



University of HUDDERSFIELD

University of Huddersfield Repository

Kulik, Anastasija

Characterisation of Polymerase δ -interacting Protein 2 (POLDIP2), a novel human protein involved in modulating DNA replication/repair and other physiological functions

Original Citation

Kulik, Anastasija (2019) Characterisation of Polymerase δ -interacting Protein 2 (POLDIP2), a novel human protein involved in modulating DNA replication/repair and other physiological functions. Masters thesis, University of Huddersfield.

This version is available at <http://eprints.hud.ac.uk/id/eprint/35267/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

<http://eprints.hud.ac.uk/>



Characterisation of Polymerase δ -interacting Protein 2
(POLDIP2), a novel human protein involved in
modulating DNA replication/repair and other
physiological functions

By: Anastasija Kulik BSc (Hons) Biochemistry

Supervisors: Dr Christopher Cooper and Dr Richard Bingham

A thesis submitted to the University of Huddersfield (Department of Biological
and Geographical Sciences) in partial fulfilment of the requirements for the degree
of Masters by Research (MRes) in Biological Sciences

September 2019

I confirm that, unless indicated otherwise, I undertook all of the work presented in this report.

Any work performed by other people is fully acknowledged.

Acknowledgments

With many thanks to my supervisor, Dr Christopher Cooper, for teaching me a wide variety of lab techniques used in molecular biology/biochemistry. Moreover, for his general support regarding non-academic issues, trust, encouragement and 24 hour availability. Also, I would like to thank my co-supervisor, Dr Richard Bingham for his patience during the explanation of complex structural biology topics, as well as, for his fast response. Additionally, I would like to thank my laboratory colleagues for their support in the lab: Naomi Nabi, Carla Asquith, Banu Tokatligil and Eden Conroy. Mum, dad and brother- thanks for love, belief, help and patience.

Abstract

Polymerase δ -interacting Protein 2 (POLDIP2) is composed of 368 amino acids, with a molecular weight of the 42 kDa and 37 kDa, respectively, before and after post-translational cleavage of the mitochondrial-targeting sequence (Hernandes, Lassègue and Griendling, 2017; Liu *et al.*, 2003). POLDIP2 was shown to be involved in the replication of cell cycle, lung functioning, maintenance of cellular homeostasis (Hernandes *et al.*, 2017). Also, POLDIP2 is involved in bone and DNA maintenance, neurodegenerative diseases, cancer and proliferation (Hernandes *et al.*, 2017). Although, the POLDIP2 is a multifunctional protein it remains unknown how POLDIP2 can bind so many different proteins in different subcellular compartments. Moreover, it remains unknown if the two POLDIP2 forms (37 kDa and 42 kDa) have different functions. The POLDIP2 biochemistry remains completely untouched, therefore this project focused on structural and functional POLDIP2 characterisation.

All stages of recombinant POLDIP2 structure determination pipeline were attempted; starting from the POLDIP2 gene sub-cloning and ending with protein structure crystallization. The POLDIP2 structure was nearly solved (~99.8%) at 2.9 Å. The obtained POLDIP2 3D structure was used for the surface charge distribution and *B* factor analysis. The POLDIP2 *B* factor analysis and predicted/known secondary structure elements superimposition revealed that both rigid and flexible POLDIP2 secondary structure elements are involved in protein-protein interactions. The POLDIP2 *B* factor analysis was superimposed with known/predicted secondary POLDIP2 structure elements, revealing that both rigid and flexible POLDIP2 secondary structure elements are involved in protein-protein interactions.

Moreover, it was attempted to crystallise the POLDIP2 in complex with the proliferating cell nuclear antigen (PCNA), as POLDIP2 structure with a binding partner is required for the detailed understanding of POLDIP2 function. Furthermore, *in vitro* crosslinking and size exclusion chromatography were used to understand the PCNA-POLDIP2 binding stoichiometry. Finally, the initial negative grid staining of the PCNA-POLDIP2 complex for further cryogenic electron microscopy (cryo-EM) studies was attempted giving future work directions.

List of the major abbreviations

Alzheimer's disease (AD)

Base excision repair (BER)

Carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1)

Carcinoembryonic antigen-related cell adhesion molecule 2 (CEACAM2)

Caseinolytic peptidase (Clp)

Cryogenic electron microscopy (cryo-EM)

Cyclin-dependent kinase (CDK)

F-box only protein 3 (FBxo3)

Fibroblast growth factor 2-FGF2

Homologous recombination (HR)

Human papillomavirus 16 E7 (HPV 16 E7)

Lipoate-activating enzyme Ac-CoA synthetase medium-chain family member 1 (ACSM1)

Matrix metalloproteinase 9 (MMP-9)

Microhomology-mediated end joining (MMEJ)

Mismatch repair (MMR)

Mitochondrial targeting peptide (mTP)

Molecular replacement (MR)

Mouse aortic smooth muscle cells (MASMC)

Mouse embryonic fibroblasts (MEFs)

Multiple Isomorphous Replacement (MIR)

Multi-wavelength Anomalous Dispersion (MAD)

NADPH oxidase 4 (Nox4)

Non-homologous end joining (NHEJ)

Non-small cell lung cancer (NSCLC)

Nucleotide excision repair (NER)

Peripheral blood mononuclear cell (PBMC)

Polymerase (Pol)

Polymerase delta-interacting protein 2 (POLDIP2)

Proliferating cell nuclear antigen (PCNA)

Reactive oxygen species (ROS)

Real-time PCR (RT-PCR)

Replication protein A (RPA)

Single-stranded binding proteins (SSBs)

Small interfering RNA (siRNA)

Translesion synthesis (TLS)

Tumor necrosis factor alpha (TNF- α)

Type 2 diabetes mellitus (T2DM)

Ubiquitin-binding domain (UBD)

Vascular smooth muscle cells (VSMC)

Table of Contents

| | |
|---|----|
| | 0 |
| Acknowledgments | 1 |
| Abstract | 2 |
| List of the major abbreviations | 4 |
| 1. Introduction | 13 |
| 1.1. POLDIP2, a poorly characterized cellular protein | 13 |
| 1.2. POLDIP2 predicted structure..... | 14 |
| 1.3. POLDIP2 cellular localization and tissue expression | 16 |
| 1.4. A general overview of DNA replication and the role of POLDIP2 | 20 |
| 1.5. DNA repair and POLDIP2 role in faithful genetic information transmission | 22 |
| 1.6. Unique human polymerase and POLDIP2..... | 25 |
| 1.7. POLDIP2 and cancer | 27 |
| 1.8. POLDIP2 and neurodegenerative disease..... | 29 |
| 1.9. POLDIP2 and bone physiology | 30 |
| 1.10. POLDIP2 and cellular homeostasis | 31 |
| 1.11. POLDIP2 and vascular function | 32 |
| 1.12. POLDIP2 and cell cycle regulation | 33 |
| 1.13. Project aims..... | 34 |
| 2. Materials and methods | 35 |

| | |
|---|----|
| 2.1. Materials | 35 |
| 2.2. Molecular Biology | 35 |
| 2.2.1. Polymerase chain reaction (PCR) | 35 |
| 2.2.2. DNA Agarose gel electrophoresis..... | 36 |
| 2.2.3. Competent <i>E. coli</i> cells | 36 |
| 2.2.4. <i>E. coli</i> culture preparation..... | 38 |
| 2.2.5. Transformation..... | 38 |
| 2.2.6. <i>E. coli</i> glycerol stock preparation | 38 |
| 2.2.7.1. pGTVL2 vector restriction digest | 39 |
| 2.2.7.2. <i>DpnI</i> treatment of PCR products..... | 39 |
| 2.2.7.3. T4 DNA polymerase treatment of pGTVL2 vector and PCR fragment | 40 |
| 2.2.7.4. Whole cell PCR colony screen..... | 41 |
| 2.2.9. DNA quantitation..... | 42 |
| 2.3.1. Small scale recombinant protein expression | 43 |
| 2.3.2. Small scale <i>E. coli</i> cell lysis..... | 43 |
| 2.3.3. Ni-NTA affinity chromatography small scale..... | 44 |
| 2.3.4. Large scale recombinant protein expression | 44 |
| 2.3.5. Large scale <i>E. coli</i> cell lysis..... | 45 |
| 2.4. Large scale protein purification | 46 |
| 2.4.1. Ni-NTA Affinity Chromatography (IMAC) and Tobacco Etch Virus (TEV)- Ni-NTA Affinity Chromatography. | 46 |

| | |
|---|----|
| 2.4.1.1. Ni-NTA Affinity Chromatography (IMAC) | 46 |
| 2.4.1.2. Reverse Ni-NTA Affinity Chromatography | 46 |
| 2.4.2. Ion exchange and heparin chromatography (IEC) | 47 |
| 2.4.3. Size Exclusion Chromatography (SEC)..... | 47 |
| 2.5. Protein Dialysis..... | 48 |
| 2.6. Protein Concentration | 48 |
| 2.7. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) | 48 |
| 2.7.1. Gel casting | 48 |
| 2.7.2. SDS-PAGE protein sample preparation..... | 49 |
| 2.7.3. Gel running conditions and staining | 49 |
| 2.9. Protein characterisation..... | 50 |
| 2.9.1. Mass spectrometry (MS)..... | 50 |
| 2.9.2. Analytical Size Exclusion Chromatography (SEC) | 51 |
| 2.9.3. <i>In vitro</i> protein crosslinking..... | 51 |
| 2.10. Structural Characterisation..... | 52 |
| 2.10.1. Circular Dichroism (CD) | 52 |
| 2.10.2. Electron microscopy (EM)..... | 52 |
| 2.10.3. Crystallography..... | 53 |
| 2.10.3.1. Vapour diffusion crystallography | 53 |
| 2.10.3.2. Screen design and optimization | 53 |

| | |
|--|-----------|
| 2.10.3.3. Crystal mounting and screening..... | 54 |
| 2.10.3.4. Protein structure refinement..... | 54 |
| 2.10.3.5. Protein structure visualization in PyMOL..... | 56 |
| 3. Results-discussion | 57 |
| 3.1. Ligation independent cloning (LIC) | 57 |
| 3.1.1 The use of highly parallelized LIC for POLDIP2 construct cloning | 57 |
| 3.1.2. PCR product analysis..... | 60 |
| 3.1.3. Vector restriction digest..... | 62 |
| 3.1.4 PCR colony screen..... | 62 |
| 3.2. Protein expression and purification..... | 64 |
| 3.2.1. Small scale protein expression analysis using 2 growth media and 2 vector systems | 64 |
| 3.2.2 Large scale POLDIP2 protein expression..... | 67 |
| 3.2.3.1. POLDIP2 (1-368) IMAC purification..... | 69 |
| 3.2.3.2. POLDIP2 (1-368) IEC purification..... | 71 |
| 3.2.3.3. POLDIP2 (1-368) SEC purification..... | 73 |
| 3.2.3.4. POLDIP2 (51-368) IMAC purification..... | 74 |
| 3.2.3.5. POLDIP2 (51-368) SEC purification..... | 76 |
| 3.3. Structural protein characterization | 79 |
| 3.3.1. POLDIP2 (51-368) secondary structure characterisation by circular dichroism | 79 |
| 3.3.2. Protein Crystallography towards POLDIP2 structural determination..... | 81 |

| | |
|--|------------|
| 3.3.2.1 Crystallisation of POLDIP2 (51-368) | 81 |
| 3.3.2.2. POLDIP2 (51-368) crystallisation optimisation | 85 |
| 3.3.2.3. Crystal mounting and X-ray diffraction screening..... | 89 |
| 3.3.2.4. POLDIP2 (51-368) structure solution..... | 91 |
| 3.3.2.4.1. Molecular replacement of POLDIP2 (51-368) | 91 |
| 3.3.3.1. POLDIP2 (51-368) secondary structure comparison between PRISPED predicted and x-ray solved model. | 97 |
| 3.3.3.2. POLDIP2 (51-368) secondary structure elements and bound molecules. | 101 |
| 3.3.3.3. POLDIP2 (51-368) labeling by <i>B</i> factor | 102 |
| 3.3.3.4. POLDIP2 (51-368) surface charge distribution | 103 |
| 3.3.3.5. POLDIP2 (51-368) statistics..... | 106 |
| 3.3.3.6. Relating POLDIP2 (1-368) structure to protein interaction partners..... | 107 |
| 3.4. Production of PCNA-POLDIP2 complexes..... | 111 |
| 3.4.1. PCNA purification | 111 |
| 3.4.1.1. PCNA IMAC purification..... | 111 |
| 3.4.1.2. PCNA SEC purification..... | 113 |
| 3.4.2. Analysis of PCNA-POLDIP2 (1-368) complex formation by <i>in vitro</i> crosslinking..... | 114 |
| 3.4.3. PCNA-POLDIP2 (1-368) complex co-purification | 116 |
| 3.4.4. PCNA-POLDIP2 complex studies..... | 119 |
| 3.4.5. Negative staining EM for PCNA-POLDIP2 (1-368) complex, towards cryo-EM. | 122 |
| 4. Conclusion | 124 |

| | |
|---|------------|
| 4.1. Conclusion and future work suggestions | 124 |
| Appendices | 130 |
| References | 135 |

1. Introduction

1.1. POLDIP2, a poorly characterized cellular protein

Polymerase delta-interacting protein 2 (also known as POLDIP2, Mitogenin I and PDIP38) - is a multifunctional human protein, initially, discovered as a binding partner of DNA polymerase δ and proliferating cell nuclear antigen (PCNA). This suggested its role in DNA repair and replication (Hernandes *et al.*, 2017; Naokatu *et al.*, 2006; Liu *et al.*, 2003). POLDIP2 is very conserved in metazoa, but it is absent in prokaryotes, fungi and plants (Hernandes *et al.*, 2017). Interestingly, the *POLDIP2* gene encoding for POLDIP2 is a unique gene, which does not have any paralogous genes (Hernandes *et al.*, 2017; Sutliff *et al.*, 2013). As a consequence POLDIP2 does not belong to any protein family, thus being a unique member, and making its function completely unpredictable (Hernandes *et al.*, 2017; Sutliff *et al.*, 2013). POLDIP2 is composed of 368 amino acids, with a molecular weight of 42 kDa and 37 kDa, respectively, before and after post-translational cleavage of the mitochondrial-targeting sequence (Hernandes *et al.*, 2017; Liu *et al.*, 2003). Structurally, POLDIP2 comprises of the N-terminal mitochondrial targeting sequence, a hemimethylated DNA binding domain (YccV-like domain) and a DUF525 domain and (Figure 1) (Fukasawa *et al.*, 2015; Claros and Vincens, 1996; Emanuelsson, Nielsen, Brunak and Heijne, 2000).

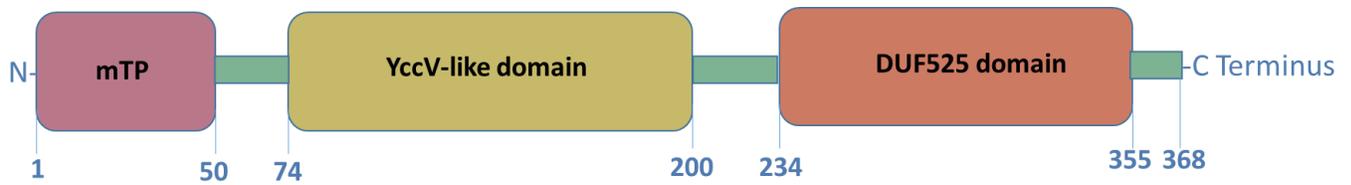


Figure 1. General representation of predicted POLDIP2 domains. According to TargetP, Mitoprot and MitoFates online software tools, the N-terminal POLDIP2 region from residue 1-49 was predicted to be a cleavable mitochondrial targeting peptide (mTP) (Fukasawa *et al.*, 2015; Claros and Vincens, 1996; Emanuelsson *et al.*, 2000). National Centre for Biotechnology Information (NCBI) identified the homology to 2 domains: low homology YccV-like domain residues 74-200 ($E = 1.3e-12$) and higher homology YccV-like domain 234-355 residues ($E = 1.3e-12$) (Hernandes *et al.*, 2017; Figure created using PowerPoint).

1.2. POLDIP2 predicted structure

Determination of the POLDIP2 structure/structure of complexes is required to uncover its function. Although, nothing is known about POLDIP2 3D structure, two domains were predicted to be present in POLDIP2 using informatics tools: YccV-like domain and DUF525 domain (Hernandes *et al.*, 2017). The YccV-like domain is often found in eukaryotic metabolic enzymes including cysteine proteases and is also present in prokaryotic heat shock protein Q (HspQ) (Brown *et al.*, 2014; Shimuta *et al.*, 2004; Nishimura *et al.*, 2015). Moreover, YccV domain was found in Clp proteases, which are located in prokaryotic/plant mitochondria and/or chloroplasts depending on species (Nishimura *et al.*, 2015).

In *E. coli* the YccV protein is known to stimulate protein degradation (for example DnaA) during the heat shock stress, as well as, it is known to bind preferentially hemimethylated DNA thus stabilising other proteins during the replication initiation process (Brown *et al.*, 2014; Shimuta *et al.*, 2004; d'Alençon *et al.*, 2003; d'Alençon *et al.*, 2003). Maga *et al.*, (2013) based on the knowledge that POLDIP2 shares 34% sequence identity to the YccV domain which binds DNA

performed an electrophoretic shift assay. The performed electrophoretic shift assay revealed that POLDIP2 does not bind to the undamaged or damaged DNA (Maga *et al.*, 2013).

The ApaG domain (also known as DUF525) is a very rare motif which is found in humans (Krzysiak *et al.*, 2016). The ApaG domain appears only in 2 human proteins: POLDIP2 and F-box only protein 3 (FBxo3) (Krzysiak *et al.*, 2016). These 2 proteins are similar to the bacterial ApaG protein. However, in contrast to bacterial ApaG domain, human ApaG domain found in FBxo3 protein seems to have a different function (Krzysiak *et al.*, 2016). Human ApaG domain does not have an ability to bind metal ions (Mg^{2+} ; Co^{2+}), whereas the bacterial ApaG protein seems to be responsible for the maintenance of divalent cation homeostasis in the cell and potentially could be involved in protein-protein interactions (Krzysiak *et al.*, 2016). Although, the structure of human ApaG domain was determined, its biochemical role remains unknown (Krzysiak *et al.*, 2016; Ilyin, Riolland, Pigeon and Guguen-Guillouzo, 2000). ApaG protein found in human FBxo3 protein has a very similar structure to the bacterial ApaG protein (Krzysiak *et al.*, 2016). Same as in bacteria the human ApaG domain is comprised of 7 β -sheet strands which are surrounded by 4 extending loops (Krzysiak *et al.*, 2016). Hernandez *et al.*, (2017) using bioinformatics tools, such as Phyre2 and I-TASSER software and 5 available ApaG homologous proteins in the PDB, predicted the 3D structure of POLDIP2 DUF525 (ApaG) domain which is shown in Figure 2. Also, the bioinformatics tools, predict the POLDIP2 YccV-like domain to have mainly β -sheet secondary structure with a terminating α -helix (Brown *et al.*, 2014; Shimuta *et al.*, 2004; Nishimura *et al.*, 2015).

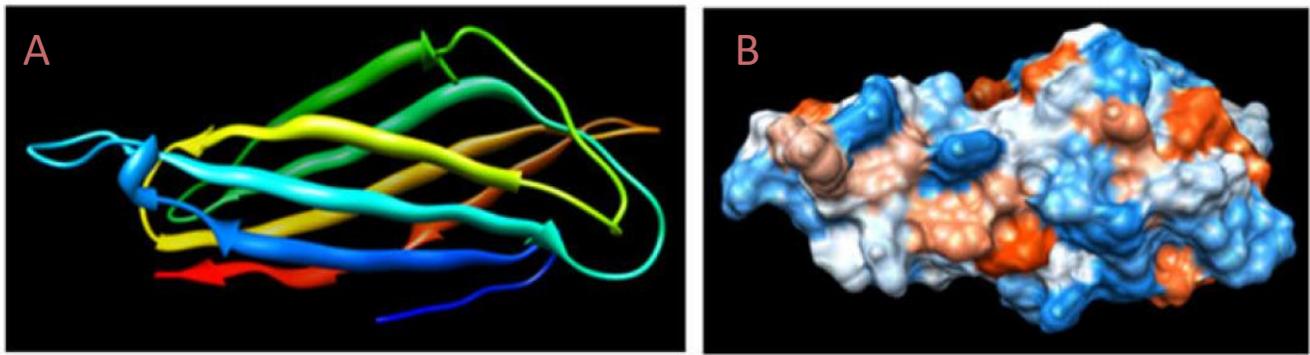


Figure 2. C-terminal POLDIP2 DUF525 domain 3D structure prediction. Ribbon DUF525 representation is shown on the panel A and DUF525 surface charge representation is shown on the panel B (I-TASSER and Phyre2 software) (Hernandes *et al.*, 2017). Red, blue and grey (surface charge figure on the right) correspond respectively to positive, negative and neutral amino acids surface charge.

Interestingly, POLDIP2 protein has a highly conserved GXGXXG motif which is known to be responsible for the binding to DNA and known to be present in kinases (Liu *et al.*, 2003; Aravind and Koonin, 1998). The presence of such motif suggests that POLDIP2 might be able to bind pyrophosphate or nucleotide triphosphate (Liu *et al.*, 2003; Aravind and Koonin, 1998; Ryazanov *et al.*, 1997). Also, POLDIP2 has a GXGVVGXXPX motif, which actually corresponds to the NAD- and FAD-binding proteins GXGXXG motif, suggesting POLDIP2 potential involvement in NAD and/or FAD binding (Liu *et al.*, 2003; Rossmann, Moras and Olsen, 1974; Dym and Eisenberg, 2001).

1.3. POLDIP2 cellular localization and tissue expression

Human POLDIP2 is found in the nucleus, mitochondria, cell membrane or cytoplasm depending on cell type and can bind many different proteins (Figure 3) (Hernandes *et al.*, 2017). POLDIP2 expression in humans was reported in different cell types, for example, POLDIP2 expression was found in aortic smooth muscle cells, vascular smooth muscle cells, astrocytes, epithelial cells, brain endothelial cells, kidney fibroblasts and cortical neurons (Lyle *et al.*, 2009; Hernandes *et al.*,

2017). This suggests that POLDIP2 can be involved in many different processes, depending on its subcellular localization and the binding partners available in specific cell types (Figure 3). At first glance it may seem unlikely- to have so many functions. However, several proteins which switch their function depending on the nucleus and cytoplasmic localization have been found (*e.g.* some families of LIM domain proteins) (Hernandes *et al.*, 2017). Therefore, POLDIP2 has many functions, which seem to be a consequence of a specific interaction with one of many binding partners (Figure 4).

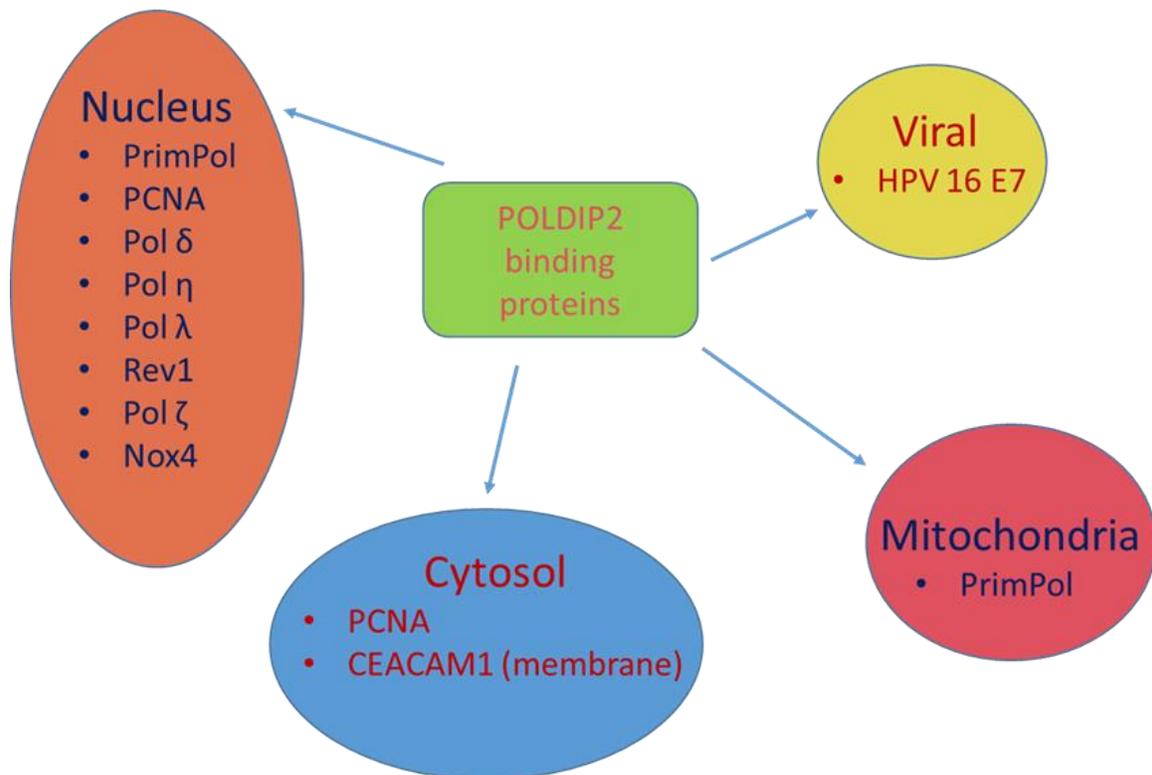


Figure 3. POLDIP2 binding partners. POLDIP2 interaction with other proteins is dependent on the subcellular compartment.

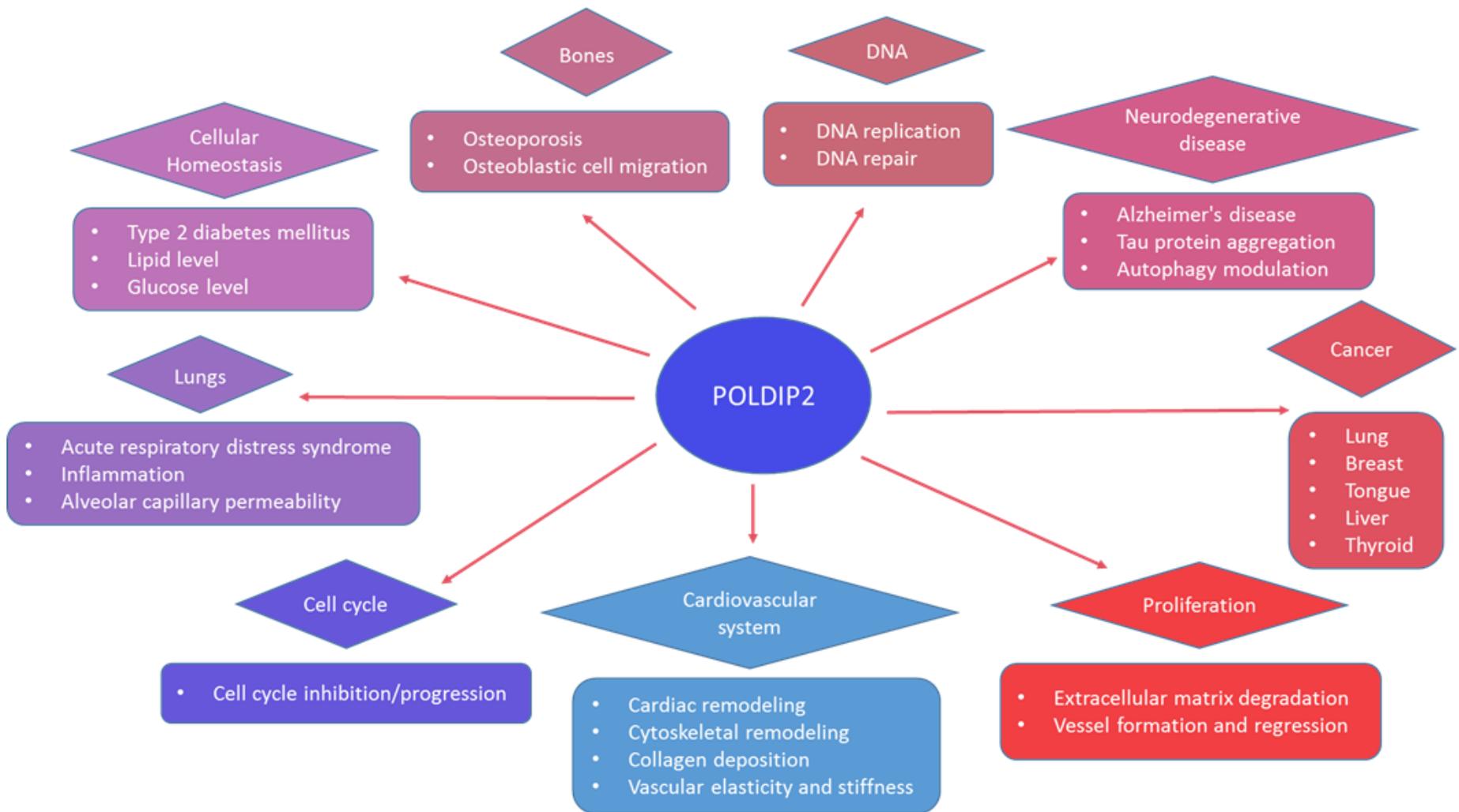


Figure 4. Multiple POLDIP2 functions. POLDIP2 plays role in cell cycle, lungs functioning, maintains cellular homeostasis, involved in bone and DNA maintenance, neurodegenerative diseases, cancer and proliferation.

Upon the entry to the mitochondria the POLDIP2 mitochondrial targeting sequence (mTS) is cleaved (Xie *et al.*, 2005). Thus, truncated POLDIP2 forms tagged with C-terminal GFP were used to reveal the mTS sequence (Xie *et al.*, 2005). The result showed that the construct containing the first 35 N-terminal residues are sufficient for the POLDIP2 targeting into mitochondria (Xie *et al.*, 2005). Moreover, the first 35 N-terminal residues tagged to the GFP alone, show the mitochondrial localization, further confirming the importance of this sequence (Xie *et al.*, 2005). However, even though an actual mTS cleavage site corresponds to the first 35 N-terminal residues, the consensus MPP cleavage recognition region predicted by the TargetP and MITOPROTII programs is located, respectively between 50-51 and 58-59 N-terminal POLDIP2 residues and is predicted to form an usual N-terminal mTS α -helix (Xie *et al.*, 2005). Both predictions could be correct as mitochondrial proteins could be cleaved in two steps, first by mitochondrial processing peptidase (MPP) at 58-59 residues and then by major intrinsic protein (MIP) at 50-51 residues (Xie *et al.*, 2005). Thus, the first approximately 50 residues are believed to be cleaved by MPP upon the POLDIP2 entry into the mitochondria (Xie *et al.*, 2005).

Although, the mTP sequence is predicted to target the newly synthesised proteins to the mitochondria, Klaile *et al.*, (2007) showed that there is another molecule which is also responsible for POLDIP2 cellular translocation. Yeast-two hybrid screen, co-immunoprecipitation, surface plasmon resonance, indirect immunofluorescence and differential centrifugation proved that carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1) can bind POLDIP2 (Klaile *et al.*, 2007). *In vivo* studies using specific antibodies against POLDIP2 revealed its interaction with both long and short CEACAM1 isoforms (CEACAM1-L and CEACAM1-S) (Klaile *et al.*, 2007). CEACAM1 can bind to POLDIP2 and translocate POLDIP2 to another

cellular compartment, for example from the cytoplasm to the nucleus and vice versa (Klaile *et al.*, 2007). Using truncated CEACAM1-L and POLDIP2 forms it was found that cytosolic C-terminal domain of CEACAM1-L and C-terminal residues 330–348 –POLDIP2 are essential for this interaction and therefore for POLDIP2 translocation into another cellular compartment (Klaile *et al.*, 2007).

POLDIP2 mRNA expression level varies between different types of rat tissue (Lyle *et al.*, 2009). For instance, POLDIP2 expression levels are extremely low in spleen and thymus, while the highest POLDIP2 expression level is observed in the diaphragm (Lyle *et al.*, 2009). Besides that, the variability between different POLDIP2 forms is observed in different tissue (Lyle *et al.*, 2009). Notably, 42 kDa POLDIP2 is mainly expressed in heart and aorta, whereas 37 kDa POLDIP2 has higher expression level in kidney, diaphragm and lung (Lyle *et al.*, 2009). Such diverse POLDIP2 expression makes this protein highly multifunctional (Lyle *et al.*, 2009).

1.4. A general overview of DNA replication and the role of POLDIP2

In human two main DNA polymerases (Pol), Pol δ and Pol ϵ are involved at the DNA replication fork (Lujan, Williams and Kunkel, 2016; Tsurimoto, Melendy and Stillman, 1990; Sun *et al.*, 2015). Pol δ is responsible for the DNA synthesis in 5'- to 3-' direction predominantly on the lagging DNA strand, but it is also able to carry the DNA synthesis on the leading DNA strands (Tsurimoto, Melendy and Stillman, 1990; Johnson, Klassen, Prakash and Prakash, 2015; Nick McElhinny *et al.*, 2008; Garbacz *et al.*, 2018). Pol δ with the aid of the PCNA synthesizes the lagging DNA strand discontinuously in 5'- to 3'- direction, by forming Okazaki fragments until it reaches the 5-' end of another Okazaki fragment (Stodola and Burgers, 2016; Lujan, Williams and

Kunkel, 2016). Whereas, the leading DNA strand is synthesised continuously by Pol ϵ bound to the PCNA (Foley, Couto, Rauf and Boyke, 2019; Pursell *et al.*, 2007; Stodola and Burgers, 2017; Hedglin, Pandey and Benkovic, 2016). However, currently the DNA replication is not properly understood and the exact way how Pol δ and Pol ϵ synthesise the DNA remains an active area of research (Foley *et al.*, 2019; Aria and Yeeles, 2018).

Human Pol δ is a pentameric protein, made of p125, p50, p68 and dimeric p12 subunits, (Khandagale, Peroumal, Manohar and Acharya, 2019). Lie *et al.*, (2003), used both 125 kDa and 50 kDa Pol δ subunits as a bait in a yeast two hybrid screen. Among the 22 identified 50 kDa subunit (Pol δ) binding proteins one of them was POLDIP2 (Liu *et al.*, 2003). Also, co-immunoprecipitation assay, yeast two hybrid screen, size exclusion chromatography and crosslinking identified that PCNA binds to 125 kDa Pol δ subunit (Zhang *et al.*, 1999). Moreover, using the pairwise yeast two hybrid assays, pull down assay and western blotting it was found that POLDIP2 binds to PCNA (Liu *et al.*, 2003). Liu *et al.*, (2003) identified 3 putative POLDIP2 motifs 81-88, 151-158 and 193-200 residues which could be potentially involved in the interaction with the PCNA QXXZXXF(F/Y) (Z-aliphatic residue) motif (Cox, 1997). As Pol δ is a multimeric protein and both PCNA and POLDIP2 bind to different Pol δ subunits and POLDIP2 can bind both PCNA and Pol δ , this could suggest that such Pol δ interaction with PCNA might be indirect during DNA replication/repair/cell cycle regulation process and could be mediated by POLDIP2 (Liu *et al.*, 2003; Klaile, Kukalev, Obrink and Müller, 2008). Moreover, POLDIP2 and PCNA binding to different Pol δ subunits suggest that POLDIP2 and PCNA could be in close proximity to each other (Liu *et al.*, 2003). Maga *et al.*, (2013), showed that POLDIP2 stimulates Pol δ by increasing Pol δ affinity for the PCNA binding, suggesting that POLDIP2 might mediate the Pol δ and PCNA interaction by acting as a chaperone.

Interestingly, POLDIP2 also interacts with human papillomavirus 16 E7 (HPV 16 E7) oncoprotein (Xie *et al.*, 2005). HPV 16 E7 can replicate viral DNA both *in vivo* and *in vitro* and the presence of p68 and p50 subunits of Pol δ enhances the DNA replication (Xie *et al.*, 2005). This raises the possibility that Poldip2 might mediate viral DNA replication via Pol δ and human papillomavirus 16 E7 oncoprotein, acting as a bridging protein (Xie *et al.*, 2005).

1.5. DNA repair and POLDIP2 role in faithful genetic information transmission

To achieve faithful transmission of genetic information from one generation to another, DNA replication accuracy is tightly controlled (Guo, Kosarek-Stancel, Tang and Friedberg, 2009). One of such ways, which allows to maintain the accurate transmission of genetic information is the presence of Pol δ and pol ϵ 3'- to 5'- exonuclease activity (Shevelev and Hübscher, 2002). Furthermore, protective mechanisms exist in our cells to ensure accurate DNA replication following damaging insults to the DNA (Dahal, Dubey and Raghavan, 2018; Carell, 2015; Guo *et al.*, 2009). Such mechanisms which maintain the DNA integrity are the following: base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), non-homologous end joining (NHEJ), homologous recombination (HR) and microhomology-mediated end joining (MMEJ) (Dahal *et al.*, 2018; Carell, 2015). However, Pol ϵ and Pol δ could also participate in these mechanisms (Johnson *et al.*, 2015; Hübscher and Berg, 1993). For example, Pol δ could be involved in MMR, BER, HR and NER, whereas Pol ϵ could participate in BER and HR (Johnson *et al.*, 2015; Jessberger, Podust, Hübscher and Berg, 1993; Loeb and Monnat, 2008; Longley, Pierce and Modrich, 1997; Maloisel, Fabre and Gangloff, 2008; Prindle and Loeb, 2012).

However, some lesions can escape these DNA repair pathways, resulting in a DNA synthesis prohibition and putting in danger cells viability (Guo *et al.*, 2009). One of the ways to remove such lesions is the use of special translesion synthesis (TLS) DNA polymerases (Guo *et al.*, 2009). TLS polymerases gather on the DNA upon the removal of Pol δ and Pol ϵ which are unable to bypass many DNA lesions causing DNA synthesis arrest (Schmitt, Matsumoto and Loeb, 2009; Hirota *et al.*, 2016; Schmitt, Matsumoto and Loeb, 2009). The TLS polymerases such as Rev1, Pol η , Pol ι , and Pol κ belong to Y-family (Guo *et al.*, 2009; Barnes, Hile, Lee and Eckert, 2017). Such polymerases have less spatially constrained active site and as a result can accommodate/fit disordered/bulky bases exhibiting much more weaker proofreading activity or do not possess proofreading ability in comparison to the DNA replicative polymerases (Hirota *et al.*, 2016; Sale, 2013; McCulloch and Kunkel, 2008; Prakash *et al.*, 2005). Therefore, TLS polymerases are recruited onto the Dna only for short DNA stretches synthesis (Prakash, Johnson and Prakash, 2005).

A physical interaction of POLDIP2 and Pol λ was confirmed by co-immunoprecipitation (Maga *et al.*, 2013). Additionally, POLDIP2 physical interaction with Pol η was confirmed by pull down and yeast two hybrid screens (Maga *et al.*, 2013; Tissier *et al.*, 2010). It was found out that, POLDIP2 enhances the activity of Pol λ and Pol η , whereas it has no effect on Pols β or Pol ι (Maga *et al.*, 2013). In the presence of POLDIP2 and the Pol λ or Pol η the synthesised DNA length is increased, while Pols β or Pol ι synthesised DNA length remains the same (Maga *et al.*, 2013). Such result suggests that binding of POLDIP2 to Pol η and/or Pols λ induces conformational changes in the polymerase which result in the ability to bind the DNA more tightly (Maga *et al.*, 2013).

Interestingly, enzymatic assay shows that Pol η and Pol λ containing POLDIP2 after the addition of PCNA and Replication protein A (RPA), synthesise even longer DNA fragments (Maga and Hubscher, 2003; Maga *et al.*, 2013). Although, the presence of POLDIP2, PCNA and RPA enhance both pol η and Pols λ , the effect on pol η is weaker than on Pols λ (Maga *et al.*, 2013; Mondol, Stodola, Galletto and Burgers, 2019). Such results could suggest that POLDIP2 changes PCNA and RPA conformation and, as a result, enabling PCNA to stabilise the polymerase on the DNA more tightly (Maga *et al.*, 2013). Moreover, this result shows that POLDIP2, PCNA and RPA have more influence on the most accurate human TLS Pol λ capable to bypass the 8-oxo-7,8-dihydroguanine lesion in comparison to Pol η (Maga *et al.*, 2013; Johnson, Prakash and Prakash, 1999). A similar result was obtained using the truncated Pol λ 255-574, corresponding to the Pol λ region which possesses the catalytic activity and is still capable of TLS (Maga *et al.*, 2013). Truncated Pol λ 255-574 synthesises the DNA more slowly than the full-length Pol λ , however, POLDIP2 addition increased the Pol λ 255-574 processivity (Maga *et al.*, 2013). This points out that the Pol λ 255-574 region which represents the catalytic Pol λ domain could be the binding region of POLDIP2 (Maga *et al.*, 2013).

Using various truncated Pol η forms the Pol η region binding the POLDIP2 was revealed (Tissier *et al.*, 2010). It was established that POLDIP2 (38 kDa) directly interacts with ubiquitin-binding zinc finger domain (UBZ) of Pol η (Tissier *et al.*, 2010). Pol η mutagenesis studies show that parts of UBZ motif which is involved in zinc chelation does not interact with POLDIP2 (Tissier *et al.*, 2010). Pol η mutation of H650A and H654A (two hydrogens which interact with Zn^{2+}) or residue D (D652A) which are involved in the C-terminal α -helix formation lead to an inability to bind ubiquitin, but such Pol η mutation does not perturb POLDIP2 binding to the polymerase (Plosky *et al.*, 2006; Tissier *et al.*, 2010). Therefore, it can be concluded that the residues which are the

most important for the ubiquitin binding are not the residues which are involved in the interaction with the POLDIP2 (Tissier *et al.*, 2010).

POLDIP2 is also capable of interacting with Rev1 and Pol ζ (via Rev7) (Tissier *et al.*, 2010). However, no interaction was observed between POLDIP2 and Pol κ or Pol ι (Tissier *et al.*, 2010). The interaction region between POLDIP2 and other DNA polymerases was not studied yet, but for example, Rev7 does not have a ubiquitin-binding domain (UBD), yet is still able to interact with POLDIP2, suggesting that except UBD domain there are other important unidentified domains involved in POLDIP2 interaction with other TLS polymerases (Tissier *et al.*, 2010). Some translesion DNA synthesis polymerases, for example, REV1 and Pol η bind Lys-164 mono-ubiquitinated PCNA which is ubiquitinated by ligase RAD18 (Garg and Burgers, 2005; Choe and Moldovan, 2017; Bienko *et al.*, 2005; Kannouche, Wing and Lehmann, 2004). Such TLS polymerase interaction with Lys-164 mono-ubiquitinated PCNA induces conformational proteins changes, which increase TLS polymerase's activity and on their way facilitate DNA lesion bypass process (Garg and Burgers, 2005; Choe and Moldovan, 2017; Bomar, Pai, Tzeng, Li and Zhou, 2007; Sabbioneda *et al.*, 2009). Hence, as UBD is also important in binding ubiquitylated PCNA it could serve as a crosstalk, which switches replicative polymerases to TLS polymerases then lesions are encountered at the replication fork (Garg and Burgers, 2005; Choe and Moldovan, 2017).

1.6. Unique human polymerase and POLDIP2

DNA-directed primase/polymerase protein (PrimPol) is a protein with both DNA polymerase and DNA primase activities, found across prokaryotes, eukaryotes, archaea and even mini-viruses (Guilliam, Bailey, Brissett and Doherty, 2016; Gupta, Lad, Ghodke, Pradeepkumar and

Kondabagil, 2019). Unlike other primases found in human genome, PrimPol is able to start primer synthesis straight away using dNTPs instead of NTPs, in a such way eliminating the need to ‘erase’ the short ribonucleotide primers made by RNA polymerase (Guilliam *et al.*, 2016; Baranovskiy *et al.*, 2015; Torregrosa-Muñumer *et al.*, 2017; Perera *et al.*, 2013). Interestingly, but this is a unique PrimPol feature as none of other known 17 human DNA polymerases is able to synthesise the primer, thus initially requiring the RNA polymerase (primase) (García-Gómez *et al.*, 2013; Frick and Richardson, 2001).

