
The molecular biology of DNA replication in the archaeon *Sulfolobus solfataricus*

Thomas R. Beattie

Green Templeton College and
Sir William Dunn School of Pathology
University of Oxford



Thesis submitted for the degree of Doctor of Philosophy
Trinity Term 2012

Table of contents

Table of contents	i
Acknowledgments	iv
Abbreviations	v
Abstract	viii
Chapter 1 – Introduction	1
DNA replication	1
Archaea – the third domain of life	3
The replisome	7
Okazaki fragment maturation	10
Proliferating cell nuclear antigen (PCNA)	20
Thesis aims	31
Chapter 2 – Materials and Methods	33
Strains	33
Cloning	34
Protein expression and purification	38
Protein labelling	44
Substrate preparation	45
Enzyme activity assays	47
DNA binding assays	49
Protein-protein interaction assays	51
<i>Sulfolobus</i> growth and analysis	52
Data visualisation and analysis	54

Chapter 3 – <i>In vitro</i> reconstitution of archaeal Okazaki fragment maturation	55
Abstract	55
Introduction	55
Results	57
Discussion	81
Chapter 4 – The mechanism of multiple enzyme coordination by PCNA in Okazaki fragment maturation	84
Abstract	84
Introduction	84
Results	87
Discussion	96
Chapter 5 – Conformational changes in Lig1 during Okazaki fragment maturation	100
Abstract	100
Introduction	100
Results	105
Discussion	125
Chapter 6 – Identification of two novel PolB1-interacting proteins	128
Abstract	128
Introduction	128
Results	135
Discussion	156
Chapter 7 – Discussion	162
Main conclusions	162
Implications in a broader context	163

Table of contents

Limitations and future work	169
References	174
Appendix I – Oligonucleotide sequences	197
Primers used for cloning	197
Oligonucleotides used for substrates	198

Acknowledgments

So many people have contributed to making my DPhil such a fantastic experience over the past four years.

First and foremost I would like to thank my supervisor, Steve Bell, from whom I have learnt so much. Steve's knowledge, enthusiasm and insight have been invaluable throughout my DPhil. He has always made himself available when I've needed input and advice, and I am hugely grateful for his constant support.

The Bell lab has been such a great place to work, both intellectually and socially, and I would like to thank all the people that have made this possible. Thanks to Alessandro, Iain, Liza, Matt, and Rachel for making me feel so welcome from the beginning, and to Ben, Cat, Hara, Helen, Li-Phing, Marek, Marleen, Maria, Nelly, Thomas, and Tamzin for continuing to make the lab both stimulating and fun. Thank you all for your moral and technical support, knowledge, and humour.

I would also like to thank our various collaborators for their help and expertise. Thanks to Achillefs Kapanidis, Johannes Holbein and Robert Crawford for their advice and technical help with FRET and fluorescence anisotropy experiments. Thanks to Laura Spagnolo and Giuseppe Cannone for their enthusiasm and tenacity in cracking such a tough structural problem. Thanks also to Susan Lea, Steven Johnson and Joe Caesar for their advice and technical help with various protein purification and biophysical techniques.

To my parents and my brother Jonathan, thank you for being such a wonderful family; for supporting me in whatever I do, for never, ever pushing me and for always taking an interest.

Finally, Claire, it is impossible to describe how much your continual support has meant to me over the last four years. Your patience and encouragement have been amazing, and your intelligence and humour have been a constant source of inspiration. Thank you for everything.

Abbreviations

9-1-1	Rad9-Rad1-Hus1
AAA+	ATPases associated with diverse cellular activities
AD	GAL4 activation domain
AdD	Adenylation domain (Lig1)
ALEX	Alternating laser excitation spectroscopy
ATP	Adenosine triphosphate
AU	Absorbance units
BD	GAL4 DNA binding domain
BER	Base excision repair
BLAST	Basic local alignment search tool
BRCT	BRCA1 C-terminus
BSA	Bovine serum albumin
CMG	Cdc45-Mcm2–7-GINS
dATP	Deoxyadenosine triphosphate
DBD	DNA binding domain (Lig1)
dC	Deoxycytosine
dCTP	Deoxycytosine triphosphate
ddC	Dideoxycytosine
dGMP	Deoxyguanosine monophosphate
dGTP	Deoxyguanosine triphosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
ds	Double-stranded
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
EDTA	Ethylenediaminetetraacetic acid

Abbreviations

EMSA	Electrophoretic mobility shift assay
ExoIII	Exonuclease III
Fen1	Flap endonuclease 1
FRET	Fluorescence resonance energy transfer
gDNA	Genomic DNA
GIN5	Go-ichi-ni-san; Sld5-Psf1-Psf2-Psf3
GST	Glutathione S-transferase
HhH	Helix-hairpin-helix
IDCL	Interdomain connector loop
IPTG	Isopropyl β -D-thiogalactopyranoside
Kb	Kilobases
K _D	Dissociation constant
kDa	Kilodalton
LC	Liquid chromatography
Lig1	DNA ligase I
LUCA	Last universal common ancestor
MCM	Minichromosome maintenance
Min	Minutes
MMR	Mismatch repair
MS	Mass spectrometry
NAD	Nicotinamide adenine dinucleotide
nm	Nanometres
nt	Nucleotides
NTA	Nitrilotriacetic acid
OB	Oligonucleotide/oligosaccharide -binding
OBD	OB-fold domain (Lig1)
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis

Abbreviations

PBP1	PolB1-binding protein 1
PBP2	PolB1-binding protein 2
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDB	Protein databank
PDE I	Phosphodiesterase I
PIP	PCNA-interacting protein
Pol	DNA polymerase
PolB1	DNA polymerase B1
RFC	Replication factor C
RNA	Ribonucleic acid
RPA	Replication protein A
RPC	Replisome progression complex
rpm	Revolutions per minute
rRNA	Ribosomal RNA
SAXS	Small-angle x-ray scattering
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
ss	Single-stranded
SSB	Single-stranded DNA-binding protein
SUMO	Small ubiquitin-like modifier
TBE	Tris-borate EDTA
TBP	TATA-box binding protein
TCA	Trichloroacetic acid
UDG1	Uracil DNA glycosylase 1
UV	Ultraviolet
WT	Wild-type
XPF	Xeroderma pigmentosum complementation group F

Abstract

DNA replication is essential for the propagation of all living organisms. The ability of a cell to accurately duplicate its entire genome is dependent upon the activity of numerous proteins. Identifying the molecular mechanisms by which these proteins act, and determining how they are physically and functionally coordinated at sites of active DNA replication, is central to understanding this essential cellular process. Archaea possess a DNA replication machinery which is ancestral to the one present in eukaryotes, and thus these organisms serve as simplified model systems for understanding the complexities of eukaryotic DNA replication.

This thesis investigates the molecular mechanisms underlying Okazaki fragment maturation in the crenarchaeon *Sulfolobus solfataricus*, which is essential to the completion of lagging strand DNA replication. Reconstitution of Okazaki fragment maturation *in vitro* demonstrated that the activities of three enzymes – PolB1, Fen1, and Lig1 – are required for this process in *S. solfataricus*. Furthermore, it was shown that optimum coordination of their three distinct activities is dependent on the ability of PolB1, Fen1 and Lig1 to simultaneously interact with a single PCNA ring, providing evidence for a mechanism of multi-enzyme coordination which may be universally employed by DNA sliding clamp proteins. The importance of protein flexibility in the accommodation of multiple proteins around a single PCNA was also investigated. Finally, the physical coordination of one of these key maturation enzymes – PolB1 – with other replisome proteins was examined. It was demonstrated that PolB1 exists in a trimeric complex *in vivo* with two previously unidentified factors, raising the possibility of uncharacterised activities and interactions for this crucial enzyme.

Taken together, these data provide new insights into functionally important protein-protein interactions within the archaeal replisome, and facilitate a greater understanding of the DNA replication machinery in both archaea and eukaryotes.

Chapter 1 – Introduction

DNA REPLICATION

Deoxyribonucleic acid (DNA) dictates the properties and activities of every living cell and therefore its replication is fundamental to the propagation of all living organisms. A large number of protein factors have evolved to catalyse DNA replication and overcome the numerous mechanistic obstacles intrinsic to this essential cellular process. Topological problems which are inherent to the manipulation and synthesis of DNA molecules must be resolved by the replication machinery (Vos et al, 2011). Replication must also be able to respond appropriately to the presence of DNA damage and competing DNA transactions such as transcription (Branzei & Foiani, 2010), and furthermore, the chromatin structure of DNA in eukaryotic cells must not only be overcome, but preserved during replication to facilitate epigenetic as well as genetic inheritance (Alabert & Groth, 2012). Despite these difficulties, it is essential that DNA replication is extremely accurate and efficient; the 3 billion bases of the human genome for example must be replicated, without error, in as little as one hour (Bell & Dutta, 2002). The demands upon the protein machinery of DNA replication are therefore considerable, and the molecular mechanisms underlying its essential functions are of significant interest.

Although DNA was first isolated from human cells in 1869 (Miescher, 1871), it was not until the middle of the twentieth century that the fundamental importance of this macromolecule was demonstrated. The concept of a genetic material that determines the inheritance of specific phenotypes through discrete units had first been proposed by Gregor Mendel (Mendel, 1865), but the molecular basis of this genetic material remained unclear for many years. However, pioneering experiments in 1944 demonstrated that purified DNA was sufficient to mediate the transformation of a non-virulent *Streptococcus pneumonia* strain into

a virulent form and thus was proposed as the genetic material (Avery et al, 1944). This hypothesis was confirmed by experiments which demonstrated that bacteriophage T2 infection of *Escherichia coli* involved transmission of only the DNA component of the virus into the host cell (Hershey & Chase, 1952).

The first insights into the mechanism of DNA function were provided by the landmark determination of the structure of the DNA double helix by Watson and Crick in 1953 (Watson & Crick, 1953a), when the molecular basis of specific base pairing between two antiparallel DNA strands was observed for the first time (Figure 1.1). This served as the foundation for a large body of subsequent work which elucidated the ‘central dogma of molecular biology’ as the mechanism by which DNA operates as the genetic material within the cell: DNA acts as a template for messenger RNA synthesis which subsequently encodes protein synthesis and thus determines the structure and function of the cell (Crick, 1970).

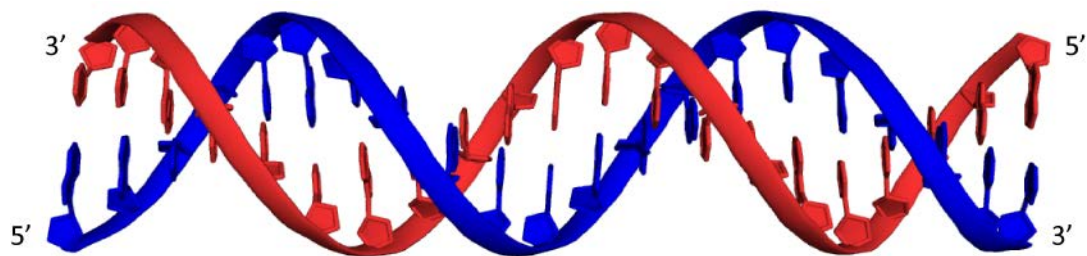


Figure 1.1 – The structure of double-stranded DNA

Crucially, the structure of DNA also potentially revealed the mechanistic basis of its heritability between subsequent generations. As Watson and Crick described, “*It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material*”. It was hypothesised that double-stranded (ds) DNA provides two template strands upon each of which a new strand is synthesised – directed by specific base-pairing – thus giving rise to two identical daughter

duplexes (Watson & Crick, 1953b). Key experiments subsequently confirmed that chromosomal DNA is indeed replicated through this semi-conservative mechanism, with each new daughter DNA molecule composed of one strand of parental DNA, and one newly synthesised strand (Meselson & Stahl, 1958).

One original hypothesis was that replication by this mechanism was an intrinsic property of the DNA molecule. However, it subsequently became clear that DNA replication is an enzyme-dependent process when Arthur Kornberg and colleagues identified DNA polymerase (Pol) I as an enzyme capable of synthesising DNA from deoxyribonucleotides in a template-dependent manner (Bessman et al, 1958; Lehman et al, 1958). In the following decades numerous additional protein factors were identified as essential for performing the huge number of activities required during initiation, elongation and termination of chromosomal DNA replication. Interestingly, despite the absolute structural conservation of DNA and the universal chemistry of DNA synthesis, two distinct cellular DNA replication machineries have evolved (Edgell & Doolittle, 1997); thus while bacteria possess one set of protein factors to catalyse the different stages of DNA replication, eukaryotes possess an entirely distinct set of proteins which perform analogous functions. It has become apparent over the last decade that archaea, the third domain of life, possess DNA replication proteins that are homologous to those of eukaryotes rather than bacteria.

ARCHAEA – THE THIRD DOMAIN OF LIFE

A pivotal discovery was made by Carl Woese and colleagues in 1977, when they discovered by sequencing ribosomal RNA (rRNA) molecules that several methanogens, previously characterised as bacteria, possessed completely unique sequence characteristics (Fox et al, 1977; Woese & Fox, 1977). It was therefore proposed that three domains of life have evolved during life on earth: bacteria, eukaryotes and the third domain, archaea. Many archaea are

morphologically similar to bacteria – hence their relatively late discovery – but in support of their distinct sequence features, archaea also possess some notable phenotypic differences from bacteria: for example, archaeal cell membranes uniquely contain ether-linked lipids, while archaeal cell walls lack the peptidoglycan present in bacteria (Kletzin, 2007). Currently, the best accepted model for the phylogeny of all living organisms proposes a tree of life, whereby the archaea and eukaryotes share a common ancestor which diverged from the bacteria earlier in time (Figure 1.2) (Woese et al, 1990). Alternative models such as the ring of life – in which the earliest eukaryote was the result of a fusion between an archaeal and a bacterial organism – have also been proposed (Rivera & Lake, 2004).

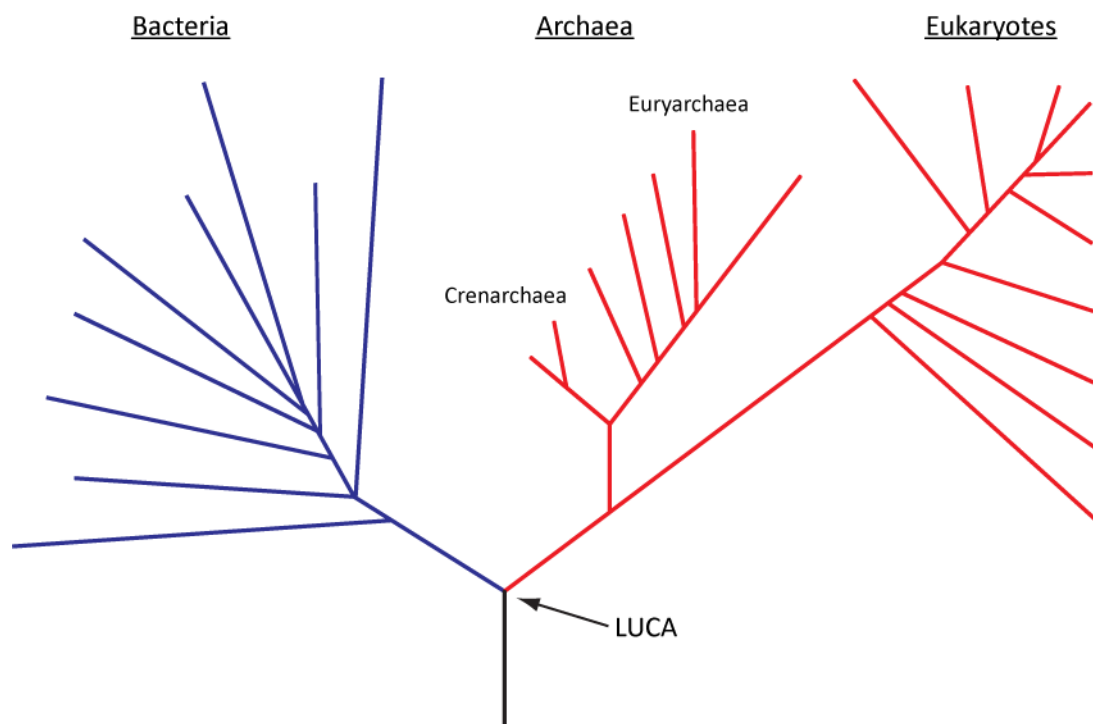


Figure 1.2 – The universal phylogenetic tree of life

Branch positions and branch lengths are based on rRNA sequence comparison. The three domains of life are indicated, and the main phyla within the archaea are also shown. The proposed last universal common ancestor (LUCA) of all cellular life is indicated. Adapted from (McGeoch & Bell, 2008).

Initial rRNA analyses demonstrated the existence of two phyla within the archaeal population, defined as the crenarchaea and euryarchaea (Woese et al, 1990); this phylogeny has subsequently been confirmed by abundant archaeal genome sequences together with various experimentally determined phenotypic differences (Brochier-Armanet et al, 2011). Additionally, two further archaeal phyla have been proposed on the basis of sequence analysis – korarchaea (Barns et al, 1996) and thaumarchaea (Brochier-Armanet et al, 2008) – although it should be noted that there is still significant debate about the exact placement of these phyla within archaeal evolution. A fifth phylum, nanoarchaea, was also previously proposed based on the sequence of a single representative, *Nanoarchaeum equitans*, but this species is now widely regarded to be a highly diverged member of the euryarchaea most closely related to the thermococcales (Brochier et al, 2005). While archaea are often characterised solely as extremophiles, in recent years it has become apparent – particularly with improvements in the sequencing of environmental samples – that archaea exist in a huge variety of different ecosystems, and are certainly not restricted to extreme environments. Notably however, no archaeal pathogens have been identified to date (Robertson et al, 2005).

The first complete archaeal genome sequence – that of the euryarchaeon *Methanococcus jannaschii* – demonstrated that archaea possess proteins with homology to the DNA replication machinery of eukaryotes, but not to bacterial replication proteins (Bult et al, 1996). Various theories have been proposed for how two distinct replication machineries have evolved across the three domains of life (reviewed in McGeoch & Bell, 2008). One proposal is that the last universal common ancestor (LUCA) of all extant cellular life forms possessed an RNA genome and thus a DNA genome and the machinery to replicate it evolved twice independently, subsequent to the divergence of bacteria and the common ancestor of archaea and eukaryotes. Alternatively, it is possible that LUCA possessed a mosaic replication apparatus with components of both current machineries, and this was subsequently differentially streamlined in the bacterial and archaeal-eukaryotic lineages. Finally, the

‘replicon takeover hypothesis’ proposes that both lineages inherited the same DNA replication machinery from LUCA, but this was subsequently displaced in just one of the lineages, for example by integration of viral sequences into a host chromosome. Regardless of its evolution, it is clear that the archaeal DNA replication machinery is a simplified, ancestral form of that present in eukaryotes, and studies in the third domain of life have thus provided significant insight into the more complex process of DNA replication in eukaryotic organisms (Barry & Bell, 2006).

In this respect, a particularly well-studied and useful model organism is the crenarchaeon *Sulfolobus solfataricus*. *S. solfataricus* is a hyperthermophilic acidophile, initially isolated from hot acidic springs (Zillig et al, 1980). This organism undergoes a eukaryotic-like cell cycle, strictly alternating between a single copy of the chromosome in a short G1 phase following cell division, and two copies of the chromosome in a long G2 phase following DNA replication, during which chromosomes appear to undergo a period of post-replicative pairing (Robinson et al, 2007; Samson & Bell, 2011). These cell cycle characteristics may be common to many crenarchaeal species (Lundgren et al, 2008). As with all archaea identified to date, *S. solfataricus* possesses a single circular chromosome. However, the *S. solfataricus* chromosome possesses three origins of replication (Lundgren et al, 2004; Robinson et al, 2004; Robinson & Bell, 2007), and thus DNA replication initiates from multiple sites in this organism, similarly to eukaryotes but distinct from bacteria and some species of archaea which possess only a single origin of replication (Myllykallio et al, 2000). The genome of *S. solfataricus* has been fully sequenced (She et al, 2001) and furthermore cell cycle synchronisation (Duggin et al, 2008), microscopy (Samson et al, 2008), and genetic tools (Wagner et al, 2009b) have recently been developed for *S. solfataricus* and related species. Therefore this organism serves as an excellent model system for investigating the complex mechanisms underpinning eukaryotic DNA replication.

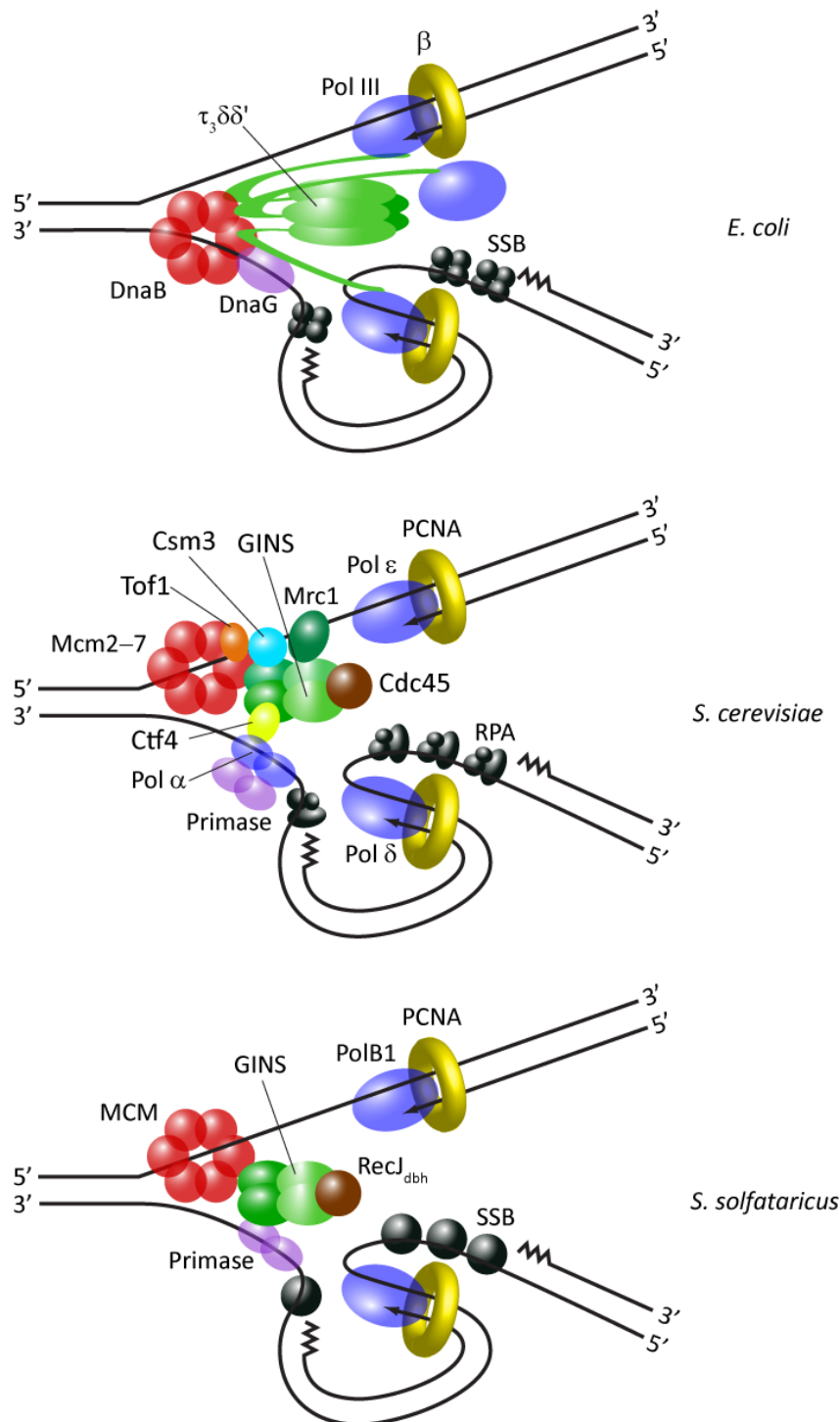
THE REPLISOME

The mechanism of semi-conservative DNA replication common to all cellular life forms provides an inherent directionality problem: DNA polymerases are only able to synthesise DNA in the 5'-3' direction (Kornberg & Baker, 1992), and yet duplex DNA is anti-parallel, thus providing two template strands of opposing polarity as it is unwound. The solution to this problem was demonstrated by the identification of short DNA strand intermediates that arise during chromosomal replication (Sakabe & Okazaki, 1966). Thus one strand – the leading strand – is synthesised continuously while the other strand – the lagging strand – is synthesised discontinuously in a series of Okazaki fragments; the lagging strand effectively being synthesised 'backwards' in short steps. Furthermore, the inability of DNA polymerases to initiate DNA synthesis *de novo* (Kornberg & Baker, 1992) requires that each of these Okazaki fragments is initiated by a short RNA primer, the 3' end of which can subsequently be extended. Okazaki fragments have been detected in all cellular organisms, although notably they are significantly larger in bacteria (1–2 kilobases) than in eukaryotes and archaea (100–200 nucleotides) (Edenberg & Huberman, 1975; Matsunaga et al, 2003).

Coupled leading and lagging strand synthesis at a single unwinding replication fork requires the coordination of numerous enzymatic activities. On the basis of early studies on bacteriophage T4 replication, it was proposed that such coordination could be mediated by the physical association of all the required protein factors into a single molecular machine (Sinha et al, 1980). This has subsequently been confirmed in bacteriophages T4 and T7 and the bacterium *E. coli*, with the multi-protein DNA replication machine termed the replisome (Benkovic et al, 2001). In *E. coli* for example, all the proteins essential for semi-discontinuous DNA synthesis have been shown to physically associate through a series of well characterised protein-protein interactions; these include the DnaB helicase for separating template DNA strands, the DnaG primase for synthesising initiating RNA primers, two Pol III cores for

performing leading and lagging strand DNA synthesis, the β sliding clamp together with the $\tau_3\delta\delta'\psi\chi$ clamp loader for supporting Pol III activity, and single-stranded (ss) DNA-binding protein (SSB) for protecting and resolving ssDNA generated during replication (Reyes-Lamothe et al, 2010). Furthermore, coordinated synthesis of the leading and lagging strands by two physically associated DNA polymerases is proposed to be mediated by the cyclical formation of a 'trombone' loop in the lagging strand, which allows template unwinding to continue while Okazaki fragments are synthesised in the opposite direction (Sinha et al, 1980). The presence of such loops has been confirmed, at least in bacterial systems, by electron microscopy and *in vitro* single molecule studies (Park et al, 1998; Hamdan et al, 2009).

Replisome structure in eukaryotic organisms is still poorly understood. Large multi-protein complexes termed replisome progression complexes (RPCs) have been isolated from *Saccharomyces cerevisiae*, comprising a number of key proteins which function in chromosomal DNA replication (Gambus et al, 2006). However, the precise architecture of these complexes is still largely unknown, and notably neither of the replicative DNA polymerases δ or ϵ have been isolated as components of RPCs. It is therefore still unclear how DNA polymerases are coordinated with other key activities at the eukaryotic replication fork, particularly in the context of coordinated leading and lagging strand synthesis (Johansson & Macneill, 2010). The global architecture of the archaeal replisome is similarly poorly understood, and has been the subject of few studies to date. Existing evidence however, suggests that interactions within the archaeal replisome are simplified relative to eukaryotes: the GINS complex in *S. solfataricus* for example acts as a direct bridge between the minichromosome maintenance (MCM) helicase on the leading strand and primase on the lagging strand (Marinsek et al, 2006), while in contrast eukaryotic primase appears to be tethered to the replisome through its interactions with Pol α , which in turn requires the eukaryote-specific Ctf4 protein for a weak interaction with GINS (Gambus et al, 2009; Kilkenny et al, 2012). Thus to better define the interactions that dictate the assembly and architecture



DNA strands are shown in black with ribonucleotides indicated by triangles. For simplicity, DNA sliding clamps and DNA polymerases are represented by single subunits. Adapted from (Barry & Bell, 2006; Gambus et al, 2009; Reyes-Lamothe et al, 2010).

OKAZAKI FRAGMENT MATURATION

Maturation in bacteria

Discontinuous Okazaki fragments must be processed into a single covalent lagging strand to maintain genome integrity. The coordinated action of several enzymes is therefore required to remove the RNA primers used to initiate each Okazaki fragment, synthesise DNA to replace these primers, and covalently ligate processed Okazaki fragments together. In bacteria, Pol I exchanges with the replicative Pol III on the lagging strand, synthesising DNA from the 3' end of terminated Okazaki fragments and utilising a unique 5'-3' exonuclease activity to simultaneously degrade downstream RNA primers by the process of nick translation (Kelly et al, 1970). RNaseHII, an enzyme capable of specifically cleaving ribonucleotides from RNA/DNA hybrid duplexes, may also provide a backup function to Pol I in removing Okazaki fragment primers (Ogawa & Okazaki, 1984; Funnell et al, 1986).

In *E. coli* there is evidence to support the operation of two distinct mechanisms facilitating the recycling of the lagging strand polymerase between successive Okazaki fragments. In the collision model, Pol III dissociates from template DNA upon encountering the 5' end of a downstream Okazaki fragment; this is mediated through the ability of Pol III to detect the elimination of template ssDNA (Georgescu et al, 2009). In contrast, the signalling model proposes that an uncharacterised signal derived from primer synthesis triggers dissociation of Pol III from the current Okazaki fragment, even if synthesis is incomplete (Li & Marians, 2000). Although there is still some debate over which is the dominant pathway *in vivo*, *in vitro* single molecule studies have demonstrated that both mechanisms may be equally operative, with redundancy ensuring efficient recycling (Hamdan et al, 2009). Importantly, both mechanisms are compatible with the role of Pol I in Okazaki fragment maturation; Pol I can either exchange with Pol III at a nick upon collision or, when Pol III has dissociated

prematurely, Pol I can complete Okazaki fragment synthesis prior to performing nick translation.

Following complete RNA primer removal by Pol I, lagging strand synthesis in bacteria is completed by the ligation of processed Okazaki fragments by an NAD-dependent DNA ligase (López de Saro & O'Donnell, 2001). Thus Okazaki fragment maturation in bacteria involves a relatively small complement of proteins. Eukaryotes and archaea, however, lack a homologue of Pol I, and thus have evolved distinct, more complex mechanisms to perform this crucial process.

The primary pathway of maturation in eukaryotes

In eukaryotes, chromosome replication is performed by three different DNA polymerases. Genetic studies in yeast have revealed a specific division of function between these polymerases; Pol ϵ performs the majority of DNA synthesis on the leading strand (Pursell et al, 2007), while Pol δ is responsible for lagging strand synthesis (Nick McElhinny et al, 2008; Miyabe et al, 2011). Pol α , in complex with a heterodimeric primase, synthesises RNA-DNA primers which initiate each Okazaki fragment on the lagging strand and are subsequently extended by Pol δ (Pavlov et al, 2006).

In vitro reconstitution using purified proteins from *S. cerevisiae* has shown that upon completion of Okazaki fragment synthesis, Pol δ is able to continue DNA synthesis while displacing the 5' end of the previously completed downstream Okazaki fragment into a flap structure, in a process termed strand displacement synthesis. This has the desirable effect of rendering the initiating RNA of Okazaki fragments suitable for cleavage. However, as Pol δ lacks an intrinsic 5'-3' nuclease like bacterial Pol I, additional factors are required to perform this activity.

The key protein responsible for this cleavage is Flap endonuclease 1 (Fen1). Fen1 is a structure-specific endonuclease capable of cleaving DNA or RNA at the base of a 5' flap; the

optimal substrate for Fen1 is actually a double flap structure, which additionally possesses an overlapping single nucleotide 3' flap (Kao et al, 2002). Structures of human and archaeal Fen1 in complex with DNA have revealed the structural rationale for this substrate specificity: a specific pocket on the surface of Fen1 accommodates the 3' flap and together with a 100° bend imposed on the DNA substrate, correctly positions the 5' flap in the active site (Figure 1.4). Induced duplex unwinding at the base of the 5' flap ensures that cleavage occurs at the correct position to generate a nick structure competent for subsequent ligation (Chapados et al, 2004; Tsutakawa et al, 2011).

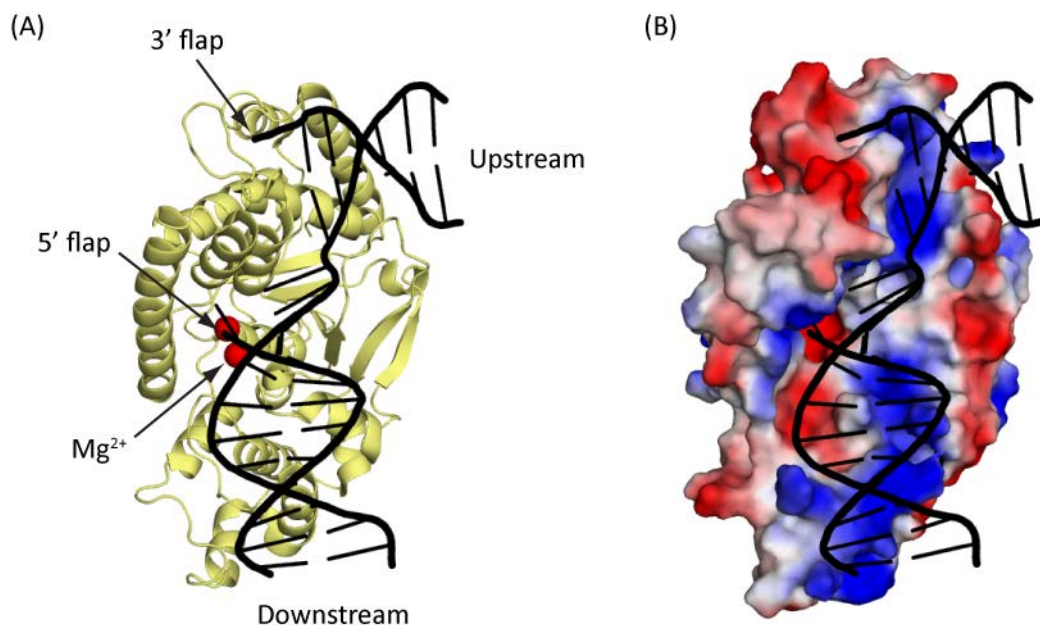


Figure 1.4 – The structure of human Fen1 bound to a double-flap DNA substrate

(A) Cartoon representation of human Fen1 (yellow) bound to DNA (black). The orientation of DNA in the context of a lagging strand template is annotated. The 5' flap is engaged in the active site of Fen1, indicated by the presence of two magnesium ions. Image was generated from PDB file 3Q8K.

(B) The contribution of electrostatic interactions to substrate positioning. The structure is viewed from the same orientation as (A), but Fen1 is displayed with electrostatic surface rendering; a path of basic residues which bind duplex DNA and force it to adopt a 100° bend is shown in blue.

Genetic studies on *S. cerevisiae* Fen1 (Rad27) support a role in Okazaki fragment maturation: although Rad27 null mutants are viable, they exhibit temperature-dependent

growth arrest, sensitivity to DNA damage and loss of genome stability, indicative of defects in DNA replication (Reagan et al, 1995; Sommers et al, 1995). Furthermore, the high frequency of duplications observed at repetitive sequences in this mutant has been proposed to arise from aberrant ligation of unprocessed flaps in the absence of Fen1-mediated cleavage (Tishkoff et al, 1997b). The synthetic lethality of Δ Rad27 with Δ Exo1 – a member of the same structure-specific nuclease superfamily – suggests that Exo1 may be capable of cleaving displaced flaps with low efficiency in the event of Rad27 dysfunction (Tishkoff et al, 1997a).

When *S. cerevisiae* Fen1 is combined with Pol δ in *in vitro* strand displacement assays, cleavage products are released from the downstream Okazaki fragment primer, confirming that Fen1 is indeed able to cleave flap structures as they are generated by Pol δ (Garg et al, 2004; Rossi & Bambara, 2006); in this way the two separate polypeptides combine to effect nick translation, similarly to bacterial Pol I. Presumably, 5' flap structures generated by Pol δ must undergo limited re-annealing prior to cleavage, to satisfy the substrate requirement of Fen1 for a 3' flap. Interestingly, biochemical studies on Pol δ suggest that a signalling mechanism is not employed to facilitate polymerase recycling during Okazaki fragment maturation in eukaryotes, unlike in bacteria (Langston & O'Donnell, 2008). This is consistent with a requirement for Pol δ to remain associated with nascent Okazaki fragments until downstream primers have been processed in cooperation with Fen1.

Okazaki fragment ligation

Cleavage of flap structures by Fen1 generates a nicked DNA substrate which is subsequently ligated by the ATP-dependent DNA ligase I (Lig1) (Ayyagari et al, 2003). Lig1 is specifically unable to ligate substrates containing RNA, which serves as a key mechanism in ensuring that RNA primers are fully removed and are not incorporated into chromosomal DNA. The structural basis of this property derives from the fact that Lig1 fully encircles its substrate during catalysis, and can discriminate between B-form dsDNA and A-form DNA/RNA duplex

(Pascal et al, 2004). However, to initiate Okazaki fragments, Pol α -primase actually synthesises approximately 30 nucleotides of DNA following a primer of approximately 10 ribonucleotides (Nethanel et al, 1988). Importantly, Pol α lacks the proofreading activity present in the other replicative polymerases and therefore this DNA must be removed in addition to the initiating RNA to maintain the fidelity of chromosomal DNA replication. Indeed, genetic interactions have demonstrated that Pol δ corrects errors produced by Pol α , consistent with strand displacement synthesis on the lagging strand being used to remove regions of both DNA and RNA synthesised by Pol α -primase (Pavlov et al, 2006). It is currently unclear how the appropriate degree of primer removal is achieved when removal of RNA alone is sufficient to render Okazaki fragments competent for ligation by Lig1. One possibility is that the efficient coupling of Pol δ and Fen1 activities during nick translation serves to limit Lig1 access to the substrate, thus favouring larger regions of Okazaki fragment repair. Consistent with this, replacement of DNA significantly beyond initiating RNA-DNA has been observed *in vitro*, even in the presence of Lig1 (Ayyagari et al, 2003). How excessive nick translation is limited under these conditions is unclear however.

One intriguing possibility has been proposed from the recently observed correlation between Okazaki fragment length and nucleosome spacing in *S. cerevisiae*. It has been suggested that the inability of Pol δ to perform strand displacement in the presence of newly-deposited histones triggers termination of nick translation at the midpoint of the nucleosome, facilitating subsequent Okazaki fragment ligation (Smith & Whitehouse, 2012). It should be noted however that Okazaki fragments were isolated in this study by inactivation of Lig1, and thus the balance of nick translation and ligase activities *in vivo* is inferred rather than directly observed. Furthermore, crenarchaeal Okazaki fragments have been shown to be similar in length to those in *S. cerevisiae*, despite lacking histone-like chromatin (Matsunaga et al, 2003), and thus the possibility remains that Okazaki fragment length is determined by properties intrinsic to the replication machinery.

The two-nuclease pathway

While the majority of cleavage products in *in vitro* Okazaki fragment maturation assays are short and consistent with a pathway involving Pol δ and Fen1 only, observations from a partially reconstituted system suggest some displaced flaps can escape immediate cleavage as they are produced by Pol δ , reaching sufficient length to become bound by the eukaryotic ssDNA binding protein, Replication Protein A (RPA) (Rossi & Bambara, 2006). Significantly, RPA has been shown to inhibit Fen1 activity *in vitro* (Bae et al, 2001). Evidence suggests another protein, Dna2, may play a key role in such circumstances. Dna2 is an essential protein in *S. cerevisiae*, and furthermore interacts both physically and genetically with Rad27 (Budd & Campbell, 1995; Budd & Campbell, 1997). Similarly to Fen1, Dna2 is a 5' flap structure-specific endonuclease, although in contrast to Fen1, Dna2 is stimulated by RPA and preferentially cleaves above the base of the flap (Bae & Seo, 2000; Stewart et al, 2010). Consistent with these complementary properties, Dna2 has been shown to be capable of acting in a sequential manner with Fen1 *in vitro*. Thus, RPA binds long flaps and stimulates their cleavage by Dna2. Shortening of the flap reduces the ability of RPA to bind, relieving the inhibition of Fen1 and allowing it to cleave to the base of the flap, generating a ligatable nick (Bae et al, 2001). Sequential handoff between these two nucleases is further enhanced by the apparent ability of Dna2 to actively displace RPA, and Fen1 to actively displace Dna2 (Stewart et al, 2008; Stewart et al, 2009). The inability of Dna2 to cleave RNA requires that flap cleavage by this enzyme occurs in the DNA downstream of initiating RNA (Bae & Seo, 2000), and thus this long-flap pathway has been proposed as another potential mechanism which favours removal of all nucleotides synthesised by Pol α -primase (Zheng & Shen, 2011).

While this two-nuclease pathway was originally proposed as the predominant mode of Okazaki fragment maturation, more recent data suggests it is most likely to operate as a backup to the pathway involving Pol δ and Fen1 only. The fact that Dna2 is essential in *S. cerevisiae* but Fen1 is not, has recently been shown to be due to interactions of Dna2 with

components of the DNA damage response, rather than representing the relative importance of these two factors in Okazaki fragment maturation (Budd et al, 2011). Analysis of cleavage products generated during *in vitro* Okazaki fragment maturation indicates that the vast majority of flaps do not reach sufficient length for RPA binding (Rossi & Bambara, 2006). Furthermore, while RPA is indeed strongly inhibitory to Fen1 cleavage of static flap substrates, when flaps are generated dynamically by the activity of Pol δ , RPA is far less inhibitory to Fen1 activity, suggesting that Fen1 is usually able to rapidly cleave emerging flaps before they become accessible to RPA (Rossi et al, 2008). Additionally, Pol δ has been shown to undergo 'idling' activity at the junction between two Okazaki fragments: repeated cycles of nucleotide addition and removal from the 3' end of the nascent Okazaki fragment (Garg et al, 2004). This activity has been proposed to minimise uncoupled strand displacement synthesis and the creation of large flaps by Pol δ , thus facilitating close coordination with Fen1 cleavage. Notably, the leading strand Pol ϵ is unable to perform this activity, consistent with its relevance to Okazaki fragment maturation (Garg et al, 2004).

Several additional activities have recently been demonstrated for Dna2, which further disputes a primary role in Okazaki fragment maturation. In combination with the Sgs1 helicase, Dna2 has been shown to be a key nuclease in the resection of dsDNA breaks during homologous recombination in *S. cerevisiae* (Zhu et al, 2008; Cejka et al, 2010; Niu et al, 2010). *Schizosaccharomyces pombe* Dna2 has been demonstrated to process structures that arise from regression of stalled replication forks, thus preventing their potentially mutagenic role in promoting aberrant homologous recombination (Hu et al, 2012). In human cells, it has been shown that a significant proportion of Dna2 localises to the mitochondria, and cooperates with other mitochondrial proteins in DNA synthesis and repair (Zheng et al, 2008; Duxin et al, 2009).

Additionally, there is some evidence to suggest that the relative importance of Fen1 and Dna2 nucleases in Okazaki fragment maturation may be different between higher eukaryotes and yeast. It has been reported that while human Dna2 is important for DNA

replication, its function is independent of Okazaki fragment maturation, in contrast to *S. cerevisiae* (Duxin et al, 2012). Furthermore, while a Rad27 deletion mutant exhibits defective growth but is viable, homozygous deletion of mouse Fen1 results in embryonic lethality (Kucherlapati et al, 2002).

Additional factors in Okazaki fragment maturation

Pif1, a 5'-3' helicase which is conserved in all eukaryotes and is implicated in diverse cellular roles such as mitochondrial DNA maintenance and telomere regulation (Bochman et al, 2010), has been identified as another potential factor in Okazaki fragment maturation. Deletion of *S. cerevisiae* Pif1 was found to suppress the lethality of Dna2 deletion, and this mutant was further rescued by a mutation in Pol δ which reduces its ability to perform strand displacement synthesis. Thus it was proposed that the Pif1 helicase cooperates with Pol δ to generate large displaced flaps which require Dna2, rather than Fen1, for resolution (Budd et al, 2006). This model is supported by *in vitro* reconstitution experiments (Rossi et al, 2008; Pike et al, 2009). However, it has been demonstrated that a key role of Pif1 is to facilitate the passage of replication forks by resolving template G-quadruplex structures (Paeschke et al, 2011) and persistent protein-DNA barriers such as those present at rRNA loci (Sabouri et al, 2012). Consistent with this, more recent biochemical studies suggest the role of Pif1 in Okazaki fragment maturation is more likely to involve displacement of Okazaki fragments that form secondary structures and impede the maturation machinery (Pike et al, 2010). Thus it is possible that lengthening of strand displaced flaps may simply represent a bystander activity of Pif1 that Dna2 has evolved to process in addition to its other diverse cellular functions.

