides Support for Flap Cleavage and Exonucleolytic Pathways in Eukaryotic Cells', *The Journal of Biological Chemistry*, **292**(12), pp. 4777–4788. DOI: <https://doi.org/10.1074/jbc.M116.758599>

- Liu, D.N., Zhou, Y.F., Peng, A.F. et al. (2017), 'HELQ reverses the malignant phenotype of osteosarcoma cells via CHK1-RAD51 signaling pathway', *Oncology Reports*, **37**(2), pp. 1107–1113. DOI: <https://doi.org/10.3892/or.2016.5329>
- Liu, L., Mo, J., Rodriguez-Belmonte, E.M. et al. (2000), 'Identification of a fourth subunit of mammalian DNA polymerase delta', *The Journal of Biological Chemistry*, **275**(25), pp. 18739–18744. DOI: <https://doi.org/10.1074/jbc.M001217200>
- Liu, P., Liang, S.Q., Zheng, C. et al. (2021), 'Improved prime editors enable pathogenic allele correction and cancer modelling in adult mice', *Nature Communications*, **12**(1), pp. 2121. DOI: <https://doi.org/10.1038/s41467-021-22295-w>
- Liu, J., Ehmsen, K.T., Heyer, W.D. et al. (2011), 'Presynaptic filament dynamics in homologous recombination and DNA repair', *Critical Reviews in Biochemistry and Molecular Biology*, **46**(3), pp. 240–270. DOI: <https://doi.org/10.3109/10409238.2011.576007>
- Liu, T., Shen, X., Ren, Y. et al. (2024), 'Genome-wide mapping of native co-localized G4s and R-loops in living cells', *eLife*, **13**, RP99026. DOI: <https://doi.org/10.7554/eLife.99026>
- Long, J., Zhu, J.Y., Liu, Y.B. et al. (2018), 'Helicase POLQ-like (HELQ) as a novel indicator of platinum-based chemoresistance for epithelial ovarian cancer', *Gynecologic Oncology*, **149**(2), pp. 341–349. DOI: <https://doi.org/10.1016/j.ygyno.2018.03.006>
- Lopez, M.J., and Mohiuddin, S.S. (2024), 'Biochemistry, Essential Amino Acids', in: *StatPearls [Internet]*. Updated 30th April 2024. Treasure Island (FL): StatPearls Publishing. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK557845/>
- Lou-Hing, A. (2024), 'The molecular annealing mechanisms of *H. sapiens* HelQ', MRes Thesis, University of Nottingham. Available at: <https://eprints.nottingham.ac.uk/77470/>. Accessed: 18/8/25
- Lu, C., Kuang, J., Shao, T. et al. (2022), 'Prime Editing: An All-Rounder for Genome Editing', *International Journal of Molecular Sciences*, **23**(17), pp. 9862. DOI: <https://doi.org/10.3390/ijms23179862>

- Lu, J., Wu, T., Zhang, B. et al. (2021), 'Types of nuclear localization signals and mechanisms of protein import into the nucleus', *Cell Communication and Signaling: CCS*, **19**(1), pp. 60. DOI: <https://doi.org/10.1186/s12964-021-00741-y>
- Lucas, I., Germe, T., Chevrier-Miller, M. et al. (2001), 'Topoisomerase II can unlink replicating DNA by precatenane removal', *The EMBO Journal*, **20**(22), pp. 6509–6519. DOI: <https://doi.org/10.1093/emboj/20.22.6509>
- Lujan, S.A., Williams, J.S., and Kunkel, T.A. (2016), 'DNA Polymerases Divide the Labor of Genome Replication', *Trends in Cell Biology*, **26**(9), pp. 640–654. DOI: <https://doi.org/10.1016/j.tcb.2016.04.012>
- Lundgren, M., Andersson, A., Chen, L. et al. (2004), 'Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination', *Proceedings of the National Academy of Sciences of the United States of America*, **101**(18), pp. 7046–7051. DOI: <https://doi.org/10.1073/pnas.0400656101>
- Ma, C.J., Gibb, B., Kwon, Y. et al. (2017), 'Protein dynamics of human RPA and RAD51 on ssDNA during assembly and disassembly of the RAD51 filament', *Nucleic Acids Research*, **45**(2), pp. 749–761. DOI: <https://doi.org/10.1093/nar/gkw1125>
- Makarova, K.S., Koonin, E.V., and Kelman, Z. (2012), 'The CMG (CDC45/RecJ, MCM, GINS) complex is a conserved component of the DNA replication system in all archaea and eukaryotes', *Biology Direct*, **7**, pp. 7. DOI: <https://doi.org/10.1186/1745-6150-7-7>
- Makarova, K.S., and Koonin, E.V. (2013), 'Archaeology of eukaryotic DNA replication', *Cold Spring Harbor Perspectives in Biology*, **5**(11), a012963. DOI: <https://doi.org/10.1101/cshperspect.a012963>
- Makarova, K.S., Krupovic, M., and Koonin, E.V. (2014), 'Evolution of replicative DNA polymerases in archaea and their contributions to the eukaryotic replication machinery', *Frontiers in Microbiology*, **5**, pp. 354. DOI: <https://doi.org/10.3389/fmicb.2014.00354>
- Makiela-Dzibenska, K., Maslowska, K.H., Kuban, W. et al. (2019)', Replication fidelity in *E. coli*: Differential leading and lagging strand effects for dnaE antimutator alleles', *DNA Repair*, **83**, 102643. DOI: <https://doi.org/10.1016/j.dnarep.2019.102643>

- Mammen, M., Choi, S.K., and Whitesides, G.M. (1998), 'Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors', *Angewandte Chemie (International ed. in English)*, **37**(20), pp. 2754–2794. DOI: [https://doi.org/10.1002/\(SICI\)1521-3773\(19981102\)37:20<2754::AID-ANIE2754>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1521-3773(19981102)37:20<2754::AID-ANIE2754>3.0.CO;2-3)
- Mao, Z., Bozzella, M., Seluanov, A. et al. (2008), 'DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells', *Cell Cycle (Georgetown, Tex.)*, **7**(18), pp. 2902–2906. DOI: <https://doi.org/10.4161/cc.7.18.6679>
- Marescal, O., and Cheeseman, I.M. (2020), 'Cellular Mechanisms and Regulation of Quiescence', *Developmental Cell*, **55**(3), pp. 259–271. DOI: <https://doi.org/10.1016/j.devcel.2020.09.029>
- Maric, M., Maculins, T., De Piccoli, G. et al. (2014), 'Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication', *Science (New York, N.Y.)*, **346**(6208), pp. 1253596. DOI: <https://doi.org/10.1126/science.1253596>
- Marín-López, M.A., Planas-Iglesias, J., Aguirre-Plans, J. et al. (2018), 'On the mechanisms of protein interactions: predicting their affinity from unbound tertiary structures', *Bioinformatics (Oxford, England)*, **34**(4), pp. 592–598. <https://doi.org/10.1093/bioinformatics/btx616>
- Marini, F., and Wood, R.D. (2002), 'A human DNA helicase homologous to the DNA cross-link sensitivity protein Mus308', *The Journal of Biological Chemistry*, **277**(10), pp. 8716–8723. DOI: <https://doi.org/10.1074/jbc.M110271200>
- Marszalek, J., and Kaguni, J. M. (1994), 'DnaA protein directs the binding of DnaB protein in initiation of DNA replication in *Escherichia coli*', *The Journal of Biological Chemistry*, **269**(7), pp. 4883–4890.
- Martínez-Carranza, M., Vialle, L., Madru, C. et al. (2024), 'Communication between DNA polymerases and Replication Protein A within the archaeal replisome', *Nature Communications*, **15**(1), pp. 10926. DOI: <https://doi.org/10.1038/s41467-024-55365-w>
- Masuda, M., Freguia, S., Wang, Y.F et al. (2010), 'Flavins contained in yeast extract are exploited for anodic electron transfer by *Lactococcus lactis*', *Bioelectrochemistry (Amsterdam, Netherlands)*, **78**(2), pp. 173–175. DOI: <https://doi.org/10.1016/j.bioelechem.2009.08.004>

Matson, S.W. (1986), 'Escherichia coli helicase II (urvD gene product) translocates unidirectionally in a 3' to 5' direction', *The Journal of Biological Chemistry*, **261**(22), pp. 10169–10175.

McCarthy, B.J., and Holland, J.J. (1965), 'Denatured DNA as a direct template for *in vitro* protein synthesis', *Proceedings of the National Academy of Sciences of the United States of America*, **54**(3), pp. 880–886. DOI: <https://doi.org/10.1073/pnas.54.3.880>

McHenry, C.S. (2011), 'Breaking the rules: bacteria that use several DNA polymerase IIIs', *EMBO Reports*, **12**(5), pp. 408–414. DOI: <https://doi.org/10.1038/embor.2011.51>

McIntosh, J.R. (2016), 'Mitosis', *Cold Spring Harbor Perspectives in Biology*, **8**(9), a023218. DOI: <https://doi.org/10.1101/cshperspect.a023218>

McKie, S.J., Neuman, K.C., and Maxwell, A. (2021), 'DNA topoisomerases: Advances in understanding of cellular roles and multi-protein complexes via structure-function analysis', *BioEssays: news and reviews in molecular, cellular and developmental biology*, **43**(4), e2000286. DOI: <https://doi.org/10.1002/bies.202000286>

McMillan, S.D., and Keck, J.L. (2024), 'Biochemical characterization of *Escherichia coli* DnaC variants that alter DnaB helicase loading onto DNA', *The Journal of Biological Chemistry*, **300**(5), pp. 107275. DOI: <https://doi.org/10.1016/j.jbc.2024.107275>

Meir, A., and Greene, E.C. (2021), 'Srs2 and Pif1 as Model Systems for Understanding Sf1a and Sf1b Helicase Structure and Function', *Genes*, **12**(9), pp. 1319. DOI: <https://doi.org/10.3390/genes12091319>

Menon, V., and Gueble, S.E. (2025), 'DNA repair helicases: from mechanistic understanding to therapeutic implications', *NAR Cancer*, **7**(4), zcaf034. DOI: <https://doi.org/10.1093/narcan/zcaf034>

Mercadante, A.A., Dimri, M., and Mohiuddin, S.S. (2023), 'Biochemistry, Replication and Transcription', in: *StatPearls [Internet]*. Updated 14th August 2023. Treasure Island (FL): StatPearls Publishing. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK540152/>

Meselson, M., and Stahl, F.W. (1958), 'THE REPLICATION OF DNA IN *ESCHERICHIA COLI*', *Proceedings of the National Academy of Sciences of the United States of America*, **44**(7), pp. 671–682. DOI: <https://doi.org/10.1073/pnas.44.7.671>

Meyer, R.R., and Laine, P.S. (1990), 'The single-stranded DNA-binding protein of *Escherichia coli*', *Microbiological Reviews*, **54**(4), pp. 342–380. DOI: <https://doi.org/10.1128/mr.54.4.342-380.1990>

Middleton, C.L., Parker, J.L., Richard, D.J. et al., (2004), 'Substrate recognition and catalysis by the Holliday junction resolving enzyme Hje', *Nucleic Acids Research*, **32**(18), pp. 5442–5451. DOI: <https://doi.org/10.1093/nar/gkh869>

Midgley-Smith, S.L., Dimude, J.U., Taylor, T. et al. (2018), 'Chromosomal over-replication in *Escherichia coli* recG cells is triggered by replication fork fusion and amplified if replicore symmetry is disturbed', *Nucleic Acids Research*, **46**(15), pp. 7701–7715. DOI: <https://doi.org/10.1093/nar/gky566>

Miller, H.I., Riggs, A.D., and Gill, G.N. (1973), 'Ribonuclease H (hybrid) in *Escherichia coli*. Identification and characterization', *The Journal of Biological Chemistry*, **248**(7), pp. 2621–2624.

Miller, K.E., Kim, Y., Huh, W.K et al. (2015), 'Bimolecular Fluorescence Complementation (BiFC) Analysis: Advances and Recent Applications for Genome-Wide Interaction Studies', *Journal of Molecular Biology*, **427**(11), pp. 2039–2055. DOI: <https://doi.org/10.1016/j.jmb.2015.03.005>

Minchin, S., and Lodge, J. (2019), 'Understanding biochemistry: structure and function of nucleic acids', *Essays in Biochemistry*, **63**(4), pp. 433–456. DOI: <https://doi.org/10.1042/EBC20180038>

Mirica, K.A., Lockett, M.R., Snyder, P.W. et al. (2012), 'Selective precipitation and purification of monovalent proteins using oligovalent ligands and ammonium sulfate', *Bioconjugate Chemistry*, **23**(2), pp. 293–299. DOI: <https://doi.org/10.1021/bc200390q>

Mojica, F.J., Juez, G., and Rodríguez-Valera, F. (1993), 'Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified PstI sites', *Molecular Microbiology*, **9**(3), pp. 613–621. DOI: <https://doi.org/10.1111/j.1365-2958.1993.tb01721.x>

- Mojica, F.J., Díez-Villaseñor, C., García-Martínez, J. et al. (2005), 'Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements', *Journal of Molecular Evolution*, **60**(2), pp. 174–182. DOI: <https://doi.org/10.1007/s00239-004-0046-3>
- Mondol, T., Stodola, J.L., Galletto, R. et al. (2019), 'PCNA accelerates the nucleotide incorporation rate by DNA polymerase δ ', *Nucleic Acids Research*, **47**(4), pp. 1977–1986. DOI: <https://doi.org/10.1093/nar/gky1321>
- Montgomery, M.K., Xu, S., and Fire, A. (1998), 'RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*', *Proceedings of the National Academy of Sciences of the United States of America*, **95**(26), pp. 15502–15507. DOI: <https://doi.org/10.1073/pnas.95.26.1550>
- Moolenaar, G.F., Moorman, C., and Goosen, N. (2000), 'Role of the *Escherichia coli* nucleotide excision repair proteins in DNA replication', *Journal of Bacteriology*, **182**(20), pp. 5706–5714. DOI: <https://doi.org/10.1128/JB.182.20.5706-5714.2000>
- Moolman, M.C., Tiruvadi Krishnan, S., Kerssemakers, J.W. et al. (2016), 'The progression of replication forks at natural replication barriers in live bacteria', *Nucleic Acids Research*, **44**(13), pp. 6262–6273. DOI: <https://doi.org/10.1093/nar/gkw397>
- Moreno, S.P., Bailey, R., Campion, N. et al. (2014), 'Polyubiquitylation drives replisome disassembly at the termination of DNA replication', *Science (New York, N.Y.)*, **346**(6208), pp. 477–481. DOI: <https://doi.org/10.1126/science.1253585>
- Moreno, S.P., and Gambus, A. (2020), 'Mechanisms of eukaryotic replisome disassembly', *Biochemical Society Transactions*, **48**(3), pp. 823–836. DOI: <https://doi.org/10.1042/BST20190363>
- Morris, R., Black, K.A., and Stollar, E.J. (2022), 'Uncovering protein function: from classification to complexes', *Essays in Biochemistry*, **66**(3), pp. 255–285. DOI: <https://doi.org/10.1042/EBC20200108>
- Moynahan, M.E., and Jasin, M. (2010), 'Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis', *Nature Reviews. Molecular Cell Biology*, **11**(3), pp. 196–207. DOI: <https://doi.org/10.1038/nrm2851>