An interaction between POLDIP2 and PrimPol was identified using pull-down assays (Guilliam *et al.*, 2015). Among POLDIP2, other PrimPol binding partners involved in DNA replication and repair were identified, such as: single-stranded binding proteins (SSBs), SSBs mitochondrial equivalent protein (mtSSB) and replication protein A (RPA) (Guilliam *et al.*, 2015; Martínez-Jiménez, Lahera and Blanco, 2017). Surprisingly, PrimPol-PCNA interaction was not observed in the pull down-assay, while POLDIP2 is known to directly interact with the PCNA (Liu *et al.*, 2003; Guilliam *et al.*, 2015). PrimPol directly binds to POLDIP2 and this was confirmed using mass spectrometry and cross-linking (Guilliam *et al.*, 2016). Crosslinking studies revealed that the highest crosslinking was between N-terminal POLDIP2 residues 1-8 and Lysine 70 in the PrimPol (Guilliam *et al.*, 2016). Interestingly, the PrimPol region which binds to POLDIP2 has strong sequence similarity to the Pol η region which also binds to POLDIP2 (Guilliam *et al.*, 2016; Klaile *et al.*, 2007; Tissier *et al.*, 2010). Both PrimPol and Pol η bind to the POLDIP2 N-terminus which is known to be the mitochondrial targeting sequence (Guilliam *et al.*, 2016; Klaile *et al.*, 2007). Moreover, the truncated POLDIP2 51-368 does not enhance PrimPol processivity, activity and DNA binding further suggesting the importance of POLDIP2 mitochondrial targeting sequence for the binding with these 2 proteins (Guilliam *et al.*, 2016; Klaile *et al.*, 2007).

POLDIP2 binding enhances PrimPol capacity to bind to DNA and reinforce polymerase activity (Guilliam *et al.*, 2016). Such POLDIP2 ability to enhance PrimPol efficiency/fidelity is very similar to the PCNA ability to enhance Pol η and Pol λ efficiency/fidelity (Maga *et al.*, 2013). Moreover, POLDIP2 binding to PrimPol enhances PrimPol ability to bypass pyrimidine 6-4 pyrimidone photoproducts (6-4PPs) and 8-oxo-7,8-dihydrodeoxyguanosine (8-oxoG) DNA lesions (Guilliam *et al.*, 2016; García-Gómez *et al.*, 2013). Furthermore, in the presence of POLDIP2, PrimPol can synthesise longer/extended primers (Guilliam *et al.*, 2016). Electrophoretic mobility shift assay showed that full-length PrimPol in the presence of POLDIP2 has a higher affinity for DNA (Guilliam *et al.*, 2016). This suggests that POLDIP2 is capable of binding to the PrimPol on the DNA and such interaction occurs via the PrimPol catalytic domain, while POLDIP2 alone is not able to bind the DNA (Guilliam *et al.*, 2016). It is of significant importance to further investigate the role of PrimPol-POLDIP2 interactions in mitochondria. Current research on PrimPol-POLDIP2 proves their mitochondrial localization and prompts that Pol γ and PrimPol are not the sole polymerase/primase involved in the human mitochondrial DNA replication, suggesting mitochondrial genome integrity is of considerable greater complexity than currently thought (Falkenberg, Larsson and Gustafsson, 2007; Guilliam *et al.*, 2016).

1.7. POLDIP2 and cancer

POLDIP2 could be involved in different cancer types. Chen *et al.*, (2018), studies show that POLDIP2 gene might be implicated in one of the most common cancer types, the lung cancer. *In vitro* and *in vivo* studies show that POLDIP2 knockdown reduced tumorigenicity and metastasis in non-small cell lung cancer (NSCLC) which accounts for 80% of lung cancers (Chen *et al.*, 2018). It was observed that POLDIP2 knockdown down-regulates other genes involved in cancer

such as: Cyclin D1, SLUG, CDH2 and TWIST (Chen *et al.*, 2018; Burch and Heintz, 2005). Chian *et al.*, (2016), using peripheral blood mononuclear cells (PBMC) containing nucleic acid markers, also proved POLDIP2 involvement in NSCLC. Using real-time PCR (RT-PCR) and statistical analysis of 187 and 310 blood samples, from unhealthy and healthy humans respectively, 8 genes were potentially identified to be tumorigenic and POLDIP2 was one among them (Chian *et al.*, 2016). Moreover, 14 NSCLS tissue samples obtained from patients revealed decreased POLDIP2 mRNA expression levels (Chen *et al.*, 2018). Nevertheless, even if there are some indications that POLDIP2 might be involved in the lung cancer formation, it remains unclear how POLDIP2 plays a role in lung cancer formation (Chen *et al.*, 2018).

Additionally, POLDIP2 could be involved in liver cancer as decreased NADPH oxidase 4 (Nox4) protein levels are observed in liver tumor cell lines derived from the mouse and POLDIP2 is known to interact directly with Nox4 protein (Crosas-Molist *et al.*, 2014; Weyemi *et al.*, 2012; Lyle *et al.*, 2009; Manuneehi, Cartland and Kavurma, 2017). Increased levels of Nox4 expression leading to an imbalance of H₂O₂ homeostasis (increased levels) have been recently proposed to be the cause of thyroid cancer (Ameziane-El-Hassani, Schlumberger and Dupuy, 2016; Salmeen, Park and Meyer, 2010). Moreover, Rodrigo *et al.*,(2014), showed that POLDIP2 binding to Nox 4 increases reactive oxygen species (ROS) production by 3 fold in rat thyrocytes and could potentially lead to the DNA damage causing the carcinogenic effect. As both POLDIP2 and Nox4 have high expression level in thyrocytes, this suggests that POLDIP2 could be involved in thyroid cancer formation caused by DNA damaging (Ameziane-El-Hassani, Schlumberger and Dupuy, 2016; Rodrigo *et al.*, 2014).

1.8. POLDIP2 and neurodegenerative disease

POLDIP2 may be involved in the progression of neurodegenerative disease formation (Kim *et al.*, 2015). Alzheimer's disease (AD) is a neurodegenerative disorder resulting in dementia and inability to learn (Kumar, Singh and Ekavali, 2015). Tau protein (accumulation of which is associated with neurodegenerative diseases) was observed to form aggregates when POLDIP2 is overexpressed in neuronal cells (Kim *et al.*, 2015; Kumar *et al.*, 2015; De Felice *et al.*, 2008). Interestingly, mouse neuron treatment with different pathogenic factors which are known to be involved in the neurodegenerative diseases such as amyloid beta, tumor necrosis factor alpha (TNF- α) or H₂O₂ increased the expression of POLDIP2 (Kam *et al.*, 2013; Kim *et al.*, 2015). However, in SH-SY5Y human neuroblastoma cells, POLDIP2 knock-down reduces H₂O₂ production level leading to Tau protein aggregation (Kim *et al.*, 2015). This suggests, that POLDIP concentration should be tightly regulated in the cell, as POLDIP2 concentration imbalance either way leads to Tau aggregation (Kim *et al.*, 2015).

Two truncated POLDIP2 forms, which have YccV-like domain and DUF525 domain were used to examine if both domains are necessary for the autophagy and Tau aggregation (Kim *et al.*, 2015). Results revealed, that DUF525 POLDIP2 domain sufficiently increases Tau aggregation and is crucial for the autophagy inhibitory function modulation (Kim *et al.*, 2015). It can be concluded that POLDIP2 regulates the autophagy flux through the DUF525 POLDIP2 domain (Kim *et al.*, 2015). Also, it should be taken into account that Alzheimer's disease is a multifactorial disease as no single gene induces this disease, meaning that POLDIP2 is not the only protein which controls the Tau aggregation process (Kim *et al.*, 2015; Kumar *et al.*, 2015; De Felice *et al.*, 2008).

1.9. POLDIP2 and bone physiology

Osteoporosis is one of the most morbidity associated disorders, in which bone mass and strength are reduced, such as increasing bone fracture risk (Katsumura *et al.*, 2016). The two cell types playing a crucial role in bone homeostasis: osteoblasts and osteoclasts (Schröder, 2019). Osteoblasts build the organic bone matrix by being responsible for collagen, osteocalcin and osteopontin protein synthesis, while osteoclasts secrete collagenase and are the major source of matrix metalloproteinase 9 (MMP-9) (which could be also involved in bone resorption) (Bruni-Cardoso, Johnson, Vessella, Peterson and Lynch, 2010; Schröder, K, 2019). In osteoporosis, the balance between bone formation and resorption, by osteoblasts and osteoclasts respectively is abolished (Katsumura *et al.*, 2016).

Both POLDIP2 and fibroblast growth factor 2 (FGF2) (responsible for morphogenesis, metabolism, fracture healing and bone growth mineralization, are expressed in MC3T3-E1 osteoblast cells (Katsumura *et al.*, 2016; Coffin, Homer-Bouthiette and Hurley, 2018). Katsumura *et al.*, (2016), found that FGF2 silences POLDIP2 expression in osteoblasts. POLDIP2 small interfering RNA (siRNA) silencing significantly reduced osteoblastic MC3T3-E1 cell migration in *in vitro* scratch wound healing assays (Katsumura *et al.*, 2016). These data suggest that POLDIP2 might be involved in the migration of osteoblastic cells, potentially leading to osteoporosis (Katsumura *et al.*, 2016).

The ROS can play a role in osteoporosis pathogenesis (Manolagas, 2010). Nox4, a constitutively active enzyme responsible for ROS production (particularly of H₂O₂) also contributes to the development of osteoporosis (Goettsch *et al.*, 2013; Schröder, 2019). It was observed that increased Nox4 expression level in human patient showed to exhibit increased osteoclast activity leading to higher bone resorption and reduced bone density (Goettsch *et al.*, 2013). As POLDIP2

directly interacts with the p22phox subunit of Nox4, its role in the aforementioned process cannot be neglected (Lyle *et al.*, 2009).

1.10. POLDIP2 and cellular homeostasis

It has been proposed that POLDIP2 could have a key role during hypoxia and cancer cell metabolic adaptation, as lipoylation of the pyruvate dehydrogenase (PDH) and α -ketoglutarate dehydrogenase (α KDH) complexes may be regulated by POLDIP2 in the mitochondria (Paredes *et al.*, 2018). POLDIP2 protein levels decrease under hypoxia or cancer, leading to the inhibition of lipoylation of the pyruvate and α -KDH complexes, as well as, during mitochondrial dysfunction (Paredes, Williams and San-Martin, 2017; Paredes *et al.*, 2018). Therefore, the POLDIP2 could be involved in the regulation of the caseinolytic peptidase (Clp)-protease complex, as well as, in the degradation of the lipoate-activating enzyme Ac-CoA synthetase medium-chain family member 1 (ACSM1) (Paredes *et al.*, 2018; Golias, Kery, Radenkovic and Papandreou, 2019). This, in turn, leads to decreased lipoylation and α KDH levels resulting in the PDH inhibition (Paredes *et al.*, 2018; Golias *et al.*, 2019).

Jiang *et al.* (2019), has also found that POLDIP2 deficiency in type 2 diabetes mellitus (T2DM) dampens lipid and glucose homeostasis. POLDIP2 expression levels decreased in T2DM humans, while Nox4 level was not significantly changed (Jiang *et al.*, 2019). Under normoglycemic conditions, insulin and leptin do not effect POLDIP2 expression (Jiang *et al.*, 2019). Hence, POLDIP2 appears to also play a central role in the control of aerobic metabolism.

1.11. POLDIP2 and vascular function

POLDIP2 maintains vascular structure, as in POLDIP2^{+/-} mice, the disorganisation of aortic structures is observed (Sutliff *et al.*, 2013; Amanso *et al.*, 2014). POLDIP2^{+/-} mice have aortas with an increased collagen deposition, as well as, an excess matrix and disordered and broken elastic laminae (Sutliff *et al.*, 2013). Such aortic changes become the risk factors for the development of cardiovascular diseases, as they lead to the loss of vascular smooth muscle (arterial) contractility and can cause arterial stiffness (Laurent, Alivon, Beaussier and Boutouyrie, 2012).

Compared to POLDIP2^{+/+} mice, the aorta extracted from POLDIP2^{+/-} mice had stiffer arteries, reduced elasticity and reduced ability to generate arterial force (Sutliff *et al.*, 2013; Griendling, 2018). Moreover, the abnormally increased collagen I and fibronectin secretion levels were detected in heterozygous POLDIP2 mice in the vascular smooth muscle cells (VSMC), as well as, in mouse aortic smooth muscle cells (MASMC) (Sutliff *et al.*, 2013; Fujii, Amanso, Abrahão, Lassègue and Griendling, 2016; Griendling, 2018). The increase in fibronectin and collagen secretion levels were also observed in homozygous POLDIP2^{+/+} mice, treated with siRNA against POLDIP2 (Sutliff *et al.*, 2013). Interestingly, collagen I secretion can be restored upon the addition of H₂O₂, suggesting that this reaction might be mediated by ROS (Sutliff *et al.*, 2013). Matsushima *et al.*, (2016), proposed that in VSMCs, POLDIP2 negatively regulates focal adhesion turnover and cytoskeletal remodelling through the redox-dependent activation of PI3K/Akt/mTOR signalling pathway. Also, Matsushima *et al.*, (2016), propose that POLDIP2 downregulation could lead to the collagen accumulation in VSMC through the activation of PI3K/Akt/mTOR signalling. Sutliff *et al.*, (2013), showed that in VSMC cells POLDIP2 increases the activity of NADPH oxidase 4 (Nox4), which is involved in the reactive oxygen species (ROS) production, POLDIP2 *in vivo* knockdown studies using siRNA against POLDIP2 revealed that decreased POLDIP2 level

diminishes H₂O₂ level. Such results suggest that in VSMC the moderate H₂O₂ production protects aortic dilatation by an increase of collagen and fibronectin secretion which makes vascular wall more elastic, thus increasing the ability to generate arterial force (Sutliff *et al.*, 2013; Amantea *et al.*, 2015; Cook-Mills, Marchese and Abdala-Valencia, 2001). It is interesting and important to note that decrease in Nox4, as well as, POLDIP2 are observed during human aortic valve stenosis, suggesting that these two proteins might be involved in myocardium remodelling (Moreno *et al.*, 2013).

1.12. POLDIP2 and cell cycle regulation

POLDIP2 could be involved in the cell cycle control via regulation of p21 (Datla *et al.*, 2019). p21 is involved in cell cycle regulation (Waga, Hannon, Beach & Stillman, 1994; Datla *et al.*, 2019). Cyclin-dependent kinase (CDK) complexes, positively regulate cell cycle progress (Ekholm and Reed, 2000). CDK complexes together with p21 (inhibitors) negatively regulate cell cycle (Ekholm and Reed, 2000). PCNA expression was significantly reduced and p21 (cell cycle inhibitor) expression was increased causing the delay in cell cycle progression in POLDIP2^{+/-} mice in comparison to POLDIP2^{+/+} (Datla *et al.*, 2019).

Flow cytometry showed that POLDIP2 protein is arrested/delayed in G2/M and G1 phases (Hernandes *et al.*, 2017). Such POLDIP2 delaying in mouse embryonic fibroblasts (MEFs), results in the decrease of cell number during the S phase (Hernandes *et al.*, 2017). Also, the POLDIP2 deficiency increases p53 phosphorylation level during G1 and G2/M phases (Hernandes *et al.*, 2017).

1.13. Project aims

POLDIP2 appears to be involved in the cell cycle, proliferation, cancer, bone homeostasis, DNA replication/repair, neurodegenerative and cardiovascular system control/function (Hernandes *et al.*, 2017). Although, POLDIP2 is a multifunctional protein it remains unknown how POLDIP2 can bind so many different proteins in different subcellular compartments and to be involved in different protein-protein interactions. Moreover, it remains unknown if two POLDIP2 forms (37 kDa and 42 kDa) have different functions. Also, POLDIP2 biochemistry remains completely untouched, as POLDIP2 3D structure alone or with binding partners has not been yet discovered. In order to understand POLDIP2 function, POLDIP2 3D structure is required. Ideally, POLDIP2 structure with a binding partner is required for the detailed understanding of POLDIP2 function. Therefore, several objectives were addressed in this research. Objectives: A) To sub-clone already existing POLDIP2 constructs into pGTVL2 vector containing GST-tag, as well as, to express all constructs in TB/LB media to optimize POLDIP2 expression levels. B) Purify POLDIP2 constructs which will have good expression level using IMAC/IEC/SEC and to optimize POLDIP2 constructs purification. C) Subject POLDIP2 into crystallization trials and attempt to reveal the protein's 3D structure. D) Attempt to solve the POLDIP2 structure in combination with one of its known binding partners (PCNA) using cryo-EM and X-ray crystallography, as well as, to determine PCNA-POLDIP2 complex stoichiometry.

2. Materials and methods

2.1. Materials

All chemicals were from Fisher Scientific (Ipswich, MA, USA). All enzymes were from New England Biolabs (Ipswich, MA, USA) and all protein purification columns were from GE Healthcare.

2.2. Molecular Biology

2.2.1. Polymerase chain reaction (PCR)

Touch-down PCR (BIORAD, California, United States) was used to amplify full-length POLDIP2, as well as, its 11 truncated forms (Appendix 1-3) (Korbie and Mattick, 2008). A cloned proliferating cell nuclear antigen (PCNA) construct was used as a positive PCR control. The PCR mix for 1 reaction was the following: 5x Phusion HF buffer (Ipswich, MA, USA), 2 U Phusion DNA Polymerase, 0.2 mM dNTP mix (0.2 mM for each dNTP), nuclease-free H₂O, to a final volume of 25 μ L. After the master mix was aliquoted, the following reagents were added separately into each tube: 1 μ L of 2.5 ng/ μ L of template DNA and 2.5 μ M primers (2.5 μ M of each forward and reverse, synthesised by Eurofins, Luxembourg) (see Appendices 1-3, Table 1). PCRs were analysed on an agarose gel (section 2.2.2).

Table 1. Touchdown PCR conditions. Information on times and temperatures is provided for each PCR cycle.

| | |
|-----------------------------|---|
| Initial Denaturation | 95°C, 10 min |
| | (Denaturation : 95°C, 30 sec; Primer annealing : 68°C, 30 sec; Extension : 72°C, 80 sec*) x5 cycles (Denaturation : 95°C, 30 sec; Primer annealing : 60°C, 30 sec; Extension : 72°C, 80 sec*) x5 cycles (Denaturation : 95°C, 30 sec; Primer annealing : 55°C, 30 sec; Extension : 72°C, 80 sec*) x5 cycles (Denaturation : 95°C, 30 sec; Primer annealing : 50°C, 30 sec; Extension : 72°C, 80 sec*) x10 cycles |
| Final Extension | 68°C, 10 min |
| Hold | 15°C |
| Note: * | 1 min per 1 kb fragment. The longest extension time was used if multiple fragments with a varying length were amplified. |

2.2.2. DNA Agarose gel electrophoresis

50 mL of 1.5 % agarose gel (w/v) made in 1x TBE buffer was used for the estimation of the correct DNA fragment amplification. Agarose was mixed with 1x TBE buffer and was microwaved until an agarose melted. 2 µL (4% v/v) of SYBR-safe (Invitrogen, California, United States) per 50 mL gel was added after cooling and poured unto a gel casting tray with combs.

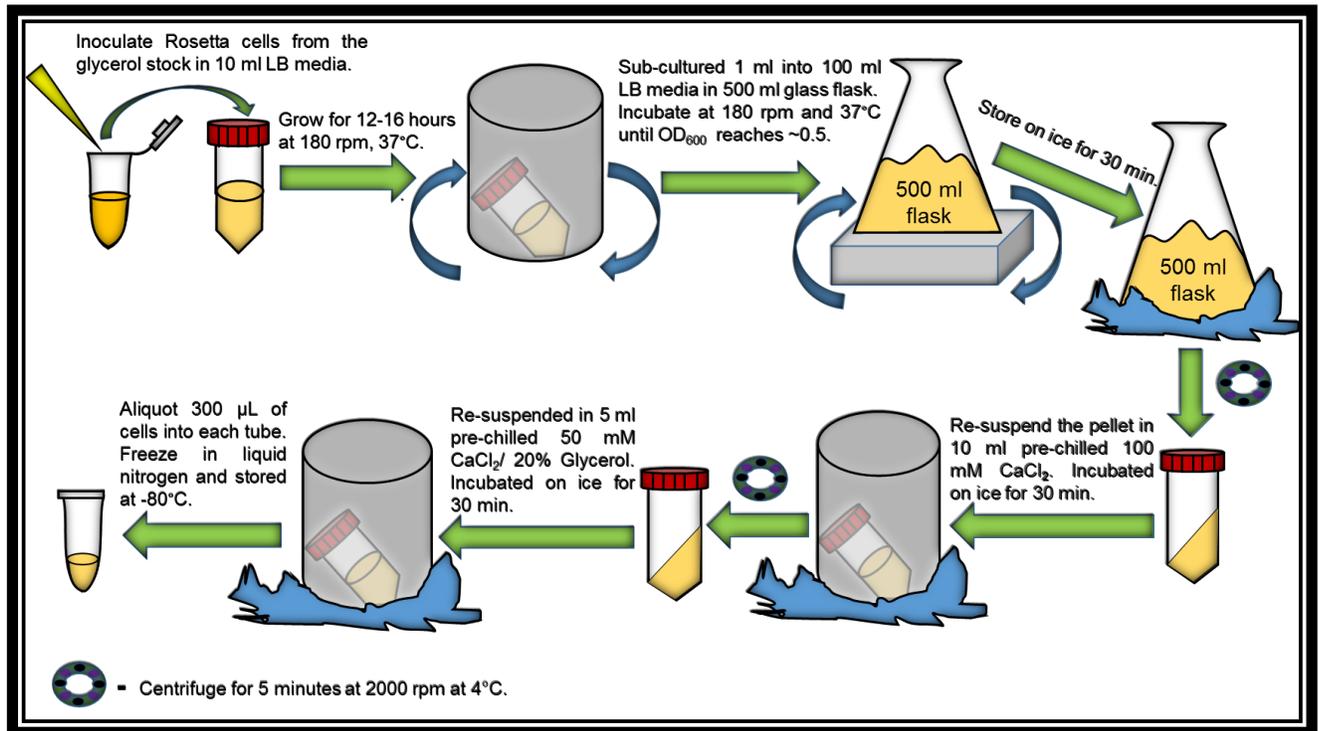
Samples were mixed with 6x DNA loading dye (ThermoFisher, Massachusetts, United States) and loaded, alongside a 2-LOG DNA ladder (NEB, Massachusetts, United States). Gels were resolved for 60 min at 80 V. The DNA was visualized/analysed under UV transilluminator and image capture software (SynGene).

2.2.3. Competent *E. coli* cells

Rosetta2 cells (*E. coli* BL21 (DE3) containing pRARE2) cells (Appendix 4) required for the protein expression and purification purposes were grown from the glycerol stock/colony in 10 mL LB media (without antibiotics) overnight at 180 rpm, 37°C (IOX400 XX1C, Gallenkamp,

Loughborough, UK) (Figure 5). 1 mL from the overnight culture was sub-cultured into 100 mL LB media in 500 mL flask. Then the OD₆₀₀ has reached ~0.5 cells were placed on ice for 30 min. After incubation Rosetta2 cells were centrifuged for 5 minutes at 2000 rpm (Mistral 3000i, MSE, East Sussex, UK). The supernatant was discarded and the pellet was gently re-suspended in ice cold, sterile 100 mM CaCl₂. Sample was incubated on ice for 30 min, centrifuged and re-suspended in 5 mL pre-chilled 50 mM CaCl₂/ 20% (v/v) glycerol as described above. Rosetta2 cells were incubated on ice for additional 30 min and 300 µL of cells were aliquoted in pre-chilled PCR tubes. Competent Rosetta2 cells were flash frozen in liquid nitrogen and stored at -80°C.

Figure 5. Flow chart of competent *E. coli* preparation. All steps, starting from the overnight *E. coli* growth and finishing with *E. coli* storage are described.



2.2.4. *E. coli* culture preparation

Overnight cultures were prepared from the *E. coli* glycerol stocks/colonies. MACH1/Rosetta2 *E. coli* cell were grown in 5 mL LB media with appropriate antibiotics (Kan for the MACH1; Kan/Chl for the Rosetta2) in 50 mL falcon tubes at 37°C, 180 rpm overnight for 12-16 hours (Appendix 4). Kan/Chl concentration were made respectively at 50 µg/mL and 34 µg/mL and they were added at 1:1000 antibiotic:media ratio.

2.2.5. Transformation

MACH1/Rosetta2 cells were thawed on ice. 2 µL (50-150 ng) of DNA (pGTVL2 constructs; PCNA was used as a positive control) were added to 25-40 µL cells gently pipetting once to mix the sample and incubated on ice for 1 hour. *E. coli*-DNA mix was heat-shocked in water bath for 45 seconds at 42°C. After, heat-shock cells were immediately placed on ice for 5 minutes. 100 µL of pre-warmed at 37°C LB media (without antibiotics) was added under aseptic conditions. Cells were incubated for 60 minutes at 180 rpm at 37°C. 60-100 µL culture was gently spread on pre-warmed (37°C) LB-Kan/5% Sucrose (w/v) or LB-Kan/Chl agar plates, respectively if MACH1 or Rosetta2 cells were used for the transformation. Plates were incubated overnight for 12-16 hours in the incubator (MIR-262 Sanyo, Panasonic, Osaka, Japan) at 37°C.

2.2.6. *E. coli* glycerol stock preparation

1 mL of overnight *E. coli* culture grown in 5 mL LB/antibiotics was mixed gently by pipetting up and down with 0.333 mL of 60% (v/v) filter sterilised glycerol. If samples were prepared after

transformation they were stored in eppendorf tubes in -20°C, until the colony PCR screen showed positive results. Then, the cells were transferred into cryogenic tubes and stored in -80°C for the overnight culture preparation.

2.2.7. Ligation independent cloning (LIC)

2.2.7.1. pGTVL2 vector restriction digest

5 µg of vector (pGTVL2), 10x NEBuffer™ 3.1 and 30 U *BsaI* enzyme were mixed in order to linearise the vector. Nuclease free water was added to make the total reaction volume 100 µL. The mix was incubated at 50°C in a heating block overnight. If restriction digest was not successful in the first instance the reaction was repeated prolonging the incubation time up to 3 days. The *BsaI* enzyme was inactivated by incubation at 65°C for 20 minutes. Plasmid digestion was checked on an agarose gel.

2.2.7.2. *DpnI* treatment of PCR products.

The PCR products were *DpnI* treated in order to remove the original DNA template. 35 µL of non-purified PCR DNA mix, 10x NEB 2.1 buffer, 5 U of *DpnI* and the nuclease free water was combined to a total volume of 50 µL and incubated for 1 hour at 37°C. Following the incubation, samples were heated for 20 minutes at 80°C for the enzyme inactivation. The PCR products were purified using QIAquick PCR purification kit (QIAGEN, Hilden, Germany), as was described in the manufacturers protocol (see Figure 6).

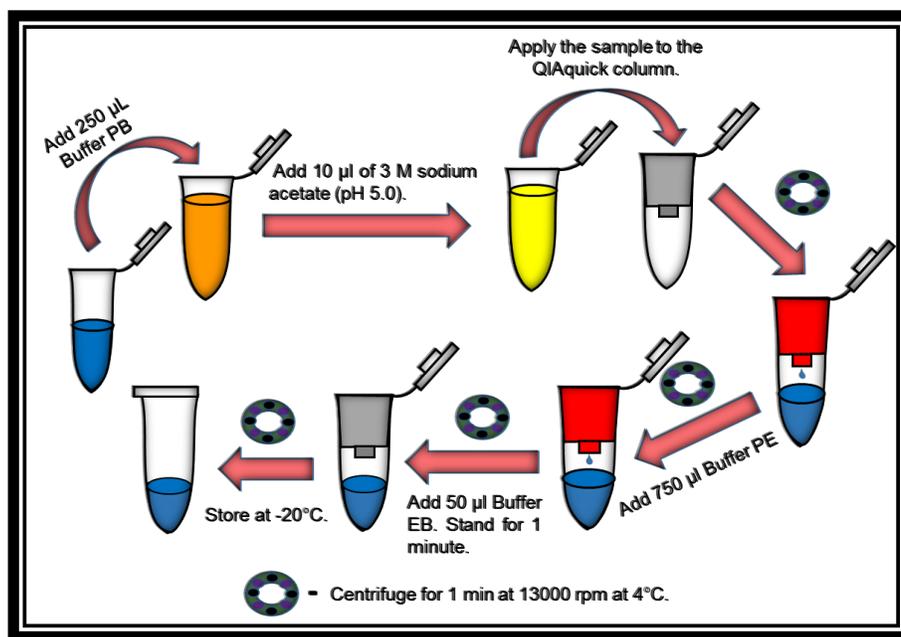


Figure 6. Flow chart of the DNA purification. All DNA purification steps are shown.

2.2.7.3. T4 DNA polymerase treatment of pGTVL2 vector and PCR fragment

For vector preparation, 50 µL of *Bsa*I digested plasmid (50-60 ng/µL), 7.5 U T4 DNA polymerase, 100 mM DTT, 10x NEB buffer TM 2.1, and 25 mM dGTP were combined and water was adjusted to make total volume of 100 µl. For PCR T4 treatment, 5ul of PCR product was combined with the same components (substituting dGTP for dCTP), to a total volume of 10 µL. The mix was incubated at room temperature for 30 minutes. Enzyme was inactivated following the incubation at 75°C for 20 minutes in a bench heating block. For the annealing, 1 µL of treated plasmid was mixed with 2 µL of the treated PCR insert and incubated for 30-40 min at room temperature. The transformation was performed as described in section 2.2.5. and *E. coli* cells were spread on the LB-Kan/5% (w/v) D-Sucrose plates. If the transformation was not successful for the first time, it

was repeated again using 1:2 and 1:4 volume ratios (vector:PCR product). Next day, the PCR colony screen was performed.

2.2.7.4. Whole cell PCR colony screen

Colonies were screened by PCR to check for the correct plasmid construct. Initially, one colony was screened for each construct. However, for those with apparently missing constructs (for example due to contamination), several other colonies were screened to detect the presence of plasmid. 5-7 ng of cloned PCNA was used as a positive PCR control. The reaction mix comprised 10 μ M pLICfor/pLICrev primers (synthesised by Eurofins, Luxembourg; Appendix 2), 5x Red buffer, 3% DMSO (v/v), 3% glycerol (v/v), 1 U MyTaq Red DNA polymerase (Bioline, UK), to a total of 20 μ L. 0.5 μ L of overnight culture grown from a single plate colony or alternatively 10 ng of plasmid DNA were added separately into each tube. The PCR conditions used for the colony screen are shown in Table 2 and the thermocycler was pre-heated to 95°C prior to inserting the samples. 10 μ L of PCR reaction was run on an agarose gel, as described in section 2.2.2, but without the 6x loading dye. Colonies which contained the plasmid with the expected size were grown overnight, following glycerol stock preparation (section 2.2.6; 1 mL used) and plasmid isolation (section 2.2. 8; 4 mL used).

Table 2. PCR running conditions. Time and temperature are shown for each PCR cycle.

| | |
|-----------------------------|--|
| Initial Denaturation | 95°C, 10 min |
| | (Denaturation: 95°C, 30 sec; Primer annealing: 50°C, 30 sec; Extension: 72°C, 130 sec*) x30 cycles |
| Final Extension | 72°C, 5 min |
| Hold | 15°C |
| Note: * | 1 min per 1 kb fragment. The longest extension time was used if multiple fragments with a varying length were amplified. |

2.2.8. Plasmid miniprep

4 mL of overnight culture were centrifuged for 20 min at 2400 rpm at 4°C. The pellet was re-suspended in 250 µL of P1 buffer, transferred into small eppendorf tubes and stored at -20°C until the PCR colony screen showed positive results. Next, the plasmids were isolated using QIAquick PCR purification kit (QIAGEN, Hilden, Germany) according to the manufacturers protocol (Figure 7). Plasmids were prepared and sent for the sequencing (by Dr Christopher Cooper).

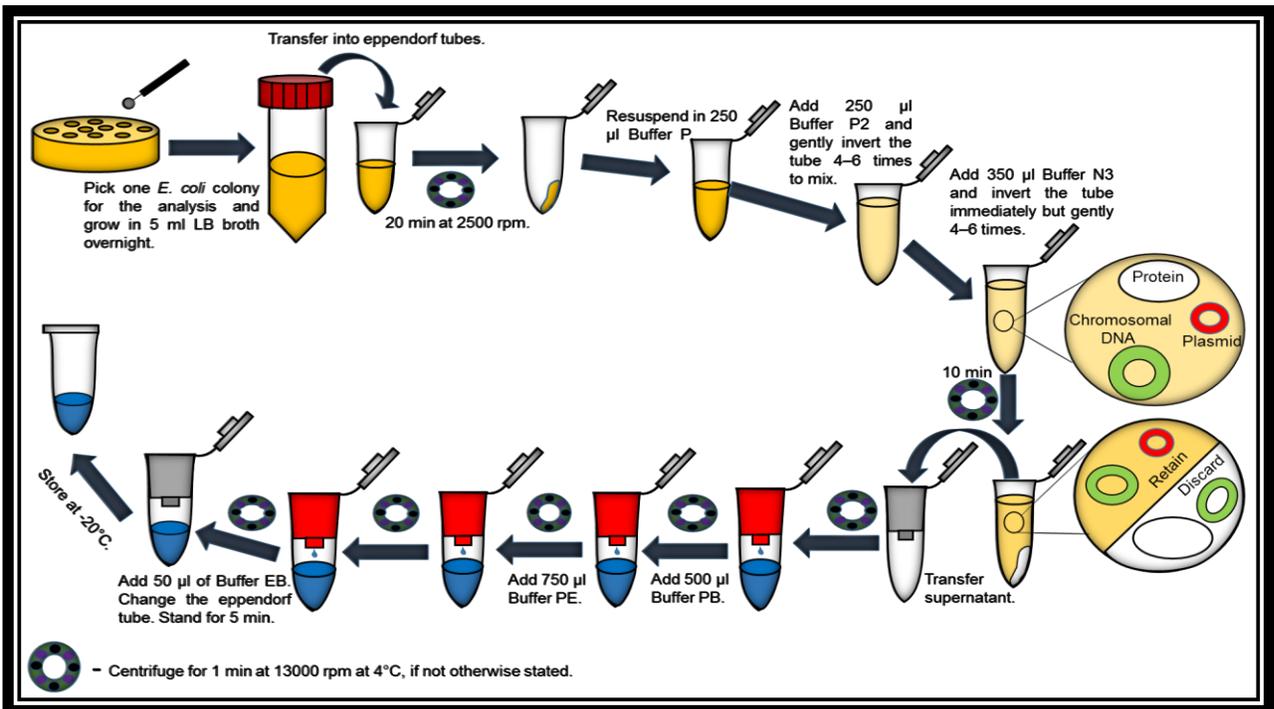


Figure 7. Plasmid DNA isolation. All plasmid DNA isolation and purification steps described in the QIA prep Spin Miniprep Kit manufacturer's protocol are shown.

2.2.9. DNA quantitation

The DNA concentration and purity was measured based on the DNA absorbance at 260 nm and A260/A280 nm absorbance ratios (Sambrook and Russell, 2001). A260/A280 absorbance ratio of

1.8-2 served as an indication of DNA purity (Sambrook and Russell, 2001). TE buffer was used as a blank. 2 μ L of purified DNA were used for the NanoDrop.

2.3. Overexpression and purification of recombinant protein

2.3.1. Small scale recombinant protein expression

Rosetta2 cells transformed with plasmid expression constructs were grown overnight (section 2.2.6). In the morning, 500 μ L of the overnight culture was inoculated into 50 mL LB-Kan or TB-Kan media in 250 mL plastic baffled flask (Ultra Yield Generon). Rosetta2 cells were incubated for 3-4 hours at 37°C, 180 rpm until OD₆₀₀ reached ~0.6-0.8 or 1.5-2, for LB and TB media, respectively. When the cells have reached the right OD they were cooled for 15-30 min at 18°C, with the occasional shaking. Protein expression was induced with 50 μ L of 100 mM IPTG and cells were incubated at 18 °C, 180 rpm for 12-16 hours. Cultures were transferred to 50 mL Falcon tubes and harvested by centrifugation for 30 min at 4000 rpm, 4°C. Harvested pellets were stored at -80°C.

2.3.2. Small scale *E. coli* cell lysis

Cells were thawed in a 37°C water bath for 2 minutes. From this point, all samples and solutions were kept on ice. Bacterial pellets were re-suspended in 3 mL lysis buffer 1 (LB1) (Appendix 5) by vortexing until the pellet was completely resuspended. Cells were sonicated (Vibra-Cell, VCX 130 PB, Connecticut, United States) on ice for 2-3 minutes (20 seconds on, 20 seconds off; amplitude 60%) respectively, for LB and TB media. The re-suspended cells were centrifuged at 13500 rpm, 4°C for 30 min.

2.3.3. Ni-NTA affinity chromatography small scale

3 mL of supernatant was transferred to a 15 mL falcon tube and 1 volume of lysis buffer 2 (LB2) was added, as well as, 200 μ L of 50% Ni-NTA beads (v/v). The sample was incubated for 1 hour at 4°C and 100 rpm. The lysate was applied into a 30 mL Econo-Pac column and allowed to drip through. The sample was washed 5 times with 2 mL wash buffer (WB) (Appendix 5). 200 μ L of elution buffer (EB) were added to each column. The eluted protein was stored at -20°C.

Small scale Ni-NTA beads preparation: 1 volume of beads (200 μ L) was washed with 1 volume of Milli-Q-H₂O. Next, Ni-NTA beads were equilibrated with 1 volume of lysis buffer 0 (LB0).

2.3.4. Large scale recombinant protein expression

10 mL of overnight culture were poured to 1L LB/TB-Kan media (no Chl for the large scale expression) in a 5 L flask (Figure 8). The flask was shaken at 180 rpm at 37°C for approximately 3-4 hours until the OD₆₀₀ had reached 1.5-2 (culture and blank were diluted 1/5, so the expected OD₆₀₀ reading was 0.3-0.4). The LB/TB media were used as a reference. When the media reached the right OD₆₀₀ flasks were put on the bench to cool down for 45-60 minutes and a short shake was given every 10 minutes while the flasks were cooling. Overnight expression (12-16 hours) was induced with 0.1 mM IPTG (final concentration) at 18°C and 180 rpm. Next day the samples were centrifuged for 30 min at 4°C and 7,800 g in JLA 10.50 rotor (Beckman Coulter Avanti^RJ-26XPI). The supernatant was discarded and the cells were re-suspended in 1x PBS buffer. Cells were transferred into 50 mL tubes and re-centrifuged at 4000 rpm for 1 hour at 4°C and the pellet stored at -80°C.

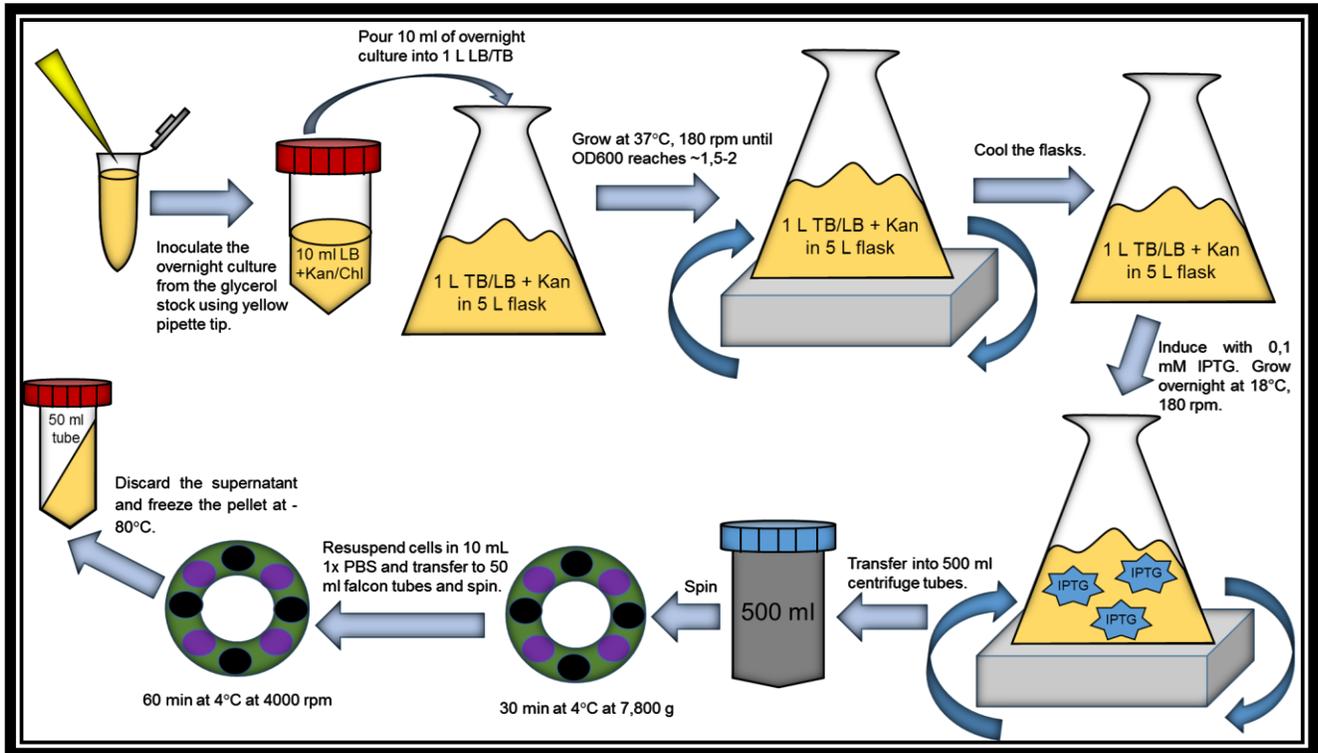


Figure 8. Flow diagram of large scale protein expression. All steps starting from *E. coli* inoculation and ending up with the *E. coli* pellet storage are shown.

2.3.5. Large scale *E. coli* cell lysis

Frozen cells were thawed in water bath for 20 min at 36-39°C. From this step all buffers and cells were kept on the ice at all times. 50 mL LBa buffer (Appendix 6) was used per litre of cell culture to re-suspend the pellet. Cells were fully re-suspended by vortexing, following the incubation at 4°C, 100 rpm for 1 hour or longer, until sample viscosity was reduced. Next, cells were disrupted by pulsed sonication (80% amplitude; 10 minutes; 10 seconds on and 10 seconds off). Cells were sonicated on ice for 5-6 minutes sonication process with mixing. Lysates were clarified by centrifugation for 1 hour at 4°C at 33.000 g in JA-25.50 rotor.

2.4. Large scale protein purification

2.4.1. Ni-NTA Affinity Chromatography (IMAC) and Tobacco Etch Virus (TEV)-Ni-NTA Affinity Chromatography.

2.4.1.1. Ni-NTA Affinity Chromatography (IMAC)

Ni-NTA affinity chromatography was used to purify 6x His-tagged recombinant proteins from *E. coli* (Crowe *et al.*, 1994). 2 mL bed volume of Ni-NTA agarose beads were applied to a 30 mL Econo-Pac column. The sample was allowed to flow through by gravity. 10 column volume (CV) of ddH₂O were passed through the column to remove the beads storage solution which contained 20% (v/v) ethanol. 10 CV of LB equilibration buffer was then applied to the column. Followed by clarified *E. coli* supernatant. Flow-through, 10 CV LB, 10 CV, W30 and 3 x 4 mL E300 elution fractions were collected. All samples were run on the SDS-PAGE gel. SDS-PAGE gel was analysed and protein dialysis was set (section 2.5.). Minimum of 1:20 TEV:protein mass ratio was used for the dialysis (gift from Structural Genomics Consortium (Oxford), purified by Dr Christopher Cooper. 6x His-tagged so it is removed during the reverse Ni-NTA). However, usually 600 µL of TEV protease (4-5 mg/mL) were added per 4 L of POLDIP 2 (51-368)/POLDIP2 (178-368) (E4/E7); 8 L of POLDIP2 (1-368) (E1) and 4 L of PCNA.

2.4.1.2. Reverse Ni-NTA Affinity Chromatography

30 mL Econo-Pac column was prepared as described above. Dialysed protein was applied to the column and allowed to drip through. The protein was passed through the column twice. 2 CV elutions of lysis buffer containing 20 mM, 40 mM, 100 mM and 300 mM imidazole pH 8 were collected. All samples were run and analysed on SDS-PAGE.