Similarly to bacteria, type 2 RNaseH enzymes have also been implicated in RNA primer removal during Okazaki fragment maturation in eukaryotes. *In vitro* reconstitution of mammalian Okazaki fragment maturation originally proposed a role whereby RNaseH2 removes the majority of initiator RNA, cleaving one ribonucleotide 5' to the RNA-DNA junction,

with Pol δ and Fen1 subsequently removing the remaining ribonucleotide (Turchi et al, 1994). However, subsequent genetic analysis in *S. cerevisiae* demonstrated that RNaseH2 is not essential, and probably functions redundantly with Fen1 as a backup pathway (Qiu et al, 1999). Recent studies in both yeast and higher eukaryotes have demonstrated an essential role for RNaseH2 enzymes in removing single ribonucleotides which are erroneously incorporated into genomic DNA by replicative polymerases with surprising frequency (Nick McElhinny et al, 2010; Lazzaro et al, 2012; Reijns et al, 2012; Sparks et al, 2012). Thus, as with Dna2 and Pif1, it is unlikely that Okazaki fragment primer removal is the primary function of RNaseH2 *in vivo*, with the majority of Okazaki fragments being processed by the coordinated activities of Pol δ and Fen1.

Nonetheless, the involvement of additional factors such as Pif1, Dna2 and RNaseH2 in possible backup pathways of maturation demonstrates the importance of the maturation machinery being able to resolve a range of different structures *in vivo* (Figure 1.5). This is illustrated most simply by the fact that the GC content of Okazaki fragments has been shown to affect the efficiency of primer excision by Pol δ and Fen1 (Rossi & Bambara, 2006). Interestingly, both Fen1 and Dna2 can be acetylated and, at least *in vitro*, this has opposing effects on their activity, inhibiting Fen1 while stimulating Dna2. This raises the intriguing possibility that the balance between different pathways may be actively regulated, for example in response to specific DNA structures (Balakrishnan et al, 2010).

Archaeal Okazaki fragment maturation

Archaeal Okazaki fragments have been demonstrated to be similar in structure to those in eukaryotes (Matsunaga et al, 2003). However, the mechanism of Okazaki fragment maturation in archaea is still poorly characterised and to date has largely only been inferred on the basis of homology to eukaryotic proteins; archaea possess functional homologues of Pol δ , together with orthologues of Fen1 and Lig1, but a role for these proteins in Okazaki fragment

maturation has not been directly tested. Given the utility of archaea as a simplified model system for studying DNA replication, it will be important to determine the exact factor requirements for Okazaki fragment maturation in this domain of life.

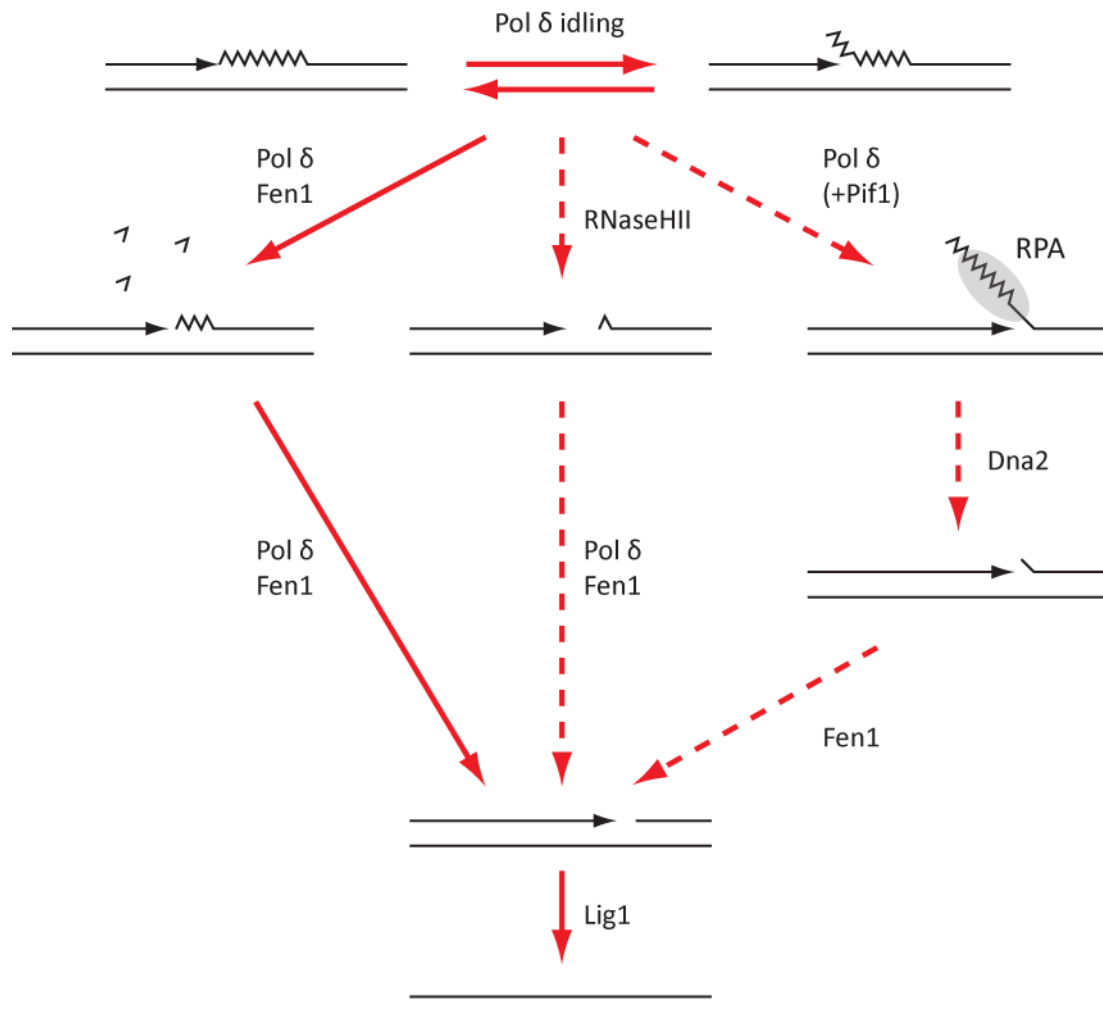


Figure 1.5 – Eukaryotic Okazaki fragment maturation pathways

DNA strands are shown schematically in black; triangles represent ribonucleotides. Minor or back-up pathways are indicated by dashed lines.

PROLIFERATING CELL NUCLEAR ANTIGEN (PCNA)**DNA sliding clamps**

Proliferating cell nuclear antigen (PCNA) was originally identified as a factor that demonstrated increased abundance in human cells undergoing DNA replication (Miyachi et al, 1978), and was subsequently shown by *in vitro* reconstitution studies to be an essential component of the eukaryotic DNA replication machinery (Prelich et al, 1987a). The essential role of PCNA was shown to be in binding Pol δ and dramatically enhancing its ability to perform processive DNA synthesis, as required during chromosomal replication (Tan et al, 1986; Prelich et al, 1987b). This stimulatory role for PCNA has subsequently been observed for both leading and lagging strand polymerases, ϵ and δ (Chilkova et al, 2007). A similarly essential role in facilitating processive synthesis was demonstrated for the β subunit of the *E. coli* replicative Pol III enzyme (Wickner, 1976). Strikingly, despite lacking any apparent primary sequence homology, PCNA, β , and polymerase processivity factors identified in bacteriophages all exhibit a very similar architecture; a ring-shaped oligomer with pseudo-hexagonal symmetry and a large central pore (Figure 1.6) (Kong et al, 1992; Krishna et al, 1994; Shamoo & Steitz, 1999; Moarefi et al, 2000). Notably given this conserved structure, the PCNA ring is a trimer, while β is dimeric. Biochemical and structural analysis has confirmed that these proteins are able to freely slide along dsDNA, which passes through their central pore (Stukenberg et al, 1991; Georgescu et al, 2008; McNally et al, 2010). These processivity factors thus function by tethering DNA polymerases to their template, enabling them to synthesise greater quantities of DNA without dissociating. As such, these proteins are often referred to as DNA sliding clamps.

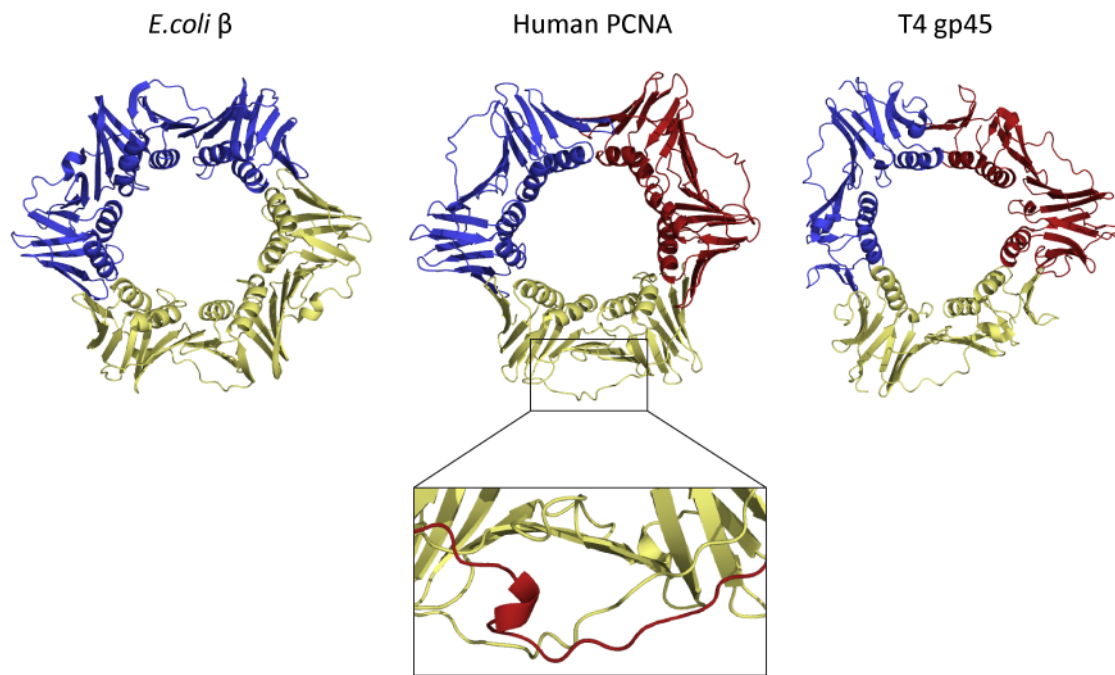


Figure 1.6 – DNA sliding clamps

Crystal structures of bacterial (*E. coli* β), eukaryotic (Human PCNA) and bacteriophage (T4 gp45) DNA sliding clamps. Although all subunits within each clamp are identical, each subunit is coloured differently for clarity. The expanded view of human PCNA shows the PIP motif from human p21 (red) bound to a human PCNA subunit (yellow). Images were generated from PDB files 2POL, 1AXC and 1CZD.

Consistent with their shared DNA replication machinery, archaea possess DNA sliding clamps orthologous to eukaryotic PCNA. All euryarchaea possess a single PCNA orthologue which assembles into a trimeric ring similarly to eukaryotic PCNA. Only a single exception has been identified, with *Thermococcus kodakaraensis* possessing two functional PCNA orthologues; this has likely arisen from a recent lateral gene transfer event (Ladner et al, 2011). In contrast, crenarchaea encode multiple PCNA orthologues, proposed to have evolved through duplications of a single PCNA gene ancestral to both archaea and eukaryotes (Chia et al, 2010). *S. solfataricus* possesses three distinct but related PCNA proteins which are all essential for viability and assemble into a heterotrimer *in vitro* and *in vivo* (Dionne et al, 2003; Zhang et al, 2010). This heterotrimer displays very similar architecture and dimensions to eukaryotic PCNA (Pascal et al, 2006; Williams et al, 2006; Hlinkova et al, 2008). The three

subunits of *S. solfataricus* PCNA assemble in a defined manner, with PCNA1 and PCNA2 forming a heterodimer which subsequently recruits PCNA3. Furthermore, this PCNA can only assemble through a single arrangement of the three subunits. This defined stereochemistry is imposed by differences in charge complementarity and hydrogen bonding patterns at each of the different subunit interfaces, together with a distinct curvature of each subunit (Doré et al, 2006; Pascal et al, 2006). Assembly of a heterotrimeric PCNA has also been observed in *Aeropyrum pernix*, although notably one of the PCNA orthologues in this organism has also been proposed to homotrimerise (Imamura et al, 2007).

Clamp loading

Given their closed ring-shaped architecture, DNA sliding clamps must be physically manipulated to facilitate their opening and encirclement of DNA. Furthermore, the positioning of sliding clamps on DNA must be specifically directed towards primer-template structures, on which DNA polymerases can initiate catalysis. This clamp loading function is performed by related pentameric AAA+ ATPase complexes in all three domains of life (Indiani & O'Donnell, 2006): bacteria require the $\tau_3\delta\delta'$ complex (which additionally associates with $\psi\chi$), while eukaryotes and archaea utilise the Replication Factor C (RFC) complex. Eukaryotic RFC is composed of a large subunit, Rfc1, and four small subunits, Rfc2–5. In contrast, RFC is simpler in almost all archaeal organisms, consisting of a large subunit, RFC_L, and four copies of a single small subunit, RFC_S (Barry & Bell, 2006).

Structural studies over recent years have begun to elucidate the elegant mechanism employed by these machines to perform their clamp loading function (Figure 1.7). The five subunits of the bacteriophage T4 clamp loader have been shown to adopt a right-handed spiral architecture in the presence of ATP, and this imposes a similar structure upon the T4 sliding clamp through numerous direct contacts, thus forcing out-of-plane clamp opening (Kelch et al, 2011). Consistent with a similar mechanism of opening for PCNA, disruption of a single

interface in an archaeal PCNA has been shown to be sufficient for loading onto DNA (Dionne et al, 2008). Furthermore, a low resolution structure of an archaeal RFC-PCNA-DNA complex, molecular dynamics simulations, and fluorescence resonance energy transfer (FRET) analysis all support PCNA opening through adoption of a right-handed spiral (Kazmirski et al, 2005; Miyata et al, 2005; Zhuang et al, 2006). Structural studies have revealed that individual subunits of both β and PCNA display greater curvature when assembled in a ring compared to their monomeric forms (Jeruzalmi et al, 2001; Hlinkova et al, 2008). It has therefore been proposed that closed sliding clamps are under 'spring tension', with clamp opening favoured upon disruption of a single interface by the clamp loader.

Specific positioning of an open clamp onto a primer-template structure is mediated by the DNA binding capacity of the clamp loader. Complementarity between the right-handed spiral of the clamp loader and the helical pitch of dsDNA facilitates the accommodation of dsDNA into a central cavity in the clamp loader. A protein collar at the top of this cavity ensures that only ssDNA can exit the clamp loader, thus generating specificity for primer-template regions where there is a transition from ds- to ssDNA (Bowman et al, 2004; Simonetta et al, 2009; Kelch et al, 2011). Notably, the extent of clamp opening observed in structural and biophysical studies is insufficient for the passage of dsDNA, and it is therefore likely that clamps are initially loaded onto ssDNA before the clamp loader slides them into position on dsDNA. Appropriate loading of the clamp on DNA has been proposed to facilitate the re-positioning of ATPase residues in the clamp loader, thus stimulating ATP hydrolysis, loss of spiral architecture, and dissociation from the loaded clamp and DNA (Kelch et al, 2011).

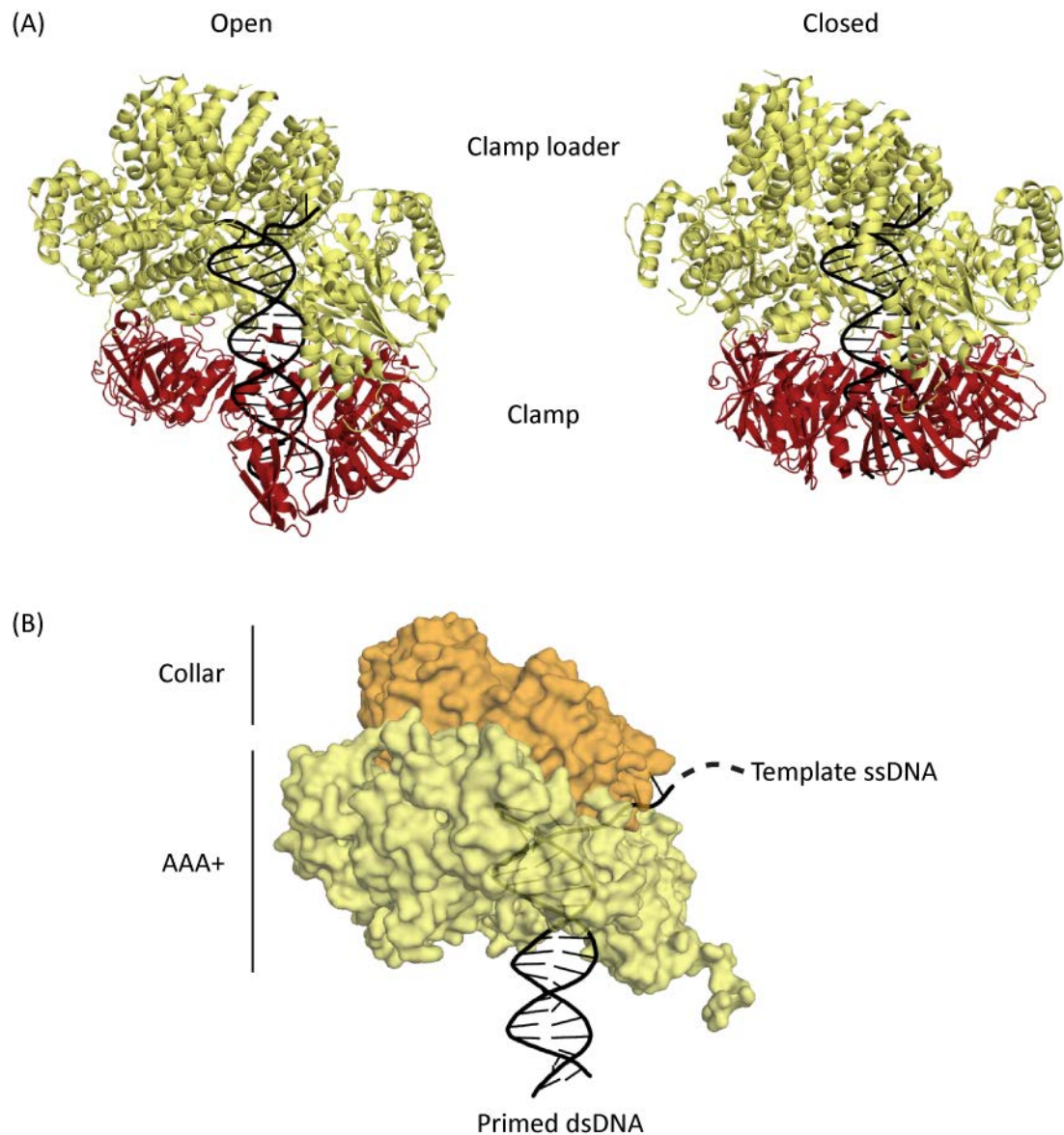


Figure 1.7 – The mechanism of DNA sliding clamp loading

(A) The structural basis of DNA sliding clamp opening. The bacteriophage T4 clamp loader (yellow) captured in complex with an open (left) or closed (right) gp45 DNA sliding clamp (red), highlighting the induced-spiral mechanism of opening. DNA is shown in black. The clamp loader complex is composed of four gp44 subunits and one distinct gp62 subunit, but for clarity all subunits are coloured identically. Images were generated from PDB files 3U60 (open) and 3U61 (closed).

(B) The structural basis of clamp loader substrate binding. The gp44/62 clamp loader complex is shown with surface rendering, with the different domains coloured as indicated. DNA, shown in black, is accommodated as a duplex in the central cavity formed between AAA+ domains, but is restricted to single-stranded form when exiting through the collar domains. The gp45 clamp is omitted for clarity. Image was generated using PDB file 3U60.

A structure of the β clamp bound to DNA has unexpectedly revealed the presence of specific interactions between aromatic residues on the surface of the ring and the single-stranded region of primer-template DNA (Georgescu et al, 2008). Biochemical analysis has shown these residues to be important for clamp loading, and thus it has been proposed that following placement on DNA by the clamp loader, specific clamp-DNA interactions serve to hold the clamp in position at the primer-template junction until a functional interaction with DNA polymerase can be established.

PCNA as a protein recruitment platform

It has become apparent in recent years that in addition to its classical role in supporting processive synthesis by DNA polymerases, PCNA also interacts with a vast range of other proteins involved in DNA replication, DNA repair, and cell cycle control (Moldovan et al, 2007). The majority of these proteins possess a short consensus sequence for interaction with PCNA, typically located at the N- or C-terminus, termed the PCNA-interacting protein (PIP) motif: Q-x-x-h-x-x-a-a, where 'h' represents a hydrophobic residue, and 'a' represents an aromatic residue. Only a small number of PIP-independent interactions with PCNA have been reported to date; the mouse Rev1 DNA polymerase for example contacts PCNA using a BRCT motif (Guo et al, 2006). Numerous structural studies have demonstrated a highly conserved mode of interaction between PIP-containing proteins and PCNA (Figure 1.6): the residues of the PIP motif form a 3_{10} helix, positioning hydrophobic and aromatic residues into a hydrophobic binding pocket located on the surface of PCNA, formed largely by the interdomain connector loop (IDCL) (Gulbis et al, 1996). Each subunit of the PCNA trimer possesses one such binding site. Proteins which bind the β -clamp also use a broadly similar interaction in which a shortened PIP-like motif adopts an extended peptide conformation and docks into a hydrophobic pocket on the surface of β formed by features analogous to the IDCL (Dalrymple et al, 2001; Bunting et al, 2003).

In eukaryotic organisms, covalent modification of PCNA further expands the repertoire of PCNA-interacting proteins by generating additional docking sites for certain factors (Hoegge et al, 2002). In response to DNA damage, PCNA can be enzymatically modified by ubiquitin on a conserved lysine residue, K164. This facilitates the recruitment of specialised translesion DNA polymerases to PCNA at stalled replication forks, via a specific ubiquitin binding module, in combination with a PIP motif. These error-prone polymerases allow sites of damage in template DNA to be bypassed (Stelter & Ulrich, 2003; Bienko et al, 2005). Alternatively, conjugation of poly-ubiquitin to K164 of PCNA serves to recruit the ZRANB3 recombinase, which facilitates error-free bypass of DNA damage through homologous recombination (Ciccio et al, 2012; Yuan et al, 2012). The ubiquitin-like protein, SUMO, can also be covalently attached to K164, which further contributes to genome stability through the recruitment of an antirecombinase, Srs2 (Pfander et al, 2005). Covalent modification of various factors by ubiquitin-like proteins has been observed in the euryarchaeon *Haloferax volcanii* (Humbard et al, 2010), while a complete eukaryotic-like ubiquitin modification system has recently been identified in the thaumarchaeon *Calditerrivibrio nitroreducens* (Nunoura et al, 2011). Thus there is potential for similar PCNA-modification systems to operate in archaea, although this has yet to be directly demonstrated.

Protein recruitment to PCNA is particularly pertinent to Okazaki fragment maturation, as the enzymes central to this process – Pol δ , Fen1 and Lig1 – have all been shown to interact with PCNA in eukaryotes (Li et al, 1995; Zhang et al, 1995; Levin et al, 1997). Photobleaching experiments in mammalian cells have demonstrated that while PCNA persists at DNA replication foci for extended periods of time, Fen1 and Lig1 exhibit more transient residence in these regions, invoking a model whereby PCNA acts as a platform for the cyclical recruitment of Okazaki fragment processing enzymes (Sporbert et al, 2005). This model is further supported by disruption of the PCNA-Fen1 interaction in mice, which results in loss of localization of Fen1 from replication foci, deficient Okazaki fragment maturation and

ultimately, embryonic lethality (Zheng et al, 2007). Similarly, disruption of the interaction between Lig1 and PCNA in human cells causes a loss of Lig1 replication foci localisation and defective Okazaki fragment maturation (Montecucco et al, 1998; Levin et al, 2000). A functional advantage of co-ordinating lagging strand enzymes in this way seems to be indicated by the fact that the β -clamp also interacts with the key enzymes involved in bacterial Okazaki fragment maturation – Pol I and DNA ligase (López de Saro & O'Donnell, 2001). Because there is no evidence to support Dna2 or Pif1 binding to PCNA, it is tempting to speculate that this is further evidence for a backup role only for these proteins in Okazaki fragment maturation.

An interaction with PCNA has been shown to significantly stimulate the catalytic activity of many enzymes *in vitro*, including DNA polymerases, Fen1 and Lig1 (Tom et al, 2000; Tom et al, 2001; Dionne et al, 2003). These analyses suggest enzyme stimulation by PCNA is generally mediated by an enhancement of the affinity of partner proteins for their DNA substrates, consistent with the topological interaction between PCNA and DNA. Interestingly however, kinetic analysis of the 3' flap endonuclease XPF from *S. solfataricus* has demonstrated that PCNA stimulates this enzyme by increasing its catalytic rate rather than altering its substrate affinity (Hutton et al, 2008). This has subsequently been shown to derive from the ability of PCNA to act as a fulcrum upon which XPF can exert force on its DNA substrate, distorting it into a suitable conformation for cleavage (Hutton et al, 2010).

The toolbelt model

Consistent with the presence of multiple protein binding sites within the oligomeric DNA sliding clamps, it has been proposed that multiple protein factors can simultaneously bind to a single sliding clamp ring; the clamp thus acting as a 'toolbelt' to coordinate multiple enzymatic activities (Indiani et al, 2005). Interestingly, each of the distinct subunits of *S. solfataricus* heterotrimeric PCNA possesses distinct specificity for binding different partner proteins,

including putative Okazaki fragment maturation factors. Thus PCNA1 specifically interacts with Fen1, XPF, the translesion DNA polymerase Dpo4, and the Holliday junction endonuclease Hjc; PCNA2 interacts with DNA polymerase B1 (PolB1); and PCNA3 interacts with Lig1, XPF and the uracil DNA glycosylase UDG1 (Dionne et al, 2003; Roberts et al, 2003; Dionne & Bell, 2005; Dorazi et al, 2006; Dionne et al, 2008). This specificity arises from subtle structural differences between the hydrophobic binding pockets of the different PCNA subunits, enabling them to discriminate between different partner protein PIP motifs, which possess distinct variations of the consensus sequence (Figure 1.8) (Pascal et al, 2006).

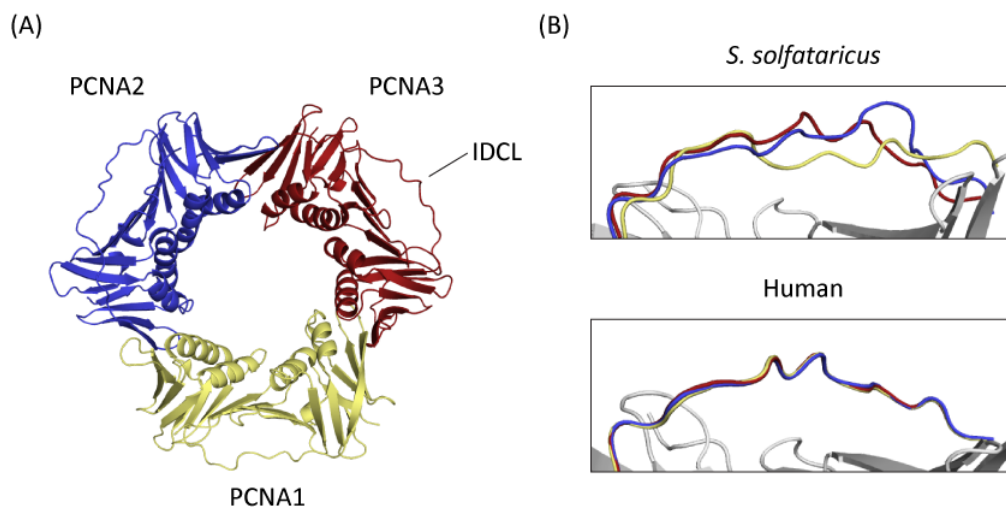


Figure 1.8 – The structural basis of partner protein binding by *S. solfataricus* PCNA

(A) The structure of *S. solfataricus* PCNA. The IDCL, which forms the protein binding pocket, is indicated. Image was generated from PDB file 2HII.

(B) Comparison of IDCL configuration between different PCNA subunits. Overlay of the three subunits of *S. solfataricus* PCNA, coloured as in (A), highlights significant variation in IDCL structure (top), conferring the ability to discriminate between distinct PIP motifs. The three subunits within a single human PCNA ring however display almost invariant IDCL structure (bottom). Images were generated from PDB files 2HII and 1AXC.

It has therefore been hypothesised that the toolbelt model may be particularly relevant to Okazaki fragment maturation in *S. solfataricus* given that the proposed maturation factors PolB1, Fen1 and Lig1 each specifically interact with a different subunit of the

heterotrimeric ring. Furthermore, these three factors have been shown to be bridged by PCNA *in vitro*, supporting their simultaneous binding to a single PCNA ring (Dionne et al, 2003). This crenarchaeal model system may therefore provide a unique opportunity to test whether a toolbelt architecture is important for the coordination of multiple enzyme activities by a DNA sliding clamp.

Clamp unloading

Early *in vitro* experiments using *E. coli* proteins suggested that Pol III dissociates from β upon the completion of each Okazaki fragment, associating with a new β molecule recruited from solution at a primed site upstream, and thus leaving β to accumulate on the newly replicated lagging strand (Stukenberg et al, 1994; Yuzhakov et al, 1996). It was noted however, that the predicted number of Okazaki fragments synthesised in a single round of replication significantly exceeds the estimated number of β molecules within a cell (Leu et al, 2000). Active unloading and recycling of β is therefore presumably essential, particularly given that the dissociation of β from dsDNA is relatively slow, with a half-life on DNA of approximately one hour (Yao et al, 1996). Biochemical analysis has demonstrated that the Pol III holoenzyme is capable of unloading as well as loading clamps; furthermore, the δ subunit of the clamp loader in isolation may be sufficient to perform clamp unloading (Leu et al, 2000).

Recent characterisation of *E. coli* replisome stoichiometry *in vivo* has demonstrated that there are approximately 3 β clamps associated with each replisome, suggesting that clamps are recycled, or at least unloaded, soon after the completion of Okazaki fragment synthesis (Reyes-Lamothe et al, 2010). This hypothesis is supported by *in vitro* measurements demonstrating that a single *E. coli* replisome can synthesise multiple Okazaki fragments in the absence of additional β clamps from solution, thus requiring recycling of every clamp upon Okazaki fragment completion (Tanner et al, 2011). *In vivo* measurements in the Gram-positive bacterium *Bacillus subtilis* have demonstrated a contrasting situation, with β clamps observed

to accumulate behind active replication forks (Su'etsugu & Errington, 2011). The reason for this difference between species is currently unclear, but one proposal is that accumulated clamps may be specifically required in *B. subtilis* to facilitate recruitment of factors for transactions on newly synthesised DNA (Su'etsugu & Errington, 2011).

The fate of PCNA following Okazaki fragment maturation in eukaryotes is less clear. Photobleaching studies in mammalian cells have demonstrated that PCNA turnover at replication foci is considerably slower than the time required to synthesise an Okazaki fragment, and therefore some PCNA recycling must occur within these foci (Sporbert et al, 2002). It has been calculated that in HeLa cells, as in *E. coli*, the estimated number of PCNA molecules is less than the predicted number of Okazaki fragments produced in a single round of DNA replication (Morris & Mathews, 1989). Together with the fact that PCNA has a relatively long half-life on DNA of approximately 30 minutes (Yao et al, 1996), it is clear that active PCNA unloading and recycling is required for lagging strand synthesis. The efficiency of PCNA recycling is currently unclear however; the resolution of photobleaching experiments was insufficient to distinguish between unloading of PCNA immediately upon completion of Okazaki fragment maturation, and recycling from a pool of accumulated PCNA molecules on completed Okazaki fragments (Sporbert et al, 2002).

It is currently unknown which proteins are responsible for the critical unloading of PCNA *in vivo*. Several alternative RFC complexes have been described, in which different proteins replace Rfc1 of the clamp loader and form a complex with Rfc2–5. Interestingly, both the Ctf18-RFC complex – involved in sister chromatid cohesion – and the Rad17-RFC complex – implicated in the loading of the Rad9-Rad1-Hus1 (9-1-1) checkpoint clamp onto DNA – have been shown to be capable of efficiently unloading PCNA *in vitro* (Bylund & Burgers, 2005; Yao et al, 2006). PCNA unloading by unmodified *S. cerevisiae* RFC has also been observed *in vitro*, although this appears to be inefficient unless the large N-terminal extension of Rfc1 – the *in vivo* function of which is unclear – is removed (Yao et al, 2006). No alternative clamp loaders

have been detected in archaea to date, and furthermore the RFC_L subunit of archaeal RFC lacks the inhibitory N-terminal extension present in eukaryotic Rfc1. Thus it is possible that a single RFC complex is responsible for both loading and unloading of PCNA in archaea.

THESIS AIMS

This thesis investigates a number of key questions regarding the molecular mechanisms of DNA replication in the model archaeon *S. solfataricus*. The factor requirements of Okazaki fragment maturation in eukaryotic organisms have been well studied but have only been inferred in archaea on the basis of protein homology and the activity of individual proteins *in vitro*. The mechanisms and protein requirements for complete maturation of archaeal Okazaki fragments are investigated in Chapter 3 by fully reconstituting this essential multi-step process *in vitro*.

Furthermore, while the DNA sliding clamp PCNA has been shown to interact with and stimulate a number of Okazaki maturation factors in both eukaryotes and *S. solfataricus*, the precise mechanistic role of this protein in Okazaki fragment maturation has not been determined. Chapter 4 utilises the reconstituted Okazaki fragment maturation system to test the toolbelt model of multiple enzyme coordination around a single PCNA ring.

Numerous structural studies have suggested that protein flexibility upon PCNA may be important for the physical accommodation of multiple enzymes around a single PCNA during toolbelt coordination. However, a functional role for such protein flexibility has yet to be directly tested. Chapter 5 describes the development of an assay to measure Lig1 conformational changes in the context of its coordination with other Okazaki fragment maturation factors around PCNA.

In contrast to bacteria, the precise architecture of the replisome in eukaryotes and archaea is still relatively poorly understood. In particular, the molecular basis of the

coordination of replicative DNA polymerases with other components of the replisome has not been determined. Chapter 6 investigates the protein-protein interactions of the *S. solfataricus* replicative DNA polymerase, PolB1, to determine its quaternary composition and molecular relationship to the core of the archaeal replisome.

Chapter 2 – Materials and Methods

STRAINS

All cloning was performed using *Escherichia coli* XL-10 Gold. All protein expression was performed using *E. coli* Rosetta. Other strains were used as indicated.

Table 2.1 – Strains used in this study

Strain	Genotype
Bacteria	
<i>Escherichia coli</i> XL-10 Gold (Stratagene)	Tet ^R $\Delta(mcrA)183 \Delta(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacI^qZ\Delta M15 Tn10(Tet^R)Amy Cm^R]$
<i>Escherichia coli</i> Rosetta (Novagen)	F ⁻ <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm</i> (DE3) pLysSRARE (Cam ^R)
Yeast	
<i>Saccharomyces cerevisiae</i> AH109 (Clontech)	MATa <i>trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1_{UAS}-GAL1_{TATA}-HIS3 GAL2_{UAS}-GAL2_{TATA}-ADE2 URA3::MEL1_{UAS}-MEL1_{TATA}-lacZ</i>
<i>Saccharomyces cerevisiae</i> Y187 (Clontech)	MAT α <i>ura3-52 his3-200 ade2-101 trp1-901 leu2-3,112 gal4Δ met⁻ gal80Δ URA3::GAL1_{UAS}-GAL1_{TATA}-lacZ</i>
Archaea	
<i>Sulfolobus solfataricus</i> P2 (DSMZ)	Wild-type (DSM1617)
<i>Sulfolobus acidocaldarius</i> SacTK (Gristwood et al, 2012)	<i>pyrE ΔamyA::Pae2453-FLAG</i>
<i>Sulfolobus acidocaldarius</i> SacPolB1 (T. Gristwood, unpublished)	<i>pyrE ΔamyA::Pae2453-FLAG Saci_1537-c-myc-6his</i>

CLONING

Protein expression constructs

All oligonucleotides were supplied by Eurogentec (see Appendix I for sequences). Site directed mutagenesis was performed using Quikchange (Stratagene) with the exception that polymerase chain reaction (PCR) amplification employed a single primer 5' phosphorylated by T4 polynucleotide kinase (NEB) and was supplemented with 0.5 mM ATP and approximately 5 pmol *S. solfataricus* Lig1 to yield closed circular products.

Table 2.2 – Expression constructs used in this work

Construct	Description	Source
pET33b-PCNA1	PCNA1 with C-terminal 6xHis tag	(Dionne et al, 2003)
pET30a-PCNA2	PCNA2 with C-terminal 6xHis tag	(Dionne et al, 2003)
pET30a-PCNA3	PCNA3 with C-terminal 6xHis tag	(Dionne et al, 2003)
pET33b-PCNA1-2-3	PCNA1-2-3 with C-terminal 6xHis tag	(Dionne et al, 2008)
pET33b-PCNA Δ 1-2-3	PCNA1-2-3 with PCNA1 A246E mutation	This work
pET33b-PCNA1- Δ 2-3	PCNA1-2-3 with PCNA2 A242E mutation	This work
pET33b-PCNA1-2- Δ 3	PCNA1-2-3 with PCNA3 A241E mutation	This work
pET33b-PolB1	PolB1 with C-terminal 6xHis tag	(Dionne et al, 2003)
pET33b-PolB1-notag	PolB1 untagged	This work
pET33b-PolB1-PIP	PolB1 with LF7AA mutation	This work
pET33b-Fen1	Fen1 untagged	(Dionne et al, 2003)
pET33b-Fen1-PIP	Fen1 with WF350AA mutation	This work
pET30a-Lig1	Lig1 with C-terminal 6xHis tag	(Dionne et al, 2003)
pET30a-Lig1-PIP	Lig1 with FL110AA mutation	This work
pET30a-Lig1-Cys	Lig1 with C483S, C544S, C545S mutations	This work
pET30a-Lig1-I464C	Lig1-Cys with I464C mutation	This work
pET30a-Lig1-K463C	Lig1-Cys with K463C mutation	This work

pET30a-Lig1-V507C	Lig1-Cys with V507C mutation	This work
pET30a-Lig1-K515C	Lig1-Cys with K515C mutation	This work
pET30a-Lig1-M516C	Lig1-Cys with M516C mutation	This work
pET30a-Lig1-E2C	Lig1-Cys with E2C mutation	This work
pET30a-Lig1-E2C.M516C	Lig1-Cys with E2C and M516C mutations	This work
pET30a-PBP1	PBP1 untagged	This work
pET30a-PBP2	PBP2 with C-terminal 6xHis tag	This work
pET33b-RNaseHII	RNaseHII untagged	This work
pET22b-Dpo4	Dpo4 untagged	(Boudsocq et al, 2001)
pET19b-SSB	SSB untagged	(Wadsworth & White, 2001)

PCNA Δ 1-2-3, PCNA1- Δ 2-3, and PCNA1-2- Δ 3 constructs were generated by introducing mutations into pET33b-PCNA1-2-3 using primers Δ P1, Δ P2 and Δ P3 respectively.

A stop codon was introduced into pET33b-PolB1 immediately following the PolB1 open reading frame, using primer PolB1_stop, to generate untagged PolB1. A mutation in the PIP motif of PolB1 ($_7$ LF $_8$ to AA) was introduced into pET33b-PolB1-notag with primer PolB1_PIP.

A mutation in the PIP motif of Fen1 ($_{350}$ WF $_{351}$ to AA) was introduced into pET33b-Fen1 with primer Fen1_PIP.

A mutation in the PIP motif of Lig1 ($_{110}$ FL $_{111}$ to AA) was introduced into pET30a-Lig1 using primer Lig1_PIP. Lig1-Cys was generated by sequential mutagenesis of pET30a-Lig1 by primers C483S and CC544SS. Single cysteine residues were introduced into the pET30a-Lig1-Cys using primers I464C, K463C, V507C, K515C or M516C. A cysteine at position 2 was introduced into pET30a-Lig1-Cys or pET30a-Lig1-M516C by PCR amplification of the relevant construct using primers E2C_for and T7R before ligation into NdeI and XhoI sites of pET30a (Novagen).

pET30a-PBP1 was generated by PCR amplification of Sso0150 from *S. solfataricus* P2 genomic DNA using primers Sso0150_for and Sso0150_rev before ligation into NdeI and XhoI sites of pET30a.

pET30a-PBP2 was generated by PCR amplification of Sso6202 from *S. solfataricus* P2 genomic DNA using primers Sso6202_for and Sso6202_rev before ligation into NdeI and XhoI sites of pET30a.

pET33b-RNaseHII was generated by PCR amplification of Sso2384 from *S. solfataricus* P2 genomic DNA using primers RNase_for and RNase_rev before ligation into NcoI and XhoI sites of pET33b (Novagen).

GST-fusion constructs

Table 2.3 – GST-fusion constructs used in this work

Construct	Description	Source
pGEX-PCNA1	GST-PCNA1	(Dionne et al, 2003)
pGEX-PCNA2	GST-PCNA2	(Dionne et al, 2003)
pGEX-PCNA3	GST-PCNA3	(Dionne et al, 2003)
pET42a-RNaseHII	GST-RNaseHII	This work
pET42a-PBP1	GST-PBP1	This work
pET42a-PBP2	GST-PBP2	This work

pET42a-RNaseHII was generated by PCR amplification of Sso2384 from *S. solfataricus* P2 genomic DNA using primers RNase_for and RNase_rev before ligation into NcoI and XhoI sites of pET42a (Novagen). pET42a-PBP1 was generated by PCR amplification of Sso0150 from *S. solfataricus* P2 genomic DNA using primers Sso0150_for2 and Sso0150_rev before ligation into NcoI and XhoI sites of pET42a. pET42a-PBP2 was generated by PCR amplification of Sso6202

from *S. solfataricus* P2 genomic DNA using primers Sso6202_for2 and Sso6202_rev2 before ligation into EcoRI and XhoI sites of pET42a.

Yeast two-hybrid constructs

Table 2.4 – Yeast two-hybrid constructs used in this work

Construct	Description	Source
pGAD-T7-PCNA1	GAL4 activation domain (AD)-PCNA1	(Dionne et al, 2008)
pGBK-T7-PCNA1	GAL4 DNA binding domain (BD)-PCNA1	(Dionne et al, 2003)
pGAD-T7-PCNA2	AD-PCNA2	(Dionne et al, 2008)
pGBK-T7-PCNA2	BD-PCNA2	(Dionne et al, 2003)
pGAD-T7-PCNA3	AD-PCNA3	(Dionne et al, 2008)
pGBK-T7-PCNA3	BD-PCNA3	(Dionne et al, 2003)
pGAD-T7-RNaseHII	AD-RNaseHII	This work
pGBK-T7-RNaseHII	BD-RNaseHII	This work
pGAD-T7-PolB1	AD-PolB1	This work
pGBK-T7-PolB1	BD-PolB1	This work
pGAD-T7-PBP1	AD-PBP1	This work
pGBK-T7-PBP1	BD-PBP1	This work
pGAD-T7-PBP2	AD-PBP2	This work
pGBK-T7-PBP2	BD-PBP2	This work
pGAD-T7-GINS15	AD-GINS15	(Marinsek et al, 2006)
pGBK-T7-GINS15	BD-GINS15	(Marinsek et al, 2006)
pGAD-T7-GINS23	AD-GINS23	(Marinsek et al, 2006)
pGBK-T7-GINS23	BD-GINS23	(Marinsek et al, 2006)
pGAD-T7-RecJ _{dbh}	AD- RecJ _{dbh}	(Marinsek et al, 2006)
pGBK-T7- RecJ _{dbh}	BD- RecJ _{dbh}	(Marinsek et al, 2006)

RNaseHII constructs were generated by PCR amplification of Sso2384 from *S. solfataricus* P2 genomic DNA using primers RNase_for and RNase_rev before ligation into NcoI and XhoI sites of pGAD-T7 or NcoI and SalI sites of pGBK-T7 (Clontech). PolB1 constructs were generated by PCR amplification of Sso0552 from *S. solfataricus* P2 genomic DNA using primers PolB1_for and PolB1_rev before ligation into NcoI and XhoI sites of pGAD-T7 or NcoI and SalI sites of pGBK-T7. PBP1 constructs were generated by PCR amplification of Sso0150 from *S. solfataricus* P2 genomic DNA using primers Sso0150_for and Sso0150_rev before ligation into NdeI and XhoI sites of pGAD-T7 or NdeI and SalI sites of pGBK-T7. PBP2 constructs were generated by PCR amplification of Sso6202 from *S. solfataricus* P2 genomic DNA using primers Sso6202_for and Sso6202_rev before ligation into NdeI and XhoI sites of pGAD-T7 or NdeI and SalI sites of pGBK-T7.