Mulcair, M.D., Schaeffer, P.M., Oakley, A.J. et al. (2006), 'A molecular mousetrap determines polarity of termination of DNA replication in *E. coli*', *Cell*, **125**(7), pp. 1309–1319. DOI: <https://doi.org/10.1016/j.cell.2006.04.040>

Mulugu, S., Potnis, A., Shamsuzzaman et al. (2001), 'Mechanism of termination of DNA replication of *Escherichia coli* involves helicase-contrahelicase interaction', *Proceedings of the National Academy of Sciences of the United States of America*, **98**(17), pp. 9569–9574. DOI: <https://doi.org/10.1073/pnas.171065898>

Mustafin, R.N., and Khusnutdinova, E.K. (2019), 'The Role of Reverse Transcriptase in the Origin of Life', *Biochemistry. Biokhimiia*, **84**(8), pp. 870–883. DOI: <https://doi.org/10.1134/S0006297919080030>

Nagai, T., Ibata, K., Park, E.S. et al. (2002), 'A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications', *Nature Biotechnology*, **20**(1), pp. 87–90. DOI: <https://doi.org/10.1038/nbt0102-87>

Nakamura, T., Zhao, Y., Yamagata, Y. et al. (2013), 'Mechanism of the nucleotidyl-transfer reaction in DNA polymerase revealed by time-resolved protein crystallography', *Biophysics (Nagoya-shi, Japan)*, **9**, pp. 31–36. DOI: <https://doi.org/10.2142/biophysics.9.31>

Nasheuer, H.P., Meaney, A.M., Hulshoff, T. et al. (2024), 'Replication Protein A, the Main Eukaryotic Single-Stranded DNA Binding Protein, a Focal Point in Cellular DNA Metabolism', *International Journal of Molecular Sciences*, **25**(1), pp. 588. DOI: <https://doi.org/10.3390/ijms25010588>

National Library of Medicine (2019), 'Molecular Biology', in *Collection Development Guidelines of the National Library of Medicine [Internet]*. Updated 25th June 2003. Bethesda (MD): National Library of Medicine (US). Available at: <https://www.ncbi.nlm.nih.gov/books/NBK518759/>

Naufer, M.N., Murison, D.A., Rouzina, I. et al. (2017), 'Single-molecule mechanochemical characterization of *E. coli* pol III core catalytic activity', *Protein Science: a publication of the Protein Society*, **26**(7), pp. 1413–1426. DOI: <https://doi.org/10.1002/pro.3152>

NEB (2014), *CRISPR/Cas9 & Targeted Genome Editing: New Era in Molecular Biology*. Available at: <https://www.neb.com/en/tools-and-resources/feature-articles/crispr-cas9-and-targeted-genome-editing-a-new-era-in-molecular-biology>. Accessed: 4/1/26

NCBI (2025a), *CCDS Report For Consensus CDS*. Available at: <https://www.ncbi.nlm.nih.gov/projects/CCDS/CcdsBrowse.cgi?REQUEST=ALLFIELDS&DATA=CCDS3603.1&ORGANISM=0&BUILDS=CURRENTBUILDS>. Accessed: 24/8/25

NCBI (2025b), *HELQ helicase, POLQ like [Homo sapiens (human)]*. Available at: <https://www.ncbi.nlm.nih.gov/gene/113510>. Accessed: 16/11/25

NCBI (2025c), *POLD1 DNA polymerase delta 1, catalytic subunit [Homo sapiens (human)]*. Available at: <https://www.ncbi.nlm.nih.gov/gene/5424>. Accessed: 27/11/25

NCBI (2025d), *POLD2 DNA polymerase delta 2, accessory subunit [Homo sapiens (human)]*. Available at: <https://www.ncbi.nlm.nih.gov/gene/5425>. Accessed: 27/11/25

NCBI (2025e), *POLD3 DNA polymerase delta 3, accessory subunit [Homo sapiens (human)]*. Available at: <https://www.ncbi.nlm.nih.gov/gene/10714>. Accessed: 27/11/25

NCBI (2025f), *POLD4 DNA polymerase delta 4, accessory subunit [Homo sapiens (human)]*. Available at: <https://www.ncbi.nlm.nih.gov/gene/57804>. Accessed: 27/11/25

NCBI (2025g), *RPA2 replication protein A2 [Homo sapiens (human)]*. Available at: <https://www.ncbi.nlm.nih.gov/gene/6118>. Accessed: 30/12/25

NCBI (2025h), *RPA3 replication protein A3 [Homo sapiens (human)]*. Available at: <https://www.ncbi.nlm.nih.gov/gene/6119/>. Accessed: 30/12/25

NEB (2025a), *AccI*. Available at: <https://www.neb.com/en/products/r0161-acci>. Accessed: 25/8/25

NEB (2025b), *BsaI-HFv2*. Available at: <https://www.neb.com/en/products/r3733-bsai-hf-v2>. Accessed: 25/8/25

NEB (2025c), *BsmBI-v2*. Available at: <https://www.neb.com/en/products/r0739-bsmbi-v2>. Accessed: 25/8/25

NEB (2025d), *DpnI*. Available at: <https://www.neb.com/en/products/r0176-dpni>. Accessed: 2/9/25

NEB (2025e), *EcoRI-HF*. Available at: <https://www.neb.com/en/products/r3101-ecori-hf>. Accessed: 25/8/25

NEB (2025f), *HindIII-HF*. Available at: <https://www.neb.com/en/products/r3104-hindiii-hf>. Accessed: 25/8/25

NEB (2025g), *KpnI-HF*. Available at: <https://www.neb.com/en/products/r3142-kpni-hf>. Accessed: 30/8/25

NEB (2025h), *NheI-HF*. Available at: <https://www.neb.com/en/products/r3131-nhei-hf>. Accessed: 25/8/25

NEB (2025i), *PvuI-HF*. Available at: <https://www.neb.com/en/products/r3150-pvui-hf>. Accessed: 25/8/25

NEB (2025j), *Q5 High-Fidelity DNA Polymerase*. Available at: <https://www.neb.com/en/products/m0491-q5-high-fidelity-dna-polymerase>. Accessed: 30/8/25

NEB (2025k), *Quick CIP*. Available at: <https://www.neb.com/en/products/m0525-quick-cip>. Accessed: 25/8/25

NEB (2025l), *T4 DNA Ligase*. Available at: <https://www.neb.com/en/products/m0202-t4-dna-ligase>. Accessed: 25/8/25

NEB (2025m), *T4 Polynucleotide Kinase*. Available at: <https://www.neb.com/en/products/m0201-t4-polynucleotide-kinase>. Accessed: 25/8/25.

NEB (2025n), *XhoI*. Available at: <https://www.neb.com/en/products/r0146-xhoi>. Accessed: 25/8/25

NEB (2026a), *BamHI-HF*. Available at: <https://www.neb.com/en/products/r3136-bamhi-hf>. Accessed: 31/1/26

NEB (2026b), *NotI-HF*. Available at: <https://www.neb.com/en/products/r3189-noti-hf>. Accessed: 31/1/26.

Neuwald, A.F. (2005), 'Evolutionary clues to eukaryotic DNA clamp-loading mechanisms: analysis of the functional constraints imposed on replication factor C AAA+ ATPases', *Nucleic Acids Research*, **33**(11), pp. 3614–3628. DOI: <https://doi.org/10.1093/nar/gki674>

Ngounou Wetie, A.G., Sokolowska, I., Woods, A.G. et al. (2014), 'Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches', *Cellular and Molecular Life Sciences: CMLS*, **71**(2), pp. 205–228. DOI: <https://doi.org/10.1007/s00018-013-1333-1>

Nishimura, K., Ishiai, M., Horikawa, K. et al. (2012), 'Mcm8 and Mcm9 form a complex that functions in homologous recombination repair induced by DNA interstrand crosslinks', *Molecular Cell*, **47**(4), pp. 511–522. DOI: <https://doi.org/10.1016/j.molcel.2012.05.047>

Nooren, I.M., and Thornton, J. M. (2003), 'Diversity of protein-protein interactions', *The EMBO Journal*, **22**(14), pp. 3486–3492. DOI: <https://doi.org/10.1093/emboj/cdg359>

Northall, S.J., Buckley, R., Jones, N. et al. (2017), 'DNA binding and unwinding by Hel308 helicase requires dual functions of a winged helix domain', *DNA Repair*, **57**, pp. 125–132. DOI: <https://doi.org/10.1016/j.dnarep.2017.07.005>

O'Donnell, M., Langston, L., and Stillman, B. (2013), 'Principles and concepts of DNA replication in bacteria, archaea, and eukarya', *Cold Spring Harbor Perspectives in Biology*, **5**(7), a010108. DOI: <https://doi.org/10.1101/cshperspect.a010108>

O'Donnell, M.E., and Li, H. (2018), 'The ring-shaped hexameric helicases that function at DNA replication forks', *Nature Structural & Molecular Biology*, **25**(2), pp. 122–130. DOI: <https://doi.org/10.1038/s41594-018-0024-x>

Ohad, N., Shichrur, K., and Yalovsky, S. (2007), 'The analysis of protein-protein interactions in plants by bimolecular fluorescence complementation', *Plant Physiology*, **145**(4), pp. 1090–1099. DOI: <https://doi.org/10.1104/pp.107.107284>

Olson, O., Pellicciari, S., Heron, E.D. et al. (2024), 'A common mechanism for recruiting the Rrm3 and RTEL1 accessory helicases to the eukaryotic replisome', *The EMBO Journal*, **43**(18), pp. 3846–3875. DOI: <https://doi.org/10.1038/s44318-024-00168-4>

- Oshige, M., Aoyagi, N., Harris, P.V. et al. (1999), 'A new DNA polymerase species from *Drosophila melanogaster*: a probable mus308 gene product', *Mutation Research*, **433**(3), pp. 183–192. DOI: [https://doi.org/10.1016/s0921-8777\(99\)00005-1](https://doi.org/10.1016/s0921-8777(99)00005-1)
- Oyama, T., Oka, H., Mayanagi, K. et al. (2009), 'Atomic structures and functional implications of the archaeal RecQ-like helicase Hjm', *BMC Structural Biology*, **9**, pp. 2. DOI: <https://doi.org/10.1186/1472-6807-9-2>
- Oyama, T., Ishino, S., Fujino, S. et al. (2011), 'Architectures of archaeal GINS complexes, essential DNA replication initiation factors', *BMC Biology*, **9**, pp. 28. DOI: <https://doi.org/10.1186/1741-7007-9-28>
- Ozawa, K., Horan, N.P., Robinson, A. (2013), 'Proofreading exonuclease on a tether: the complex between the *E. coli* DNA polymerase III subunits α , epsilon, θ and β reveals a highly flexible arrangement of the proofreading domain', *Nucleic Acids Research*, **41**(10), pp. 5354–5367. DOI: <https://doi.org/10.1093/nar/gkt162>
- Pace, C.N., Fu, H., Fryar, K.L. et al. (2011), 'Contribution of hydrophobic interactions to protein stability', *Journal of Molecular Biology*, **408**(3), pp. 514–528. DOI: <https://doi.org/10.1016/j.jmb.2011.02.053>
- Paliy, O., and Gunasekera, T.S. (2007), 'Growth of *E. coli* BL21 in minimal media with different gluconeogenic carbon sources and salt contents', *Applied Microbiology and Biotechnology*, **73**(5), pp. 1169–1172. DOI: <https://doi.org/10.1007/s00253-006-0554-8>
- Pan, S., and Chen, R. (2022), 'Pathological implication of protein post-translational modifications in cancer', *Molecular Aspects of Medicine*, **86**, 101097. DOI: <https://doi.org/10.1016/j.mam.2022.101097>
- Pang, M., McConnell, M., and Fisher, P.A. (2005), 'The *Drosophila mus 308* gene product, implicated in tolerance of DNA interstrand crosslinks, is a nuclear protein found in both ovaries and embryos', *DNA Repair*, **4**(9), pp. 971–982. DOI: <https://doi.org/10.1016/j.dnarep.2005.04.020>
- Park, J.S., Wang, M., Park, S.J. et al. (1999), 'Zinc finger of replication protein A, a non-DNA binding element, regulates its DNA binding activity through redox', *The Journal of Biological Chemistry*, **274**(41), pp. 29075–29080. DOI: <https://doi.org/10.1074/jbc.274.41.29075>

Parker, M.W., Botchan, M.R., and Berger, J.M. (2017), 'Mechanisms and regulation of DNA replication initiation in eukaryotes', *Critical Reviews in Biochemistry and Molecular Biology*, **52**(2), pp. 107–144. DOI: <https://doi.org/10.1080/10409238.2016.1274717>

Parkes, A. (2024), 'Molecular biology of the DExD-box helicases DDX49 and DDX52', PhD Thesis, University of Nottingham. Available at: <https://eprints.nottingham.ac.uk/77577>. Accessed: 3/1/26

Patel, J.A., and Kim, H. (2023), 'The TIMELESS effort for timely DNA replication and protection', *Cellular and Molecular Life Sciences: CMLS*, **80**(4), pp. 84. DOI: <https://doi.org/10.1007/s00018-023-04738-3>

Paudel, S., Severin, G.B., Pirani, A. et al. (2025), 'Multiplexed PCR to measure *in situ* growth rates of uropathogenic *E. coli* during experimental urinary tract infection', *Applied and Environmental Microbiology*, **91**(8), e0238224. DOI: <https://doi.org/10.1128/aem.02382-24>

Paul, T., Lee, I. R., Pangeni, S. et al. (2025), 'Mechanistic insights into direct DNA and RNA strand transfer and dynamic protein exchange of SSB and RPA', *Nucleic Acids Research*, **53**(12), gkaf642. DOI: <https://doi.org/10.1093/nar/gkaf642>

Peacock, J., and Jaynes, J.B. (2017), 'Using competition assays to quantitatively model cooperative binding by transcription factors and other ligands', *Biochimica et Biophysica Acta. General Subjects*, **1861**(11 Pt A), pp. 2789–2801. DOI: <https://doi.org/10.1016/j.bbagen.2017.07.024>

Peng, X., Wang, J., Peng, W. (2017), 'Protein-protein interactions: detection, reliability assessment and applications', *Briefings in Bioinformatics*, **18**(5), pp. 798–819. DOI: <https://doi.org/10.1093/bib/bbw066>

Perera, H.M., and Trakselis, M.A. (2022), 'Determining translocation orientations of nucleic acid helicases', *Methods (San Diego, Calif.)*, **204**, pp. 160–171. DOI: <https://doi.org/10.1016/j.ymeth.2021.11.001>

Pettersen, E.F., Goddard, T.D., Huang, C.C. et al. (2021), 'UCSF ChimeraX: Structure visualization for researchers, educators, and developers', *Protein science: a publication of the Protein Society*, **30**(1), pp. 70–82. DOI: <https://doi.org/10.1002/pro.3943>