2.4.2. Ion exchange and heparin chromatography (IEC)

An ÄKTA Pure FPLC (GE Healthcare) was used for the IEC purification. 3 different columns were used for the purification: HiTrap-Q/HiTrap-SP/HiTrap-Heparin, depending on the protein and the experiment. Proteins isoelectric point (pI) was estimated using the Expasy, ProtParam program (<http://www.expasy.ch/tools/protparam.html>) (Gasteiger *et al.*, 2003). If protein pI was above 7.5 HiTrap-SP (- charge) column was used, and if pI was below 7.5 HiTrap-Q (+ charge) column was used (for pH 7.5 buffers). For the use of HiTrap Heparin column the pI determination was not necessary. The appropriate column was equilibrated with the 20 CV of the appropriate buffer (Appendix 7). The protein was concentrated (section 2.3.8) and diluted 5 fold (5% glycerol (v/v), 20 mM Tris/Hepes (pH 7.5) for Q/SP-IEC respectively) to bring the final NaCl concentration to 100 mM before loading onto the column. 2/5 mL concentrated protein was loaded into 2/5 mL washed capillary loop depending on volume of sample. The sample was then eluted in 100-1000 mM NaCl elution gradient with a flow rate of 1 mL/min at 9°C in 2 mL fractions. The column was stored in Milli-Q-H₂O or 20% ethanol (v/v) at 9°C until needed. Samples were analysed by SDS-PAGE.

2.4.3. Size Exclusion Chromatography (SEC)

2/5 mL concentrated protein was loaded onto equilibrated 2/5 mL capillary loop, respectively, following removal of aggregates by syringe filtering through 0.2 µm filters. Protein was run on 1.5 CV pre-equilibrated S200 16/600 column (20 mM HEPES pH 7.5, 500/100 mM NaCl, 5% glycerol (v/v) and 1mM DTT, added immediately before use). Protein was eluted at 1 mL/min at 9°C in 2 mL fractions. The column was stored in 20% ethanol (v/v) or Milli-Q-H₂O at 9°C until needed.

After, SEC SDS-PAGE gel was run, pooled fractions were concentrated. Usually, the fresh protein was used further, however sometimes the protein was flash frozen in the liquid nitrogen and stored at -80°C until was used.

2.5. Protein Dialysis

Protein was dialysed in pre-chilled dialysis buffer (Appendix 8). Protein:dialysis buffer ratio was 1:167. Protein dialysis was carried in snakeskin dialysis membrane (3500 Da MWCO) for 1-3 days in a cold cabined at 9°C with stirring. Before sealing the dialysis bag air bubbles were removed.

2.6. Protein Concentration

30 kDa or 10 kDa protein concentrators (Vivaspin® 20; 10,000 MWCO; PES Membrane (Generon)) were washed with Milli-Q-H₂O following the centrifugation for 5 minutes, at 4°C and 2500 rpm. The protein was applied to the concentration column and centrifuged at 2500 or 4000 rpm, respectively for 15 or 10 minutes at 4°C until the desired proteins concentration was reached.

2.7. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE)

2.7.1. Gel casting

Spacers, combs and glass slides were cleaned with 70% (v/v) ethanol and dried. 10% resolving gel (v/v) and 4% staking gel (v/v) were used for the protein visualization unless stated otherwise. Reagents were added in the order provided in the Table 3. After 15-25 min the clear band was

formed between the gel and water layer. The water was completely removed using filter paper. 4% stacking gel (v/v) was poured to the very top of the edges and the well-forming combs were inserted. The gels were stored at 4°C until they used.

Table 3. The recipe for the resolving and stacking gels is shown. SDS-PAGE gel components.

| SDS-PAGE gel recipe/ gel percentage | 10% resolving gel | 12% gel resolving | 4% stacking gel |
|-------------------------------------|-------------------|-------------------|-----------------|
| Mill-Q water | 19.4 ml | 17.4 ml | 9.5 ml |
| 1.5 M Tris pH 8.8 | 10 ml | 10 ml | — |
| 0.5 M Tris pH 6.8 | — | — | 3.78ml |
| 10% SDS | 400 µl | 400 µl | 150 µl |
| 40% Acrylamide/bis | 10 ml | 12 ml | 1.5 ml |
| TEMED | 20 µl | 20 µl | 15 µl |
| 10% APS (made fresh every 7 days) | 200 µl | 200 µl | 75 µl |

2.7.2. SDS-PAGE protein sample preparation

4x SDS loading dye (4x LDS sample buffer, 0.2 M DTT) was mixed with protein sample. 5-20 µL of sample were loaded for all other gels, unless stated otherwise. Precision plus protein dual color standard marker (Bio-Rad) was used for molecular weight comparisons.

2.7.3. Gel running conditions and staining

Gels were resolved for ~100 min at 100 V in 1x TGS buffer, until the dye has reached the bottom of the gel. The gel was stained with InstantBlue™ protein stain overnight at room temperature with gentle agitation. In the morning the gel was de-stained with dionised water.

2.8. Protein spectrophotometry

ExPASy ProtParam program (<http://www.expasy.ch/tools/protparam.html>) was used to calculate E1, E4 and E7 (POLDIP2 constructs) molecular weights (MW) and Extinction coefficients (EC), see Table 4. The average EC and MW were used to estimate E1, E4 and E7 concentration at 280 nm wavelength. The PCNA concentration was determined at 280 nm assuming that 1 Abs = 1 mg/ml.

Table 4. POLDIP2 constructs EC and MW. ProtParam predicted values are shown for E1, E4 and E7 POLDIP2 constructs.

| POLDIP construct name: | Extinction coefficient (if all Cys residues form cystines) $M^{-1} cm^{-1}$ | Extinction coefficient (if all Cys residues are reduced) $M^{-1} cm^{-1}$ | Average Extinction coefficient value used ($M^{-1} cm^{-1}$) | Molecular weight (Da) |
|------------------------|--|---|--|-----------------------|
| E1 | 80120 | 79870 | 79995 | 42033.28 |
| E4 | 63495 | 63370 | 63433 | 36750.22 |
| E7 | 63495 | 63370 | 63433 | 35753.14 |

2.9. Protein characterisation

2.9.1. Mass spectrometry (MS)

Tandem mass spectrometry (MS/MS): bands were cut from the gel and stored in 200 μ L of 10% methanol (v/v). Intact mass spectrometry (MS): protein concentration was 0.02 mg/mL and the sample was diluted in 0.1% formic acid (v/v) making the total volume of 200 μ L. Both MS/MS and intact MS samples were stored at 4°C until analysis (performed by Dr Rod Chalk, University of Oxford).

2.9.2. Analytical Size Exclusion Chromatography (SEC)

Analytical SEC was performed as described in section 2.4.3. but was run against Bio-Rad SEC standards (1350 MW, 17000 MW, 44000 MW, 157000 MW, 670000 MW). Also, for the PCNA-POLDIP2 complex salt concentration was lowered to 100 mM NaCl (samples were co-incubated for 3 hours before SEC purification).

2.9.3. *In vitro* protein crosslinking

For crosslinking 1 μg of protein was used (1 μg of each for the complex) in a total. Final volume was 10 μL . Protein was diluted in SEC buffer (varying NaCl concentration for a final 140 mM NaCl concentration). 2 crosslinkers were used BS3 and glutaraldehyde. From the 50 mM/1% BS3/Glutaraldehyde stock the following final concentrations were made: 10, 5, 1, 0.5, 0.1, 0.05, 0 mM (BS3 (w/v)) and 0.2%, 0.1%, 0.04%, 0.02%, 0.01%, 0.002%, 0% (Glutaraldehyde (v/v)). 2.5 μL of BS3/Glutaraldehyde was added into a tube containing 10 μL of protein and mixed gently by pipetting up and down. The mixture was incubated on ice for 1 hour and then 2.5 μL of the quenching solution (250 mM Tris-HCl, pH 7.5) was added. The sample was incubated for 15 minutes on ice. 4 x SDS loading dye was added and the sample incubated for 2 min at 95°C. 20 μL was loaded on the SDS-PAGE gel.

2.10. Structural Characterisation

2.10.1. Circular Dichroism (CD)

The protein was dialysed overnight in 8 kDa membrane Mini GeBaFlex tube Dialysis Kit, in a non-chiral 50 mM KHPO₄ buffer (pH 7.5) overnight at 10°C. Protein was spun down for 5 min at 13300 rpm at 4°C to remove aggregates and quantitated. CD was performed with the kind assistance of G. Nasir Khan (University of Leeds). CD Chirascan Plus (Applied Photophysics) spectropolarimeter was used with the grant code 094232-Wellcome Trust. 50 mM KHPO₄ buffer was used as a blank. 200 µL of 0.450 mg/mL was used. Protein was scanned 2 times at 5°C. Far-UV Protein scans were conducted in 180-260 nm range with scan-time of 2 minutes. The AutpPM millidegree was used, with 2 nm band-width, 1 second response time, 1 nm resolution and 0.1 cm path length. 2 reads for each wavelength were averaged following the buffer blank spectra subtraction. The noisy data below 190 nm was not taken for the data analysis. The obtained data was analysed using DICHROWEB software suite using reference set 4 and CONTIN program (Whitmore and Wallace, 2004).

2.10.2. Electron microscopy (EM)

Negative stain EM was performed with the kind assistance of Dr Emma Hesketh (University of Leeds). PCNA-POLDIP2 complex was run in the SEC buffer (section 2.9.2.), except that 100 mM NaCl was used. Protein complex was spun down for 3 min at 14000 rpm, 20°C. Few drops of 5.15 mg/mL were used. The grid was glow discharged for 50 seconds at 15mAmp and 0.38 mBar using a PELCO easiGlow, before being washed with 2-3 droplets of water twice and then treated twice

with 2-3 droplets of 1% (w/v) uranyl acetate. The excess liquid was removed from the grid using a filter paper. After the second addition of 1% (w/v) uranyl acetate the sample was allowed to dry at room temperature for 5 minutes. FEI T12 microscope with 120 keV Lab6 electron source was used for the carbon-coated, 400 mesh Cu grid visualization. Micro-graphs were recorded on Gatan US4000/SP 4k x 4k CCD camera.

2.10.3. Crystallography

2.10.3.1. Vapour diffusion crystallography

96 well MRC crystal plates were set manually using the sitting drop vapor diffusion method at varying proteins concentration and incubated at 4° or 20°C (Appendix 9). 4 screens were used: Morpheus® HT-96, JCSG-*plus*TM, PACT *premier*TM and The BCS Screen HT-96 (Molecular Dimensions). 50 µL of mother liquor was pipetted into the reservoir. 1 µL mother liquor was mixed with 1 µL of protein. The plates were sealed, incubated and checked after 1,2, 3, 5, 14 and then each 30 days.

2.10.3.2. Screen design and optimization

For the optimization, a grid screen based on the mother liquor components with a finer screen was prepared in house. The same sitting drop vapor diffusion method was used and the detailed method description is given in section 2.10.3.1.and Appendix 9.

2.10.3.3. Crystal mounting and screening

Protein crystals were extracted from the crystallization drop under stereomicroscope. (Trinocular Stereomicroscope with Transmitted and Incident Light Base (Molecular Dimensions)). Single crystals were usually picked, however clumps of crystals were broken and screened as well. Cryo-protectant was not typically added, as all conditions screened contained sufficient cryoprotectant. Crystals were flash frozen in the liquid nitrogen, placed on the goniometer (D8 VENTURE experimental setup) and centered by spinning on the Phi axis. A copper X-ray radiation source (D8 Venture Bruker) was used, with a wavelength of 1.5 Å. The following screening parameters were applied: 2theta=0; omega=0; Phi=0 or 90° (always checked at two angles); image width=0.5; exposure time=10-20 seconds; detector distance=100 mm; temp=120K. If ice rings appeared or the diffraction pattern showed high mosaicity (spots overlap/not resolved/streaking) the crystal was annealed by warming the flash frozen crystal at room temperature for 2-5 seconds, and then flash cooled again under the liquid nitrogen stream.

2.10.3.4. Protein structure refinement

Data Collection: a single dataset consisting of 360 images, each of 0.5° oscillation was collected resulting in a total rotation of 180°. Data reduction (integration and scaling) was performed using Proteum3 (Bruker) and default settings. The resultant unmerged reflection file in XDS format (.hkl) was then converted to CCP4 (.mtz) file format using POINTLESS (CCP4). SCALA was then used to merge the data and add a FreeR column for 5% of the data. Systematic absences were analysed to determine the space group. The SCALA log file suggested a P62 screw axis (probability 0.994). P62 and P64 are indistinguishable by systematic absences, and so the molecular replacement was attempted in both space groups.

As part of the SCALA run, the program TRUNCATE was used to convert intensities to structure factor amplitudes. MATTHEWS was then used to calculate the Matthews coefficient, which was 2.79 for 1 molecule per asymmetric unit. Models for molecular replacement were identified by a BLAST search of the PDB and by Fold and Function Assignment System (FFAS). PDB accession code 1VBV was used as a model for the N-domain. Two structures of the C-domain were identified as potential models, 5HDW and 2F1E, which differ slightly in loop regions and the N- and C-termini. Structures of 5HDW and 2F1E were superimposed in PyMOL and loop regions that were visibly different between the two structures were deleted from the model. All side-chains were then deleted from both models using PDBSET to generate a main-chain of poly-alanine ready for MR.

Molecular replacement was conducted in PHASER using these polyalanine models based on 5HDW (32% sequence identity) and 1VBV (15% sequence identity) prepared as described above (loops and side-chains deleted). Phaser was instructed to search for 1 copy of 5HDW then 1 copy of 1VBV.

Structure refinement then proceeded by alternating between manual real-space refinement in COOT and reciprocal space refinement in REFMAC. Polyalanine was built into areas of positive difference density in Fo-Fc maps. Density modification was conducted using DM before automatic chain tracing in BUCCANEER. This allowed 73% of the structure to be built (211 residues). The remainder of the structure was modeled manually by alternating between COOT and REFMAC. CCP4 Program Suite v7.0.078 including SHELX and COOT v0.8.9.2 was used and molecular replacement and refinement was conducted by Dr Richard Bingham.

2.10.3.5. Protein structure visualization in PyMOL

Protein 3D structure visualization and figure production was made using PyMOL (The PyMOL Molecular Graphics System, Version 2.3.1 Schrödinger, LLC).

3. Results-discussion

All stages of recombinant POLDIP2 structure determination pipeline were attempted. POLDIP2 constructs were sub-cloned into pGTVL2 vector containing a GST tag. POLDIP2 constructs expression levels were determined and samples with the acceptable expression levels were selected for the further large scale protein purification and crystallisation trials. Obtained POLDIP2 crystals were screened and the protein which had the best diffraction was used to solve POLDIP2 structure. POLDIP2 structure was nearly fully solved (~99.8%) at ~2.9 Å. Further, POLDIP2 *B* factor and surface charge distribution were analysed. POLDIP2 3D structure was related to predicted and known binding partners, revealing the secondary structure elements which are involved in protein-protein interaction. Moreover, attempts to crystallise PCNA-POLDIP2 complex were made. An attempt to understand PCNA-POLDIP2 stoichiometry was made using *in vitro* crosslinking and size exclusion chromatography. Also, an initial negative staining EM for PCNA-POLDIP2 complex, towards cryo-EM complex structure determination was attempted.

3.1. Ligation independent cloning (LIC)

3.1.1 The use of highly parallelized LIC for POLDIP2 construct cloning

Existing constructs were sub-cloned into a new expression vector (pGTVL2 containing an N-terminal TEV protease cleavable 6xHis-GST affinity tag), in order to test if protein expression level could be enhanced in comparison to previously cloned constructs in pNH-TrxT (6xHis-Thioredoxin tag, cloned by Klaudia Maruszczak). Not only the full length POLDIP2, but truncated protein fragments were cloned using the highly parallelised LIC approach, as protein fragmentation is very often the key to obtaining highly-expressing/crystallising proteins (Savitsky

et al., 2010). Truncated proteins (or domains) could be more stable and rigid, hence more likely to form crystal contacts (Dale, Oefner and D'Arcy, 2003). The schematic representation of the LIC process is shown to aid the subsequent process understanding described further in this thesis (Figure 9).

LIC cloning as a cloning method eliminates the need to check insert's insertion orientation, as well as preventing plasmid self-ligation due to the presence of long non-complementary overhangs (Aslanidis and Jong, 1990). Generation of longer cohesive ends allows the formation of more stable insert-vector complex through the formation of multiple hydrogen bonds, allowing the omission of the ligation step in comparison to the traditional sticky end cloning which utilizes short cohesive ends (Li and Evans, 1997). Non-covalently associated vector-insert molecules can be repaired and covalently joined upon the entry into the *E. coli* cell after successful survival of the transformation process, as linear DNA in *E. coli* will be degraded by heterotrimer RecBCD *E. coli* exonuclease (Aslanidis and Jong, 1990; Kuzminov and Stahl, 1997; Benoit *et al.*, 2016).

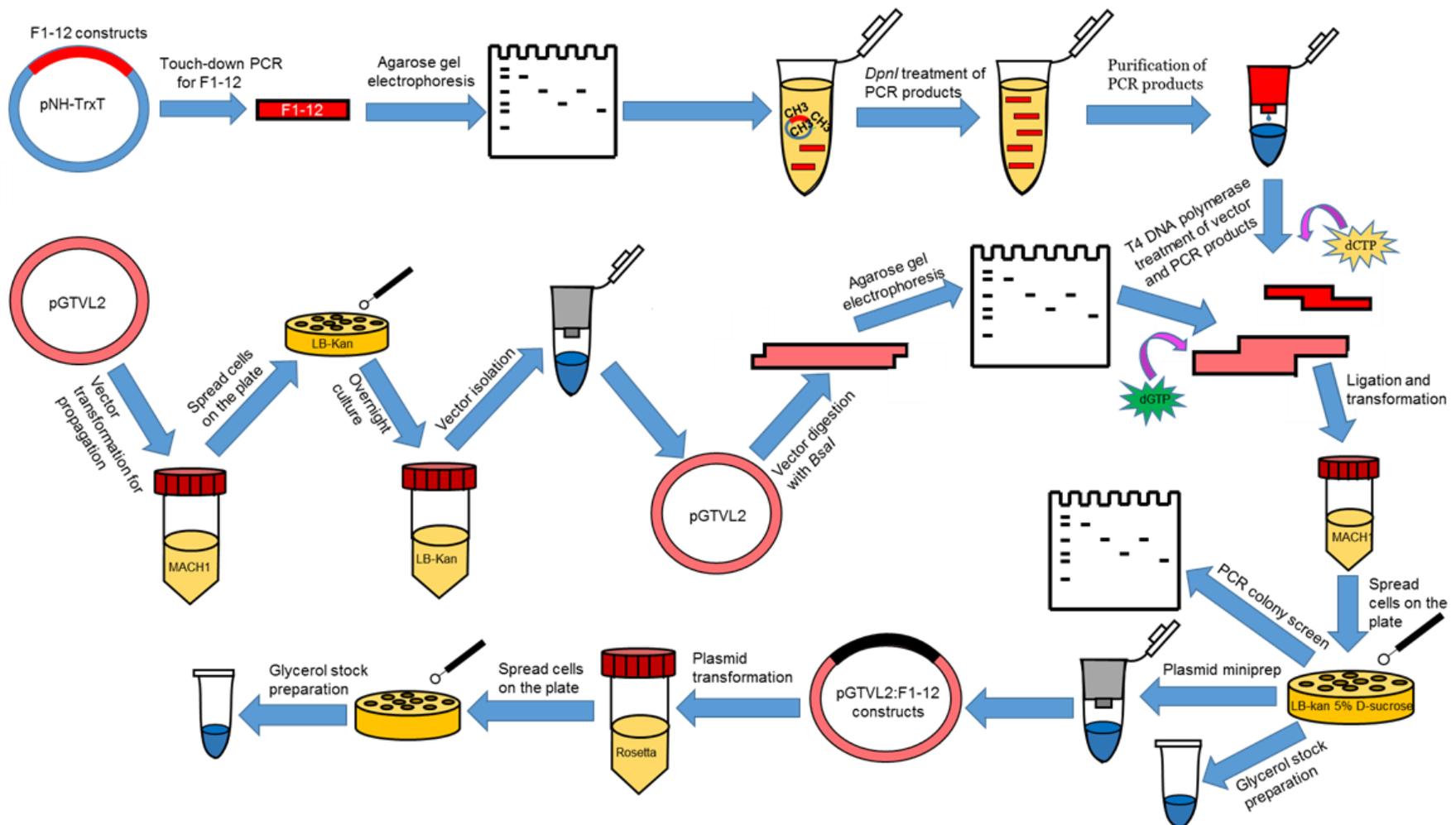


Figure 9. The schematic representation of the Ligation independent cloning and pGTVL2 vector preparation. The PCR product preparation is represented in the first top row. The second row shows pGTVL2 vector preparation, afterwards merging with the PCR products in the ligation reaction.

3.1.2. PCR product analysis

F1-12 POLDIP2 constructs (Table 5) were successfully amplified by touch-down PCR (Figure 10) from the pNH-TrxT vector which has full length POLDIP2 insertion (cloned by Klaudia Maruszczak). The touch-down PCR reduces non-specific primer binding (Don, Cox, Wainwright, Baker and Mattick, 1991). Don *et al.*, 1991; Korbie and Mattick, 2008). This in turn increases the amplification rate of the desired PCR product and avoids non-specific PCR product formation (Don *et al.*, 1991; Korbie and Mattick, 2008). Next, after the touch-down PCR the amplification of the correct DNA fragment was confirmed by running 1.5% agarose gel, confirming the presence of PCR products of specific DNA amplification. No extra fragments were detected on the gel, pointing out to the specific DNA amplification. The final gel with two combined PCR reactions is shown in Figure 10. Further, PCR products were *DpnI* treated before purification in order to digest the original template pNH-Trxt vector at homomethylated or hemimethylated (cleaved at much slower rate) palindromic 5'-GATC-3' sites (Siwek, Czapinska, Bochtler, Bujnicki & Skowronek, 2012; Mierzejewska *et al.*, 2014). Such vector treatment eliminates the appearance of false positive colonies, which can appear as a result of contaminating template pNH-TrxT vector transformation in *E. coli*, as both vectors encode kanamycin resistance (Siwek *et al.*, 2012; Mierzejewska *et al.*, 2014).

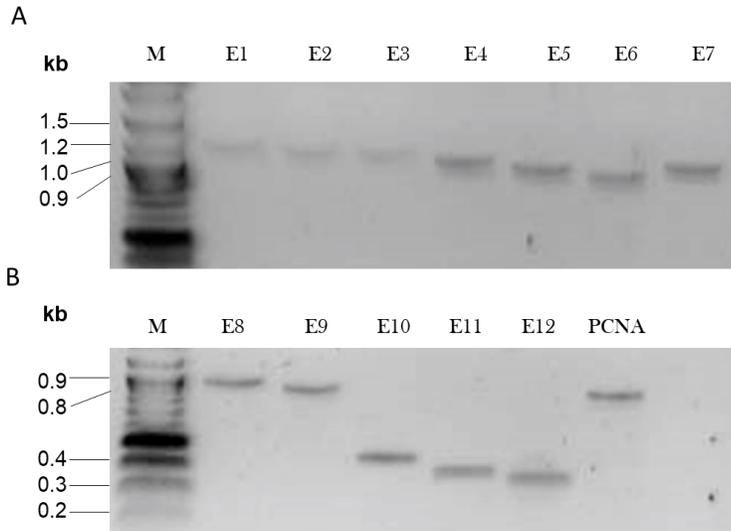


Figure 10. PCR amplification POLDIP2.

A: 1.5% agarose gel containing all expected POLDIP2 (E1-12) amplification fragments, as well as, the PCNA which was used as amplification positive control is shown on the left. M-marker; E- POLDIP2 construct cloned into pGTVL2 vector.

| Construct | Length of the PCR product (bp) | Construct name and a.a. sequence boundaries for (E1-12) |
|-----------|--------------------------------|---|
| E1/F1 | 1104 | POLDIP2 (1-368) |
| E2/F2 | 1044 | POLDIP2 (1-348) |
| E3/F3 | 1014 | POLDIP2 (1-338) |
| E4/F4 | 954 | POLDIP2 (51-368) |
| E5/F5 | 894 | POLDIP2 (51-348) |
| E6/F6 | 864 | POLDIP2 (51-338) |
| E7/F7 | 927 | POLDIP2 (178-368) |
| E8/F8 | 867 | POLDIP2 (178-348) |
| E9/F9 | 837 | POLDIP2 (178-338) |
| E10/F10 | 384 | POLDIP2 (221-368) |
| E11/F11 | 324 | POLDIP2 (221-348) |
| E12/F12 | 294 | POLDIP2 (221-338) |
| PCNA | 783 | FL |

Table 5. Expected POLDIP2 construct amplification protein size (bp) and respective protein boundaries (residues).

The first column includes construct name; second column states construct size (bp); third column shows construct residue boundaries. In all other experiments which will be further discussed in this work (except LIC section and if not otherwise stated) the annotation of the third column (on the right) will be related to E1-12 construct size.

3.1.3. Vector restriction digest

pGTVL2 vector was successfully linearized by treatment with the *BsaI* restriction enzyme (Figure 11). The expected restriction digest fragments are 1931 bp and 6016 bp in length and can be seen on an agarose gel. Precipitated, circular or supercoiled forms of non-digested pGTVL2 vector were not observed on the final agarose gel allowing further T4 DNA polymerase treatment and transformation.

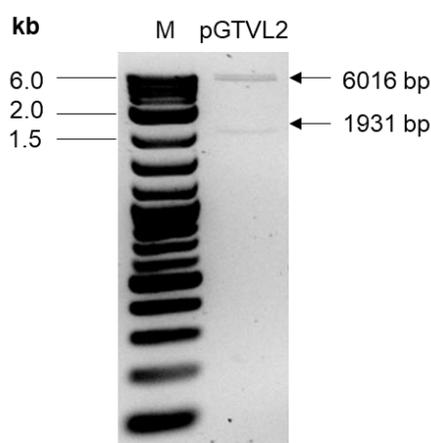


Figure 11. pGTVL2 vector restriction digest with *BsaI*. 3 μ L of the purified DNA was loaded on 1.5 % agarose gel. 2-Log DNA Ladder (NEB) is shown on the left. Black arrows on the right indicate the presence of the predicted vector digestion fragments.

3.1.4 PCR colony screen

Following ligation and cloning of purified POLDIP2 PCR products into pGTVL2, F1-12 constructs (Table 6), as well as, the PCNA were successfully amplified by colony PCR. The PCR was performed with the addition of DMSO/Glycerol and amplified PCR products, matching the expected size. An agarose gel which was run to verify the presence of the correct DNA fragments is shown in Figure 12.

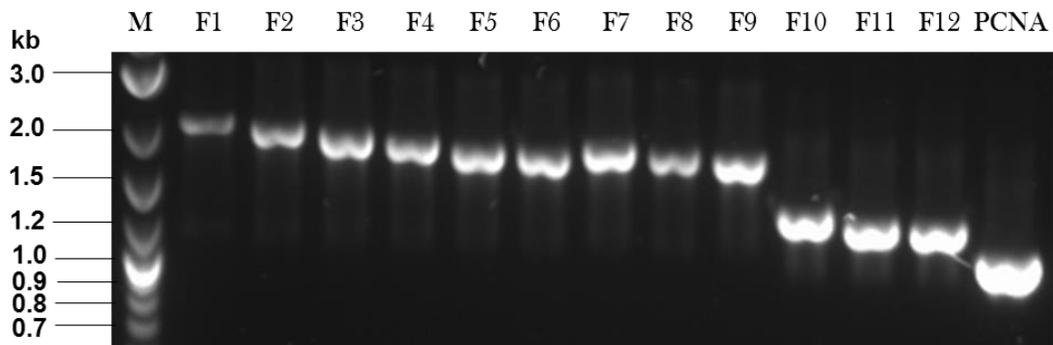


Figure 12. PCR colony screen agarose gel. The expected F1-12 PCR fragments (Table 6) can be seen on 1.5% agarose gel on the left. 2-Log DNA Ladder (NEB) is shown on the left.

| Construct/ Positive Control | Length of the PCR product (bp) | Length of the PCR product (bp)+ N-/C- terminal location of vector screening primers p) |
|-----------------------------------|---|--|
| F1 | 1104 | 2070 |
| F2 | 1044 | 2010 |
| F3 | 1014 | 1980 |
| F4 | 954 | 1920 |
| F5 | 894 | 1860 |
| F6 | 864 | 1830 |
| F7 | 927 | 1893 |
| F8 | 867 | 1833 |
| F9 | 837 | 1803 |
| F10 | 384 | 1350 |
| F11 | 324 | 1290 |
| F12 | 294 | 1260 |
| PCNA | 783 | 783 |

Table 6. POLDIP2 F1-12 and PCNA construct sizes (bp). Construct name is shown on the left; construct length in bp is shown in the middle column; construct length with the N-/C- terminal primer amplification region addition size (285 + 681 bp) is shown on the right.

3.2. Protein expression and purification

3.2.1. Small scale protein expression analysis using 2 growth media and 2 vector systems

Successfully cloned pGTVL2-POLDIP2 constructs (Table 7) were expressed in LB/TB media, comparing with the previously-cloned equivalent POLDIP2 constructs in the pNH-TrxT background. 50 ml test expression clearly shows how the presence of different tags (TrxT/GST) and media (LB/TB) makes an impact on protein's expression (Figure 13). Comparing POLDIP2 E1-12 constructs (pNH-TrxT) and PCNA expression in LB (A) and TB (C) it can be seen that the expression levels of these constructs is generally higher in TB media than in LB media. Also, in all 4 gels the PCNA is highly expressed, whereas CHDL1 has low expression level, as expected. CHDL1 expression levels are known to be poor, a common feature for the large human DNA binding proteins (Dr Christopher Cooper, *pers. comm.*). Generally, higher expression levels of PCNA (especially) and CHD1L are observed in TB media than in LB media. TB media is richer than LB, as it contains glycerol (also, additional inducing nutrients) and potassium phosphates, which respectively provide additional energy source and help to maintain appropriate pH (Lessard, 2013). This consequently leads to better cell growth and higher bacterial cell density accumulation (Lessard, 2013).

Table 7. POLDIP2 constructs length and size with and without TrxT/GST tags. Table shows POLDIP2 constructs boundaries which corresponds to the expected untagged, as well as, TrxT or GST tagged protein. E-POLDIP2 constructs cloned into pGTVL2 vector; F-POLDIP2 constructs cloned into pNH-TrxT vector.

| Protein name | E1 /F1 | E2/F2 | E3/F3 | E4/F4 | E5/F5 | E6/F6 | E7/F7 | E8/F8 | E9/F9 | E10/F10 | E11/F11 | E12/F12 | PCNA (+/control) | CHD1L (-/control) |
|---|--------|-------|-------|--------|--------|--------|---------|---------|---------|---------|---------|---------|---------------------|----------------------|
| POLDIP2 construct boundaries (a.a.) | 1-368 | 1-348 | 1-338 | 51-368 | 51-348 | 51-338 | 178-368 | 178-348 | 178-338 | 221-368 | 221-348 | 221-338 | FL | FL |
| Protein size without tag (kDa) | 42 | 40 | 39 | 37 | 35 | 33 | 36 | 34 | 32 | 15 | 13 | 12 | 29 | 113 |
| Approximate protein size with TrxT-tag (kDa) | 54 | 52 | 51 | 49 | 47 | 45 | 48 | 46 | 44 | 27 | 25 | 24 | N/A | N/A |
| Approximate protein size with GST-tag (kDa) | 70 | 68 | 67 | 65 | 63 | 61 | 64 | 62 | 60 | 43 | 41 | 40 | N/A | N/A |

Very weak expression of E2, E5 and E6 constructs is seen in LB media, whereas E3, E8-12 construct expression is not observed. E3 has a low expression in TB media. The E8-12 in TB media correspond to contaminating *E. coli* proteins which has bound to the purification beads, as the size of the bands does not match the predicted construct's size. Also, elutions which do not seem to have the expression of the predicted construct can be seen in F6, F8-9 and F12 constructs which are fused with GST-tag in pGTVL2 and expressed in LB/TB media, such assumption is based on the comparison to other nearby elution. For example, in F6 the band which is located nearby the F5 construct (which has an expected size, hence the correct protein is expressed) is located a little bit above the observed F5 band. However, F5 is 63 kDa and F6 is 31 kDa suggesting that the presence of F6 construct should be below the F5 construct due to the smaller proteins size.

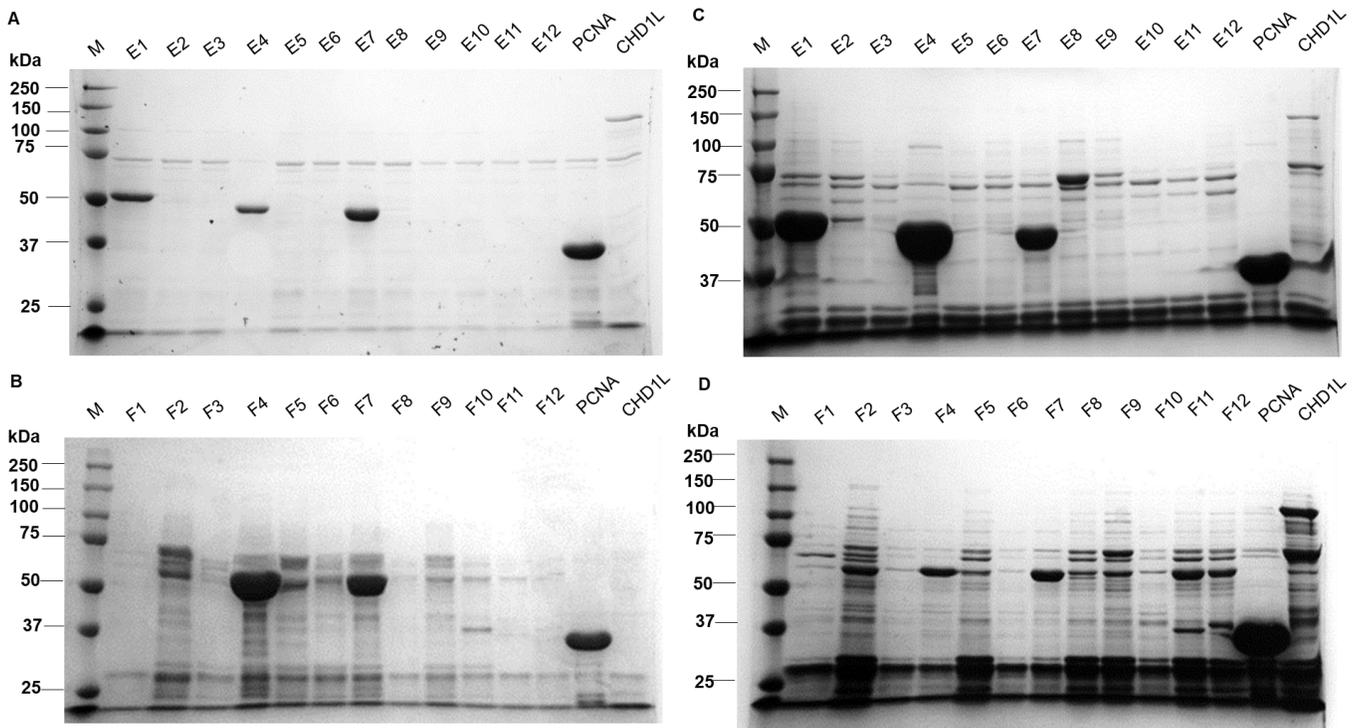


Figure 13. POLDLIP2 constructs expression in TB/LB media using TrxT/GST- tags. The expected POLDLIP2 constructs sizes are shown in Table 7. The first well in all gels (A-D) always contained Precision Plus Protein Dual Colour Standards Marker. POLDLIP2 F1-12 constructs containing GST-tag expressed in TB and LB media, respectively are shown in (B) and (D). The same POLDLIP2 constructs containing the TrxT-tag and expressed in TB and LB media, respectively are shown in (A) and (C). PCNA was used as a high expressing positive control, while CHD1L was used as expressing positive control. Other co-purified contaminants (proteins) can be seen in all wells and in all gels represented in this figure. A/B were expressed in LB media and C/D were expressed in TB media.

Although, E1, E4 and E7 have the better expression with a TrxT-tag in TB media, some protein constructs such as F2, F3, F5 and F10 are better expressed with the GST-tag in LB/TB media. F11 construct is only expressed with GST-tag in TB media. Interestingly, most POLDLIP2 F1-12 constructs were better expressed in LB than TB media, except the full-length POLDLIP (F1), F11 and PCNA, but such difference could be due to the poor bacterial cell lysis and experimental variability, as there is a small amount of background proteins present in these elutions, comparing

to the nearby elutions. Also, low protein levels, lower cell density and slower bacterial growth is observed then recombinant protein has toxic effect and interferes with normal cell's homeostasis (Miroux and Walker, 1996). For example, by binding to other proteins recombinant protein could interfere with important biological functions, such as bacterial proliferation (Doherty, Connolly and Worrall, 1993; Dong, Nilsson and Kurland, 1995).

GST and TrxT fusion tags can not only act as chaperones by stabilizing and solubilizing the recombinant protein, but can also protect the recombinant protein from proteolytic cleavage (Jacquet *et al.*, 1999; Martinez *et al.*, 1995; Kapust and Waugh, 1999; Davis, Elisee, Newham and Harrison, 1999; Esposito and Chatterjee, 2006). GST tag has the poorest recombinant protein solubility enhancement capability (Bird, 2011; Hammarström, Woestenenk, Hellgren, Härd and Berglund, 2006). Whereas, thioredoxin is one of the best and smallest available protein tags, as it displays much better recombinant protein solubility enhancement and often does not interfere with proteins activity (Costa, Almeida, Castro, Domingues and Besir, 2013; Rosano and Ceccarelli, 2014). For the reasons mentioned above TrxT samples were selected for the further experiments. As proteins yield was high for the E4/E7 and full-length POLDIP2 (E1) had a significant expression levels, these constructs were selected for the further study and were sent for the DNA sequencing.

3.2.2 Large scale POLDIP2 protein expression

POLDIP2 constructs (POLDIP2 (1-368) (E1), POLDIP2 (51-368) (E4), POLDIP2 (178-368) (E7)) resulting in a high protein yield were selected for the large scale purification. Several reasons which will be discussed below had an impact on high recombinant protein expression levels. First

of all, the vector contains strong T7 RNA promoter which leads to high protein expression levels (Kar and Ellington, 2018; Studier and Moffatt, 1986; Tabor and Richardson, 1985; Rosenberg *et al.*, 1987). Also, for the successful human POLDIP2 protein expression the Rosetta2 *E. coli* strain was used. Rosetta2 strain contains a plasmid which encodes for rare tRNAs (Sørensen, Sperling-Petersen and Mortensen, 2003; Sørensen and Mortensen, 2005). Such rare tRNA supplementation is particularly important for the start of recombinant proteins synthesis/production and it helps to overcome codon bias problems which can lead to a premature translational termination, frameshift mutations and/or even wrong amino acid incorporation (McNulty *et al.*, 2003; Kurland and Gallant, 1996; Sørensen, Sperling-Petersen and Mortensen, 2003; Sørensen and Mortensen, 2005). Moreover, several bacterial colonies were screened for protein expression, as gene expression variability is observed between different colonies containing the same recombinant gene (Mitarai, Jensen and Semsey, 2015). In some cases such variability can be caused by the presence of chromosomal T7 RNA polymerase mutation(s), which can lead to diminished polymerase expression levels and as a result to decreased recombinant protein expression (Vethanayagam and Flower, 2005; Kar and Ellington, 2018). Interestingly, even toxic human proteins show expression level variability between different colonies (Sivashanmugam *et al.*, 2009). Additionally, recombinant protein expression was carried out at low temperature, as low temperatures slow down hydrophobic protein-protein interactions preventing inclusion body formation (San-Miguel, Pérez-Bermúdez and Gavidia, 2013; Schein, 1989; Vasina and Baneyx, 1996; Baldwin, 1986). Likewise, slower rates of protein synthesis allow newly transcribed protein to fold correctly, preventing loss of activity (Vera, González-Montalbán, Arís and Villaverde, 2007).

3.2.3. Large scale protein purification

3.2.3.1. POLDIP2 (1-368) IMAC purification

Protein purity is the key step which determines the success of crystallisation and the quality of the solved protein three dimensional structure (McPherson, 2004). A highly pure and homogeneous protein sample is required, as other proteins (contaminants) may interfere with the packing of desirable protein into a crystal lattice leading to the poor crystal diffraction (McPherson, 2004). Hence, more ordered crystal lattice has higher diffraction power giving rise to a higher resolution of solved crystal structure (McPherson, 2004).

During the first purification step towards the purification of POLDIP2 (1-368) protein from the crude bacterial extract, Ni-NTA affinity chromatography was performed. From the first purification (Panel A, Figure 14) gel it can be seen that the majority of *E. coli* proteins did not bind to the column successfully passing through. In order to improve POLDIP2 (1-368) purification an imidazole (10 mM and 30 mM, Appendix 6) was used to eliminate weakly bound protein contaminants which were bound to the column (Gräslund *et al.*, 2008; GE Healthcare, 2006; Robichon, Luo, Causey, Benner and Samuelson, 2011). However, some bacterial proteins likely containing patches of histidines or even nonconsecutive histidines, as well as, functionally relevant metal binding sites which are located on proteins 3D surface can also bind to the column and sometimes such protein-column interaction can be quite strong (Frances, 1991; Cheung, Wong and Ng, 2012; Porath, Carlsson, Olsson, and Belfrage, 1975). Thus, these proteins were bound to the column and therefore were eluted together with 6xHis-tagged recombinant POLDIP2 (1-368) protein contaminating our sample (Frances, 1991; Porath *et al.*, 1975).

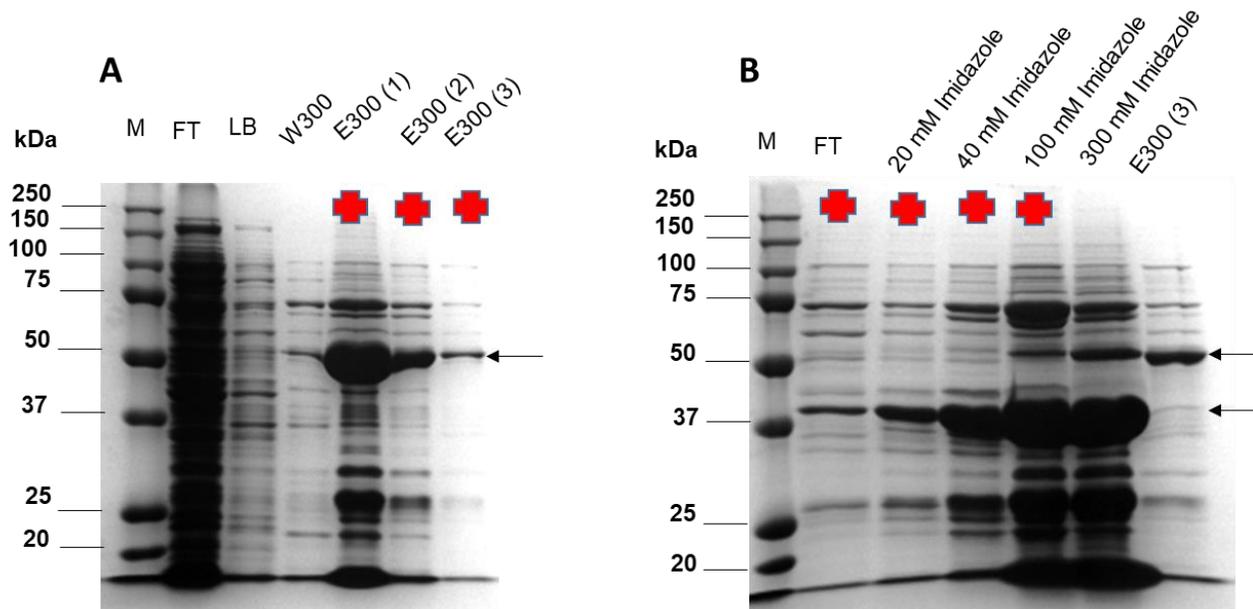


Figure 14. POLDIP2 (1-368) (E1) purification (8 L lysed). 20 μ L of the protein sample was loaded onto the first IMAC gel (Panel A) and 5 μ L of protein sample was loaded onto the second gel (Panel B). The black arrow points to the fragments which correspond to 42 kDa and 54 kDa, respectively indicating cleaved and TrxT-tag containing POLDIP2 (1-368) protein. Samples which were pooled and taken for further purification step are denoted with the red cross. The above explained legend relates to all gel purification figures which will be discussed in this thesis. A. This gel shows the initial protein purification by Ni-NTA Affinity Chromatography. B. The gel shows the second purification step, Ni-NTA Affinity Chromatography performed after dialysis and TrxT-tag cleavage with TEV protease. M-marker; FT-flow through; LB-lysis buffer; W300- wash buffer; E300 (1-3)- elution buffer. The following legend applies for all purification gels in this thesis. S200 16/600 column was used for all SEC experiments and all gels were 12% unless otherwise stated.