PROTEIN EXPRESSION AND PURIFICATION

Protein expression

Transformed *E. coli* Rosetta cells were grown to an OD₆₀₀ of 0.4–0.6 before induction with 1 mM IPTG for 3 hours at 37°C. Cells were harvested by centrifugation at 4000 rpm for 15 minutes. To test for protein expression and solubility, a 50 ml culture was treated as above. 1 ml samples were withdrawn before and after induction, centrifuged for 1 minute at 13000 rpm, and the pellet was resuspended to an OD₆₀₀ of 6 in SDS-PAGE loading buffer. The remainder of the culture was centrifuged and the pellet resuspended in 5 ml 20 mM Tris pH 8, 300 mM NaCl and lysed by sonication (6 x 10 seconds). 1 ml of lysate was withdrawn and centrifuged for 10 minutes at 13000 rpm. The supernatant was analysed as the soluble fraction and the pellet was resuspended in 1 ml SDS-PAGE loading buffer and analysed as the insoluble fraction. All samples were analysed by SDS-PAGE.

Purification of Lig1

All mutants of Lig1 were purified identically. Cells were lysed by grinding in liquid nitrogen and resuspended in 20 mM Tris pH 8, 300 mM NaCl and complete EDTA-free protease inhibitors (Roche). Extract was then incubated at 75°C for 25 minutes before centrifugation at 17000 rpm for 30 minutes to remove cell debris and heat-denatured protein. The supernatant was applied to 2 ml Ni-NTA agarose resin (Qiagen), pre-equilibrated in 20 mM Tris pH 8, 300 mM NaCl. Bound protein was washed with 15 mM imidazole and eluted with 500 mM imidazole. Eluted fractions were loaded onto a HiLoad Superdex 200 26/60 gel filtration column (GE Healthcare) pre-equilibrated in 10 mM Tris pH 8, 150 mM NaCl, 14 mM β -mercaptoethanol. Protein-containing fractions were diluted two-fold in 10 mM Tris pH 8, 14 mM β -mercaptoethanol, loaded onto a 1 ml MonoQ column (GE Healthcare) pre-equilibrated in 10 mM Tris pH 8, 75 mM NaCl, 14 mM β -mercaptoethanol, and eluted over a 15 ml linear gradient of 75–1000 mM NaCl. Protein-containing fractions were pooled, snap frozen in liquid nitrogen and stored at -80°C. Protein concentration was calculated from the absorbance at 280 nm using an extinction coefficient of $63260 \text{ M}^{-1}\text{cm}^{-1}$.

Purification of PCNA

All mutants of the PCNA1-2-3 fusion were purified identically. Cells were resuspended in 20 mM Tris pH 8, 300 mM NaCl and protease inhibitors, and lysed by passage through a French press at 20000 psi before centrifugation at 17000 rpm for 30 minutes to remove cell debris. The supernatant was incubated at 75°C for 25 minutes before further centrifugation at 17000 rpm for 30 minutes to remove heat-denatured protein. The supernatant was applied to 2 ml Ni-NTA agarose resin, pre-equilibrated in 20 mM Tris pH 8, 300 mM NaCl. Bound protein was washed with 15 mM imidazole and eluted with 500 mM imidazole. Eluted fractions were loaded onto a HiLoad Superdex 200 26/60 gel-filtration column pre-equilibrated in 20 mM Tris pH 8, 300 mM NaCl. Protein-containing fractions were pooled, snap frozen in liquid nitrogen

and stored at -80°C. Protein concentration was calculated from the absorbance at 280 nm using an extinction coefficient of 35300 M⁻¹cm⁻¹.

To prepare non-covalent PCNA heterotrimer, cell pellets from individual PCNA1, PCNA2 and PCNA3 overexpression were combined, and purification was performed exactly as described for PCNA1-2-3, with the exception that cell lysate was incubated at 85°C for 25 minutes. Protein concentration was calculated from the absorbance at 280 nm using an extinction coefficient of 35300 M⁻¹cm⁻¹.

Purification of Fen1

Fen1 and the Fen1 PIP mutant were purified identically. Cells were resuspended in 20 mM MES pH 6, 100 mM NaCl, 1 mM EDTA, 1 mM DTT and protease inhibitors, and lysed by passage through a French press at 20000 psi before centrifugation at 17000 rpm for 30 minutes. The supernatant was incubated at 75°C for 15 minutes before further centrifugation at 17000 rpm for 30 minutes. The supernatant was purified by passage through a 5 ml HiTrap Heparin column (GE Healthcare) and a HiLoad Superdex 75 26/60 gel filtration column (GE healthcare) exactly as described previously (Hutton et al, 2008). Protein-containing fractions were pooled, snap frozen in liquid nitrogen and stored at -80°C. Protein concentration was calculated from the absorbance at 280 nm using an extinction coefficient of 28880 M⁻¹cm⁻¹ for wild-type (WT) Fen1 and 23380 M⁻¹cm⁻¹ for the PIP mutant.

Purification of PolB1

PolB1 and the PolB1 PIP mutant were purified identically. Cells were resuspended in buffer A (10 mM HEPES pH 7.5, 100 mM NaCl, 1 mM DTT) with protease inhibitors, lysed by passage through a French press at 20000 psi and centrifuged at 17000 rpm for 30 minutes. The supernatant was incubated at 65°C for 20 minutes before further centrifugation at 17000 rpm for 30 minutes. The supernatant was loaded onto a 5 ml HiTrap Heparin column pre-

equilibrated in buffer A, and eluted over a 75 ml linear gradient of 100–1000 mM NaCl. Protein-containing fractions were loaded onto a HiLoad Superdex 200 26/60 column pre-equilibrated in 10 mM HEPES pH 7.5, 500 mM NaCl, 1 mM DTT. Eluted PolB1 was diluted five-fold in 10 mM HEPES pH 7.5, 1 mM DTT and loaded onto a 1 ml MonoS column (GE healthcare) pre-equilibrated in buffer A and eluted over a 15 ml linear gradient of 100–1000 mM NaCl. Protein-containing fractions were pooled, snap frozen in liquid nitrogen and stored at -80°C. Protein concentration was calculated from the absorbance at 280 nm using an extinction coefficient of 118165 M⁻¹cm⁻¹ for WT PolB1 and 118040 M⁻¹cm⁻¹ for the PIP mutant.

Purification of Dpo4

Cells were resuspended in buffer A (10 mM Tris pH 8, 150 mM NaCl, 1 mM DTT) with protease inhibitors, lysed by passage through a French press at 20000 psi and centrifuged at 17000 rpm for 30 minutes. The supernatant was incubated at 75°C for 25 minutes before further centrifugation at 17000 rpm for 30 minutes. The supernatant was loaded onto a 5 ml HiTrap Heparin column pre-equilibrated in buffer A and eluted over a 75 ml linear gradient of 150–1000 mM NaCl. Protein-containing fractions were loaded onto a HiLoad Superdex 200 26/60 column pre-equilibrated in 10 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT. Eluted Dpo4 was snap frozen in liquid nitrogen and stored at -80°C. Protein concentration was calculated from the absorbance at 280 nm using an extinction coefficient of 22350 M⁻¹cm⁻¹.

Purification of PBP1

Cells were resuspended in 20 mM Tris pH 8, 300 mM NaCl, 1 mM DTT and protease inhibitors, lysed by passage through a French press at 20000 psi and centrifuged at 17000 rpm for 30 minutes. The supernatant was incubated at 75 °C for 25 minutes before further centrifugation at 17000 rpm for 30 minutes. The supernatant was diluted two-fold in 10 mM Tris pH 8, 1 mM DTT, applied to a 5 ml HiTrap Q column (GE Healthcare) equilibrated in 10 mM Tris pH 8, 150

mM NaCl, and eluted over a 75 ml linear gradient of 150–1000 mM NaCl. Protein-containing fractions were concentrated to 5 ml and applied to a HiLoad Superdex 75 26/60 column equilibrated in 20 mM Tris pH 8, 300 mM NaCl, 1 mM DTT. Eluted PBP1 was snap frozen in liquid nitrogen and stored at -80°C. Protein concentration was calculated from the absorbance at 280 nm using an extinction coefficient of 25440 M⁻¹cm⁻¹.

Purification of PBP2

Cells were resuspended in 20 mM Tris pH 8, 300 mM NaCl and protease inhibitors, lysed by passage through a French press at 20000 psi and centrifuged at 17000 rpm for 30 minutes. The supernatant was incubated at 75 °C for 25 minutes before further centrifugation at 17000 rpm for 30 minutes. The supernatant was applied to 2 ml Ni-NTA agarose resin, pre-equilibrated in 20 mM Tris pH 8, 300 mM NaCl. Bound protein was washed with 15 mM imidazole and eluted with 500 mM imidazole. Protein-containing fractions were applied to a HiLoad Superdex 75 26/60 column equilibrated in 20 mM Tris pH 8, 300 mM NaCl, 1 mM DTT. Eluted PBP2 was snap frozen in liquid nitrogen and stored at -80°C. Protein concentration was calculated from the absorbance at 280 nm using an extinction coefficient of 4470 M⁻¹cm⁻¹.

Purification of SSB

SSB was purified exactly as described previously (Wadsworth & White, 2001). Protein concentration was calculated from the absorbance at 280 nm using an extinction coefficient of 12490 M⁻¹cm⁻¹.

Purification of GST-fusion proteins

All GST-fusion proteins were purified identically. Cells were resuspended in 20 mM Tris pH 8, 300 mM NaCl and protease inhibitors, lysed by passage through a French press at 20000 psi and centrifuged at 17000 rpm for 30 minutes. The supernatant was applied to 1 ml glutathione sepharose high performance resin (GE healthcare) pre-equilibrated in 20 mM Tris

pH 8, 300 mM NaCl, and incubated with agitation for approximately 1 hour. Resin was washed 3 times with 20 mM Tris pH 8, 300 mM NaCl and resuspended in 20 mM Tris pH 8, 300 mM NaCl to a final volume of approximately 2 ml. Proteins were stored bound to the resin at 4°C.

Purification of endogenous *S. acidocaldarius* PolB1

Approximately 3 grams of *S. acidocaldarius* biomass was obtained from exponentially growing cultures of SacTK or SacPolB1 as indicated. Cells were resuspended in 20 ml 10 mM Tris pH 8, 100 mM NaCl, 10 mM MgCl₂, 1 mM CaCl₂, 14 mM β-mercaptoethanol, 0.1 % TritonX-100 and complete protease inhibitors, and incubated on ice for 10 minutes. Cells were lysed by two passages through a French press at 20000 psi, before addition of 600 units of DNaseI (Roche) and incubation for a further 30 minutes on ice. Extracts were centrifuged at 17000 rpm for 30 minutes to remove cell debris. The buffer composition of the supernatant was adjusted to 10 mM Tris pH 8, 300 mM NaCl, 10 mM imidazole, 14 mM β-mercaptoethanol before application to 2 ml of Ni-NTA resin, equilibrated in the same buffer. Bound protein was washed with 20 mM imidazole before elution with 500 mM imidazole. PolB1-containing fractions were diluted six-fold in 10 mM Tris pH 8, 1 mM DTT, applied to a 1 ml MonoQ column and eluted over a 15 ml linear gradient of 50–1000 mM NaCl. PolB1-containing fractions were concentrated by TCA precipitation and analysed by SDS-PAGE. Indicated gel bands were excised with a scalpel, digested with trypsin and analysed by LC-MS/MS (Sir William Dunn School of Pathology Central Proteomics Facility). The highest scoring protein identified by MASCOT is indicated for each band.

Western blotting

Following separation by SDS-PAGE, protein samples were transferred to nitrocellulose membrane by wet transfer at 250 mA for 1 hour in 50 mM Tris pH 8.3, 384 mM glycine, 20 % methanol. The membrane was blocked in 5 % milk dissolved in TBST (10 mM Tris pH 8, 150

mM NaCl, 0.1 % Tween20) for 1 hour at room temperature (or overnight at 4°C). The membrane was then incubated with the indicated primary antibody in 5 % milk/TBST for 1 hour at room temperature. Following washing in TBST (3 x 10 minutes) the membrane was incubated with anti-IgG (anti-rabbit or mouse as appropriate) horseradish peroxidase-conjugated secondary antibody (Pierce) for 1 hour at room temperature. The membrane was washed in TBST (3 x 10 minutes), and chemiluminescence detection performed using ECL (GE Healthcare).

PROTEIN LABELLING

Ellman's test

15 nmol of Lig1 E2C.M516C was incubated with 14 mM β -mercaptoethanol for 1 hour at 4°C, before buffer exchange into reaction buffer (100 mM sodium phosphate pH 8, 1 mM EDTA) using a MicroSpin G-25 column (GE Healthcare). The sample was immediately incubated with 200 μ g Ellman's reagent (Pierce) in a final volume of 2.5 ml reaction buffer for 15 minutes at room temperature. The concentration of free sulfhydryls was calculated from the absorbance of the reaction product, TNB, at 412 nm using an extinction coefficient of $14150 \text{ M}^{-1}\text{cm}^{-1}$.

Lig1 fluorophore modification

For small-scale labelling to determine cysteine reactivity, 0.5 nmol of indicated Lig1 protein was incubated with 14 mM β -mercaptoethanol for 1 hour at 4°C, before buffer exchange into 20 mM Tris pH 7, 300 mM NaCl using a MicroSpin G-25 column. Proteins were incubated with 5 nmol Cy3B-maleimide (GE Healthcare) or 5 nmol Alexa647-maleimide (Molecular Probes) for 2 hours at room temperature with gentle agitation. Reactions were quenched by addition of 14 mM β -mercaptoethanol and excess unreacted fluorophore was removed by repeated spin concentration using an Ultracel YM-30 filter (Millipore). The concentration of Lig1, Cy3B and Alexa647 after reaction was measured by absorbance at 280 nm, 559 nm, and 650 nm

respectively across a 1 mm pathlength, using extinction coefficients of $63260 \text{ M}^{-1}\text{cm}^{-1}$, $130000 \text{ M}^{-1}\text{cm}^{-1}$, and $240000 \text{ M}^{-1}\text{cm}^{-1}$ respectively. When calculating protein concentration, correction factors of 0.08 for Cy3B and 0.03 for Alexa647 were applied to take into account absorbance of these fluorophores at 280 nm.

For analysis by chromatography, 5 nmol E2C.M516C was reacted with 0, 5 or 20 nmol Alexa647-maleimide as above. Following removal of unreacted fluorophores, samples were diluted four-fold in 10 mM Tris pH 8, applied to a 1 ml MonoQ column and eluted over a linear gradient of 75–1000 mM NaCl. The extent of protein labelling for protein-containing fractions was measured as above.

For large-scale labelling, 15 nmol E2C.M516C was incubated with 14 mM β -mercaptoethanol for 1 hour at 4°C, before buffer exchange into 20 mM Tris pH 7, 300 mM NaCl using a 5 ml HiTrap desalting column (GE healthcare). Protein was incubated with 15 nmol Alexa647-maleimide and 30 nmol Cy3B-maleimide for 3 hours at room temperature with gentle agitation. The reaction was quenched by the addition of 14 mM β -mercaptoethanol and excess unreacted fluorophore was removed by buffer exchange into 20 mM Tris pH 8, 300 mM NaCl using a 5 ml HiTrap desalting column. Sample was diluted four-fold in 10 mM Tris pH 8, applied to a 1 ml MonoQ column and eluted over a linear gradient of 75–1000 mM NaCl. Protein labelling was analysed as above and fractions corresponding to protein with a single conjugated Alexa647 were collected.

SUBSTRATE PREPARATION

All oligonucleotides were supplied by Eurogentec (see Appendix I for sequences). Where indicated, oligonucleotides were either 5' end labelled with [γ - ^{32}P]-ATP (PerkinElmer) using T4 polynucleotide kinase (NEB) or 3' end labelled with [α - ^{32}P]-cordycepin (PerkinElmer) using

Terminal Transferase (NEB). RNase inhibitor (NEB) was included when generating RNA-containing substrates.

Polymerase substrate

10 pmol of 5' ³²P-labelled 'Upstream' was annealed to 20 pmol of 'Template' by incubating at 95°C for 10 minutes before gradual cooling to room temperature in 10 mM Tris pH 8, 50 mM NaCl, 1 mM EDTA. Unincorporated radionucleotide was removed by passage through a MicroSpin G-25 column before storage at -20°C. All annealing reactions described below were performed identically.

Flap substrate

10 pmol of 5' ³²P-labelled 'Downstream_flap' was annealed 20 pmol of 'Template' and 40 pmol of 'Upstream_1nt'.

Nick substrates

For hairpin substrates, 10 pmol of 'Nick2' was 5' phosphorylated with unlabelled ATP using T4 polynucleotide kinase and annealed to 10 pmol of 5' ³²P-labelled 'Nick1'. For non-ligatable substrates, 'Nick1' was replaced with 'Nick1ddC'. For fluorescence anisotropy substrates, 'Nick2' was replaced by 'Nick2_internal' or 'Nick2_3' as indicated.

For comparison of DNA or RNA ligation, 10 pmol 5' ³²P-labelled 'Downstream_DNA' or 'Downstream_RNA' as indicated was annealed to 20 pmol of 'Template' and 40 pmol of 'Upstream_nick'.

Okazaki fragment substrates

To visualise upstream events, 10 pmol of 5' ³²P-labelled 'Upstream' was annealed to 20 pmol of 'Template' and 40 pmol of 'Downstream_RNA' or 'Downstream_DNA' as indicated. To visualise downstream events, 10 pmol of 5' or 3' ³²P-labelled 'Downstream_RNA' or

‘Downstream_DNA’, as indicated, was annealed to 20 pmol of ‘Template’ and 40 pmol of ‘Upstream’. For pull-down assays, substrates were constructed as above with ‘Template’ replaced by ‘B-Template’.

For idling assays, 10 pmol of 5' ^{32}P -labelled or unlabelled ‘Upstream_I’ was annealed to 20 pmol of ‘Template_I’ and 40 pmol of ‘Downstream_I’ (if present).

ENZYME ACTIVITY ASSAYS

Polymerase assay

Reaction mixtures (20 μl) contained 50 mM Tris pH 7.5, 50 mM KCl, 10 mM MgCl_2 , 100 μM dNTPs, 1 mM ATP, 100 $\mu\text{g/ml}$ BSA, 10 mM DTT, 0.1 pmol ^{32}P -labelled substrate, and the indicated amounts of PolB1 and PCNA. Reactions were incubated at 50°C for 15 minutes before being quenched by the addition of 20 μl loading dye (8 M urea, 1X TBE, 0.05 % Bromophenol Blue) and incubated for 5 minutes at 99°C. Reaction products were resolved by electrophoresis through a 12 % polyacrylamide, 8 M urea, 1X TBE gel. Gels were dried and visualised by autoradiography.

Flap cleavage assay

Reaction mixtures (20 μl) contained 70 mM Tris pH 7.6, 50 mM KCl, 10 mM MgCl_2 , 100 $\mu\text{g/ml}$ BSA, 5 mM DTT, 0.1 pmol ^{32}P -labelled substrate, and the indicated amounts of Fen1 and PCNA. Reactions were incubated at 50°C for 15 minutes before being quenched by the addition of 20 μl loading dye (8 M urea, 1X TBE, 0.05 % Bromophenol Blue) and incubated for 5 minutes at 99°C. Reaction products were resolved by electrophoresis through a 12 % polyacrylamide, 8 M urea, 1X TBE gel. Gels were dried and visualised by autoradiography.

Ligation assay

Reaction mixtures (20 μ l) contained 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 1 mM ATP, 100 μ g/ml BSA, 10 mM DTT, 0.1 pmol ^{32}P -labelled substrate and the indicated amounts of Lig1 and PCNA. Reactions were incubated at 70°C for 30 minutes before being quenched by the addition of 20 μ l loading dye (8 M urea, 1X TBE, 0.05 % Bromophenol Blue) and incubated for 5 minutes at 99°C. Reaction products were resolved by electrophoresis through a 12 % polyacrylamide, 8 M urea, 1X TBE gel. Gels were dried and visualised by autoradiography.

Maturation assay

Reaction mixtures (20 μ l) contained 50 mM Tris pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 100 μ M dNTPs, 4 mM $MnCl_2$, 1 mM ATP, 100 μ g/ml BSA, 10 mM DTT, 0.1 pmol ^{32}P -labelled substrate, 0.1 pmol PolB1 (or 50 fmol Dpo4), 0.1 pmol Fen1, 1 pmol Lig1 and 20 pmol PCNA as indicated. Reactions were incubated at 50°C for 15 minutes before being quenched by the addition of 20 μ l loading dye (95 % formamide, 5 mM EDTA, 0.05 % Bromophenol Blue) and incubated for 5 minutes at 99°C. Reaction products were resolved by electrophoresis through a 12 % polyacrylamide, 8 M urea, 1X TBE gel. Gels were dried and visualised by autoradiography.

To analyse flap cleavage size, reaction products were separated on a 20 % polyacrylamide, 8 M urea, 1X TBE gel and visualised by autoradiography. For reference, 0.1 pmol 5' ^{32}P -labelled 'Downstream_RNA' oligonucleotide was digested with 10 μ g Phosphodiesterase I (Sigma) for 5 minutes at 37°C in reaction buffer before quenching.

Idling assay

Assays were adapted from those described previously (Garg et al, 2004). Reaction mixtures (20 μ l) contained 50 mM Tris pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 4 mM $MnCl_2$, 1 mM ATP, 100 μ g/ml BSA, 10 mM DTT, 0.1 pmol substrate, 100 μ M dATP, dCTP, dTTP, 10 μ M dGTP, and 50 nM [α - ^{32}P]-dGTP (PerkinElmer). 0.1 pmol PolB1, 50 fmol Dpo4, 0.1 pmol Fen1, and 20 pmol

PCNA were added as indicated, and reactions were incubated at 50°C for 60 minutes before being quenched by the addition of 10 µl 25 mM EDTA, 1 % SDS, 5 mM dGTP and 5 mM dGMP. Where indicated, reaction products were digested with 100 units of Exonuclease III (NEB) at 37°C for 30 minutes. Reaction products were resolved by thin layer chromatography on PEI/cellulose in 0.5 M LiCl, 0.5 M formic acid and visualised by autoradiography.

DNA BINDING ASSAYS

Electrophoretic mobility shift assay (EMSA)

For SSB, reaction mixtures (20 µl) containing approximately 0.1 pmol ³²P-labelled substrate and the indicated amount of SSB were incubated for 15 minutes at 50°C in 50 mM Tris pH 7.5, 100 mM KCl, 10 mM MgCl₂, 100 µM dNTPs, 4 mM MnCl₂, 1 mM ATP, 100 µg/ml BSA, 10 mM DTT, 5 % glycerol. For PolB1, 0.1 pmol ³²P-labelled substrate and the indicated amount of PolB1 was incubated for 10 minutes at room temperature in 70 mM Tris pH 7.6, 50 mM KCl, 10 mM MgCl₂, 100 µg/ml BSA, 5 mM DTT, 5 % glycerol. For Lig1, reaction mixtures (10 µl) containing 0.1 pmol ³²P-labelled substrate and the indicated amount of Lig1 were incubated for 10 minutes at room temperature in 50 mM Tris pH 7.5, 5 mM DTT, 50 µg/ml BSA, 5 % glycerol. Protein-DNA complexes were resolved by electrophoresis through a 6 % polyacrylamide, 0.5X TBE gel (0.25X for Lig1) run at 175 V. The gel was dried and visualised by autoradiography.

For western blotting of Lig1 EMSA, reactions were performed as above with the addition of 500 nM unlabelled substrate. Following electrophoresis, DNA and protein were simultaneously transferred to Zetaprobe (Bio-Rad) and nitrocellulose respectively, as described previously (Demczuk et al, 1993). DNA was detected by autoradiography. Lig1 was detected by western blotting using a 1:1000 dilution of Lig1 antiserum (Dionne et al, 2003).

Fluorescence anisotropy

Binding reactions (80 μ l) contained 50 mM Tris pH 7.5, 5 mM DTT, 50 μ g/ml BSA and 50 nM Cy3B-labelled substrate. Lig1 was titrated into reactions in small volumes; dilution was taken into account when calculating final protein concentrations. After each protein addition reactions were mixed and incubated for 2 minutes at room temperature before measurement of anisotropy at room temperature. Measurements were performed by an L-format spectrofluorimeter fitted with polarisers. Excitation and emission wavelengths were 540 nm and 580 nm respectively. Standard G-factor correction was applied automatically by the instrument. Each measurement consisted of 10 data points over 60 seconds which were averaged. The change in anisotropy at each protein concentration was determined by subtracting the value of anisotropy in the absence of protein. Three independent titrations were performed. The dissociation constant (K_D) of Lig1 binding to substrate was calculated by fitting a single-site binding model to the data using GraphPad Prism.

DNA pull-down

To immobilise substrate, 25 μ g of Dynabeads M280 streptavidin (Invitrogen) were incubated with 0.1 pmol biotin-modified substrate in 5 μ l binding buffer (5 mM Tris pH 7.5, 1 M NaCl, 0.5 mM EDTA) for 10 minutes at room temperature in a LoBind tube (Eppendorf). Using a magnetic tube rack, beads were washed once with 100 μ l binding buffer and twice with 100 μ l reaction buffer (50 mM Tris pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 100 μ M dGTP, 100 μ M dTTP, 4 mM $MnCl_2$, 1 mM ATP, 100 μ g/ml BSA, 10 mM DTT) and resuspended in 20 μ l reaction buffer. All subsequent steps were performed at 50°C in a water bath. Immobilised substrate was incubated with 0.1 pmol PolB1, 0.1 pmol Fen1, and 20 pmol PCNA as indicated for 5 minutes, before washing twice with 100 μ l reaction buffer and resuspending in 20 μ l reaction buffer. 100 μ M dATP and 100 μ M dCTP were added and reactions incubated for 15 minutes before being quenched by the addition of 20 μ l loading dye (95 % formamide, 5 mM EDTA, 0.05 %

Bromophenol Blue) and incubated for 5 minutes at 99°C. Reaction products were resolved by electrophoresis through a 12 % polyacrylamide, 8 M urea, 1X TBE gel. Gels were dried and visualised by autoradiography

PROTEIN-PROTEIN INTERACTION ASSAYS

GST pull-down

Approximately 10 µg of GST-fusion protein was incubated with 10 µg of interaction partner in 100 µl TBST (10 mM Tris pH 8, 150 mM NaCl, 0.1 % Tween20) for 30 minutes at room temperature with agitation. TBSTM was used when testing GST-PCNA fusion proteins (TBST supplemented with 5 mM MgCl₂). Beads were recovered by centrifugation and washed three times with 500 µl TBST/TBSTM. Beads were resuspended in 1X SDS loading buffer and proteins were resolved on a 10 or 12 % polyacrylamide gel and visualised with Coomassie staining or western blotting with the indicated antisera: 1:1000 Lig1, 1:5000 Dpo4 or 1:2000 Vps4 (Dionne et al, 2003; Samson et al, 2008). For use in control reactions where indicated, Vps4 was provided by Rachel Samson (Samson et al, 2008).

Yeast two-hybrid assay

S. cerevisiae haploid strains AH109 or Y187 were transformed according to the Matchmaker protocol (Clontech) with plasmids encoding the indicated proteins fused to the GAL4 activation domain (pGAD-T7) or GAL4 DNA-binding domain (pGBK-T7). Transformants were selected by growth on SC-Leu (pGAD-T7) or SC-Trp (pGBK-T7) agar after 3 days at 30°C. Appropriate haploid strains were mated on YPED agar overnight at 30°C. Diploid strains possessing both pGAD-T7 and pGBK-T7 were selected by growth on SC-Trp-Leu after 3 days at 30°C. To test for interactions, each diploid strain was replica plated onto both SC-Trp-Leu and SC-Trp-Leu-His agar; this was achieved either by direct re-streaking of a single colony, or by plating 5 µl of a single colony grown in YPED media to an OD₆₀₀ of 0.1. Growth is shown after 3 days at 30°C;

growth on media lacking histidine after this time is considered confirmation of a protein-protein interaction.

To test for fusion-protein expression, indicated diploid strains were grown overnight in YPED media and cells harvested by centrifugation at 3000 rpm for 3 minutes. Cells were resuspended in 300 µl 20 % TCA and vortexed with glass beads. Extracts were centrifuged at 3000 rpm for 10 minutes and the pellet resuspended in 1X SDS loading buffer. Samples were analysed by western blotting with a 1:1000 dilution of anti-c-Myc (Millipore) or 1:10000 PCNA1, 1:5000 PCNA2 and 1:10000 PCNA3 antisera (Dionne et al, 2003).

Analytical gel filtration

Samples (100 µl) containing 20 mM Tris pH 8, 300 mM NaCl and the indicated proteins were incubated for 10 minutes at room temperature, before loading onto a Superdex 200 HR 10/30 column equilibrated in 20 mM Tris pH 8, 300 mM NaCl. 1 ml fractions were collected and analysed by SDS-PAGE on a 4–12 % gradient gel (Invitrogen). Molecular weight calibration was performed using a low molecular calibration kit (GE healthcare).

Protein thermostability

Samples (50 µl) containing 20 mM Tris pH 8, 300 mM NaCl and 0.5 nmol PolB1, 0.5 nmol PBP1, and 0.5 nmol PBP2 as appropriate were incubated at the indicated temperatures for 25 minutes. Samples were centrifuged for 10 minutes at 13000 rpm to remove heat-denatured protein and the supernatant was analysed by SDS-PAGE.

SULFOLOBUS GROWTH AND ANALYSIS

***Sulfolobus* cell growth and synchronisation**

S. solfataricus P2 cells were grown at 75°C in Brock's medium pH 3.2 supplemented with 0.2 % (w/v) tryptone. *S. solfataricus* P2 cells were synchronised as described previously (Duggin et

al, 2008), except G1 cells were collected after 5 hours. To analyse protein expression, 2 ml samples were collected at the indicated time-points during synchronous growth. Cells were centrifuged at 13000 rpm for 10 minutes, resuspended in 1X SDS loading buffer, and analysed by western blotting with 1:10000 PCNA1, 1:5000 PCNA2, 1:10000 PCNA3, or 1:1000 TBP antisera as indicated.

S. acidocaldarius cells were grown at 75°C in Brock's medium pH 3.2 supplemented with 0.1 % (w/v) tryptone, 0.2 % (w/v) xylose and 0.025 mg/ml uracil. To analyse tagged PolB1, *S. acidocaldarius* SacTK or PolB1 strains as indicated were grown to an OD₆₀₀ of 0.1–0.2, and a 1 ml aliquot withdrawn and centrifuged at 13000 rpm for 10 minutes. Cells were resuspended in 1X SDS loading buffer and the samples were analysed by western blotting with 1:1000 diluted anti-c-myc (Millipore) or 1:1000 diluted anti-His antibodies (AbD Serotec).

Flow cytometry

400 µl of cells were fixed with 800 µl ice-cold 80 % ethanol, washed, and resuspended in 10 mM Tris pH 7.4, 10 mM MgCl₂, 100 µg/ml RNaseA, 10 µM SYTOX green (Invitrogen). Samples were analysed on a Dako CyAn instrument (Beckman Coulter) as described previously (Robinson et al, 2007).

Genomic DNA analysis

Approximately 0.5 grams of *S. acidocaldarius* biomass was resuspended in 10 mM Tris pH 8, 100 mM NaCl, 10 mM MgCl₂, 1 mM CaCl₂, 0.1 % Triton X-100 and complete protease inhibitors, and incubated on ice for 10 minutes. The indicated amount of DNaseI was added and samples incubated on ice for a further 30 minutes. DNA was extracted twice with phenol:chloroform and once with chloroform, before ethanol precipitation. Samples were run on a 0.5 % agarose gel and DNA visualised with SafeView (NBS Biologicals).

DATA VISUALISATION AND ANALYSIS

All visualisation and quantification of autoradiograms was performed using Advanced Image Data Analyzer (AIDA). Graphs were generated by Microsoft Excel, except fluorescence anisotropy binding data which was analysed by GraphPad Prism. Multiple sequence alignments were performed by Clustal W2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2>) and visualised using BOXSHADE (http://www.ch.embnet.org/software/BOX_form). Mapping of residue conservation onto Lig1 was performed by the ConSurf server (<http://consurf.tau.ac.il>). All visualisation of proteins structures was performed with PyMOL (<http://www.pymol.org>). Protein structure prediction was performed by Phyre² (<http://www.sbg.bio.ic.ac.uk/phyre2>). Visualisation and analysis of chromatograms was performed using UNICORN (GE Healthcare).

Chapter 3 – *In vitro* reconstitution of archaeal Okazaki fragment maturation

ABSTRACT

Chromosomal DNA replication requires one daughter strand – the lagging strand – to be synthesised as a series of discontinuous, RNA-primed Okazaki fragments, which must subsequently be matured into a single covalent DNA strand. In eukaryotic organisms, numerous factors have been shown to be involved in this essential multi-step process, but the requirements for archaeal Okazaki fragment maturation are still poorly understood. To understand this process in more detail in archaea, Okazaki fragment maturation was reconstituted *in vitro* from purified components derived from *Sulfolobus solfataricus*. Six proteins were necessary and sufficient for coupled DNA synthesis, RNA primer removal and DNA ligation. PolB1, Fen1 and Lig1 provided the required catalytic activities, with coordination of their activities dependent upon the heterotrimeric DNA sliding clamp, PCNA.

INTRODUCTION

An extensive series of biochemical and genetic studies have elucidated multiple distinct pathways operating in eukaryotic organisms to resolve discontinuous Okazaki fragments into a single covalent daughter strand (see Chapter 1 for more detail). However, the mechanistic basis of Okazaki fragment maturation in archaea is still poorly understood and lacking in direct experimental insight. Consistent with a eukaryotic-like DNA replication machinery, archaeal Okazaki fragments have been demonstrated to be approximately 100 nucleotides (nt) in length, and possess a 5' RNA primer of approximately 10 ribonucleotides, similar to eukaryotes (Matsunaga et al, 2003). However, the heterodimeric primase which catalyses primer synthesis in archaea is not associated with a Pol α orthologue, and therefore archaeal primers

are likely to comprise simply of RNA rather than the RNA-DNA chimera present in eukaryotes (Liu et al, 2001; Lao-Sirieix & Bell, 2004). The model crenarchaeon *S. solfataricus* appears to possess orthologues of just a subset of the Okazaki fragment maturation factors present in eukaryotes: a B-type DNA polymerase (PolB1), flap endonuclease 1 (Fen1), and the ATP-dependent DNA ligase I (Lig1). These factors have therefore been proposed to mediate Okazaki fragment maturation in this organism, although this has not been directly demonstrated.

Furthermore, *S. solfataricus* also possesses a DNA sliding clamp orthologous to eukaryotic proliferating cell nuclear antigen (PCNA). Previous work has revealed that, unusually, *S. solfataricus* PCNA is a heterotrimer composed of three distinct subunits which assemble in a defined manner (Dionne et al, 2003). Interestingly, these subunits have preferred interaction partners. Most intriguingly with relevance to Okazaki fragment maturation, PCNA1 specifically interacts with Fen1, PCNA2 is bound by PolB1, and PCNA3 preferentially interacts with Lig1. Furthermore, previous work revealed that the PCNA heterotrimer can act as a bridge between Fen1, PolB1 and Lig1 (Dionne et al, 2003). Taken together, these data suggest that PCNA may organise these three enzymes into a complex to facilitate Okazaki fragment maturation.

The following two chapters investigate the role of PolB1, Fen1 and Lig1 in DNA replication in archaea and the functional relevance of their observed interactions with heterotrimeric PCNA. This chapter tests the hypothesis that PolB1, Fen1, Lig1 and PCNA are involved in Okazaki fragment maturation in *S. solfataricus*. Reconstitution of this process *in vitro* from purified proteins demonstrated that the enzyme activities of PolB1, Fen1 and Lig1 were necessary and sufficient to provide the range of activities required to process discontinuous Okazaki fragments into a single covalent daughter strand. Furthermore, an interaction of each of these enzymes with PCNA was required for completion of this complex multistep process.

RESULTS

All three subunits of *S. solfataricus* heterotrimeric PCNA are expressed throughout S-phase

Previous biochemical and structural analyses have demonstrated that the *S. solfataricus* PCNA subunits PCNA1, PCNA2 and PCNA3 assemble into a stable heterotrimer *in vitro* and *in vivo* (Dionne et al, 2003; Pascal et al, 2006). Furthermore, all three PCNA genes are essential for viability in the related archaeon *Sulfolobus islandicus* (Zhang et al, 2010). It was therefore investigated if this unique assembly does indeed support DNA replication in *S. solfataricus* cells. By adapting a cell cycle synchronisation method previously developed for *Sulfolobus acidocaldarius* (Duggin et al, 2008), *S. solfataricus* cells were synchronised in the G1 phase of the cell cycle and synchronous growth was followed through S-phase and into G2 (Figure 3.1A). Analysis of protein expression during synchronous growth demonstrated that PCNA1, PCNA2 and PCNA3 were all present throughout S-phase (Figure 3.1B). This therefore further supports the hypothesis that *S. solfataricus* heterotrimeric PCNA facilitates DNA replication in this organism. With reference to the loading control TATA-box binding protein (TBP), it is noteworthy that there does not appear to be any significant cell cycle modulation of PCNA protein levels across the time-course tested.

PolB1, Fen1 and Lig1 demonstrate the catalytic activities required for Okazaki fragment maturation

To test whether *S. solfataricus* PolB1, Fen1 and Lig1 are involved in Okazaki fragment maturation, recombinant versions of all three proteins were purified (Figure 3.2A, C and E). To ensure these purified proteins possess their predicted activities, they were each tested individually on model oligonucleotide substrates *in vitro* and reaction products were separated on the basis of size by denaturing gel electrophoresis.

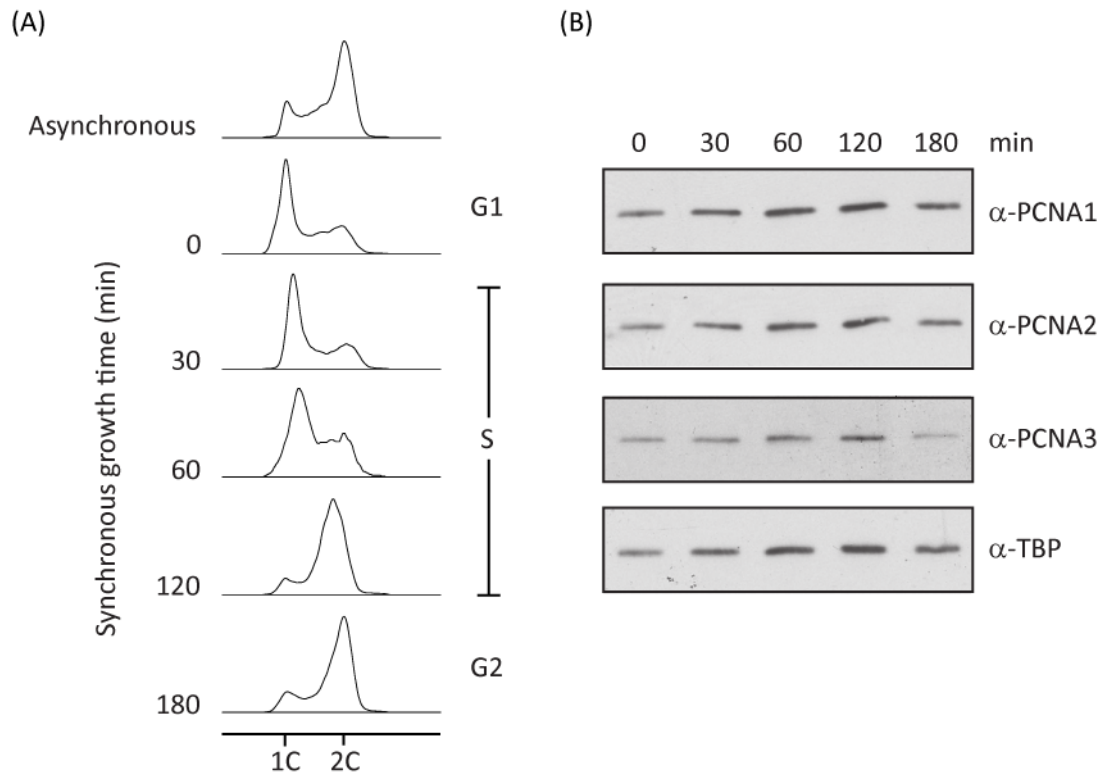


Figure 3.1 – Expression of *S. solfataricus* heterotrimeric PCNA throughout S-phase

(A) Cell cycle synchronisation of *S. solfataricus* cells. An asynchronous culture was applied to the 'Baby machine' apparatus, G1 phase cells were collected at time 0 and subsequently grown synchronously at 75°C. The DNA content of cells was analysed at the indicated time-points by flow cytometry; '1C' and '2C' indicate fluorescent signal corresponding to 1 or 2 copies of genomic DNA, respectively.

(B) Analysis of PCNA expression throughout the cell cycle. Samples were analysed at the indicated time-points during synchronous growth by western blotting with the indicated antisera. TATA-box binding protein (TBP) was analysed as a loading control.

PolB1 demonstrated robust primer extension activity to generate duplex DNA (Figure 3.2B), confirming its function as a DNA polymerase. Similarly, Fen1 cleaved the 5' flap from an optimal double-flap substrate determined previously (Kao et al, 2002), demonstrating its expected structure-specific nuclease activity (Figure 3.2D). Finally, as predicted, Lig1 covalently ligated a model nicked duplex substrate (Sriskanda et al, 2000) (Figure 3.2F). Thus these purified proteins possess the DNA polymerase, nuclease and ligase activities required for Okazaki fragment maturation.

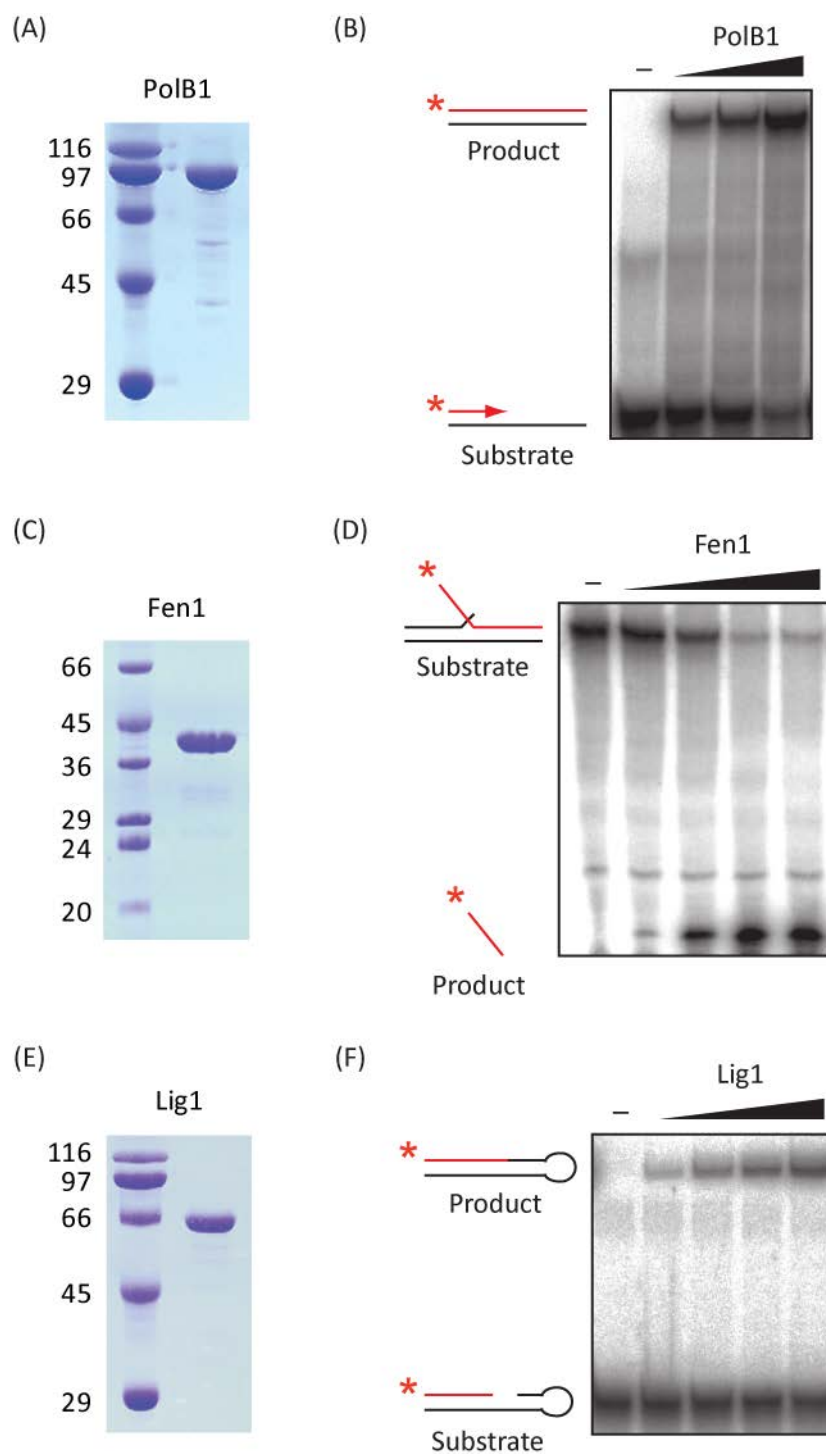


Figure 3.2 – Catalytic activities of PolB1, Fen1 and Lig1 *in vitro*

(A) Purified recombinant PolB1, analysed by SDS-PAGE and visualised by Coomassie staining. Molecular weight markers are shown on the left, sizes are in kDa.