- Pienaar, J.A., Singh, A., and Barnard, T.G. (2016), 'The viable but non-culturable state in pathogenic *Escherichia coli*: A general review', *African Journal of Laboratory Medicine*, **5**(1), pp. 368. DOI: <https://doi.org/10.4102/ajlm.v5i1.368>
- Pinilla-Redondo, R., Russel, J., Mayo-Muñoz, D. (2022), 'CRISPR-Cas systems are widespread accessory elements across bacterial and archaeal plasmids', *Nucleic Acids Research*, **50**(8), pp. 4315–4328. DOI: <https://doi.org/10.1093/nar/gkab859>
- Podust, V.N., Podust, L.M., Müller, F. et al. (1995), 'DNA polymerase delta holoenzyme: action on single-stranded DNA and on double-stranded DNA in the presence of replicative DNA helicases', *Biochemistry*, **34**(15), pp. 5003–5010. DOI: <https://doi.org/10.1021/bi00015a011>
- Podust, V.N., Chang, L.S., Ott, R. et al. (2002), 'Reconstitution of human DNA polymerase delta using recombinant baculoviruses: the p12 subunit potentiates DNA polymerizing activity of the four-subunit enzyme', *The Journal of Biological Chemistry*, **277**(6), pp. 3894–3901. DOI: <https://doi.org/10.1074/jbc.M109684200>
- Pohler, J.R., Otterlei, M., and Warbrick, E. (2005), 'An *in vivo* analysis of the localisation and interactions of human p66 DNA polymerase delta subunit', *BMC Molecular Biology*, **6**, pp. 17. DOI: <https://doi.org/10.1186/1471-2199-6-17>
- Pommier, Y., Sun, Y., Huang, S.N. et al. (2016), 'Roles of eukaryotic topoisomerases in transcription, replication and genomic stability', *Nature Reviews. Molecular cell Biology*, **17**(11), pp. 703–721. DOI: <https://doi.org/10.1038/nrm.2016.111>
- Pommier, Y., Nussenzweig, A., Takeda, S. et al. (2022), 'Human topoisomerases and their roles in genome stability and organization', *Nature Reviews. Molecular Cell Biology*, **23**(6), pp. 407–427. DOI: <https://doi.org/10.1038/s41580-022-00452-3>
- Pontén, J., and Saksela, E. (1967), 'Two established *in vitro* cell lines from human mesenchymal tumours', *International Journal of Cancer*, **2**(5), pp. 434–447. DOI: <https://doi.org/10.1002/ijc.2910020505>
- Postow, L., Crisona, N.J., Peter, B.J. et al. (2001), 'Topological challenges to DNA replication: conformations at the fork', *Proceedings of the National Academy of Sciences of the United States of America*, **98**(15), pp. 8219–8226. DOI: <https://doi.org/10.1073/pnas.111006998>

- Prado, F. (2014), 'Homologous recombination maintenance of genome integrity during DNA damage tolerance', *Molecular & Cellular Oncology*, **1**(2), e957039. DOI: <https://doi.org/10.4161/23723548.2014.957039>
- Pratt, O.S., Elliott, L.G., Haon, M. et al. (2025), 'AlphaFold 2, but not AlphaFold 3, predicts confident but unrealistic β -solenoid structures for repeat proteins', *Computational and Structural Biotechnology Journal*, **27**, pp. 467–477. DOI: <https://doi.org/10.1016/j.csbj.2025.01.016>
- Prindle, M.J., and Loeb, L.A. (2012), 'DNA polymerase delta in DNA replication and genome maintenance', *Environmental and Molecular Mutagenesis*, **53**(9), pp. 666–682. DOI: <https://doi.org/10.1002/em.21745>
- Promega (2009), *Wizard PlusSV Minipreps DNA Purification System Quick Protocol*. Available at: https://www.promega.com/-/media/files/resources/protcards/wizard-plus-sv-minipreps-dna-purification-system-quick-protocol.pdf?rev=8293c52acfca4de495f1df90a7753d50&sc_lang=en. Accessed: 31/8/25
- Promega (2024), *Wizard SV Gel and PCR Clean-Up System Quick Protocol*. Available at: <https://www.promega.com/-/media/files/resources/protcards/wizard-sv-gel-and-pcr-clean-up-system-quick-protocol.pdf>. Accessed: 1/9/25
- Prusiner, S.B. (1982), 'Novel proteinaceous infectious particles cause scrapie', *Science (New York, N.Y.)*, **216**(4542), pp. 136–144. DOI: <https://doi.org/10.1126/science.6801762>
- Prusiner, S.B. (1998), 'Prions', *Proceedings of the National Academy of Sciences of the United States of America*, **95**(23), pp. 13363–13383. DOI: <https://doi.org/10.1073/pnas.95.23.13363>
- Prusiner, S.B., Scott, M.R., DeArmond, S.J. et al. (1998), 'Prion protein biology', *Cell*, **93**(3), pp. 337–348. DOI: [https://doi.org/10.1016/s0092-8674\(00\)81163-0](https://doi.org/10.1016/s0092-8674(00)81163-0)
- Pustovalova, Y., Magalhães, M.T., D'Souza, S. et al. (2016), 'Interaction between the Rev1 C-Terminal Domain and the PolD3 Subunit of Pol ζ Suggests a Mechanism of Polymerase Exchange upon Rev1/Pol ζ -Dependent Translesion Synthesis', *Biochemistry*, **55**(13), pp. 2043–2053. DOI: <https://doi.org/10.1021/acs.biochem.5b01282>

Pyo, J., Lee, W., Choi, E.Y. et al. (2023), 'Qualitative Research in Healthcare: Necessity and Characteristics', *Journal of Preventive Medicine and Public Health = Yebang Uihakhoe Chi*, **56**(1), pp. 12–20. DOI: <https://doi.org/10.3961/jpmph.22.451>

Qutbi, M. (2024), 'Quantitative Performance Evaluation of Commonly Used Colormaps for Image Display in Myocardial Perfusion Imaging: Analysis based on Perceptual Metrics', *Molecular Imaging and Radionuclide Therapy*, **33**(2), pp. 94–105. DOI: <https://doi.org/10.4274/mirt.galenos.2024.34711>

Radford, H.M., Toft, C.J., Sorenson, A.E., and Schaeffer, P. M. (2023), 'Inhibition of Replication Fork Formation and Progression: Targeting the Replication Initiation and Primosomal Proteins', *International Journal of Molecular Sciences*, **24**(10), pp. 8802. DOI: <https://doi.org/10.3390/ijms24108802>

Radhakrishnan, S.K., Jette, N., and Lees-Miller, S.P. (2014), 'Non-homologous end joining: emerging themes and unanswered questions', *DNA Repair*, **17**, pp. 2–8. DOI: <https://doi.org/10.1016/j.dnarep.2014.01.009>

Rajamani, D., Thiel, S., Vajda, S. et al. (2004), 'Anchor residues in protein-protein interactions', *Proceedings of the National Academy of Sciences of the United States of America*, **101**(31), pp. 11287–11292. DOI: <https://doi.org/10.1073/pnas.0401942101>

Ramamurthy, T., Ghosh, A., Pazhani, G.P. et al. (2014), 'Current Perspectives on Viable but Non-Culturable (VBNC) Pathogenic Bacteria', *Frontiers in Public Health*, **2**, pp. 103. DOI: <https://doi.org/10.3389/fpubh.2014.00103>

Rao, V.S., Srinivas, K., Sujini, G.N. et al. (2014), 'Protein-protein interaction detection: methods and analysis', *International Journal of Proteomics*, **2014**, 147648. DOI: <https://doi.org/10.1155/2014/147648>

Rashid, F., and Berger, J.M. (2025), 'How bacteria initiate DNA replication comes into focus', *BioEssays: News and reviews in molecular, cellular and developmental biology*, **47**(1), e2400151. DOI: <https://doi.org/10.1002/bies.202400151>

Rehman, I., and Botelho, S. (2024), 'Biochemistry, Tertiary Protein Structure', in: *StatPearls [Internet]*. Updated 25th May 2024. Treasure Island (FL). Available at: <https://www.ncbi.nlm.nih.gov/books/NBK470269/>

- Ren, H., Ou, Q., Pu, Q. et al. (2024), 'Comprehensive Review on Bimolecular Fluorescence Complementation and Its Application in Deciphering Protein-Protein Interactions in Cell Signaling Pathways', *Biomolecules*, **14**(7), pp. 859. DOI: <https://doi.org/10.3390/biom14070859>
- Reyes, G.X., Schmidt, T.T. Kolodner, R.D. et al. (2015), 'New insights into the mechanism of DNA mismatch repair', *Chromosoma*, **124**(4), pp. 443–462. DOI: <https://doi.org/10.1007/s00412-015-0514-0>
- Reyes, G.X., Kolodziejczak, A., Devakumar, L.J.P.S. et al. (2021), 'Ligation of newly replicated DNA controls the timing of DNA mismatch repair', *Current Biology: CB*, **31**(6), pp. 1268–1276.e6. DOI: <https://doi.org/10.1016/j.cub.2020.12.018>
- Richards, J.D., Johnson, K.A., Liu, H. et al. (2008), 'Structure of the DNA repair helicase hel308 reveals DNA binding and autoinhibitory domains', *The Journal of Biological Chemistry*, **283**(8), pp. 5118–5126. DOI: <https://doi.org/10.1074/jbc.M707548200>
- Ringe, D. (1995), 'What makes a binding site a binding site?', *Current Opinion in Structural Biology*, **5**(6), pp. 825–829. DOI: [https://doi.org/10.1016/0959-440x\(95\)80017-4](https://doi.org/10.1016/0959-440x(95)80017-4)
- Robertson, A.B., Klungland, A., Rognes, T. et al. (2009), 'DNA repair in mammalian cells: Base excision repair: the long and short of it', *Cellular and Molecular Life Sciences: CMLS*, **66**(6), pp. 981–993. DOI: <https://doi.org/10.1007/s00018-009-8736-z>
- Rodgers, K., and McVey, M. (2016), 'Error-Prone Repair of DNA Double-Strand Breaks', *Journal of Cellular Physiology*, **231**(1), pp. 15–24. DOI: <https://doi.org/10.1002/jcp.2505>
- Romanowski, P., and Madine, M.A. (1996), 'Mechanisms restricting DNA replication to once per cell cycle: MCMS, pre-replicative complexes and kinases', *Trends in Cell Biology*, **6**(5), pp. 184–188. DOI: [https://doi.org/10.1016/0962-8924\(96\)10015-5](https://doi.org/10.1016/0962-8924(96)10015-5)
- Rose, D., Thomas, W., and Holm, C. (1990), 'Segregation of recombined chromosomes in meiosis I requires DNA topoisomerase II', *Cell*, **60**(6), pp. 1009–1017. DOI: [https://doi.org/10.1016/0092-8674\(90\)90349-j](https://doi.org/10.1016/0092-8674(90)90349-j)

Rouillon, C., and White, M.F. (2011), 'The evolution and mechanisms of nucleotide excision repair proteins', *Research in Microbiology*, **162**(1), pp. 19–26. DOI: <https://doi.org/10.1016/j.resmic.2010.09.003>

Sahu, S.R., Utkalaja, B.G., Patel, S.K. et al. (2023), 'Spot Assay and Colony Forming Unit (CFU) Analyses-based sensitivity test for *Candida albicans* and *Saccharomyces cerevisiae*', *Bio-Protocol*, **13**(21), e4872. DOI: <https://doi.org/10.21769/BioProtoc.4872>

Saier, M.H., Jr. (2019), 'Understanding the Genetic Code', *Journal of Bacteriology*, **201**(15), e00091-19. DOI: <https://doi.org/10.1128/JB.00091-19>

Saito, H., Kashida, S., Inoue, T. et al. (2007), 'The role of peptide motifs in the evolution of a protein network', *Nucleic Acids Research*, **35**(19), pp. 6357–6366. DOI: <https://doi.org/10.1093/nar/gkm692>

Saleh-Gohari, N., Bryant, H.E., Schultz, N. et al. (2005), 'Spontaneous homologous recombination is induced by collapsed replication forks that are caused by endogenous DNA single-strand breaks', *Molecular and Cellular Biology*, **25**(16), pp. 7158–7169. DOI: <https://doi.org/10.1128/MCB.25.16.7158-7169.2005>

Samson, R.Y., Xu, Y., Gadelha, C. et al. (2013), 'Specificity and function of archaeal DNA replication initiator proteins', *Cell Reports*, **3**(2), pp. 485–496. DOI: <https://doi.org/10.1016/j.celrep.2013.01.002>

Samson, R.Y., Abeyrathne, P.D., and Bell, S.D. (2016), 'Mechanism of Archaeal MCM Helicase Recruitment to DNA Replication Origins', *Molecular Cell*, **61**(2), pp. 287–296. DOI: <https://doi.org/10.1016/j.molcel.2015.12.005>

Sancar, A. (1996), 'DNA excision repair', *Annual Review of Biochemistry*, **65**, pp. 43–81. DOI: <https://doi.org/10.1146/annurev.bi.65.070196.000355>

Sanchez-Pulido, L., and Ponting, C.P. (2011), 'Cdc45: the missing RecJ ortholog in eukaryotes?', *Bioinformatics (Oxford, England)*, **27**(14), pp. 1885–1888. DOI: <https://doi.org/10.1093/bioinformatics/btr332>

- San Martin, M.C., Stamford, N.P., Dammerova, N. et al. (1995), 'A structural model for the Escherichia coli DnaB helicase based on electron microscopy data', *Journal of Structural Biology*, **114**(3), pp. 167–176. DOI: <https://doi.org/10.1006/jsbi.1995.1016>
- Sardiu, M.E., and Washburn, M.P. (2011), 'Building protein-protein interaction networks with proteomics and informatics tools', *The Journal of Biological Chemistry*, **286**(27), pp. 23645–23651. DOI: <https://doi.org/10.1074/jbc.R110.174052>
- Sarker, A.H., Trego, K.S., Zhang, W. et al. (2020), 'Thirdhand smoke exposure causes replication stress and impaired transcription in human lung cells', *Environmental and Molecular Mutagenesis*, **61**(6), pp. 635–646. DOI: <https://doi.org/10.1002/em.22372>
- Satoh, H., Sasatani, Y., Miyazaki, K. et al. (2022), 'Significance of Polar Charged Amino Acids in Compound Mutations in EGFR-mutated Patients Treated With First-line Afatinib', *In Vivo (Athens, Greece)*, **36**(4), pp. 1829–1833. DOI: <https://doi.org/10.21873/invivo.12899>
- Sarvari, G., and Boehr, D.D. (2026), 'Structure, Function and Inhibition of Helicases Involved in Virus Infection', *Biomolecules*, **16**(2), pp. 273. DOI: <https://doi.org/10.3390/biom16020273>
- Saxena, S., and Zou, L. (2022), 'Hallmarks of DNA replication stress', *Molecular Cell*, **82**(12), pp. 2298–2314. DOI: <https://doi.org/10.1016/j.molcel.2022.05.004>
- Schaeffer, R.D., Kinch, L.N., Liao, Y. et al. (2016), 'Classification of proteins with shared motifs and internal repeats in the ECOD database', *Protein Science: a publication of the Protein Society*, **25**(7), pp. 1188–1203. DOI: <https://doi.org/10.1002/pro.2893>
- Schafer, K.A. (1998), 'The cell cycle: a review', *Veterinary Pathology*, **35**(6), pp. 461–478. DOI: <https://doi.org/10.1177/030098589803500601>
- Schärer, O.D. (2013), 'Nucleotide excision repair in eukaryotes', *Cold Spring Harbor Perspectives in Biology*, **5**(10), a012609. DOI: <https://doi.org/10.1101/cshperspect.a012609>
- Schilder, J., and Ubbink, M. (2013), 'Formation of transient protein complexes', *Current Opinion in Structural Biology*, **23**(6), pp. 911–918. DOI: <https://doi.org/10.1016/j.sbi.2013.07.009>