After dialysis and 6xHis-tag removal, POLDIP2 (1-368) was purified further (Panel B, Figure 14). POLDIP2 sample was still quite contaminated, because cleaved POLDIP2 strongly interacts with the column even at 100 mM imidazole levels. Such strong protein interaction with the column could be due to the presence of additional His patches or nonconsecutive His or non-specific ion exchange effects with the resin. In an ideal situation TEV-cleaved POLDIP2 (1-368) should pass through the column, whereas His-rich proteins, TrxT tag and TEV protease would be stuck to the

column resulting in a POLDIP2 (1-368) purification (Gräslund *et al.*, 2008). However, the TEV cleavage was successful, as the change of proteins size from 54 kDa to 42 kDa can be observed on the gel. On the gel POLDIP2 (1-368) decrease in size by ~12 kDa was observed, indicating removal of the ~12 kDa TrxT tag. This suggests no need to move TrxT tags position from N-terminus to C-terminus, as N-terminal TrxT tag does not exhibit conformational hindrance and allows TEV protease to cleave the TrxT tag (Block *et al.*, 2009; Raran-Kurussi and Waugh, 2016; Sørensen and Mortensen, 2005). The TrxT-tag was not cleaved completely however, resulting in the inability to pool the 300 mM imidazole elution (it has quite high amount of uncleaved POLDIP2 (1-368), as priority was given for the achievement of high proteins purification quality versus quantity). Imperfect TrxT tag cleavage could be due to short dialysis time, presence of high POLDIP2 (1-368) concentration and/or low TEV protease amount (Sørensen and Mortensen, 2005).

3.2.3.2. POLDIP2 (1-368) IEC purification

IEC is a purification method which separates proteins according to their net surface charge, which influences protein Ni-NTA binding ability (Alpert *et al.*, 2010). Increase of ionic strength of the buffer allows gradient proteins unbinding from the column resulting in a gradual protein separation (Alpert *et al.*, 2010). Moreover, even if proteins have the same net charge, they will have different surface charge distribution and that influences protein binding to the IEC column resulting in a better proteins separation (Alpert *et al.*, 2010). Three different columns (HiTrap Q; HiTrap SP and HiTrap Heparin), as well as, different pH ranges (6.5; 7; 7.5) were used for the POLDIP2 (1-368) IEC optimisation. The use of SP-IEC column with a combination of buffer pH 7 gave the best POLDIP2 (1-368) purification results, as in comparison to other tested conditions higher POLDIP2 percentage was bound to the column (data not shown). Thus, POLDIP2 (1-368) with net positive charge pI 8.8 was purified using the negatively charged SP-IEC column (pH 7 buffer) for all other

experiments, but still some POLDIP2 (1-368) protein did not bind to the column resulting in protein loss as can be seen in the (Figure 15, Panel A-B, (samples 1-3)). POLDIP2 (1-368) IEC chromatogram (Panel C, Figure 15) shows 2 peaks, the first one corresponds to the unbound POLDIP2 (1-368) and the second peak is POLDIP2 (1-368) bound to the column. The first fractions containing POLDIP2 (1-368) could not be taken as they contained predominantly contaminating *E. coli* proteins which pass through the negatively charged SP-IEC column, thus leading to our bound to the column POLDIP2 (1-368) purification (Schwartz, Ting, and King, 2001). Samples 13-14 (Panel A-B), were not taken as they contain quite large amount of 25 kDa contaminants in comparison to POLDIP2 (1-368) amount.

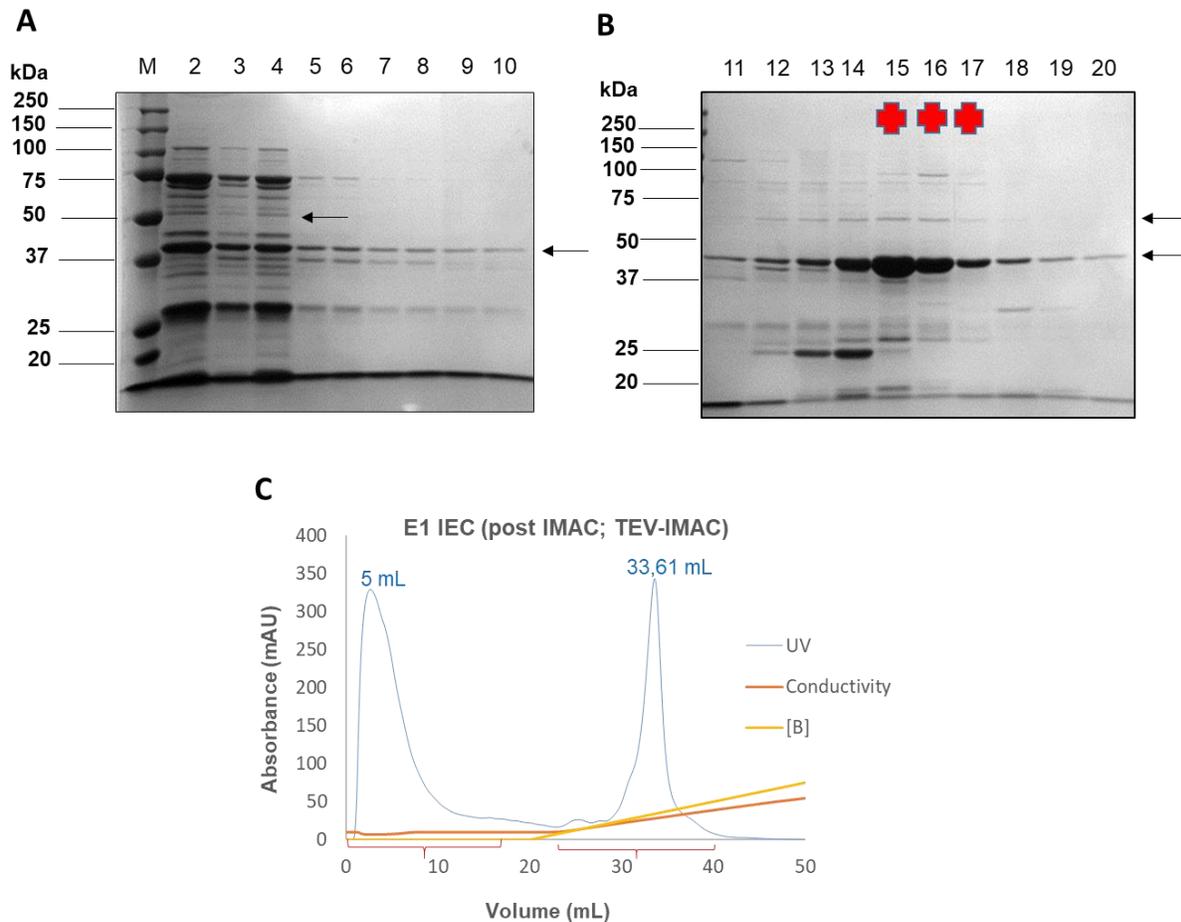


Figure 15. POLDIP2 (1-368) construct IEC purification. 5 μ L of fractions were loaded on gels. Panel A and B represent the gels after IEC using SP-IEC column. Panel C represents the chromatogram from ÄKTA Pure for IEC. The POLDIP2 elutes in the range between ~200-300 mM NaCl.

3.2.3.3. POLDIP2 (1-368) SEC purification

The size exclusion chromatography was used as the final POLDIP2 (1-368) polishing step. POLDIP2 (1-368) SEC resulted in very successful removal of small <30 kDa proteins (the fragment close to 25 kDa could be the product of proteolytic POLDIP2 cleavage), as well as, some other contaminants suggesting that only minor contamination remained (Panel A, Figure 16). The final POLDIP2 (1-368) gel is approximated to be >97% pure. Additionally, SEC can be used not only for the protein purification, but also for the protein oligomerization state determination (Sakashita, Kiyoi, Naoe and Urano, 2018; Gavrilov and Monteiro, 2015). Although, SEC separates proteins based on their size, sometimes proteins 3D shape can have an influence on proteins retention time (Sakashita *et al.*, 2018; Gavrilov and Monteiro, 2015).

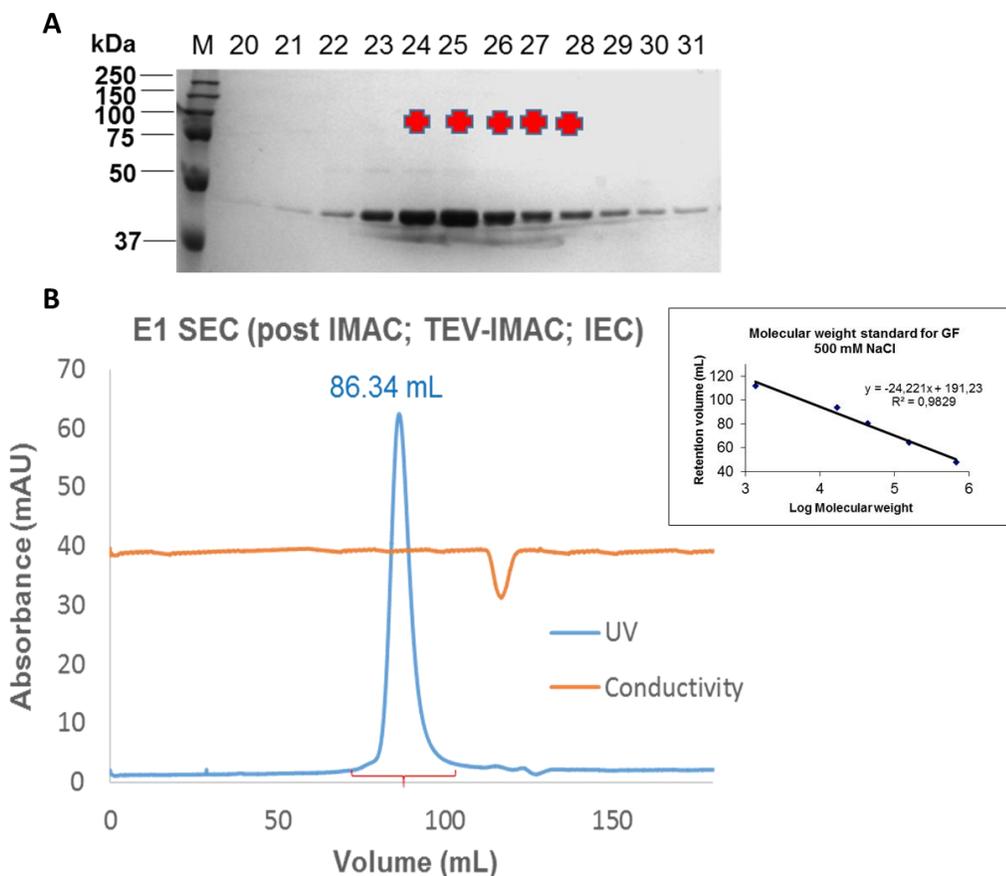


Figure 16. POLDIP2 (1-368) construct SEC purification (8 L lysed). 5 μ L of the protein sample were loaded on the gel. (A). Gel represents the final purification step by size exclusion chromatography. (B). The chromatography obtained after SEC with the calibration plot inset. The calibration plot was made using molecular size markers consisting of 5 proteins for POLDIP2 (1-368) (E1) oligomeric state estimation. Both protein and markers were run in the same SEC buffer.

3.2.3.4. POLDIP2 (51-368) IMAC purification

For the first POLDIP2 (51-368) purification step the Ni-NTA affinity chromatography was applied. The first purification gel (Panel A, Figure 17) shows that 6xHis-tagged POLDIP2 (51-368) is bound to the column (E300 (1)), while the large quantity of contaminating *E. coli* proteins passed through the column (FT, LB). After dialysis and 6xHis-tag removal, the POLDIP2 (51-368) proteins construct was applied to the Ni-NTA affinity column again (Panel B, Figure 17) and the major fraction of cleaved POLDIP2 (51-368) successfully passed through the column, but

some cleaved POLDIP2 (51-368) was bound to the column (probably due to the presence of His on proteins 3D surface). POLDIP2 (51-368) TrxT tag cleavage, can be confirmed by the POLDIP2 (51-368) change in mass. For example, FT and E300 (3) shown in Panel B, Figure 17, correspond to 49 kDa and 37 kDa. Difference in 12 kDa corresponds to TrxT tags size, indicating tag cleavage. Thus, POLDIP2 (51-368) N-terminal TrxT tag does not produce any steric hindrance and TEV protease is capable of cleaving the tag (Block *et al.*, 2009). However, TrxT tag is not fully cleaved and that could be due to: high POLDIP2 (51-368) concentration; small TEV protease concentration; too short dialysis time) (Sørensen and Mortensen, 2005). As POLDIP2 (51-368) was not fully cleaved fractions which contained un-cleaved POLDIP2 (51-368) were not taken, as TrxT-tagged POLDIP2 (51-368) may interfere with further assays/experiments (Sørensen and Mortensen, 2005). Thus, 100 and 300 mM Imidazole samples (Panel B, Figure 17) were not taken for further experiments as they contained quite large fraction of un-cleaved POLDIP2 (51-368), as well as, Ni-NTA bound contaminating *E. coli* proteins (Gräslund *et al.*, 2008). However, most contaminants remained bound to the column at this step, suggesting that only one additional purification step is needed. Bands near 25 kDa (Panel B, Figure 17) could be POLDIP2 (51-368) degradation bands. Therefore, one of those bands was cut out from the SDS-page gel and was sent for the mass spectrometry analysis (performed by Dr Rod Chalk, University of Oxford).

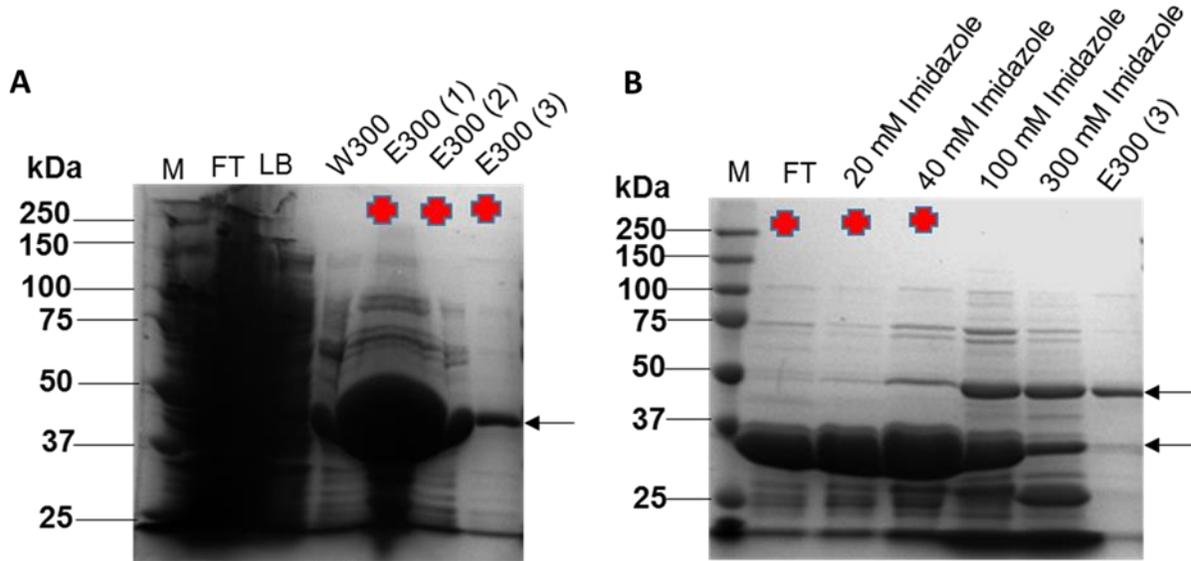


Figure 17. POLDIP2 (51-368) IMAC purification (4 L lysed). 20 μ L of the protein sample was loaded onto the IMAC gel shown on the left (Panel A) and 5 μ L of protein sample was loaded onto the gel shown on the right (Panel B). Panel A gel shows the initial protein purification by Ni-NTA Affinity Chromatography. Panel B gel shows the second purification step, Ni-NTA Affinity Chromatography performed after dialysis and TrxT-tag cleavage with TEV protease. The black arrow points to the fragments which correspond to 37 kDa and 49 kDa, respectively indicating cleaved and TrxT-tag containing POLDIP2 (51-368) protein.

3.2.3.5. POLDIP2 (51-368) SEC purification

The size exclusion chromatography (Figure 18) was used as the final ‘polishing’ step and resulted in further removal of contaminants suggesting that only minor contamination remained, and to ensure homogeneity. It can be approximated that the taken fractions are >97% pure, as the gel is overloaded. POLDIP2 (51-368) oligomerization state was determined based on the observed MW from the calibration (Panel B, Figure 18) and calculated MW from the proteins sequence. POLDIP2 (51-368) observed and calibration calculated MW ratio is 1.04, which can be rounded up to 1. Therefore, from the obtained calibration graph it was estimated that POLDIP2 (51-368) is

a monomer. However, full length (FL) POLDIP2 could be oligomer, as N-terminal POLDIP2 terminus could be involved in the oligomer formation (discussed in POLDIP2 (1-368) SEC purification). Lower bands near 25 kDa, could be POLDIP2 (51-368) degradation fragments. Samples 19-21 were not taken as they contained un-cleaved POLDIP2 (51-368) protein (Panel A, Figure 18).

Interestingly, POLDIP2 (1-368) has later retention time than the truncated POLDIP2 (51-368). However, POLDIP2 (1-368) is larger than POLDIP2 (51-368) and should come out from the column before POLDIP2 (51-368). Such non-standard POLDIP2 (1-368) behaviour, could be presumably due to the unstructured/flexible N-terminus which causes such non-ideal behaviour (*e.g.* difference in proteins 3D shape) (Kopaciewicz and Regnier, 1982). Also, POLDIP2 (1-368) N-terminus could potentially interact with the SEC column matrix, as a result causing the difference in the retention/elution time. From the obtained calibration graph it was observed that FL POLDIP2 is smaller than the calculated monomeric size (Panel B, Figure 18), while POLDIP2 (51-368) came out as a monomer. As the FL POLDIP2 (1-368) is larger than POLDIP2 (51-368) it should be expected to resolve larger than a monomer. Such, estimation is based on the observed MW from the calibration (Panel B, Figure 18) calculated MW from the proteins sequence using 4 different molecular size markers of known Mr. POLDIP2 (1-368) and POLDIP2 (51-368) observed and calibration calculated Mr ratios were, respectively 0.77 and 1.04. Also, it is important to note that then POLDIP2 (1-368) was resolved at 100 mM NaCl, the protein precipitated on the column making impossible to analyze the POLDIP2 (1-368) oligomerisation state at low ionic concentration. Thus, POLDIP2 (1-368) could potentially be oligomeric (dimer, trimer *etc.*) as the high salt concentration prevents electrostatic protein-protein interaction.

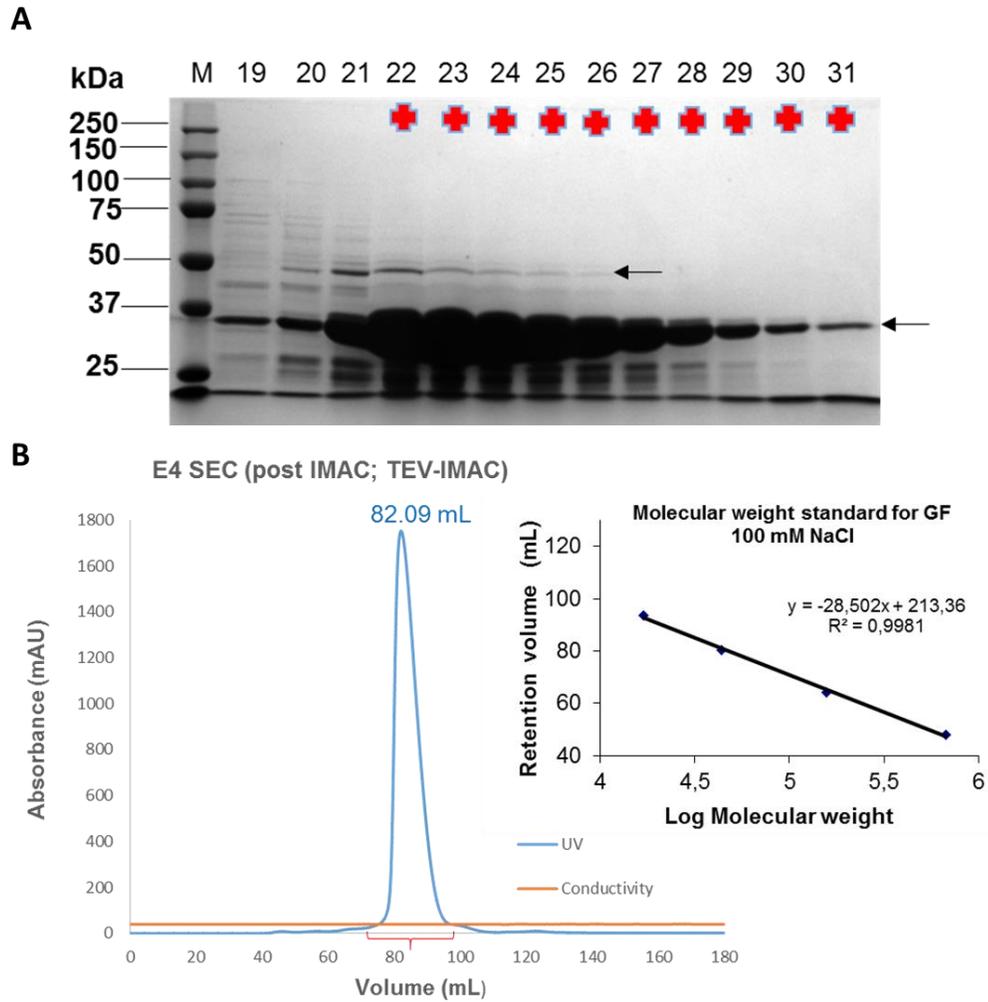


Figure 18. POLDIP2 (51-368) SEC purification. Panel A gel represents the final purification step by size exclusion chromatography. Panel B represents the chromatography obtained after SEC with the calibration curve inset. Panel B shows the calibration plot made using molecular size markers consisting of 4 proteins for the POLDIP2 (1-368) oligomeric state estimation. The 1350 Da protein in the marker was the outlier, hence unsuitable for the S200 SEC column, therefore was excluded from the calibration plot. Both protein and markers were run in the same SEC buffer.

3.3. Structural protein characterization

3.3.1. POLDIP2 (51-368) secondary structure characterisation by circular dichroism

As the circular dichroism (CD) spectrum of a protein is the sum of the spectra that represent the secondary structural elements, CD spectra can be used to approximate the secondary structural content found in the protein (Greenfield, 1999). From the obtained far-UV spectra it is clearly seen that POLDIP2 (51-368), also referred as E4 is not a disordered protein (Figure 19), but mainly it has well-structured regions, as no positive peak was observed in the region near 217 nm and no negative peak observed near 200 nm region (Greenfield, 1999; Bannister and Bannister, 1974).

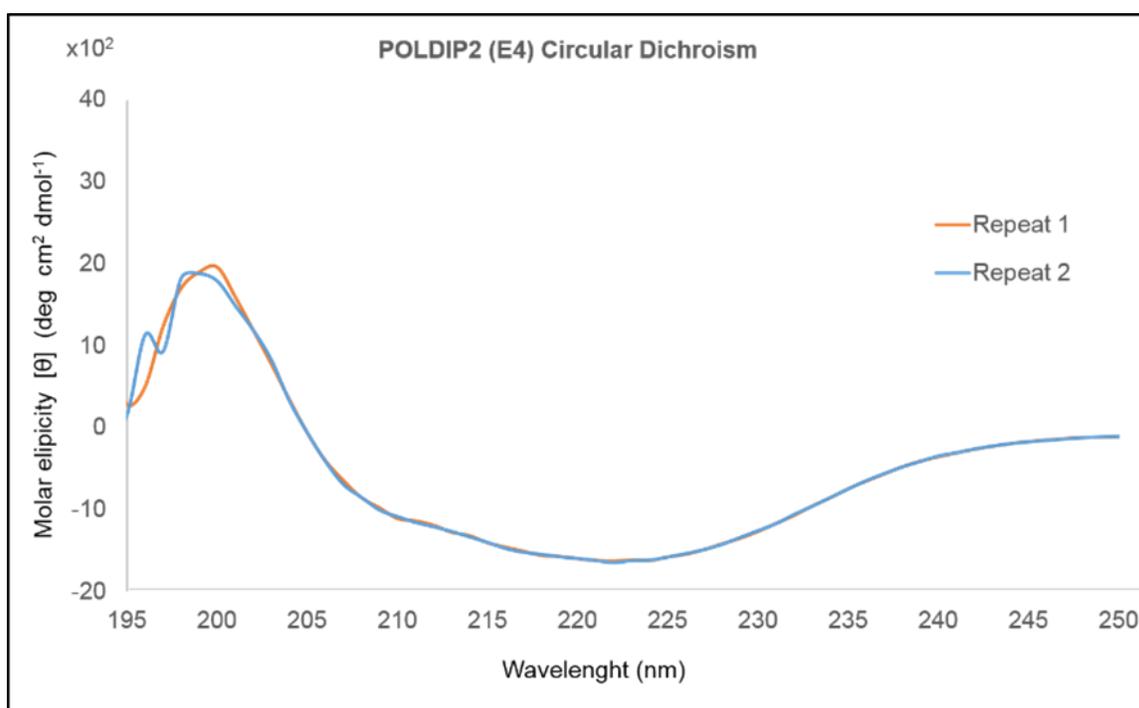


Figure 19. Far-UV circular dichroism spectra of POLDIP2 (51-368, E4). 2 repeats were performed. The slight mismatch between two repeats in the 195-203 nm region is potentially due to POLDIP2 (51-368) degradation (Nasir Khan, *pers. comm.*).

Indeed, the peak is mostly in the β -strand range, as the broad negative band is located near 218 nm and a positive band is present near 195 nm (Whitmore and Wallace, 2004). The presence of high β -strand content in POLDIP2 was confirmed by DICHROWEB software which predicted it to be 57.2% (Whitmore and Wallace, 2004). The region of 220 nm in β -sheets is usually less ‘smooth,’ suggesting the presence of some α -helical content (Whitmore and Wallace, 2004). Structures as α -helices contain a very long and smooth peak near 222 nm (Greenfield, 1999; Bannister and Bannister, 1974). Such prediction does not contradict the DICHROWEB software prediction, as the estimated α -helical content in POLDIP2 is 23.4% (Whitmore and Wallace, 2004). The observations mentioned above were made with an aid of an idealized proteins CD spectra, taken from the Figure 20. Generally, β -sheets contain intermediate loops which join 2 β -strands together, thus it would be expected to have some flexible loops (Whitmore and Wallace, 2008). The presence of flexible loops/turns was predicted to be 19.3% using DICHROWEB software (Whitmore and Wallace, 2004).

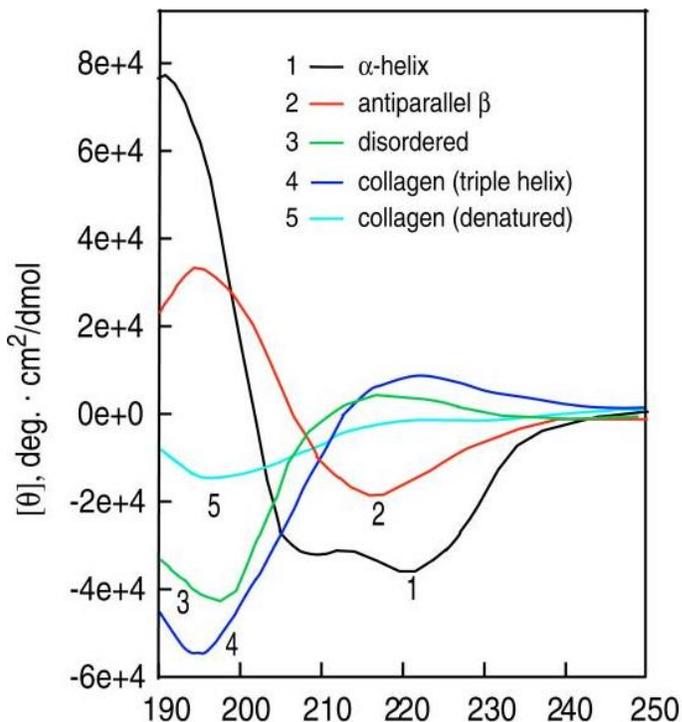


Figure 20. Idealized CD proteins spectra.

Figure shows peaks which corresponds to proteins secondary structure. The figure taken from Greenfield (2006).

3.3.2. Protein Crystallography towards POLDIP2 structural determination

The schematic representation of the protein structure determination process is shown in Figure 21, to aid the subsequent process understanding described further in this thesis.

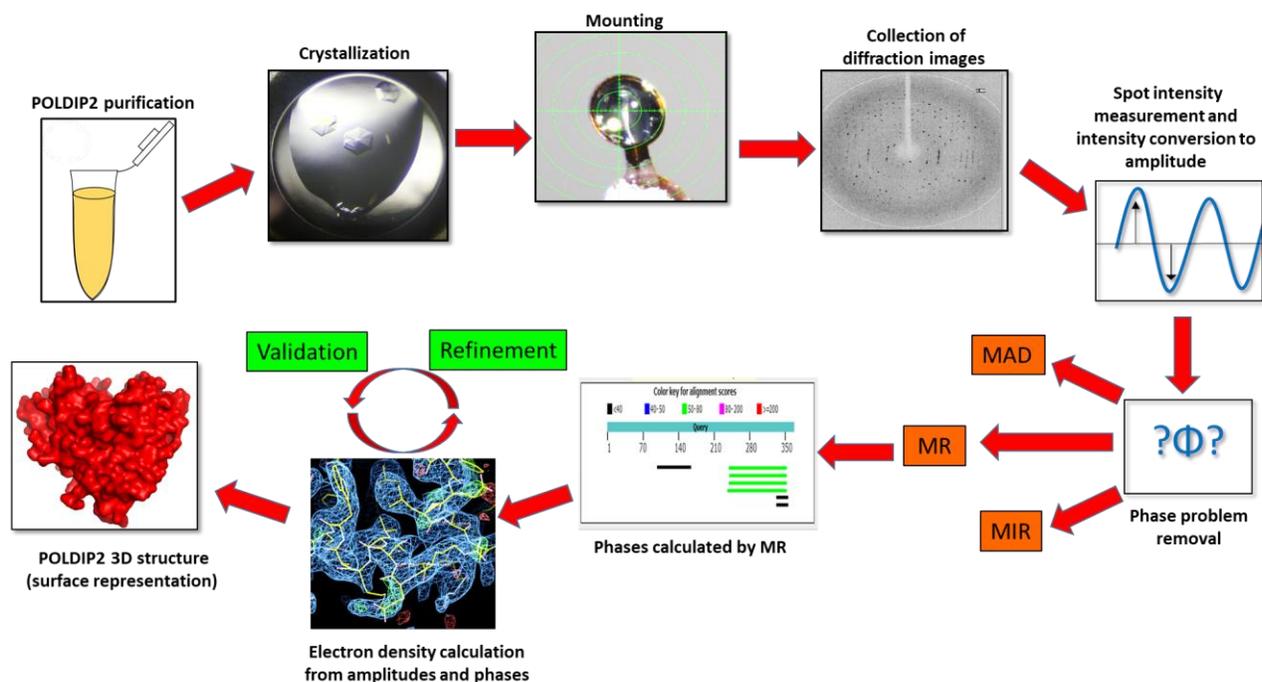


Figure 21. Schematic representation of steps leading to POLDIP2 (51-368) structure solution. MAD- Multi-wavelength Anomalous Dispersion; **MR-** Molecular replacement; **MIR-** Multiple Isomorphous Replacement.

3.3.2.1 Crystallisation of POLDIP2 (51-368)

There is no any correlation between proteins 3D structure and crystal forming condition (Chayen and Saridakis, 2008). Crystallisation conditions cannot be predicted *a priori*, hence multiple conditions should be tried (Jancarik and Kim, 1991). The sitting drop technique was used to obtain protein crystals using commercial JCSG, PACT, BCS and Morpheus kits, hence, covering a wide range of chemical ‘space’ sampling, which could potentially promote proteins crystallisation

(Snell *et al.*, 2008). All these screens are designed, based on the observed conditions under which many other proteins were crystalized (Jancarik and Kim, 1991). Then conditions under which protein crystallises are found, such proteins crystallisation condition can be used as a “hint” for the crystallisation optimisation. Protein crystals usually grow at a preferable temperature which is however in not known for a new protein (Liu *et al.*, 2011). Thus, the first crystal plates with JCSG and BSC were set at 4°C and 20°C. POLDIP2 (51-368) was sensitive to the change in the temperature, as difference in crystal growth was observed (data not shown). Plates which were incubated at 4°C resulted in very small microcrystals, which did not grow any further even then plates were checked after 5 months (Table 8; Figure 22, A(1)). Although, small crystals can have very good diffraction (there is no correlation between proteins size and diffraction quality), the observed crystals were too small to be scooped from the mother liquor for the screening purposes (Sitsel and Raunser, 2019). It can be concluded that probably even if evaporation has occurred during this 5 months, there was no protein left in the solution to form any larger crystals due to access nucleation. Also, it is likely that POLDIP2 (51-368) was exhausted with the multiple nucleations, so once nucleated POLDIP2 (51-368) crystals could not grow larger (McPherson, 2004). High nucleation rate could occur due to the presence of dust or denatured particles (Chayen and Saridakis, 2008). Protein could be filtered through the syringe filter before setting crystal plates, but in this case there is a possibility that nucleation might be not observed as these particles might be important for the crystal growth (Chayen and Saridakis, 2008). It was attempted to crystallise the POLDIP2 (178-368) construct using the same screens JCSG/BCS (using 4°C and 20° C), however, POLDIP2 (178-368) construct did not yield any crystals (Appendix 9). This suggests that the slight N-terminal extension of POLDIP2 (51-368) is important for crystal packing. Moreover, this suggests that POLDIP2 (51-368) 3D structure could be more stable/tightly

packed, presumably due to the additional secondary structure formation in comparison to POLDIP2 (178-368) which could be less tightly packed/less stable.

However, the same condition incubated at 20°C resulted in a three hexagonal mono-crystals (crystals not fused to each other) Figure 22, (A (1)-(3)). These three crystals are most likely to be located in the region of the 'super solubility curve' - the best area to grow large crystals, the area where the lower nucleation zone meets with the upper metastable zone (Luft, Wolfley, and Snell, 2011; McPherson and Gavira, 2013). The crystallisation of POLDIP2 (51-368) (Figure 22, Panel A) was reproducible, as the frozen protein from the same preparation, as well, as the protein from a new preparation, both resulted in the formation of very similar crystals. Under other conditions (B-E) small crystals appeared probably being in the nucleation zone. Even the crystals (A) were observed in the first set plate, further optimisation was required in order to gain a crystal which could give better diffraction.

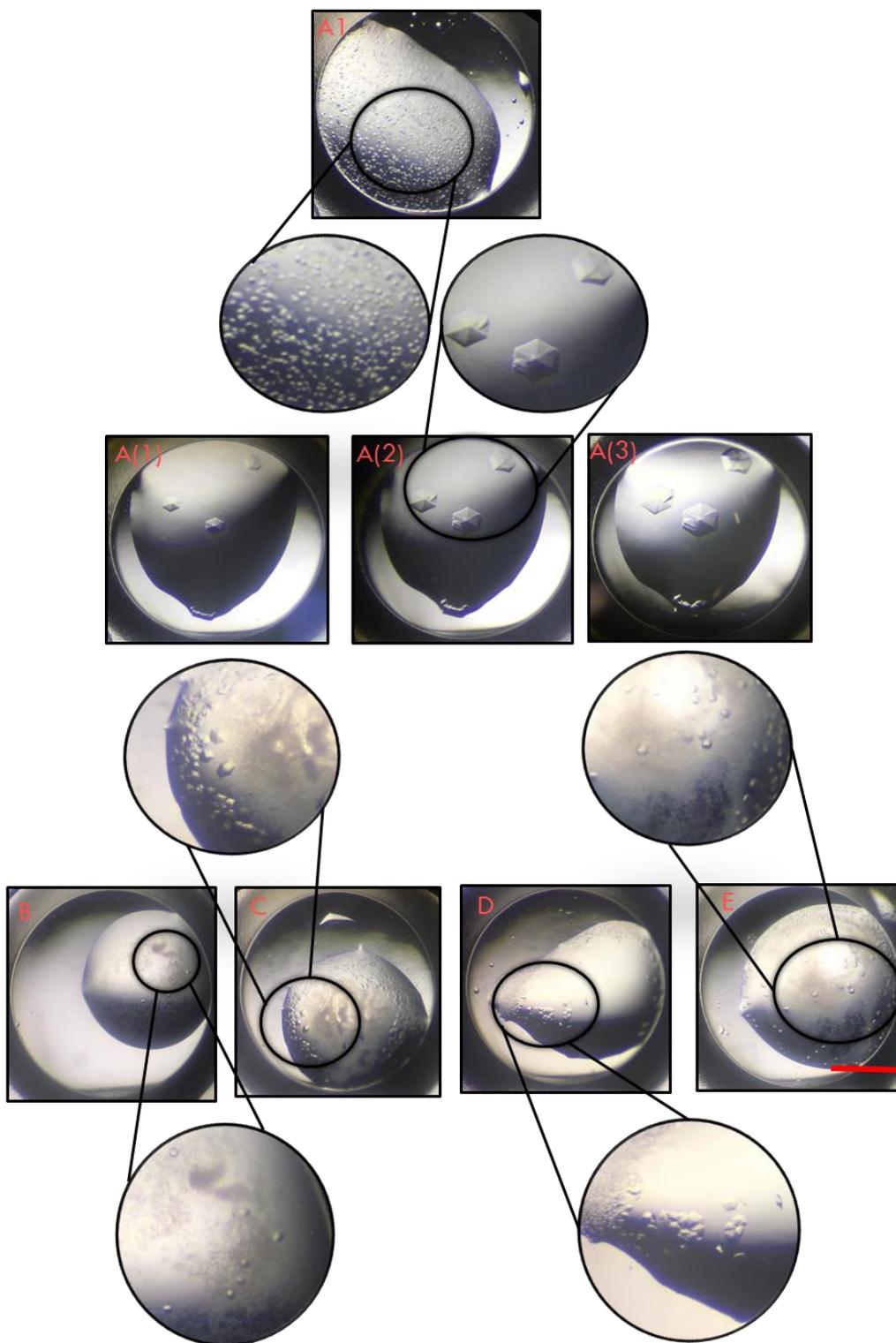


Figure 22. POLDIP2 (51-368) protein crystals obtained using commercial kits. A corresponds to the same protein crystallisation conditions, but A1 was incubated at 4°C while A (1)-(3) at 20°C. A (1)-(3), represents the same crystals growth after 21 hours, after 2 days and 7 days (crystal growth stopped). The

A1 figure was taken after 21 hours, however, the crystal growth stopped at this stage. Figure 22, B-E shows other conditions were POLDIP2 (51-368) crystals were observed. The drop volume was 2 μ l, protein and buffer were mixed in equal volumes. Scale bar represent approx. 0.92 mm and is applicable for all wells. Conditions under which the illustrated POLDIP2 (51-368) crystals were observed are shown in Table 8.

Table 8. Conditions under which POLDIP2 (51-368) protein crystals were obtained using commercial kits.

| Number | Screen | Salt | Conc. Buffer | pH | Precipitant |
|--------|----------|--|------------------------------------|-----|---|
| A | JCSG | 0.2 mM Calcium acetate hydrate | 0.1M Sodium cacodylate | 6.5 | 40% v/v PEG 300 |
| B | BCS | 40 mM Calcium chloride dehydrate 40 mM Sodium formate | 0.1 M Tris | 8 | 0,25% v/v PEG Smear Low (12.5%v/v PEG 400 12.5%v/v PEG 500 MME 12.5%v/v PEG 600 12.5%w/v PEG 1000) |
| C | Morpheus | Magnesium chloride/ Calcium chloride | 0.1 M Sodium Hepes; MOPS (acid) | 7.5 | 30% PEGMME 550; PEG 20K |
| D | Morpheus | Magnesium chloride/ Calcium chloride | 0.1 M Sodium Hepes; MOPS (acid) | 7.5 | 30% Ethylene glycol; PEG 8K |
| E | Morpheus | Magnesium chloride/ Calcium chloride | 0.1 M Sodium Hepes; MOPS (acid) | 7.5 | 37,5% PEGMME 550; PEG 20K |

3.3.2.2. POLDIP2 (51-368) crystallisation optimisation

Following the initial observation of crystals, new crystals screens were designed, taking into the account conditions under which initial crystal formation was observed. 2 plates were designed combining different conditions shown in Table 8. POLDIP2 (51-368) optimisation was made in order to obtain crystals which could have higher internal order (crystalline order), hence leading to a better diffraction and resolution (McPherson, 2004). This is because slight variation in polypeptide side chain conformations can be observed between different crystals (McPherson, 2004). Although, POLDIP2 (51-368) successfully crystallises in the presence of cacodylic acid, the use of this reagent was avoided in the optimization, due to the reagent's toxicity. Optimisation

involved varying salt concentrations (100 and 200 mM CaCl₂) the pH (6.5; 7; 7.5; 8), buffer (Tris, Hepes and MOPS) different precipitants (PEG300; PEG500 MME; ethylene glycol and MPD), different precipitant percentages which were above and below the observed percentage under which crystals formed. CaCl₂ was used as it was observed in 3/5 conditions where POLDIP2 (51-368) crystals formed and Ca²⁺ was observed in the rest 2/5 conditions (see Table 8). Cations can dehydrate proteins as they both compete for the water, thus molecules which are deprived of solvent will seek to associate with one another (Collins, 2004). All used precipitants except MPD were observed under conditions where POLDIP2 (51-368) crystallised. However, MPD is often found to be important for the precipitation and can act as a cryo-protectant, thus was also used for the POLDIP2 (51-368) crystallisation (Jancarik and Kim, 1991). Ideally cryo-protectant should be present in the crystallisation solution in order to avoid crystal damage during cryo-protectant addition (Garman, 1999; Garman and Schneider, 1997). Precipitants usually do not directly interact with proteins, but they can produce volume exclusion effect causing protein molecules to separate from the solution and hence potentially form a crystal (Ingham, 1990). For the optimization, conditions under which POLDIP2 (51-368) crystallization was observed were further optimized (pH and buffers).

The same sitting drop method was used during the optimisation. POLDIP2 (51-368) crystal formation was observed in the wide range of different pH 6.5-8. Interestingly, hexagonal POLDIP2 (51-368) monocrystals were observed in the pH 6.5 which is very close to POLDIP2 (51-368) pI 6.45 (A1-3; Figure 23; Table 9). A1-3 crystals are more likely to be present in the region of the ‘super-solubility curve’. No particular crystal formation pattern was observed, except for A1-6 conditions. In A1-6 larger crystals were formed under the lower PEG300 concentration, while the subsequent increase in PEG300 concentration lead to the smaller crystal formation, finally causing

protein precipitation. POLDIP2 (51-368) crystals were observed when higher molecular weight PEG500 was used at higher pH 7.5-8 (F1-11). However, no particular crystal formation pattern was observed and some wells did not contain any crystals. Higher molecular weight PEG500 MME produced crystals at 100 mM Ca^{2+} , but at 200 mM Ca^{2+} protein precipitation increased and if crystals were present they were quite small. For the PEG300 at 200 mM Ca^{2+} smaller crystals were observed, but less precipitation in comparison to PEG500 MME. In general, more crystals and bigger crystals were observed at 100 mM Ca^{2+} than 200 mM Ca^{2+} . No crystals were observed in wells where MPD or ethylene glycol was present.