(B) PolB1 *in vitro* activity on a model primer-template substrate. Substrate and product are indicated on the left with radiolabelled strand indicated in red and label position with an asterisk. Substrate was incubated with 0, 2.5, 5, 10 fmol PolB1 and reaction products were separated by denaturing PAGE.

(C) Purified recombinant Fen1, analysed by SDS-PAGE.

(D) Fen1 *in vitro* activity on an optimal double-flap substrate. Substrate was incubated with 0, 16, 258, 4125, 66000 fmol Fen1 and reaction products were separated by denaturing PAGE.

(E) Purified recombinant Lig1, analysed by SDS-PAGE.

(F) Lig1 *in vitro* activity on a model nicked substrate. Substrate was incubated with 0, 3.4, 6.7, 13.5, 27 pmol Lig1 and reaction products were separated by denaturing PAGE.

Previous studies have identified PCNA-interacting protein (PIP) motifs present in PolB1, Fen1 and Lig1 responsible for mediating their specific interaction with the different subunits of PCNA (Dionne et al, 2003; Pascal et al, 2006). To further confirm the identity of these PIP motifs, and to facilitate functional analyses described below, purified recombinant versions of PolB1, Fen1 and Lig1 possessing point mutations in their predicted PIP motifs were also generated. These point mutations did not appear to globally disrupt PolB1, Fen1 or Lig1, as measured by their identical purification and basal function on model substrates *in vitro* compared to wild-type (WT) proteins (Figure 3.3B, D and F). Substitution of two hydrophobic or aromatic residues in the PIP motif with alanine however, was found to disrupt interactions observed by GST pull-down between PolB1 and PCNA2 (Figure 3.3A), Fen1 and PCNA1 (Figure 3.3C), and Lig1 and PCNA3 (Figure 3.3E). These different PIP motifs highlight the considerable sequence variability in crenarchaeal PCNA binding motifs compared to the largely invariant eukaryotic motif. Notably, PolB1 appears to utilise a shortened PIP motif to bind PCNA2, which resembles the bacterial β clamp binding motif (Doré et al, 2006). Additionally, the Lig1 PIP motif is unusually located internally, rather than at the N- or C-terminus of the protein.

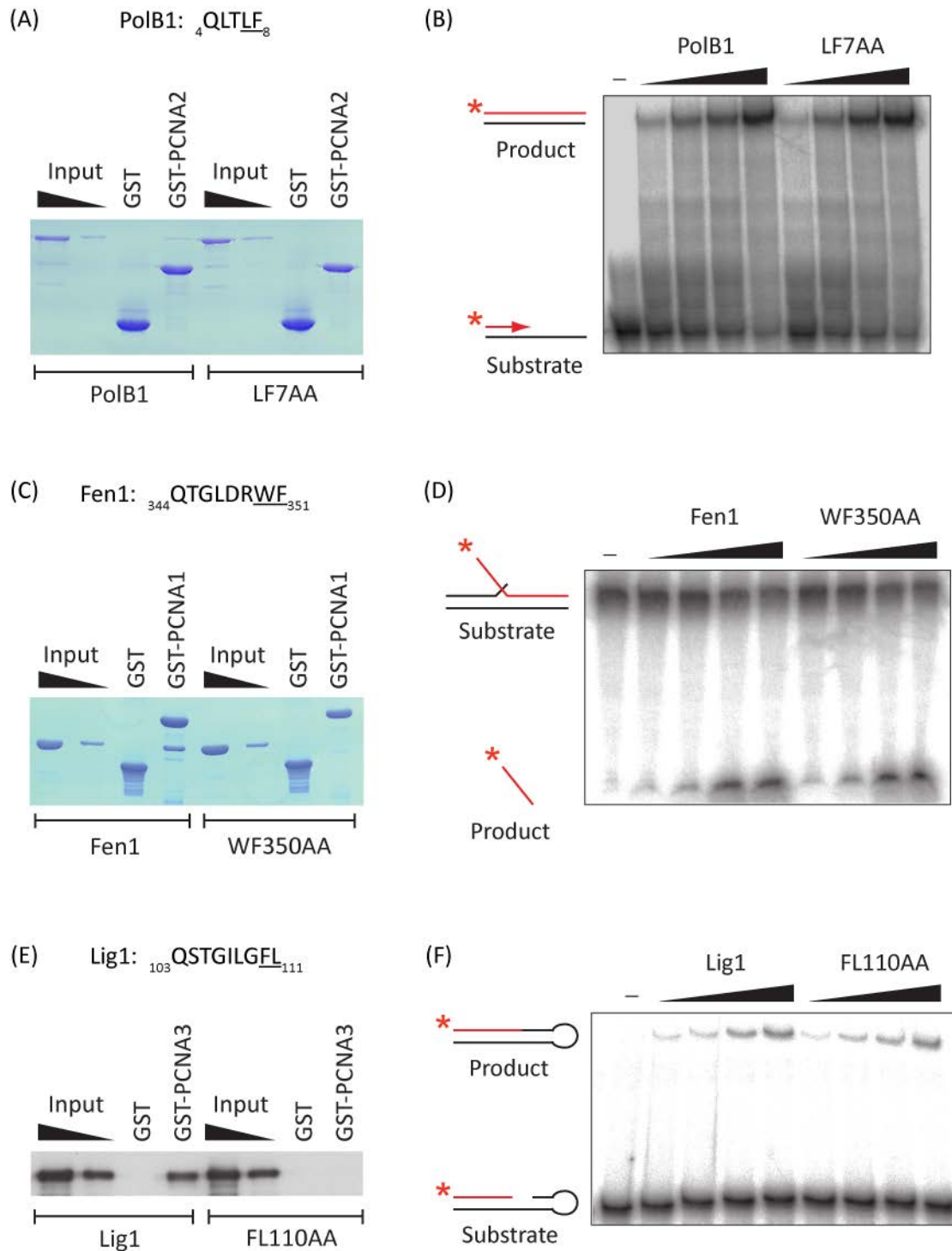


Figure 3.3 – Characterisation of PolB1, Fen1 and Lig1 PIP mutants

(A) Pull-down of WT or PIP mutant (LF7AA) PolB1 by GST-tagged PCNA2. The PIP motif sequence is shown at the top, with mutated residues underlined. Reactions were analysed by SDS-PAGE and visualised with Coomassie staining. 50 % and 10 % PolB1 inputs are shown.

(B) Comparison of the *in vitro* activity of WT or PIP mutant (LF7AA) PolB1 on an optimal primer-template substrate. Substrate and product are indicated on the left with radiolabelled strands indicated in red

and label position by an asterisk. Substrate was incubated with 0, 3.5, 7, 14, 28 fmol PolB1 proteins and reaction products were separated by denaturing PAGE.

(C) Pull-down of WT or PIP mutant (WF350AA) Fen1 by GST-tagged PCNA1. The PIP motif sequence is shown at the top. 50 % and 10 % Fen1 inputs are shown.

(D) Comparison of the *in vitro* activity of WT or PIP mutant (WF350AA) Fen1 on an optimal double-flap substrate. Substrate was incubated with 0, 55, 219, 875, 3500 fmol Fen1 proteins.

(E) Pull-down of WT or PIP mutant (FL110AA) Lig1 by GST-tagged PCNA3. The PIP motif sequence is shown at the top. Reactions were analysed by western blotting with Lig1 antiserum. 50 % and 10 % Lig1 inputs are shown.

(F) Comparison of the *in vitro* activity of WT or PIP mutant (FL110AA) Lig1 on a model nicked substrate. Substrate was incubated with 0, 1.37, 2.75, 5.5, 11 pmol Lig1.

The replicative polymerase, PolB1, performs strand displacement synthesis on a model lagging strand substrate

Previous bioinformatic analysis has demonstrated that while crenarchaeal species possess multiple B-type DNA polymerases, many of these may in fact be catalytically inactive, as judged by the presence of mutations in several key catalytic motifs (Iwai et al, 2000; Rogozin et al, 2008). To better understand the role of *S. solfataricus* PolB1 in DNA replication, a similar bioinformatic analysis was performed focussing just on *Sulfolobus* species (Figure 3.4). In *S. solfataricus*, only PolB1 appears to possess the full complement of conserved residues required for 5'-3' DNA polymerase and 3'-5' exonuclease activities, with *S. solfataricus* PolB2 and PolB3 possessing several putative inactivating substitutions. Interestingly, while this also appears to be the case with *S. acidocaldarius* polymerases, *S. tokodaii* PolB3 possesses all the required polymerase and exonuclease catalytic residues, suggesting variability may exist between organisms in the number of active polymerases. Indeed, activity from two distinct B-type polymerases has been observed in *Aeropyrum pernix* and *Pyrodictium occultum*, although the role of these polymerases *in vivo* has not been characterised (Uemori et al, 1995; Cann et al, 1999). Surprisingly, recent biochemical analysis has demonstrated *in vitro* activity for both *S. solfataricus* PolB2 and PolB3, raising the intriguing possibility that they may only require a

subset of catalytic residues to perform certain functions *in vivo* (Choi et al, 2011; Bauer et al, 2012). The properties demonstrated for these enzymes – such as weak polymerase and DNA binding activities, low processivity, and the ability to bypass some template lesions – suggests they are not capable of acting as replicative DNA polymerases and may play a minor genome maintenance function *in vivo*. *S. solfataricus* PolB1 is therefore presumed to act as the primary DNA polymerase on both the leading and lagging strands, although this remains to be demonstrated *in vivo*.

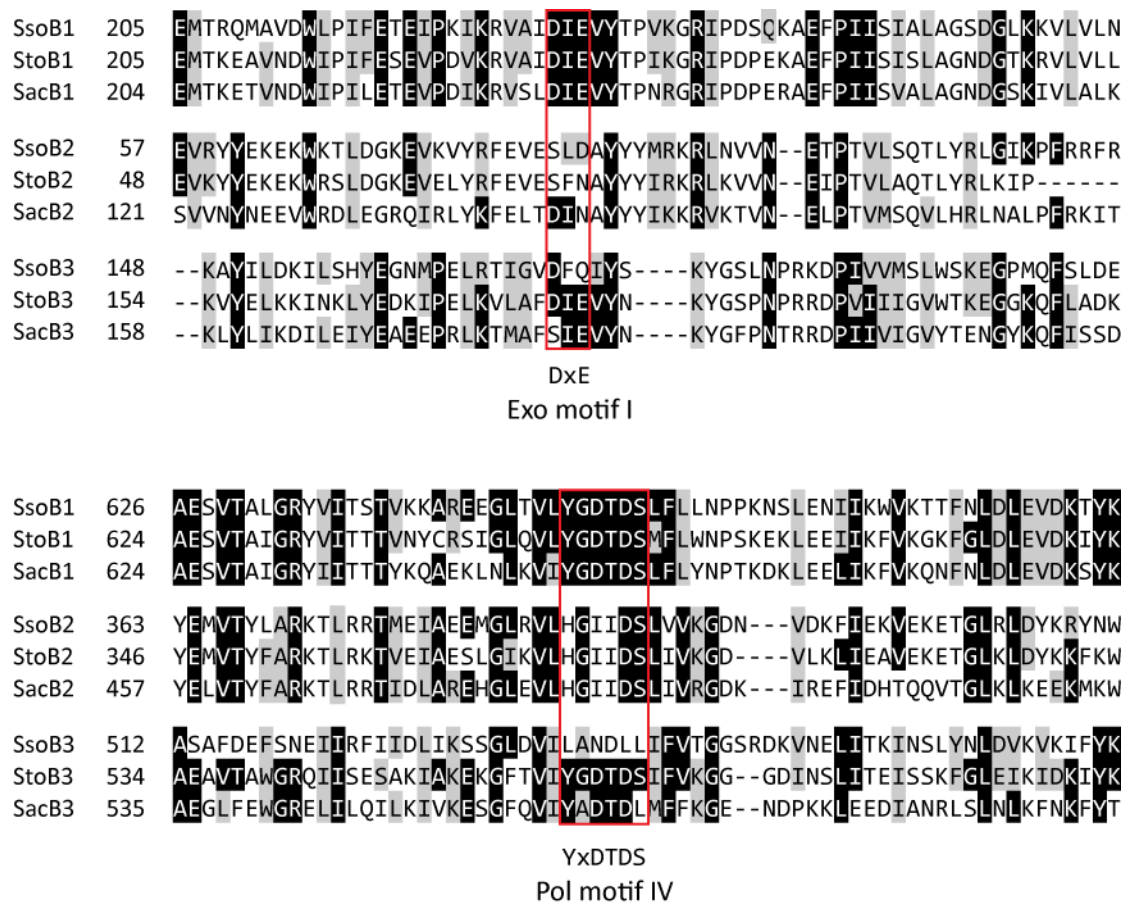


Figure 3.4 – Sequence comparison of B-type polymerases in *Sulfolobus* species

Multiple sequence alignment of B-type polymerases from *S. solfataricus* (Sso), *S. tokodaii* (Sto), and *S. acidocaldarius* (Sac). Residues with identity to the consensus are highlighted in black, residues with similarity to the consensus are highlighted in grey. Sequences were aligned along their entire length but only selected regions are shown for clarity; the first residue shown of each protein sequence is numbered on the left. Important exonuclease (exo) and polymerase (pol) catalytic motifs are boxed in red with the consensus sequence indicated below.

In order to analyse the process of Okazaki fragment maturation *in vitro* a synthetic lagging strand substrate was constructed, consisting of a 100 nt template strand annealed to two complementary strands mimicking the known structure of archaeal Okazaki fragments (Matsunaga et al, 2003): an upstream strand with a free 3' end available for DNA synthesis and a downstream strand primed by 13 nt of RNA (Figure 3.5A). The activity of the putative replicative polymerase PolB1 was first measured on this lagging strand substrate in the presence of dNTPs, with reaction products separated on the basis of size by denaturing gel electrophoresis (Figure 3.5B). In the absence of a downstream Okazaki fragment, all concentrations of PolB1 employed were sufficient to fully-extend the primer across the entire region of un-replicated template. When a downstream Okazaki fragment was present however, a significant number of intermediate-length DNA synthesis products were observed, consistent with the progress of PolB1 being impeded by the presence of downstream dsDNA. Notably, the majority of these extension products did proceed into the region occupied by the downstream Okazaki fragment (>42 nt) with some full-length synthesis also observed, particularly at higher concentrations of PolB1. The fact that this was not accompanied by any degradation of the downstream fragment (Figure 3.5B, right-hand panel) suggests that PolB1 performed strand displacement synthesis, excluding the downstream Okazaki fragment into a flap structure. Interestingly, this strand displacement synthesis activity appeared to be more active on a physiological RNA-primed Okazaki fragment than on a fragment composed entirely of DNA (Figure 3.5C); similar properties have been observed for the *S. cerevisiae* lagging strand polymerase Pol δ (Stith et al, 2008). PolB1 therefore possesses the properties required to displace unwanted RNA primers on the lagging strand.

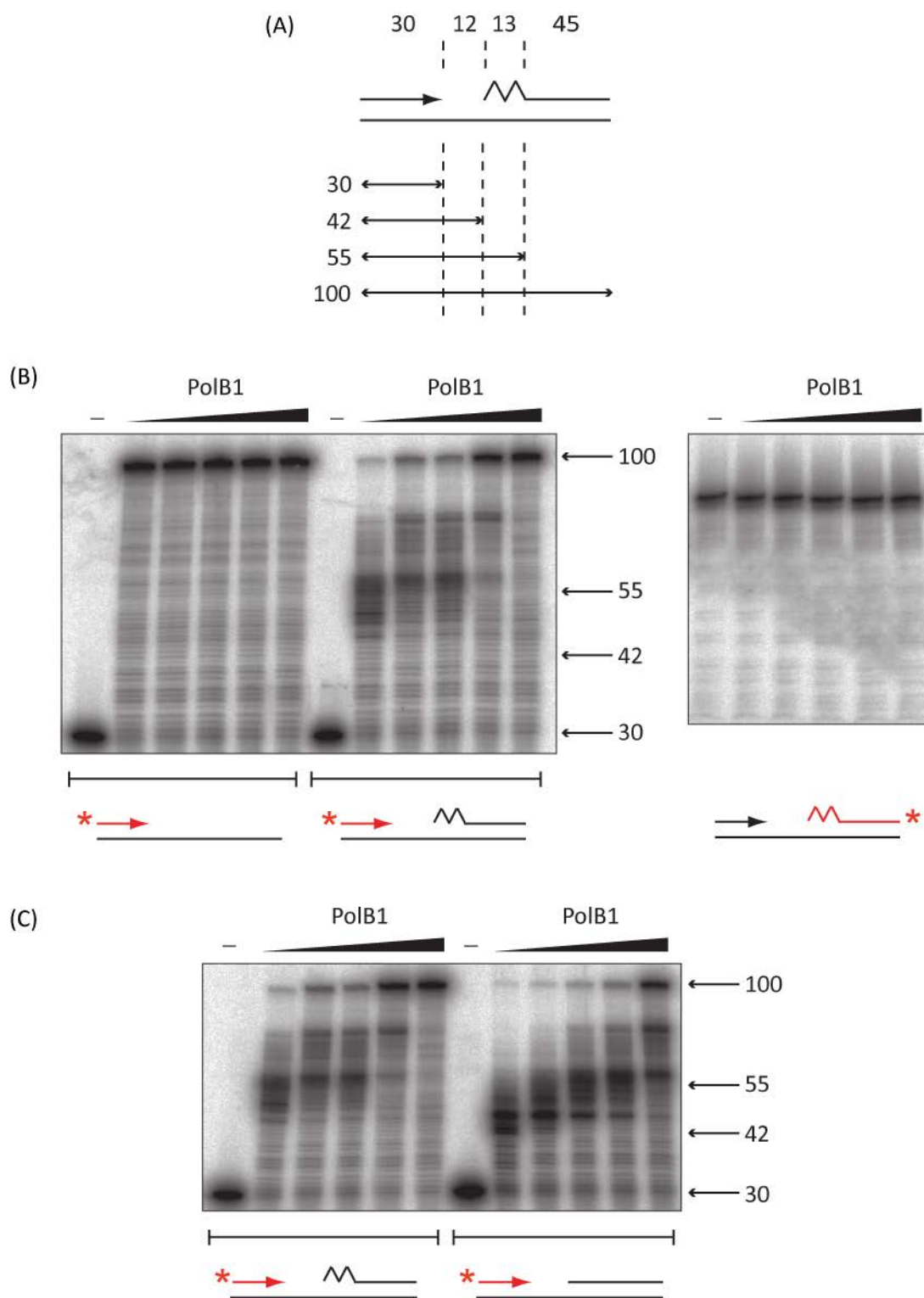


Figure 3.5 – Characterisation of PolB1 strand displacement activity

(A) A schematic of the lagging strand substrate used in *in vitro* assays. The jagged region of the substrate denotes a ribonucleotide primer. The size of each section, in nucleotides, is indicated above, and the size of notable reaction products is indicated below. For sequences of oligonucleotides see Appendix I.

(B) Comparison of strand displacement activity by PolB1 on substrates possessing or lacking a downstream Okazaki fragment. Reactions contained 0, 15, 30, 60, 120, 240 fmol PolB1. Reaction products were separated by denaturing PAGE. Substrates are indicated below gels, the radiolabelled strand is indicated in red and the position of the label indicated by an asterisk. The migration of notable DNA synthesis products is indicated on the right.

(C) Comparison of PolB1 activity on substrates containing RNA-primed or DNA-primed Okazaki fragments. Experiments were performed as in (B).

PolB1 coordinates with Fen1 in a PCNA-dependent manner to excise RNA primers from Okazaki fragments

The effect of other factors on this strand displacement activity by PolB1 was next examined. In the presence of PCNA, PolB1 activity was significantly stimulated, further extending the primer into the region occupied by the downstream Okazaki fragment, with a 50 % increase in full-length DNA synthesis observed (Figure 3.6A, compare lanes 3 and 4). For simplicity, a fully-functional version of heterotrimeric PCNA was used in which the three subunits are fused in a single polypeptide (Dionne et al, 2008); further implications of this are discussed below and in Chapter 4. Given that strand displacement by PolB1 was predicted to give rise to a downstream flap structure, the activity of the flap endonuclease, Fen1, was also tested on the lagging strand substrate. While the addition of Fen1 and PCNA alone had no effect (Figure 3.6A, lanes 2 and 8), the addition of Fen1 to PolB1 and PCNA resulted in cleavage of the downstream Okazaki fragment (Figure 3.6A, lane 11), suggesting that flap structures generated by PolB1 during strand displacement synthesis were indeed a substrate for subsequent Fen1 activity. Substitution of the Fen1 PIP mutant into the primer excision assay (Figure 3.6A, lane 12) resulted in a four-fold reduction in downstream flap cleavage. The interaction of Fen1 with PCNA is thus crucial to coupling its activity to that of PolB1, which in turn is significantly stimulated by PCNA.

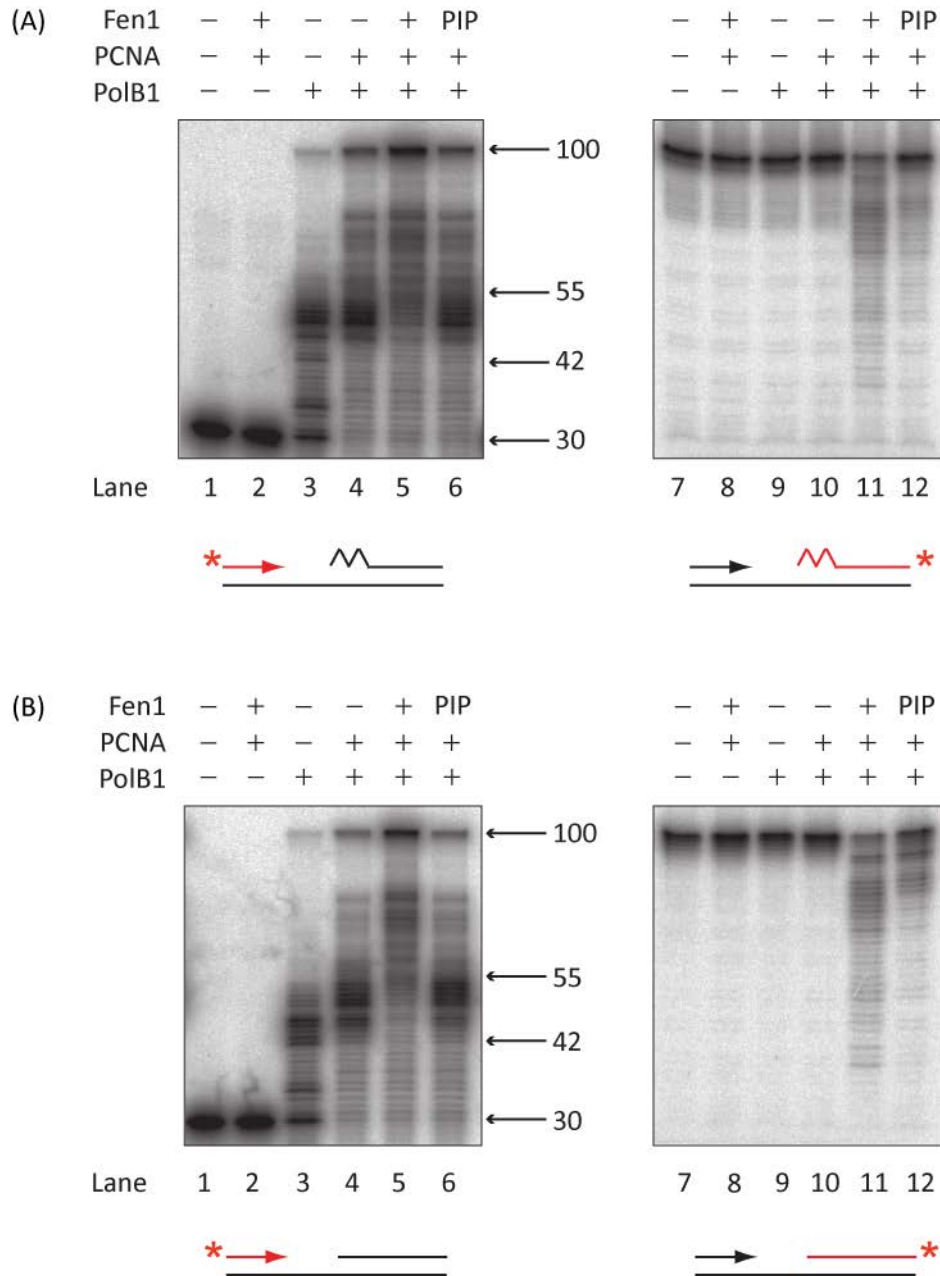


Figure 3.6 – PCNA-dependent coupling of PolB1 and Fen1 activities *in vitro*

(A) Activity of purified PolB1, Fen1 and PCNA on an *in vitro* lagging strand substrate. Reactions contained 0.1 pmol PolB1, 0.1 pmol Fen1 and 20 pmol PCNA. Reaction products were separated by denaturing PAGE. The migration of notable DNA synthesis products is indicated on the right, sizes are in nucleotides. 'PIP' denotes the substitution of a Fen1 mutated in the PIP motif that is defective in PCNA binding. Substrates are indicated below gels, with the radiolabelled strand indicated in red and the position of the label indicated by an asterisk.

(B) Activity of purified PolB1, Fen1 and PCNA on an *in vitro* lagging strand substrate lacking a downstream RNA primer. Experiments were performed as in (A).

Interestingly, the presence of Fen1 resulted in a further 15 % increase in the yield of full-length DNA synthesis product (Figure 3.6A, lane 5). The fact that this was eliminated when the Fen1 PIP mutant was substituted into the reactions (Figure 3.6A, lane 6) suggests that this was not due to a direct physical interaction between PolB1 and Fen1, but rather was mediated by the activity of Fen1 on their shared substrate. Thus it is proposed that following flap cleavage, PolB1 subsequently re-engages the upstream primer to initiate further RNA primer displacement and excision, the activities of PolB1 and Fen1 becoming coupled to perform nick translation. Such activity is ideally suited to the requirement of replacing initiating RNA with error-free DNA.

The ability of PolB1 and Fen1 to excise the 5' end of the downstream Okazaki fragment appeared to be equally efficient on RNA-primed fragments as on DNA-only fragments (Figure 3.6B). This suggests that there is not a requirement for additional factors to specifically excise ribonucleotides. The similar extent of 5' RNA or DNA excision, despite an apparent preference of PolB1 to displace RNA over DNA, can be explained by the more comparable strand displacement observed in the presence of PCNA, consistent with its classical role in tethering polymerases to their substrate (Figures 3.6A and B, compare lanes 3 and 4).

The coupled activities of PolB1 and Fen1 resulted in a range of cleaved products, with up to 30 nt of the downstream Okazaki fragment removed. These could be the result of multiple rounds of flap creation and cleavage or alternatively could arise from single flap cleavage events with a broad size distribution. To distinguish between these mechanisms, the initial flap cleavage products were measured by labelling the downstream Okazaki fragment at the 5' end and performing maturation assays as before (Figure 3.7). Most cleavage products were relatively short (<4nt), in agreement with measurements on *S. cerevisiae* proteins (Garg et al, 2004; Rossi & Bambara, 2006), suggesting the RNA primer excision observed was the result of multiple cycles of coordinated PolB1 and Fen1 activity. Interestingly, when this analysis was performed in the presence of the Fen1 PIP mutant, together with the expected

decrease in the magnitude of flap cleavage, there was also an increase in the peak flap cleavage length from 1–2nt to 2–3nt. This supports a mechanistic role for PCNA whereby co-localisation of PolB1 and Fen1 facilitates more rapid flap cleavage as it emerges from the strand displacing polymerase.

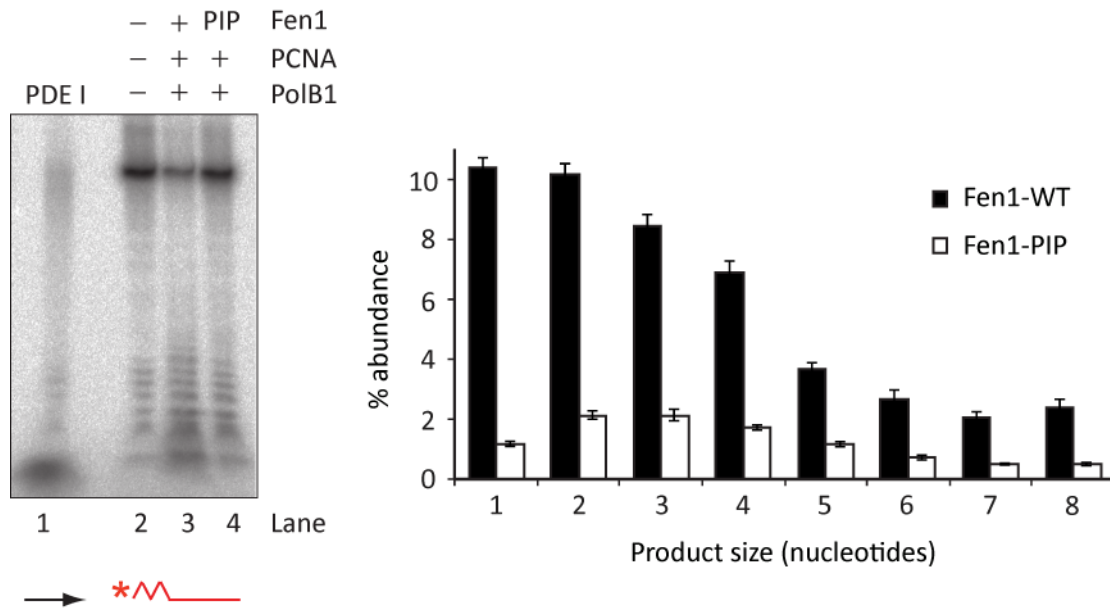


Figure 3.7 – Analysis of cleavage products generated by PolB1 and Fen1.

Reactions were performed as in Figure 3.6. Substrate is indicated below gel, with the radiolabelled strand indicated in red and the position of the label indicated by an asterisk. PDE I indicates substrate digestion by snake phosphodiesterase I to indicate the migration of monoribonucleotides. Quantification of product abundance as a percentage of total DNA after subtraction of background degradation (lane 2) is illustrated on the right. Values are mean $\pm$ standard error of the mean (SEM) (n=3).

PolB1 undergoes limited idling at a nick

It has previously been reported that the lagging strand DNA polymerase in *S. cerevisiae*, Pol δ , can maintain its position at the nick between two Okazaki fragments by undergoing a process termed idling: multiple cycles of nucleotide addition and accompanying strand displacement

followed by exonuclease-mediated backtracking. This process has been proposed to prevent excessive strand displacement synthesis and maintain a suitable nick structure for subsequent Okazaki fragment ligation (Garg et al, 2004).

To test if PolB1 also undergoes idling, a variant lagging strand substrate was employed in which dGTP could only be incorporated upon strand displacement of the first two nucleotides of the downstream Okazaki fragment. Exonuclease-mediated backtracking specifically at this site, indicative of idling, could then be measured by the release of dGMP. Relative to background exonuclease activity observed on a primer-template substrate, a specific increase in nucleotide turnover by PolB1 was observed in the presence of a downstream Okazaki fragment, suggesting PolB1 is capable of idling at a nick (Figure 3.8A and B). As a control, no significant nucleotide turnover was observed in the presence of the *S. solfataricus* Y-family DNA polymerase Dpo4, which lacks 3'-5' exonuclease activity. Furthermore, PolB1 idling was reduced by 30 % in the presence of Fen1. Consistent with the previous analysis, this suggests that cleavage of strand displaced RNA by Fen1 allows PolB1 to perform further DNA synthesis, rather than backtracking. However, the rate of PolB1 idling observed was almost three orders of magnitude lower than that observed with *S. cerevisiae* Pol δ . This low rate of idling was not due simply to efficient strand displacement by PolB1, as a significant amount of pausing at the nick site was observed under these conditions (Figure 3.8C). Therefore, while PolB1 does appear able to undergo limited idling, it is not clear whether this is an important mechanism underlying Okazaki fragment maturation in *S. solfataricus*.

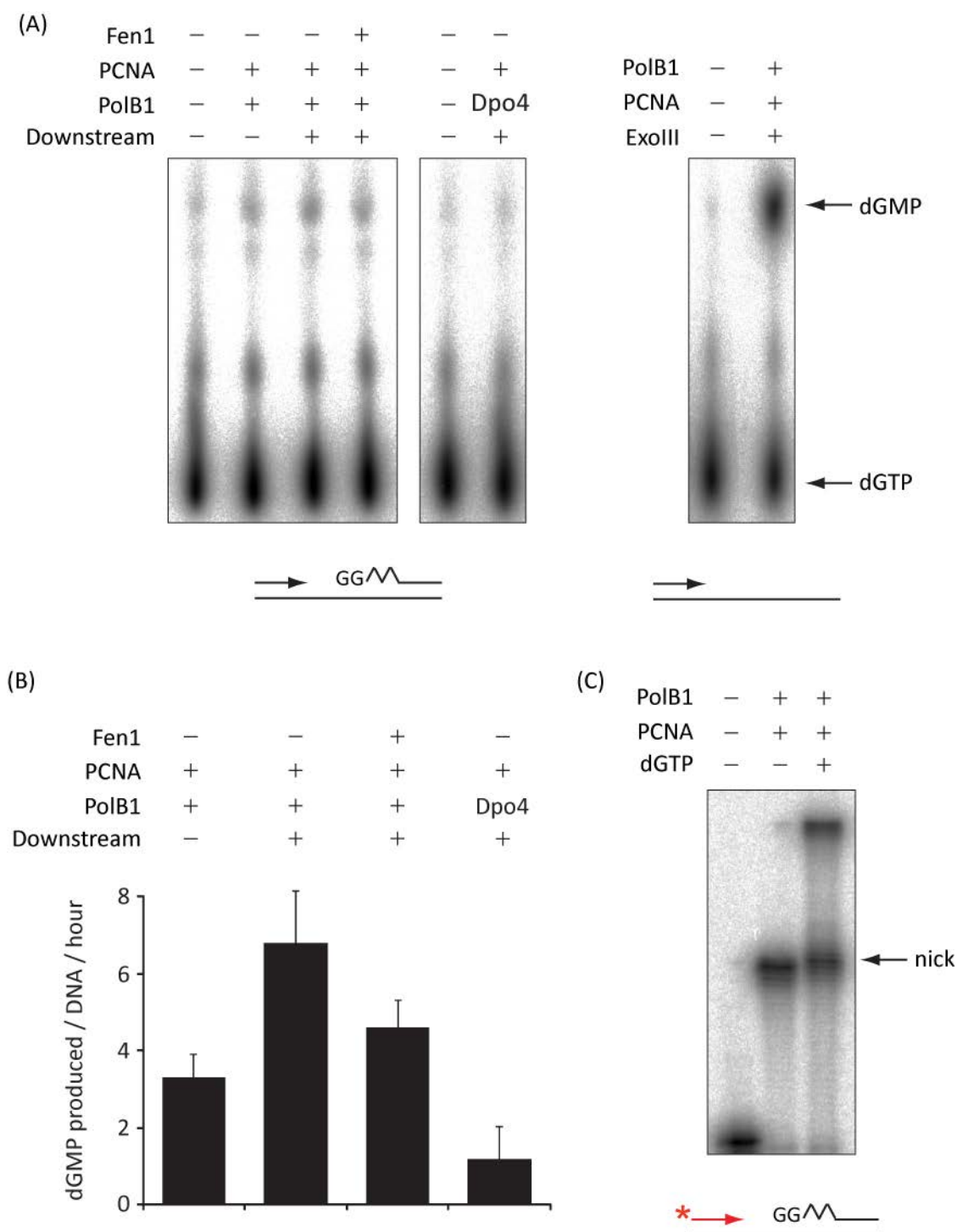


Figure 3.8 – Analysis of PolB1 idling at a nick

(A) dGTP conversion to dGMP by DNA polymerases. Reactions contained 0.1 pmol PolB1 or 50 fmol Dpo4, 20 pmol PCNA, and 0.1 pmol Fen1 as indicated. 'Downstream' indicates the presence of a downstream Okazaki fragment, 'Dpo4' indicates the substitution of PolB1 with Dpo4. Substrate, shown at the bottom, contains two guanosine residues at the 5' end of the downstream Okazaki fragment. Turnover of ^{32}P -GTP at these sites was determined by separation of reaction products by thin layer chromatography. As shown on the right, migration of dGMP was determined by PolB1-PCNA mediated

dGTP incorporation into a primer-template substrate, shown at the bottom, followed by product digestion by Exonuclease III (ExoIII).

(B) Quantification of experiments performed as in (A). Values are mean $\pm$ SEM (n=3).

(C) Strand displacement synthesis by PolB1 and PCNA on the idling substrate. Reactions were performed as in (A), but reaction products were separated by denaturing PAGE. Substrate is shown below the gel, radiolabelled strand is shown in red with the label position indicated by an asterisk. The migration of DNA synthesis products extended to the nick site is indicated.

Lig1 coordinates with PolB1 and Fen1 in a PCNA-dependent manner to ligate processed Okazaki fragments

Coupled strand displacement synthesis and RNA cleavage by PolB1 and Fen1 was predicted to generate a DNA nick competent for ligation. To test this hypothesis, the activity of Lig1 on the lagging strand substrate was tested (Figure 3.9A). Incubation of Lig1 with the full complement of proteins required for RNA primer excision (PolB1, PCNA and Fen1, lane 4) resulted in the production of a slower migrating DNA molecule which, under the denaturing conditions of the electrophoresis, could only arise through covalent joining of the labelled downstream Okazaki fragment to the upstream Okazaki fragment (Figure 3.9A, lane 5). Notably, omission of Fen1 eliminates not only RNA primer excision but also ligation in the presence of Lig1, despite the fact that gap filling by PolB1 is predicted to generate a nick intermediate (Figure 3.9A, lane 3). It was demonstrated that this was due to the inability of Lig1 to ligate an RNA-containing nick intermediate which would arise as a result of PolB1 activity alone; robust DNA ligase activity was observed under identical conditions (Figure 3.9B). This discrimination has been demonstrated for other DNA ligases and is believed to act as a control mechanism to prevent incorporation of RNA into the nascent lagging strand (Pascal et al, 2004). This mechanism is further supported by the fact that in this reconstituted system, only downstream fragments cleaved beyond the 13 nt of initiating RNA were converted to the ligated form (Figure 3.9A, compare lanes 4 and 5).

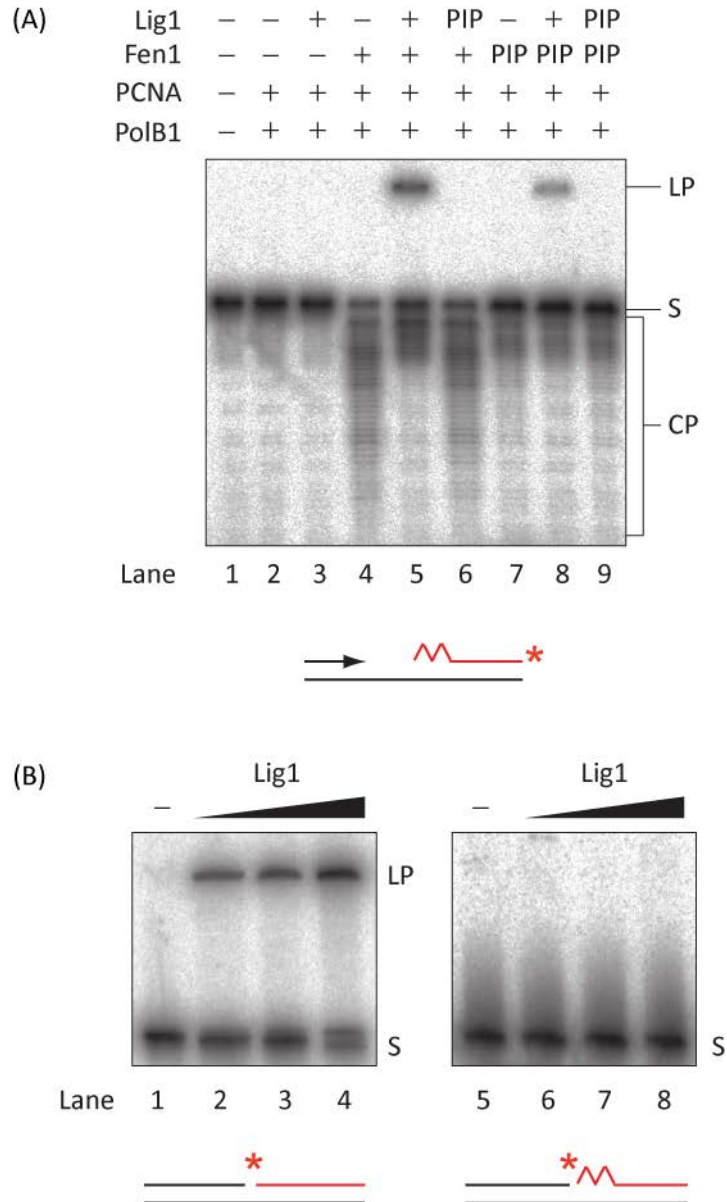


Figure 3.9 – Completion of Okazaki fragment maturation *in vitro* by PolB1, Fen1 and Lig1

(A) Activity of PolB1, Fen1, Lig1 and PCNA on an *in vitro* lagging strand substrate. Reactions contained 0.1 pmol PolB1, 0.1 pmol Fen1, 1 pmol Lig1 and 20 pmol PCNA. Reaction products were separated by denaturing PAGE. The migration of substrate (S), cleavage products (CP) and the ligation product (LP) is indicated. Substrate is indicated below gel, with radiolabelled strand indicated in red and label position with an asterisk. The jagged region of the substrate denotes a 13 ribonucleotide primer. 'PIP' denotes the substitution of Fen1 and/or Lig1 which are mutated in the PIP motif and are defective in PCNA binding.

(B) Substrate specificity of Lig1. Reactions contained 0, 0.625, 1.25, 2.5 pmol Lig1. Reaction products were separated by denaturing PAGE. The migration of substrate (S) and ligated product (LP) are indicated. The presence of an S+1 species at high Lig1 concentrations indicates 5' adenylation of the nick, an intermediate in the ligation reaction.

Thus, fully coupled Okazaki fragment processing was observed *in vitro* with ligation absolutely dependent on prior RNA removal, which in turn was absolutely dependent on prior strand displacement synthesis. This analysis demonstrates that the catalytic activities of three proteins – PolB1, Fen1 and Lig1 – are necessary and sufficient for Okazaki fragment maturation in *S. solfataricus*.

Substitution of the Fen1 PIP mutant into this reaction resulted in reduced RNA primer excision, and thus a 5.5-fold reduction in final ligation (Figure 3.9A, lanes 7 and 8). Furthermore, substitution of the Lig1 PIP mutant either alone (Figure 3.9A, lane 6) or in combination with Fen1 PIP (Figure 3.9A, lane 9) completely abolished ligation of Okazaki fragments. Thus interaction of all three essential enzymes – PolB1, Fen1 and Lig1 – with PCNA is crucial to the efficiency of Okazaki fragment maturation.

It was confirmed that a non-fused PCNA complex assembled from PCNA1, PCNA2 and PCNA3 (Figure 3.10A) also supported coordinated PolB1, Fen1 and Lig1 maturation activity *in vitro* and thus these observations were not affected by using a fused heterotrimeric PCNA (Figure 3.10B, lanes 5 and 6). Maturation in the presence of the non-covalent heterotrimer was 2.4-fold less efficient than with the fused PCNA. This may have been the result of subunit dissociation which is prevented in the covalent fusion. In this respect it is noteworthy that PCNA3 has been shown to assemble into the ring with a much lower affinity than that of the PCNA1-2 dimer and thus may be more prone to dissociation from the PCNA ring *in vitro* (Dionne et al, 2003). Notably, in the absence of either form of PCNA, no detectable Okazaki fragment ligation was catalysed by PolB1, Fen1 and Lig1, confirming a requirement for PCNA in addition to these enzymatic activities (Figure 3.10B, lane 4).



(B) Comparison of non-fused and fused heterotrimeric PCNA in stimulating Okazaki fragment maturation *in vitro*. Reactions contained 0.1 pmol PolB1, 0.1 pmol Fen1, 1 pmol Lig1 and 20 pmol PCNA (total heterotrimer). Reaction products were separated by denaturing PAGE. Substrate is indicated below gel with radiolabelled strand indicated in red and label position with an asterisk. The jagged region of the substrate denotes a 13 ribonucleotide primer. The migration of substrate (S), cleavage products (CP) and the ligation product (LP) is indicated.