- Schillinger, C., Boisguerin, P., and Krause, G. (2009), 'Domain Interaction Footprint: a multi-classification approach to predict domain-peptide interactions', *Bioinformatics (Oxford, England)*, **25**(13), pp. 1632–1639. DOI: <https://doi.org/10.1093/bioinformatics/btp264>
- Schmitt, M.W., Matsumoto, Y., and Loeb, L.A. (2009), 'High fidelity and lesion bypass capability of human DNA polymerase delta', *Biochimie*, **91**(9), pp. 1163–1172. DOI: <https://doi.org/10.1016/j.biochi.2009.06.007>
- Schofield, M.J., and Hsieh, P. (2003), 'DNA mismatch repair: molecular mechanisms and biological function', *Annual Review of Microbiology*, **57**, pp. 579–608. DOI: <https://doi.org/10.1146/annurev.micro.57.030502.090847>
- Schreiber, G., Haran, G., and Zhou, H.X. (2009), 'Fundamental aspects of protein-protein association kinetics', *Chemical Reviews*, **109**(3), pp. 839–860. DOI: <https://doi.org/10.1021/cr800373w>
- Schreiber, G., and Keating, A.E. (2011), 'Protein binding specificity versus promiscuity', *Current Opinion in Structural Biology*, **21**(1), pp. 50–61. DOI: <https://doi.org/10.1016/j.sbi.2010.10.002>
- Schumacher, S.B., Stucki, M., and Hübscher, U. (2000), 'The N-terminal region of DNA polymerase delta catalytic subunit is necessary for holoenzyme function', *Nucleic Acids Research*, **28**(2), pp. 620–625. DOI: <https://doi.org/10.1093/nar/28.2.620>
- Schvartzman, J.B., Hernández, P., Krimer, D.B. et al. (2019), 'Closing the DNA replication cycle: from simple circular molecules to supercoiled and knotted DNA catenanes', *Nucleic Acids Research*, **47**(14), pp. 7182–7198. DOI: <https://doi.org/10.1093/nar/gkz586>
- Sclafani, R.A., and Holzen, T.M. (2007), 'Cell cycle regulation of DNA replication', *Annual Review of Genetics*, **41**, pp. 237–280. DOI: <https://doi.org/10.1146/annurev.genet.41.110306.130308>
- Scott, O.B., Gu, J., and Chan, A.W.E. (2022), 'Classification of Protein-Binding Sites Using a Spherical Convolutional Neural Network', *Journal of Chemical Information and Modeling*, **62**(22), pp. 5383–5396. DOI: <https://doi.org/10.1021/acs.jcim.2c00832>

- Seitz, E.M., Brockman, J.P., Sandler, S.J. et al. (1998), 'RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange', *Genes & Development*, **12**(9), pp. 1248–1253. DOI: <https://doi.org/10.1101/gad.12.9.1248>
- Seki, M., Marini, F., and Wood, R.D. (2003), 'POLQ (Pol theta), a DNA polymerase and DNA-dependent ATPase in human cells', *Nucleic Acids Research*, **31**(21), pp. 6117–6126. DOI: <https://doi.org/10.1093/nar/gkg814>
- Sekimizu, K., Bramhill, D., and Kornberg, A. (1987), 'ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome', *Cell*, **50**(2), pp. 259–265. DOI: [https://doi.org/10.1016/0092-8674\(87\)90221-2](https://doi.org/10.1016/0092-8674(87)90221-2)
- Seo, K., and Ichihashi, N. (2023), 'Investigation of Compatibility between DNA Replication, Transcription, and Translation for *in Vitro* Central Dogma', *ACS Synthetic Biology*, **12**(6), pp. 1813–1822. DOI: <https://doi.org/10.1021/acssynbio.3c00130>
- Serdyuk, A., and Allers, T. (2025), 'DNA Replication in Time and Space: The Archaeal Dimension', *DNA*, **5**(2), pp. 24. DOI: <https://doi.org/10.3390/dna5020024>
- Seychell, B.C., and Beck, T. (2021), 'Molecular basis for protein-protein interactions', *Beilstein Journal of Organic Chemistry*, **17**, pp. 1–10. DOI: <https://doi.org/10.3762/bjoc.17.1>
- Shah Punatar, R., Martin, M.J., Wyatt, H D. et al. (2017), 'Resolution of single and double Holliday junction recombination intermediates by GEN1', *Proceedings of the National Academy of Sciences of the United States of America*, **114**(3), pp. 443–450. DOI: <https://doi.org/10.1073/pnas.1619790114>
- Shang, A., Shao, S., Zhao, L. et al. (2024), 'Far-Red Fluorescent Proteins: Tools for Advancing *In Vivo* Imaging. *Biosensors*, **14**(8), pp. 359. DOI: <https://doi.org/10.3390/bios14080359>
- Sharief, F.S., Vojta, P.J., Ropp, P.A. et al. (1999), 'Cloning and chromosomal mapping of the human DNA polymerase theta (POLQ), the eighth human DNA polymerase', *Genomics*, **59**(1), pp. 90–96. DOI: <https://doi.org/10.1006/geno.1999.5843>
- Shaw, M.K., Marr, A.G., and Ingraham, J.L. (1971), 'Determination of the minimal temperature for growth of *Escherichia coli*', *Journal of Bacteriology*, **105**(2), pp. 683–684. DOI: <https://doi.org/10.1128/jb.105.2.683-684.1971>

- Shereda, R.D., Kozlov, A.G., Lohman, T.M. et al. (2008), 'SSB as an organizer/mobilizer of genome maintenance complexes', *Critical Reviews in Biochemistry and Molecular Biology*, **43**(5), pp. 289–318. DOI: <https://doi.org/10.1080/10409230802341296>
- Shin, Y., Hedglin, M., and Murakami, K.S. (2025), 'Cryo-EM structure of apo-form human DNA polymerase δ elucidates its minimal DNA synthesis activity without PCNA', *The Journal of Biological Chemistry*, **301**(4), pp. 108342. DOI: <https://doi.org/10.1016/j.jbc.2025.108342>
- Shmakov, S.A., Utkina, I., Wolf, Y.I. et al. (2020), 'CRISPR Arrays Away from *cas* Genes', *The CRISPR Journal*, **3**(6), pp. 535–549. DOI: <https://doi.org/10.1089/crispr.2020.0062>
- Sigurdson, C.J., Bartz, J.C., and Glatzel, M. (2019), 'Cellular and Molecular Mechanisms of Prion Disease', *Annual Review of Pathology*, **14**, pp. 497–516. DOI: <https://doi.org/10.1146/annurev-pathmechdis-012418-013109>
- Simonetta, K.R., Kazmirski, S.L., Goedken, E.R. et al. (2009), 'The mechanism of ATP-dependent primer-template recognition by a clamp loader complex', *Cell*, **137**(4), pp. 659–671. DOI: <https://doi.org/10.1016/j.cell.2009.03.044>
- Singleton, M.R., and Wigley, D.B. (2002), 'Modularity and specialization in superfamily 1 and 2 helicases', *Journal of Bacteriology*, **184**(7), pp. 1819–1826. DOI: <https://doi.org/10.1128/JB.184.7.1819-1826.2002>
- Singleton, M.R., Dillingham, M.S., and Wigley, D.B. (2007), 'Structure and mechanism of helicases and nucleic acid translocases', *Annual Review of Biochemistry*, **76**, pp. 23–50. DOI: <https://doi.org/10.1146/annurev.biochem.76.052305.115300>
- Sishc, B.J., and Davis, A.J. (2017), 'The Role of the Core Non-Homologous End Joining Factors in Carcinogenesis and Cancer', *Cancers*, **9**(7), pp. 81. DOI: <https://doi.org/10.3390/cancers9070081>
- Sleight, S.C., Wigginton, N.S., and Lenski, R.E. (2006), 'Increased susceptibility to repeated freeze-thaw cycles in *Escherichia coli* following long-term evolution in a benign environment', *BMC Evolutionary Biology*, **6**, pp. 104. DOI: <https://doi.org/10.1186/1471-2148-6-104>

Smith, M.C., Nasser, N.S., Morin, E.J. et al. (2024), 'Redox-active molecules in bacterial cultivation media produce photocurrent', *Biosensors and Bioelectronics: X*, **21**, 100558. DOI: <https://doi.org/10.1016/j.biosx.2024.100558>

SnapGene (2025), *pEGFP-C1*. Available at: https://www.snapgene.com/plasmids/fluorescent_protein_genes_and_plasmids/pEGFP-C1.

Accessed: 24/8/25

Sneed, J.L., Grossi, S.M., Tappin, I. et al. (2013), 'Reconstitution of recombination-associated DNA synthesis with human proteins', *Nucleic Acids Research*, **41**(9), pp. 4913–4925. DOI: <https://doi.org/10.1093/nar/gkt192>

Snider, J., Thibault, G., and Houry, W.A. (2008), 'The AAA+ superfamily of functionally diverse proteins', *Genome Biology*, **9**(4), pp. 216. DOI: <https://doi.org/10.1186/gb-2008-9-4-216>

Song, H.Y., Shen, R., Mahasin, H. et al. (2023), 'DNA replication: Mechanisms and therapeutic interventions for diseases', *MedComm*, **4**(1), e210. DOI: <https://doi.org/10.1002/mco2.210>

Sonneville, R., Moreno, S.P., Knebel, A. et al. (2017), 'CUL-2^{LRR-1} and UBXN-3 drive replisome disassembly during DNA replication termination and mitosis' *Nature Cell Biology*, **19**(5), pp. 468–479. DOI: <https://doi.org/10.1038/ncb3500>

Spasov, D.S. (2024), 'Binding Affinity Determination in Drug Design: Insights from Lock and Key, Induced Fit, Conformational Selection, and Inhibitor Trapping Models', *International Journal of Molecular Sciences*, **25**(13), pp. 7124. DOI: <https://doi.org/10.3390/ijms25137124>

Speck, C., Chen, Z., Li, H. et al. (2005), 'ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA', *Nature Structural & Molecular Biology*, **12**(11), pp. 965–971. DOI: <https://doi.org/10.1038/nsmb1002>

Speer, S.L., Zheng, W., Jiang, X. et al. (2021), 'The intracellular environment affects protein-protein interactions', *Proceedings of the National Academy of Sciences of the United States of America*, **118**(11), e2019918118. DOI: <https://doi.org/10.1073/pnas.2019918118>

Spengelink, L.M., Lewis, J.S., Jergic, S. et al. (2019), 'Recycling of single-stranded DNA-binding protein by the bacterial replisome', *Nucleic Acids Research*, **47**(8), pp. 4111–4123. DOI: <https://doi.org/10.1093/nar/gkz090>

- Spinks, R.R., Spenkelink, L.M., Stratmann, S.A. et al. (2021), 'DnaB helicase dynamics in bacterial DNA replication resolved by single-molecule studies', *Nucleic Acids Research*, **49**(12), pp. 6804–6816. DOI: <https://doi.org/10.1093/nar/gkab493>
- Spivak, G. (2015), 'Nucleotide excision repair in humans', *DNA Repair*, **36**, pp. 13–18. DOI: <https://doi.org/10.1016/j.dnarep.2015.09.003>
- Srimani, S., Schmidt, C.X., Gómez-Serranillos, M.P. et al. (2022), 'Modulation of Cellular Circadian Rhythms by Secondary Metabolites of Lichens', *Frontiers in Cellular Neuroscience*, **16**, 907308. DOI: <https://doi.org/10.3389/fncel.2022.907308>
- Stamereilers, C., Wong, S., and Tsourkas, P.K. (2021), 'Characterization of CRISPR Spacer and Protospacer Sequences in *Paenibacillus larvae* and Its Bacteriophages', *Viruses*, **13**(3), pp. 459. DOI: <https://doi.org/10.3390/v13030459>
- Stano, N.M., Jeong, Y.J., Donmez, I. et al. (2005), 'DNA synthesis provides the driving force to accelerate DNA unwinding by a helicase', *Nature*, **435**(7040), pp. 370–373. DOI: <https://doi.org/10.1038/nature03615>
- Stark, R. (2023), 'Protein-mediated interactions in the dynamic regulation of acute inflammation', *Biocell: official journal of the Sociedades Latinoamericanas de Microscopia Electronica ... et. al*, **47**(6), pp. 1191–1198. DOI: <https://doi.org/10.32604/biocell.2023.027838>
- Stauffer, M.E., and Chazin, W.J. (2004), 'Physical interaction between replication protein A and Rad51 promotes exchange on single-stranded DNA', *The Journal of Biological Chemistry*, **279**(24), pp. 25638–25645. DOI: <https://doi.org/10.1074/jbc.M400029200>
- Stein, A., Panjkovich, A., and Aloy, P. (2009), '3did Update: domain-domain and peptide-mediated interactions of known 3D structure', *Nucleic Acids Research*, **37**(Database issue), pp. D300–D304. DOI: <https://doi.org/10.1093/nar/gkn690>
- Stillman, B. (2008), 'DNA polymerases at the replication fork in eukaryotes', *Molecular Cell*, **30**(3), pp. 259–260. DOI: <https://doi.org/10.1016/j.molcel.2008.04.011>
- Stollar, E.J., and Smith, D.P. (2020), 'Uncovering protein structure', *Essays in Biochemistry*, **64**(4), pp. 649–680. DOI: <https://doi.org/10.1042/EBC20190042>

Sugino, A., Hirose, S., and Okazaki, R. (1972), 'RNA-linked nascent DNA fragments in *Escherichia coli*', *Proceedings of the National Academy of Sciences of the United States of America*, **69**(7), pp. 1863–1867. DOI: <https://doi.org/10.1073/pnas.69.7.1863>

Suhasini, A.N., and Brosh, R.M., Jr (2010), 'Mechanistic and biological aspects of helicase action on damaged DNA', *Cell Cycle (Georgetown, Tex.)*, **9**(12), pp. 2317–2329. DOI: <https://doi.org/10.4161/cc.9.12.11902>

Sun, P.D., Foster, C.E., and Boyington, J.C. (2004), 'Overview of protein structural and functional folds', *Current Protocols in Protein Science*, **17**(1), pp. 1711–17.1. DOI: <https://doi.org/10.1002/0471140864.ps1701s35>

Swart, J.R., and Griep, M.A. (1993), 'Primase from *Escherichia coli* primes single-stranded templates in the absence of single-stranded DNA-binding protein or other auxiliary proteins. Template sequence requirements based on the bacteriophage G4 complementary strand origin and Okazaki fragment initiation sites', *The Journal of Biological Chemistry*, **268**(17), pp. 12970–12976.