Table 9. Conditions under which crystals were observed

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-----------------------|------------------------------|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | MPD 20% | 23% | 26% | 29% | 32% | 35% | 20% | 23% | 26% | 29% | 32% | 35% |
| All wells have: AK018 | | Other 25% | 30% | 35% | 40% | 45% | 50% | 25% | 30% | 35% | 40% | 45% | 50% |
| A | 100 mM CaCl2 PEG 300 | YY | NY | YY | | YY | NY | | NY | | YY | YY | YN |
| B | 100 mM CaCl2 PEG 500 MME | YY | NY | NY | | YY | YN | | NY | YY | YY | YY | YY |
| C | 100 mM CaCl2 Ethylene Glycol | | | | | | | | | | | | |
| D | 100 mM CaCl2 MPD | | | | | | | | | | | | |
| E | 100 mM CaCl2 PEG 300 | | | YN | NY | | YY | | YN | | YN | YY | YY |
| F | 100 mM CaCl2 PEG 500 MME | YY | YY | | | | | | NY | YY | NY | | YN |
| G | 100 mM CaCl2 Ethylene Glycol | | | | | | | | | | | | |
| H | 100 mM CaCl2 MPD | | | | | | | | | | | | |

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-----------------------|------------------------------|-----------|-----|-----|-----|--------|------|-----|-----|-----|-----|-----------|------|
| | | MPD 20% | 23% | 26% | 29% | 32% | 35% | 20% | 23% | 26% | 29% | 32% | 35% |
| All wells have: AK019 | | Other 25% | 30% | 35% | 40% | 45% | 50% | 25% | 30% | 35% | 40% | 45% | 50% |
| A | 200 mM CaCl2 PEG 300 | YN | | | | NY | YY | | | | | YN | YN |
| B | 200 mM CaCl2 PEG 500 MME | YY | YY | YY | Pre | Cipita | Tion | YN | YY | YY | YY | Precipita | tion |
| C | 200 mM CaCl2 Ethylene Glycol | | | | | | | | | | | | |
| D | 200 mM CaCl2 MPD | | | | | | | | | | | | |
| E | 200 mM CaCl2 PEG 300 | | | NY | | | | | | | | | |
| F | 200 mM CaCl2 PEG 500 MME | YY | YY | YY | | | | YY | YN | YN | YY | Precipita | tion |
| G | 200 mM CaCl2 Ethylene Glycol | | | | | | | | | | | | |
| H | 200 mM CaCl2 MPD | | | | | | | | | | | | |

 MOPS pH 6,5
 Hepes pH 7,5
 Hepes pH 7
 Tris pH 8

Y-crystals present
 N-no crystals
 The letter in the left -upper well
 The letter in the right -bottom well

 MOPS pH 6,5
 Hepes pH 7,5
 Hepes pH 7
 Tris pH 8

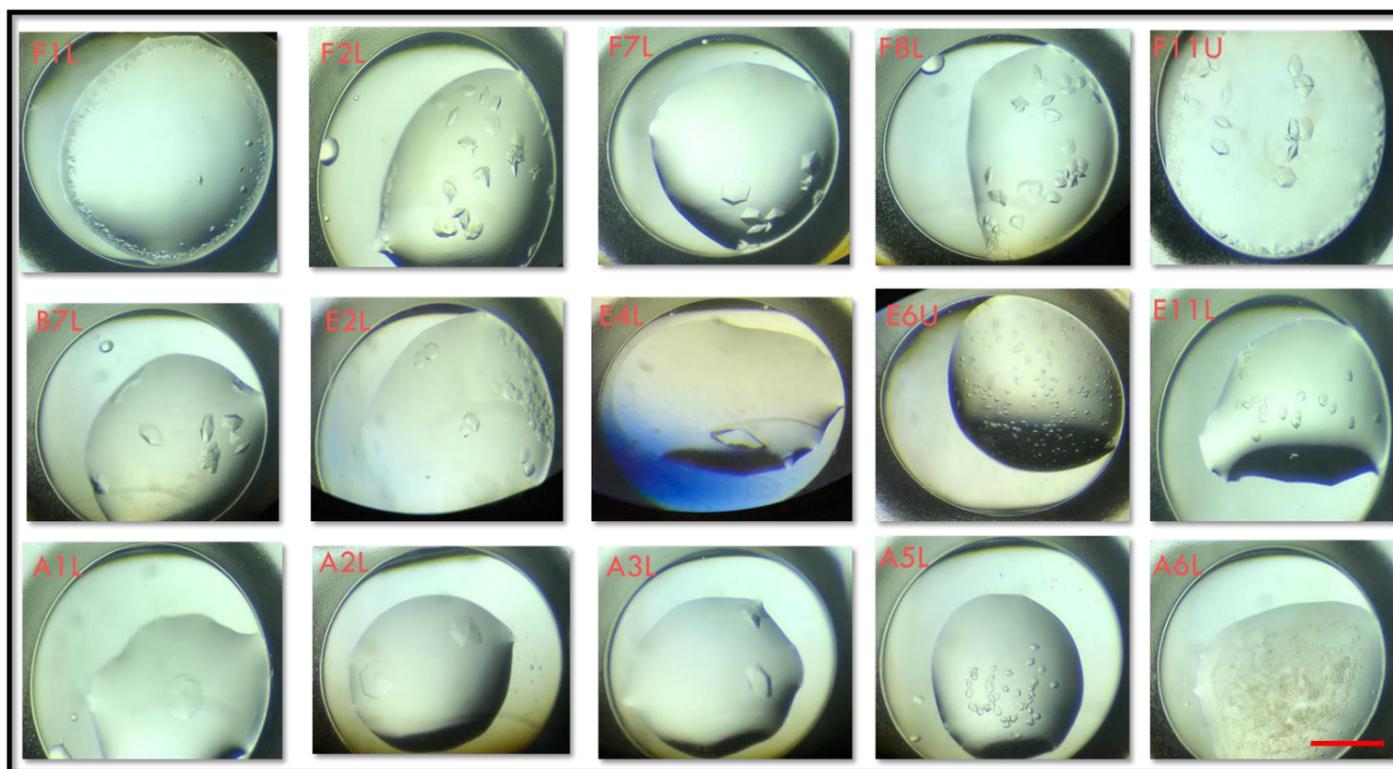


Figure 23. POLDIP2 (51-368) optimisation under conditions stated in Table 9. All shown crystals were observed after 24 hours. All crystals are from 100 mM Ca^{2+} plate. Most crystals shown in the figure appear to be in the metastable zone, while F1, E6, E11, A5 and A6 appear to be in the nucleation zone. Scale bar represent approx. 0.92 mm and is applicable for all wells. L, lower well; U, upper well.

3.3.2.3. Crystal mounting and X-ray diffraction screening

A loop containing a POLDIP2 (51-368) protein crystal was mounted and flash frozen, prior to placing on a goniometer (Figure 24). Loops were positioned precisely within the X-ray beam by spinning it at Φ 90° and 180° until the crystal appeared to be in the centre in all directions. Applied cryo-cooling on the goniometer can cause lattice disorder (increased mosaicity) and ice ring formation leading to poor crystal diffraction (Kriminski, Caylor, Nonato, Finkelstein and Thorne, 2002). For this reason crystals were warmed up for few seconds and the frozen again

(Kriminski *et al.*, 2002). This was done in order to promote crystal lattice relaxation and ice ring disappearance leading to a better crystal diffraction (Kriminski *et al.*, 2002). The crystal generally showed diffraction pattern with low mosaicity, as no spot overlap was visually identified and spots were resolved (Sarvestani, Walenta, Busetto, Lausib & Fourmec, 1998). The diffraction spots which were furthestmost from the diffraction pattern centre were identified in the higher resolution region to be ~ 2.89 Å. Although, multiple crystals were screened the POLDIP2 (51-368) which was crystallized at 20°C (condition A, Table 8; Appendix 9) had the better predicted diffraction and was further used for the molecular replacement.

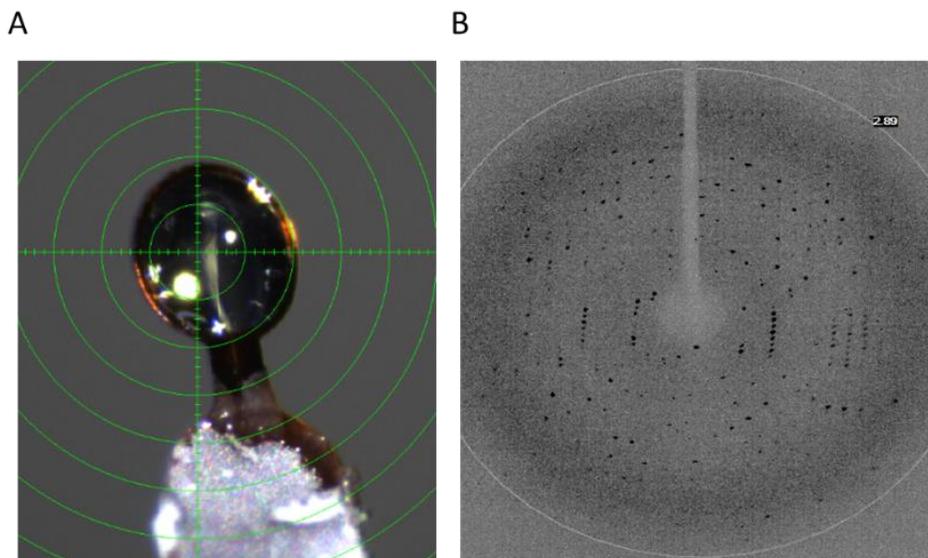


Figure 24. Images of POLDIP2 (51-368) protein crystal and corresponding X-ray diffraction pattern.

The crystal in the loop, as well as, the diffraction pattern belongs to one of three crystals which were shown in Figure 22, A (1)-(3) images. However, only the small part of the crystal is present in the loop as the crystal broke into smaller parts during the harvesting. A. POLDIP2 (51-368) crystal in the loop. B. Black spots are the result of constructive x-ray waves interference, while the space in between belongs to destructive x-ray waves interference. The white ring indicates the resolution rings. The right Panel shows diffraction to ~ 2.89 Å resolution.

3.3.2.4. POLDIP2 (51-368) structure solution

3.3.2.4.1. Molecular replacement of POLDIP2 (51-368)

Similar sequences were searched using BLASTP for POLDIP2 (51-368) to solve proteins structure using the Molecular Replacement (MR) method (Evans and McCoy, 2008). The models of these proteins were used to obtain the phases, and to build the electron density map in the unit cell by combining them with the experimentally observed amplitudes (Evans and McCoy, 2008). Five similar sequences were identified for the POLDIP2 (51-368) sequence (Figure 25). The obtained sequences showed low sequence identity. However, the amino acids composition was ~32-33% identical to the POLDIP2 (51-368) sequence and the identified sequence appeared to be quite short covering only ~30-31% of the POLDIP2 (51-368) sequence. All 5 found protein were found to cover POLDIP2 (51-368) sequence from residues ~237-241 to 353, the region which shows the sequence similarity to YccV- like domain. However, the level of positive substitutions between the POLDIP2 (51-368) and similar sequences was quite high ranging from 43% to 52%. All these 5 similar sequences were used further separately for the molecular replacement (MR) using Phaser (undertaken by Dr Richard Bingham).

Results of MATHEWS Coefficient:

| Nmol/asym | Matthews Coeff | %solvent | P(2.69) | P(tot) |
|-----------|----------------|----------|---------|--------|
| 1 | 2.79 | 55.95 | 0.99 | 0.99 |
| 2 | 1.40 | 11.89 | 0.01 | 0.01 |

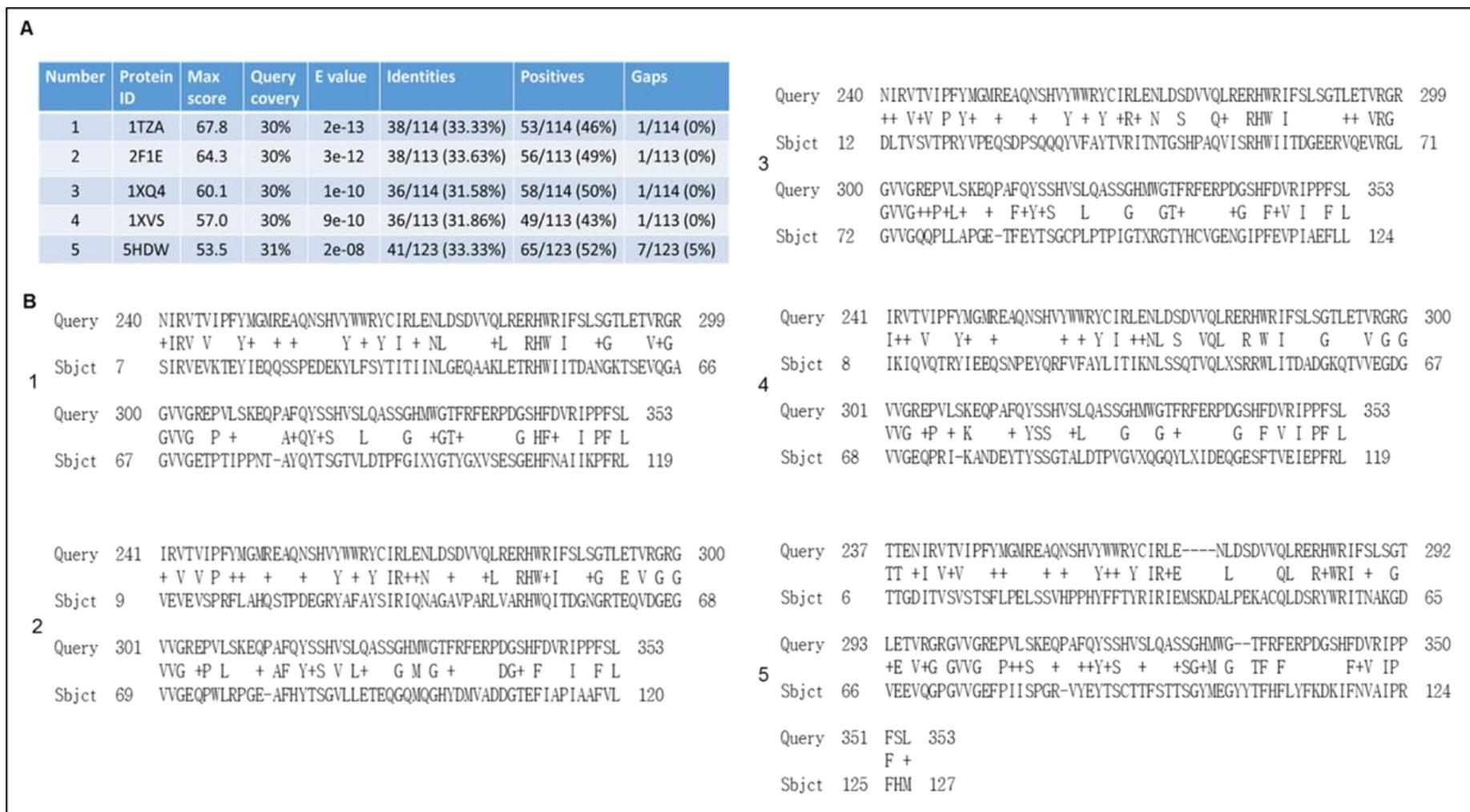


Figure 25. Similar sequences alignment for POLDIP2 (51-368). BLASTP for POLDIP2 (51-368) was run against protein database (PDB) using BLASTP BLASTP algorithm. A. The obtained values describing proteins sequence homology are shown. B. Query (POLDIP2 (51-368)) sequence alignment with homologous proteins sequence. Conserved amino acids are represented with +; a gap is shown for non-conserved amino acids; identical amino acids are represented by one letter amino acid code.

3.3.2.4.2. POLDIP2 (51-368) refinement in COOT

Coot/REFMAC was used for model building/refinement which involves the manual fitting of amino acids to the electron density map based on the prior knowledge of the protein's primary sequence (Emsley, Lohkamp, Scott and Cowtan, 2010). As a result, this process helps to optimise the calculated structure factor diminishing the mismatch between the calculated and visually observed differences in the structure factor (Emsley *et al.*, 2010). An electron density map which is used for the manual amino acid fitting is built by combining the phases (taken from the molecular replacement protein model) and amplitudes (taken from collected diffraction data) (Taylor, 2003).

Figure 26 shows POLDIP2 (51-368) real space refinement in COOT. In Panel A, Figure 26 the Glu215 does not fit into this electron density region, because Glu215 does not reach the positive electron density (yellow dashed circle 2), as well as, Glu215 leaves behind the clear positive electron density (yellow dashed circle 1). The Glu215 displacement (into yellow dashed region 1) allows the addition of Arg214, which successfully occupies the positive green electron density. As a result, both Glu215 and Arg214 fit into the green electron density allowing further real-space refinement and amino acid addition. Figure 26 was produced during the refinement process, thus amino acids which are not labelled in the figure were not yet modeled. For example (Panel A, Figure 26) the density region circled in violet indicates the presence of an aromatic side chain. Therefore, Phe, Tyr or His should be fitted into this electron density region (based on the POLDIP2 (51-368) amino acid sequence knowledge) replacing Leu.

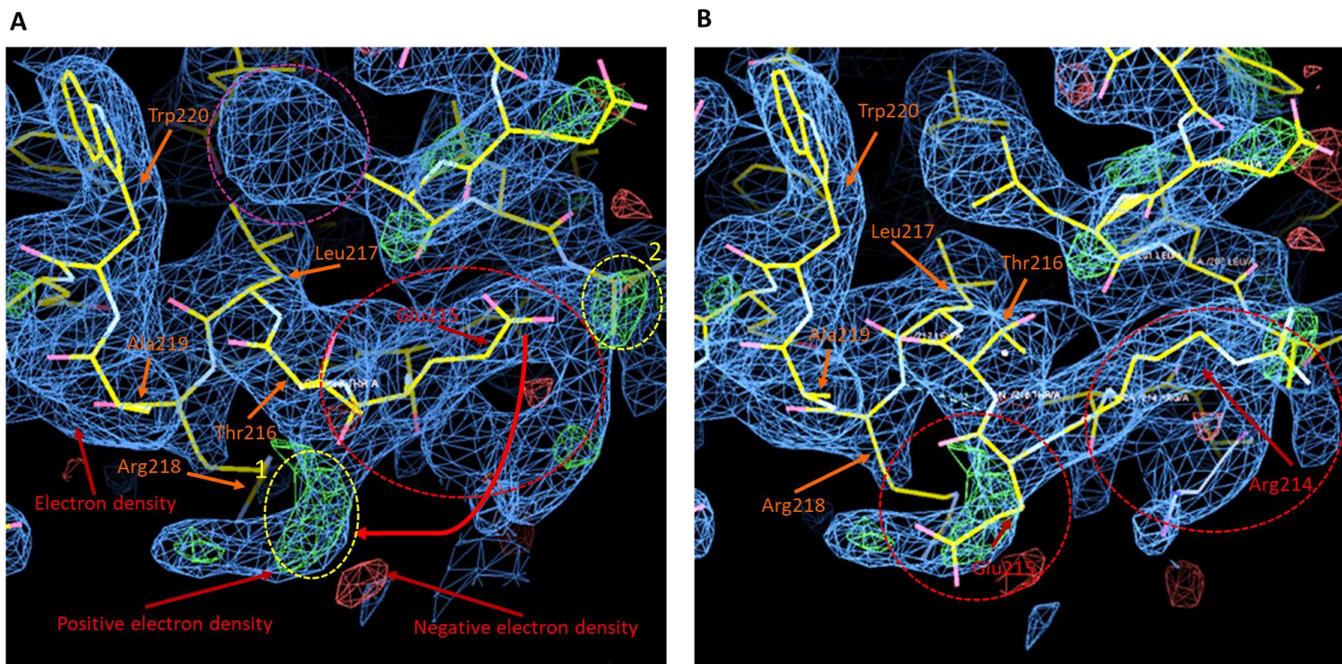


Figure 26. Example of Real-Space Refinement in COOT. Positive density on the Fo-Fc electron density map (green) indicates the correct positions of Glu215 and Arg214. The model was adjusted accordingly (red arrow). The α -helix has been extended by one residue (Glu215). Orange arrow points to amino acids which are modeled. On Panel A, 1 and 2 yellow circle dash line indicates areas with the positive electron density, where according to POLDIP2 (51-368) sequence Glu215 (1) and Arg214 (2) should be located. On Panel B, correctly positioned Glu215 and Arg214 are highlighted in red dashed circles. The electron density is shown in blue.

During real-space refinement process a strong positive electron density was identified near Cys266 (Figure 27). As cacodylic acid was present in POLDIP2 (51-368) crystallisation conditions, an assumption that it could react with Cys266 was made. This was further proved based on the knowledge that cacodylic acid can react with thiols (R-SH group), hence with protein cysteines (Jacobson, Murphy and Das-Sarma, 1972). Thereby, cacodylic acid was placed into the strong positive (green) electron density containing area attached to Cys266.

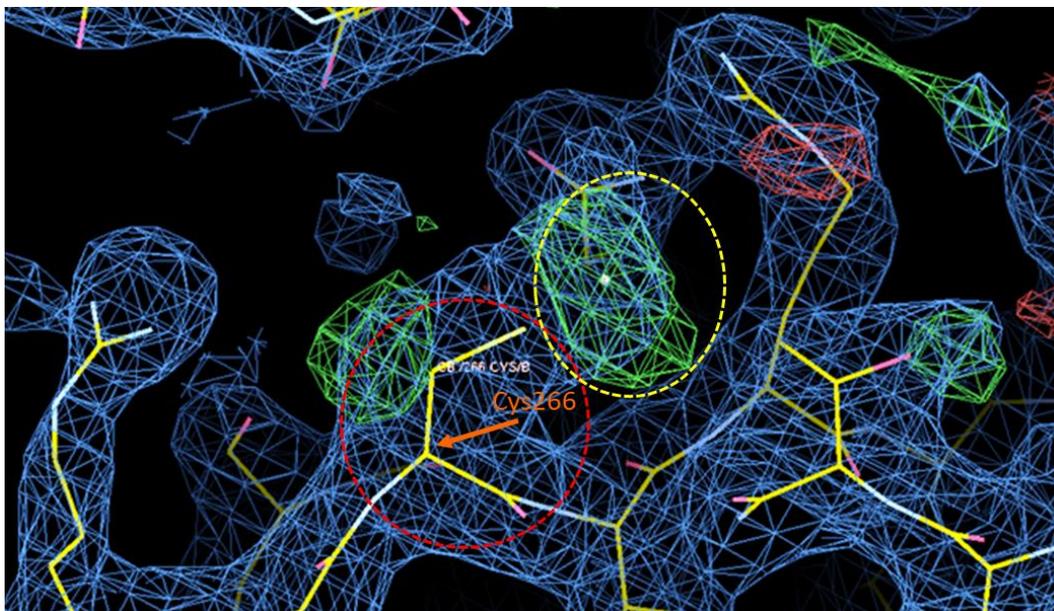
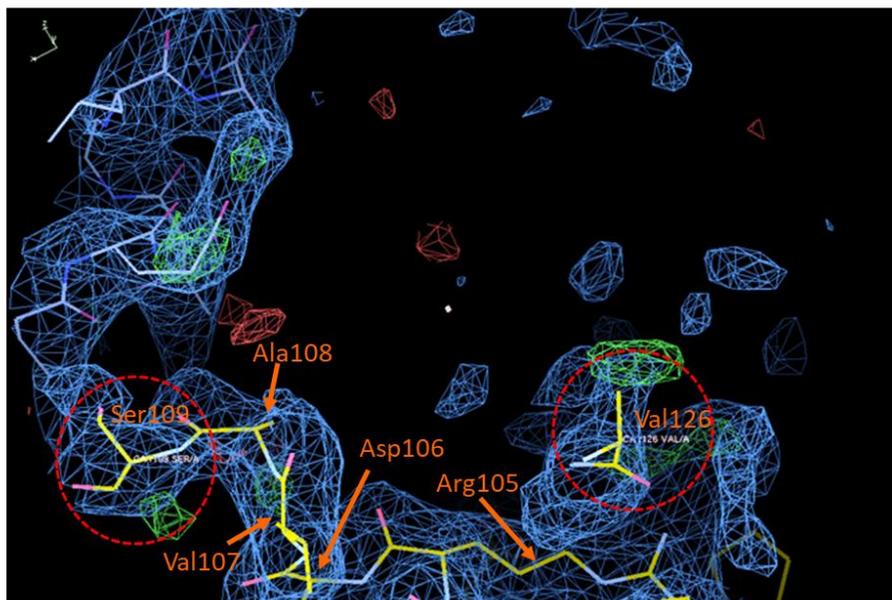


Figure 27. POLDIP2 (51-368) real-space refinement in COOT and the presence of positive electron density near Cys266. Positive density on the Fo-Fc electron density map (green) indicates the presence of electrons which should occupy this region. The region where cacodylic acid should be positioned is highlighted in yellow dashed circle. Cys266 to which cacodylic acid is attached is highlighted in red dashed circle. The electron density is shown in blue.

Loops often have weak/diffuse electron densities due to their mobile nature (Emsley and Cowtan, 2004). The POLDIP2 (51-368) structure is nearly fully resolved. However, the problem of modeling 2 loops arose due to poor electron density. Figure 28 was produced when all residues (except Asp141- Ser166 and Ser109-Val126 regions) were modeled into electron density regions. However, it is difficult to rebuild regions Asp141- Ser166 and Ser109-Val126 due to poor electron density and low resolution. The electron density in Panel A, Figure 28, is very poor (no electron density can be seen). The electron density in Panel B, Figure 28 is better as positive electron density regions are present (appeared after manual refinement in COOT and run of REFMAC5). However, it is still poor and it remains difficult to build amino acids accurately into this region (Murshudov, Vagin and Dodson, 1997). Although, in Panel B, Figure 28 positive electron density

appeared after refinement in COOT and REFMAC5, the observed electron density can be used only to build the proteins backbone. No electron density can be seen for the side chain location making it difficult to build correctly. However, the electron density in Asp141- Ser166 region (Panel B, Figure 28) could improve by further refinement in COOT and REFMAC5 making it possible to build Asp141- Ser166 region more accurately (Dr Richard Bingham *pers. comm.*).

A



B

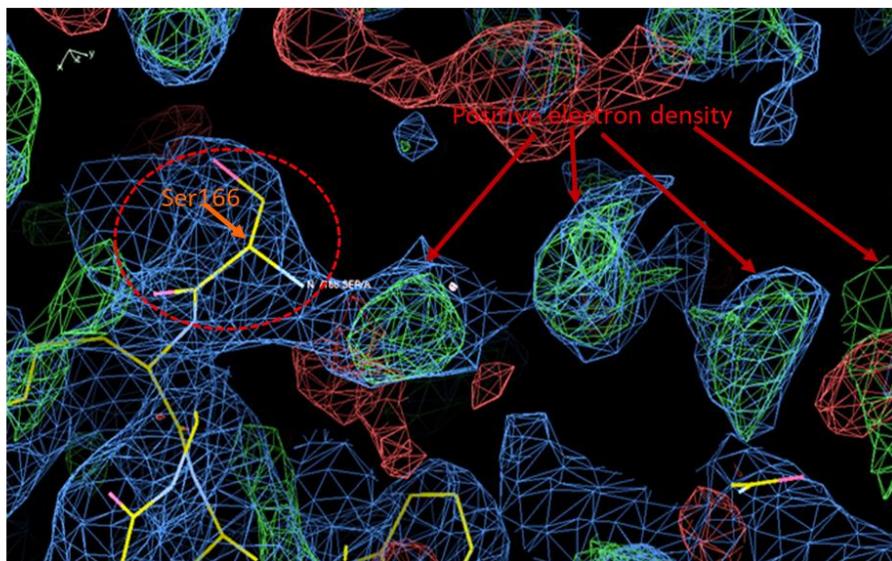


Figure 28. POLDIP2 (51-368) real space refinement in COOT and weak loop electron density. Panel A: Loop region between Ser109 and Val126 is disordered. There is no significant electron density on the map. Panel B: Loop between Asp141- Ser166 which is not yet modelled yet. Ser109, Val126 and Ser166 amino acids highlighted in the red dashed circle represent the loop region(s) after/before which modeling become impossible due to the poor electron density. Orange arrows points to the amino acid side chains. The electron density is shown in blue. Positive density on the Fo-Fc electron density map is shown in green.

3.3.3. Characterisation of the POLDIP2 (51-368) protein structure

3.3.3.1. POLDIP2 (51-368) secondary structure comparison between PRISPRED predicted and x-ray solved model.

PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) predicted POLDIP2 (51-368) region to contain 3 α -helices and 19 β -strands. 2 such strands are very short might be excluded due to inaccuracy of prediction for the short regions, namely, the second β -strand in the 60-70 residue region, as well as, β -strand in the 300-310 residue region (Panel A, Figure 29). 1 α -helix and 3 short β -strands located in POLDIP2 (1-50) will be not taken into account, as this region was not crystallised. In comparison to PRISPRED predicted secondary structure, the solved protein structure revealed POLDIP2 (51-368) secondary structure consists of 2 α -helices and 17 β -stands, which is similar to the PSIPRED prediction.

2 β -strands predicted by PSIPRED to be located in the 60-70 residue region are missing in the POLDIP2 (51-368) observed structure. However, the β -strand predicted to be located around residues 80-90 is present in both the resolved structure and PSIPRED predicted model (2 amino acids longer in solved structure). The first 2 β -strands in the 90-100 amino acid region predicted by PSIPRED actually form 1 long β -strand (β_2 in the Panel C). However, the β_3 strand observed in the solved POLDIP2 (51-368) structure was predicted correctly by PSIPRED in 100-110 residue region. The β -strand predicted by PSIPRED in the residue 120-140 region is also present in the solved structure (β_4 in Panel C, Figure 29), but the β -strand in the solved structure is shorter by 3 amino acids. β -strand in the 160-170 region and α -helix in 190-200 predicted by PSIPRED is missing from the solved structure. 2 observed short β -strands (β_5 ; β_6) and 1 short α -helix (α_1) and in the 170-190 region, as well as, 1 short β -strand (β_7) in the 200-210 residue range were not predicted by PSIPRED, but are present in the solved structure (Figure 29). An α -helix predicted to

be in the 190-200 residue region is not present in the solved POLDIP2 (51-368) structure. β -strand (β 8) in the 210-220 residue region is present in the observed and predicted structure. Also, the predicted α -helix (α 2) is present in the predicted and observed structure, but in the solved structure it is longer by approximately 9 amino acids. PSIPRED predicted α -helix in the 230-240 residue range, however in the solved POLDIP2 structure this region corresponds to the β -strand (β 9). Both PSIPRED predicted and POLDIP2 (51-368) solved structures have β -strand (β 10) in the 240-250 residue range, but in the PSIPRED prediction this β -strand is longer. β -strand (β 11) in the 250-270 residue region is also present in both predicted and solved structure, but this β -strand is longer in the PSIPRED predicted version. β -strand (β 12) located in the 270-290 residue region is present in both. β -strand (β 13) in the 290-300 residue region is also present in both, but it is longer in the observed structure. The short β -strand in 300-310 amino acid region is not observed in the solved structure. However, β -strand (β 14) is also present in both, but is much shorter in PSIPRED than in the solved structure. β 15 located in 320-340 region is present in predicted and observed sequences. β 16 located in region is present in both, but is shorter in the PSIPRED predicted version. Observed short β 17 strand in the 350-360 residue region is absent in the PSIPRED prediction. This result, suggests that PSIPRED can predict proteins structure to some extent, however in order to prove such predictions proteins structure should be solved by wet lab techniques (NMR, Cryo-EM, X-ray crystallography).

In the solved POLDIP2 (51-368) structure two domains can be seen. The N-terminal domain will be referred as domain A and C-terminal domain will be referred as domain B, respectively to the domain shown on the right and on the left (Panel C, Figure 29). Domain A has 2 α -helices, 8 β -strands. Domain B is composed of 9 antiparallel β strands. Domain A has higher loop content in the comparison to the domain B. Disulfide bonds were not observed in the POLDIP2 structure.

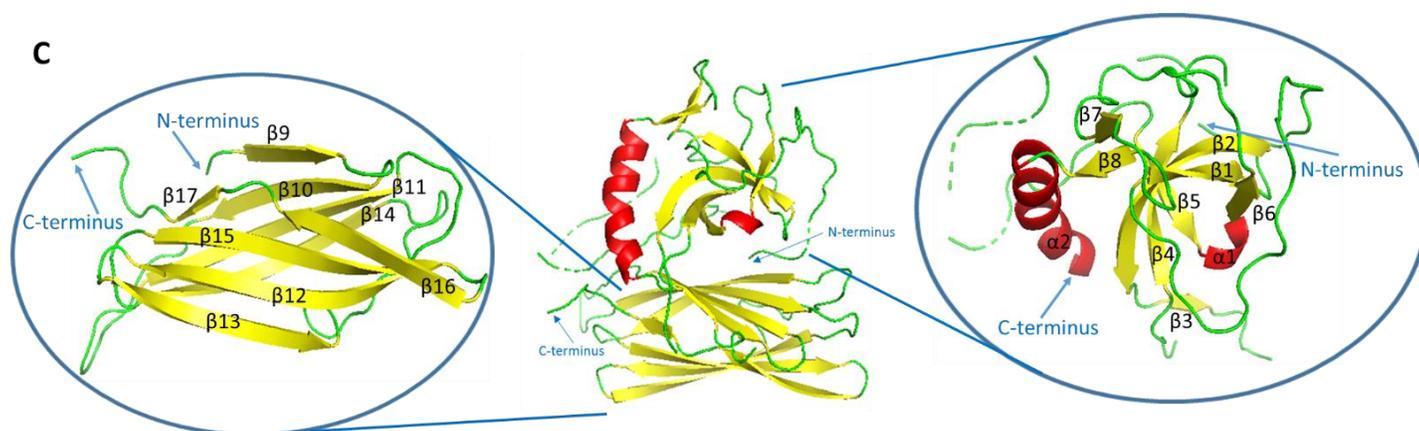
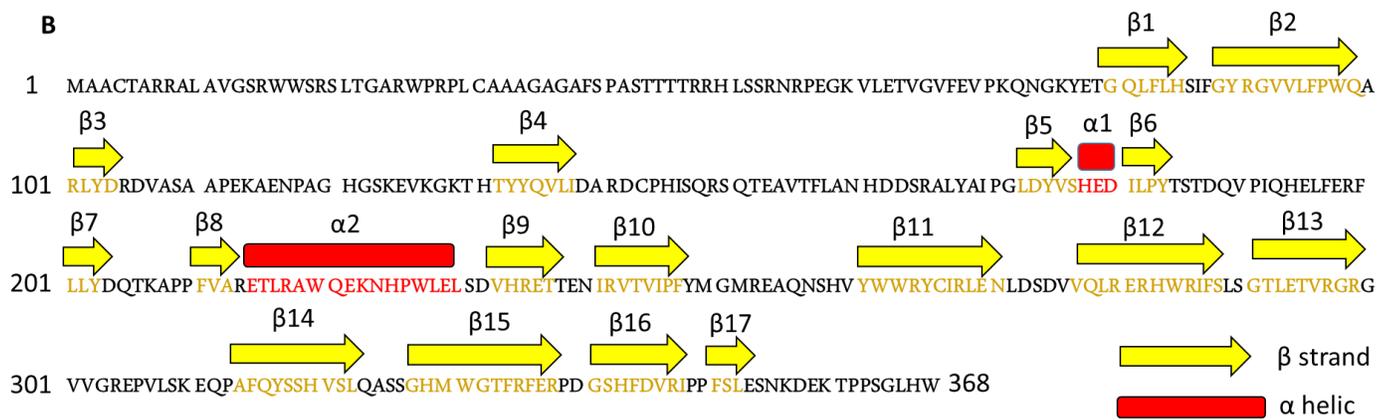


Figure 29. POLDIP2 predicted and obtained secondary structure composition comparison, followed by POLDIP2 3D structure cartoon representation. Panel A: predicted POLDIP2 (1-368) secondary structure motifs from the primary sequence by PSIPRED. Panel B: experimentally obtained POLDIP2 (51-368) secondary structure elements (X-ray crystallography). The secondary structure of the first 50 residues was not experimentally determined as this region was not crystallised. β -strands are labeled in yellow and α -helices are labeled in red. Panel C: POLDIP2 (51-368) 3D cartoon structure. Two POLDIP2 (51-368) domains are shown on the left and on the right side. POLDIP2 (51-368) structure is colored accordingly proteins secondary structure composition. Green represents loops; yellow represents β -strands and red color represents α -helices. Visually observed secondary structure elements in POLDIP2 (51-368) (Panel C) are numbered accordingly POLDIP2 (51-368) sequence starting from the N-terminus (Panel B).

3.3.3.2. POLDIP2 (51-368) secondary structure elements and bound molecules.

In Figure 30, POLDIP2 (51-368) with reduced loop content is shown. In domain A, 2 α -helices (very short $\alpha 1$ and quite large $\alpha 2$ helices) are shown (Panel A, Figure 30). Panel B, Figure 30 shows the cacodylic acid (CAD) bound to the Cys266 residue, which could be necessary for the POLDIP2 (51-368) structure stabilization. Panel C Figure 30, shows visible A and B domain interface. In A domain $\beta 3$ strand interacts with $\beta 9$ strand present in the B domain, forming double stranded parallel β -sheet. In POLDIP2 (51-368) structure the positively charged Na^+ ion is bound to the negatively charged Glu94 residue, potentially stabilizing the structure. Na^+ ion interaction with negatively charged Glu94 might be necessary to make POLDIP2 (51-368) molecules more hydrophobic. Neutralization of the net negative-charges in POLDIP2 (51-368) could weaken the repulsive intermolecular interactions between POLDIP2 (51-368) molecules leading to a protein crystallisation (Collins, 2004).

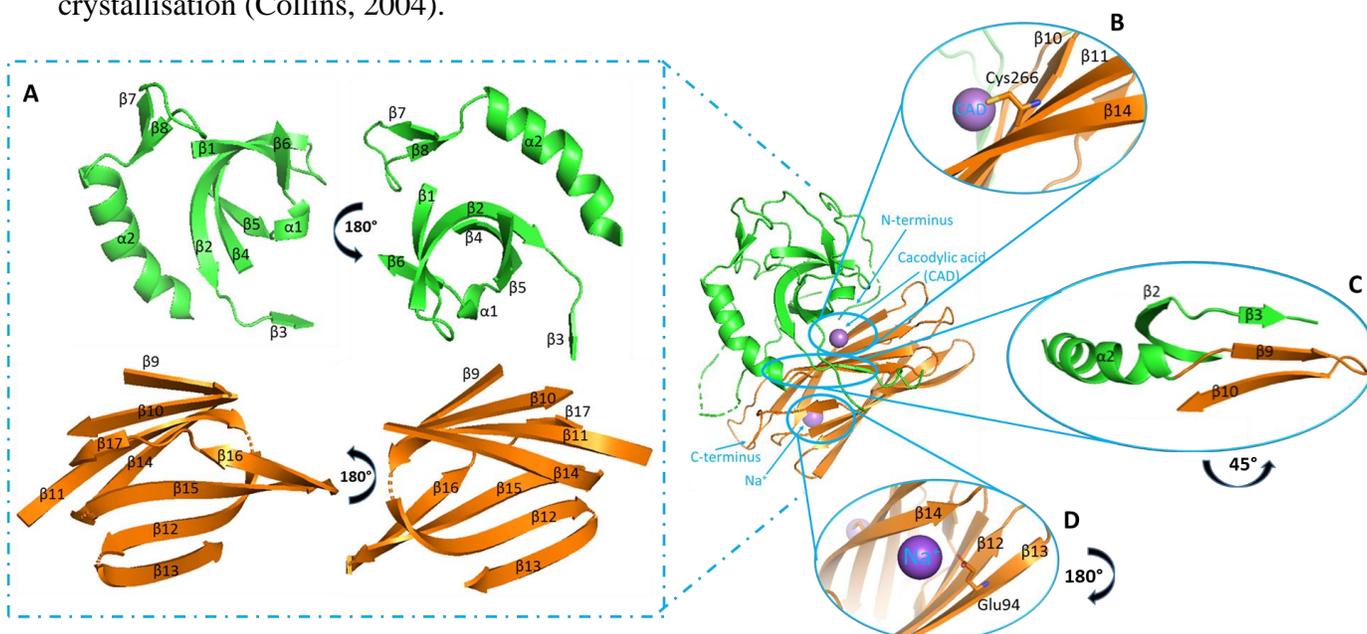


Figure 30. POLDIP2 (51-368) domain A and B cartoon domain representation. POLDIP2 (51-368) domains A and B are shown respectively in green and orange. A. Shows separated POLDIP2 (51-368) domains with a reduced flexible part content, emphasizing on the secondary structure elements (α -helices

and β -strands). B Cacodylic acid (CAD) bond to Cys266. C. Visible domain A and B interface. D. Na^+ bound to Glu94 residue.

3.3.3.3. POLDIP2 (51-368) labeling by *B* factor

B factors help to measure the flexibility/mobility which is associated with specific atoms (Henzler-Wildman and Kern, 2007). However, *B* factors cannot be interpreted simply as the amplitude of atomic fluctuations, as not only the intra-molecular motion, but crystal contacts with lattice disorder contribute to them (Henzler-Wildman and Kern, 2007). Generally, lower *B* factor values are observed for the β -sheet forming secondary structures as they are more stable and less flexible in comparison to protein loops (Emsley and Cowtan, 2004). This can be also seen in the POLDIP2 (51-368) structure colored by the *B* factor (Figure 31). On the whole, low *B* factors are observed for β -strands/sheets in the POLDIP2 (51-368) structure. Increase in mobility is observed in $\alpha 2$ helix in comparison to the β -strands. Higher flexibility levels are observed in loops (mostly yellow-red). This could suggest that $\alpha 2$ and loop dynamics might help in protein-protein interaction.

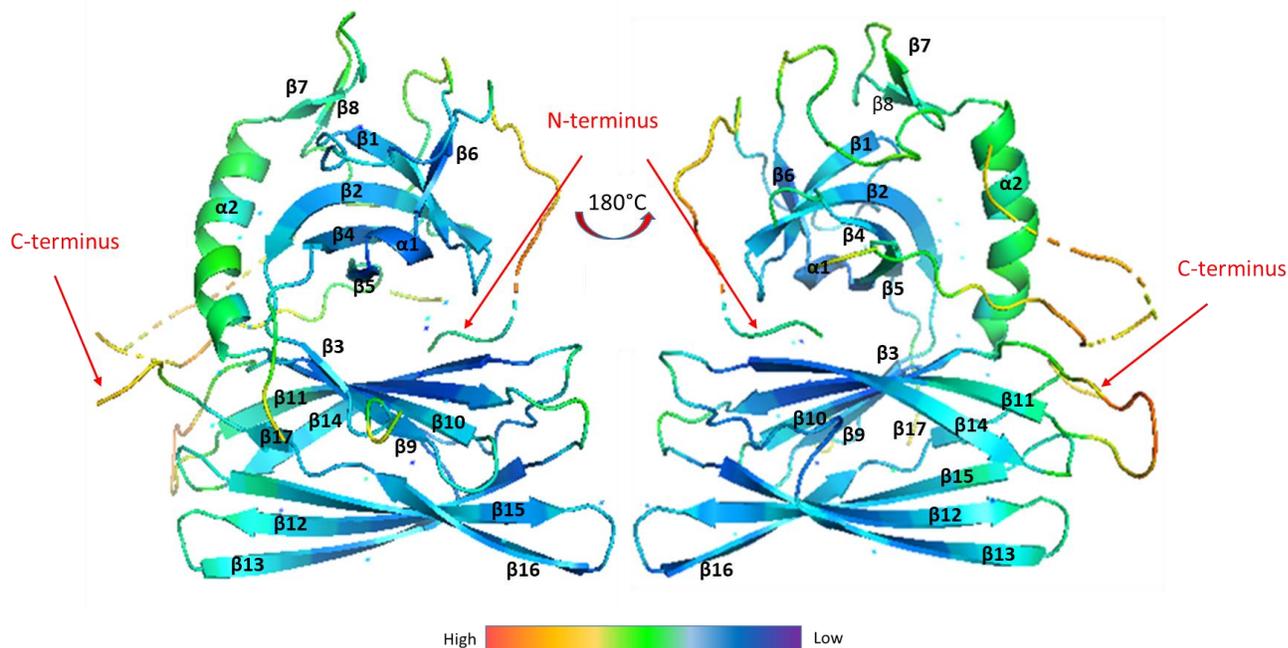


Figure 31. POLDIP2 (51-368) colored by *B* factor. Atoms with low temperature values are colored in blue and high temperature values are colored in red. Intermediate color such as light blue, green, yellow represent the scale of increasing temperature values.