In all cellular organisms, ssDNA is bound by essential ssDNA-binding proteins (SSBs) which play a central role in numerous DNA transactions (Wold, 1997). In *S. solfataricus*, this function is performed by the monomeric SSB protein; notably, while the domain structure of this protein

resembles bacterial SSB, structurally it is actually more similar to eukaryotic RPA (Wadsworth & White, 2001; Haseltine & Kowalczykowski, 2002; Kerr et al, 2003). Given the prevalence of ssDNA template on the lagging strand during replication, it is possible that *S. solfataricus* SSB may play a role in Okazaki fragment maturation. To test this hypothesis, purified recombinant SSB was produced (Figure 3.11A), and it was first determined if SSB was able to bind to the lagging strand substrate by electrophoretic mobility shift assay (EMSA) (Figure 3.11B). Under the same reaction conditions used to measure Okazaki fragment maturation, an interaction of SSB with the lagging strand substrate was observed, with the pattern of binding broadly consistent with previous estimates of a dissociation constant for SSB in the low nanomolar range (Wadsworth & White, 2001). At lower concentrations of SSB a single protein-DNA complex was observed, presumably resulting from coating of the entire 12 nt region of ssDNA; this is likely to be mediated by 2–3 molecules of SSB (Wadsworth & White, 2001; Haseltine & Kowalczykowski, 2002). However, at higher concentrations, larger, more heterogeneous protein-DNA complexes were observed, suggesting SSB was able to extend its ssDNA binding footprint by melting dsDNA, as has been observed previously (Cubeddu & White, 2005; Marsh et al, 2006). Indeed, at the highest concentrations of SSB a decrease in the size of protein-DNA complexes relative to the starting substrate was observed, indicating complete SSB-mediated dissociation of the labelled Okazaki fragment from the template.

The effect of SSB on full maturation by PolB1, Fen1, Lig1 and PCNA was then tested, across the range of concentrations at which substrate binding but not full melting was observed (Figure 3.11C). Addition of SSB had a modest inhibitory effect on maturation, resulting in a two-fold reduction in ligation at the maximum concentration tested. This may have been the result of residual SSB-mediated melting or breathing of the substrate, even at the lower concentrations of SSB employed here, which could have potentially disrupted the ability of PolB1 to engage the substrate. Notably, the *S. solfataricus* chromatin protein, Alba, has been shown to prevent duplex melting by SSB *in vitro*, and it is likely that disruption of

DNA replication substrates *in vivo* is more controlled (Cubeddu & White, 2005). It does not appear from these data therefore that SSB is essential for Okazaki fragment maturation in *S. solfataricus*, at least not in the final stages of the process reconstituted here. Given the negligible effect of SSB under these assay conditions, it was omitted from subsequent experiments.

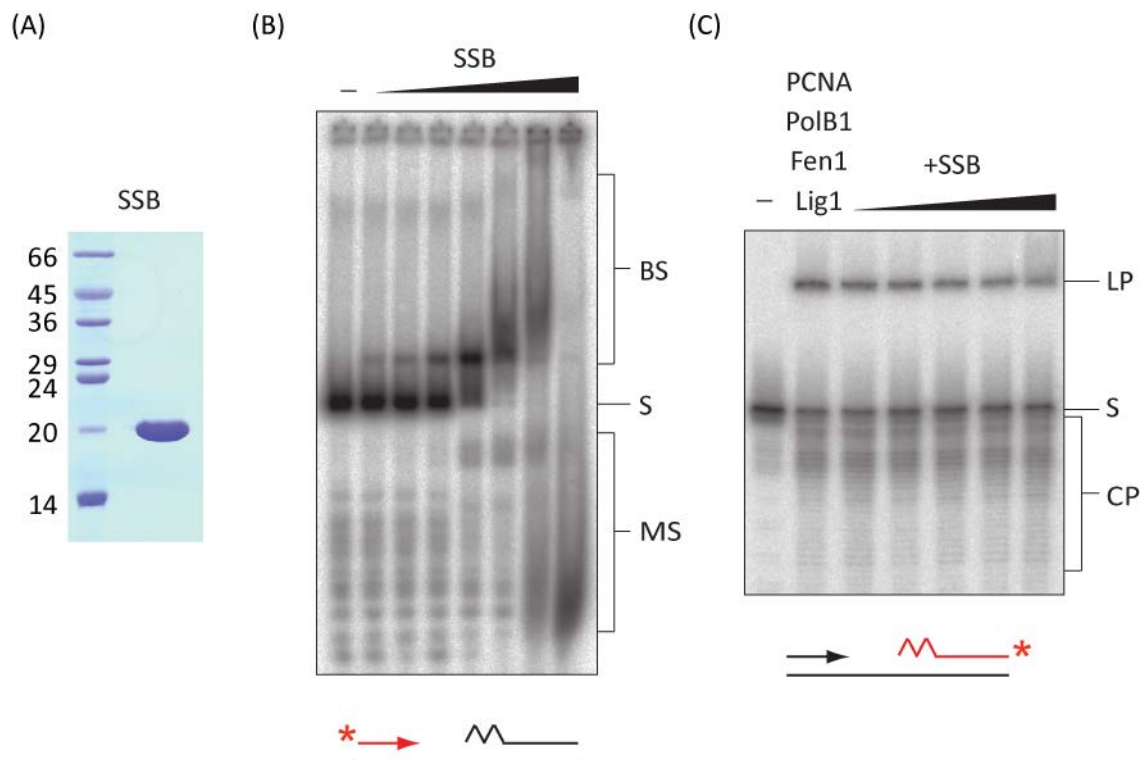


Figure 3.11 – The effect of SSB on Okazaki fragment maturation *in vitro*

(A) Analysis of purified SSB by SDS-PAGE, visualised by Coomassie staining. Molecular weight markers are labelled on the left, sizes are in kDa

(B) SSB binding to a lagging strand substrate. Reactions contained 0, 0.14, 5.6, 22.3, 89.1, 356, 1425, 5700 fmol of SSB. Reaction products were separated by native gel electrophoresis, and the migration of substrate (S), bound substrate (BS), and melted substrate (MS) is indicated on the right. Substrate is shown below the gel, radiolabelled strand is indicated in red and label position with an asterisk.

(C) The effect of SSB on Okazaki fragment maturation. Reactions contained 0.1 pmol PolB1, 0.1 pmol Fen1, 1 pmol Lig1, 20 pmol PCNA, and 0, 0.14, 5.6, 22.3, 89.1, 356 fmol of SSB. Reaction products were separated by denaturing PAGE. The migration of substrate (S), cleavage products (CP) and the ligation product (LP) is indicated

RNaseHII is not a core component of the *S. solfataricus* Okazaki fragment maturation machinery

Type 2 RNaseH enzymes have been implicated in RNA removal during Okazaki fragment maturation in eukaryotes (Turchi et al, 1994; Qiu et al, 1999) and furthermore have been shown to interact with PCNA in mammalian cells and euryarchaea (Meslet-Cladière et al, 2007; Chon et al, 2009; Bubeck et al, 2011). *S. solfataricus* possesses a predicted RNaseHII gene and so to test a potential role of this protein in Okazaki fragment maturation, its ability to interact with PCNA was investigated by yeast two-hybrid assay. No interactions were detected between RNaseHII and any of the PCNA subunits (Table 3.1 and Figure 3.12A), despite the fusion proteins of RNaseHII, PCNA1, PCNA2 and PCNA3 all being expressed in *S. cerevisiae* in at least one of the clones tested (Figure 3.12B). Consistent with this, no PIP motif could be detected in RNaseHII by sequence analysis.

In order to test this interaction by other means and to potentially test the effect of RNaseHII in the reconstituted Okazaki fragment maturation system, recombinant RNaseHII was overexpressed in *E. coli* for purification. However, no soluble protein could be recovered, despite a number of expression conditions being tested (Figure 3.12C). Fusion of RNaseHII to an N-terminal glutathione S-transferase (GST) tag did enable recovery of soluble protein and this was subsequently immobilised and tested in its ability to pull-down individual PCNA subunits (Figure 3.12D). Similarly to the yeast two-hybrid assay, no interactions could be detected between RNaseHII and PCNA. It is therefore unlikely that RNaseHII is a core component of the Okazaki fragment maturation machinery with PolB1, Fen1 and Lig1. However, it cannot be ruled out that RNaseHII is incorrectly folded in the assays described here, particularly given the problems encountered with heterologous expression. Due to this inability to isolate purified protein, the functional role of RNaseHII in *S. solfataricus* Okazaki fragment maturation remains unclear.

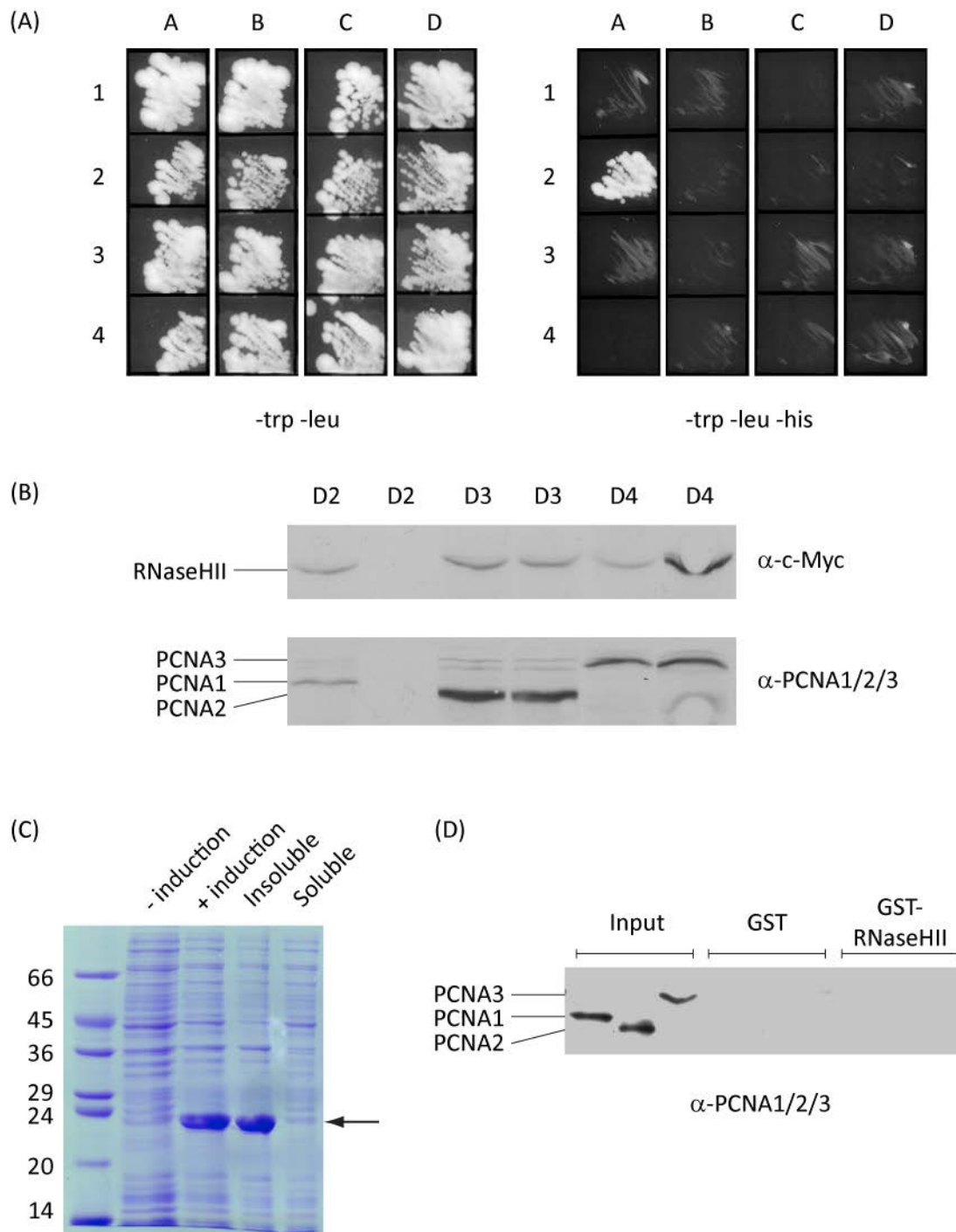


Figure 3.12 – Biochemical analysis of *S. solfataricus* RNaseHII

(A) Yeast two-hybrid analysis of RNaseHII-PCNA interactions. RNaseHII or PCNA subunits fused to the GAL4 DNA binding domain (BD) or GAL4 activation domain (AD) were transformed into yeast, as described in Table 3.1. Successful transformation of yeast is indicated by growth on media lacking tryptophan and leucine (left hand panel). Interaction between the fusion proteins is indicated by growth on media lacking tryptophan, leucine, and histidine (right hand panel). The pattern of interactions shown here was consistent across four independent clones tested for each interaction.

(B) Analysis of fusion protein expression in yeast. Yeast extracts were prepared from the indicated diploid yeast strains and the expression of RNaseHII-BD or PCNA-AD fusions was detected by western blotting with anti-c-myc or PCNA antibodies respectively. Migration of individual PCNA fusion proteins is indicated on the left.

(C) Expression and solubility of *S. solfataricus* RNaseHII in *E. coli*, analysed by SDS-PAGE and visualised by Coomassie staining. Molecular weight markers are shown on the left, sizes are in kDa. The migration of RNaseHII at the expected molecular weight of 24 kDa is indicated with an arrow.

(D) Pull-down of PCNA subunits by GST-tagged RNaseHII. Reactions were analysed by western blotting with PCNA antisera. 10 % PCNA inputs are shown, and the migration of different subunits is indicated on the left.

Table 3.1 – Protein interactions tested by yeast two-hybrid assay in Figure 3.12.

Grid reference	DNA binding domain	Activation domain	Interaction
A1	–	–	–
A2	Murine p53	SV40 large T antigen	+
A3	PCNA1	–	–
A4	PCNA2	–	–
B1	PCNA3	–	–
B2	–	PCNA1	–
B3	–	PCNA2	–
B4	–	PCNA3	–
C1	–	RNaseHII	–
C2	RNaseHII	–	–
C3	PCNA1	RNaseHII	–
C4	PCNA2	RNaseHII	–
D1	PCNA3	RNaseHII	–
D2	RNaseHII	PCNA1	–
D3	RNaseHII	PCNA2	–
D4	RNaseHII	PCNA3	–

DISCUSSION

By reconstituting Okazaki fragment maturation *in vitro*, it was determined that the enzymes PolB1, Fen1 and Lig1 are necessary and sufficient to catalyse coupled processing of discontinuous RNA-primed Okazaki fragments into a single covalent daughter strand in *S. solfataricus*. Furthermore, an interaction of all three of these enzymes with their cognate subunits of heterotrimeric PCNA is required for this multistep processing to occur.

While several different pathways have been proposed to mediate Okazaki fragment maturation in eukaryotic systems, the central role of PolB1, Fen1 and Lig1 in *S. solfataricus* appears to correlate with their orthologues constituting the core pathway for maturation in eukaryotes (Ayyagari et al, 2003). Interestingly, Pif1 and Dna2, two additional factors proposed to play a central role in this process in eukaryotes (Rossi et al, 2008), appear to be absent from *S. solfataricus* and indeed largely from archaea. These additional factors may therefore have evolved after the divergence of archaea and eukaryotes to function redundantly with the core polymerase-Fen1-Lig1 pathway. A Dna2 homologue has been identified in a single archaeal species to date, *Pyrococcus horikoshii* (Higashibata et al, 2003). It is interesting to note that this apparently exceptional case of archaeal Dna2 correlates with the presence of a heterotrimeric RPA in *Pyrococcus* species, orthologous to the eukaryotic complex shown to influence Dna2 and Fen1 activities (Komori & Ishino, 2001). In contrast, the architecture of SSBs in other archaeal species is highly diverse. Some euryarchaea possess three RPA proteins which function independently as homodimers rather than forming a heterotrimeric complex (Robbins et al, 2004), while in other euryarchaeal species individual RPAs may be associated with additional unique subunits (Stroud et al, 2012). In contrast, all crenarchaea possess a monomeric bacterial-like SSB (Wadsworth & White, 2001), with the exception of the thermoproteales which utilise a unique protein lacking an OB-fold domain, termed ThermoDBP (Paytubi et al, 2011). Thus it may be speculated that Dna2 has only

evolved in the minority of archaeal species where a eukaryotic-like RPA is able to bind displaced Okazaki fragments and prevent primer cleavage by Fen1.

One factor previously implicated in Okazaki fragment maturation that does have identifiable orthologues in archaea is RNaseHII. No interaction was observed between *S. solfataricus* RNaseHII and any subunits of PCNA, suggesting it is not a component of the core Okazaki fragment maturation machinery identified here. Indeed, the unrestricted ability of PolB1, PCNA and Fen1 to remove ribonucleotides from the 5' end of Okazaki fragments suggests there is not an essential requirement for RNaseHII, or indeed any other additional factors, in facilitating RNA removal. Genetic studies in the euryarchaeon *Haloferax volcanii* have proposed the dominant role of RNaseHII to be in the resolution of RNA misincorporated into chromosomal DNA, rather than during DNA replication (Meslet-Cladière et al, 2007) and consistent with this, biochemical studies of archaeal RNaseHII enzymes have demonstrated robust activity in removing single ribonucleotides embedded in DNA duplexes (Le Laz et al, 2010; Bubeck et al, 2011). Indeed, recent studies in both yeast and higher eukaryotes have demonstrated an essential function for RNaseH enzymes in removing ribonucleotides erroneously incorporated into genomic DNA by replicative polymerases (Nick McElhinny et al, 2010; Lazzaro et al, 2012; Reijns et al, 2012). Nonetheless, it cannot be ruled out that RNaseHII plays a minor role in Okazaki fragment maturation in *S. solfataricus*; a backup role to Fen1 has been proposed from genetic studies in *S. cerevisiae* (Qiu et al, 1999), while recent biochemical analysis suggests that primer removal by RNaseHII prior to processing by the Pol-Fen1-Lig1 machinery may constitute an alternative pathway in euryarchaea (Henneke, 2012). Unfortunately, this hypothesis could not be tested in the reconstituted system established here due to the insolubility of heterologously expressed *S. solfataricus* RNaseHII.

Also investigated here was the function of SSB, with a modest inhibitory effect on reconstituted Okazaki fragment maturation observed. Because the substrate employed here is designed to study the final stages of Okazaki fragment maturation only, the amount of ssDNA

present is relatively small. The amount of ssDNA generated during lagging strand replication *in vivo*, where synthesis of Okazaki fragments will lag behind template production by the replicative helicase MCM, is anticipated to be much greater, and it is likely that SSB may have a more important role in this context. Indeed, consistent with this hypothesis a recent study has shown a stimulatory effect of SSB on PolB1 when replicating longer stretches of ssDNA template (Choi et al, 2011). Previous studies have demonstrated that *S. solfataricus* SSB physically interacts with numerous proteins, including reverse gyrase, several DNA repair proteins, and RNA polymerase, the latter specifically via a C-terminal acidic tail (Richard et al, 2004; Cubeddu & White, 2005; Napoli et al, 2005; Rolfsmeier & Haseltine, 2010). Analogously, *E. coli* SSB is physically linked to the replicative Pol III holoenzyme through an interaction of its C-terminal acidic tail with the χ subunit of the clamp loader complex (Kelman et al, 1998). However, the minimal effect of SSB on Okazaki fragment maturation observed here suggests that it may not share any direct physical interactions with PolB1, Fen1 or Lig1, and rather any role SSB plays in lagging strand replication may indeed be restricted to its classical function in binding and resolving ssDNA structures.

The reconstituted Okazaki fragment maturation system established here can be of considerable use in future studies in establishing the role of further candidate factors in this complex multistep process, such as additional chromatin proteins and the clamp loading complex, RFC. One factor demonstrated to be essential for the activities of PolB1, Fen1 and Lig1 in Okazaki fragment maturation is PCNA, and the mechanistic basis for this role is explored in more detail in Chapter 4.

Chapter 4 – The mechanism of multiple enzyme coordination by PCNA in Okazaki fragment maturation

ABSTRACT

PCNA is a central component of the DNA replication machinery in eukaryotes and archaea, acting as recruitment platform for numerous DNA replication factors. It has been proposed that PCNA can act as molecular toolbelt, simultaneously interacting with multiple proteins to coordinate their distinct enzyme activities. *S. solfataricus* PCNA is a heterotrimer, with each subunit displaying distinct specificity for binding PolB1, Fen1 or Lig1, raising the possibility that these enzymes may be coordinated using a toolbelt mechanism during Okazaki fragment maturation. By employing the *in vitro* Okazaki fragment maturation system developed in Chapter 3, the functional effect of preventing PolB1, Fen1 and Lig1 from simultaneously interacting with a single PCNA ring was determined. The resulting data support a toolbelt model whereby PolB1, Fen1 and Lig1 must be simultaneously bound to a single heterotrimeric PCNA ring for optimum efficiency of Okazaki fragment maturation.

INTRODUCTION

In addition to its well-documented role in acting as a processivity factor for DNA polymerases, PCNA has been shown to interact with a huge number of different proteins involved in various DNA transactions (Moldovan et al, 2007). PCNA effectively acts as a molecular platform, encircling duplex DNA and recruiting partner proteins to their sites of action on DNA through protein-protein interactions. Most of these partner proteins have been shown to interact with PCNA through a binding pocket located on the outer surface of the PCNA ring, and crucially, as PCNA is trimeric, each ring-shaped molecule presents three protein binding pockets (Gulbis et al, 1996). It has therefore been hypothesised that PCNA is capable of interacting

simultaneously with multiple protein factors, acting as a toolbelt molecule to coordinate different enzyme activities (Indiani et al, 2005).

There is now significant evidence to support the simultaneous binding of multiple protein factors to a single oligomeric DNA sliding clamp. Three, albeit identical, molecules of human Fen1 have been co-crystallised with a PCNA trimer, supporting the concept that sterically at least, multiple proteins can be simultaneously accommodated on a single PCNA ring (Sakurai et al, 2005). Similar results have been obtained for an archaeal RNaseHII bound to PCNA (Bubeck et al, 2011). Functional studies have provided evidence for two different *E. coli* DNA polymerases, Pol III and Pol IV, simultaneously binding the β clamp, in a manner that facilitates their exchange on DNA to bypass template lesions (Indiani et al, 2005; Heltzel et al, 2009). Consistent with this model, an engineered β clamp which possesses only a single functional binding pocket appears unable to support switching to Pol IV-mediated DNA synthesis *in vivo*, suggesting simultaneous rather than alternating binding of DNA polymerases to β is required for this process (Sutton et al, 2010). Simultaneous binding of multiple different proteins to homotrimeric human PCNA (Doré et al, 2009) and euryarchaeal PCNA (Kiyonari et al, 2009; Shi et al, 2012) has also been observed *in vitro*.

Alternatively, it has been argued that the three identical binding sites presented by a single homotrimeric PCNA can only support the competitive binding of different factors, thus disputing the toolbelt model (Moldovan et al, 2007). Indeed, it has been demonstrated that the binding of PCNA by Lig1 *in vivo* (Subramanian et al, 2005) or *in vitro* (Levin et al, 1997) is sufficient to prevent the interaction of other factors with PCNA. However, both these studies employed non-physiological, elevated concentrations of Lig1; conditions that would favour the saturation of PCNA binding sites. It cannot be ruled out that simultaneous binding of different factors to PCNA may be possible *in vivo* when local protein concentrations, and the pool of factors available to bind PCNA at a given time, may be more controlled.

Additionally, it has been proposed that the specific mode of interaction of some factors with PCNA may prevent the recruitment of additional proteins to the same PCNA ring, preventing the formation of a toolbelt. Small-angle x-ray scattering (SAXS) analysis of PCNA from the euryarchaeon *Pyrococcus abyssi* in complex with the DNA endonuclease NucS has revealed binding of a single NucS molecule to the trimeric ring, despite the availability of three identical binding sites (Creze et al, 2012). However, given that NucS is observed to project away from the PCNA ring, and thus does not physically obstruct the remaining two protein binding sites, the structural basis for this mechanism is unclear; PCNA has been shown by numerous structural studies to adopt an invariant conformation regardless of whether it is bound by partner proteins, and thus it appears unlikely that any form of allosteric communication between protein binding sites exists within PCNA. Nonetheless, if such a mechanism does exist, it is clearly not universal for all PCNA-binding proteins, as demonstrated by the binding of multiple Fen1 or RNaseHII molecules to PCNA, as described above.

In vitro reconstitution of Okazaki fragment maturation in Chapter 3 demonstrated an essential role for *S. solfataricus* PCNA, with each of the necessary enzymes – PolB1, Fen1 and Lig1 – requiring an interaction with PCNA to facilitate coupled processing activity. A similarly essential role for PCNA has been observed in eukaryotic Okazaki fragment maturation (Ayyagari et al, 2003). Each subunit of *S. solfataricus* PCNA displays a distinct specificity for binding a different maturation factor, and furthermore, pull-down data suggest that this heterotrimeric PCNA can simultaneously interact with PolB1, Fen1 and Lig1 (Dionne et al, 2003). It has been thus been proposed that *S. solfataricus* PCNA may coordinate enzyme activities during Okazaki fragment maturation by acting as a toolbelt, physically co-localising PolB1, Fen1 and Lig1. A functional requirement for such coordination has yet to be directly tested however.

In this chapter, the *in vitro* Okazaki fragment maturation system developed in Chapter 3 is employed to test the mechanistic basis of multiple enzyme stimulation by PCNA. By

employing novel PCNA mutants, maturation is shown to be less efficient when PolB1, Fen1 and Lig1 are prevented from binding the same PCNA ring. Furthermore, forcing DNA polymerase and Fen1 to compete for binding to the same, rather than different, PCNA subunits also decreases the efficiency of maturation. Thus a model is proposed whereby efficient Okazaki fragment maturation requires the simultaneous binding of each of the required enzymes to their respective PCNA subunits, with the heterotrimeric PCNA ring effectively organising PolB1, Fen1 and Lig1 into a processing complex on the lagging strand.

RESULTS

Okazaki fragment maturation is most efficient when PolB1, Fen1 and Lig1 are coordinated around a single PCNA molecule

S. solfataricus PolB1, Fen1 and Lig1 are all dependent on an interaction with PCNA to catalyse Okazaki fragment maturation *in vitro*. To determine if coordinated activity between these three factors requires an interaction with the same PCNA ring, the effect of separating the PolB1, Fen1 and Lig1 binding sites across different PCNA rings was tested. A fused PCNA heterotrimer (Dionne et al, 2008) was used to generate mutant rings in which the protein binding site in just one of the three subunits of PCNA was disrupted at a time. A conserved alanine residue lining the hydrophobic binding site of each PCNA subunit was replaced with a glutamic acid to sterically and electrostatically exclude PIP motif binding (Figure 4.1).

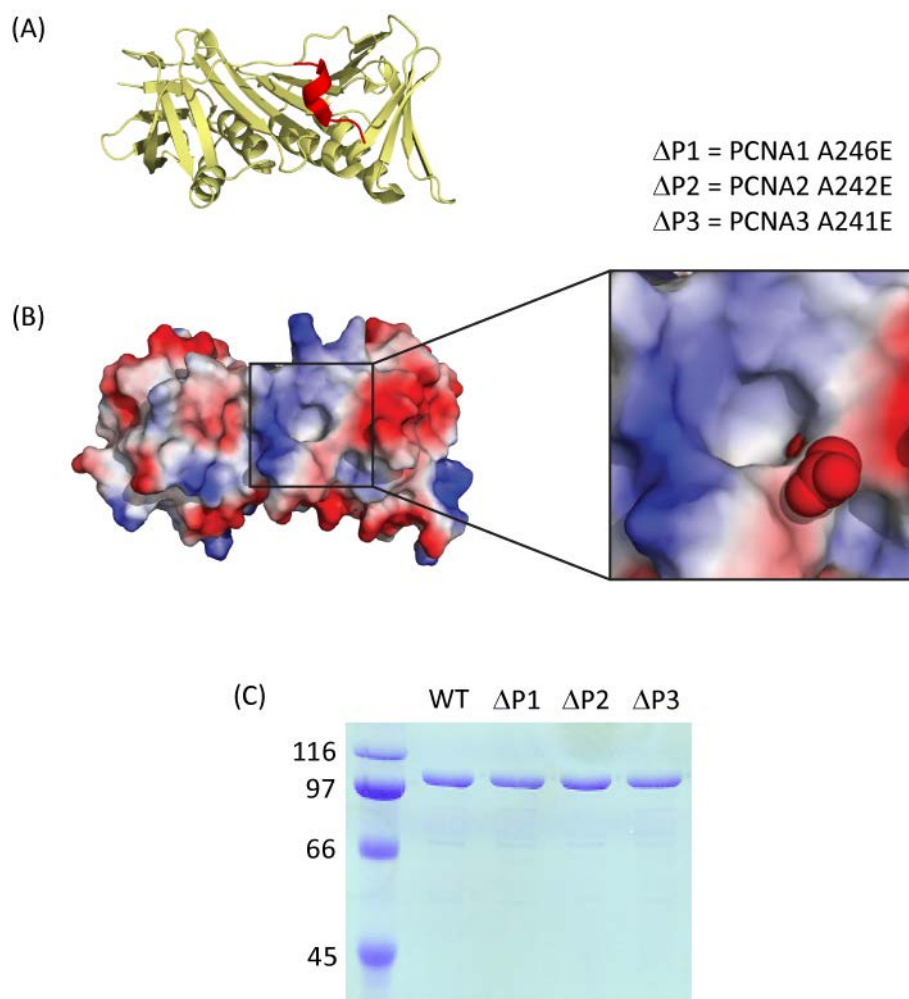


Figure 4.1 – The structural basis of PCNA mutagenesis to disrupt protein binding pockets

(A) Cartoon representation of PCNA1 (yellow) in complex with the PIP motif of Fen1 (red).

(B) Electrostatic surface view of the same PCNA1 molecule highlighting the boxed hydrophobic binding pocket. The Fen1 PIP motif is omitted for clarity. The magnified section models a mutation of alanine 246 to glutamate, rendered in space-filling mode. The corresponding mutations in the other subunits of PCNA are listed above. Each of these mutations was introduced individually into the covalent PCNA1-2-3 fusion. Images were generated from PDB file 2IZO.

(C) Purified recombinant PCNA fusion mutants analysed by SDS-PAGE and visualised by Coomassie staining. Molecular weight markers on the left, sizes are in kDa.

These mutants were validated by comparing their ability to stimulate PolB1, Fen1 or Lig1 individually on their respective optimal model substrates (Figure 4.2); the lack of activity observed in the presence of PCNA alone confirmed that the effect of PCNA on each enzyme was genuinely stimulatory, rather than PCNA possessing any intrinsic catalytic activity.

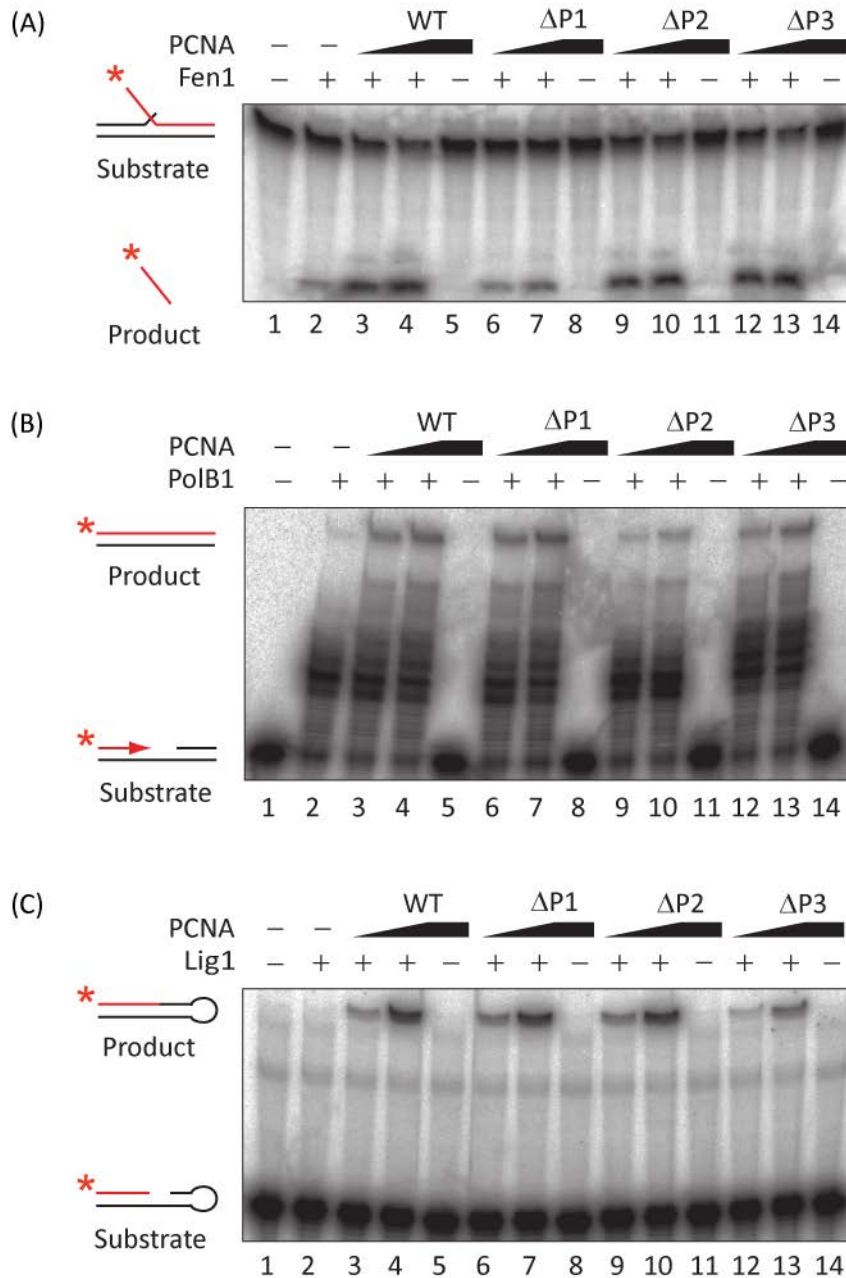


Figure 4.2 – The effect of PCNA mutants on stimulating the activities of PolB1, Fen1 and Lig1 *in vitro*.

(A) Stimulation of Fen1 activity on an optimal double-flap substrate *in vitro* by different PCNA mutants. Substrates and products are indicated on the left, with radiolabelled strands indicated in red and label position by an asterisk. Reactions contained 65 fmol Fen1 and 0, 0.77, 7.7 pmol of each PCNA protein. Reaction products were separated by denaturing PAGE.

(B) Stimulation of PolB1 activity on a model primer-template substrate *in vitro* by different PCNA mutants. Reactions contained 30 fmol PolB1 and 0, 0.77, 7.7 pmol of each PCNA protein. Assays were performed as described in Materials and Methods with the exception that reactions contained 100mM KCl and 4mM MnCl₂.

(C) Stimulation of Lig1 activity on a model nicked substrate by different PCNA mutants *in vitro*. Reactions contained 10 pmol Lig1 and 0, 0.77, 7.7 pmol of each PCNA protein. Reactions were performed as described in Materials and Methods with the exception that reactions were supplemented with 50mM KCl and were incubated for 15 minutes at 50°C.

Each PCNA mutant was defective in stimulating the intended enzyme activity ($\Delta P1 = \text{Fen1}$, $\Delta P2 = \text{PolB1}$, $\Delta P3 = \text{Lig1}$) but retained its ability to stimulate the remaining two enzymes, eliminating the possibility that the PCNA mutations had disrupted neighbouring binding sites or the overall ring architecture of the protein. This requirement restricted mutagenesis to relatively subtle substitutions, and therefore the impact of the mutated binding sites was relatively small. For example, mutation of the PIP motif in Lig1 completely abrogated PCNA-stimulated activity (Figure 3.9), while in contrast, a 4.5-fold reduction in Lig1 stimulation was exerted by the $\Delta P3$ mutant at the highest concentration of PCNA tested (Figure 4.2C, lanes 4 and 13). The effects of these PCNA mutations were nonetheless significant and reproducible.

The mutant PCNA proteins were substituted into the *in vitro* Okazaki fragment maturation assay developed in Chapter 3, and the extent of Okazaki fragment ligation was quantified (Figure 4.3). Substitution of each mutant into the system individually resulted in a significant reduction in maturation activity compared to WT PCNA, consistent with PolB1, Fen1 and Lig1 all requiring an interaction with their respective PCNA binding sites for activity. Crucially, when a mix of all three mutants was substituted into the system, such that the molarity of functional PCNA1, PCNA2 and PCNA3 binding sites was equal to WT but with all three sites never on the same PCNA molecule, maturation activity was not restored to WT levels. By using the fused heterotrimer, the possibility of subunit exchange to generate a fully WT heterotrimer upon mixing of mutants could be eliminated. Thus, the data presented here indicates a requirement for the three distinct enzymes PolB1, Fen1 and Lig1 to be coordinated around a single PCNA molecule for maximum efficiency of Okazaki fragment maturation.

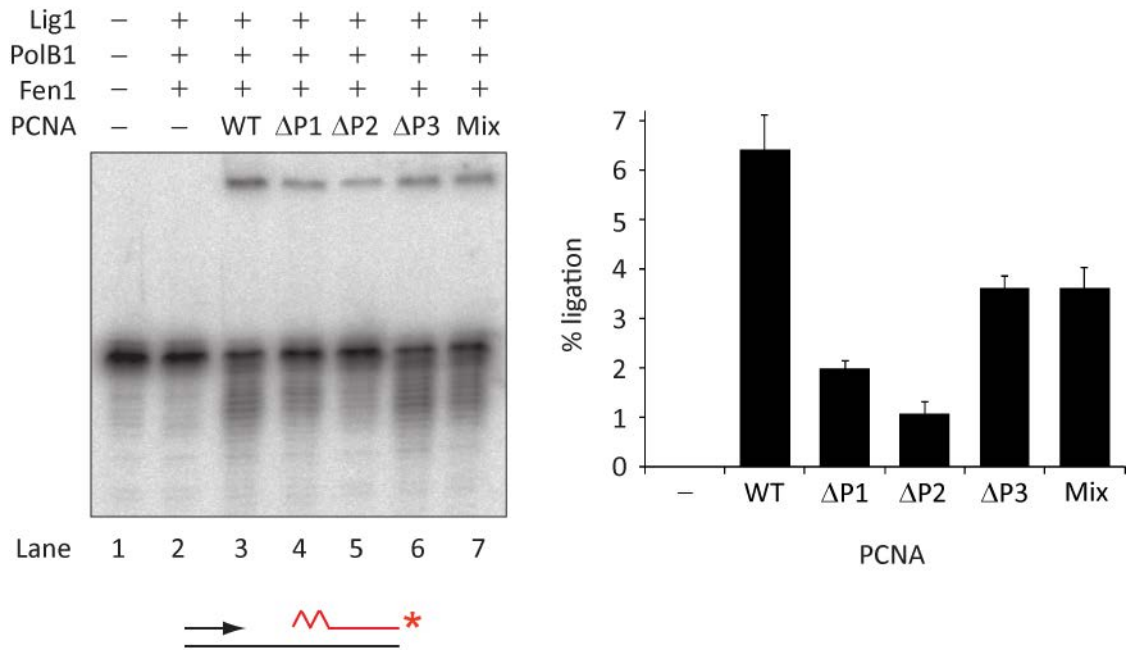


Figure 4.3 – The effect on Okazaki fragment maturation of preventing PolB1, Fen1 and Lig1 from binding the same PCNA ring

Reactions contained 0.1 pmol PolB1, 0.1 pmol Fen1, 1 pmol Lig1 and 20 pmol PCNA. Reaction products were separated by denaturing PAGE. Substrate is indicated below gel, with radiolabelled strand indicated in red and label position with an asterisk. The jagged region of the substrate denotes a 13 ribonucleotide primer. Quantification of the upper ligated band – a measure of completed Okazaki fragment maturation – is shown on the right. Values are mean $\pm$ SEM (n=3).

Efficient coordination of DNA polymerase and Fen1 activities requires simultaneous access to PCNA

Coordination around a single PCNA could be a dynamic process, involving the sequential access of PolB1, Fen1 and Lig1 to their respective binding sites, or alternatively could involve simultaneous binding of the three enzymes to PCNA. To distinguish between these mechanisms, the effect on RNA primer excision of forcing the required enzymatic activities to alternate in their access to PCNA was analysed. This was achieved by using the *S. solfataricus* translesion DNA polymerase Dpo4, which interacts with the same PCNA1 binding site as Fen1 (Dionne et al, 2008).

Recombinant Dpo4 was purified (Figure 4.4A) and substituted into the Okazaki fragment maturation assay in place of PolB1 (Figure 4.4B). Dpo4 possessed similar catalytic activity to PolB1 on the lagging strand substrate, performing PCNA-stimulated strand displacement synthesis (Figure 4.4B, lanes 2 and 3). Consistent with its known lack of 3'-5' exonuclease activity, Dpo4 was unable to undergo any significant idling at a nick (Figure 3.8) and therefore potentially employed a different mechanism to PolB1 to perform strand displacement synthesis. However, idling activity by PolB1 was very low, and together with the very similar pattern of strand displacement observed for both PolB1 and Dpo4 (compare Figure 4.4B, lane 3 with Figure 3.6A, lane 4) this suggests that both enzymes generated comparable displaced RNA flap structures under these experimental conditions.

Addition of Fen1 to Dpo4 failed to result in the coordinated enzyme activity observed with PolB1; downstream fragment cleavage was minimal (Figure 4.4B, lane 9, compare with Figure 3.6A, lane 11) while strand displacement by Dpo4 was actually decreased rather than increased in the presence of Fen1 (Figure 4.4B, lane 4, compare with Figure 3.6A, lane 5). Dpo4 did still exhibit some PCNA-stimulated strand displacement in the presence of Fen1, demonstrating that Dpo4 access to PCNA was decreased rather than totally blocked. Similarly, these effects were not observed when the Fen1 PIP mutant was substituted into the reaction, confirming that they were facilitated by the ability of WT Fen1 to access PCNA (Figure 4.4B, lanes 5 and 10). Therefore, despite Dpo4 and Fen1 both being able to access their shared binding site on PCNA and possessing the appropriate catalytic activities, alternating access to PCNA was not sufficient for effective coupling of their activities. This is in agreement with the data presented in Figure 3.7, which suggested a requirement for co-localisation of PolB1 and Fen1 on the same PCNA, to enable rapid handoff of flap substrates between the two enzymes.

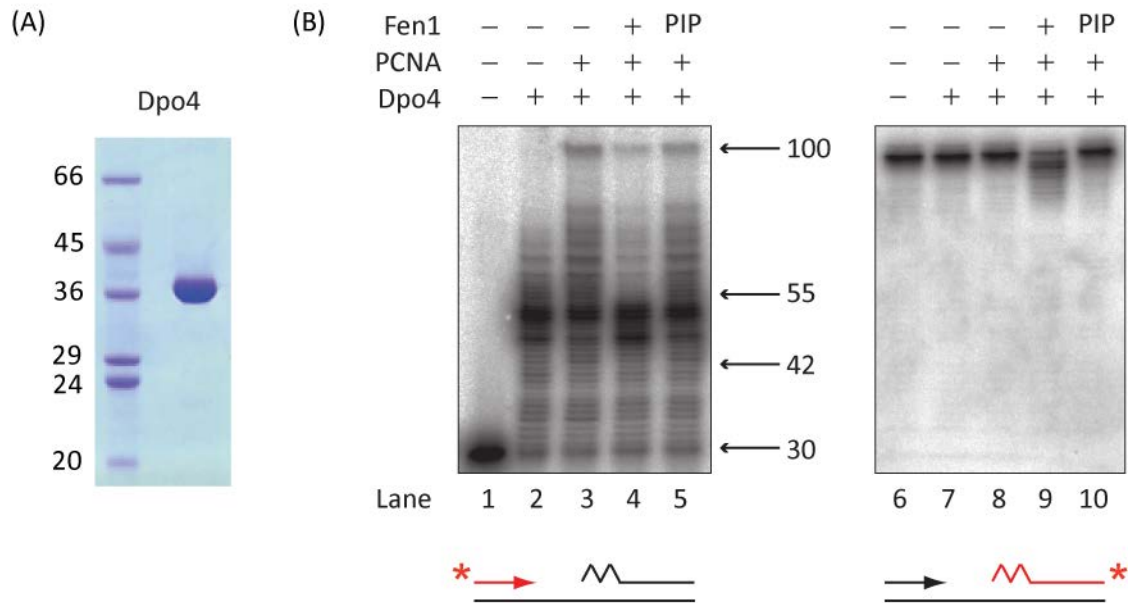


Figure 4.4 – The effect on primer excision of preventing DNA polymerase and Fen1 from simultaneously binding the same PCNA ring

(A) Analysis of purified Dpo4 by SDS-PAGE, visualised by Coomassie staining. Molecular weight markers are labelled on the left, sizes are in kDa.