Syeda, A.H., Dimude, J.U., Skovgaard, O. et al. (2020), 'Too Much of a Good Thing: How Ectopic DNA Replication Affects Bacterial Replication Dynamics', *Frontiers in Microbiology*, **11**, pp. 534. DOI: <https://doi.org/10.3389/fmicb.2020.00534>

Takara, T.J., and Bell, S.P. (2011), 'Multiple Cdt1 molecules act at each origin to load replication-competent Mcm2-7 helicases', *The EMBO Journal*, **30**(24), pp. 4885–4896. DOI: <https://doi.org/10.1038/emboj.2011.394>

Takata, K., Reh, S., Tomida, J. et al. (2013), 'Human DNA helicase HELQ participates in DNA interstrand crosslink tolerance with ATR and RAD51 paralogs', *Nature Communications*, **4**, pp. 2338. DOI: <https://doi.org/10.1038/ncomms3338>

Talavera, D., Robertson, D.L., and Lovell, S.C. (2011), 'Characterization of protein-protein interaction interfaces from a single species', *PLOS One*, **6**(6), e21053. DOI: <https://doi.org/10.1371/journal.pone.0021053>

Tang, N., Wen, W., Liu, Z. et al. (2023), 'HELQ as a DNA helicase: Its novel role in normal cell function and tumorigenesis (Review)', *Oncology Reports*, **50**(6), pp. 220. DOI: <https://doi.org/10.3892/or.2023.8657>

Teilum, K., Olsen, J.G., and Kragelund, B.B. (2021), 'On the specificity of protein-protein interactions in the context of disorder', *The Biochemical Journal*, **478**(11), pp. 2035–2050. DOI: <https://doi.org/10.1042/BCJ20200828>

Temin, H.M., and Mizutani, S. (1970), 'RNA-dependent DNA polymerase in virions of Rous sarcoma virus', *Nature*, **226**(5252), pp. 1211–1213. DOI: <https://doi.org/10.1038/2261211a0>

Thean, R.K., Ong, D.X., Heng, Z.S. et al. (2021), 'To Plate or to Simply Unfreeze, That Is the Question for Optimal Plasmid Extraction', *Journal of Biomolecular Techniques: JBT*, **32**(2), pp. 57–62. DOI: <https://doi.org/10.7171/jbt.20-3203-001>

Thermo Fisher (2007), *PureLink Genomic DNA Mini Kit*. Available at: https://documents.thermofisher.com/TFS-Assets/LSG/manuals/purelink_genomic_mini_man.pdf. Accessed: 25/8/25

Thermo Fisher (2025a), *DH5α Competent Cells*. Available at: <https://www.thermofisher.com/uk/en/home/life-science/cloning/competent-cells-for-transformation/competent-cells-strains/dh5a-competent-cells.html>. Accessed: 17/6/25

Thermo Fisher (2025b), *Genotypes and Genetic Markers of Competent E. coli Cells*. Available at: <https://www.thermofisher.com/uk/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/molecular-cloning/transformation/competent-cell-genotypes-genetic-markers.html>. Accessed: 17/6/25

Thermo Fisher (2025c), *Proteinase K Solution (20 mg/mL), RNA grade*. Available at: <https://www.thermofisher.com/order/catalog/product/25530049>. Accessed: 25/8/25

Thermo Fisher (2025d), *One Shot BL21-AI Chemically Competent E. coli*. Available at: <https://www.thermofisher.com/order/catalog/product/C607003>. Accessed: 17/6/25

Thomas, P., Terradot, G., Danos, V. et al. (2018), 'Sources, propagation and consequences of stochasticity in cellular growth', *Nature Communications*, **9**(1), pp. 4528. DOI: <https://doi.org/10.1038/s41467-018-06912-9>

Tian, Y., Zhang, X.Y., Aslam, M.M. et al. (2025), 'Unraveling competitive protein interactions: A critical step toward improved production and therapeutics', *Journal of Advanced Research*,

S2090-1232(25)00909-9. Advance online publication. DOI:

<https://doi.org/10.1016/j.jare.2025.11.014>

Torres, R., Carrasco, B., Ayora, S. et al. (2025), 'Hallmarks of DNA replication stress responses in *Escherichia coli* and *Bacillus subtilis*', *FEMS Microbiology Reviews*, **49**, fuaf041. DOI: <https://doi.org/10.1093/femsre/fuaf041>

Trojanowski, D., Hołowka, J., and Zakrzewska-Czerwińska, J. (2018), 'Where and When Bacterial Chromosome Replication Starts: A Single Cell Perspective', *Frontiers in Microbiology*, **9**, pp. 2819. DOI: <https://doi.org/10.3389/fmicb.2018.02819>

Tsegay, P.S., Lai, Y., and Liu, Y. (2019), 'Replication Stress and Consequential Instability of the Genome and Epigenome', *Molecules (Basel, Switzerland)*, **24**(21), pp. 3870. DOI: <https://doi.org/10.3390/molecules24213870>

Tsegaye, S., Dedefo, G., and Mehdi, M. (2021), 'Biophysical applications in structural and molecular biology', *Biological Chemistry*, **402**(10), pp. 1155–1177. DOI: <https://doi.org/10.1515/hsz-2021-0232>

Tumini, E., Barroso, S., -Calero, C.P. et al. (2016), 'Roles of human POLD1 and POLD3 in genome stability', *Scientific Reports*, **6**, pp. 38873. DOI: <https://doi.org/10.1038/srep38873>

Tuteja, N., and Tuteja, R. (2004), 'Prokaryotic and eukaryotic DNA helicases. Essential molecular motor proteins for cellular machinery', *European Journal of Biochemistry*, **271**(10), pp. 1835–1848. DOI: <https://doi.org/10.1111/j.1432-1033.2004.04093.x>

Tuteja, N., and Tuteja, R. (2006), 'Helicases as molecular motors: An insight', *Physica A*, **372**(1), pp. 70–83. DOI: <https://doi.org/10.1016/j.physa.2006.05.014>

Tuttle, A.R., Trahan, N.D., and Son, M.S. (2021), 'Growth and Maintenance of *Escherichia coli* Laboratory Strains', *Current Protocols*, **1**(1), e20. DOI: <https://doi.org/10.1002/cpz1.20>

Umbricht, C.B., Erdile, L.F., Jabs, E.W. et al. (1993), 'Cloning, overexpression, and genomic mapping of the 14-kDa subunit of human replication protein A', *The Journal of Biological Chemistry*, **268**(9), pp. 6131–6138.

Umbricht, C.B., Griffin, C.A., Hawkins, A.L. et al. (1994), 'High-resolution genomic mapping of the three human replication protein A genes (RPA1, RPA2, and RPA3)', *Genomics*, **20**(2), pp. 249–257. DOI: <https://doi.org/10.1006/geno.1994.1161>

Uversky V.N. (2020), 'New technologies to analyse protein function: an intrinsic disorder perspective', *F1000Research*, **9**, F1000 Faculty Rev-101. DOI: <https://doi.org/10.12688/f1000research.20867.1>

Vafabakhsh, R., and Ha, T. (2012), 'Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization', *Science (New York, N.Y.)*, **337**(6098), pp. 1097–1101. DOI: <https://doi.org/10.1126/science.1224139>

Vare, D., Groth, P., Carlsson, R. et al. (2012), 'DNA interstrand crosslinks induce a potent replication block followed by formation and repair of double strand breaks in intact mammalian cells', *DNA Repair*, **11**(12), pp. 976–985. DOI: <https://doi.org/10.1016/j.dnarep.2012.09.010>

Varga, T., and Aplan, P.D. (2005), 'Chromosomal aberrations induced by double strand DNA breaks', *DNA Repair*, **4**(9), pp. 1038–1046. DOI: <https://doi.org/10.1016/j.dnarep.2005.05.004>

Veaute, X., Delmas, S., Selva, M. et al. (2005), 'UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in *Escherichia coli*', *The EMBO Journal*, **24**(1), pp. 180–189. DOI: <https://doi.org/10.1038/sj.emboj.7600485>

Verhoef, M.J., and Casebeer, A.L. (1997), 'Broadening horizons: Integrating quantitative and qualitative research', *The Canadian Journal of Infectious Diseases = Journal Canadien des Maladies Infectieuses*, **8**(2), pp. 65–66. DOI: <https://doi.org/10.1155/1997/349145>

Vos, S.M., Tretter, E.M., Schmidt, B.H. et al. (2011), 'All tangled up: how cells direct, manage and exploit topoisomerase function', *Nature Reviews. Molecular cell biology*, **12**(12), pp. 827–841. DOI: <https://doi.org/10.1038/nrm3228>

Walther, A.P., Gomes, X.V., Lao, Y. et al. (1999), 'Replication protein A interactions with DNA. 1. Functions of the DNA-binding and zinc-finger domains of the 70-kDa subunit', *Biochemistry*, **38**(13), pp. 3963–3973. DOI: <https://doi.org/10.1021/bi982370u>

- Wang, D., and Gribskov, M. (2005), 'Examining the architecture of cellular computing through a comparative study with a computer', *Journal of the Royal Society, Interface*, **2**(3), pp. 187–195. DOI: <https://doi.org/10.1098/rsif.2005.0038>
- Wang, J.D., and Levin, P.A. (2009), 'Metabolism, cell growth and the bacterial cell cycle', *Nature Reviews. Microbiology*, **7**(11), pp. 822–827. DOI: <https://doi.org/10.1038/nrmicro2202>
- Wang, Z. (2021), 'Regulation of Cell Cycle Progression by Growth Factor-Induced Cell Signaling', *Cells*, **10**(12), pp. 3327. DOI: <https://doi.org/10.3390/cells10123327>
- Wardell, K., Haldenby, S., Jones, N. et al. (2017), 'RadB acts in homologous recombination in the archaeon *Haloferax volcanii*, consistent with a role as recombination mediator', *DNA Repair*, **55**, pp. 7–16. DOI: <https://doi.org/10.1016/j.dnarep.2017.04.005>
- Watson, J.D., and Crick, F.H. (1953), 'The structure of DNA', *Cold Spring Harbor Symposia On Quantitative Biology*, **18**, pp. 123–131. DOI: <https://doi.org/10.1101/sqb.1953.018.01.020>
- Wee, J., and Wei, G.W. (2024), 'Benchmarking AlphaFold3's protein-protein complex accuracy and machine learning prediction reliability for binding free energy changes upon mutation', *ArXiv*. arXiv: 2406.03979v1.
- Weedon, M.N., Ellard, S., Prindle, M.J. et al. (2013), 'An in-frame deletion at the polymerase active site of POLD1 causes a multisystem disorder with lipodystrophy', *Nature Genetics*, **45**(8), pp. 947–950. DOI: <https://doi.org/10.1038/ng.2670>
- Wei, C., and Zhao, X. (2018), 'Induction of Viable but Nonculturable *Escherichia coli* O157:H7 by Low Temperature and Its Resuscitation', *Frontiers in Microbiology*, **9**, pp. 2728. DOI: <https://doi.org/10.3389/fmicb.2018.02728>
- Weinreb, C., Wolock, S., Tusi, B.K. et al. (2018), 'Fundamental limits on dynamic inference from single-cell snapshots', *Proceedings of the National Academy of Sciences of the United States of America*, **115**(10), pp. E2467–E2476. DOI: <https://doi.org/10.1073/pnas.1714723115>
- Wendel, B.M., Hernandez, A.J., Courcelle, C.T. et al. (2021), 'Ligase A and RNase HI Participate in Completing Replication on the Chromosome in *Escherichia coli*', *DNA*, **1**(1), pp. 13–25. DOI: <https://doi.org/10.3390/dna1010003>

West, S.C. (2009), 'The search for a human Holliday junction resolvase', *Biochemical Society Transactions*, **37**(Pt 3), pp. 519–526. DOI: <https://doi.org/10.1042/BST0370519>

Westermarck, J., Ivaska, J., and Corthals, G.L. (2013), 'Identification of protein interactions involved in cellular signaling', *Molecular & Cellular Proteomics: MCP*, **12**(7), pp. 1752–1763. DOI: <https://doi.org/10.1074/mcp.R113.027771>

White, M.A., Azeroglu, B., Lopez-Vernaza, M.A. et al. (2018), 'RecBCD coordinates repair of two ends at a DNA double-strand break, preventing aberrant chromosome amplification', *Nucleic Acids Research*, **46**(13), pp. 6670–6682. DOI: <https://doi.org/10.1093/nar/gky463>

White, M.F., and Allers, T. (2018), 'DNA repair in the archaea-an emerging picture', *FEMS Microbiology Reviews*, **42**(4), pp. 514–526. DOI: <https://doi.org/10.1093/femsre/fuy020>

Williams, S.L., Casas-Delucchi, C.S., Raguseo, F. et al. (2023), 'Replication-induced DNA secondary structures drive fork uncoupling and breakage', *The EMBO Journal*, **42**(22), e114334. DOI: <https://doi.org/10.15252/embi.2023114334>

Wohl, S., Gilron, Y., and Zheng, W. (2025), 'Structural and Functional Relevance of Charge-Based Transient Interactions inside Intrinsically Disordered Proteins', *ACS Physical Chemistry Au*, **5**(4), pp. 356–366. DOI: <https://doi.org/10.1021/acspchemau.5c00005>

Wold, M.S. (1997), 'Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism', *Annual Review of Biochemistry*, **66**, pp. 61–92. DOI: <https://doi.org/10.1146/annurev.biochem.66.1.61>

Wolffe, A.P. (1995), 'The cold-shock response in bacteria', *Science Progress*, **78**(Pt 4), pp. 301–310.

Woodman, I.L., Briggs, G.S., and Bolt, E.L. (2007), 'Archaeal Hel308 domain V couples DNA binding to ATP hydrolysis and positions DNA for unwinding over the helicase ratchet', *Journal of Molecular Biology*, **374**(5), pp. 1139–1144. DOI: <https://doi.org/10.1016/j.jmb.2007.10.00>

Woodman, I.L., and Bolt, E.L. (2011a), 'Winged helix domains with unknown function in Hel308 and related helicases', *Biochemical Society Transactions*, **39**(1), pp. 140–144. DOI: <https://doi.org/10.1042/BST0390140>

- Woodman, I.L., Brammer, K., and Bolt, E.L. (2011b), 'Physical interaction between archaeal DNA repair helicase Hel308 and Replication Protein A (RPA)', *DNA Repair*, **10**(3), pp. 306–313. DOI: <https://doi.org/10.1016/j.dnarep.2010.12.001>
- Wong, E.V. (2009), 'DNA', in: *Cells: Molecules and Mechanisms*. Louisville (KY): Axolotl Academic Publishing Company.
- Wozniak, K.J., and Simmons, L.A. (2022), 'Bacterial DNA excision repair pathways', *Nature Reviews. Microbiology*, **20**(8), pp. 465–477. DOI: <https://doi.org/10.1038/s41579-022-00694-0>
- Wrede, P. (2007), 'Molecular biology: self-sustaining chemistry', *Chemistry Central Journal*, **1**, pp. 25. DOI: <https://doi.org/10.1186/1752-153X-1-25>
- Wright, P.E., and Dyson, H.J. (2015), 'Intrinsically disordered proteins in cellular signalling and regulation', *Nature Reviews. Molecular Cell Biology*, **16**(1), pp. 18–29. DOI: <https://doi.org/10.1038/nrm3920>
- Wright, W.D., Shah, S.S., and Heyer, W.D. (2018), 'Homologous recombination and the repair of DNA double-strand breaks', *The Journal of Biological Chemistry*, **293**(27), pp. 10524–10535. DOI: <https://doi.org/10.1074/jbc.TM118.000372>
- Wyszomirski, K.H., Curth, U., Alves, J. et al. (2012), 'Type III restriction endonuclease EcoP15I is a heterotrimeric complex containing one Res subunit with several DNA-binding regions and ATPase activity', *Nucleic Acids Research*, **40**(8), pp. 3610–3622. DOI: <https://doi.org/10.1093/nar/gkr1239>
- Xia, Y. (2021), 'The Fate of Two Unstoppable Trains After Arriving Destination: Replisome Disassembly During DNA Replication Termination', *Frontiers in Cell and Developmental Biology*, **9**, 658003. DOI: <https://doi.org/10.3389/fcell.2021.658003>
- Xu, D., Ayyamperumal, S., Zhang, S. et al. (2025), 'The p12 Subunit Choreographs the Regulation and Functions of Two Forms of DNA Polymerase δ in Mammalian Cells', *Genes*, **16**(2), pp. 188. DOI: <https://doi.org/10.3390/genes16020188>