3.3.3.4. POLDIP2 (51-368) surface charge distribution

POLDIP2 (51-368) surface charge distribution shows that POLDIP2 (51-368) could be involved in the interaction with proteins which have both/or negatively/positively charged surfaces (Figure 32), as proteins usually interact on the surface interface (Ma, Elkayam, Wolfson and Nussinov, 2003). From the Figure 32 below it can be seen that POLDIP2 (51-368) top and bottom areas mainly have red patches (POLDIP2 is negatively charged at these areas). Hence, these POLDIP2 parts could be involved in binding proteins which have a positive surface charge.

Also, POLDIP2 (51-368) has a positively charged spanning ring (Figure 32). This suggests that POLDIP2 (51-368) could bind (be wrapped around) circle shaped molecules with a negatively charged hole inside, as both proteins shape and chemical complementarity should be considered (Ivarsson and Jemth, 2019). Hence these areas could be involved in the binding with the negatively charged DNA/RNA, poly(ADP-ribose) or protein molecules (Lipfert, Doniach, Das and Herschlag, 2014; Chang, Jacobson and Mitchison, 2004; Bolden, Fry, Muller, Citarella and Weissbach, 1972). One POLDIP2 (51-368) side (Figure 32, 3rd orientation from the left) has mainly negatively charged surface (blue), suggesting that this area could be involved in the interaction with positively charged protein areas.

Additionally, some hydrophobic areas located at the interface between positively and negatively charged areas are present in POLDIP2 (51-368) suggesting its possible involvement in hydrophobic protein-protein interactions. Interestingly, strong protein interaction sites are usually

more hydrophobic, while sites involved in transient interactions usually tend to have more polar residues (La *et al.*, 2013). Hence, POLDIP2 (51-368) could be involved both in transient/weak and stronger protein-protein interactions, however from visual surface assessment POLDIP2 (51-368) has more charged surface areas than hydrophobic. This could reflect the ability of POLDIP2 to interact with multiple protein partners. Additionally, the charge distribution in PyMol is based only on the POLDIP2 amino acid sequence, so it does not take into account protein's crystallisation conditions, as well as, subtle pH changes which occur within protein grooves and pockets (Dr Georgios Psakis, *pers. comm.*). Therefore, POLDIP2 charge distribution will be analyzed with molecular dynamics simulation which predicts protein's surface charge more accurately, when protein's structure will be completely solved (Dr Georgios Psakis, *pers. comm.*).

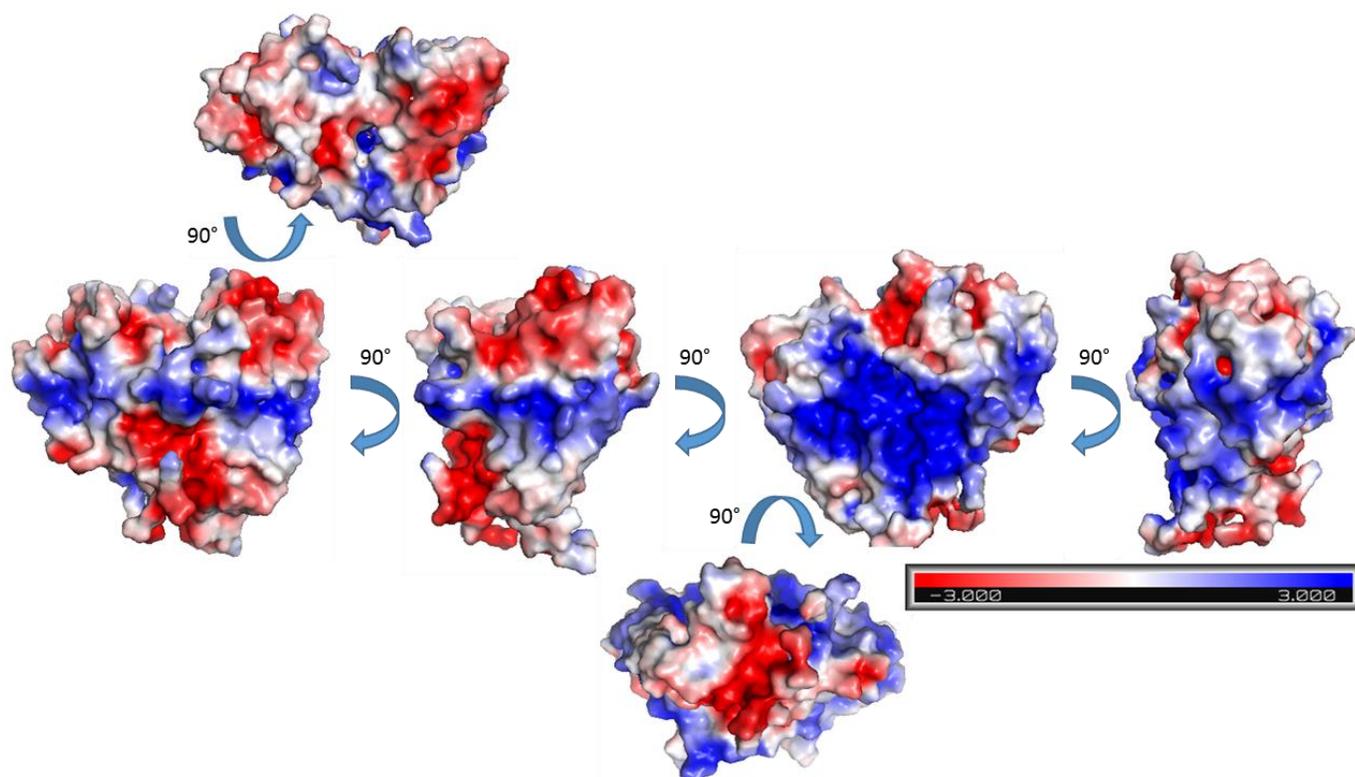


Figure 32. POLDIP2 (51-368) surface electrostatic potential. In blue are the negatively charged patches, in red are positively charged patches and grey corresponds to the hydrophobic patches. Surface electrostatic potential was calculated using PyMol.

3.3.3.5. POLDIP2 (51-368) statistics

A

| Data collection | |
|-----------------------------|--------------------------|
| Source | D8 Venture (Bruker) |
| Temperature (K) | 120 |
| Wavelength (Å) | 1.54 |
| Resolution range (Å) | 23.87–2.90 (3.06–2.90) |
| No. of measured reflections | 73526 (6303) |
| No. of unique reflections | 9222 (1326) |
| Multiplicity | 8.0 (4.8) |
| Completeness (%) | 99.7 (99.8) |
| Mean I/σ(I) | 15.8 (2.5) |
| R _{merge} | 0.132 (0.65) |
| Space group | P62 |
| a, b, c (Å) | 120.14, 120.14, 49.52 |
| α, β, γ (°) | 90.00, 90.00, 120.00 |
| Refinement | |
| Resolution range (Å) | 23.88–2.90 (2.976–2.900) |
| No. of reflections | 8763 (641) |
| Completeness (%) | 99.64 (99.55) |
| R factor | 0.21 (0.32) |
| R _{free} | 0.32 (0.47) |
| R.m.s.d., bond lengths (Å) | 0.008 |
| R.m.s.d., bond angles (°) | 1.648 |
| Ramachandran plot | |
| Preferred regions (%) | 233 (89.6) |
| Allowed regions (%) | 23 (8.9%) |
| Outliers (%) | 4 (1.5%) |
| B factors (Å ²) | |
| Mean B Value | 41.15 |
| From Wilson Plot | 33.91 |

B

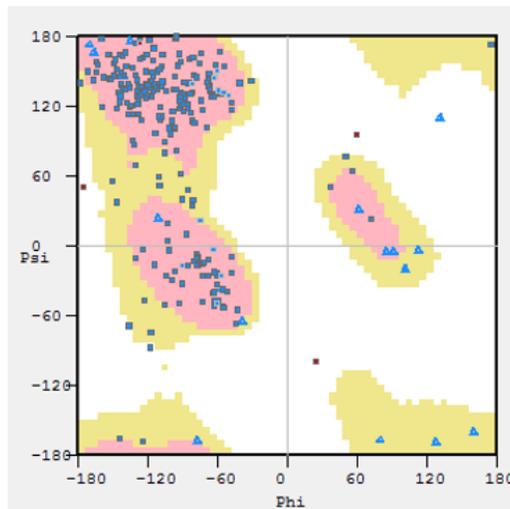


Table 10. Structural refinement statistics and Ramachandran plot. A. Data collection information, refinement statistics, Ramachandran plot and B factors values. R-factor calculated using equation $R = (\sum | |F_{obs} - |F_{calc} | |) / \sum |F_{obs} |$, R_{free} based on 5% of reflections calculated using the same equation. B. The Ramachandran plot. Psi and Phi angles in degrees are shown, respectively on y and x axis. Preferred regions are colored in pink, allowed regions are shown in yellow and disallowed regions are shown in the white background.

The statistics are within expected ranges for X-ray data of this resolution (2.9 Å). Most of residues (233) fall into a preferred region (89.6% are located in the pink area) (Table 10, Panel A and B). Small amount of amino acids (23) are located in allowed region (8.9% are in the yellow region) and only 4 amino acids are located in disallowed region (1.5% are in the white background) (Panel B, Table 10). However, POLDIP2 structure is only partially refined and will be further refined.

POLDIP2 structure is nearly fully solved (~99.8%). The Rwork and Rfree should decrease further with more cycles of refinement in COOT/REFMAC5 (Dr Richard Bingham *pers. comm.*).

3.3.3.6. Relating POLDIP2 (1-368) structure to protein interaction partners

PrimPol-POLDIP2 (1-368) crosslinking mass spectrometric studies showed that interaction was mediated by POLDIP2 residues: 1-8, 115-129 and 215-223 (Guilliam *et al.*, 2016). This suggests that multiple POLDIP2 (1-368) regions could be involved in the binding with PrimPol. The whole POLDIP2 (215-223) region corresponds to α 2-helix (Panel A, Figure 33). Another POLDIP2 V126-K129 region which interacts with PrimPol forms a loop (after β 4), however 115-126 region is missing in the figure as it is located in the flexible loop which was not yet modeled (Panel B, Figure 33). Loops and α -helices are more commonly involved in protein-protein interactions (account for 92%) than β -strand interactions (Yu and Guo, 2008). Interestingly, loops upon binding to protein can undergo disorder to order transition, gaining secondary structure (α -helix; β -strand) or can remain structurally heterogeneous (Ivarsson and Jemth, 2019). Also, such gain of secondary structure elements can increase proteins affinity for the binding partner(s) (Ivarsson and Jemth, 2019). In yeast two hybrid screens using truncated POLDIP2 and CAECAM1 constructs, it was revealed that C-terminal POLDIP2 residues 330-348 are sufficient to bind CAECAM1 (Panel C, Figure 33) (Klaile *et al.*, 2007). The use of truncated CAECAM1 constructs revealed that C-terminal CAECAM1 450-519 residues are involved in the interaction with POLDIP2 (Klaile *et al.*, 2007).

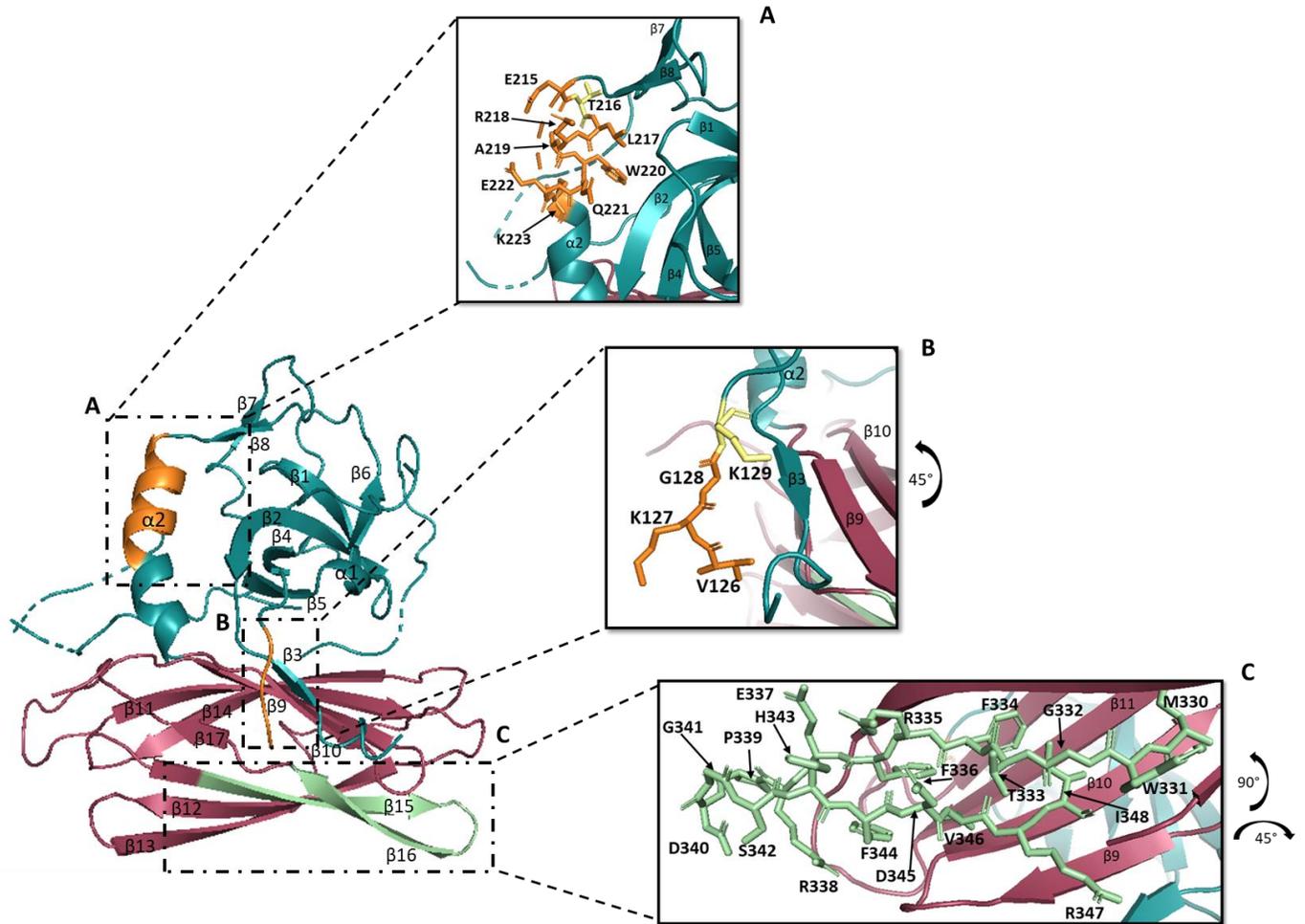


Figure 33. POLDIP2 (51-368) regions involved in PrimPol and CEACAM1 interaction represented as sticks. POLDIP2 (51-368) domain A is shown in blue and domain B is shown in purple. A-B Regions which are potentially involved in PrimPol-POLDIP2 interaction are shown in orange. K129 and T216 which were crosslinked with the PrimPol are highlighted in yellow. PrimPol K232 was bound to the POLDIP2 (1-368) T216. PrimPol T235 was bound to POLDIP2 (1-368) K129 (Guilliam *et al.*, 2016). POLDIP 1-8 region could not be shown as it not present in the solved POLDIP2 (51-368) expression construct. C. POLDIP2 (1-368) region corresponding to β 15-loop- β 16 structure involved in CAECAM1 binding is shown in green.

Liu *et al.*, (2003) predicted that POLDIP2 (1-368) could interact with PCNA via 3 motifs: Q81-F88 (Panel B, Figure 34), Q193-F200 (Panel C, Figure 34) and Q151-L158 (not modeled). Also, the G59 shown in Panel A, Figure 34 is predicted to be cleaved in the mitochondria (Liu *et al.*,

2003). Moreover, Panel E, Figure 34 shows the R299-V307 region which contains a potential NAD and FAD binding motif, as well as, the predicted DNA binding motif is shown as well (Panel D, Figure 34) (Liu *et al.*, 2003; Rossmann, Moras and Olsen,1974; Dym and Eisenberg, 2001). From the statistical point of view, conserved residues/regions are often important for the function, but it does not mean that POLDIP2 (1-368) will be able to bind such ligands (Halperin, Wolfson and Nussinov, 2004). For example, the binding could be not possible due to steric 3D hindrance, hence all bioinformatically/ statistically predicted assumptions must be verified using biochemical and biophysical analysis.

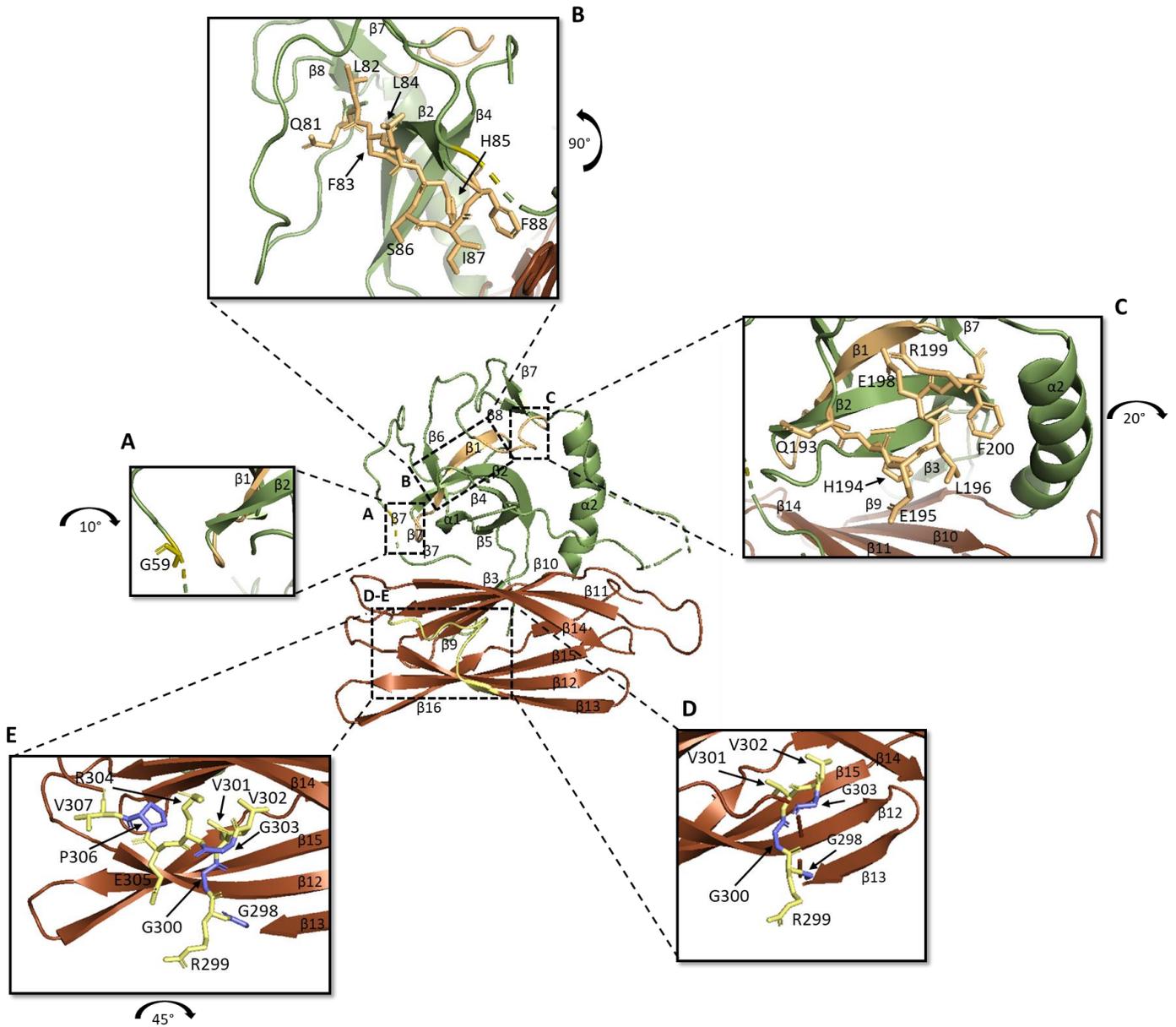


Figure 34. POLDIP2 (51-368) regions predicted by bioinformatics tools to be involved in the protein-protein interaction. Predicted regions involved in the binding are represented using stick model. A. The region between 58-59 amino acids is predicted to be cleaved. E. R299-V307 region contains a putative NAD and FAD binding motif. Conserved residues are highlighted in blue. D. R299-G303 region contains predicted DNA binding motif. Conserved residues are highlighted in blue. Panel B-C shows predicted PCNA-POLDIP2 interaction regions: Q81-F88 (Panel B) and Q193-F200 (Panel C). The third predicted interaction region Q151-L158 is not shown as it is missing from the structure.

3.4. Production of PCNA-POLDIP2 complexes

3.4.1. PCNA purification

3.4.1.1. PCNA IMAC purification

PCNA and POLDIP2 interaction was confirmed using the pairwise yeast two hybrid assays, pull down assay and western blotting (Liu *et al.*, 2003). Thus, further PCNA-POLDIP2 complex studies would help to uncover the mechanism of interplay between PCNA and POLDIP2 in regulating TLS polymerases and common binding partners.

Ni-NTA affinity chromatography was used for the initial human PCNA purification, using a C-terminally tagged 6xHis-FLAG PCNA fusion (from Dr Christopher Cooper). From the first purification gel (Panel A, Figure 35) it can be seen that 6xHis-tagged PCNA is bound to the column, while the majority of *E. coli* proteins successfully passed through the column. However, some bacterial proteins likely containing patches of histidines or even nonconsecutive histidines, as well as, functionally relevant metal binding sites which are located on proteins 3D surface can also be seen bound to the column contaminating our E300 1-3 samples (Panel A, Figure 35) (Frances, 1991; Cheung, Wong and Ng, 2012; Porath *et al.*, 1975).

After dialysis and 6xHis-tag removal, the PCNA protein was purified further (Panel B, Figure 35) and the major fraction of cleaved PCNA (FT) successfully passed through the column, but some cleaved PCNA molecules were still interacting with the column (40; 100; 300 mM imidazole) (Panel B, Figure 35). However, FT and 20 mM imidazole samples (Panel B, Figure 35) were reasonably pure, as the majority of *E. coli* proteins, as well, the TEV protease and FLAG tag remained bound to the column and did not elute, before Imidazole was added. As most contaminants remained bound to the column at this step, this suggested that only one additional

purification step is needed. Also, TEV cleavage was successful, as the change in proteins size from 30 kDa to 29 kDa can be seen. The 1 kDa difference corresponds to the 6xHis-FLAG tags size. The E300 fraction containing FLAG tagged PCNA (used as a control in Panel B, Figure 35) is located above all other fractions (FT-300 mM imidazole) which correspond to the successfully cleaved PCNA protein (Panel B, Figure 35). Samples for further purification were selected, balancing protein yield with purity. Although, 40, 100 and 300 mM imidazole fractions (Panel B, fig x) contained quite large quantities of PCNA, those were relatively impure, and thus were subjected for the further purification steps.

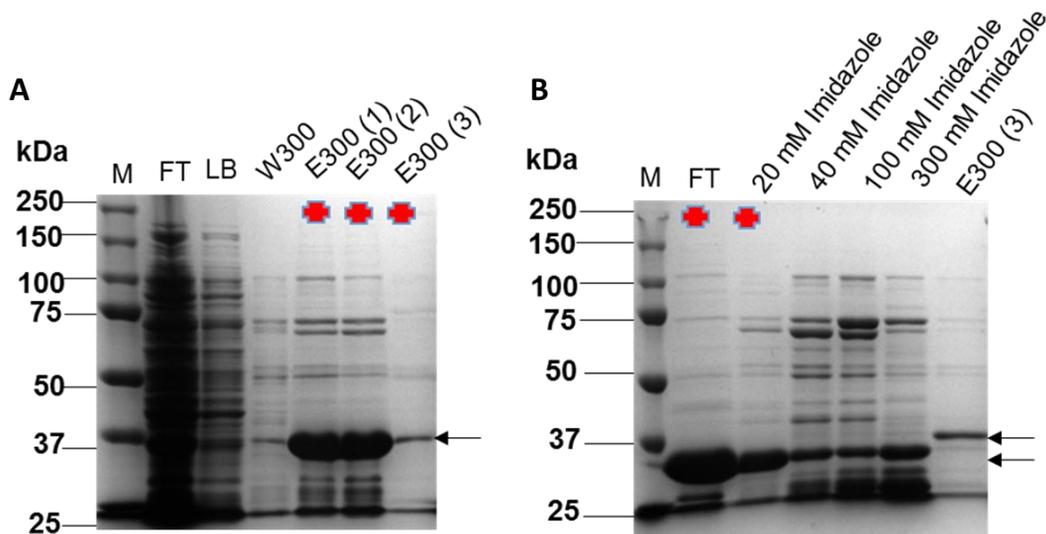


Figure 35. IMAC PCNA purification (4 L lysed). 20 μ L of sample was loaded onto the first IMAC gel (Panel A) and 5 μ L of protein sample was loaded onto the second gel (Panel B). Panel A shows the initial PCNA purification by Ni-NTA Affinity Chromatography. Panel B shows the second purification step, Ni-NTA Affinity Chromatography performed after dialysis and FLAG fusion-tag cleavage with TEV protease.

3.4.1.2. PCNA SEC purification

The size exclusion chromatography (Figure 36) was used as the final polishing step and resulted in further removal of contaminants suggesting that only minor contamination remained. It can be approximated that the gel is >95% pure, as the gel is overloaded. From the obtained calibration graph it was estimated/observed that the PCNA forms a trimer (observed 97 kDa; actual 89 kDa). Such, estimation is based on the observed MW from the calibration (Panel B, Figure 36) and calculated molecular weight (MW) from the protein sequence using 5 different molecular size markers of known MW. PCNA observed and calibration calculated MW ratio was 3.28 and that can be rounded to 3 (non-ideal behavior could be due to the PCNA torroidal shape which influences elution rate). Such results agree with March *et al.*, (2017), where they solved PCNA structure using X-ray crystallography and PCNA was show to form a trimer. Samples 15-19 were taken for further experiments, as the priority was given for the achievement of high proteins purification quality versus protein quantity (Panel A, Figure 36).

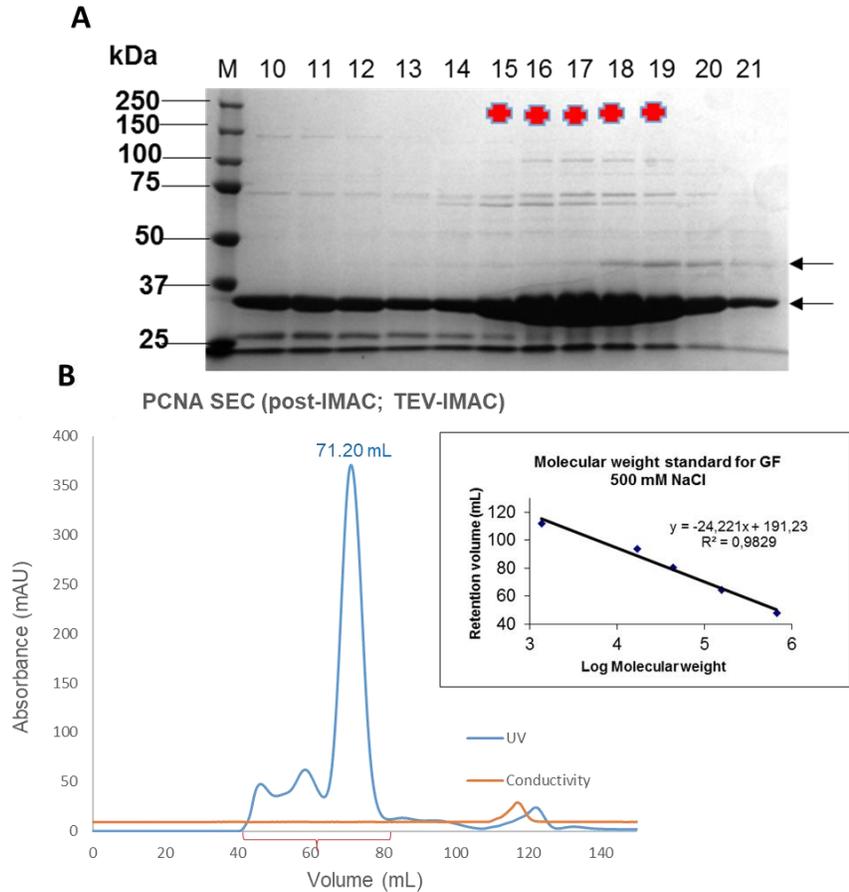


Figure 36. PCNA purification (4 L lysed). 5 μ L of the protein sample was loaded. (A) Gel represents the final purification step by size exclusion chromatography. (B) The chromatography obtained after SEC with the calibration curve inset. The calibration plot was made using molecular size markers consisting of 4 proteins for PCNA oligomeric state estimation. The 1350 Da protein in the marker was the outlier, hence unsuitable for the S200GF column used, therefore was excluded from the calibrations. Both protein and markers were run in the same SEC buffer.

3.4.2. Analysis of PCNA-POLDIP2 (1-368) complex formation by *in vitro* crosslinking

Crosslinking was performed to determine the oligomeric state of POLDIP2 following inconsistent SEC results (section 3.4.3.), as well as, to probe potential POLDIP2-PCNA protein-protein

interactions (Figure 37). Then BS3 and glutaraldehyde crosslinkers were used, in the gel (A) and (C) Figure 37, a weak band was observed between 75 and 100 kDa, suggesting that POLDIP2 (1-368, 42 kDa) alone could potentially form a dimer of ~84 kDa size. Also, between 75 kDa and 100 kDa in the same figure cross-linked PCNA (29 kDa) can be seen forming a trimer (~87 kDa). It is also supported by the analytical PCNA SEC data, which suggests that PCNA is a trimer. Moreover, Gulbis *et al.*, (1996) using X-ray crystallography confirmed that human PCNA is a trimer. Two bands can be seen in the PCNA trimer formation region possibly due to different PCNA forms (closed torroidal PCNA and open linear PCNA). The upper band could correspond to the PCNA trimer in which only two BS3 crosslinking molecules hold together 3 domains, as a result, making the molecule linear. The bottom band could be fully closed circular PCNA molecules, where PCNA domains are connected together via 3 BS3 cross-linkers, making the protein more compact and allowing to move faster through the gel. In the gel (B) the PCNA trimer and potential POLDIP2 (1-368) dimer location overlap. In the bottom part of the gel, high amounts of non cross-linked POLDIP2 (upper band) and PCNA (lower band) can be seen. In the gel (D) the POLDIP2 (1-368) (dimer) or PCNA (trimer) crosslinking was not observed. However, gel (D) and gel (B) have a band between 100 and 150 kDa, which could be the PCNA-POLDIP2 (1-368) complex formation in 1:3 ratio (129 kDa). In general BS3 cross-linked proteins of interest better than glutaraldehyde, but still, most of the PCNA-POLDIP2 (1-368) remained non cross-linked. This could result from the accessible surface lysine residues not being close enough for glutaraldehyde to modify and crosslink (spacer arm length is 5 Å), as BS3 has longer 11.4 Å spacer arm (Turk and Chapkin, 2015; Krohn, Yanagisawa and Grasser, 2002). A small fraction of PCNA-POLDIP2 (1-368) crosslinking could suggest the formation of a weakly-bound complex, but western blotting

or band excision followed by mass spectrometric identification would validate if this is a true POLDIP2-PCNA complex.

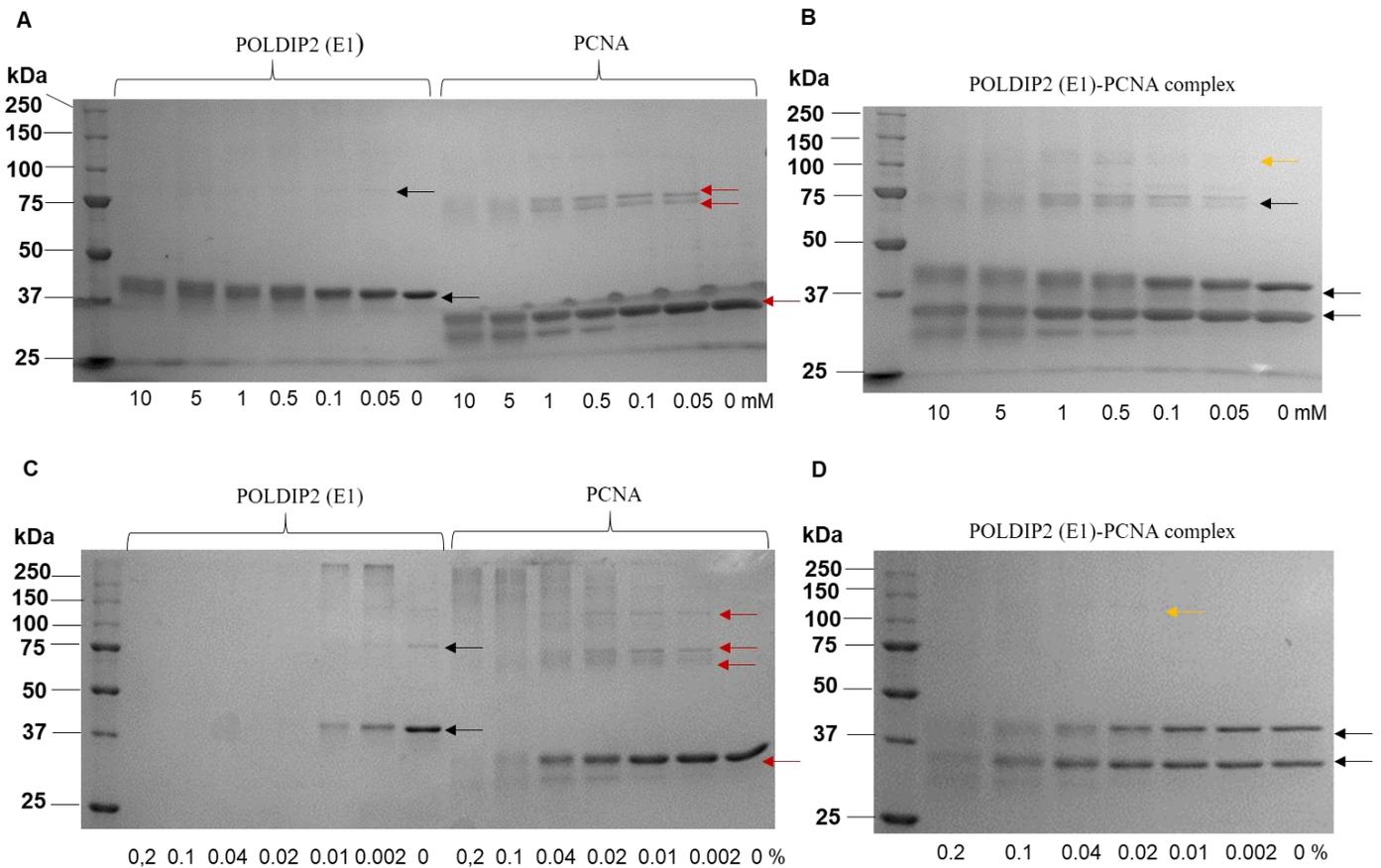


Figure 37. Investigation of POLDIP2-PCNA stoichiometry by *in vitro* crosslinking with BS3 and glutaraldehyde. Gel (A) and (B) represent the crosslinking with BS3 cross-linker. Gel (C) and (D) represent the crosslinking with the glutaraldehyde crosslinker. Samples were cross-linked with decreasing concentrations of the cross-linker. The final cross-linker concentrations are shown on the bottom of the gel. Black arrows indicate the non cross-linked or crosslinked (dimer) POLDIP2, red arrows indicate non-crosslinked or crosslinked (trimer) PCNA and orange indicate potential POLDIP2/PCNA complex.

3.4.3. PCNA-POLDIP2 (1-368) complex co-purification

As POLDIP2 crosslinking with PCNA suggested complex formation an effort to reconstitute this complex biochemically using co-purification SEC, further proving that POLDIP2 interacts with the PCNA was made. Low salt concentration was used in order not to interfere with protein-protein

complex formation by causing electrostatic interaction disruption (Curtis, Montaser, Prausnitz and Blanch, 1998).

POLDIP2 (1-368) protein which does not bind the PCNA at 100 mM NaCl, is lost after SEC due to high precipitation at low salt concentration. Also, the PCNA-POLDIP2 (1-368) binding stoichiometry is not clear, as in samples 9-17 it could be possible that complex forms at 3:1 ratio, respectively for the PCNA and POLDIP2 (1-368) proteins. In this case a PCNA trimer could bind one POLDIP2 (1-368) molecule (Gulbis *et al.*, 1996). Such, estimation is based on the PCNA-POLDIP2 observed MW (120 kDa) from the calibration graph (using 4 different protein markers of known size) and calculated molecular weight MW (129 kDa) from the protein sequence (Panel B, Figure 38). PCNA-POLDIP2 complex observed and calibration calculated MW ratio was 0.93 and that can be rounded to 1 (non-ideal behavior could be due to the PCNA-POLDIP2 complex shape which influences elution rate). However, visual examination of 18-22 samples suggests that protein interaction occurs at 1:1 ratio. Thus, further studies are required to explain PCNA-POLDIP2 (1-368) stoichiometry. Likewise, it is also possible that PCNA-POLDIP2 (1-368) interaction could be transient. In such case if PCNA-POLDIP2 (1-368) interaction is transient/weak the POLDIP2 (1-368) protein which unbinds from the PCNA could precipitate (fall out of the solution) making it impossible to rebind the PCNA molecule again, as it will be lost.

Interestingly, when PCNA-POLDIP2 (51-368) were co-purified (data not shown) two peaks were present on the SEC chromatograph and gel clearly showed two distinct regions for the PCNA and POLDIP2 (51-368). Such result suggests that N-terminal POLDIP2 (1-368) terminus is required for the interaction with PCNA. Also, this data supports the Guillian *et al.*, (2016) data, where it was showed using crosslinking that the N-terminal POLDIP2 1-8 region is involved in the protein binding (PrimPol). Moreover, although POLDIP2 (1-368) expression level is high, upon TrxT tag

removal POLDIP2 (1-368) precipitates and after the final purification step there is approximately 40 μ l (5 mg/ml) remaining. This suggests that POLDIP2 (51-368) is more rigid and stable due to the lack of N-terminal region as in comparison, high quantities of POLDIP2 (51-368) were obtained even after TrxT tag removal. The purified PCNA-POLDIP2 (1-368) complex from the same preparation (shown in Figure 38) was taken for the further crystallisation and cryo-EM studies.

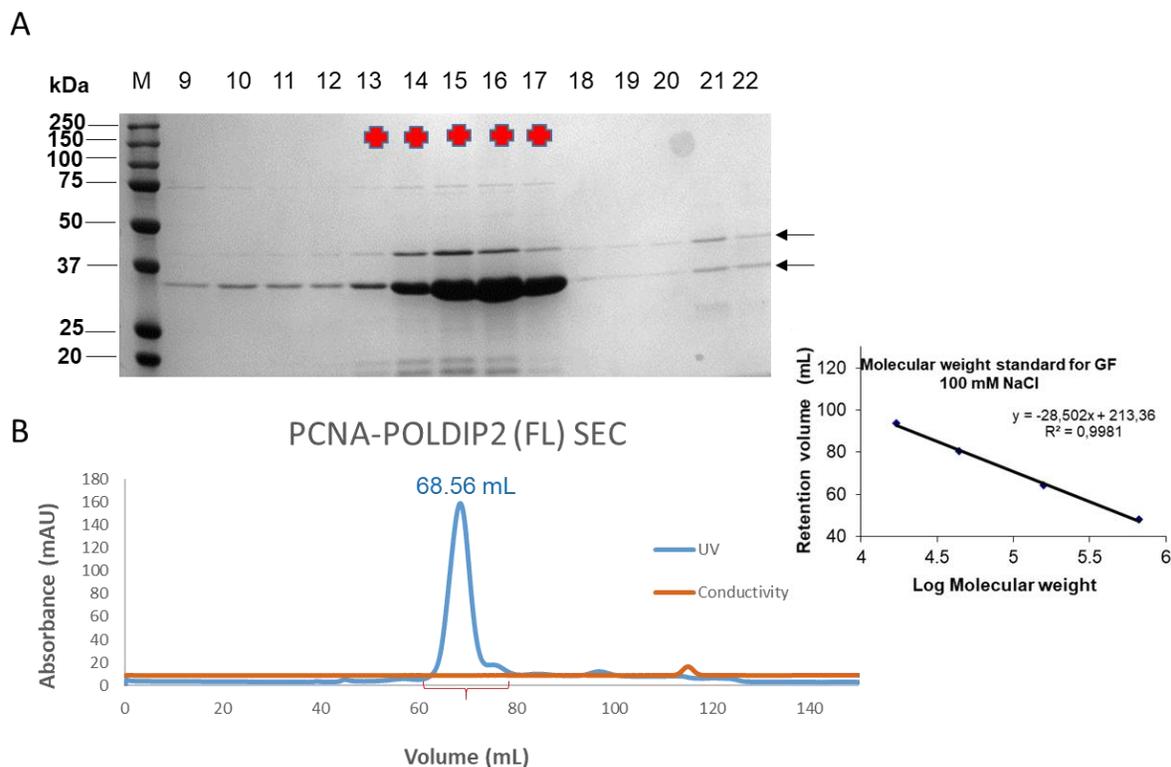


Figure 38. PCNA and POLDIP2 (1-368) SEC. 5 μ L of the protein sample was loaded. A. Gel represents the PCNA-POLDIP2 (1-368) co-purification step by size exclusion chromatography. B. The chromatograph obtained for PCNA-POLDIP2 (1-368) potential complex after SEC (100 mM NaCl). The calibration plot was made using molecular size markers consisting of 4 proteins for PCNA-POLDIP2 complex oligomeric state estimation. Both proteins and markers were run in the same SEC buffer.

3.4.4. PCNA-POLDIP2 complex studies

Usually protein co-crystallisation results in better crystallisation in comparison than if full length protein is attempted to be crystallised alone (McPherson, 2004). Protein-protein interaction may stabilize the desired protein, while protein alone might be more flexible and not stable complicating the process of crystal formation process (McPherson, 2004). Moreover, protein complex can yield us the detailed information regarding protein-protein interaction allowing to understand such process/mechanism in details (McPherson, 2004). Therefore, as POLDIP2 (51-

368) contains loops which were not modeled due to poor electron density it would be useful to co-crystallize the protein in order to obtain the better electron density for these loop regions. Additionally, POLDIP2 (1-368) could be more stable in complex structure than alone leading to the crystal formation. Additionally, POLDIP2 (1-368) structure in complex could lead to the better understanding of POLDIP2 (1-368) ability to bind different proteins. Moreover, if interaction occurs via flexible loops, upon binding such loop could adopt secondary structure, potentially making it easier to crystallise (Tompa and Fuxreiter, 2008).

Liu *et al.*, (2003) using yeast two hybrid screen, co-immunoprecipitation, GST pull down, immunoaffinity chromatography and native gel showed that POLDIP2 interacts with PCNA. As PCNA was the most studied POLDIP2 binding partner this protein was selected for further PCNA-POLDIP2 complex studies. Sitting drop crystallisation technique was applied in this experiment and two screens were used JCSG/Morpheus, with potential PCNA-POLDIP2 (1-368) complex in the upper well and PCNA alone in the lower well. Nearly all drops remained clear, suggesting the insufficient protein concentration in JCSG screen, while nearly all conditions in Morpheus screen resulted in protein precipitation. Figure 39, A and B show the presence of nucleation only in the well which contains PCNA-POLDIP2 (1-368) complex, but the drops which have only PCNA remained clear (Table 11). C and D have crystals for PCNA-POLDIP2 (1-368) complex, but not for the PCNA alone. Figure 39, C and D crystals with undefined shape were observed after 1 month. This is likely to be due to water evaporation, which increased protein concentration leading to the crystal contact formation (nucleation) and crystal growth (Luft, Wolfley and Snell, 2011). These crystals will be mounted and screened in the near future.