(B) Activity of purified Dpo4, Fen1 and PCNA on an *in vitro* lagging strand substrate. Reactions contained 50 fmol Dpo4, 0.1 pmol Fen1 and 20 pmol PCNA. Reaction products were separated by denaturing PAGE. The migration of notable DNA synthesis products is indicated, sizes are in nucleotides. Substrates are indicated below gels, with radiolabelled strand indicated in red and label position with an asterisk. The jagged region of the substrate denotes a 13 ribonucleotide primer. 'PIP' denotes the substitution of a Fen1 mutated in the PIP motif that is defective in PCNA binding.

PCNA does not form a stable functional complex with PolB1, Fen1 and Lig1 on DNA

It was next investigated whether heterotrimeric PCNA genuinely assembles with all three partner proteins simultaneously into a stable functional complex on DNA. To test this, an assay was developed whereby PCNA, PolB1, Fen1 and Lig1 were allowed to assemble on an immobilised lagging strand substrate before the removal of unbound proteins by extensive washing. The composition of remaining stable complexes on DNA was then determined by their ability to perform complete Okazaki fragment maturation, which is dependent on the presence of all four proteins. To facilitate immobilisation and washing, the lagging strand

substrate used previously was first conjugated to magnetic beads using a biotin-streptavidin linkage; addition of the bead did not inhibit the ability of any proteins to access the substrate, as bulk Okazaki fragment maturation was unaffected (Figure 4.5A).

To validate the methodology of this assay, the formation of a PolB1-PCNA complex on DNA was first tested. To assemble PolB1 on DNA in a functional manner, PolB1 was incubated with the substrate at the reaction temperature of 50°C in the absence of dATP and dCTP; thus the polymerase was able to engage the 3' end of the primer and initiate extension but stalled upon encountering the first template deoxythymidine after four bases (Figure 4.5B, lanes 2 and 3). Maintaining the samples at 50°C, the substrate was then washed extensively to remove unbound protein and DNA synthesis was re-initiated by the addition of the remaining dNTPs. Under these conditions, significant full-length synthesis was observed in the presence of PolB1 and PCNA (Figure 4.5B, lane 5), demonstrating the ability of these proteins to form a stable functional complex on this primed substrate. The much reduced synthesis performed by PolB1 alone (Figure 4.5B, lane 4) is consistent with the known role of PCNA as a processivity factor and may reflect the ability of PolB1-PCNA to remain more stably bound to the substrate than PolB1 during washes and/or the ability of substrate-bound PolB1-PCNA to perform more extensive primer extension than PolB1 prior to dissociation from the substrate. When these proteins were incubated with beads alone and then washed prior to substrate addition and incubation, minimal DNA synthesis was observed (Figure 4.5B, lanes 6 and 7), confirming that the washes were effective at removing free protein and that the observed primer extension by PolB1 and PCNA was genuinely performed by a stable substrate-bound complex.

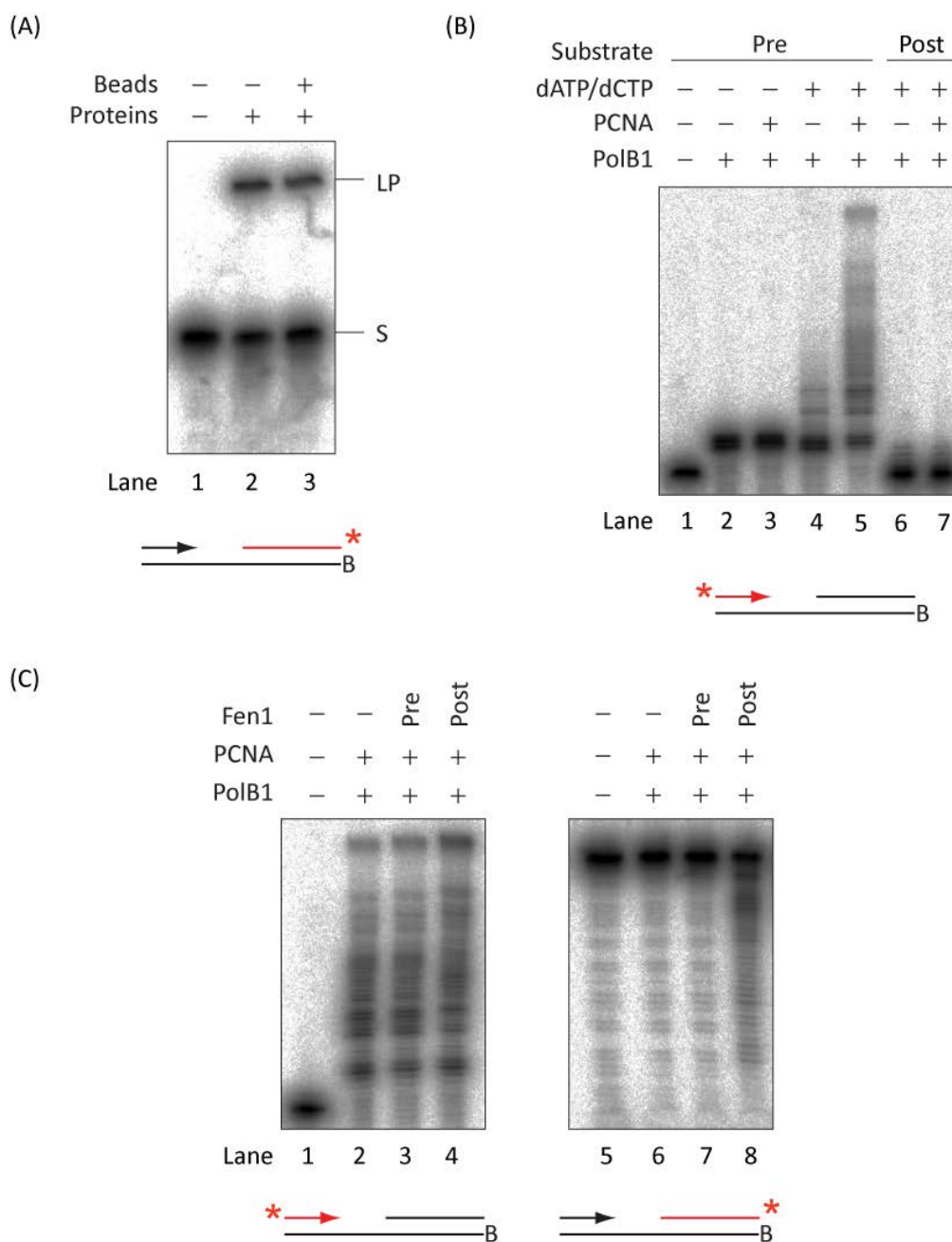


Figure 4.5 – Isolation of stable PCNA-protein complexes on DNA

(A) The effect on maturation activity of substrate conjugation to a magnetic bead. 'Proteins' indicates the addition of 0.1 pmol PolB1, 0.1 pmol Fen1, 1 pmol Lig1 and 20 pmol PCNA. Reaction products were separated by denaturing PAGE. The migration of substrate (S) and ligation product (LP) are indicated. Substrate is indicated below gel, with radiolabelled strand indicated in red, label position with an asterisk, and position of bead attachment with 'B'.

(B) Analysis of PolB1-PCNA complex formation on DNA. Reactions contained 0.1 pmol PolB1 and 20 pmol PCNA. 'Pre' and 'Post' indicate addition before or after bead washing, respectively.

(C) Analysis of PolB1-PCNA-Fen1 complex formation on DNA. Reactions contained 0.1 pmol PolB1, 0.1 pmol Fen1 and 20 pmol PCNA.

It was then tested whether Fen1 could also form a stable complex with PCNA, simultaneously with PolB1. Following the same protocol as above, no difference was observed in either primer extension activity or downstream Okazaki fragment cleavage when Fen1 was incubated with PolB1 and PCNA prior to washing and re-initiation of DNA synthesis (Figure 4.5C, lanes 3 and 7). In contrast, when Fen1 was added to the reaction after washing, the expected increase in primer extension and extensive downstream Okazaki fragment cleavage were observed (Figure 4.5C, lanes 4 and 8), confirming that the stable PolB1-PCNA complex assembled on DNA was competent to coordinate with Fen1. This therefore suggests that, under the conditions employed here, Fen1 did not form a stable complex with PCNA bound to substrate-engaged PolB1. However, given the stringent wash conditions employed here, it cannot be ruled out that a weaker, more transient complex of PCNA, PolB1, Fen1 and Lig1 can assemble on DNA, as suggested by the preceding functional data.

DISCUSSION

The data presented here, in conjunction with the previous chapter, support a model whereby the essential enzymes PolB1, Fen1 and Lig1 are coordinated around a single PCNA ring during Okazaki fragment maturation, with simultaneous binding to this ring supporting the most efficient coupling of their interdependent activities. Such coordination would tightly couple PolB1-mediated flap creation to Fen1-mediated flap cleavage. Furthermore, co-localisation of Lig1 would allow it to interrogate cleavage products generated by Fen1, facilitating ligation of Okazaki fragments immediately upon complete removal of initiating RNA; thus providing a mechanism by which covalent integrity of the genome is efficiently maintained (Figure 4.6).

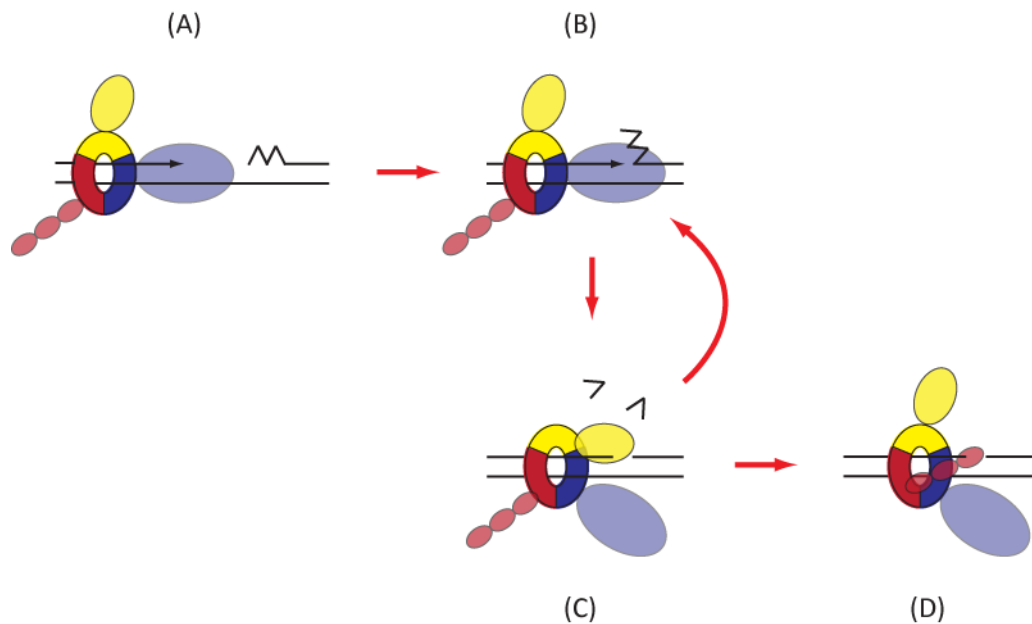


Figure 4.6 – Schematic illustration of Okazaki fragment maturation in *S. solfataricus*

(A) DNA is illustrated by black lines, RNA primer is illustrated by the jagged region. PolB1 (blue), Fen1 (yellow) and Lig (red) bind to their specific PCNA2, PCNA1 or PCNA3 binding sites on heterotrimeric PCNA (blue, yellow and red respectively). PolB1 engages the template to perform DNA synthesis while Fen1 and Lig1 are carried by PCNA.

(B) Upon encountering a downstream Okazaki fragment, PolB1 initiates strand displacement synthesis, displacing downstream RNA into a flap structure.

(C) Fen1 engages the generated flap and cleaves it, producing a nick. Multiple rounds of PCNA-coordinated PolB1 and Fen1 coupled activity enables complete removal of RNA primers.

(D) Lig1 interrogates nick structures emerging from Fen1, and following RNA removal encircles the DNA nick and catalyses Okazaki fragment ligation.

Although a pull-down assay developed to isolate stable functional complexes on DNA failed to demonstrate the presence of a PCNA-PolB1-Fen1 complex, it is unlikely that anything other than a highly stable complex would be identified under the stringent conditions employed. Indeed, under less stringent conditions, PCNA has been shown to act as a bridge between all three factors, Fen1, PolB1 and Lig1 (Dionne et al, 2003), supporting the model of coordination around a single PCNA proposed here. In addition to the core factors investigated here, it should also be noted that additional interactions have been reported between Dpo4 and PCNA1 (Dionne et al, 2008), the uracil DNA glycosylase UDG1 and PCNA3 (Dionne & Bell,

2005), the DNA repair factor XPF and PCNA1 and 3 (Roberts et al, 2003), the Holliday junction endonuclease Hjc and PCNA1 (Dorazi et al, 2006), and between the RFC complex and all three subunits of PCNA (Dionne et al, 2003). Therefore it would presumably be detrimental to cellular function for PCNA to form too stable a complex with just the Okazaki fragment maturation factors, preventing it from forming functional interactions with other key proteins. Indeed, artificially increasing the affinity of some partner proteins for PCNA in *S. cerevisiae* has been shown to severely impair cell growth (Fridman et al, 2010). The data presented here suggest that PCNA, PolB1, Fen1 and Lig1 form a weakly-associated complex which facilitates highly coordinated Okazaki fragment maturation while still allowing other factors to associate with PCNA when required, such as upon encountering DNA damage. More sensitive techniques may therefore be required to measure the formation of multi-protein PCNA complexes. In this respect it is interesting that preliminary studies using a recently developed single molecule FRET technique have observed the transient simultaneous association of two distinct euryarchaeal DNA polymerases with a single homotrimeric PCNA ring (Shi et al, 2012).

The heterotrimeric PCNA of crenarchaea has provided a unique opportunity to investigate the architectural basis of multi-enzyme coordination during lagging strand DNA replication. The orthologous factors to PolB1, Fen1 and Lig1 all interact with eukaryotic homotrimeric PCNA, and it is possible that they also may be coordinated around a single PCNA ring. Indeed, in addition to the single molecule observation described above, biochemical analysis supports the simultaneous binding of two different proteins to the homotrimeric human PCNA (Doré et al, 2009), while biophysical data has demonstrated simultaneous binding of distinct proteins to the homodimeric *E. coli* β clamp (Indiani et al, 2005) and the homotrimeric euryarchaeal PCNA (Kiyonari et al, 2009).

However, when the binding sites provided by the sliding clamp are identical, it is unclear what drives the formation of productive complexes of different enzyme. It is possible that appropriate toolbelt complexes are generated transiently, as multiple factors undergo

extensive exchange on the sliding clamp, generating different combinations of enzymes. Such exchange may be driven towards more productive combinations by mechanisms which restrict the proteins available to exchange on PCNA at given times. For example, phosphorylation and methylation have been observed to differentially affect the binding of Fen1 to human PCNA (Henneke et al, 2003; Guo et al, 2010), while phosphorylation of p21 (Scott et al, 2000) and the Pol32 subunit of Pol δ (Rahmeh et al, 2012) have also been shown to disrupt the interaction of these proteins with PCNA. Furthermore, conjugation of SUMO to K127 of PCNA can disrupt its interaction with Eco1, and possibly other partner proteins (Moldovan et al, 2006).

In contrast, crenarchaea, in which a relatively small number of PCNA-interacting proteins have been identified to date, may have evolved distinct PCNA subunits as the simplest mechanism to drive the formation of complexes composed of different enzymes. Notably, the current data cannot exclude an identical PolB1-Fen1-Lig1-PCNA complex forming on the leading strand during *S. solfataricus* chromosomal DNA replication. However, given the interaction of *S. solfataricus* PCNA with additional factors outside of those involved in lagging strand DNA replication, additional mechanisms may exist to determine which enzymes are available to interact with PCNA at different times, as with homo-oligomeric DNA sliding clamps.

One outstanding question from the model proposed here is how multiple large enzymes are sterically accommodated around a single PCNA ring. An increasing body of structural evidence has provided insight into how this may be achieved, and this is investigated in more detail in Chapter 5.

Chapter 5 – Conformational changes in Lig1 during Okazaki fragment maturation

ABSTRACT

Structural studies over recent years have demonstrated that many PCNA-bound factors are highly flexible upon PCNA. This property may be central to the ability of multiple different enzymes to undergo coordinated activity whilst simultaneously binding to a single PCNA ring. *Sulfolobus solfataricus* Lig1 is one such PCNA-binding protein that has been predicted from structural studies to undergo a large conformational change when engaging DNA to perform catalysis. To investigate how this conformational flexibility is correlated with the presence and activity of other PCNA-bound factors during Okazaki fragment maturation, a FRET assay was designed to measure intramolecular distances within Lig1, in solution. The substrate binding activity of Lig1 was characterised, and a fully-functional derivative of Lig1 containing two site-specifically introduced cysteine residues was generated. However, fluorophore conjugation at these sites to facilitate FRET measurements was found to disrupt the activity of Lig1, and thus FRET assays could not be undertaken. Nonetheless, the tools and reagents developed during this study will be useful in the future for determining the functional importance of Lig1 conformational flexibility.

INTRODUCTION

One key consideration of a toolbelt model, in which a single DNA sliding clamp simultaneously binds and organises multiple different enzymes, is how these different proteins are physically accommodated upon their shared protein platform. Numerous structural studies over recent years have begun to provide valuable insight into this question. The structure of three human Fen1 molecules bound to a single PCNA ring has been determined (Sakurai et al, 2005). Each

Fen1 molecule in this structure adopts a distinct conformation with respect to PCNA, and modelling suggests that only one of these conformations is competent to catalytically engage DNA passing through the central pore of PCNA (Doré et al, 2006). It has been proposed therefore that the remaining Fen1 conformations represent inactive or carrier configurations, whereby the enzyme can remain associated with PCNA until it is required, without interfering with the access of other PCNA-bound factors to DNA.

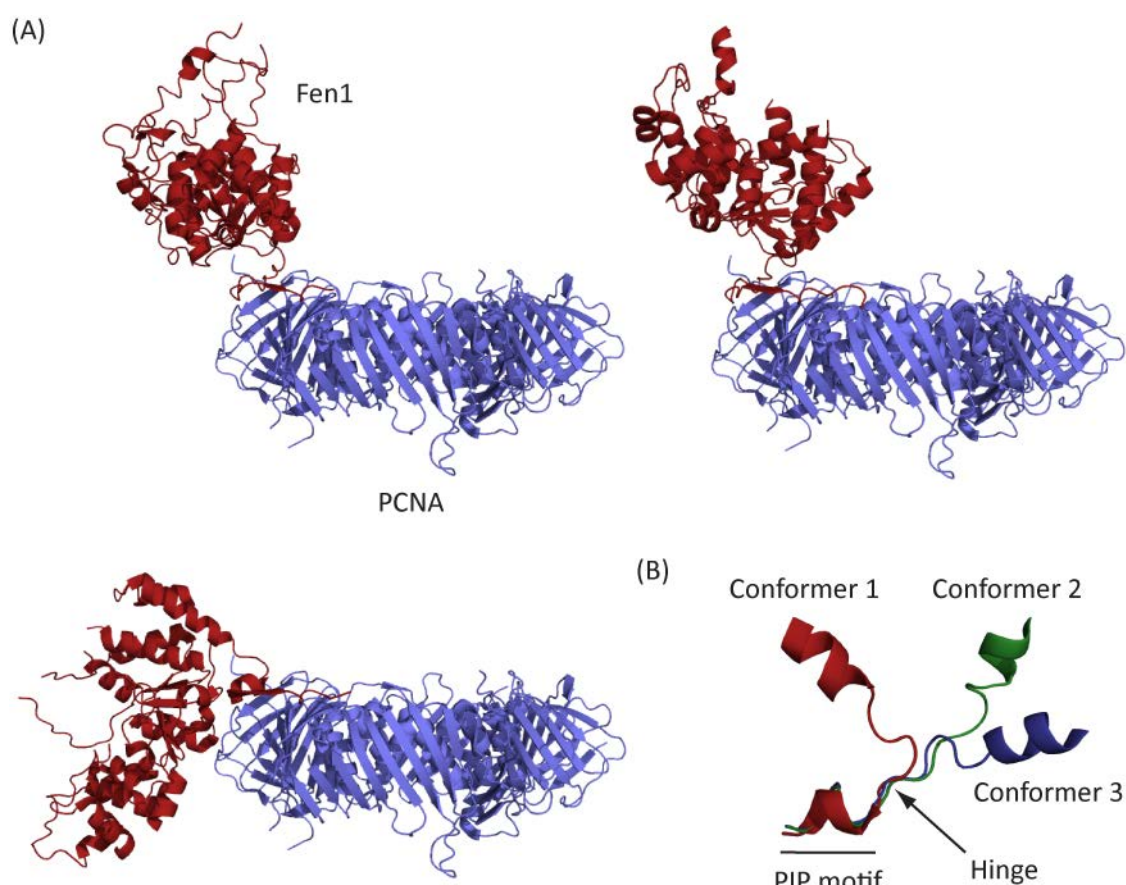


Figure 5.1 – Conformational flexibility of PCNA-bound Fen1

(A) Three conformations of human Fen1 (red) captured when crystallised bound to PCNA (blue). The PCNA ring is viewed from the side and only one Fen1 molecule is shown in each image for clarity; in the crystal three molecules of Fen1 are bound to a single PCNA

(B) Close-up of the hinge region of Fen1. The three Fen1 molecules were aligned by their C-terminal PIP motifs. Images were generated from PDB file 1UL1.

The ability of Fen1 to adopt different orientations with respect to PCNA appears to be largely mediated by a conserved flexible hinge region located between the anchoring PIP motif and the catalytic core of the protein (Figure 5.1); this hinge is important for PCNA stimulation of Fen1 catalysis *in vitro* (Sakurai et al, 2005). Furthermore, specific secondary contacts – distinct from the primary PIP interaction – appear to stabilise the different orientations of Fen1 on PCNA. This is consistent with biochemical data which has demonstrated that while the PIP interaction may be most important for the recruitment of Fen1 to PCNA from solution, additional interactions may be more important for supporting PCNA-stimulated activity on DNA (Gomes & Burgers, 2000). The recent crystallisation of three *Archaeoglobus fulgidus* RNaseHII molecules in complex with their cognate homotrimeric PCNA also demonstrated multiple configurations of RNaseHII with respect to PCNA, mediated by a flexible hinge motif, demonstrating conservation of these structural themes (Bubeck et al, 2011).

Further examples support the existence of inactive carrier configurations of PCNA-binding proteins. Structural data has shown the little-finger domain of the *E. coli* Y-family translesion DNA polymerase Pol IV to adopt a position at the side of the β sliding clamp (Bunting et al, 2003). This orientation is maintained by extensive secondary interactions distinct from the primary interaction of the clamp binding motif with a hydrophobic pocket on β , analogous to the PIP-PCNA interaction. Adoption of this position by the full-length Pol IV protein would prevent it from being able to engage DNA passing through the central pore of β . Biochemical data suggest that Pol IV can bind to the β sliding clamp simultaneously with the *E. coli* replicative polymerase Pol III, and this secondary interaction may be key in ensuring that while Pol IV is suitably localised to efficiently take over when lesions are encountered, it does not normally interfere with the processive activity of Pol III (Indiani et al, 2005; Wagner et al, 2009a). Similarly, the Y-family DNA polymerase from *S. solfataricus*, Dpo4, has also been shown to adopt an apparently inactive conformation orientated by interactions with the side of the PCNA ring (Xing et al, 2009). Again, relative to the structure of free Dpo4, the adoption

of this specific conformation in complex with PCNA appears to be mediated by the presence of a flexible hinge in proximity to the PIP motif of Dpo4.

Structural analysis of the B-type DNA polymerase (PolB) from *Pyrococcus furiosus* by crystallography and electron microscopy has demonstrated how conformational changes upon a PCNA platform can be directly linked to the catalytic activity of the protein. PolB has been observed to adopt different conformational states upon the PCNA ring to switch between engagement of DNA by its distinct polymerase and exonuclease active sites (Nishida et al, 2009; Mayanagi et al, 2011). Each conformation appears to be stabilised by a distinct set of secondary contacts between the outer surface of PCNA and the polymerase.

It is now clear that changes in the orientation of PCNA upon DNA may also play an important role in catalysis by PCNA-binding proteins. Significant variations in the tilt of DNA as it passes through PCNA have been observed in several ternary complexes of PCNA and partner proteins on DNA (Miyata et al, 2005; Mayanagi et al, 2009; Mayanagi et al, 2011), challenging the traditional view in which DNA passes perpendicularly through the central PCNA pore. This is further supported by crystal structures of both β and PCNA in complex with DNA, which demonstrate high degrees of tilting by the DNA duplex relative to the clamp pore (Georgescu et al, 2008; McNally et al, 2010). The combination of protein flexibility upon PCNA and PCNA flexibility upon DNA may therefore be an important factor in efficient catalysis. This is perhaps best illustrated by *P. furiosus* PolB, where molecular modelling suggests the conformational switching upon PCNA discussed above must be combined with variations in the tilt of DNA to enable the different catalytic sites of PolB to be engaged (Mayanagi et al, 2011). Comparative modelling of Fen1 bound to PCNA or the 9-1-1 sliding clamp also supports different Fen1 orientations on the sliding clamp, and different sliding clamp orientations on DNA, being employed to facilitate Fen1 catalysis in these two distinct settings (Querol-Audí et al, 2012).

One of the most striking proposed conformational changes in a PCNA-binding protein is that of Lig1. Human Lig1 bound to its nicked DNA substrate has been shown to encircle the

DNA duplex in a ring-shaped conformation that has been suggested to occlude the binding of other proteins to PCNA (Pascal et al, 2004). However, in the absence of a DNA substrate, the *S. solfataricus* orthologue of Lig1 adopts an elongated conformation, radiating outwards from the PCNA ring (Figure 5.2) (Pascal et al, 2006). Additionally, a low resolution electron microscopy structure of the *P. furiosus* Lig1-PCNA-DNA ternary complex has demonstrated that PCNA-bound Lig1 partially encircles its nick substrate, possibly representing an intermediate conformation prior to nick sealing (Mayanagi et al, 2009).

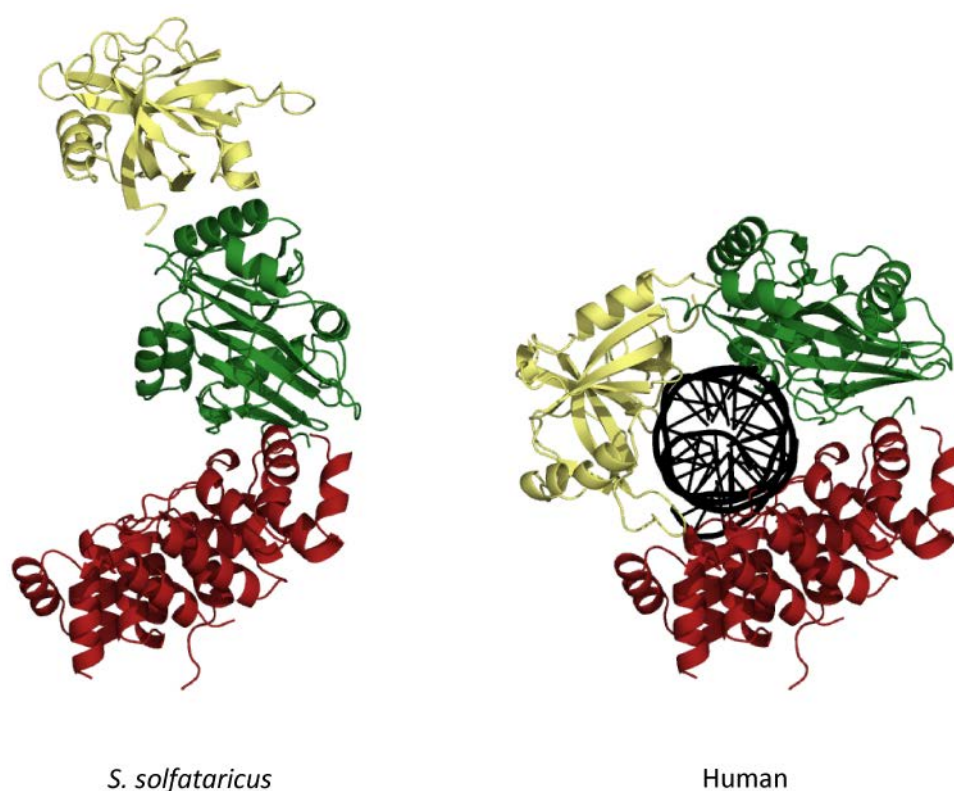


Figure 5.2 – Different conformational states of Lig1 orthologues

The structure of *S. solfataricus* Lig1 is shown on the left, and human Lig1 bound to DNA is shown on the right. The DNA-binding domain, adenylation domain, and OB-fold domain are shown in red, green and yellow respectively. DNA is shown in black. Images were generated from PDB files 2HIV and 1X9N.

It is tempting to hypothesise that conformational rearrangements of *S. solfataricus* Lig1 may be particularly pertinent to the efficient coordination of enzyme activities during

Okazaki fragment maturation. By adopting an elongated conformation, Lig1 could be carried by PCNA without interfering with the repetitive coordination of PolB1 and Fen1 required to remove RNA primers, which demands unimpeded access to the duplex passing through PCNA. Upon removal of RNA however, Lig1 would be perfectly localised to rapidly catalyse Okazaki fragment ligation, as is desirable for the maintenance of genome integrity. Given the observed flexibility of both PolB and Fen1 on PCNA, as discussed above, it can even be speculated that these proteins may remain bound to PCNA during Lig1 catalysis by adopting complementary orientations on PCNA which accommodate Lig1 as it encircles DNA.

Despite the host of structural data now available on protein-PCNA interactions, there is still little known about the dynamics of PCNA-binding proteins in solution, in particular the direct relevance of protein conformational changes upon PCNA to complex multi-enzyme processes, such as Okazaki fragment maturation. This chapter describes the initial stages in the development of an assay to investigate how the conformation of *S. solfataricus* Lig1 upon PCNA is correlated with the activity of other active enzymes simultaneously bound to PCNA, such as PolB1 and Fen1. Lig1 was characterised and engineered to enable measurement of changing intramolecular distances within the protein by an *in vitro* fluorescence resonance energy transfer (FRET) assay. While appropriate conditions to establish this assay were unable to be determined in this study, valuable techniques and reagents were developed that will be useful in future studies of Lig1 in the context of Okazaki fragment maturation.

RESULTS

Lig1 forms a stable complex with a nicked DNA substrate

As a starting point in understanding Lig1 conformational changes in solution, the large substrate-induced rearrangement of Lig1 predicted from crystal structures was initially investigated. The substrate binding activity of Lig1 was therefore first characterised in detail.

Purified recombinant *S. solfataricus* Lig1 was shown to possess DNA ligase activity on a model nicked DNA hairpin substrate *in vitro* (see Chapter 3). To facilitate sensitive binding measurements by fluorescence anisotropy, derivatives of this substrate were designed in which a Cy3B fluorophore was incorporated either internally or at the 3' end of the hairpin. To ensure incorporation of these fluorophores did not disrupt Lig1 access to the substrate, the ability of Lig1 to catalyse ligation of these substrates *in vitro* was measured (Figure 5.3A). Relative to unmodified substrate, Lig1 activity was impaired when Cy3B was incorporated internally, the fluorophore presumably sterically preventing Lig1 access to the nick. However, ligation activity was unimpaired when Cy3B was positioned at the end of the duplex, and therefore it was concluded that this moiety did not interfere with normal substrate binding by Lig1. This 3'-labelled substrate was used for subsequent analysis.

To further confirm the suitability of this substrate, the ability of PCNA to stimulate Lig1 activity on the substrate *in vitro* was measured (Figure 5.3B). As expected, the addition of PCNA to Lig1 resulted in a significant stimulation of Lig1 activity on an unlabelled substrate. Data elsewhere confirms that this was not due to an intrinsic catalytic activity possessed by PCNA (Figure 4.2). At the highest concentrations of PCNA, the amount of ligation reached saturation; the fact that this occurs at less than 100 % ligation was presumably due to a small amount of unannealed radiolabelled strand or incomplete 5' phosphorylation of the nick – a requirement for ligation. On the Cy3B-labelled substrate, PCNA stimulation of Lig1 is comparable, confirming that the presence of Cy3B on the substrate does not impair access of Lig1 or PCNA.

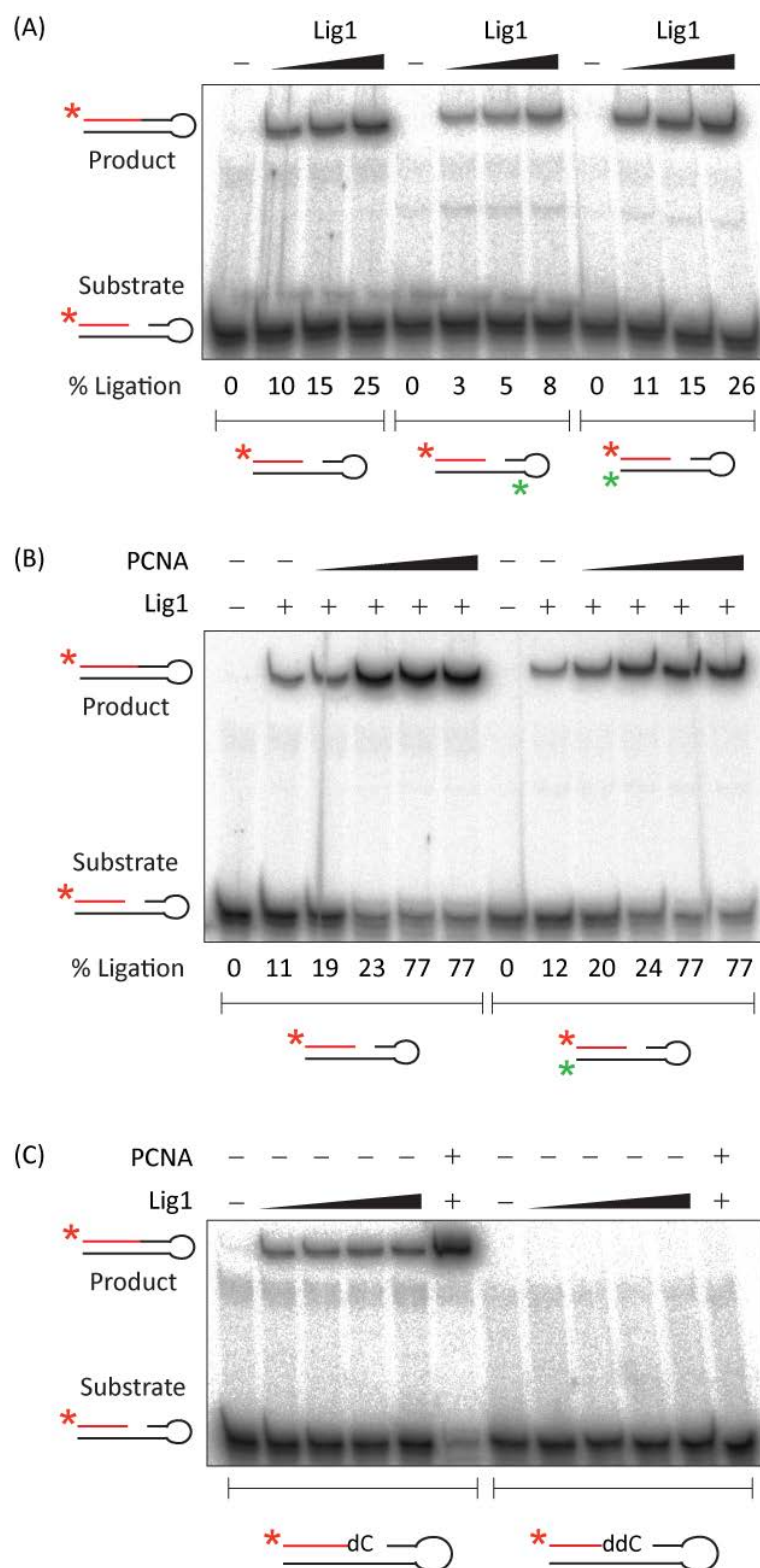


Figure 5.3 – Characterisation of substrates used for Lig1 DNA binding studies

(A) *In vitro* activity of Lig1 on fluorescently-labelled nicked DNA substrates. Reactions contained 0, 8.75, 17.5, 35 pmol Lig1. Reaction products were separated by denaturing PAGE, with migration of substrates and products indicated on the left. Quantification of products is shown underneath the gel. Different

substrates tested are shown underneath the gel, radiolabelled strands are highlighted in red with label position indicated by a red asterisk and sites of Cy3B incorporation indicated with a green asterisk.

(B) Stimulation of Lig1 *in vitro* activity by PCNA on a fluorescently-labelled nicked DNA substrate. Reactions contained 35 pmol Lig1 (where indicated) and 0, 18, 75, 300, 1200 fmol PCNA.

(C) Inhibition of Lig1 *in vitro* activity by incorporation of a dideoxynucleotide at a nick site. Reactions contained 0, 15, 30, 60, 120 pmol Lig1 and 1.5 pmol PCNA where indicated. Different substrates tested are shown underneath the gel; a 3' deoxycytosine (dC) or dideoxycytosine (ddC) at the nick is indicated.

To favour stable binding of Lig1 to its substrate rather than catalytic turnover and dissociation, the effect of creating a non-ligatable substrate was investigated. A previous study examining substrate binding by human Lig1 successfully utilised a dideoxy-nucleotide on the 3' side of a DNA nick to prevent ligation and stabilise substrate binding (Pascal et al, 2004); a 3' hydroxyl is required for phosphodiester bond formation by Lig1. The inability of *S. solfataricus* Lig1 to ligate a similar 3'-dideoxy-nucleotide-containing nick was confirmed, even in the stimulatory presence of PCNA (Figure 5.3C). Because this assay only measured final ligation, these data alone cannot rule out the possibility that substrate binding was disrupted by this substrate in addition to inhibition of the final catalytic step. However, previous biochemical data have demonstrated that the presence of a 3' hydroxyl is not required for nick substrate recognition by a viral ATP-dependent DNA ligase (Odell & Shuman, 1999). Furthermore, substrate binding was clearly maintained during co-crystallisation of human Lig1 with a similar substrate (Pascal et al, 2004).

Lig1 binding to a Cy3B-labelled non-ligatable substrate was subsequently measured by fluorescence anisotropy, and a dissociation constant of $15 \pm 1.4 \mu\text{M}$ was determined (Figure 5.4A). To confirm this result, substrate binding by Lig1 was also measured by EMSA (Figure 5.4B). Addition of Lig1 to a non-ligatable substrate resulted in the formation of a single protein-DNA complex, with the pattern of binding in agreement with a dissociation constant in the low micromolar range.

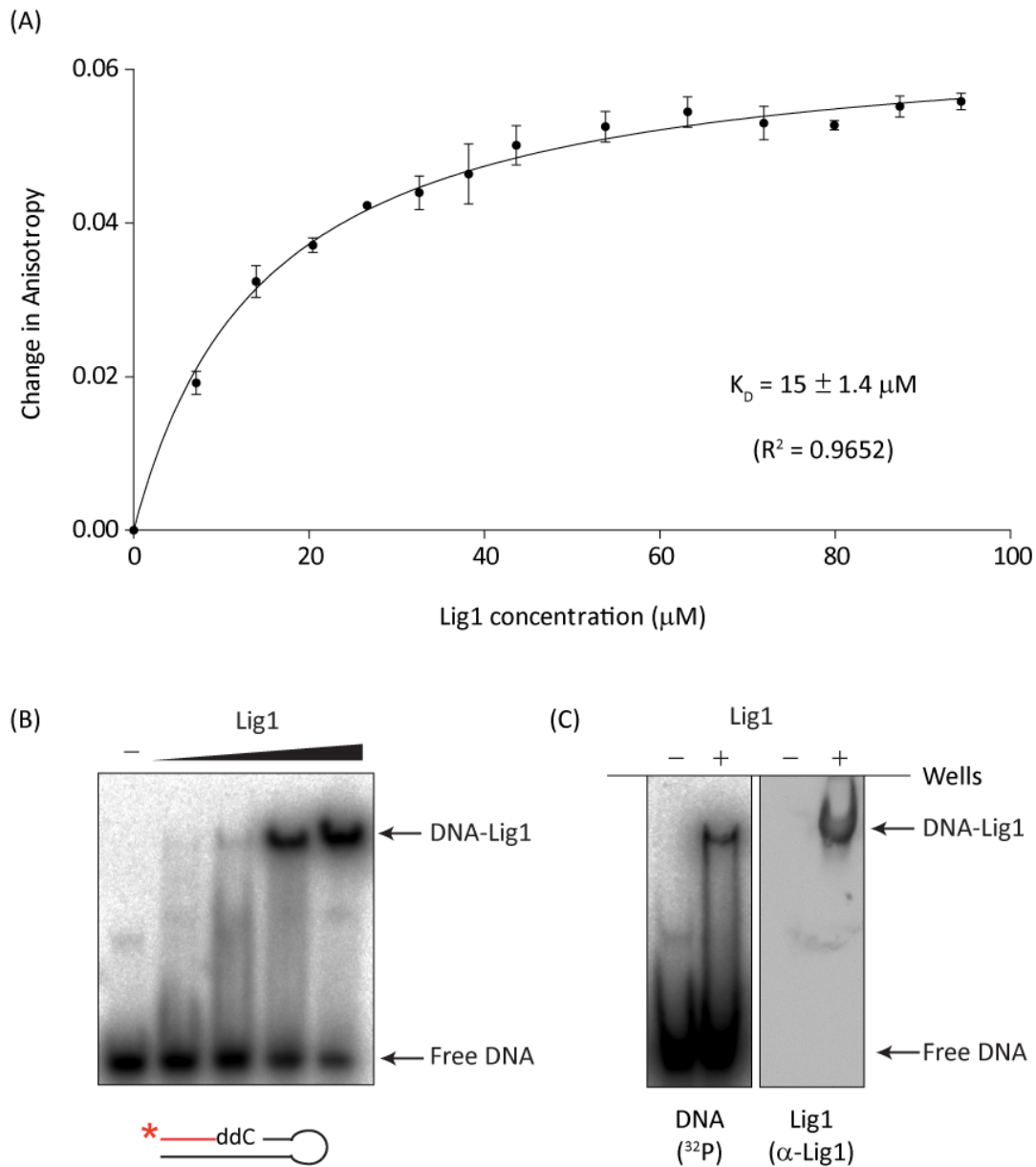


Figure 5.4 – Characterisation of the DNA binding activity of Lig1

(A) Fluorescence anisotropy measurements of Lig1 binding to a Cy3B-labelled non-ligatable DNA nick substrate. Values are mean $\pm$ SEM ($n=3$). Non-linear curve fitting was applied using GraphPad Prism software, assuming a single binding site model, and the derived dissociation constant (K_D) is presented.

(B) Analysis of Lig1 substrate binding by EMSA. Substrate was incubated with 0, 75, 150, 300, 600 pmol Lig1. Reaction products were separated by native gel electrophoresis, and migration of DNA and DNA-protein complexes is indicated on the right. Substrate is shown below gel, with the labelled strand highlighted in red and label position indicated with an asterisk.

(C) Western blotting analysis of DNA binding species. Substrate was incubated with 375 pmol Lig1. Reaction products were separated by native gel electrophoresis and DNA and protein components were detected by autoradiography and immunoblotting with Lig1 antiserum, respectively. Signals were aligned by the known position of sample wells, shown at the top.

Because relatively high protein concentrations were required to observe substrate binding by both fluorescence anisotropy and EMSA, it is possible that binding was mediated by a low abundance contaminating protein resulting from recombinant expression in *E. coli*. To exclude this possibility, both the DNA and protein components of EMSA reactions were simultaneously detected by autoradiography and immunoblotting, respectively (Demczuk et al, 1993). Overlaying the resulting signals demonstrated that Lig1 is indeed responsible for the substrate binding event observed (Figure 5.4C). Furthermore, because the substrate bound by Lig1 in EMSA assays was not modified by Cy3B, this confirms that the binding event measured by fluorescence anisotropy was genuinely due to a protein-DNA rather than aberrant protein-fluorophore interaction. Taken together these data demonstrate the characterisation and quantification of stable substrate binding by Lig1, to facilitate subsequent analysis of conformational changes in this protein.

Removal of three endogenous cysteine residues does not impair Lig1 catalytic activity

To analyse the conformation of Lig1 under different conditions, changing intramolecular distances between two different fluorophores, incorporated site-specifically into Lig1, were to be measured by FRET. One of the best available techniques for achieving site-specific protein labelling is the covalent modification of surface-exposed cysteine residues with thiol-reactive maleimide compounds (Kapanidis & Weiss, 2002). *S. solfataricus* Lig1 possesses three endogenous cysteine residues, which are tightly clustered in the oligonucleotide/oligosaccharide binding (OB)-fold domain (OBD). These were unsuitable for providing global distance measurements within Lig1 (Figure 5.5A). To enable site-specific modification at just two positions, a mutant Lig1, denoted -Cys, was first generated in which the three endogenous cysteine residues were removed by substitution with serine. Serine was chosen to minimise potential disruption to the structure of Lig1, as it possesses a side chain of broadly similar size to cysteine. The resulting protein behaved identically during purification to

wild-type (WT) Lig1, suggesting these mutations did not disrupt the overall structure of the protein (Figure 5.5B).