- Xu, H.S., Roberts, N., Singleton, F.L. et al. (1982), 'Survival and viability of nonculturable *Escherichia coli* and *Vibrio cholerae* in the estuarine and marine environment. *Microbial Ecology*, **8**(4), pp. 313–323. DOI: <https://doi.org/10.1007/BF02010671>
- Xu, N., You, Y., Liu, C. et al. (2020), 'Structural basis of DNA replication origin recognition by human Orc6 protein binding with DNA', *Nucleic Acids Research*, **48**(19), pp. 11146–11161. DOI: <https://doi.org/10.1093/nar/gkaa751>
- Xue, X., Raynard, S., Busygina, V. et al. (2013), 'Role of replication protein A in double holliday junction dissolution mediated by the BLM-Topo III α -RMI1-RMI2 protein complex', *The Journal of Biological Chemistry*, **288**(20), pp. 14221–14227. DOI: <https://doi.org/10.1074/jbc.M113.465609>
- Yang, C.L., Chang, L.S., Zhang, P. et al. (1992), 'Molecular cloning of the cDNA for the catalytic subunit of human DNA polymerase delta', *Nucleic Acids Research*, **20**(4), pp. 735–745. DOI: <https://doi.org/10.1093/nar/20.4.735>
- Yang, D., Wang, W., Zhao, L. et al. (2024), 'Resuscitation of viable but nonculturable bacteria promoted by ATP-mediated NAD⁺ synthesis', *Journal of Advanced Research*, **60**, pp. 27–39. DOI: <https://doi.org/10.1016/j.jare.2023.08.002>
- Yang, R., Hunker, O., Wise, M. et al. (2024), 'Multiple mechanisms for licensing human replication origins', *Nature*, **636**(8042), pp. 488–498. DOI: <https://doi.org/10.1038/s41586-024-08237-8>
- Yang, Y., Liu, Z., Wang, F. et al. (2015), 'FANCD2 and REV1 cooperate in the protection of nascent DNA strands in response to replication stress', *Nucleic Acids Research*, **43**(17), pp. 8325–8339. DOI: <https://doi.org/10.1093/nar/gkv737>
- Yao, N.Y., and O'Donnell, M. (2010), 'SnapShot: The replisome', *Cell*, **141**(6), pp. 1088–1088.e1. DOI: <https://doi.org/10.1016/j.cell.2010.05.042>
- Yao, N.Y., and O'Donnell, M. (2012), 'The RFC clamp loader: structure and function', *Sub-Cellular Biochemistry*, **62**, pp. 259–279. DOI: https://doi.org/10.1007/978-94-007-4572-8_14

- Yarranton, G.T., and Gefter, M.L. (1979), 'Enzyme-catalyzed DNA unwinding: studies on *Escherichia coli* rep protein', *Proceedings of the National Academy of Sciences of the United States of America*, **76**(4), pp. 1658–1662. DOI: <https://doi.org/10.1073/pnas.76.4.1658>
- Yoshida, K., and Fujita, M. (2021), 'DNA damage responses that enhance resilience to replication stress', *Cellular and Molecular Life Sciences: CMLS*, **78**(21-22), pp. 6763–6773. DOI: <https://doi.org/10.1007/s00018-021-03926-3>
- Yuan, Z., Georgescu, R., Schauer, G.D. et al. (2020), 'Structure of the polymerase ϵ holoenzyme and atomic model of the leading strand replisome', *Nature Communications*, **11**(1), pp. 3156. DOI: <https://doi.org/10.1038/s41467-020-16910-5>
- Yuan, Z., Georgescu, R., Li, H. et al. (2023), 'Molecular choreography of primer synthesis by the eukaryotic Pol α -primase', *Nature Communications*, **14**(1), pp. 3697. DOI: <https://doi.org/10.1038/s41467-023-39441-1>
- Yuan, Z., Georgescu, R., Yao, N.Y. et al. (2024), 'Mechanism of PCNA loading by Ctf18-RFC for leading-strand DNA synthesis', *Science (New York, N.Y.)*, **385**(6708), pp. eadk5901. DOI: <https://doi.org/10.1126/science.adk5901>
- Zagrovic, B., Adlhart, M., and Kapral, T.H. (2023), 'Coding From Binding? Molecular Interactions at the Heart of Translation', *Annual Review of Biophysics*, **52**, pp. 69–89. DOI: <https://doi.org/10.1146/annurev-biophys-090622-102329>
- Zambrano, M.M., Siegele, D.A., Almirón, M. et al. (1993), 'Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures', *Science (New York, N.Y.)*, **259**(5102), pp. 1757–1760. DOI: <https://doi.org/10.1126/science.7681219>
- Zawilak-Pawlik, A., Nowaczyk, M., and Zakrzewska-Czerwińska, J. (2017), 'The Role of the N-Terminal Domains of Bacterial Initiator DnaA in the Assembly and Regulation of the Bacterial Replication Initiation Complex', *Genes*, **8**(5), pp. 136. DOI: <https://doi.org/10.3390/genes8050136>
- Zechiedrich, E.L., and Osheroff, N. (1990), 'Eukaryotic topoisomerases recognize nucleic acid topology by preferentially interacting with DNA crossovers', *The EMBO Journal*, **9**(13), pp. 4555–4562. DOI: <https://doi.org/10.1002/j.1460-2075.1990.tb07908.x>

- Zegerman, P. (2013), 'DNA replication: polymerase epsilon as a non-catalytic converter of the helicase', *Current Biology: CB*, **23**(7), pp. R273–R276. DOI: <https://doi.org/10.1016/j.cub.2013.03.008>
- Zeman, M.K., and Cimprich, K.A. (2014), 'Causes and consequences of replication stress', *Nature Cell Biology*, **16**(1), pp. 2–9. <https://doi.org/10.1038/ncb2897>
- Zhang, J., Tan, C.K., McMullen, B. et al. (1995), 'Cloning of the cDNAs for the small subunits of bovine and human DNA polymerase delta and chromosomal location of the human gene (POLD2)', *Genomics*, **29**(1), pp. 179–186. DOI: <https://doi.org/10.1006/geno.1995.1229>
- Zhang, J., and Landick, R. (2016), 'A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure', *Trends in Biochemical Sciences*, **41**(4), pp. 293–310. DOI: <https://doi.org/10.1016/j.tibs.2015.12.009>
- Zhang, W., Petri, K., Ma, J. et al. (2024), 'Enhancing CRISPR prime editing by reducing misfolded pegRNA interactions', *eLife*, **12**, RP90948. DOI: <https://doi.org/10.7554/eLife.90948>
- Zhang, Y.M., Li, B., and Wu, W.Q. (2024), 'Single-molecule insights into repetitive helicases', *The Journal of Biological Chemistry*, **300**(11), pp. 107894. DOI: <https://doi.org/10.1016/j.jbc.2024.107894>
- Zhang, Z., Zabaikina, I., Nieto, C. et al. (2025), 'Stochastic gene expression in proliferating cells: Differing noise intensity in single-cell and population perspectives', *PLOS Computational Biology*, **21**(6), e1013014. DOI: <https://doi.org/10.1371/journal.pcbi.1013014>
- Zhao, J., Bacolla, A., Wang, G. et al. (2010), 'Non-B DNA structure-induced genetic instability and evolution', *Cellular and molecular life sciences: CMLS*, **67**(1), pp. 43–62. DOI: <https://doi.org/10.1007/s00018-009-0131-2>
- Zhao, X., Wei, C., Li, J. et al. (2017), 'Cell cycle-dependent control of homologous recombination', *Acta Biochimica et Biophysica Sinica*, **49**(8), pp. 655–668. DOI: <https://doi.org/10.1093/abbs/gmx055>
- Zhao, Y., Hou, K., Li, Y. et al. (2023), 'Human HELQ regulates DNA end resection at DNA double-strand breaks and stalled replication forks', *Nucleic Acids Research*, **51**(22), pp. 12207–12223. DOI: <https://doi.org/10.1093/nar/gkad940>

Zhao, Y., Hou, K., Liu, Y. et al. (2024), 'Helicase HELQ: Molecular Characters Fit for DSB Repair Function', *International Journal of Molecular Sciences*, **25**(16), pp. 8634. DOI: <https://doi.org/10.3390/ijms25168634>

Zheng, F., Georgescu, R.E., Li, H. et al. (2020), 'Structure of eukaryotic DNA polymerase δ bound to the PCNA clamp while encircling DNA', *Proceedings of the National Academy of Sciences of the United States of America*, **117**(48), pp. 30344–30353. DOI: <https://doi.org/10.1073/pnas.2017637117>

Zhong, Q., Xiao, X., Qiu, Y. et al. (2023), 'Protein posttranslational modifications in health and diseases: Functions, regulatory mechanisms, and therapeutic implications', *MedComm*, **4**(3), e261. DOI: <https://doi.org/10.1002/mco2.261>

Zhou, H.X., and Pang, X. (2018), 'Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation', *Chemical Reviews*, **118**(4), pp. 1691–1741. DOI: <https://doi.org/10.1021/acs.chemrev.7b00305>

Zhou, Y., Meng, X., Zhang, S. et al. (2012), 'Characterization of human DNA polymerase delta and its subassemblies reconstituted by expression in the MultiBac system', *PLOS One*, **7**(6), e39156. DOI: <https://doi.org/10.1371/journal.pone.0039156>

Zymo Research (2023), *ZymoPURE II Plasmid Maxiprep Kit*. Available at: https://files.zymoresearch.com/protocols/d4202_d4203_zymopure_ii_plasmid_maxiprep.pdf. Accessed: 18/9/25

Chapter 7: Appendix

Table 7.1 Statistics and details from two-tailed Welch's t-tests of POLD3 Quincunx region mutants BiFC assay FLUOstar Omega fluorescence scans.

Statistic	NmVenus + CmVenus	NmVenus-POLD3 + CmVenus N-HELQ	NmVenus-POLD3 ^{QC} + CmVenus N-HELQ	NmVenus-POLD3 ^{KS} RRR ₊ CmVenus N-HELQ	NmVenus-POLD3 ^K KSRR ₊ CmVenus N-HELQ	NmVenus-POLD3 ^K KDRR ₊ CmVenus N-HELQ	NmVenus-POLD3 ^A KRRA ₊ CmVenus N-HELQ
a.							
Fluorescence Normalised Using OD600							
n	3	3	3	3	3	3	3
M	33441	90688	90576	90485	72993	70252	108749
SD	2944	2464	10813	7094	5616	7475	8665
SEM	1700	1423	6243	4095	3242	4316	5003
Compared to NmVenus + CmVenus^b.							
MD ± SEM	N/A	57247 ± 217	57135 ± 6470	57044 ± 4434	39552 ± 3661	36811 ± 4639	75307 ± 5284
t	N/A	25.83	8.831	12.86	10.80	7.936	14.25
df	N/A	3.880	2.295	2.669	3.022	2.606	2.456
p (t-test)	N/A	< 0.0001	0.0081	0.0018	0.0016	0.0068	0.0020
t-test significance ^c	N/A	****	**	**	**	**	**
+/- ^d	N/A	+	+	+	+	+	+

95% CI	N/A	51016 to 63477	32459 to 81812	41889 to 72198	27949 to 51155	20702 to 52920	56174 to 94441
r^2	N/A	0.9942	0.9714	0.9841	0.9748	0.9603	0.9881
DFn	N/A	2	2	2	2	2	2
DFd	N/A	2	2	2	2	2	2
F	N/A	1.428	13.49	5.805	3.638	6.446	8.662
p (F-test)	N/A	0.8238	0.1380	0.2939	0.4312	0.2686	0.2070
F-test significance ^c	N/A	N/S	N/S	N/S	N/S	N/S	N/S
Compared to NmVenus-POLD3 + CmVenus-N-HELQ^b							
MD ± SEM	-57247 ± 2217	N/A	-111.6 ± 6403	-203.1 ± 4336	-17695 ± 3541	-20436 ± 4544	18061 ± 5201
t	25.83	N/A	0.01743	0.04684	4.997	4.497	3.472
df	3.880	N/A	2.207	2.476	2.743	2.430	2.321
p (t-test)	< 0.0001	N/A	0.9875	0.9662	0.0190	0.0317	0.0595
t-test significance ^c	****	N/A	N/S	N/S	*	*	N/S
+/- d.	-	N/A	N/A	N/S	-	-	N/A
95% CI	-63477 to -51016	N/A	-25326 to 25103	-15810 to 15404	-29586 to -5804	-37022 to -3849	-1597 to 37718
r^2	0.9942	N/A	0.0001376	0.0008855	0.9011	0.8927	0.8386
DFn	2	N/A	2	2	2	2	2
DFd	2	N/A	2	2	2	2	2

F	1.428	N/A	19.26	8.287	5.194	9.203	12.37
p (F-test)	0.8238	N/A	0.0987	0.2153	0.3229	0.1960	0.1496
F-test significance ^{c.}	N/S	N/A	N/S	N/S	N/S	N/S	N/S
Compared to NmVenus-POLD3^{QC} + CmVenus-N-HELQ^{b.}							
MD ± SEM	- 57135 ± 6470	111.6 ± 6403	N/A	-91.50 ± 7466	-17583 ± 7035	-20324 ± 7590	18172 ± 8000
t	8.831	0.01743	N/A	0.01225	2.499	2.678	2.271
df	2.295	2.207	N/A	3.452	3.006	3.556	3.819
p (t-test)	0.0081	0.9875	N/A	0.9909	0.0876	0.0627	0.0888
t-test significance ^{c.}	**	N/S	N/A	N/S	N/S	N/S	N/S
+/- d.	-	N/A	N/A	N/A	N/A	N/A	N/A
95% CI	-81812 to 32459	-25103 to 25326	N/A	-22185 to 22002	-39946 to 4780	-42475 to 1826	-4462 to 40806
r ²	0.9714	0.0001376	N/A	4.350x10 ⁻⁰⁰⁵	0.6752	0.6685	0.5747
DFn	2	2	N/A	2	2	2	2
DFd	2	2	N/A	2	2	2	2
F	13.49	19.26	N/A	2.324	3.707	2.092	1.557
p (F-test)	0.1380	0.0987	N/A	0.6018	0.4249	0.6468	0.7821