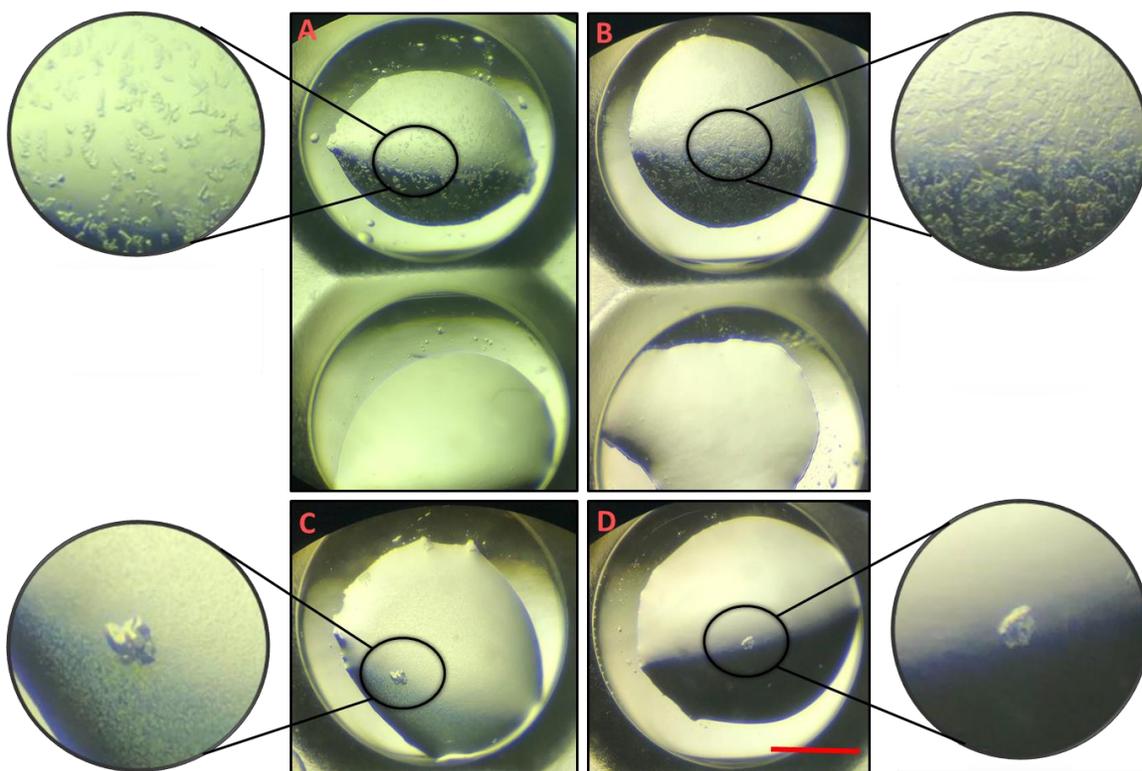


Figure 39. PCNA-POLDIP2 (51-368) crystals. Figures of obtained PCNA-POLDIP2 (1-368) complex crystals. The upper well contained PCNA-POLDIP2 (1-368) and the lower well contained only PCNA which was used as the control. For PCNA-POLDIP2 (1-368) and PCNA 5.15 mg/ml protein concentration was used.

Table 11. PCNA-POLDIP2 (1-368) crystals. Conditions under which the potential PCNA-POLDIP2 (1-368) complex could be formed shown in Figure 39.

| Number | Screen | Salt | Conc. Buffer | pH | Precipitant |
|--------|--------|--------------------------------------|-------------------------|-----|------------------|
| A | JCSG | 0.2 M Magnesium formate dihydrate | None | | 20% w/v PEG 3350 |
| B | JCSG | 0.2 M Magnesium chloride hexahydrate | 0.1 M Bis-Tris | 5,5 | 25% w/v PEG 3350 |
| C | JCSG | 0.2 M Calcium acetate hydrate | 0.1 M Sodium cacodylate | 6,5 | 40% v/v PEG 300 |
| D | JCSG | 0.2 M Magnesium chloride hexahydrate | 0.1 M Sodium HEPES | 7,5 | 30% v/v PEG 400 |

3.4.5. Negative staining EM for PCNA-POLDIP2 (1-368) complex, towards cryo-EM.

Prior to negative stain EM analysis, POLDIP2 was mixed with PCNA and incubated for 3 hours. The final SEC was run for the complex resulting in the single peak in the chromatogram, which was further confirmed by the SDS-PAGE gel, which proved the appearance of PCNA and POLDIP2 in the same fractions and the lack of other contaminants (same preparation was used as for the complex crystallisation studies (Figure 40).

The proteins solution is likely to be heterogeneous, as no homogenous proteins or PCNA-POLDIP2 (1-368) complex can be seen. Protein degradation and soluble aggregation (the sample was centrifuged just before applying it on the grid) may potentially be seen on the grid. Protein particles/proteins can be differentiated from the background by eye. However, the image contrast was not very good judging by the apparent contrast due to the glycerol presence in the buffer which strongly reduces the image background, thus making it incompatible with the cryo-EM grid preparation and further studies (Renaud *et al.*, 2018).

However, the grid shows good particle distribution. Particles are distributed randomly and are present at different parts of the grid being not centred just in the grid's edge, suggesting satisfactory protein, as well as, stain concentration (good blotting and drying time was applied) (Drulyte *et al.*, 2018).

Also, molecules adopt random orientations resulting in different projections which are required for the further computational analysis. However, such molecules cannot be identified on Negative stained cryo-EM micrographs (Drulyte *et al.*, 2018). However, a molecule highlighted in the red circle could be PCNA, due to the visible hole in the middle. Gulbis *et al.*, (1996) confirmed using X-ray crystallography that it is a trimer which has hole in the middle. Further studies will be performed to optimise negative staining, to attempt to structurally determine POLDIP2 with at least one protein binding partner.

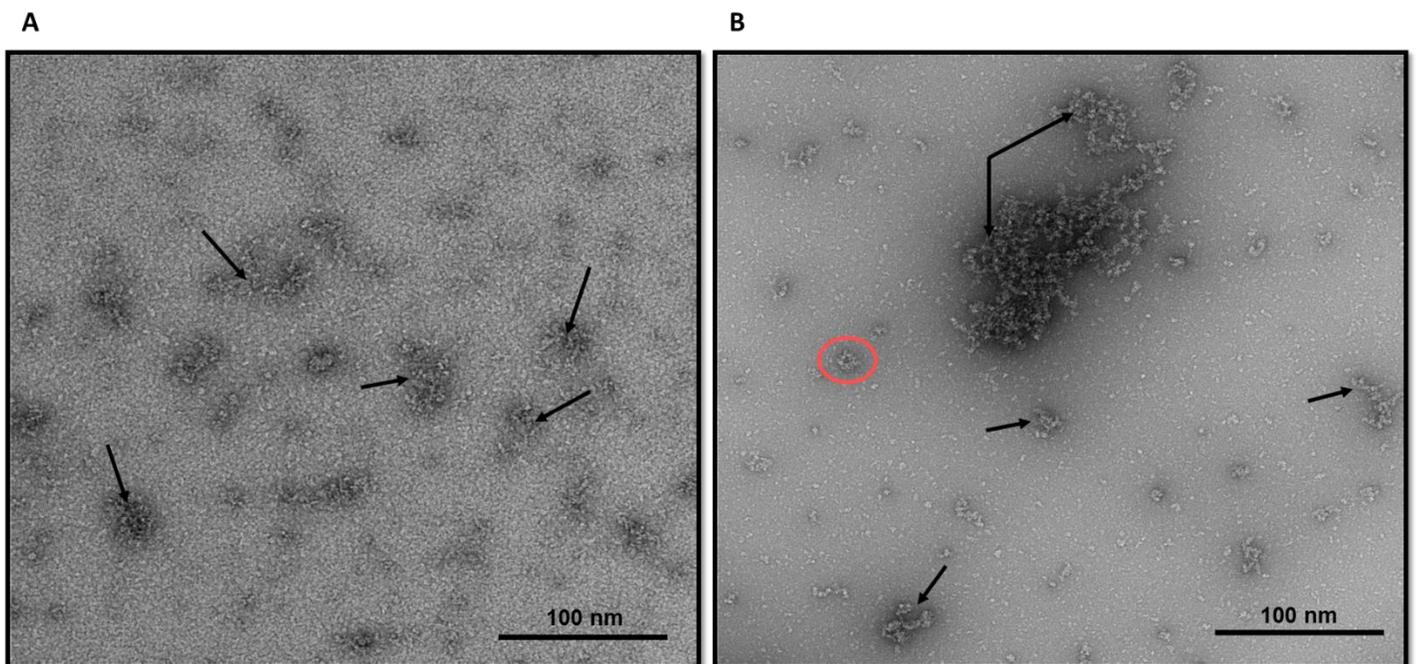


Figure 40. Negative stained cryo-EM micrographs of potential POLDIP2-PCNA complex. A. Due to the high contamination, it was impossible to observe potential POLDIP2-PCNA complex. B. The same protein as on the left diluted 10x. A potential PCNA molecule is highlighted in the red circle. The black arrow shows potential protein soluble aggregation. Scale bar represents 100 nm in length.

4. Conclusion

4.1. Conclusion and future work suggestions

In conclusion, full length and truncated POLDIP2 constructs were successfully sub-cloned into a new pGTVL2 vector containing 6xHis-GST fusion tag. 12 POLDIP2 constructs tagged with TrxT tag and 12 POLDIP2 constructs tagged with GST tag were expressed in TB/LB media. After results were analyzed three POLDIP2 constructs were selected for further studies: POLDIP2 (1-368), POLDIP2 (51-368) and POLDIP2 (178-368). Both POLDIP2 (51-368) and POLDIP2 (178-368) were subjected to crystallisation trials, however, no crystal formation was observed for the POLDIP2 (178-368) construct. Nonetheless, using commercial crystallization kits five conditions led to POLDIP2 (51-368) crystal formation. As POLDIP2 (51-368) crystallization optimization did not yield crystals with better diffraction in comparison to the first obtained crystal from the commercial screening kit, an attempt to solve POLDIP2 (51-368) structure was made using the crystal from the commercial kit. The best crystal diffraction pattern was obtained at 20°C under the following conditions: 0.2 M Calcium acetate, 0.1 M Sodium cacodylate, 40% v/v PEG300, pH 6.5.

POLDIP2 (51-368) structure is ~99.8% complete, lacking Asp141-Ser166 and Ser109-Val126 regions. With further refinement in COOT/REFMAC5 it could be possible to build the lacking amino acids into the electron density map. Latterly, positive electron density appeared for the Asp141-Ser166 region, thus an attempt to build this region could be made. Also, although multiple POLDIP2 (51-368) optimisation crystals were already screened and did not yield any better diffraction, some remained unscreened crystals should be checked, perhaps using an X-ray source stronger than the in-house source used herein, such as the Diamond synchrotron (Walter *et al.*,

1995). Moreover, data from the POLDIP2 (51-368) crystal which had the best diffraction was collected, and an attempt was made to solve its structure. X-ray crystallography lacks information about dynamic behavior, as it is a static snapshot which does not represent native conditions (Fenwich *et al.*, 2014). Therefore, NMR could be used to study POLDIP2 protein dynamics and conformational cycles, revealing the major conformational states (Fenwich *et al.*, 2014). Also, the chemical shift difference between different conformations can be taken into account and the interconversion time between different conformations can be estimated on a wide range of timescale giving the quantifying indication of protein's motion (Fenwich *et al.*, 2014; Fenwick, Bedem, Fraser and Wright, 2014). Additionally, as high-resolution NMR can be used for visualization of the flexible region, it would be good to solve POLDIP2 (1-368) structure using this technique. In such case, NMR could allow to solving POLDIP2 (51-368) unstructured Asp141- Ser166 and Ser109-Val126 regions, superimposing NMR data with already obtained POLDIP2 (51-368) x-ray solved data.

As POLDIP2 (1-368) has a predicted flexible N-terminus and is less stable in the buffer than POLDIP2 (51-368) or POLDIP2 (178-368) an attempt was made to crystallise POLDIP2 (1-368) in complex with PCNA. Before PCNA-POLDIP2 (1-368) complex crystallisation these proteins were SEC co-purified and *in vitro* cross-linked. SEC and *in vitro* cross-linking data suggest that PCNA-POLDIP2 (1-368) could potentially form a complex. However, these experiments did not reveal a definitive PCNA-POLDIP2 (1-368) complex stoichiometry, suggesting that complex formation might be transient. Therefore, the following techniques can be used in the future to reveal POLDIP2 (1-368) or PCNA-POLDIP2 (1-368) complex oligomerization state: mass spectrometry (on cross-linked samples), small-angle x-ray scattering, multi-angle light scattering analysis combined with size exclusion chromatography, analytical ultracentrifugation, dynamic

light scattering and isothermal titration calorimetry (Korasic and Tanner, 2018; Wohlgemuth, Lenz and Urlaub, 2015; Gell, Grant and Mackay, 2012). Also, in order to test if PCNA-POLDIP2 (1-368) association is transient surface plasmon resonance, fluorescence polarization and isothermal titration calorimetry could be used (Vedadi *et al.*, 2010). PCNA-POLDIP2 (1-368) SEC co-purification and *in vitro* crosslinking suggested complex formation, hence the co-purified proteins were set to crystallization trials. Using PCNA as a control several conditions were identified where the potential PCNA-POLDIP2 (1-368) crystal complex could be formed. These crystals will be screened in the near future. Also, an attempt was made to assess the formation of PCNA-POLDIP2 (1-368) complexes using negative staining, towards cryo-EM structural determination. However, the obtained results showed that complex was not stable even though the buffer contained glycerol. Thus, the buffer should be optimized (get rid of glycerol) in the future to make the complex more stable and suitable for the cryo-EM studies (Thompson *et al.*, 2016).

Experimentally-determined and bioinformatically predicted POLDIP2 protein interaction regions were analysed in the resolved POLDIP2 (51-368) structure. The related experimentally-determined regions involved in the protein-protein interaction turned out to be $\alpha 2$, $\beta 15$ -loop- $\beta 16$ and loop (followed after $\beta 4$ strand). The predicted POLDIP2 (1-368) binding proteins are located in different cellular compartments (Introduction section, Figure 3). This suggests, that POLDIP2 (1-368) function could be different in different cellular compartments. Also, compartmentalisation eliminates the presence of some potential binding partners and even higher POLDIP2 binding specificity could be achieved due to the presence of different POLDIP2 isoforms (Stein, Pache, Bernadó, Pons & Aloy, 2009; Perkins, Diboun, Dessailly, Lees and Orengo, 2010). It is also possible that PCNA-POLDIP2 (1-368) complex could recruit third common binding partner

(Figure 41). Interestingly, POLDIP2 (1-368) is not the only protein which is present in different cellular compartments and is able to bind PCNA. The Apurinic/aprimidinic (AP) endonuclease 2 (APE2) protein is also present in both nuclear and mitochondrial compartments and can bind the PCNA (Tsuchimoto *et al.*, 2001). To find new binding partners: yeast two-hybrid screen and affinity blotting can be used (Phizicky and Fields, 1995). Also, *in vivo* crosslinking, co-immunoprecipitation, western-blot/mass spectrometry could help to identify new POLDIP2 (1-368) or POLDIP2 (51-368) binding proteins, verifying the POLDIP2 (1-368) N-terminal importance in protein-protein interaction (Phizicky and Fields, 1995). Likewise, POLDIP2 (1-368) and POLDIP2 (51-368) construct function can be studied in different cellular compartments separately, as POLDIP2 is found in nucleus, cytoplasm and mitochondria. Moreover, POLDIP2 (1-368) could be studied in the eukaryotic organism, as POLDIP2 (1-368) protein is a human protein and hence could be *in vivo* post-translationally modified. Post-translational modifications could be one of the POLDIP2 (1-368) protein-protein interaction controlling points, suggesting that POLDIP2 (1-368) could bind different proteins depending on the post-translational modification.

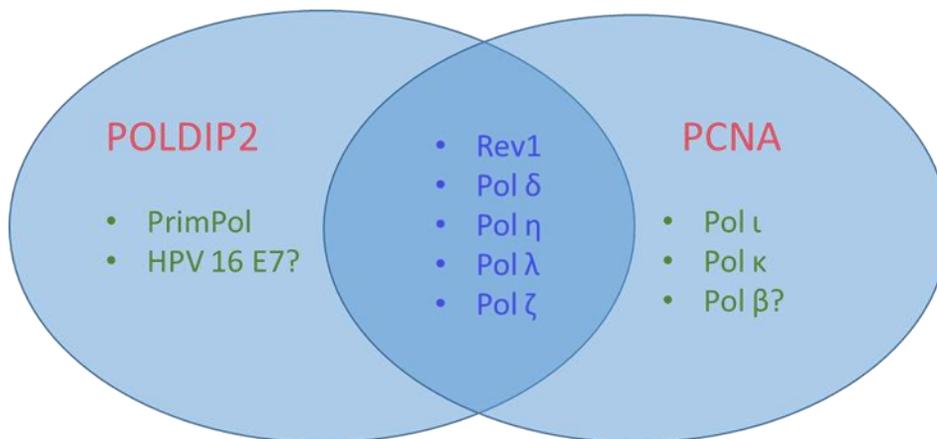


Figure 41. Common POLDIP2 (1-368) and PCNA binding proteins. Proteins which both POLDIP2 (1-368) and PCNA can bind are shown at the intersection. From all studied literature no evidence was found if PCNA can directly interact with HPV 16 E7, as well as, no confirmation was found it POLDIP2 (1-368)

can directly interact with Pol β (Moldovan, Pfander and Jentsch, 2007; Stoimenov and Helleday, 2009; Boehm, Goldenberg and Washington, 2016; Guillian *et al.*, 2015; Maga and Hubscher, 2003; Witko-Sarsat *et al.*, 2010; Tsuchimoto *et al.*, 2001; Haracska *et al.*, 2001; Haracska *et al.*, 2002; Naryzhny and Lee, 2010; Sykora *et al.*, 2017; Vaisman, Masutani, Hanaoka, & Chaney, 2000; Vidal *et al.*, 2004; Zhang, Yuan, Wu & Wang, 2000; Zhuang *et al.*, 2008). Figure created using PowerPoint.

Xie *et al.*, (2005), performed an electrophoretic shift assay using POLDIP2 (1-368) and showed that POLDIP2 (1-368) does not interact with the DNA. Such assumption might be wrong, as POLDIP2 (1-368) YccV domain shares 34% homology with the bacterial YccV like domain (Maga *et al.*, 2013). Bacterial YccV like domain predominantly binds hemimethylated DNA and as a result becomes able to stabilize other proteins during the DNA replication initiation process (Brown *et al.*, 2014; Shimuta *et al.*, 2004; d'Alencon *et al.*, 2003). Hence, an electrophoretic shift assay should be performed using the hemimethylated DNA. Also, it should be taken into account that POLDIP2 (1-368) could bind specific hemimethylated DNA sequences. Additionally, our POLDIP2 (51-368) surface charge representation model also confirms the presence of positively charged patches suggesting POLDIP2 (1-368) might be able to bind negatively charged DNA.

To sum up, project aims were met, with a nearly fully resolved POLDIP2 protein structure (99.8%). As was predicted by circular dichroism and PSIPRED, X-ray crystallography proved that POLDIP2 is mainly β sheet protein. POLDIP2 surface charge distribution analysis revealed that POLDIP2 contains hydrophobic, positively and negatively charged regions, suggesting and partially explaining its ability to interact with a numerous number of binding partners. POLDIP2 has large positively charged surface area, which could be involved in the interaction with negatively charged molecules, such as DNA, RNA or poly(ADP-ribose) (Chang, Jacobson and Mitchison, 2004). Additionally, *B* factor analysis suggested that $\alpha 2$ and loops are more flexible

(these regions are involved in the PrimPol interaction) in the structure, whereas the part forming β sheet structure is more rigid (POLDIP2 β 15-loop- β 16 secondary structure is involved in CAECAM1 binding). Therefore, both rigid regions, as well as, regions which exhibit some flexibility can be involved in protein-protein interaction. Interestingly, co-purification SEC shows that truncated POLDIP2 (51-368) form does not bind PCNA, whereas PCNA-POLDIP2 complex formation is observed then full length POLDIP2 protein is used. POLDIP2 structure solution is the first part of “journey” and in order to better understand how POLDIP2 binds multiple partners, POLDIP2 structure should be solved in combination with other binding partners. Therefore, future work will focus on POLDIP2-TLS polymerase structure determination, as well as, further PCNA-POLDIP2 complex studies in order to unravel POLDIP2 role at the DNA replication fork.

Appendices

Appendix 1- Table 1. POLDIP2 construct length, used primers and gene length.

| Sample Name | Construct length (a.a) | Sample code | Forward Primer | Reverse Primer | Gene length (bp) |
|-------------|------------------------|-------------|----------------|----------------|------------------|
| F01 | POLDIP2 (1-368) | AK001 | F000 | R000 | 1104 |
| F02 | POLDIP2 (1-348) | AK002 | F000 | R001 | 1044 |
| F03 | POLDIP2 (1-338) | AK003 | F000 | R002 | 1014 |
| F04 | POLDIP2 (51-368) | AK004 | F001 | R000 | 954 |
| F05 | POLDIP2 (51-348) | AK005 | F001 | R001 | 894 |
| F06 | POLDIP2 (51-338) | AK006 | F001 | R002 | 864 |
| F07 | POLDIP2 (178-368) | AK007 | F002 | R000 | 927 |
| F08 | POLDIP2 (178-348) | AK008 | F002 | R001 | 867 |
| F09 | POLDIP2 (178-338) | AK009 | F002 | R002 | 837 |
| F10 | POLDIP2 (721-368) | AK010 | F003 | R000 | 384 |
| F11 | POLDIP2 (721-348) | AK011 | F003 | R001 | 324 |
| F12 | POLDIP2 (721-338) | AK012 | F003 | R002 | 294 |

Appendix 2-Table 2. POLDIP2 primer sequences.

| Primer Name | Primer Sequence (with highlighted LIC overhangs) |
|-------------|--|
| Forward 000 | 5'- <u>TACTTCCAATCCATGGCAGCCTGTACAGCC</u> -3' |
| Forward 001 | 5'- <u>TACTTCCAATCCATGCTCTCGTCCCGAAACCGAC</u> -3' |
| Forward 002 | 5'- <u>TACTTCCAATCCATGAAAGTGTTGGAGACAGTT</u> -3' |
| Forward 003 | 5'- <u>TACTTCCAATCCATGATACGTGTC</u> ACTGTCATC-3' |
| Reverse 000 | 5'- <u>TATCCACCTTTACTGTCA</u> CCAGTGAAGGCCTGAGGG-3' |
| Reverse 001 | 5'- <u>TATCCACCTTTACTGTCA</u> AATCCGAACATCAAAGTGG-3' |
| Reverse 002 | 5'- <u>TATCCACCTTTACTGTCA</u> CTTTCAAAGCGGAACGTG-3' |
| pLICfor | 5'-TGTGAGCGGATAACAATTCC-3' |
| pLICrev | 5'-AGCAGCCAACCTCAGCTTCC-3' |

Appendix 7. Table 6. Columns and buffers used for IEC.

| Buffer/Column used | HiTrap Q | HiTrap SP | HiTrap Heparin |
|--------------------|---|---|---|
| A buffer | 20 mM Tris-Cl pH 7.5 100 mM NaCl 5% glycerol 1 mM DTT | 20 mM HEPES pH 7.5 100 mM NaCl 5% glycerol 1 mM DTT | 20 mM HEPES pH 7.5 100 mM NaCl 5% glycerol 1 mM DTT |
| B buffer | 20 mM Tris-Cl pH 7.5 1000 mM NaCl 5% glycerol 1 mM DTT | 20 mM HEPES pH 7.5 1000 mM NaCl 5% glycerol 1 mM DTT | 20 mM HEPES pH 7.5 1000 mM NaCl 5% glycerol 1 mM DTT |

Appendix 8. Table 7. Dialysis buffer recipe.

| | |
|------------------------|--|
| Dialysis buffer | 20 mM HEPES pH 7,5 500 mM NaCl 5% Glycerol 4 mM β -ME |
|------------------------|--|

Appendix 9. Table 8. All POLDIP2, POLDIP2 optimization and POLDIP2-PCNA complex plates.

| Construct (aa) | Construct code | Compound solvent | Protein frozen/fresh | Cleaved? | Temp (°C) | Screen | Concentration of protein (mg/ml) | Size drops (µl) | Drop ratio upper well (protein:ML µl) | Drop ratio lower well (protein:ML µl) |
|----------------|----------------|------------------|----------------------|----------|-----------|---------------------------------------|----------------------------------|-----------------|---------------------------------------|---------------------------------------|
| E04 (151-368) | CC083 | SEC buffer | fresh | Yes | 20 | JCSG | 20,0 | 2 | 1:1 | N/A |
| E07 (178-368) | CC086 | SEC buffer | fresh | Yes | 20 | JCSG | 20,0 | 2 | 1:1 | N/A |
| E07 (178-368) | CC086 | SEC buffer | fresh | Yes | 20 | PACT | 20,0 | 2 | 1:1 | N/A |
| E04 (151-368) | CC083 | SEC buffer | fresh | Yes | 20 | PACT | 20,0 | 2 | 1:1 | N/A |
| E7 (178-368) | CC086 | SEC buffer | fresh | Yes | 4 | PACT | 20,0 | 2 | 1:1 | N/A |
| E7 (178-368) | CC086 | SEC buffer | fresh | Yes | 4 | JCSG | 20,0 | 2 | 1:1 | N/A |
| E4 (151-368) | CC083 | SEC buffer | fresh | Yes | 4 | PACT | 20,0 | 2 | 1:1 | N/A |
| E4 (151-368) | CC083 | SEC buffer | fresh | Yes | 4 | JCSG | 20,0 | 2 | 1:1 | N/A |
| E4 (151-368) | CC083 | SEC buffer | fresh | Yes | 20 | JCSG | 20 and 10 | 2 | 1:1 (20 mg/ml) | 1:1 (10 mg/ml) |
| E4 (151-368) | CC083 | SEC buffer | fresh | Yes | 20 | JCSG | 20 and 15 | 2 | 1:1 (20 mg/ml) | 1:1 (15mg/ml) |
| E4 (151-368) | CC083 | SEC buffer | frozen | Yes | 20 | JCSG | 20 and 15 | 2 | 1:1 (20 mg/ml) | 1:1 (15mg/ml) |
| E7 (178-368) | CC086 | SEC buffer | fresh | Yes | 20 | JCSG | 15 | 2 | 1:1 (15 mg/ml) | N/A |
| E4 (151-368) | CC083 | SEC buffer | frozen | Yes | 20 | BCS | 20 and 10 | 2 | 1:1 (20 mg/ml) | 1:1 (10 mg/ml) |
| E4 (151-368) | CC083 | SEC buffer | frozen | Yes | 20 | Morpheus | 20 and 10 | 2 | 1:1 (20 mg/ml) | 1:1 (10 mg/ml) |
| E4 (151-368) | CC083 | SEC buffer | frozen | Yes | 4 | BCS | 20 and 10 | 2 | 1:1 (20 mg/ml) | 1:1 (10 mg/ml) |
| PCNA-E1/PCNA | N/A | SEC buffer | frozen | Yes | 20 | JCSG | 5,15/5,15 | 2 | 1:1 (5,15 mg/ml) | 1:1 (5, 15 mg/ml) |
| PCNA-E1/PCNA | N/A | SEC buffer | frozen | Yes | 20 | PACT | 5,15/5,15 | 2 | 1:1 (5,15 mg/ml) | 1:1 (5,15 mg/ml) |
| E4 (151-368) | CC083 | SEC buffer | fresh | Yes | 20 | Optimization 100 mM CaCl ₂ | 20/24 | 2 | 1:1 (20 mg/ml) | 1:1 (24 mg/ml) |
| E4 (151-368) | CC083 | SEC buffer | fresh | Yes | 20 | Optimization 200 mM CaCl ₂ | 20/24 | 2 | 1:1 (20 mg/ml) | 1:1 (24 mg/ml) |

References

- Alpert, A.J., Petritis, K., Kangas, L., Smith, R.D., Mechtler, K., Mitulović, G. ... Mohammed, A.J.R. (2010). Peptide Orientation Affects Selectivity in Ion-Exchange Chromatography. *Analytical Chemistry*, 82 (12), 5253-5259. . doi: 10.1021/ac100651k
- Amanso, A.M., Lassègue, B., Joseph, G., Landázuri, N., Long, J.S., Weiss, D. ... Griendling, K.K. (2014). Poldip2 promotes postischemic neovascularization of the mouse hindlimb. *Arteriosclerosis Thrombosis Vascular Biology*, 34 (7), 1548-1555. doi: i:10.1161/ATVBAHA.114.303873.
- Amantea, D., Micieli, G., Tassorelli, C., Cuartero, M.I., Ballesteros, I., Certo, M. ... Bagetta, G. (2015). Rational modulation of the innate immune system for neuroprotection in ischemic stroke. *Frontiers in Neuroscience*, 147 (9), . doi: 10.3389/fnins.2015.00147.
- Ameziane-El-Hassani, R., Schlumberger, M., & Dupuy, C. (2016). NADPH oxidases: new actors in thyroid cancer? *Nature Reviews. Endocrinology*, 12 (8), 485-494. doi: 10.1038/nrendo.2016.64.
- Aravind, L. & Koonin, E.V. (1998). Phosphoesterase domains associated with DNA polymerases of diverse origins.. *Nucleic Acids Research*, 26 (26), 3746-3752. doi: 10.1093/nar/26.16.3746.
- Aria, V. & Yeeles, J.T.P. (2018). Mechanism of Bidirectional Leading-Strand Synthesis Establishment at Eukaryotic DNA Replication Origins. *Molecular Cell*, 2765 (18), 30879-30887. doi: 10.1016/j.molcel.2018.10.019.
- Aslanidis, C. & Jong, P.J. (1990). Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Research*, 18 (20), 6069-6074. doi: 10.1093/nar/18.20.6069.
- Baldwin, R.L. (1986). Temperature dependence of the hydrophobic interaction in protein folding. *Proceedings of the National Academy of Sciences*, 83 (21), 8069-8072. . doi: 10.1073/pnas.83.21.8069
- Bannister, W.H. & Bannister, J.V. (1974). Circular dichroism and protein structure. *International Journal of Biochemistry*, 5 (9), 673-677. doi: 10.1016/0020-711X(74)90052-4.
- Baranovskiy, A.G., Zhang, Y., Suwa, Y., Babayeva, N.D., Gu, J., Pavlov, Y.I., & Tahirov, T.H. (2015). Crystal Structure of the Human Primase. *The Journal of Biological Chemistry*, 290 (9), 5635-5646. doi: 10.1074/jbc.M114.624742.
- Barnes, R.P., Hile, S.E., Lee, M.Y., & Eckert, K.A. (2017). DNA polymerases eta and kappa exchange with the polymerase delta holoenzyme to complete common fragile site synthesis. *DNA Repair*, 57 (1), 1-11. doi: 10.1016/j.dnarep.2017.05.006.
- Benoit, R.M., Ostermeier, C., Geiser, M., Su Zhou Li, J., Widmer, H., & Auer, M. (2016). Seamless Insert-Plasmid Assembly at High Efficiency and Low Cost. *PLoS ONE*, 11 (4), 1-13. doi: 10.1371/journal.pone.0153158.
- Bienko, M., Green, C.M., Crosetto, N., Rudolf, F., Zapart, G., Coull, B. ... Dikic, I. (2005). Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science*, 310 (5755), 1821-1824. doi: 10.1126/science.1120615.

- Bird, L.E. (2011). High throughput construction and small scale expression screening of multi-tag vectors in *Escherichia coli*. *Methods*, 55 (1), 29-37. . doi: 10.1016/j.ymeth.2011.08.002
- Block, H., Maertens, B., Spriestersbach, A., Brinker, N., Kubicek, J., Fabis, R. ... Schäfer, F. (2009). Immobilized-Metal Affinity Chromatography (IMAC): A Review. *Methods in Enzymology*, 463 (1), 439-473. . doi: 10.1016/S0076-6879(09)63027-5
- Boehm, E.M., Goldenberg, M.S., & Washington, M.T. (2016). The Many Roles of PCNA in Eukaryotic DNA Replication. *The Enzymes*, 39 (1), 231-254. doi: 10.1016/bs.enz.2016.03.003.
- Bolden, A., Fry, M., Muller, R., Citarella, R., & Weissbach, A. (1972). The presence of a polyriboadenylic acid-dependent DNA polymerase in eukaryotic cells. *Archives of Biochemistry and Biophysics*, 153 (1), 26-33. doi: 10.1016/0003-9861(72)90416-x.
- Bomar, M.G., Pai, M.T., Tzeng, S.R., Li, S.S., & Zhou, P. (2007). Structure of the ubiquitin-binding zinc finger domain of human DNA Y-polymerase eta. *EMBO Reports*, 8 (3), 247-251. doi: 10.1038/sj.embor.7400901.
- Brown, D. I., Lassègue, B., Lee, M., Zafari, R., Long, J. S., Saavedra, H. I., & Griendling, K. K. (2014). Poldip2 knockout results in perinatal lethality, reduced cellular growth and increased autophagy of mouse embryonic fibroblasts. *PloS One*, 9(5), e96657. 10.1371/journal.pone.0096657
- Bruni-Cardoso, A., Johnson, L.C., Vessella, R.L., Peterson, T.E., & Lynch, C.C. (2010). Osteoclast derived matrix metalloproteinase-9 directly impacts angiogenesis in the prostate tumor-bone microenvironment. *Molecular Cancer Research*, 8 (4), 459-470. doi: 10.1158/1541-7786.MCR-09-0445.
- Burch, P.M. & Heintz, N.H. (2005). Redox regulation of cell-cycle re-entry: cyclin D1 as a primary target for the mitogenic effects of reactive oxygen and nitrogen species. *Antioxidant and redox signaling*, 7 (5-6), 741-751. doi: 10.1089/ars.2005.7.741.
- Carell, T. (2015). DNA repair. *Angewandte Chemie International Edition*, 54(51),15330-15333. doi:10.1002/anie.201509770
- Chang, P., Jacobson, M.K., & Mitchison, T.J. (2004). Poly(ADP-ribose) is required for spindle assembly and structure. *Nature*, 432 (7017), 645-649. . doi: 10.1038/nature03061
- Chayen, N.E. & Saridakis, E. (2008). Protein crystallisation: from purified protein to diffraction-quality crystal. *Nature Methods*, 5 (2), 147-153. doi: 10.1038/nmeth.f.203.
- Chen, Y.C., Kuo, C.C., Chian, C.F., Tzao, C., Chang, S.Y., Shih, Y.L. ... Su, H.Y. (2018). Knockdown of POLDIP2 suppresses tumor growth and invasion capacity and is linked to unfavorable transformation ability and metastatic feature in non-small cell lung cancer. *Experimntal Cell Research* , 368 (1), 42-49. doi: 10.1016/j.yexcr.2018.04.011.
- Cheung, R.C., Wong, J.H., & Ng, T.B. (2012). Immobilized metal ion affinity chromatography: a review on its applications. *Applied Microbiology and Biotechnology* , 96 (6), 1411-1420. . doi: 10.1007/s00253-012-4507-0
- Chian, C.F., Hwang, Y.T., Terng, H.J., Lee, S.C., Chao, T.Y., Chang, H. ... Perng, W.C. (2016). Panels of tumor-derived RNA markers in peripheral blood of patients with non-small cell lung

cancer: their dependence on age, gender and clinical stages. *Oncotarget*, 7 (31), 50582-50595. doi: 10.18632/oncotarget.10558.

Choe, K.N. & Moldovan, G.L. (2017). Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork. *Molecular Cell*, 65 (3), 380-392. doi: 10.1016/j.molcel.2016.12.020.

Claros, M. G., & Vincens, P. (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. *European Journal of Biochemistry*, 241(3), 779-786. doi: 10.1111/j.1432-1033.1996.00779.x

Coffin, J.D., Homer-Bouthiette, C., & Hurley, M. (2018). Fibroblast Growth Factor 2 and Its Receptors in Bone Biology and Disease. *Journal of the Endocrine Society*, 2 (7), 657-671. doi: 10.1210/js.2018-00105.

Collins, K.D. (2004). Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallisation process. *Methods*, 34 (3), 300-311. . doi: 10.1016/j.ymeth.2004.03.021

Cook-Mills, J.M., Marchese, M.E., & Abdala-Valencia, H. (2001). Vascular cell adhesion molecule-1 expression and signaling during disease: regulation by reactive oxygen species and antioxidants. *Antioxidants and Redox Signaling*, 15 (6), 1607-1638. doi: 10.1089/ars.2010.3522.

Costa, S.J., Almeida, A., Castro, A., Domingues, L., & Besir, H. (2013). The novel Fh8 and H fusion partners for soluble protein expression in *Escherichia coli*: a comparison with the traditional gene fusion technology. *Applied Microbiology and Biotechnology*, 97 (15), 6779–6791. . doi: 10.1007/s00253-012-4559-1

Cox, L.S. (1997). Who binds wins: Competition for PCNA rings out cell-cycle changes. *Trends in Cell Biology*, 7 (12), 493-498. doi: 10.1016/S0962-8924(97)01170-7.

Crosas-Molist, E., Bertran, E., Sancho, P., López-Luque, J., Fernando, J., Sánchez, A. ... Fabregat, I. (2014). The NADPH oxidase NOX4 inhibits hepatocyte proliferation and liver cancer progression. *Free Radical Biology and Medicine*, 69 (1), 338-347. doi: 10.1016/j.freeradbiomed.2014.01.040.

Crowe, J., Döbeli, H., Gentz, R., Hochuli, E., Stüber, D., & Henco, K. (1994). 6xHis-Ni-NTA chromatography as a superior technique in recombinant protein expression/purification. *Methods in Molecular Biology*, 31 (1), 1371-1387. . doi: 10.1385/0-89603-258-2:371

Curtis, R.A., Montaser, A., Prausnitz, J.M., & Blanch, H.W. (1998). Protein-protein and protein-salt interactions in aqueous protein solutions containing concentrated electrolytes. *Biotechnology and Bioengineering*, 58 (4), 451. . doi: <https://library.hud.ac.uk/>

Dahal, S., Dubey, S., & Raghavan, S. C. (2018). Homologous recombination-mediated repair of DNA double-strand breaks operates in mammalian mitochondria. *Cellular and Molecular Life Sciences*, 75(9), 1641-1655. doi:10.1007/s00018-017-2702-y

Dale, G.E., Oefner, C., & D'Arcy, A. (2003). The protein as a variable in protein crystallisation. *Journal of Structural Biology*, 142 (1), 88-97. doi: doi:10.1016/S1047-8477(03)00041-8.

d'Alençon, E., Taghbalout, A., Bristow, C., Kern, R., Aflalo, R., & Kohiyama, M. (2003). Isolation of a new hemimethylated DNA binding protein which regulates dnaA gene expression. *Journal of Bacteriology*, 185(9), 2967-2971. doi: 10.1128/JB.185.9.2967-2971.2003

Datla, S.R., Hilenski, L.L., Seidel-Rogol, B., Dikalova, A.E., Harousseau, M., Punkova, L. ... Griendling, K.K. (2019). Poldip2 knockdown inhibits vascular smooth muscle proliferation and neointima formation by regulating the expression of PCNA and p21. *Nature*, 99 (3), 387-398. doi: 10.1038/s41374-018-0103-y.

Davis, G.D., Elisee, C., Newham, D.M., & Harrison, R.G. (1999). New fusion protein systems designed to give soluble expression in Escherichia coli. *Biotechnology and Bioengineering*, 65 (4), 382-388. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/10506413>

De Felice, F. G., Wu, D., Lambert, M. P., Fernandez, S. J., Velasco, P. T., Lacor, P. N., Klein, W. L. (2008). Alzheimer's disease-type neuronal tau hyperphosphorylation induced by A β oligomers. *Neurobiology of Aging*, 29(9), 1334-1347. doi: 10.1016/j.neurobiolaging.2007.02.029

Doherty, A.J., Connolly, B.A., & Worrall, A.F. (1993). Overproduction of the toxic protein, bovine pancreatic DNaseI, in Escherichia coli using a tightly controlled T7-promoter-based vector. *Gene*, 136 (1-2), 337-340. doi: 10.1016/0378-1119(93)90491-K

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K., & Mattick, J.S. (1991). 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research*, 19 (14), 4008-4008. doi: 10.1093/nar/19.14.4008. original no malenkij...

Dong, H., Nilsson, L., & Kurland, C.G. (1995). Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction. *Journal of Bacteriology*, 177 (6), 1497-1504. doi: 10.1128/jb.177.6.1497-1504.1995

Dym, O. & Eisenberg, D. (2001). Sequence-structure analysis of FAD-containing proteins. *Protein Science*, 10 (9), 1712- 1728. doi: 10.1110/ps.12801.

Ekholm, S.V. & Reed, S.I. (2000). Regulation of G(1) cyclin-dependent kinases in the mammalian cell cycle. *Current opinion in cell biology*, 12 (6), 676-684. doi: 10.1016/S0955-0674(00)00151-4.

Emanuelsson, O., Nielsen, H., Brunak, S., & Heijne, G. (2000). Predicting Subcellular Localization of Proteins Based on their N-terminal Amino Acid Sequence. *Journal of Molecular Biology*, 300(4), 1005-1016. doi: 10.1006/jmbi.2000.3903

Emsley P, Cowtan K (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D*60, 2126-2132.

Emsley, P., Lohkamp, B., Scott, W.G., & Cowtan, K. (2010). Features and development of Coot. *Biological Crystallography*, 66 (4), 486-501. doi: 10.1107/S0907444910007493.

Esposito, D. & Chatterjee, D.K. (2006). Enhancement of soluble protein expression through the use of fusion tags. *Current Opinion in Biotechnology*, 17 (4), 353-358. doi: 10.1016/j.copbio.2006.06.003

Evans, P. & McCoy, A. (2008). An introduction to molecular replacement. *Biological Crystallography*, 64 (1), 1-10. doi: 10.1107/S0907444907051554

- Falkenberg, M., Larsson, N.G., & Gustafsson, C.M. (2007). DNA replication and transcription in mammalian mitochondria. *Annual review of Biochemistry*, 76 (1), 679-699. doi: 10.1146/annurev.biochem.76.060305.152028.
- Fenwick, R.B., Bedem, H.V.D., Fraser, J.S., & Wright, P.E. (2014). Integrated description of protein dynamics from room-temperature X-ray crystallography and NMR. *Proceedings of the National Academy of Sciences*, 111 (4), E445-E454. . doi: 10.1073/pnas.1323440111
- Foley, M.C., Couto, L., Rauf, S., & Boyke, A. (2019). Insights into DNA polymerase δ 's mechanism for accurate DNA replication. *Journal of Molecular modeling*, 25 (3),. doi: 10.1007/s00894-019-3957-z.
- Frances, A.H. (1991). Metal-Affinity Separations: A New Dimension in Protein Processing. *Nature Biotechnology*, 9 (1), 151-156. doi: 10.1038/nbt0291-151.
- Frick, D.N. & Richardson, C.C. (2001). DNA primases. *Annual review of Biochemistry*, 70 (1), 39-80. doi: 10.1146/annurev.biochem.70.1.39.
- Fujii, M., Amanso, A., Abrahão, T.B., Lassègue, B., & Griendling, K.K. (2016). Polymerase delta-interacting protein 2 regulates collagen accumulation via activation of the Akt/mTOR pathway in vascular smooth muscle cells. *Journal of Molecular and Cellular Cardiology*, 92 (1), 21-29. doi: 10.1016/j.yjmcc.2016.01.016.
- Fukasawa, Y., Tsuji, J., Fu, S. C., Tomii, K., Horton, P., & Imai, K. (2015). MitoFates: Improved Prediction of Mitochondrial Targeting Sequences and Their Cleavage Sites. *Molecular and Cellular Proteomics*, 14(4), 1113–1126. doi: 10.1074/mcp.M114.043083
- Garbacz, M.A., Lujan, S.A., Burkholder, A.B., Cox, P.B., Wu, Q., Zhou, Z.X. ... Kunkel, T.A. (2018). Evidence that DNA polymerase δ contributes to initiating leading strand DNA replication in *Saccharomyces cerevisiae*. *Nature Communications*, 9 (1), . doi: 10.1038/s41467-018-03270-4.
- García-Gómez, S., Reyes, A., Martínez-Jiménez, M. I., Chocrón, E. S., Mourón, S., Terrados, G., Blanco, L. (2013). PrimPol, an archaic primase/polymerase operating in human cells. *Molecular Cell*, 52(4), 541-553. doi:10.1016/j.molcel.2013.09.025
- Garg, P. & Burgers, P.M. (2005). DNA polymerases that propagate the eukaryotic DNA replication fork. *Critical reviews in biochemistry and molecular biology*, 40 (2), 115-128. doi: 10.1080/10409230590935433 .
- Garg, P. & Burgers, P.M. (2005). Ubiquitinated proliferating cell nuclear antigen activates translesion DNA polymerases η and REV1. *Proceedings of the National Academy of Sciences of the United States of America*, 102 (51), 18361-18366. doi: 10.1073/pnas.0505949102.
- Garman, E. (1999). Cool data: quantity AND quality. *Biological Crystallography*, 55 (10), 1642-1653. doi: 10.1107/S0907444999008653.
- Garman, E.F. & Schneider, T.R. (1997). Macromolecular Cryocrystallography. *Journal of Applied Crystallography*, 30 (3), 211-237. doi: 10.1107/S0021889897002677.

- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, *31*(13), 3784-3788. doi: 10.1093/nar/gkg563
- Gavrilov, M. & Monteiro, M.J. (2015). Derivation of the molecular weight distributions from size exclusion chromatography. *European Polymer Journal*, *65* (1), 191-196. . doi: 10.1016/j.eurpolymj.2014.11.018
- GE Healthcare (2006). Addition of imidazole during binding improves purity of histidine-tagged proteins [Application Note 28-4067-41 AA]. Sweden: K. Hölscher, M. Richter-Roth, and B. Felden de Neumann.
- Gell, D.A., Grant, R.P., & Mackay, J.P. (2012). The detection and quantitation of protein oligomerization. *Advances in Experimental Medicine and Biology*, *747* (1), 19-41. . doi: 10.1007/978-1-4614-3229-6_2
- Goettsch, C., Babelova, A., Trummer, O., Erben, R.G., Rauner, M., Rammelt, S. ... Schröder, K. (2013). NADPH oxidase 4 limits bone mass by promoting osteoclastogenesis. *The Journal of Clinical Investigation*, *123* (11), 4731-4738. doi: 10.1172/JCI67603.
- Golias, T., Kery, M., Radenkovic, S., & Papandreou, I. (2019). Microenvironmental control of glucose metabolism in tumors by regulation of pyruvate dehydrogenase. *International Journal of Cancer*, *144* (4), 674-686. doi: 10.1002/ijc.31812.
- Gräslund, S., Nordlund, P., Weigelt, J., Hallberg, B.M., Bray, J., Gileadi, O. ... Gunsalus, K.C. (2008). Protein production and purification. *Nature Methods*, *5* (2), 135-146. . doi: 10.1038/nmeth.f.202
- Greenfield, N.J. (1999). Applications of circular dichroism in protein and peptide analysis. *TrAC Trends in Analytical Chemistry*, *18* (4), 236-244. doi: 10.1016/S0165-9936(98)00112-5.
- Greenfield, N.J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols*, *1* (6), 2876-2890. . doi: 10.1038/nprot.2006.202
- Griendling, K.K. (2018). Nox4, Poldip2 and vascular function. *Free Radical Biology and Medicine*, *120* (1), 11-12. doi: 10.1016/j.freeradbiomed.2018.04.051.
- Guilliam, T. A., Bailey, L. J., Brissett, N. C., & Doherty, A. J. (2016). PolDIP2 interacts with human PrimPol and enhances its DNA polymerase activities. *Nucleic Acids Research*, *44*(7), 3317-3329. doi:10.1093/nar/gkw175
- Guilliam, T.A., Jozwiakowski, S.K., Ehlinger, A., Barnes, R.P., Rudd, S.G., Bailey, L.J. ... Doherty, A.J. (2015). Human PrimPol is a highly error-prone polymerase regulated by single-stranded DNA binding proteins. *Nucleic Acids Research*, *43* (2), 1056-1068. doi: 10.1093/nar/gku1321.
- Gulbis, J.M., Kelman, Z., Hurwitz, J., O'Donnell, M., & Kuriyan, J. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell*, *87* (2), 297-306. . doi: 10.1016/s0092-8674(00)81347-1

- Guo, C., Kosarek-Stancel, J.N., Tang, T.S., & Friedberg, E.C. (2009). Y-family DNA polymerases in mammalian cells. *Cellular and Molecular Life Sciences*, 66 (14), 2363-2381. doi: 10.1007/s00018-009-0024-4.
- Guo, C., Sonoda, E., Tang, T.S., Parker, J.L., Bielen, A.B., Takeda, S. ... Friedberg, E.C. (2006). REV1 protein interacts with PCNA: significance of the REV1 BRCT domain in vitro and in vivo. *Molecular Cell*, 23 (2), 265-271. doi: 10.1016/j.molcel.2006.05.038.
- Guo, C., Tang, T.S., Bienko, M., Parker, J.L., Bielen, A.B., Sonoda, E. ... Friedberg, E.C. (2006). Ubiquitin-binding motifs in REV1 protein are required for its role in the tolerance of DNA damage. *Molecular and Cellular Biology*, 26 (23), 8892-8900. doi: 10.1128/MCB.01118-06.
- Gupta, A., Lad, S.B., Ghodke, P.P., Pradeepkumar, P.I., & Kondabagil, K. (2019). Mimivirus encodes a multifunctional primase with DNA/RNA polymerase, terminal transferase and translesion synthesis activities. *Nucleic Acids Research*, 47 (13), 6932-6945. doi: 10.1093/nar/gkz236.
- Hammarström, M., Woestenenk, E.A., Hellgren, N., Härd, T., & Berglund, H. (2006). Effect of N-terminal solubility enhancing fusion proteins on yield of purified target protein. *Journal of Structural and Functional Genomics*, 7 (1), 1-14. . doi: 10.1007/s10969-005-9003-7
- Haracska, L., Johnson, R.E., Unk, I., Phillips, B.B., Hurwitz, J., Prakash, L., & Prakash, S. (2001). Targeting of human DNA polymerase ι to the replication machinery via interaction with PCNA. *Proceedings of the National Academy of Sciences*, 98 (25), 14256-14261. doi: 10.1073/pnas.261560798.
- Haracska, L., Unk, I., Johnson, R.E., Phillips, B.B., Hurwitz, J., Prakash, L., & Prakash, S. (2002). Stimulation of DNA synthesis activity of human DNA polymerase kappa by PCNA. *Molecular and Cellular Biology*, 22 (3), 784-791. doi: 10.1128/mcb.22.3.784-791.2002.
- Hedglin, M., Pandey, B., & Benkovic, S.J. (2016). Stability of the human polymerase δ holoenzyme and its implications in lagging strand DNA synthesis. *Proceedings of the National Academy of Sciences*, 113 (13), 1777-1786. doi: 10.1073/pnas.1523653113.
- Henzler-Wildman, K. & Kern, D. (2007). Dynamic personalities of proteins. *Nature*, 450 (7172), 964-972. . doi: 10.1038/nature06522
- Hernandes, M.S., Lassègue, B., & Griendling, K.K. (2017). Polymerase δ -interacting protein 2: a multifunctional protein. *Journal of Cardiovascular Pharmacology*, 69 (6), 335-342. doi: 10.1097/FJC.0000000000000465.
- Hirota, K., Tsuda, M., Tsurimoto, T., Cohen, I.S., Livneh, Z., Kobayashi, K. ... Takeda, S. (2016). In vivo evidence for translesion synthesis by the replicative DNA polymerase δ . *Nucleic Acids Research*, 44 (15), 7242-7250. doi: 10.1093/nar/gkw439.
- Ilyin, G.P., Rialland, M., Pigeon, C., & Guguen-Guillouzo, C. (2000). cDNA cloning and expression analysis of new members of the mammalian F-box protein family. *Genomics*, 67 (1), 40-47. doi: 10.1006/geno.2000.6211.
- Ingham, K.C. (1990). Precipitation of proteins with polyethylene glycol. *Methods in Enzymology*, 182 (1), 301-306. . doi: 10.1016/0076-6879(90)82025-w

- Jacobson, K.B., Murphy, J.B., & Das Sarma, B. (1972). Reaction of cacodylic acid with organic thiols. *FEBS Letters*, 22 (1), 80-82. . doi: 10.1016/0014-5793(72)80224-2
- Jacquet, A., Daminet , V., Haumont, M., Garcia, L., Chaudoir, S., Bollen , A., & Biemans, R. (1999). Expression of a recombinant *Toxoplasma gondii* ROP2 fragment as a fusion protein in bacteria circumvents insolubility and proteolytic degradation. *Protein Expression and Purification* , 17 (3), 392-400. . doi: 10.1006/prev.1999.1150
- Jancarik, J. & Kim, S.H. (1991). Sparse Matrix Sampling: A Screening Method for Crystallisation of Proteins. *Journal of Applied Crystallography*. *Journal of Applied Crystallography*, 24 (1), 409-411. . doi: 10.1107/S0021889891004430
- Jessberger, R., Podust, V., Hübscher , U., & Berg, P. (1993). A mammalian protein complex that repairs double-strand breaks and deletions by recombination. *The Journal of Biological Chemistry*, 268 (20), 15070-15079. Retrieved from <http://www.jbc.org/content/268/20/15070.short>.
- Jiang, Z., Zhou, J., Li, T., Tian, M., Lu, J., Jia, Y. ... Chen, K. (2019). Hepatic deficiency of Poldip2 in type 2 diabetes dampens lipid and glucose homeostasis. *Metabolism: Clinical Studies*, 25 (99), 90-101. doi: 10.1016/j.metabol.2019.153948.
- Johnson, R.E., Klassen, R., Prakash, L., & Prakash, S. (2015). A Major Role of DNA Polymerase δ in Replication of Both the Leading and Lagging DNA Strands. *Molecular Cell*, 59 (2), 163-175. doi: 10.1016/j.molcel.2015.05.038.
- Johnson, R.E., Prakash, S., & Prakash, L. (1999). Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Pol ϵ . *Science*, 283 (5404), 1001-1004. doi: 10.1126/science.283.5404.1001.
- Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S., & Prakash, L. (2000). Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature*, 406 (6799), 1015-1019. doi: 10.1038/35023030.
- Kam, T.I., Song, S., Gwon, Y., Park, H., Yan, J.J., Im, I. ... Jung, Y.K. (2013). Fc γ RIIb mediates amyloid- β neurotoxicity and memory impairment in Alzheimer's disease. *The Journal of Clinical Investigation*, 123 (7), 2791-2802. doi: 10.1172/JCI66827.
- Kannouche, P.L., Wing, J., & Lehmann, A.R. (2004). Interaction of human DNA polymerase ϵ with monoubiquitinated PCNA: a possible mechanism for the polymerase switch in response to DNA damage. *Molecular Cell*, 14 (4), 491-500. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/15149598>.
- Kapust, R.B. & Waugh, D.S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science*, 8 (8), 1668-1674. . doi: 10.1110/ps.8.8.1668
- Kar, S. & Ellington, A.D. (2018). Construction of synthetic T7 RNA polymerase expression systems. *Methods*, 143 (1), 110-120. . doi: 10.1016/j.ymeth.2018.02.022
- Katsumura, S., Izu, Y., Yamada, T., Griendling, K., Harada, K., Noda, M., & Ezura, Y. (2016). FGF suppresses Poldip2 expression in osteoblasts: FGF regulates Poldip2 expression. *Journal of Cellular Biochemistry* 118(7), 1670-1677. doi:10.1002/jcb.25813

- Khandagale, P., Peroumal, D., Manohar, K., & Acharya, N. (2019). Human DNA polymerase delta is a pentameric holoenzyme with a dimeric p12 subunit. *Life Sciences Alliance*, 18 (2), . doi: 10.26508/lsa.201900323.
- Kim, Y., Park, H., Nah, J., Moon, S., Lee, W., Hong, S.H. ... Jung, Y.K. (2015). Essential role of POLDIP2 in Tau aggregation and neurotoxicity via autophagy/proteasome inhibition. . *Biochemical and Biophysical Research Communications*, 462 (2), 112-118. doi: 10.1016/j.bbrc.2015.04.084.
- Klaile, E., Kukalev, A., Obrink, B., & Müller, M.M. (2008). PDIP38 is a novel mitotic spindle-associated protein that affects spindle organization and chromosome segregation. *Cell Cycle*, 7 (20), 3180-3186. doi: 10.4161/cc.7.20.6813.
- Klaile, E., Müller, M.M., Kannicht, C., Otto, W., Singer, B.B., Reutter, W. ... Lucka, L. (2007). The cell adhesion receptor carcinoembryonic antigen-related cell adhesion molecule 1 regulates nucleocytoplasmic trafficking of DNA polymerase delta-interacting protein 38. *The Journal of Biological Chemistry*, 282 (36), 26629-26640. doi: 10.1074/jbc.M701807200.
- Kopaciewicz, W. & Regnier, F.E. (1982). Nonideal size-exclusion chromatography of proteins: Effects of pH at low ionic strength. *Analytical Biochemistry*, 126 (1), 8-16. . doi: 10.1016/0003-2697(82)90102-6
- Korasic, D.A. & Tanner, J.J. (2018). Determination of protein oligomeric structure from small-angle X-ray scattering. *Protein Science*, 27 (1), 814-824. . doi: 10.1002/pro.3376
- Korbie, D.J. & Mattick, J.S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols*, 3 (9), 1452-1456. doi: 10.1038/nprot.2008.133.
- Kriminski, S., Caylor, C.L., Nonato, M.C., Finkelstein, K.D., & Thorne, R.E. (2002). Flash-cooling and annealing of protein crystals. *Biological Crystallography*, 58 (3), 459-471. doi: 10.1107/s0907444902000112.
- Krohn, N.M., Yanagisawa, S., & Grasser, K.D. (2002). Home Current Issue Papers in Press Editors' Picks JBC Reviews Specificity of the Stimulatory Interaction between Chromosomal HMGB Proteins and the Transcription Factor Dof2 and Its Negative Regulation by Protein Kinase CK2-mediated Phosphorylation. *The Journal of Biological Chemistry*, 277 (36), 32438 – 32444. . doi: 10.1074/jbc.M203814200
- Krzysiak, T. C., Chen, B. B., Lear, T., Mallampalli, R. K., & Gronenborn, A. M. (2016). Crystal structure and interaction studies of the human FBxo3 ApaG domain. *The FEBS Journal*, 283(11), 2091-2101. 10.1111/febs.13721
- Kumar, A., Singh, A., & Ekavali. (2015). A review on alzheimer's disease pathophysiology and its management: An update. *Pharmacological Reports: PR*, 67(2), 195-203. doi:10.1016/j.pharep.2014.09.004
- Kurland, C. & Gallant, J. (1996). Errors of heterologous protein expression. *Current Opinion in Biotechnology*, 7 (5), 489-493. . doi: 10.1016/S0958-1669(96)80050-4
- Kuzminov, A. & Stahl, F.W. (1997). Stability of Linear DNA in recA Mutant Escherichia coli Cells Reflects Ongoing Chromosomal DNA Degradation. *Journal of Biotechnology* 179 (3), 880-888. doi: 10.1128/jb.179.3.880-888.1997.

- Hernandes, L. Lassègue and Griendling B. (2013). Poldip2 sustains vascular structure and function. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 33(9), doi: 10.1161/ATVBAHA.113.301913.
- Laurent, S., Alivon, M., Beaussier, H., & Boutouyrie, P. (2012). Aortic stiffness as a tissue biomarker for predicting future cardiovascular events in asymptomatic hypertensive subjects. *Annals of Medicine*, 44 (1), 93-97. doi: 10.3109/07853890.2011.653398.
- Lessard, J.C. (2013). Growth Media for E. coli. *Methods in Enzymology*, 533 (1), 181-189. . doi: 10.1016/B978-0-12-420067-8.00011-8
- Li, C. & Evans, R.M. (1997). Ligation independent cloning irrespective of restriction site compatibility. *Nucleic Acids Research*, 25 (20), 4165-4166. doi: 10.1093/nar/25.20.4165.
- Lipfert, J., Doniach, S., Das, R., & Herschlag, D. (2014). Understanding Nucleic Acid–Ion Interactions. *Annual Reviews in Biochemistry*, 83 (1), 813-841. . doi: 10.1146/annurev-biochem-060409-092720
- Liu, J., Yin, D.C., Guo, Y.Z., Wang, X.K., Xie, S.X., Lu, Q.Q., & Liu, Y.M. (2011). Selecting Temperature for Protein Crystallisation Screens Using the Temperature Dependence of the Second Virial Coefficient. *PLoS One*, 6 (3), e17950. . doi: 10.1371/journal.pone.0017950
- Liu, L., Rodriguez-Belmonte, E. M., Mazloun, N., Xie, B., & Lee, Marietta Y W T. (2003). Identification of a novel protein, PDIP38, that interacts with the p50 subunit of DNA polymerase delta and proliferating cell nuclear antigen. *The Journal of Biological Chemistry*, 278(12), 10041-10047. doi:10.1074/jbc.M208694200.
- Loeb, L.A. & Monnat, R.J. (2008). DNA polymerases and human disease. *Nature Reviews Genetics*, 9 (8), 594-604. doi: 10.1038/nrg2345.
- Longley, M.J., Pierce, A.J., & Modrich, P. (1997). DNA polymerase delta is required for human mismatch repair in vitro. *The Journal of Biological Chemistry*, 272 (16), 10917-10921. doi: 10.1074/jbc.272.16.10917.
- Luft, J.R., Wolfley, J.R., & Snell, E.H. (2011). What's in a drop? Correlating observations and outcomes to guide macromolecular crystallisation experiments. *Crystal Growth*, 11 (3), 651-663. doi: 10.1021/cg1013945.
- Lujan, S.A., Williams, J.S., & Kunkel, T.A. (2016). DNA Polymerases Divide the Labor of Genome Replication. *Trends in cell biology*, 26 (9), 640-654. doi: 10.1016/j.tcb.2016.04.012.
- Lyle, A. N., Deshpande, N. N., Taniyama, Y., Seidel-Rogol, B., Pounkova, L., Du, P., Griendling, K. K. (2009). Poldip2, a novel regulator of Nox4 and cytoskeletal integrity in vascular smooth muscle cells. *Circulation Research*, 105(3), 249–259. <http://doi.org/10.1161/CIRCRESAHA.109.193722>
- Maga, G. & Hubscher, U. (2003). Proliferating cell nuclear antigen (PCNA): a dancer with many partners. *Journal of Cell Science*, 116 (15), 3051-3060. doi: 10.1242/jcs.00653.
- Maga, G., Crespan, E., Markkanen, E., Imhof, R., Furrer, A., Villani, G. ... Loon, B. (2013). DNA polymerase δ -interacting protein 2 is a processivity factor for DNA polymerase λ during 8-oxo-

7,8-dihydroguanine bypass. *Proceedings of the National Academy of Sciences*, 110 (47), 18850-18855. doi: 10.1073/pnas.1308760110.

Maloisel, L., Fabre, F., & Gangloff, S. (2008). DNA polymerase delta is preferentially recruited during homologous recombination to promote heteroduplex DNA extension. *Molecular and Cellular Biology*, 28 (4), 1373-1382. doi: 10.1128/MCB.01651-07.

Manolagas, S.C. (2010). From estrogen-centric to aging and oxidative stress: a revised perspective of the pathogenesis of osteoporosis. *Endocrine Reviews*, 31 (3), 266-300. doi: 10.1210/er.2009-0024

Manuneehi, C.P., Cartland, S.P., & Kavurma, M.M. (2017). NADPH Oxidases, Angiogenesis, and Peripheral Artery Disease. *Antioxidants*, 6 (3), 1-12. doi: 10.3390/antiox6030056.

March, M.D., Merino, N., Barrera-Vilarmau, S., Crehuet, R., Onesti, S., Blanco, F.J., & Biasio, A.D. (2017). Structural basis of human PCNA sliding on DNA. *Nature Communications*, 8 (1), . doi: 10.1038/ncomms13935

Martinez, A., Knappskog, P.M., Olafsdottir, S., Døskeland, A.P., Eiken, H.G., Svebak, R.M. ... Flatmark, T. (1995). Expression of recombinant human phenylalanine hydroxylase as fusion protein in *Escherichia coli* circumvents proteolytic degradation by host cell proteases. Isolation and characterization of the wild-type enzyme. *The Biochemical Journal*, 306 (2), 589-597. . doi: 10.1042/bj3060589

Martínez-Jiménez, M.I., Lahera, A., & Blanco, L. (2017). Human PrimPol activity is enhanced by RPA. *Scientific reports*, 7 (783), . doi: 10.1038/s41598-017-00958-3.

Matsushima, S., Zablocki, D., Tsutsui, H., & Sadoshima, J. (2016). Poldip2 negatively regulates matrix synthesis at focal adhesions. *Journal of Molecular and Cellular Cardiology*, 94 (1), 10-12. doi: 10.1016/j.yjmcc.2016.03.001.

McCulloch, S.D. & Kunkel, T.A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Nature Publishing Group*, 18 (1), 148-161. doi: 10.1038/cr.2008.4.

McNulty, D.E., Claffee, B.A., Huddleston, M.J., Porter, M.L., Cavnar, K.M., & Kane, J.F. (2003). Mistranslational errors associated with the rare arginine codon CGG in *Escherichia coli*. *Protein Expression and Purification*, 27 (2), 365-374. . doi: 10.1016/S1046-5928(02)00610-1

McPherson, A. (2004). Introduction to protein crystallisation. *Methods*, 34 (3), 254-265. . doi: 10.1016/j.ymeth.2004.03.019

Mierzejewska, K., Siwek, W., Czapinska, H., Kaus-Drobek, M., Radlinska, M., Skowronek, K. ... (2014). Structural basis of the methylation specificity of R.DpnI. *Nucleic Acids Research*, 42 (13), 8745-8754. doi: 10.1093/nar/gku546.

Miroux, B. & Walker, J.E. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology*, 260 (3), 289-298. . doi: 10.1006/jmbi.1996.0399

- Mitarai, N., Jensen, M.H., & Semsey, S. (2015). Coupled Positive and Negative Feedbacks Produce Diverse Gene Expression Patterns in Colonies. *American Society for Microbiology*, 6 (2), e00059-15.
- Moldovan, G.L., Pfander, B., & Jentsch, S. (2007). PCNA, the maestro of the replication fork. *Cell*, 129 (4), 665-679. doi: 10.1016/j.cell.2007.05.003.
- Mondol, T., Stodola, J.L., Galletto, R., & Burgers, P.M. (2019). PCNA accelerates the nucleotide incorporation rate by DNA polymerase δ . *Nucleic Acids Research*, 47 (4), 1977-1986. doi: 10.1093/nar/gky1321.
- Moreno, M.U., Gallego, I., López, B., González, A., Fortuño, A., San José, G. ... Zalba, G. (2013). Decreased Nox4 levels in the myocardium of patients with aortic valve stenosis. *Clinical Science*, 125 (6), 291-300. doi: 10.1042/CS20120612.
- Murshudov, G.N., Vagin, A.A., & Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Biological Crystallography*, 53 (3), 240-255. doi: 10.1107/S09074444996012255.
- Naokatu, A., Takeshi, N., Akira, K., Hiroyuki, O., Yoshinori, K., Reika, A. ... Tomihiko, H. (2006). Regulation of mitochondrial morphology and cell survival by Mitogenin I and mitochondrial single-stranded DNA binding protein. *Elsevier*, 8 (1760), 1364-1372. doi:10.1016/j.bbagen.2006.05.012.
- Naryzhny, S.N. & Lee, H. (2010). Proliferating cell nuclear antigen in the cytoplasm interacts with components of glycolysis and cancer. *FEBS Letters*, 584 (20), 4292-4298. doi: 10.1016/j.febslet.2010.09.021.
- Nick McElhinny, S.A., Gordenin, D.A., Stith, C.M., Burgers, P.M., & Kunkel, T.A. (2008). Division of labor at the eukaryotic replication fork. *Molecular Cell*, 30 (2), 137-144. doi: 10.1016/j.molcel.2008.02.022.
- Nishimura, K., Apitz, J., Friso, G., Kim, J., Ponnala, L., Grimm, B., & Wijk, K. J. (2015). Discovery of a Unique Clp Component, ClpF, in Chloroplasts: A Proposed Binary ClpF-ClpS1 Adaptor Complex Functions in Substrate Recognition and Delivery. *The Plant Cell*, 27(10), 2677-2691. doi:10.1105/tpc.15.00574
- Paredes, F., Sheldon, K., Lassègue, B., Williams, H.C., Faidley, E.A., Benavides, G.A. ... San Martin, A. (2018). Poldip2 is an oxygen-sensitive protein that controls PDH and α KGDH lipoylation and activation to support metabolic adaptation in hypoxia and cancer. *Proceedings of the National Academy of Sciences*, 115 (8), 1789-1794. doi: 10.1073/pnas.1720693115.
- Paredes, F., Williams, H., & San-Martin, A. (2017). Poldip2 is an Oxygen-sensitive Mitochondrial Protein that Controls Oxidative/glycolytic Metabolism Balance and Proteasome Activity. *Free Radical Biology and Medicine*, 112 (1), 173-174. doi: 10.1016/j.freeradbiomed.2017.10.271.
- Perera, R.L., Torella, R., Klinge, S., Kilkenny, M.L., Maman, J.D., & Pellegrini, L. (2013). Mechanism for priming DNA synthesis by yeast DNA polymerase α . *eLife*, 2 (2), . doi: 10.7554/eLife.00482.

- Perkins, J.R., Diboun, I., Dessailly, B.H., Lees, J.G., & Orengo, C. (2010). Transient protein-protein interactions: structural, functional, and network properties. *Structure*, *18* (10), 1233-1243. . doi: 10.1016/j.str.2010.08.007
- Phizicky, E.M. & Fields, S. (1995). Protein-protein interactions: methods for detection and analysis. *Microbiological Reviews*, *59* (1), 94-123. . doi: PMC239356
- Plosky, B.S., Vidal, A.E., Fernández de Henestrosa, A.R., McLenigan, M.P., McDonald, J.P., Mead, S., & Woodgate, R. (2006). Controlling the subcellular localization of DNA polymerases iota and eta via interactions with ubiquitin. *The EMBO Journal*, *25* (12), 2847-2855. doi: 10.1038/sj.emboj.7601178.
- Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature*, *258* (5536), 598–599. doi: 10.1038/258598a0.
- Prakash, S., Johnson, R.E., & Prakash, L. (2005). Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annual Review of Biochemistry*, *74* (1), 317-353. doi: 10.1146/annurev.biochem.74.082803.133250.
- Prindle, M.J. & Loeb, L.A. (2012). DNA polymerase delta in DNA replication and genome maintenance. *Environmental and molecular mutagenesis*, *53* (9), 666-682. doi: 10.1002/em.21745.
- Pursell, Z.F., Isoz, I., Lundström, E.B., Johansson, E., & Kunkel, T.A. (2007). Yeast DNA Polymerase ϵ Participates in Leading-Strand DNA Replication. *Science*, *317* (5834), 127-130. doi: 10.1126/science.1144067.
- Raran-Kurussi, S. & Waugh, D.S. (2012). The ability to enhance the solubility of its fusion partners is an intrinsic property of maltose-binding protein but their folding is either spontaneous or chaperone-mediated. *PloS One*, *7* (11), e49589. . doi: 10.1371/journal.pone.0049589
- Raran-Kurussi, S. & Waugh, D.S. (2016). A Dual Protease Approach for Expression and Affinity Purification of Recombinant Proteins. *Analytical Biochemistry*, *504* (1), 30-37. . doi: 10.1016/j.ab.2016.04.006
- Robichon, C., Luo, J., Causey, T.B., Benner, J.S., & Samuelson, J.C. (2011). Engineering Escherichia coli BL21 (DE3) derivative strains to minimize E. coli protein contamination after purification by immobilized metal affinity chromatography. *Applied and Environmental Microbiology*, *77* (13), 4634-4646. . doi: 10.1128/AEM.00119-11
- Rodrigo, S.F., Andrea, C.F.F., Fabio, H., Corinne, D., & Denise, P.C. (2014). Sexual dimorphism and thyroid dysfunction: a matter of oxidative stress? *Journal of Endocrinology*, *221* (2), R31-R40. doi: DOI: 10.1530/JOE-13-0588
- Rosano, G.L. & Ceccarelli, E.A. (2014). Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in Microbiology*, *5* (172). doi: 10.3389/fmicb.2014.00172
- Rosenberg, A.H., Lade, B.N., Chui D.S., Lin, S.W., Dunn, J.J., & Studier, F.W. (1987). Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene*, *56* (1), 125-135. . doi: 10.1016/0378-1119(87)90165-x

- Rossmann, M.G., Moras, D., & Olsen, K.W. (1974). Chemical and biological evolution of nucleotide-binding protein. *Nature*, 250 (463), 194-199. doi: 10.1038/250194a0.
- Ryazanov, A.G., Ward, M.D., Mendola, C.E., Pavur, K.S., Dorovkov, M.V., Wiedmann, M. ... Hait, W.N. (1997). Identification of a new class of protein kinases represented by eukaryotic elongation factor-2 kinase. *Proceedings of the National Academy of Sciences*, 94 (10), 4884-4889. doi: 10.1073/pnas.94.10.4884.
- Sabbioneda, S., Green, C.M., Bienko, M., Kannouche, P., Dikic, I., & Lehmann, A.R. (2009). Ubiquitin-binding motif of human DNA polymerase eta is required for correct localization. *Proceedings of the National Academy of Sciences*, 106 (8), 20-20. doi: 10.1073/pnas.0812744106.
- Sakashita, G., Kiyoi, H., Naoe, T., & Urano, T. (2018). Analysis of the oligomeric states of nucleophosmin using size exclusion chromatography. *Scientific Reports*, 4008 (8). doi: :10.1038/s41598-018-22359-w
- Sale, J.E. (2013). Translesion DNA synthesis and mutagenesis in eukaryotes. *Cold Spring Harbor Perspectives in Biology*, 5 (3). doi: 10.1101/cshperspect.a012708.
- Salmeen, A., Park, B.O., & Meyer, T. (2010). The NADPH oxidases NOX4 and DUOX2 regulate cell cycle entry via a p53-dependent pathway. *Nature*, 29 (31), 4473-4484. doi: 10.1038/onc.2010.200.
- Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual (3rd ed) Cold Spring Harbor Laboratory Press N.Y, ISBN: 10 0-87969-577-3.
- San-Miguel, T., Pérez-Bermúdez, P., & Gavidia, I. (2013). Production of soluble eukaryotic recombinant proteins in *E. coli* is favoured in early log-phase cultures induced at low temperature. *SpringerPlus*, 2 (89),. doi: 10.1186/2193-1801-2-89
- Sarvestani, A., Walenta, A.A.H., Busetto, A.E., Lausib, B.A., & Fourmec, R. (1998). Simulation and Measurement of X-ray Diffraction from Single Crystals and Evaluation of Optimised Data Collection. *Journal of Applied Crystallography*, 31, 899-909. doi: 10.1107/S0021889898007833.
- Savitsky, P., Bray, J., Cooper, C.D., Marsden, B.D., Mahajan, P., Burgess-Brown, N.A., & Gileadi, O. (2010). High-throughput production of human proteins for crystallisation: The SGC experience. *Journal of Structural Biology*, 172 (1), 3-13. doi: 10.1016/j.jsb.2010.06.008.
- Schein, C.H. (1989). Production of Soluble Recombinant Proteins in Bacteria. *Nature Biotechnology*, 7 (1), 1141-1149. Retrieved from <http://www.nature.com/naturebiotechnology>
- Schmitt, M.W., Matsumoto, Y., & Loeb, L.A. (2009). High fidelity and lesion bypass capability of human DNA polymerase delta. *Biochimie*, 91 (9), 1163-1172. doi: 10.1016/j.biochi.2009.06.007.
- Schröder, K. (2019). NADPH oxidases in bone homeostasis and osteoporosis. *Free Radical Biology and Medicine*, 132 (1), 67-72. doi: 10.1016/j.freeradbiomed.2018.08.036.
- Schwartz, R., Ting, C.S., & King, J. (2001). Whole Proteome pI Values Correlate with Subcellular Localizations of Proteins for Organisms within the Three Domains of Life. *Genome Res*, 11 (5), 703-709. doi: 10.1101/gr-1587r.

Shevelev, I.V. & Hübscher, U. (2002). The 3'-5' exonucleases. *Nature reviews. Molecular Cell Biology*, 3 (5), 364-376. doi: 10.1038/nrm804.

Shimuta, T.R., Nakano, K., Yamaguchi, Y., Ozaki, S., Fujimitsu, K., Matsunaga, C., Noguchi, K., Emoto, A., Katayama, T. (2004). Novel heat shock protein HspQ stimulates the degradation of mutant DnaA protein in Escherichia coli. *Cells* 9(12), 1151–1166

Sitsel, O. & Raunser, S. (2019). Big insights from tiny crystals. *Nature Chemistry*, 11 (2), 106-108. . doi: 10.1038/s41557-019-0211-3

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., & Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein Science*, 18 (5), 936-948. . doi: 10.1002/pro.102

Siwek, W., Czapinska, H., Bochtler, M., Bujnicki, J.M., & Skowronek, K. (2012). Crystal structure and mechanism of action of the N6-methyladenine-dependent type IIM restriction endonuclease R.DpnI. *Nucleic Acid Research*, 40 (15), 7563-7572. doi: 10.1093/nar/gks428.

Snell, E.H., Nagel, R.M., Wojtaszczyk, A., Neill, H.O., Wolfley, J.L., & Luft, J.R. (2008). The application and use of chemical space mapping to interpret crystallisation screening results. *Structural Biology*, 64 (1), 1240-1249. . doi: 10.1107/s0907444908032411

Sørensen, H.P. & Mortensen, K.K. (2005). Advanced genetic strategies for recombinant protein expression in Escherichia coli. *Journal of Biotechnology*, 115 (2), 113-128. . doi: 10.1016/j.jbiotec.2004.08.004

Sørensen, H.P., Sperling-Petersen, H.U., & Mortensen, K.K. (2003). Production of recombinant thermostable proteins expressed in Escherichia coli: completion of protein synthesis is the bottleneck. *Journal of Chromatography*, 786 (2), 207-214. . doi: 10.1016/S1570-0232(02)00689-X

Stein, A., Pache, R.A., Bernadó, P., Pons, M., & Aloy, P. (2009). Dynamic interactions of proteins in complex networks: a more structured view. *The FEBS Journal*, 276 (19), 5390-405. . doi: 10.1111/j.1742-4658.2009.07251.x

Stodola, J.L. & Burgers, P.M. (2016). Resolving individual steps of Okazaki-fragment maturation at a millisecond timescale. *Nature*, 23 (5), 402-408. doi: 10.1038/nsmb.3207.

Stodola, J.L. & Burgers, P.M. (2017). Mechanism of Lagging-Strand DNA Replication in Eukaryotes. *Advances in experimental medicine and biology*, 1042 (1), 117-133. doi: 10.1007/978-981-10-6955-0_6.

Stoimenov, I. & Helleday, T. (2009). PCNA on the crossroad of cancer. *Biochemical Society Transactions*, 37 (3), 605-613. doi: 10.1042/BST0370605.

Studier, F.W. & Moffatt, B.A. (1986). Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-level Expression of Cloned Genes. *Journal of Molecular Biology*, 189 (1), 113-130. . doi: 10.1016/0022-2836(86)90385-2

Sun, J., Shi, Y., Georgescu, R.E., Yuan, Z., Chait, B.T., Li, H., & O'Donnell, M.E. (2015). The architecture of a eukaryotic replisome. *Nature Structural and Molecular Biology*, 22 (12), 976-982. doi: 10.1038/nsmb.3113.

- Sutliff, R., Hilenski, L.L., Amanso, A.M., Parastatidis, I., Dikalova, A.E., Hansen, L. ... Lassègue, B. (2013). Poldip2 sustains vascular structure and function. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 33 (9), doi: 10.1161/ATVBAHA.113.301913.
- Sykora, P., Kanno, S., Akbari, M., Kulikowicz, T., Baptiste, A. B., Leandro, G.S., Lu,H., Tian,, J., May, A., Becker, K.A., Croteau, D.L., Wilson, D.M., Sobol, R.W., Yasui, A., Bohr, V.A. (2017). DNA Polymerase Beta Participates in Mitochondrial DNA Repair. *Molecular and Cellular Biobiology*, 37(16), e00237-17
- Tabor, S. & Richardson, C.C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proceedings of the National Academy of Sciences*, 82 (4), 1074-1078. Retrieved from 10.1073/pnas.82.4.1074
- Taylor, G. (2003). The phase problem. *Biological Crystallography*, 59 (11), 1881-1890. Retrieved from <https://onlinelibrary.wiley.com/iucr/doi/10.1107/S0907444903017815>.
- Thompson, R.F., Walker, M., Siebert, C.A., Muench, S.P., & Ranson, N.A. (2016). An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods*, 100 (1), 3-15. . doi: 10.1016/j.ymeth.2016.02.017
- Tissier, A., Janel-Bintz, R., Coulon, S., Klaile, E., Kannouche, P., Fuchs, R.P., & Cordonnier , A.M. (2010). Crosstalk between replicative and translesional DNA polymerases: PDIP38 interacts directly with Pol eta. *DNA Repair*, 9 (8), 922-928. doi: 10.1016/j.dnarep.2010.04.010.
- Tompa, P. & Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends in Biochemical Sciences*, 33 (1), 2-8.. doi: 10.1016/j.tibs.2007.10.003
- Torregrosa-Muñumer, R., Forslund , J.M.E., Goffart, S., Pfeiffer, A., Stojkovič, G., Carvalho, G. ... Pohjoismäki, J.L.O. (2017). PrimPol is required for replication reinitiation after mtDNA damage. *PNAS*, 114 (43), 11398-11403. doi: 10.1073/pnas.1705367114.
- Tsuchimoto, D., Sakai, Y., Sakumi, K., Nishioka, K., Sasaki, M., Fujiwara, T., & Nakabeppu, Y. (2001). Human APE2 protein is mostly localized in the nuclei and to some extent in the mitochondria, while nuclear APE2 is partly associated with proliferating cell nuclear antigen. *Nucleic Acids Research*, 29 (11), 2349-2360. doi: 10.1093/nar/29.11.2349.
- Tsurimoto, T., Melendy, T., & Stillman, B. (1990). Sequential initiation of lagging and leading strand synthesis by two different polymerase complexes at the SV40 DNA replication origin. *Nature*, 346 (6284), 534-539. doi: 10.1038/346534a0.
- Turk, H.F. & Chapkin, R.S. (2015). Analysis of Epidermal Growth Factor Receptor Dimerization by BS3 Cross-Linking. *Methods in Molecular Biology*, 1233 (1), 25-34. . doi: 10.1007/978-1-4939-1789-1_3
- Vaisman, A., Masutani, C., Hanaoka, F., & Chaney, S.G. (2000). Efficient translesion replication past oxaliplatin and cisplatin GpG adducts by human DNA polymerase eta. *Biochemistry*, 39 (16), 4575-4580. doi: 10.1021/bi000130k.
- Vasina, J.A. & Baneyx, F. (1996). Recombinant protein expression at low temperatures under the transcriptional control of the major Escherichia coli cold shock promoter cspA.. *Applied and*

Environmental Microbiology, 62 (4), 1444-1447. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC167914/>

Vedadi, M., Arrowsmith, C.H., Allali-Hassani, A., Senisterra, G., & Wasney, G.A. (2010). Biophysical characterization of recombinant proteins: a key to higher structural genomics success. *Journal of Structural Biology*, 172 (1), 107-119. . doi: 10.1016/j.jsb.2010.05.005

Vera, A., González-Montalbán, N., Arís, A., & Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnology and Bioengineering*, 96 (6), 1101-1106. . doi: 10.1002/bit.21218

Vethanayagam, J.G.G. & Flower, A.M. (2005). Decreased gene expression from T7 promoters may be due to impaired production of active T7 RNA polymerase. *Microbial Cell Factories*, 4 (3), 1-7. . doi: 10.1186/1475-2859-4-3

Vidal, A.E., Kannouche, P., Podust, V.N., Yang, W., Lehmann, A.R., & Woodgate, R. (2004). Proliferating cell nuclear antigen-dependent coordination of the biological functions of human DNA polymerase ϵ . *The Journal of Biological Chemistry*, 279 (46), 48360-48368. doi: 10.1074/jbc.M406511200.

Waga, S., Hannon, G.J., Beach, D., & Stillman, B. (1994). The p21 inhibitor of cyclin-dependent kinases controls DNA replication by interaction with PCNA. *Nature*, 369 (6481), 574-578. doi: 10.1038/369574a0.

Walter, R.L., Thiel, D.J., Barna, S.L., Tate, M.W., Wall, M.E., Eikenberry, E.F. ... Ealick, S.E. (1995). High-resolution macromolecular structure determination using CCD detectors and synchrotron radiation. *Structure*, 3 (8), 835-844. doi: 10.1016/S0969-2126(01)00218-0.

Weyemi, U., Lagente-Chevallier, O., Boufraquech, M., Prenois, F., Courtin, F., Caillou, B. ... Dupuy, C. (2012). ROS-generating NADPH oxidase NOX4 is a critical mediator in oncogenic H-Ras-induced DNA damage and subsequent senescence. *Oncogene*, 31 (9), 1117-1129. doi: 10.1038/onc.2011.327.

Whitmore, L. & Wallace, B.A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, 32 (Web Server issue). doi: 10.1093/nar/gkh371.

Whitmore, L. & Wallace, B.A. (2008). Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases. *Biopolymers*, 89 (5), 392-400. doi: 10.1002/bip.20853.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R. ... Wilson, K.S. (2011). Overview of the CCP4 suite and current developments. *Biological Crystallography*, 67 (4), 235-242. . doi: 10.1107/S0907444910045749

Witko-Sarsat, V., Mocek, J., Bouayad, D., Tamassia, N., Ribeil, J.A., Candlish, C. ... Cassatella, M.A. (2010). Proliferating cell nuclear antigen acts as a cytoplasmic platform controlling human neutrophil survival. *The Journal of Experimental Medicine*, 207 (12), 2631-2345. doi: 10.1084/jem.20092241.

- Wohlgemuth, A., Lenz, C., & Urlaub, H. (2015). Studying macromolecular complex stoichiometries by peptide-based mass spectrometry. *Proteomics*, *15* (5), 862–879. . doi: 10.1002/pmic.201400466
- Xie, B., Li, H., Wang, Q., Xie, S., Rahmeh, A., Dai, W., & Lee, M.Y. (2005). Further characterization of human DNA polymerase delta interacting protein 38. *The Journal of Biological Chemistry*, *280* (23), 22375-22384. doi: 10.1074/jbc.M414597200.
- Yu, J.T. & Guo, M.Z. (2008). Prediction of Protein-Protein Interactions from Secondary Structures in Binding Motifs Using the Statistic Method. Fourth International Conference on Natural Computation, China, 1, 100-103. doi: 10.1109/ICNC.2008.451
- Zhang, M., Brewer, A.C., Schröder, K., Santos, C.X., Grieve, D.J., Wang, M. ... Shah, A.M. (2010). NADPH oxidase-4 mediates protection against chronic load-induced stress in mouse hearts by enhancing angiogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *107* (42), 18121-18126. doi: 10.1073/pnas.1009700107.
- Zhang, Y., Yuan, F., Wu, X., & Wang, Z. (2000). Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase iota. *Molecular and Cellular Biology*, *20* (19), 7099-7108. doi: 10.1128/mcb.20.19.7099-7108.2000.
- Zhuang, Z., Johnson, R.E., Haracska, L., Prakash, L., Prakash, S., & Benkovic, S.J. (2008). Regulation of polymerase exchange between Pol η and Pol δ by monoubiquitination of PCNA and the movement of DNA polymerase holoenzyme. *Proceedings of the National Academy of Sciences*, *105* (14), 5361- 5366. doi: 10.1073/pnas.0801310105