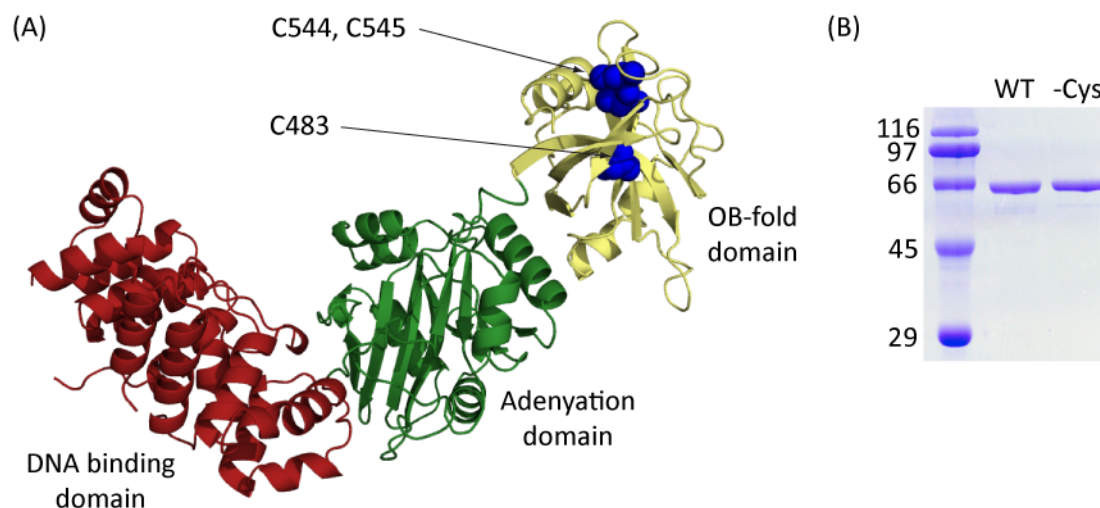


Figure 5.5 – Construction of a Lig1 mutant lacking endogenous cysteine residues

(A) *S. solfataricus* Lig1 structure highlighting the position of cysteine residues (blue) which were replaced with serine residues by site directed mutagenesis. Functional domains of Lig1 are labelled and coloured differently for clarity. Image was generated from PDB file 2HIV

(B) Purification of recombinant Lig1 lacking endogenous cysteines (-Cys) compared to wild-type (WT) Lig1. Proteins were analysed by SDS-PAGE and visualised by Coomassie staining. Molecular weight markers are on the left, sizes are in kDa.

Furthermore, -Cys was found to retain both its basal ligase activity (Figure 5.6A) and stimulation of its activity by PCNA *in vitro* (Figure 5.6B), relative to WT Lig1. It was therefore concluded that the endogenous cysteine residues of Lig1 are not essential for protein function and that -Cys is a suitable template for the reintroduction of cysteine residues in more favourable positions.

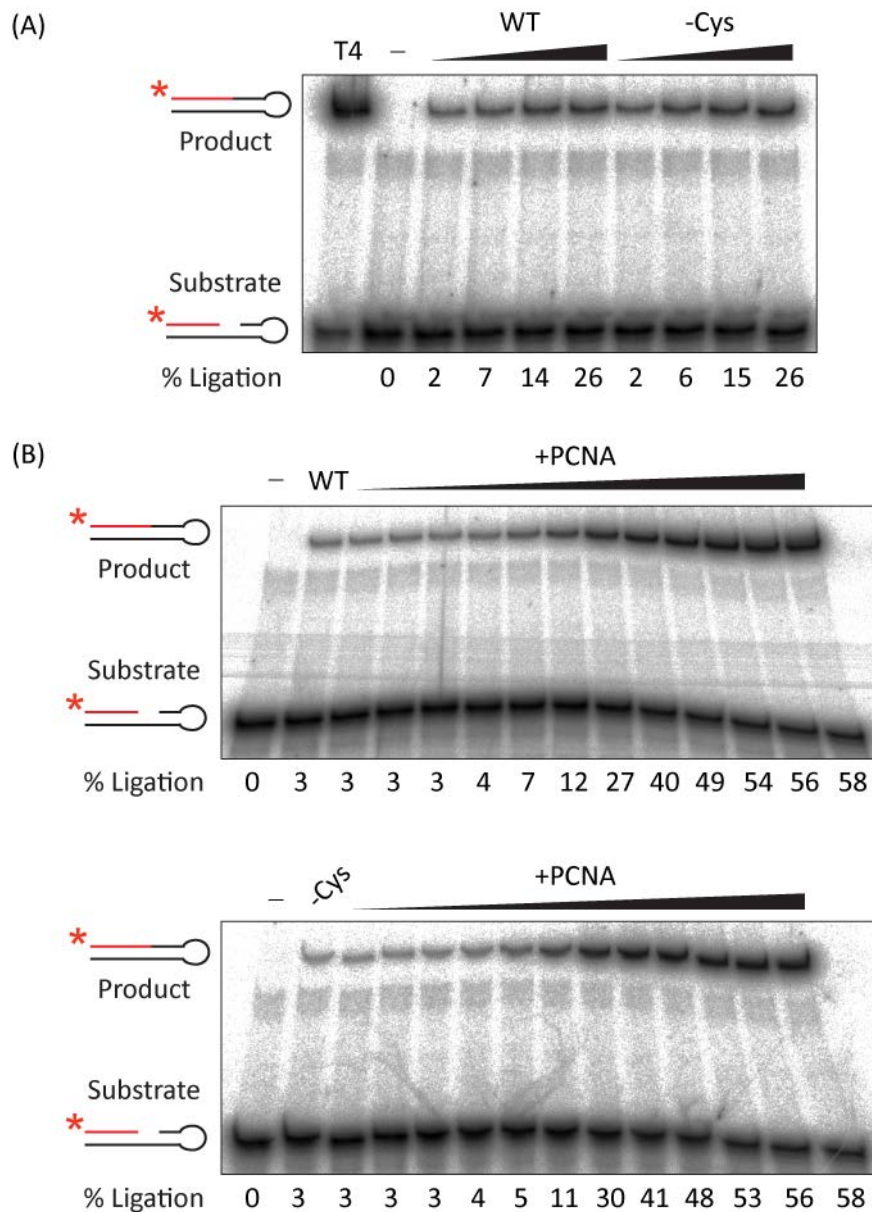


Figure 5.6 – *In vitro* activity of a Lig1 mutant lacking endogenous cysteine residues

(A) Comparison of the *in vitro* activity of WT Lig1 and Lig1 lacking endogenous cysteine residues (-Cys). Reactions contained 0, 6.25, 12.5, 25, 50 pmol indicated protein or 1 unit T4 DNA ligase (NEB) as a positive control. Reaction products were separated by denaturing PAGE, and migration of substrate and product is shown on the left; the labelled strand is highlighted in red and label position indicated with an asterisk. Quantification of products is shown underneath the gel.

(B) Stimulation of WT or -Cys *in vitro* ligase activity by PCNA. Reactions contained 12.4 pmol indicated Lig1 protein and 0, 3, 6, 12, 23, 47, 93, 187, 375, 750, 1500, 3000, 6000 fmol PCNA.

Lig1 containing two site-specifically introduced cysteine residues is fully functional

Based on large scale rearrangements suggested by structural studies (Pascal et al, 2004; Pascal et al, 2006), it was predicted that significant substrate-induced distance changes should be observable between a cysteine residue positioned in the OBD of Lig1 and a second cysteine in the DNA binding domain (DBD). To facilitate cysteine engineering, multiple sequence alignment of Lig1 orthologues was performed and the degree of residue conservation was subsequently mapped onto the known structure of *S. solfataricus* Lig1. This enabled selection of residues that are not only dispensable for protein function, as suggested by low conservation, but are also surface exposed, to enable fluorophore access (Figure 5.7).

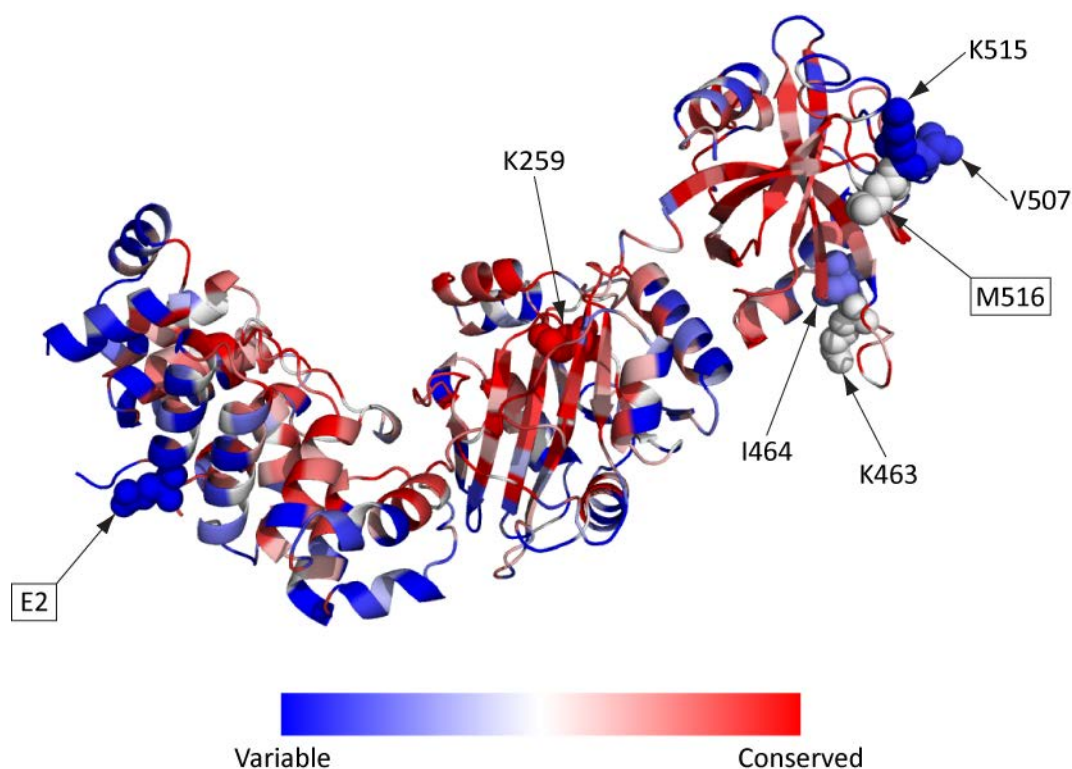


Figure 5.7 – Position and conservation of residues in Lig1 chosen for mutation to cysteine.

The degree of conservation of every Lig1 residue was mapped onto the Lig1 structure (PDB file 2HIV) using the Consurf server (<http://consurf.tau.ac.il>) and coloured according to the key underneath the structure. Residues chosen for mutagenesis are labelled and rendered in space-filling mode, successful mutants are boxed. The essential catalytic lysine, 259, is shown for reference as a highly conserved residue.

Four residues matching these criteria in the OBD (I464, K463, V507, K515) were individually mutated to cysteine, the resulting recombinant proteins were purified and their *in vitro* ligase activity characterised (Figure 5.8). All four of these mutants displayed deficient ligase activity compared to WT Lig1, suggesting that incorporation of cysteine at any of these positions disrupted Lig1 structure or function.

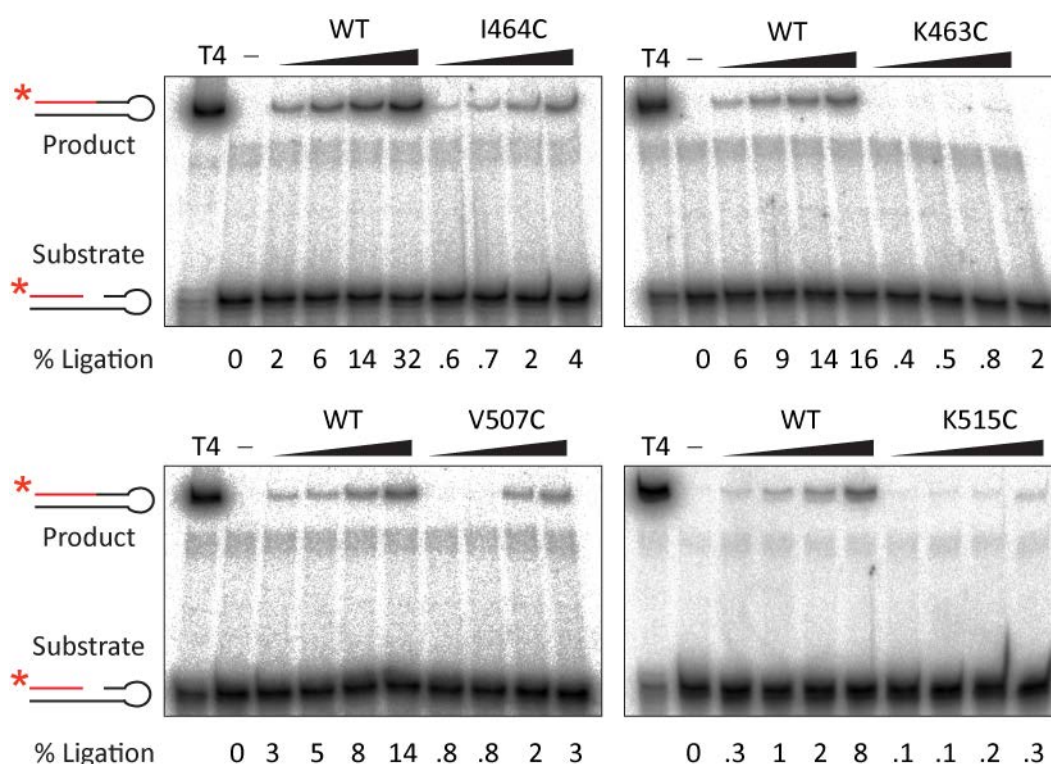


Figure 5.8 – Characterisation of *in vitro* activity of Lig1 mutants possessing single cysteine residues

Reaction products were separated by denaturing PAGE and migration of substrate and product is shown on the left; the labelled strand is highlighted in red and label position indicated with an asterisk. Quantification of products is shown beneath the gel. Reactions contained the following amounts of indicated Lig1 proteins: WT and I464C – 0, 8.75, 17.5, 35, 70 pmol. WT and K463C – 0, 8.75, 17.5, 35, 70 pmol. WT and V507C – 0, 16.9, 34, 68, 136 pmol. WT and K515C – 0, 1.12, 2.25, 4.5, 9 pmol. 1 unit of T4 DNA ligase was included in all experiments as a positive control.

Substitution of the weakly conserved methionine 516 with cysteine however, resulted in a protein that behaved identically during purification to WT Lig1 (Figure 5.9A) and retained comparable *in vitro* ligase activity to WT Lig1 (Figure 5.9B). Furthermore, PCNA stimulation of

M516C ligase activity was identical to WT Lig1 (Figure 5.9C). Therefore, cysteine incorporation at position 516 in the OBD resulted in a fully functional protein *in vitro*.

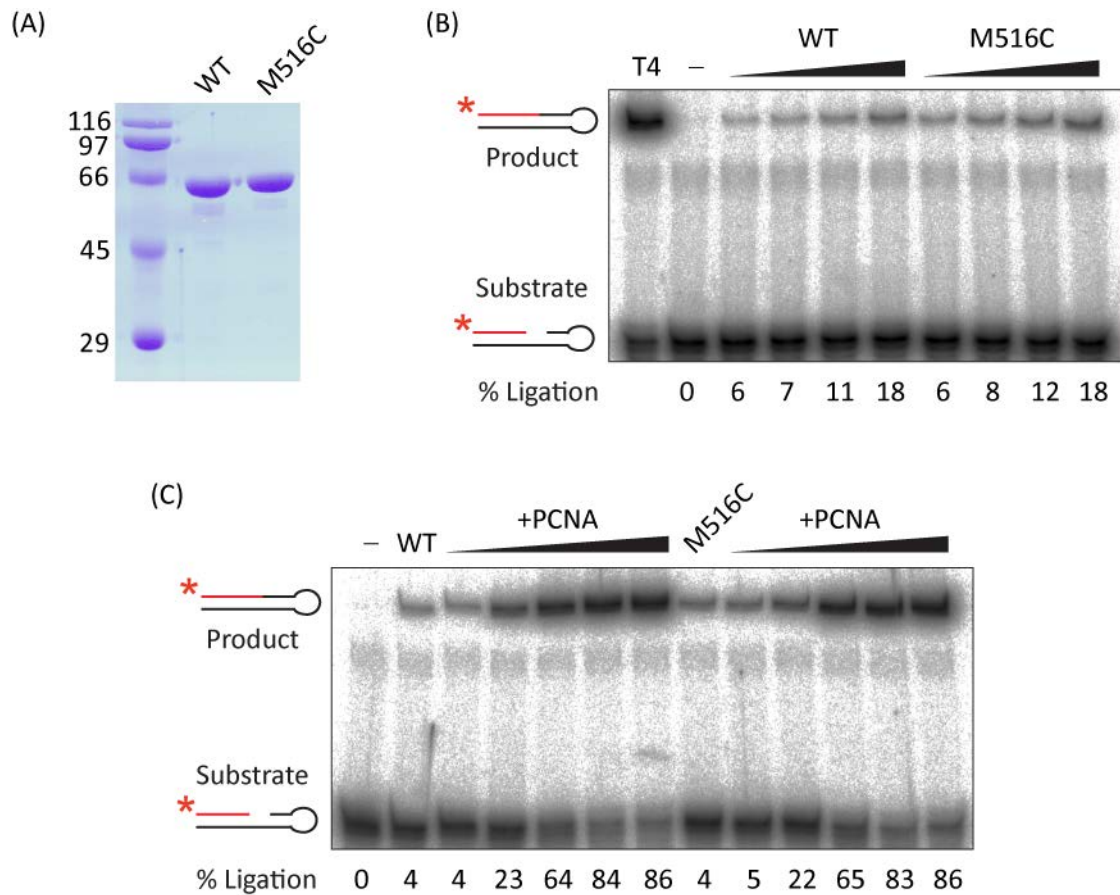


Figure 5.9 – Characterisation of a Lig1 mutant possessing a single cysteine at position 516.

(A) Purification of recombinant Lig1 containing a cysteine residue substituted at position 516 (M516C), compared to WT Lig1. Proteins were analysed by SDS-PAGE and visualised by Coomassie staining. Molecular markers are shown on the left, sizes are in kDa.

(B) Comparison of WT and M516C *in vitro* ligase activity. Reactions contained 0, 13.5, 27, 54, 108 pmol indicated protein. 1 unit T4 DNA ligase was included as a positive control. Reaction products were separated by denaturing PAGE, and migration of substrate and product is shown on the left; the labelled strand is highlighted in red and label position indicated with an asterisk. Quantification of product is shown beneath the gel.

(C) Stimulation of WT and M516C *in vitro* ligase activity by PCNA. Reactions contained 27 pmol Lig1 and 0, 23, 94, 375, 1500, 6000 fmol PCNA.

This functional M516C mutant was subsequently further modified by the substitution of a poorly conserved, solvent exposed glutamate at position 2 in the DBD with a second cysteine residue (E2C.M516C).

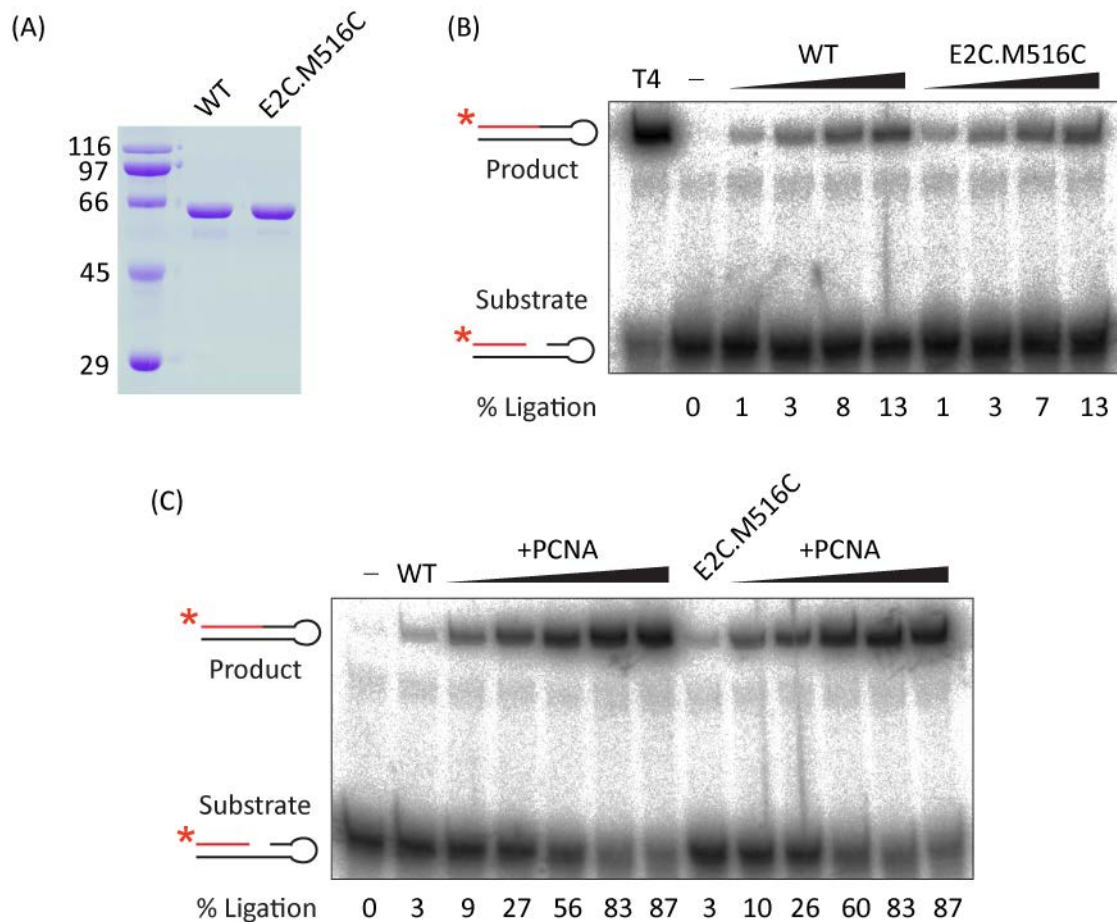


Figure 5.10 – Characterisation of a Lig1 mutant possessing cysteine residues at positions 2 and 516.

(A) Purification of recombinant Lig1 containing cysteine residues substituted at positions 2 and 516 (E2C.M516C) compared to WT Lig1. Proteins were analysed by SDS-PAGE and visualised by Coomassie staining. Molecular markers are shown on the left, sizes are in kDa.

(B) Comparison of WT and E2C.M516C *in vitro* ligase activity. Reactions contained 0, 1.37, 2.75, 5.5, 11 pmol indicated protein. 1 unit T4 DNA ligase was included as a positive control. Reaction products were separated by denaturing PAGE and migration of substrate and product is shown on the left; labelled strand is highlighted in red and label position indicated by an asterisk. Quantification of product is shown beneath the gel.

(C) Stimulation of E2C.M516C *in vitro* ligase activity by PCNA, compared to WT Lig1. Reactions contained 4.5 pmol Lig1 and 0, 23, 94, 375, 1500, 6000 fmol PCNA.

E2C.M516C behaved identically during purification to WT Lig1 (Figure 5.10A), and retained both WT *in vitro* ligase activity (Figure 5.10B) and PCNA stimulation of ligase activity (Figure 5.10C). This Lig1 mutant modified by the incorporation of two cysteine residues at distinct positions was thus concluded to be fully functional *in vitro*. Using available Lig1 structures, the distance between these two engineered cysteine residues was predicted to be 89 Å in the absence of DNA and 42 Å in the DNA-bound form of the protein; a difference that would be readily detectable by FRET.

Both engineered cysteine residues in Lig1 are accessible for modification by fluorophores

To confirm that the introduced cysteine residues were accessible to the solvent for modification, as predicted from the Lig1 structure, E2C.M516C was incubated with the sulfhydryl-reactive Ellman's reagent and the extent of cysteine modification was quantified from a resulting colour change. A molar ratio of two sulfhydryl groups per protein was reproducibly obtained, suggesting that both cysteine residues are indeed readily accessible. No reaction of Ellman's reagent with the -Cys mutant was observed, confirming the specificity of this reaction.

To facilitate FRET studies, a donor fluorophore – Cy3B – and an acceptor fluorophore – Alexa647 – were chosen with suitable overlapping emission and excitation spectra. To test if these fluorophores could access and react with both cysteine residues in E2C.M516C, variants of this protein possessing each of the two cysteine residues individually (E2C and M516C) were incubated with maleimide derivatives of Cy3B and Alexa647. After the removal of unreacted fluorophore, protein samples were quantified by UV-visible spectroscopy. In addition to protein absorbance at 280 nm, both E2C and M516C possessed significant absorbance at the wavelengths of Cy3B (559 nm) or Alexa647 (650 nm) when reacted with the relevant fluorophore (Figure 5.11A and B).

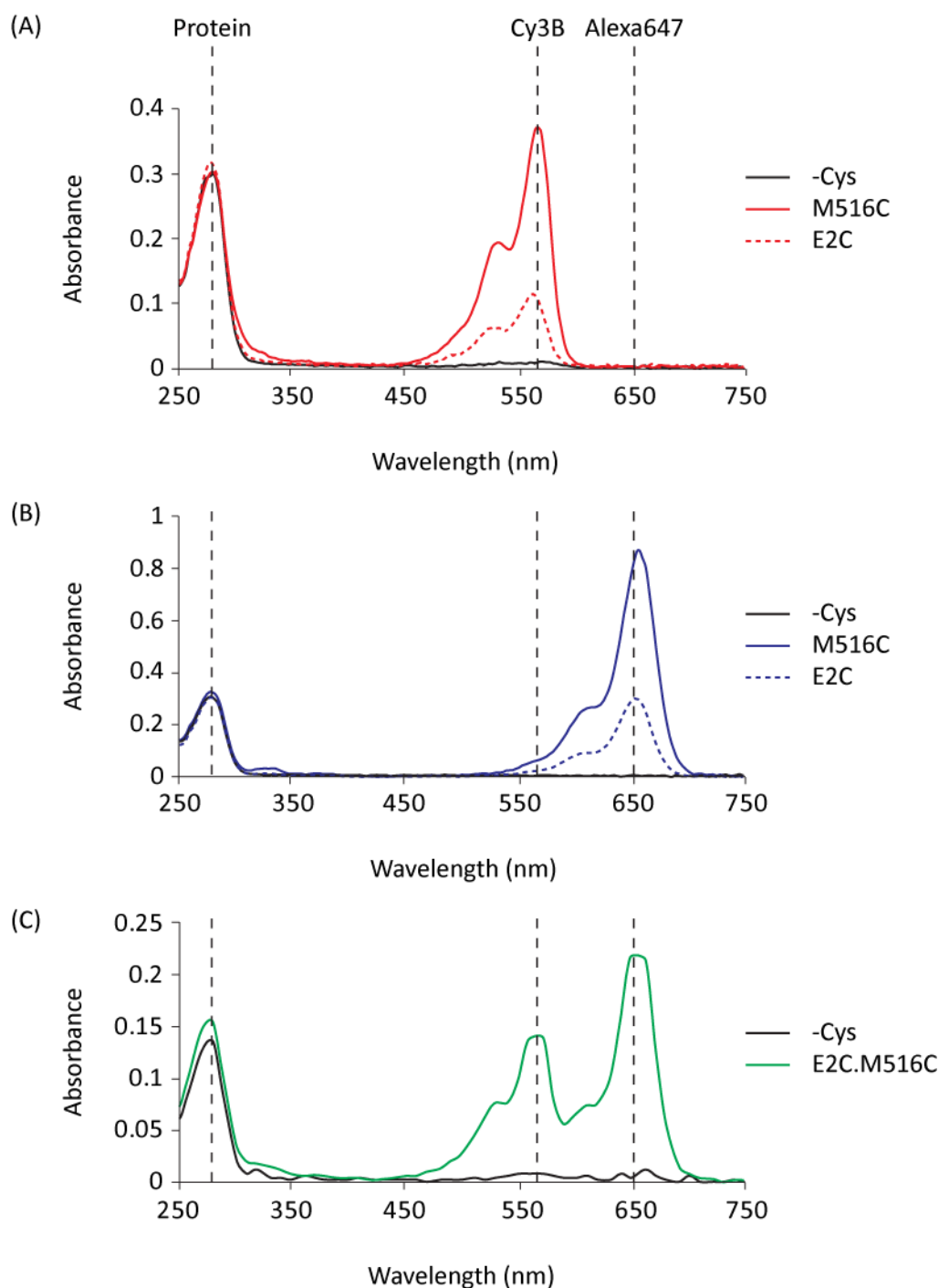


Figure 5.11 – Site-specific covalent modification of Lig1 with fluorophores

(A) Modification of Lig1 mutants by Cy3B. The indicated proteins were reacted with Cy3B-maleimide and the absorbance of reaction products was measured across the UV-visible range. The maximum absorbance wavelengths for protein (280 nm), Cy3B (559 nm) and Alexa647 (650 nm) are indicated by dashed lines.

(B) Modification of Lig1 mutants by Alexa647. Samples were analysed as in (A)

(C) Simultaneous modification of Lig1 mutants by both Cy3B and Alexa647. Samples were analysed as in (A)

Both cysteine residues were therefore reactive towards both fluorophores. Reaction of the -Cys protein with either fluorophore, performed in parallel, did not result in any detectable absorbance by Cy3B or Alexa647, demonstrating that the reactivity of these fluorophores was highly specific for cysteine residues, and that non-covalently associated fluorophore was effectively removed by the protocol employed here. Notably, the cysteine at position 2 showed less reactivity towards both fluorophores than that at position 516; this is presumably caused by decreased accessibility of the fluorophores due to steric factors, or a less compatible local chemical environment.

To generate Lig1 labelled with two different fluorophores for FRET assays, E2C.M516C was reacted simultaneously with both Cy3B and Alexa647 maleimide derivatives, and analysed by UV-visible spectroscopy as before (Figure 5.11C). Modification of E2C.M516C by both Cy3B and Alexa647 was detected, ruling out potential competition between the fluorophores. Preliminary analysis of this sample was performed by alternating laser excitation spectroscopy (ALEX), a technique that allows the determination of donor-acceptor stoichiometry as well as FRET values for single molecules (Kapanidis et al, 2005). This analysis suggested that the majority of E2C.M516C molecules were labelled by a single fluorophore – either at a single position or both positions – with only a very small population of molecules covalently modified by both Cy3B and Alexa647 (Johannes Holbein, personal communication). Therefore, despite both cysteine residues being accessible and reactive to Cy3B and Alexa647, additional procedures were required to increase the amount of dual-labelled Lig1, to improve the sensitivity of FRET assays.

Lig1 dual-labelled with Cy3B and Alexa647 can be purified using ion-exchange chromatography

Labelling reactions appeared to result in a mixed population of unlabelled, single-labelled and various dual-labelled Lig1 species. To isolate dual-labelled species which specifically possessed

one molecule of each fluorophore, the possibility of separating different labelled species by standard chromatography techniques was investigated. Alexa647 is negatively charged, due to the presence of multiple sulphonate groups, and it was therefore tested whether the presence of this fluorophore would affect the retention of E2C.M516C on a positively charged MonoQ column. In the absence of any fluorophore, E2C.M516C bound to this column and was eluted with a linear salt gradient at approximately 280 mM NaCl (Figure 5.12). As E2C.M516C was reacted with increasing concentrations of Alexa647, two additional species were detected with longer retention times. Quantification by UV-visible spectroscopy demonstrated that these correlated with discrete labelled species: E2C.M516C labelled with a single molecule of Alexa647, which eluted at approximately 320 mM NaCl, and E2C.M516C labelled with two molecules of Alexa647, which eluted at 350 mM NaCl. Under identical conditions, the elution of Cy3B-labelled E2C.M516C was unchanged, consistent with Cy3B lacking a net charge. Ion-exchange chromatography can thus be employed as a tool to separate different Alexa647-labelled species of Lig1.

To purify Lig1 dual-labelled with both fluorophores, E2C.M516C was simultaneously reacted with both Alexa647 and a slight excess of Cy3B, and protein labelled with a single Alexa647 molecule was subsequently isolated by MonoQ chromatography as above (Figure 5.13A). Analysis of collected fractions by UV-visible spectroscopy confirmed that every Lig1 molecule was covalently modified with a single Alexa647, while approximately 54 % of these molecules were additionally labelled with Cy3B (Figure 5.13B).

Taking their different extinction coefficients into account, the reactivity of both fluorophores was very similar towards a given cysteine residue (Figure 5.11), and thus it is likely that the observed dual labelling was largely stochastic, with both possible arrangements of the two fluorophores likely to exist in the purified dual-labelled population. Although not every Lig1 molecule was dual-labelled, this represented a significant enrichment and was considered satisfactory for subsequent analysis.

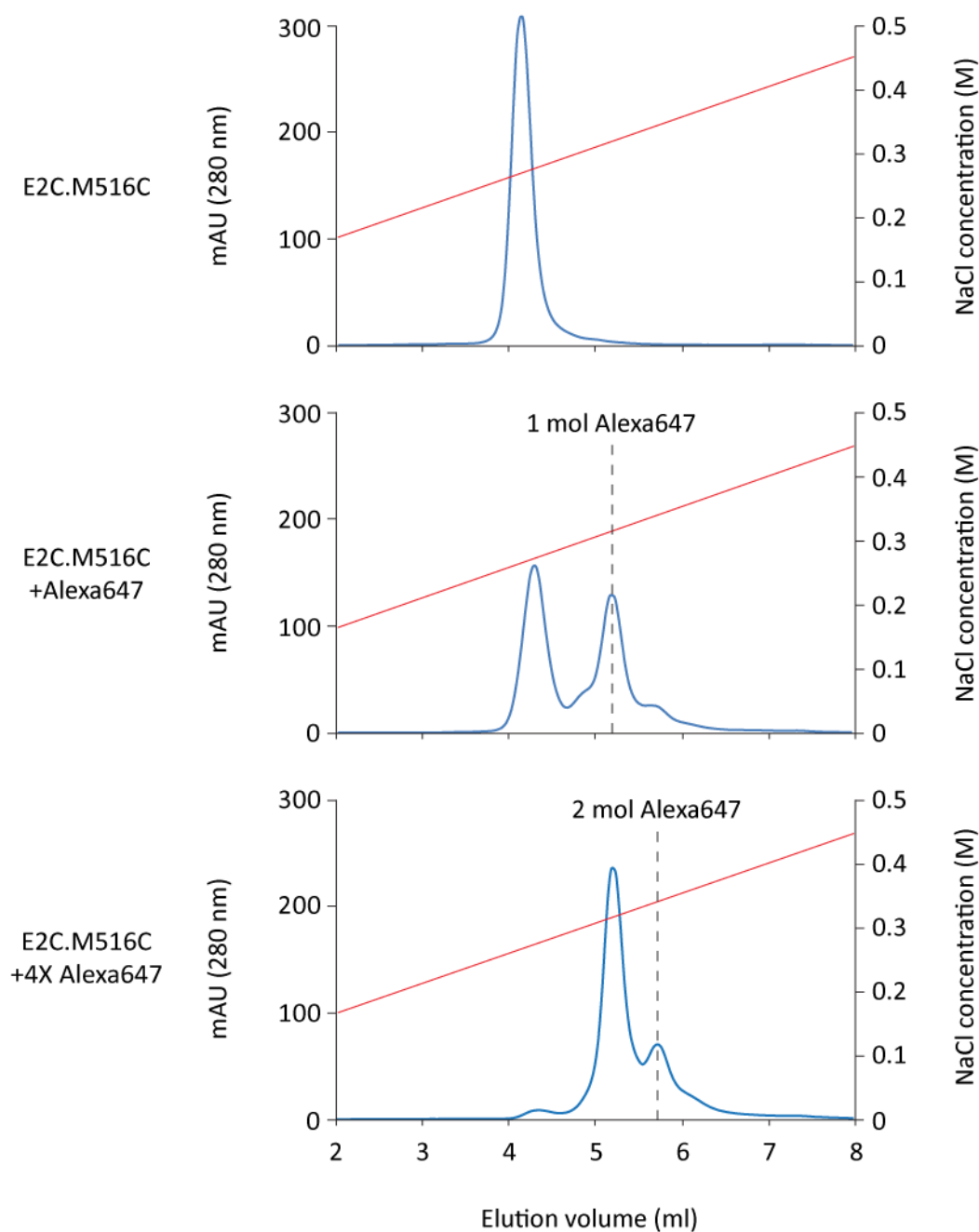


Figure 5.12 – Separation of Alexa647-labelled Lig1 species by ion-exchange chromatography

Untreated E2C.M516C Lig1 (top panel) was compared to E2C.M516C reacted with equimolar Alexa647-maleimide (middle panel) or a four-fold molar excess of Alexa647-maleimide (bottom panel). Samples were applied to a MonoQ ion-exchange column and eluted over a linear gradient of NaCl, shown by the red line and the right-hand y-axis. Protein elution was measured by absorbance at 280 nm, indicated by the blue line and the left-hand y-axis. Peak fractions indicated by the dashed lines were analysed by UV-visible spectroscopy and the amount of Alexa647 relative to 1 mol of Lig1 protein was calculated, as shown above the dashed line.

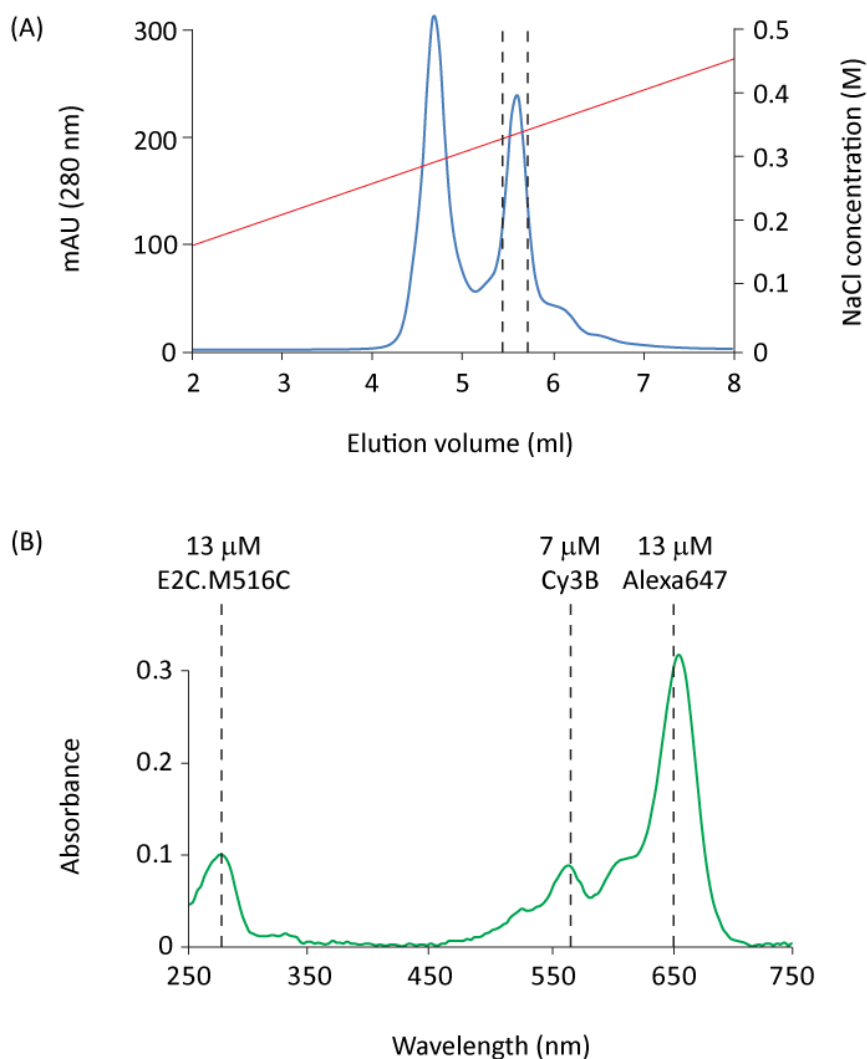


Figure 5.13 – Purification of Lig1 dual-labelled with Cy3B and Alexa647

(A) Purification of E2C.M516C labelled with a single Alexa647 molecule. After reaction with Cy3B and Alexa647, E2C.M516C was applied to a MonoQ ion-exchange column and eluted over a linear gradient of NaCl, indicated by the red line and right-hand y-axis. Protein elution was measured by absorbance at 280 nm, indicated by the blue line and left-hand y-axis. Fractions eluted between the dashed lines were collected and pooled for subsequent analysis.

(B) Analysis of purified Alexa647-labelled E2C.M516C. Absorbance of purified fractions was measured in the UV-visible range. The concentration of protein, Cy3B, and Alexa647 was calculated from the absorbance at their respective maximum absorbance wavelengths, indicated by annotated dashed lines.

Fluorophore conjugation at position 516 impairs Lig1 activity

To ensure that labelling of Lig1 to a high degree with both fluorophores did not disrupt its function, the *in vitro* activity of the purified labelled sample was tested relative to unlabelled E2C.M516C (Figure 5.14A).

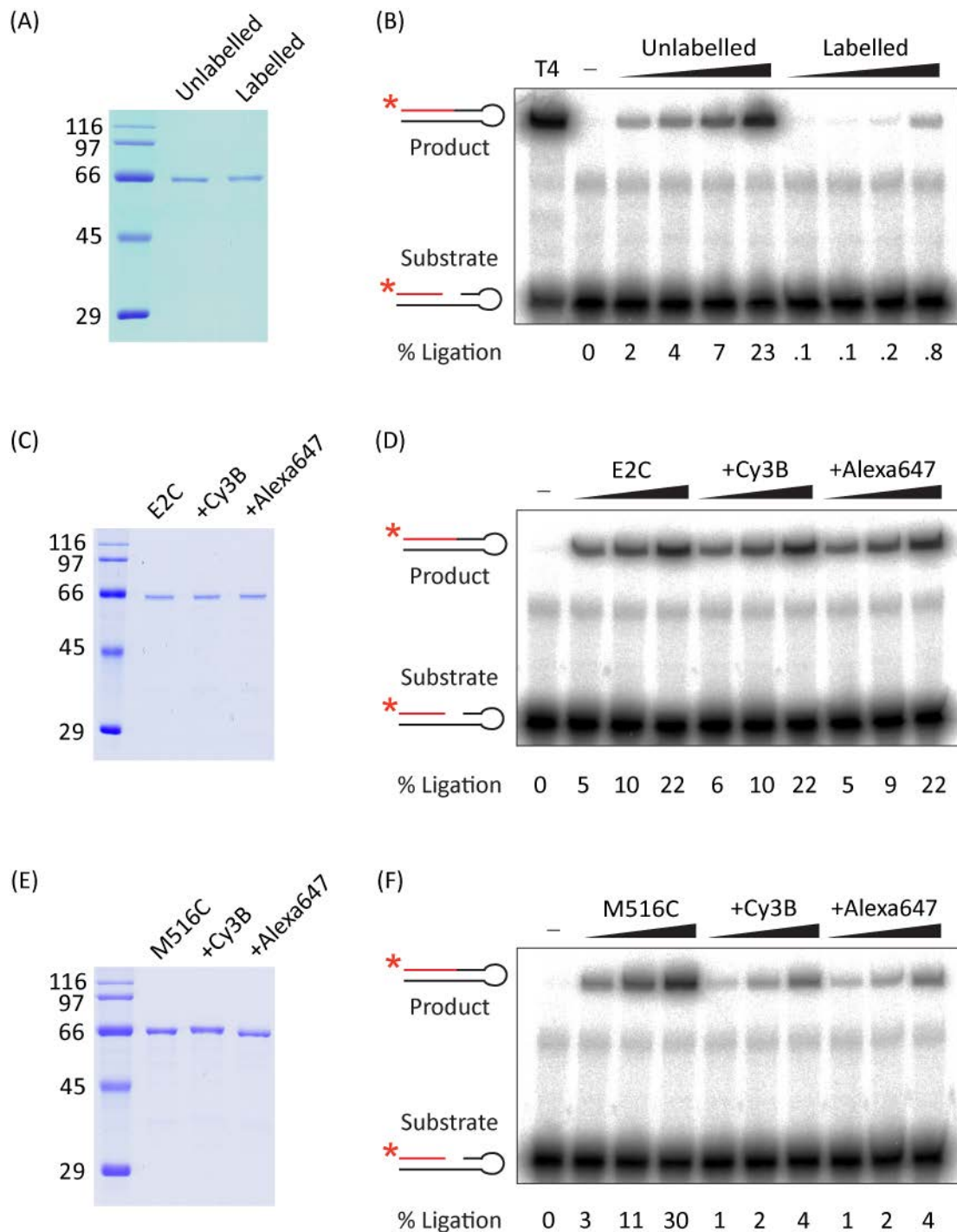


Figure 5.14 – The effect of covalent modification by fluorophores on the *in vitro* activity of Lig1

(A) Comparison of unlabelled E2C.M516C Lig1 with E2C.M516C dual-labelled with Cy3B and Alexa647. Proteins were analysed by SDS-PAGE and visualised by Coomassie staining. Molecular weight markers are on the left, sizes are in kDa.