F-test significance ^c	N/S	N/S	N/A	N/S	N/S	N/S	N/S
Fluorescence Normalised Using CFU/mL							
n	3	3	3	3	3	3	3
M	0.00107 1	7.447	4.236	0.1501	11.64	0.01210	0.00423 8
SD	0.00185 0	12.82	7.283	0.2309	20.11	0.01750	0.00088 15
SEM	0.00106 8	7.403	4.205	0.1333	11.61	0.01010	0.00050 89
Comparison to NmVenus + CmVenus^b							
MD ± SEM	N/A	7.446 ± 7.403	4.235 ± 4.205	0.1491 ± 0.1333	11.63 ± 11.61	0.01103 ± 0.01016	0.00316 7 ± 0.00118 3
t	N/A	1.006	1.007	1.118	1.002,	1.086,	2.677
df	N/A	2.000	2.000	2.000	2.000	2.045	2.864
p (t-test)	N/A	0.4204	0.4200	0.3798	0.4218	0.3889	0.0791
t-test significance ^c	N/A	N/S	N/S	N/S	N/S	N/S	N/S
+/- ^d	N/A	N/A	N/A	NA	N/A	NA	N/A
95% CI	N/A	-24.41 to 39.30	-13.86 to 22.33	-0.4244 to 0.7225	-38.31 to 61.58	- 0.03177 to 0.05383	- 13 to 0.00703 5
r ²	N/A	0.3359	0.3365	0.3846	0.3343	0.3657	0.7145

DFn	N/A	2	2	2	2	2	2
DFd	N/A	2	2	2	2	2	2
F	N/A	48050143	15502912	15578	118149 300	89.47	4.403
p (F-test)	N/A	< 0.0001	< 0.0001	0.0001	< 0.0001	0.0221	0.3701
F-test significance ^c .	N/A	****	****	***	****	*	N/S
Comparison to NmVenus-POLD3 + CmVenus-N-HELQ^b.							
MD ± SEM	-7.446 ± 7.403	N/A	-3.212 ± 8.514	-7.297 ± 7.404	4.188 ± 13.77	-7.435 ± 7.403	-7.443 ± 7.403
t	1.006, 2.000	N/A	0.3773	0.9856,	0.3042,	1.004	1.005
df	2.000	N/A	3.169	2.001	3.396	2.000	2.000
p (t-test)	0.4204	N/A	0.7298	0.4282	0.7787	0.4210	0.4206
t-test significance ^c .	N/S	N/A	N/S	N/S	N/S	N/S	N/S
+/- ^d .	N/A	N/A	N/A	N/A	N/A	N/A	N/A
95% CI	-39.30 to 24.41	N/A	-29.51 to 23.08	-39.13 to 24.54	-36.88 to 45.25	-39.29 to 24.42	-39.30 to 24.41
r ²	0.3359	N/A	0.04298	0.3268	0.02653	0.3353	0.3358
DFn	2	N/A	2	2	2	2	2
DFd	2	N/A	2	2	2	2	2
F	48050143	N/A	3.099	3084	2.459	537054	211584579
p (F-test)	< 0.0001	N/A	0.4879	0.0006	0.5782	< 0.0001	< 0.0001

F-test significance ^c	****	N/A	N/S	***	N/S	****	****
Comparison to NmVenus-POLD3^{QC} + CmVenus-N-HELQ^b							
MD ±	-4.235 ±	3.212 ±	N/A	-4.085 ±	7.400 ±	-4.224 ±	-4.231 ±
SEM	4.205	8.514		4.207	12.35	4.205	4.205
t	1.007	0.3773	N/A	0.9711	0.5994	1.004	1.006
df	2.000	3.169	N/A	2.004	2.516	2.000	2.000
p (t-test)	0.4200	0.7298	N/A	0.4338	0.5985	0.4210	0.4202
T-test significance ^c	N/S	N/S	N/A	N/S	N/S	N/S	N/S
+/- ^d	N/A	N/A	N/A	N/A	N/A	N/A	N/A
95% CI	-22.33 to 13.86	-23.08 to 29.51	N/A	-22.15 to 13.98	-36.53 to 51.33	-22.32 to 13.87	-22.32 to 13.86
r ²	0.3365	0.04298	N/A	0.3200	0.1249	0.3353	0.3361
DFn	2	2	N/A	2	2	2	2
DFd	2	2	N/A	2	2	2	2
F	1550291 2	3.099	N/A	995.2	7.621	173275	6826570 9
p (F- test)	< 0.0001	0.4879	N/A	0.0020	0.2320	< 0.0001	< 0.0001
F-test significance ^c	****	N/S	N/A	**	N/S	****	****
Fluorescence Normalised Using CFU/mL With Outliers Removed							
n	3	2	2	2	2	2	3
M	0.00107 1	0.04454	0.03066	0.01703	0.02717	0.00200 3	0.00423 8

SD	0.001850	0.02064	0.03591	0.01796	0.03197	0.0006703	0.0008815
SEM	0.001068	0.01459	0.02539	0.01270	0.02261	0.0004740	0.0005089
Comparison to NmVenus + CmVenus^b.							
MD ± SEM	N/A	0.04347 ± 0.01463	0.02959 ± 0.02541	0.01596 ± 0.01275	0.02610 ± 0.02263	0.00093 ± 0.001168	0.00316 ± 0.001183
t	N/A	2.971	1.164	1.252	1.153	0.7976	2.677
df	N/A	1.011	1.004	1.014	1.004	2.659	2.864
p (t-test)	N/A	0.2046	0.4513	0.4269	0.4541	0.4901	0.3701
t-test significance ^c .	N/A	N/S	N/S	N/S	N/S	N/S	N/S
+/- d.	N/A	N/A	N/A	N/A	N/A	N/A	N/A
95% CI	N/A	-0.1378 to 0.2248	-0.2907 to 0.3498	-0.1407 to 0.1727	-0.2584 to 0.3106	-0.003071 to 0.004935	-0.0007013 to 0.007035
r ²	N/A	0.8973	0.5746	0.6072	0.5697	0.1930	0.7145
DFn	N/A	1	1	1	1	2	2
DFd	N/A	2	2	2	2	1	2
F	N/A	124.5	376.9	94.29	298.7	7.615	4.403
p (F-test)	N/A	0.0159	0.0053	0.0209	0.0067	0.4965	0.3701

F-test significance ^{c.}	N/A	*	**	*	**	N/S	N/S
Comparison to NmVenus-POLD3 + CmVenus-N-HELQ^{b.}							
MD ± SEM ± 0.01463	-0.04347 ± 0.01463	N/A	-0.01388 ± 0.02929	-0.02751 ± 0.01935	- 0.01737 ± 0.02691	- 0.04254 ± 0.01460	- 0.04030 ± 0.01460
t	2.971	N/A	0.4740	1.422	0.6456	2.914	2.760
df	1.011	N/A	1.596	1.963	1.710	1.002	1.002
p (t-test)	0.2046	N/A	0.6923	0.2931	0.5943	0.2101	0.2208
t-test significance ^{c.}	N/S	N/A	N/S	N/S	N/S	N/S	N/S
+/- ^{d.}	N/A	N/A	N/A	N/A	N/A	N/A	N/A
95% CI to 0.1378	-0.2248 to 0.1378	N/A	-0.1760 to 0.1483	-0.1123 to 0.05726	-0.1541 to 0.1194	-0.2271 to 0.1420	-0.2248 to 0.1442
r ²	0.8973	N/A	0.1235	0.5075	0.1960	0.8944	0.8837
DFn	1	N/A	1	1	1	1	1
DFd	2	N/A	1	1	1	1	2
F	124.5	N/A	3.028	1.320	2.400	947.7	548.0
p (F-test)	0.0159	N/A	0.6641	0.9119	0.7298	0.0413	0.0036
F-test significance ^{c.}	*	N/A	N/S	N/S	N/S	*	**
Comparison to NmVenus-POLD3^{QC} + CmVenus-N-HELQ^{b.}							

MD $\pm$ SEM	-0.02959 $\pm$ 0.02541	0.01388 $\pm$ 0.02929	N/A	-0.01363 $\pm$ 0.02839	- 0.00348 8 $\pm$ 0.03400	- 0.02866 $\pm$ 0.02540	- 0.02642 $\pm$ 0.02540
t	1.164	0.4740	N/A	0.4800	0.1026	1.128	1.040
df	1.004	1.596	N/A	1.471	1.974	1.001	1.001
p (t-test)	0.4513	0.6641	N/A	0.6929	0.9278	0.4616	0.4873
t-test significance ^c	N/S	N/S	N/A	N/S	N/S	N/S	N/S
+/- ^d .	N/A	N/A	N/A	N/A	N/A	N/A	N/A
95% CI	-0.3498 to 0.2907	-0.1483 to 0.1760	N/A	-0.1893 to 0.1620	-0.1517 to 0.1447	-0.3508 to 0.2935	-0.3485 to 0.2957
r ²	0.5746	0.1235	N/A	0.1354	0.00530 6	0.5599	0.5195
DFn	1	1	N/A	1	1	1	1
DFd	2	1	N/A	1	1	1	2
F	376.9	3.028	N/A	3.997	1.262	2870	1660
p (F- test)	0.0053	0.6641	N/A	0.5905	0.9262	0.0238	0.0012
F-test significance ^c	**	N/S	N/A	N/S	N/S	*	**

Two-tailed Welch's t-tests were performed using GraphPad Prism 10 to determine if fluorescence triplicate means showed a significant different compared to NmVenus + CmVenus, NmVenus-POLD3 + CmVenus-N-HELQ, and NmVenus-POLD3^{QC} + CmVenus-N-HELQ fluorescence triplicate means. a. Statistics: n = sample size, M = mean, SD = standard deviation, SEM = standard error of mean, MD = mean difference, t = t-value, df = degrees of

freedom, p = p-value, 95% CI = 95% confidence interval, r^2 = coefficient of determination, DFn = degrees of freedom numerator, DFd = degrees of freedom denominator, F = F-value. b. Comparison is performed as mean in the column minus mean of the comparison protein pair. c. Significance level is denoted by asterisks over the column: * = p < 0.05, ** = p < 0.01, *** = p < 0.001, **** = p < 0.0001, N/S = 'no significance', and N/A = 'not applicable'. d. + indicates that that column protein pair has a higher mean than the comparative protein pair; - indicates it is lower.

Table 7.2 Statistics and details from two-tailed Welch's t-tests of POLD3 IDPRs and putative NLS BiFC assay FLUOstar Omega fluorescence scans.

Statistic ^a .	NmVen us + CmVen us	NmVen us- POLD3 + CmVen us N- HELQ	NmVenu s- POLD3 ^{ID} PR1 + CmVenus N-HELQ	NmVenus- POLD3 ^{IDP} R2 ₊ CmVenus N-HELQ	NmVenu s- POLD3 ^{ID} PR3 + CmVenus N-HELQ	NmVenu s- POLD3 ^{ID} PR4 + CmVenus N-HELQ	NmVenu s- POLD3 ^P NLS + CmVenu s N- HELQ
Fluorescence Normalised Using OD600							
n	3	3	3	3	3	3	3
M	29969	45688	60450	94799	137137	84541	42106
SD	9449	15566	2144	7013	19670	7416	9438
SDM	5455	8987	1238	4049	11357	4281	5449
Comparison to NmVenus + CmVenus^b.							
MD ± SEM	N/A	15719 ± 10513	30481 ± 5594	64830 ± 6794	107168 ± 12599	54572 ± 6935	12137 ± 7710
t	N/A	1.495	5.449	9.543	8.506	7.869	1.574
df	N/A	3.298	2.205	3.691	2.876	3.786	4.000
p (t-test)	N/A	0.2237	0.0258	0.0010	0.0040	0.0018	0.1906

t-test significance ^c .	N/A	N/S	*	***	**	**	N/S
+/- ^d .	N/A	N/A	+	+	+	+	N/A
95% CI	N/A	-16092 to 47531	8438 to 52525	45327 to 84332	66079 to 148257	34882 to 74262	-9270 to 33545
r ²	N/A	0.4040	0.9309	0.9611	0.9618	0.9424	0.3825
DFn	N/A	2	2	2	2	2	2
DFd	N/A	2	2	2	2	2	2
F	N/A	2.714	19.43	1.815	4.334	1.623	1.002
p (F-test)	N/A	0.5385	0.0979	0.7104	0.3750	0.7624	0.9989
F-test significance ^c .	N/A	N/S	N/S	N/S	N/S	N/S	N/S
Comparison to NmVenus-POLD3 + CmVenus-N-HELQ^b.							
MD ± SEM	-15719 ± 10513	N/A	14762 ± 9072	49110 ± 9857	91449 ± 14482	38852 ± 9955	-3582 ± 10510
t	1.495	N/A	1.627	4.982	6.315	3.903	0.3409
df	3.298	N/A	2.076	2.780	3.799	2.863	3.295
p (t-test)	0.2237	N/A	0.2407	0.0186	0.0038	0.0325	0.7538
t-test significance ^c .	N/S	N/A	N/S	*	**	*	N/S
+/- ^d .	N/A	N/A	N/A	+	+	+	N/A
95% CI	-47531 to 16092	N/A	-22935 to 52459	16288 to 81933	50388 to 132510	6300 to 71405	-35395 to 28230

r ²	0.4040	N/A	0.5606	0.8993	0.9130	0.8418	0.03406
DFn	2	N/A	2	2	2	2	2
DFd	2	N/A	2	2	2	2	2
F	2.714	N/A	52.73	4.926	1.597	4.406	2.720
p (F-test)	0.5385	N/A	0.0372	0.3375	0.7701	0.3700	0.5376
F-test significance ^c	N/S	N/A	*	N/S	N/S	N/S	N/S

Two-tailed Welch's t-tests were performed using GraphPad Prism 10 to determine if fluorescence triplicate means showed a significant difference compared to NmVenus + CmVenus and NmVenus-POLD3 + CmVenus-N-HELQ fluorescence triplicate means. a. Statistics: n = sample size, M = mean, SD = standard deviation, SEM = standard error of mean, MD = mean difference, t = t-value, df = degrees of freedom, p = p-value, 95% CI = 95% confidence interval, r² = coefficient of determination, DFn = degrees of freedom numerator, DFd = degrees of freedom denominator, F = F-value. b. Comparison is performed as mean in the column minus mean of the comparison protein pair. c. Significance level is denoted by asterisks over the column: * = p < 0.05, ** = p < 0.01, *** = p < 0.001, **** = p < 0.0001, N/S = 'no significance', and N/A = 'not applicable'. d. + indicates that that column protein pair has a higher mean than the comparative protein pair; - indicates it is lower.

Table 7.3 Statistics and details from two-tailed Welch's t-tests of BL21-AI *E. coli* growth media FLUOstar Omega autofluorescence scans.

Statistic ^a	500/520 nm				590-50/670 nm			
	2X	M9	Supplemente	Mu	2X	M9	Supplemente	Mu
	Minimal		d 2X M9	Broth	Minimal		d 2X M9	Broth
	Medium		Medium		Medium		Medium	
Comparison to Mu Broth Within Filter pairs – Gain 1473^b								
n	6		6	6	6		6	6
M	199832		211206	232429	1280		1378	1474
SD	3167		4866	6043	37.17		100.0	28.58

SEM	1293	1986	2467	15.17	40.83	11.67
MD ± SEM	-32597 ± 2785	-21223 ± 3167	N/A	-194.0 ± 19.14	-95.67 ± 42.46	N/A
t	11.70	6.701	N/A	10.14	2.253	N/A
df	7.554	9.565	N/A	9.381	5.811	N/A
p	0.00000 4	0.000066	N/A	0.000002	0.066600	N/A
t-test significance ^c .	****	****	N/A	****	N/S	N/A
+/- d.	-	-	N/A	-	N/A	N/A
q	0.00000 4	0.000067	N/A	0.000004	0.033633	N/A
Comparison to Mu Broth Within Filter Pairs – Gain 3025^b.						
n	6	6	6	6	6	6
M	260000	260000	260000	207758	226543	242725
SD	0.000	0.000	0.000	3507	15570	4290
SEM	0.000	0.000	0.000	1432	6357	1751
MD ± SEM	0.000 ± 0.000	0.000 ± 0.000	N/A	-34967 ± 2262	-16182 ± 6593	N/A
t	N/A	N/A	N/A	15.46	2.454	N/A
df	N/A	N/A	N/A	9.620	5.755	N/A
p	N/A	N/A	N/A	<0.00000 1	0.051225	N/A
t-test significance ^c .	N/A	N/A	N/A	****	N/S	N/A
+/- d.	N/A	N/A	N/A	-	N/A	N/A

q	N/A	N/A	N/A	<0.00000 1	0.051737	N/A
Comparison of Respective Media Between Filter Pairs - Gain 1473^e.						
n	6	6	6	6	6	6
M	199832	211206	232429	1280	1378	1474
SD	3167	4866	6043	37.17	100.0	28.58
SEM	1293	1986	2467	15.17	40.83	11.67
MD ± SEM	198552 ± 1293	209828 ± 1987	230955 ± 2467	-198552 ± 1293	-209828 ± 1987	- 230955 ± 2467
t	153.6	105.6	93.62	153.6	105.6	93.62
df	5.001	5.004	5.000	5.001	5.004	5.000
p (t-test)	<0.0001	<0.0001	<0.000 1	<0.0001	<0.0001	<0.000 1
t-test significance ^c	****	****	****	****	****	****
+/- d.	+	+	+	-	-	-
95% CI	195228 to 201875	204721 to 214934	224614 to 237296	-201875 to 195228	-214934 to - 204721	- 237296 to 224614
r ²	0.9998	0.9996	0.9994	0.9998	0.9996	0.9994
DFn	5	5	5	5	5	5
DFd	5	5	5	5	5	5
F	7259	2367	44703	7259	2367	44703
p (F-test)	<0.0001	<0.0001	<0.000 1	<0.0001	<0.0001	<0.000 1

F-test significance ^c	****	****	****	****	****	****
Comparison of Respective Media Between Filter Pairs - Gain 3025^e.						
n	6	6	6	6	6	6
M	260000	260000	260000	207758	226543	242725
SD	0.000	0.000	0.000	3507	15570	4290
SEM	0.000	0.000	0.000	1432	6357	1751
MD ± SEM	52242 ± 1432	33457 ± 6357	17275 ± 1751	-52242 ± 1432	-33457 ± 6357	-17275 ± 1751
t	36.49	5.263	9.865	36.49	5.263	9.865
df	5.000	5.000	5.000	5.000	5.000	5.000
p (t-test)	<0.0001	0.0033	0.0002	<0.0001	0.0033	0.0002
t-test significance ^c	****	**	***	****	**	***
+/- d.	+	+	+	-	-	-
95% CI	48562 to 55922	17117 to 49797	12774 to 21777	-55922 to -48562	17117 to 49797	-21777 to -12774
r ²	0.9963	0.8471	0.9511	0.9963	0.8471	0.9511
DFn	5	5	5	5	5	5
DFd	5	5	5	5	5	5
F	Infinity	Infinity	Infinity	Infinity	Infinity	Infinity
p (F-test)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

F-test significance ^c	****	****	****	****	****	****
-------------------------------------	------	------	------	------	------	------

Two-tailed Welch's t-tests were performed using GraphPad Prism 10 to determine if media autofluorescence means showed a significant difference to each other. a. Statistics: n = sample size, M = mean, SD = standard deviation, SEM = standard error of mean, MD = mean difference, t = t-value, df = degrees of freedom, p = p-value, q = q-value, 95% CI = 95% confidence interval, r^2 = coefficient of determination, DFn = degrees of freedom numerator, DFd = degrees of freedom denominator, F = F-value. b. Comparison is performed as mean in the column minus mean of Mu broth within the same filter pair and gain. c. Significance level is denoted by asterisks over the column: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$, N/S = 'no significance', and N/A = 'not applicable'. d. + indicates that that column protein pair has a higher mean than the comparative protein pair; - indicates it is lower. E. Comparison is performed as mean in the column minus mean of the same medium with the other filter pair but same gain.

Table 7.4 Statistics and details from two-tailed Welch's t-tests of RPA1 and RPA2 BiFC assay FLUOstar Omega fluorescence scans.

Statistic ^a	NmVenus CmVenus	+ NmVenus-POLD3 + CmVenus N- HELQ	NmVenus- RPA1 CmVenus N- HELQ	+ NmVenus- RPA2 CmVenus N- HELQ
Fluorescence Normalised Using OD600				
n	3	3	3	3
M	25480	30866	72927	77194
SD	3489	7980	4866	5389
SEM	2014	4607	2810	3111
Compared to NmVenus + CmVenus^b				
MD ± SEM	N/A	5386 ± 5028	47447 ± 3457	51714 ± 3706

t	N/A	1.071	13.72	13.95
df	N/A	2.738	3.626	3.426
p (t-test)	N/A	0.3694	0.0003	0.0004
t-test significance ^{c.}	N/A	N/S	***	***
+/- d.	N/A	N/A	+	+
95% CI	N/A	-11520 to 22292	37446 to 57448	40706 to 62722
r ²	N/A	0.2953	0.9811	0.9827
DFn	N/A	2	2	2
DFd	N/A	2	2	2
F	N/A	5.232	1.946	2.386
p (F-test)	N/A	0.3209	0.6790	0.5907
F-test significance ^{c.}	N/A	N/S	N/S	N/S
Compared to NmVenus-POLD3 + CmVenus-N-HELQ^{b.}				
MD ± SEM	-5386 ± 5028	N/A	42062 ± 5396	46329 ± 5559
t	1.071	N/A	7.794	8.333
df	2.738	N/A	3.307	3.510
p (t-test)	0.3694	N/A	0.0031	0.0020
t-test significance ^{c.}	N/S	N/A	**	**
+/- d.	N/A	N/A	+	+
95% CI	-22292 to 11520	N/A	25755 to 58368	30005 to 62652
r ²	0.2953	N/A	0.9484	0.9519
DFn	2	N/A	2	2
DFd	2	N/A	2	2

F	5.232	N/A	2.689	2.193
p (F-test)	0.3209	N/A	0.5421	0.6264
F-test significance ^{c.}	N/S	N/A	N/S	N/S
Compared to NmVenus-RPA1 + CmVenus-N-HELQ^{b.}				
MD ± SEM	-47447 ± 3457	-42062 ± 5396	N/A	4267 ± 4192
t	13.72	7.794	N/A	1.018
df	3.626	3.307	N/A	3.959
p (t-test)	0.0003	0.0031	N/A	0.3669
t-test significance ^{c.}	***	**	N/A	N/S
+/- ^{d.}	—	-	N/A	N/A
95% CI	-57448 to -37446	-58368 to -25755	N/A	-7420 to 15954
r ²	0.9811	0.9484	N/A	0.2074
DFn	2	2	N/A	2
DFd	2	2	N/A	2
F	1.946	2.689	N/A	1.226
p (F-test)	0.6790	0.5421	N/A	0.8983
F-test significance ^{c.}	N/S	N/S	N/A	N/S
Fluorescence Normalised Using CFU/mL				
n	3	3	3	3
M	1.177x10 ⁻⁰⁰⁸	0.01471	0.0003897	7.124x10 ⁻⁰⁰⁷
SD	4.661x10 ⁻⁰⁰⁹	0.02524	0.0001410	1.180x10 ⁻⁰⁰⁶
SEM	2.691x10 ⁻⁰⁰⁹	0.01457	8.141x10 ⁻⁰⁰⁵	6.815x10 ⁻⁰⁰⁷
Compared to NmVenus + CmVenus^{b.}				

MD ± SEM	N/A	0.01471 ± 0.01457	0.0003897 ± 8.141x10 ⁻⁰⁰⁵	7.006x10 ⁻⁰⁰⁷ ± 6.815x10 ⁻⁰⁰⁷
t	N/A	1.009	4.786	1.028
df	N/A	2.000	2.000	2.000
p (t-test)	N/A	0.4191	0.0410	0.4120
t-test significance ^{c.}	N/A	N/S	*	N/S
+/- d.	N/A	N/A	+	N/A
95% CI	N/A	-0.04799 to 0.07740	3.939x10 ⁻⁰⁰⁵ to 0.0007399	-2.231x10 ⁻⁰⁰⁶ to 3.633x10 ⁻⁰⁰⁶
r ²	N/A	0.3375	0.9197	0.3458
DFn	N/A	2	2	2
DFd	N/A	2	2	2
F	N/A	29316759259891	915103819	64127
p (F-test)	N/A	<0.0001	<0.0001	<0.0001
F-test significance ^{c.}	N/A	****	****	****
Compared to NmVenus-POLD3 + CmVenus-N-HELQ^{b.}				
MD ± SEM	-0.01471 ± 0.01457	N/A	-0.01432 ± 0.01457	-0.01471 ± 0.01457
t	1.009	N/A	0.9826	1.009
df	2.000	N/A	2.000	2.000
p (t-test)	0.4191	N/A	0.4294	0.4191
t-test significance ^{c.}	N/S	N/A	N/S	N/S

+/- d.	N/A	N/A	N/A	N/A
95% CI	-0.07740 to 0.04799	N/A	-0.07701 to 0.04837	-0.07740 to 0.04799
r ²	0.3375	N/A	0.3256	0.3375
DFn	2	N/A	2	2
DFd	2	N/A	2	2
F	29316759259891	N/A	32037	457164379
p (F-test)	<0.0001	N/A	<0.0001	<0.0001
F-test significance ^{c.}	****	N/A	****	****
Compared to NmVenus-RPA1 + CmVenus-N-HELQ ^{b.}				
MD ± SEM	-0.0003897 ± 8.141x10 ⁻⁰⁰⁵	0.01432 ± 0.01457	N/A	-0.0003890 ± 8.141x10 ⁻⁰⁰⁵
t	4.786	0.9826	N/A	4.778
df	2.000	2.000	N/A	2.000
p (t-test)	0.0410	0.4294	N/A	0.0411
t-test significance ^{c.}	*	N/S	N/A	*
+/- d.	-	N/A	N/A	-
95% CI	-0.0007399 to - 3.939x10 ⁻⁰⁰⁵	-0.04837 to 0.07701	N/A	-0.0007392 to -3.872x10 ⁻⁰⁰⁵
r ²	0.9197	0.3256	N/A	0.9194
DFn	2	2	N/A	2
DFd	2	2	N/A	2
F	915103819	32037	N/A	14270
p (F-test)	<0.0001	<0.0001	N/A	0.0001

F-test significance ^c .	****	****	N/A	***
Fluorescence Normalised Using CFU/mL With Outliers Removed				
n	3	2	2	3
M	1.177x10 ⁻⁰⁰⁸	0.0001365	0.0003270	7.124x10 ⁻⁰⁰⁷
SD	4.661x10 ⁻⁰⁰⁹	5.586x10 ⁻⁰⁰⁵	0.0001273	1.180x10 ⁻⁰⁰⁶
SEM	2.691x10 ⁻⁰⁰⁹	3.950x10 ⁻⁰⁰⁵	3.950x10 ⁻⁰⁰⁵	6.815x10 ⁻⁰⁰⁷
Compared to NmVenus + CmVenus^b.				
MD ± SEM	N/A	0.0001365 ± 3.950x10 ⁻⁰⁰⁵	0.0003270 ± 9.000x10 ⁻⁰⁰⁵	7.006x10 ⁻⁰⁰⁷ ± 6.815x10 ⁻⁰⁰⁷
t	N/A	3.455	3.633	1.028
df	N/A	1.000	1.000	2.000
p (t-test)	N/A	0.1793	0.1710	0.4120
t-test significance ^c .	N/A	N/S	N/S	N/S
+/- ^d .	N/A	N/A	N/A	N/A
95% CI	N/A	-0.0003654 to 0.0006384	-0.0008166 to 0.001471	-2.231x10 ⁻⁰⁰⁶ to 3.633x10 ⁻⁰⁰⁶
r ²	N/A	0.9227	0.9296	0.3458
DFn	N/A	1	1	2
DFd	N/A	2	2	2
F	N/A	143631286	745658333	64127
p (F-test)	N/A	<0.0001	<0.0001	<0.0001

F-test significance ^c .	N/A	****	****	****
Compared to NmVenus-POLD3 + CmVenus-N-HELQ^b.				
MD ± SEM	-0.0001365 ± 3.950x10 ⁻⁰⁰⁵	N/A	0.0001905 ± 9.829x10 ⁻⁰⁰⁵	-0.0001358 ± 3.951x10 ⁻⁰⁰⁵
t	3.455	N/A	1.938	3.437
df	1.000	N/A	1.371	1.001
p (t-test)	0.1793	N/A	0.2470	0.1801
t-test significance ^c .	N/S	N/A	N/S	N/S
+/- ^d .	N/A	N/A	N/A	N/A
95% CI	-0.0006384 to 0.0003654	N/A	-0.0004851 to 0.0008661	-0.0006371 to 0.0003655
r ²	0.9227	N/A	0.7326	0.9219
DFn	1	N/A	1	1
DFd	2	N/A	1	2
F	143631286	N/A	5.191	2240
p (F-test)	<0.0001	N/A	0.5266	0.0009
F-test significance ^c .	****	N/A	N/S	***
Compared to NmVenus-RPA1 + CmVenus-N-HELQ^b.				
MD ± SEM	-0.0003270 ± 9.000x10 ⁻⁰⁰⁵	-0.0001905 ± 9.829x10 ⁻⁰⁰⁵	N/A	-0.0003263 ± 9.000x10 ⁻⁰⁰⁵
t	3.633	1.938	N/A	3.625
df	1.000	1.371	N/A	1.000
p (t-test)	0.1710	0.2470	N/A	0.1713

t-test significance ^{c.}	N/S	N/S	N/A	N/S
+/- ^{d.}	N/A	N/A	N/A	N/A
95% CI	-0.001471 to 0.0008166	-0.0008661 to 0.0004851	N/A	-0.001470 to 0.0008170
r ²	0.9296	0.7326	N/A	0.9293
DFn	1	1	N/A	1
DFd	2	1	N/A	2
F	745658333	5.191	N/A	11628
p (F-test)	<0.0001	N/S	N/A	0.0002
F-test significance ^{c.}	****	N/A	N/A	***

Two-tailed Welch's t-tests were performed using GraphPad Prism 10 to determine if fluorescence triplicate means showed a significant difference compared to NmVenus + CmVenus, NmVenus-POLD3 + CmVenus-N-HELQ, and NmVenus-RPA1 + CmVenus-N-HELQ fluorescence triplicate means. a. Statistics: n = sample size, M = mean, SD = standard deviation, SEM = standard error of mean, MD = mean difference, t = t-value, df = degrees of freedom, p = p-value, 95% CI = 95% confidence interval, r² = coefficient of determination, DFn = degrees of freedom numerator, DFd = degrees of freedom denominator, F = F-value. b. Comparison is performed as mean in the column minus mean of the comparison protein pair. c. Significance level is denoted by asterisks over the column: * = p < 0.05, ** = p < 0.01, *** = p < 0.001, **** = p < 0.0001, N/S = 'no significance', and N/A = 'not applicable'. d. + indicates that that column protein pair has a higher mean than the comparative protein pair; - indicates it is lower.