(B) Comparison of unlabelled and dual-labelled E2C.M516C *in vitro* ligase activity. Reactions contained 0, 0.25, 0.5, 1, 2 pmol indicated protein. 1 unit of T4 DNA ligase was included as a positive control. Reaction products were separated by denaturing PAGE, and migration of substrate and product is

shown on the left; the labelled strand is highlighted in red and label position indicated with an asterisk. Quantification of product is shown underneath the gel.

(C) Comparison of unlabelled E2C Lig1 with Cy3B-labelled E2C and Alexa647-labelled E2C by SDS-PAGE.

(D) Comparison of the *in vitro* ligase activity of unlabelled E2C, Cy3B-labelled E2C and Alexa647-labelled E2C. Reactions contained 0, 2.5, 5, 10 pmol indicated protein.

(E) Comparison of unlabelled M516C Lig1 with Cy3B-labelled M516C and Alexa647-labelled M516C by SDS-PAGE.

(F) Comparison of the *in vitro* ligase activity of unlabelled M516C, Cy3B-labelled M516C and Alexa647-labelled M516C. Reactions contained 0, 1.5, 3, 6 pmol indicated protein.

Labelled E2C.M516C was found to have severely defective basal ligase activity. To determine if modification at a specific position or by a specific fluorophore was responsible for this disruption, the *in vitro* activity of individually-labelled single cysteine mutants of Lig1 was tested. Modification of E2C with either Cy3B or Alexa647 did not have any effect on the ligase activity of this protein (Figure 5.14B). It could be argued that due to the substoichiometric labelling of E2C (Figure 5.11), a defect caused by modification of this residue may have been undetectable, but given the magnitude of impairment observed with dual-labelled Lig1, together with the sensitivity of the ligase assay employed here, this is unlikely. In contrast, modification of M516C with either Cy3B or Alexa647 resulted in significantly reduced ligase activity (Figure 5.14C). This suggests that modification of position 516 was responsible for the dramatic loss of activity observed in the purified dual-labelled sample. This appears to be consistent with the difficulties encountered in identifying non-essential residues in the OBD, and perhaps suggests an important previously uncharacterised functional role for this region of Lig1. Given the significant defect in *in vitro* activity, it is unlikely that dual-labelled Lig1 fully retains its native structure or range of conformations, and therefore it was decided this sample would be unsuitable for subsequent FRET assays.

DISCUSSION

This study sought to measure conformational changes of Lig1 in solution, in order to understand how such changes may be related to the ability of a single PCNA ring to physically coordinate multiple enzymes simultaneously, most notably during Okazaki fragment maturation. To achieve this, the development of a FRET assay was undertaken, to measure Lig1 intramolecular distances under different conditions, with the effect of substrate binding to be tested initially. Unfortunately, it was not possible to fully develop this assay due to difficulties in engineering a dual fluorophore-labelled Lig1 molecule which retained its native ligase function. Nonetheless, a number of aspects of this study will prove useful for future work.

The substrate binding activity of Lig1 was characterised and quantified in this study, thus facilitating future studies on substrate-induced Lig1 conformational changes. Surprisingly, the substrate binding affinity of 15 μM determined for *S. solfataricus* Lig1 appears to be relatively weak, considering the affinity of human Lig1 for a non-ligatable nicked substrate was calculated to be more than ten-fold higher at 0.4 μM (Pascal et al, 2004). Differences in the experimental conditions between these studies do, however, make a direct comparison difficult. Notably, due to instrument limitations, the substrate binding measurements presented here were performed at room temperature rather than the physiologically relevant temperature of 70°C employed in functional assays, and thus the values derived here may represent substrate binding in significantly sub-optimal conditions.

Protocols were also developed in this study to rapidly generate Lig1 mutants and test for retention of ligase activity *in vitro*. Given that an introduced cysteine at position 2 can be modified by Cy3B or Alexa647 without a loss of function, future studies need to identify just one more site for modification using these established tools. Furthermore, a technique has been developed to enrich Lig1 dual-labelled with both Cy3B and Alexa647, thus potentially

enabling a sensitive FRET assay to be performed following identification of a second viable modification site.

One key difficulty encountered during this study was the identification of non-essential surface-exposed residues in the OBD for cysteine replacement; even once such a residue had been identified, its subsequent modification by Cy3B or Alexa647 disrupted Lig1 activity. The OBD not only plays an important role in engaging the minor groove of substrate DNA during catalysis (Pascal et al, 2004), it also appears to assist the catalytic adenylation domain (Add) during the formation of an essential ligase-AMP intermediate, which initiates the ligase catalytic cycle (Pascal, 2008). Compared to the elongated structure of *S. solfataricus* Lig1 (Pascal et al, 2006), the OBD undergoes large translational and rotational movements in the closed structures of *P. furiosus* and *A. fulgidus* Lig1, which have been proposed to represent intermediates in this protein adenylation reaction (Nishida et al, 2006; Kim et al, 2009). Similarly, the OBD adopts a further distinct orientation in the DNA-bound structure of human Lig1 (Pascal et al, 2004). Therefore, although none of the residues tested in this study are present at protein-protein interfaces in any of the observed OBD arrangements – consistent with their lack of strong conservation – it is possible that their substitution or modification may hinder the ability of the OBD to translate and rotate between its various positions throughout the Lig1 catalytic cycle. To identify potential sites for modification in the OBD, future studies may be able to use the increased structural data now available to identify residues unlikely to disrupt these structural transitions.

Interestingly, structural data suggests that large conformational rearrangements from more open elongated forms to compact DNA-encircling forms may be a unifying mechanism among DNA ligases. In addition to the examples of the ATP-dependent ligases discussed above, the structure of one of the smallest known eukaryotic DNA ligases, that of the *Chlorella* virus, has been solved in both the presence and absence of a DNA substrate, allowing direct comparison of its different conformational states (Odell et al, 2000; Nair et al, 2007). Similarly

to Lig1, *Chlorella* virus ligase encircles its substrate using OB-fold and adenylation domains, together with a novel DNA binding 'latch' which plays an analogous role to the Lig1 DBD. In the absence of DNA, this ligase adopts a more open conformation whereby the OBD is rotated approximately 180° relative to the AdD and the latch becomes disordered; this apparent interdomain flexibility is supported by recent NMR studies in solution (Piserchio et al, 2010).

Similarly, bacterial NAD-dependent ligases have also been shown to encircle their DNA substrate using an OBD-AdD core together with a helix-hairpin-helix (HhH) DNA binding domain (Nandakumar et al, 2007). Again, to engage the substrate, it appears that these domains must undergo significant rotational and translational rearrangements relative to each other (Lee et al, 2000). Finally, the structure of DNA-bound human DNA ligase III in combination with SAXS and molecular dynamics simulations suggest that this DNA ligase also displays large-scale interdomain flexibility in the absence of substrate, but fully encircles DNA upon encountering a nick (Cotner-Gohara et al, 2010).

In summary, it appears that a diverse range of different DNA ligases share this common conformational flexibility. It should be noted however, that these valuable insights have largely been provided by crystallographic studies thus far, and measurements of protein dynamics in solution – such as that attempted by this study – are still largely lacking. Such studies will be important in the future for understanding the observed conformational transitions of Lig1 in the context of its interaction with PCNA, particularly in the presence of other PCNA-bound factors.

Chapter 6 – Identification of two novel PolB1-interacting proteins

ABSTRACT

Replicative DNA polymerases in eukaryotes are multimeric enzymes, composed of primary catalytic subunits and additional accessory subunits. In contrast, the putative replicative polymerase in crenarchaea, PolB1, has previously been shown to be a monomeric protein possessing both 5'-3' polymerase and 3'-5' exonuclease activities. Furthermore, it is not known if PolB1 physically interacts with other components of the DNA replication machinery to facilitate functional coordination, as has been demonstrated in bacterial DNA replication. In this study, the composition and interactions of PolB1 within the cell were therefore investigated in more detail. Native PolB1 was purified from *Sulfolobus acidocaldarius*, and two novel PolB1 interacting proteins – PBP1 and PBP2 – were identified. These interactions were confirmed by several techniques, and the assembly of a stoichiometric trimeric complex between PolB1, PBP1 and PBP2 was demonstrated *in vitro*. This ternary complex was shown to be less active than monomeric PolB1 in DNA binding and primer extension activities *in vitro*, and it is hypothesised that PBP1 and PBP2 instead confer other yet to be determined activities upon PolB1 that may be essential for its function *in vivo*.

INTRODUCTION

DNA polymerases

DNA polymerase enzymes are central to genome duplication and maintenance in all cellular organisms, and to fulfil such essential functions they are both numerous and diverse; human cells for example possess 16 different DNA polymerases of four distinct types. Based on primary amino acid sequence, DNA polymerases are classified into one of six different families

– A, B, C, D, X or Y (Bebenek & Kunkel, 2004). In eukaryotic organisms, three B-type DNA polymerases are believed to be responsible for the bulk of chromosomal DNA replication. Elegant genetic studies in both *S. cerevisiae* and *S. pombe* utilising polymerase mutants with unique mutational spectra have demonstrated that Pol ϵ performs the majority of DNA synthesis on the leading strand (Pursell et al, 2007), while Pol δ is responsible for lagging strand replication (Nick McElhinny et al, 2008; Miyabe et al, 2011). Pol α , in complex with the heterodimeric primase complex, is responsible for synthesising RNA-DNA primers required to initiate each Okazaki fragment on the lagging strand and genetic interactions suggest that these primers are subsequently extended by DNA polymerase δ (Pavlov et al, 2006).

While the catalytic activities of Pol α , ϵ and δ reside in their large subunits (*S. cerevisiae* Pol1, Pol2 and Pol3 respectively), these polymerases are also associated with additional subunits. Each polymerase is associated with a different B-subunit (*S. cerevisiae* Pol12, Dpb2 and Pol31, respectively), which share sequence homology with one another. Despite lacking any detectable catalytic activity, these B-subunits are all essential (Bebenek & Kunkel, 2004). Furthermore, while Pol α is associated with the primase complex as discussed above, Pol ϵ is additionally associated with two subunits, Dpb3 and Dpb4, while Pol δ is associated with the unrelated Pol32 subunit. These non-catalytic subunits have all been shown to be important for DNA replication *in vivo*, but their exact functions are still unclear; they may play structural roles in stabilising the catalytic core of the enzyme or mediate protein-protein interactions with other components of the DNA replication machinery (Johansson & Macneill, 2010). In this respect it is interesting that a recent study demonstrated the presence of functional PIP motifs in all three subunits of *S. cerevisiae* Pol δ , which are proposed to act in concert to maximise polymerase processivity when in complex with PCNA (Acharya et al, 2011).

In comparison with eukaryotes, the distribution of DNA polymerases in the archaea is less uniform and their relative *in vivo* functions are currently less clear. Crenarchaeal species

appear to possess two or three distinct single-subunit B-type polymerases (Pisani et al, 1992; Prangishvili & Klenk, 1993; Edgell et al, 1997), in addition to a Y-family polymerase capable of performing lesion bypass DNA synthesis (Boudsocq et al, 2001). Interestingly, bioinformatic and biochemical analysis suggests that PolB1 orthologues are probably responsible for synthesising both leading and lagging strands during DNA replication, with PolB2 and PolB3 perhaps having a role in DNA repair; such functions *in vivo* have yet to be confirmed however (see Chapter 3 for more discussion). It is therefore unlikely that the division of function among multiple B-type polymerases observed in eukaryotes also exists in the crenarchaea.

In contrast, most euryarchaea possess a single B-type polymerase, although some species do also possess a second, potentially inactivated B-type polymerase gene (Rogozin et al, 2008). Notably however, euryarchaea also possess a D-family DNA polymerase composed of a large subunit, DP2, with 5'-3' polymerase activity and a small subunit, DP1, with 3'-5' proofreading exonuclease activity (Cann et al, 1998; Jokela et al, 2004). D-family polymerases are unique to euryarchaea, thaumarchaea and korarchaea (Brochier-Armanet et al, 2008; Elkins et al, 2008). Biochemical studies on the D-type polymerase from *Pyrococcus abyssi* have shown that, unlike PolB from this organism, it is able to extend from RNA primers and is capable of performing strand displacement synthesis, activities suited to lagging strand replication (Henneke et al, 2005). Furthermore, genetic studies in the euryarchaeon *Halobacterium* sp. NRC-1 have demonstrated that both the chromosomal B-type polymerase and the two subunit D-type polymerase in this organism are essential for viability (Berquist et al, 2007). This has led to the proposal of a model in which PolB synthesises the leading strand while PolD synthesises the lagging strand during chromosomal replication in euryarchaea.

Analysis of eukaryotic DNA polymerases in relation to the archaea has highlighted their complex evolution. While the catalytic subunits of eukaryotic B-type polymerases are most likely derived from archaeal B-type polymerases, the essential B-subunits of the eukaryotic polymerases are in fact related to the small DP1 subunit of euryarchaeal D-type polymerases

(Tahirov et al, 2009). Acquired mutations in key catalytic residues has resulted in the eukaryotic B-subunits losing exonuclease activity however (Aravind & Koonin, 1998; Makiniemi et al, 1999). The use of catalytically inactive domains derived from active DNA replication enzymes appears to be a common theme within the replication machinery: eukaryotic Pol ϵ is divided into an N-terminal active polymerase-exonuclease module derived from archaeal PolB, and a C-terminal inactive polymerase-exonuclease module, possibly of bacterial origin (Tahirov et al, 2009). Inactivated domains and subunits have been proposed to play a structural role in supporting the active components of the replication machinery (Rogozin et al, 2008). Eukaryotic B-type polymerases also appear to possess diverged, inactivated N-terminal uracil-binding pockets derived from archaeal B-type polymerases; in archaea these facilitate stalling of the polymerase upon encountering template uracil residues, preventing their erroneous replication (Wardle et al, 2008).

All eukaryotic B-type polymerases possess two C-terminal metal-coordinating cysteine clusters, proposed to form zinc fingers (Klinge et al, 2009). Further evolutionary complexity is illustrated by the similarity of these sequences to a zinc finger motif found at the C-terminus of euryarchaeal DP2 (Tahirov et al, 2009). It has been proposed that an ancestral archaeal PolB acquired the C-terminal fragment of DP2 containing its metal binding domain, and this region subsequently duplicated in the eukaryotic DNA polymerases. Indeed, this may be relevant to the acquisition by eukaryotic B-type polymerases of their B-subunits, as the zinc-finger of DP2 has been shown to be important in mediating interactions with the small DP1 subunit (Yamasaki et al, 2010). Interestingly, recent biochemical analysis has demonstrated that while one of the cysteine clusters in eukaryotic B-type polymerases is indeed likely to form a zinc finger, the more C-terminal motif, most important for assembly with the relevant B-subunit of the polymerase, is actually a [4Fe-4S] cluster (Netz et al, 2011). This further extends several recent unanticipated reports of iron-sulphur clusters present in nucleic acid processing enzymes (White & Dillingham, 2011). It will be interesting for future studies to establish

whether the putative ancestral metal binding cluster in euryarchaeal PolD is a zinc or iron binding motif.

The replisome

The multi-protein molecular machine responsible for coordinated leading and lagging strand synthesis at the replication fork is termed the replisome. One interesting outstanding question is how archaeal-eukaryotic DNA polymerases are physically linked to other replisome components, if at all. A precedent for such architecture comes from the bacterium *E. coli*, where the replicative polymerase Pol III is a stable component of the replisome. Pol III is a heterotrimer composed of the α polymerase subunit (a member of the C-family, unique to bacteria), the ϵ 3'-5' exonuclease proofreading subunit, and the non-catalytic θ subunit. In addition to its interaction with the β sliding clamp processivity factor, this $\alpha\epsilon\theta$ core associates with the τ subunit of the AAA+ clamp loader complex, which also interacts with the DnaB helicase-DnaG primase complex; thus the clamp loader serves to physically coordinate all the required activities for coupled leading and lagging strand synthesis (Johnson & O'Donnell, 2005). Interestingly, recent analysis has established both *in vivo* and *in vitro* that contrary to long established models, the clamp loader complex constitutively possesses three, not two, τ subunits and thus a single replisome is stably associated with three Pol III molecules (McInerney et al, 2007; Reyes-Lamothe et al, 2010). Single molecule studies suggest that the presence of a third polymerase in the replisome improves the efficiency of Okazaki fragment synthesis by acting as a backup lagging strand polymerase (Georgescu et al, 2011; Lia et al, 2011).

There is also evidence in bacteriophages for a network of interactions serving to physically coordinate leading and lagging strand polymerases with each other and with other replisome proteins. Bacteriophage T7 requires just four proteins for DNA replication and as such has been the subject of extensive study as a model replication machinery (Hamdan &

Richardson, 2009). The replicative polymerase, gp5, together with its processivity factor, *E. coli* thioredoxin, has been shown to directly interact with the gp4 hexameric helicase-primase (Hamdan et al, 2007). Thus gp4 is presumed to act analogously to the *E. coli* clamp loader, using its multiple subunits to bridge between leading and lagging strand gp5 molecules and coordinating them with the process of duplex unwinding. Such a model is supported by the ability of these four proteins to perform highly coordinated leading and lagging strand synthesis *in vitro* (Lee et al, 1998). Similarly, coordinated leading and lagging strand replication has been reconstituted *in vitro* from the 8 proteins required for T4 bacteriophage replication (Salinas & Benkovic, 2000). This has been proposed to be mediated by a direct dimerization of the replicative polymerase gp43, and the association of this dimeric polymerase with the gp41-gp61-gp59 primase-helicase complex (Ishmael et al, 2003).

In contrast, there is currently limited evidence to support the incorporation of the eukaryotic replicative polymerases δ and ϵ into a stable replisome (Kunkel & Burgers, 2008). Isolation of the core replisome components GINS and Mcm2–7 from *S. cerevisiae* has enabled the characterisation of the replisome progression complex (RPC), a large multi-protein assembly present at active replication forks which possesses a number of activities required for chromosomal replication (Gambus et al, 2006). Neither Pol δ nor Pol ϵ were found to be present within the RPC, although Pol α has been shown to be a weakly associating component (Gambus et al, 2009). A physical and functional interaction has been observed between Pol ϵ and the checkpoint mediator and RPC component Mrc1, raising the possibility that Pol ϵ may be associated with the replisome through this interaction (Lou et al, 2008); this has yet to be observed directly however.

Recent data has identified a fragile complex in *S. cerevisiae*, termed the pre-loading complex (pre-LC), composed of Pol ϵ , GINS, Sld2 and Dbp11 which assembles in a cell-cycle dependent manner prior to the initiation of DNA replication (Muramatsu et al, 2010). One hypothesis is that the non-catalytic C-terminal region of Pol ϵ may act as a scaffold to recruit

GIN5 to origins of replication and thus mediate assembly of the replicative Cdc45-Mcm2–7-GIN5 (CMG) helicase. It is unclear, however, what relevance this has to the association of Pol ϵ with active replication forks subsequent to replication initiation. Overall, despite the complex requirements of coordinating the action of the three DNA polymerases α , δ and ϵ between the leading and lagging strands and with other key replication proteins, the physical basis for such coordination is still relatively unclear.

The global architecture of the archaeal replisome and its relationship with DNA polymerases is currently perhaps even less well understood than that of eukaryotes, and has only been the subject of a limited number of studies to date. Screening of an *Archaeoglobus fulgidus* genomic DNA library using a yeast two-hybrid assay only identified interactions between the two subunits of PolD, and interactions of both PolB and PolD with PCNA (Motz et al, 2002). More recently, *in vivo* affinity tagging of DNA replication proteins in the euryarchaeon *Thermococcus kodakaraensis* was used to identify potential interaction partners by mass spectrometry (Li et al, 2010). Interactions were identified between PolB and RPA, MCM, and three genes of unknown function; and between PolD and PCNA, GIN15, MCM, DNA ligase, Cdc6 and one gene of unknown function. These interactions have yet to be verified by other techniques, but hint at a network of interactions which may be important for the formation of an active replisome.

An interaction of *Sulfolobus solfataricus* PolB1 with three replication initiation proteins Orc1-1, Orc1-2 and Orc1-3 has been proposed from biochemical analysis (Zhang et al, 2009b). The biological explanation for such an interaction is currently unclear, but it is possible that analogous to the pre-LC in *S. cerevisiae*, recruitment of PolB1 by initiator proteins may be an important functional step in origin activation. It is even less clear however whether such an interaction has relevance beyond initiation, to the function of PolB1 at active replication forks. Further biochemical analysis of *S. solfataricus* PolB1 has proposed a trimeric arrangement of the enzyme in the presence of DNA substrate (Mikheikin et al, 2009; Lin et al, 2012). This is

potentially relevant to the physical coordination of DNA polymerases on the leading and lagging strands as described above in bacterial and bacteriophage systems, although more work is required to confirm the functional relevance of this observed *in vitro* oligomerisation. Furthermore, the physical basis of how, if at all, *S. solfataricus* PolB1 is coordinated with other components of the DNA replication machinery – beyond its well-documented interaction with PCNA – has yet to be addressed.

This chapter investigates the composition of crenarchaeal PolB1 and its interaction with other proteins *in vivo* by isolating endogenous PolB1 from *Sulfolobus acidocaldarius* cells and identifying associated factors. Two previously uncharacterised PolB1 interacting proteins were identified and a range of techniques were used to confirm these interactions and demonstrate the formation of a stoichiometric ternary complex between PolB1 and these two factors *in vitro*. This ternary complex appears to be less active than isolated PolB1 in DNA binding and primer extension assays *in vitro*, and potential biological explanations for this are discussed.

RESULTS

S. acidocaldarius* PolB1 modified by a tandem affinity tag is fully functional *in vivo

To better understand the composition of the *Sulfolobus* DNA replication machinery *in vivo*, several *S. acidocaldarius* strains were constructed in which known DNA replication proteins were modified with an affinity tag to facilitate *in vivo* analysis and isolation for biochemical studies (Tamzin Gristwood, unpublished data). This study used a strain in which the presumed replicative DNA polymerase, PolB1, was modified: using homologous recombination-directed integration, the endogenous chromosomal PolB1 gene (Saci_1537) was replaced by a copy additionally encoding a tandem affinity tag of c-myc and hexahistidine at the C-terminus of the protein, following a short alanine linker (Figure 6.1A). Successful chromosomal integration was

confirmed by PCR amplification and DNA sequencing (Tamzin Gristwood, unpublished data). To further confirm the presence of both affinity tags in endogenous PolB1, exponentially growing cultures of both the parental strain – termed SacTK due to the prior integration of a thymidine kinase enzyme at the α -amylase locus (Gristwood et al, 2012) – and the modified strain – termed SacPolB1 – were subjected to western blotting analysis (Figure 6.1B). Both c-myc and histidine tags were detected in the modified strain only, at the correct size of 101 kDa, confirming that full-length PolB1 possesses both affinity tags.

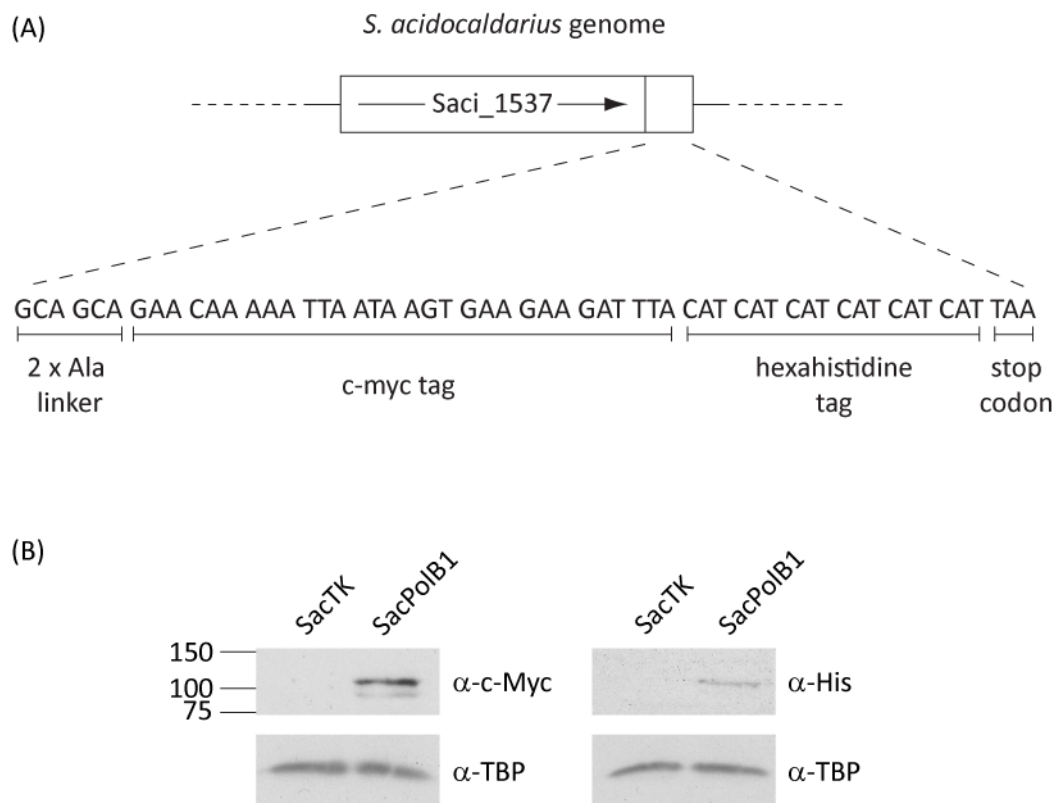


Figure 6.1 – Construction of an *S. acidocaldarius* strain in which PolB1 is modified with a dual affinity tag

(A) Schematic of chromosomal modification of *S. acidocaldarius* PolB1 (Saci_1537). The endogenous Saci_1537 gene was replaced by a copy containing the additional 3' sequences illustrated in the expanded view.

(B) Western blotting analysis of parental (SacTK) and modified (SacPolB1) strains to confirm the presence of a dual affinity tag. Migration of selected molecular weight markers is indicated on the left, antibodies used for detection are indicated on the right of each blot. TBP was used as a loading control.

To ensure that the introduction of these affinity tags into PolB1 did not affect DNA replication and thus bias subsequent *in vivo* analysis – for example by disrupting the interaction of PolB1 with DNA or other replication factors – the growth rate of the modified SacPolB1 strain was compared to the parental SacTK strain (Figure 6.2A).

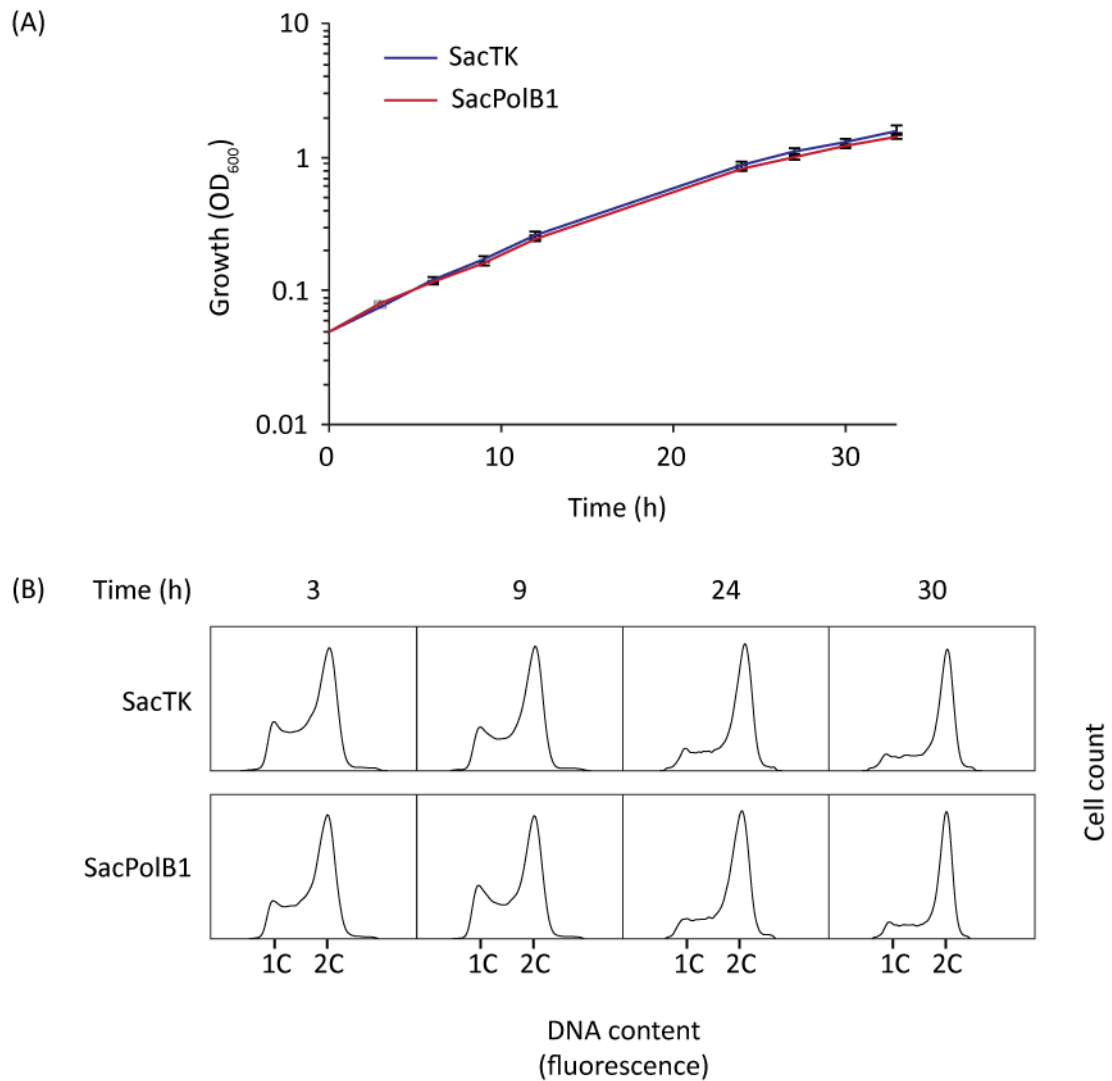


Figure 6.2 – Characterisation of an *S. acidocaldarius* strain in which PolB1 is modified with a dual affinity tag

(A) Comparison of growth of parental (SacTK) and modified (SacPolB1) *S. acidocaldarius* strains, measured by absorbance at 600 nm. The scale on the y-axis is logarithmic.

(B) Flow cytometry analysis of parental (SacTK) and modified (SacPolB1) strains. Samples were withdrawn from the time-course in (A) at the indicated time-points. '1C' and '2C' indicate fluorescent signal corresponding to 1 or 2 copies of genomic DNA, respectively

Both strains grew at the same rate, suggesting that the modified strain is able to undergo unimpaired DNA replication. Flow cytometry analysis at selected time-points during this growth also demonstrated that both strains had comparable DNA content throughout growth, supportive of identical DNA replication dynamics (Figure 6.2B). Thus PolB1 modified by a C-terminal tandem affinity tag is able to perform its normal *in vivo* function.

S. acidocaldarius* PolB1 interacts with two novel proteins *in vivo

To determine if PolB1 interacts with other known components of the *S. acidocaldarius* DNA replication machinery, or with other previously unidentified factors, PolB1 was isolated from *S. acidocaldarius* cells using its engineered histidine tag. As a potent DNA binding protein, there was potential for PolB1 to indirectly co-purify with other DNA binding factors in interactions mediated by DNA. To avoid erroneous assignment of such co-purification as direct protein-protein interactions, trial experiments were first performed to establish a purification protocol that could be performed in the absence of potential DNA bridging. To achieve this, optimum conditions were determined for the complete digestion of *S. acidocaldarius* genomic DNA in cell extracts by DNaseI (Figure 6.3A). This approach may also have potentially benefitted purification yield, as isolation of replisome components from *S. cerevisiae* was previously shown to require chromosomal digestion to release factors from chromatin and into the soluble extract (Gambus et al, 2006).

These conditions were applied to *S. acidocaldarius* biomass from the parental (SacTK) or modified strain (SacPolB1) and digested extracts were subjected to Ni²⁺-NTA chromatography to isolate histidine-tagged PolB1 (Figure 6.3B). Although this was successful in isolating a protein with the correct molecular weight of PolB1 (101 kDa), a large number of other proteins were also isolated non-specifically, as indicated by their presence in the SacTK sample. Purified material was therefore further subjected to ion-exchange chromatography, after which only a small number of proteins were present in the non-specific SacTK sample,

and several gel bands could be identified as specifically co-purifying with PolB1 in the SacPolB1 sample. These candidate gel bands were excised and analysed by mass spectrometry.

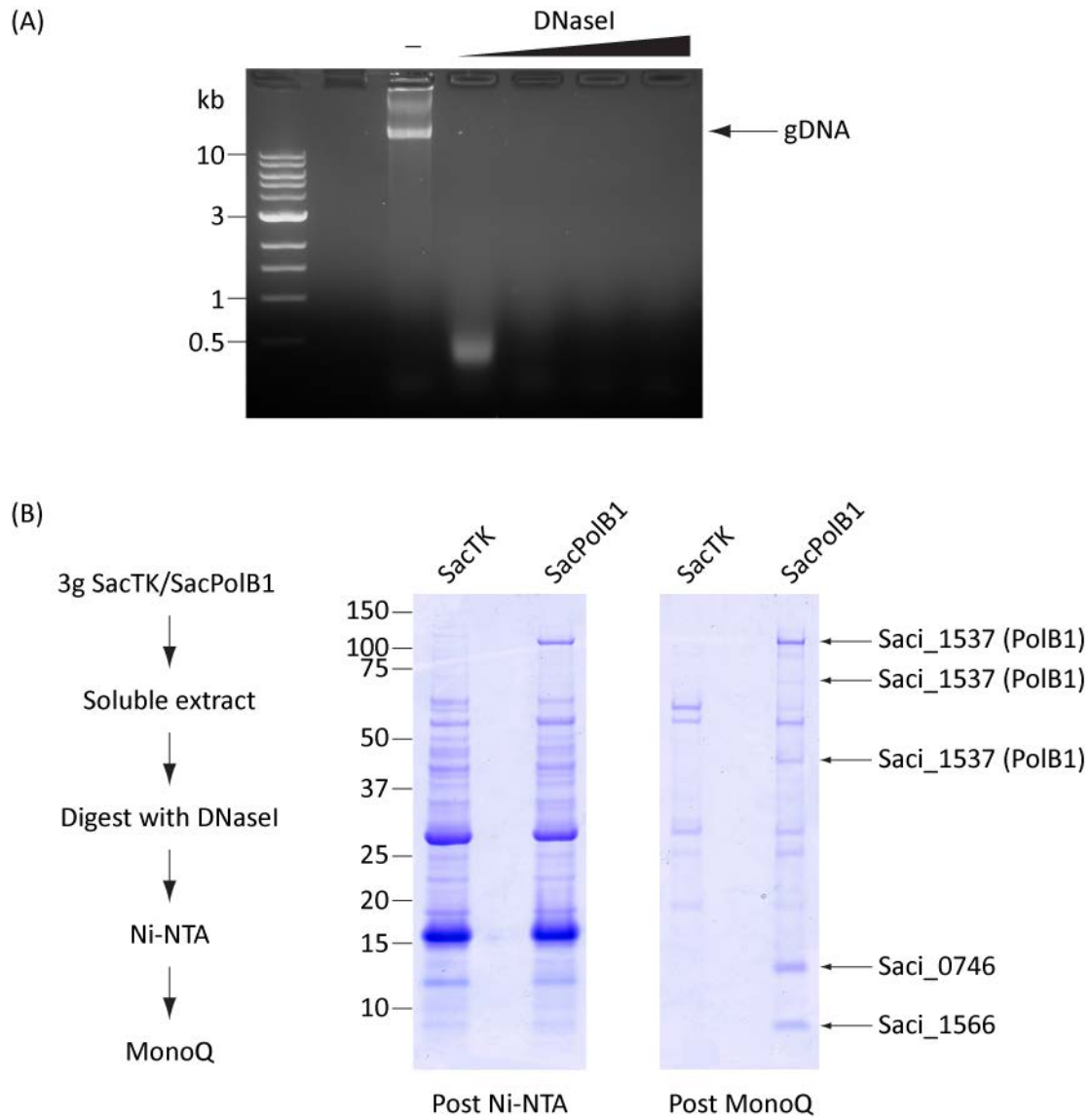


Figure 6.3 – Purification of endogenous PolB1 from *S. acidocaldarius*

(A) Trial digestion of genomic DNA in SacPolB1 cell extracts with DNaseI. Extracts were incubated with 0, 0.05, 0.5, 5, 50 units of DNaseI and purified DNA was analysed on a 0.5 % agarose gel. Migration of standards is indicated on the left, sizes are in kilobases (kb). 'gDNA' indicates the migration of undigested genomic DNA.

(B) Purification of dual affinity-tagged PolB1 from SacPolB1 cells. A schematic of the purification protocol is illustrated on the left. Samples were analysed by SDS-PAGE on 4–12 % gradient polyacrylamide gels and stained with Coomassie. Migration of molecular weight markers is indicated on the left, sizes are in kDa. Species identified by mass spectrometry are indicated on the right.

Multiple bands were identified as PolB1 – both full-length protein, identified from its expected migration at 101 kDa, and smaller products presumably resulting from protein degradation during the purification procedure; notably, recombinant *S. solfataricus* PolB1 also appears to display susceptibility to limited protein degradation during purification. Interestingly, two additional proteins were identified as specifically co-purifying with PolB1: Saci_0746 and Saci_1566. The fact that no known DNA or chromatin binding proteins were identified in this analysis supports the conclusion that the purification scheme was successful in eliminating DNA-mediated interactions, and suggests that the co-purification of PolB1 with both Saci_0746 and Saci_1566 was via direct protein-protein interactions.

Saci_0746 and Saci_1566 are crenarchaeal-specific proteins with unknown functions

Both Saci_0746 and Saci_1566 are annotated in the *S. acidocaldarius* genome (Chen et al, 2005) as hypothetical proteins of unknown function. Iterative interrogation of genome databases using PSI-BLAST identified several orthologues of both proteins (Figure 6.4A). The species distribution of orthologues of Saci_0746 and Saci_1566 is broadly the same, supportive of a related function for the two proteins. Both sets of orthologues appear to be restricted to crenarchaeal species only, and furthermore while orthologues exist in the sulfobacterales, desulfurococcales and acidilobales, none were detected in the final crenarchaeal order, the thermoproteales. The proposed early divergence of the thermoproteales (Brochier-Armanet et al, 2011) could be consistent with these two proteins only evolving subsequently, in the more closely related sulfobacterales, desulfurococcales and acidilobales. However, alternative analysis has suggested a late divergence for the thermoproteales (Yarza et al, 2010), and thus a lineage specific loss of these proteins in the thermoproteales can also currently be proposed. Notably, given the small size of both proteins and thus the limited sequence data available (97 amino acids for Sac_0746, 68 amino acids for Saci_1566) it is difficult to detect distantly related orthologues by bioinformatics alone. Therefore it cannot be ruled out that more highly

diverged orthologues of both proteins exist throughout the crenarchaea and indeed perhaps in the other archaeal phyla. Such limitations may also account for those species in which a Saci_1566 but not a Saci_0746 orthologue was detected.

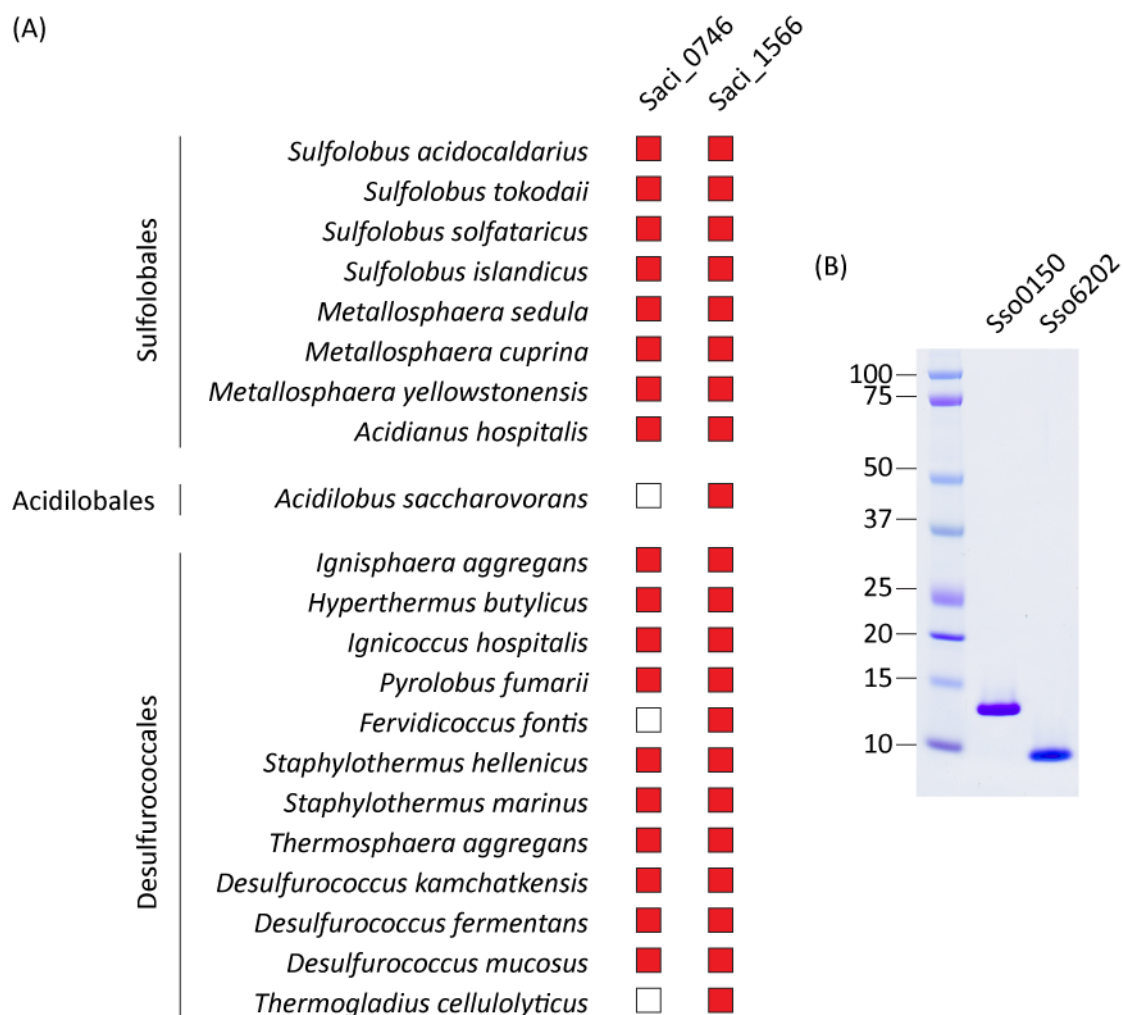


Figure 6.4 – Identification and purification of Saci_0746 and Saci_1566 orthologues

(A) Saci_0746 and Saci_1566 orthologues identified by PSI-BLAST. Presence of an orthologue in the indicated species is indicated by a red box, absence of an orthologue is indicated by a white box. The taxonomic order to which the listed species belong is indicated on the left.

(B) Purified recombinant *S. solfataricus* orthologues of Saci_0746 (Sso0150) and Saci_1566 (Sso6202) analysed by SDS-PAGE and visualised by Coomassie staining. Molecular weight markers are indicated on the left, sizes are in kDa.

There is no detectable homology between either of these proteins and three potential euryarchaeal PolB-interacting proteins identified in a previous study (Li et al, 2010). Furthermore, the incompatibility of their predicted molecular weights suggests these are not simply highly diverged orthologues, but completely different proteins.

Because no similarities to any known functional domains were identified in either protein by PSI-BLAST, three-dimensional structure prediction was performed using the Phyre server (Kelley & Sternberg, 2009) to gain possible insights into protein function. While no significant structural homology could be identified for Saci_1566, interestingly Saci_0746 was predicted to contain a Zinc finger domain with 85 % confidence (Figure 6.5B). Indeed, sequence analysis revealed the presence of two conserved cysteine and two conserved histidine residues (Figure 6.5A), typical of a C2H2 zinc finger domain (Krishna et al, 2003). These domains mediate diverse functions in DNA binding and protein-protein interactions, with the coordinated zinc ion functioning to stabilise protein folding. The presence of this domain in a putative PolB1-binding protein is particularly interesting when considering the distribution of such domains across other archaeal-eukaryotic DNA polymerases; while the catalytic subunits of both euryarchaeal D-family polymerases and all eukaryotic B-type DNA polymerases contain a C-terminal zinc finger domain, there is no evidence to date for an association between archaeal B-type DNA polymerases and a zinc finger domain.

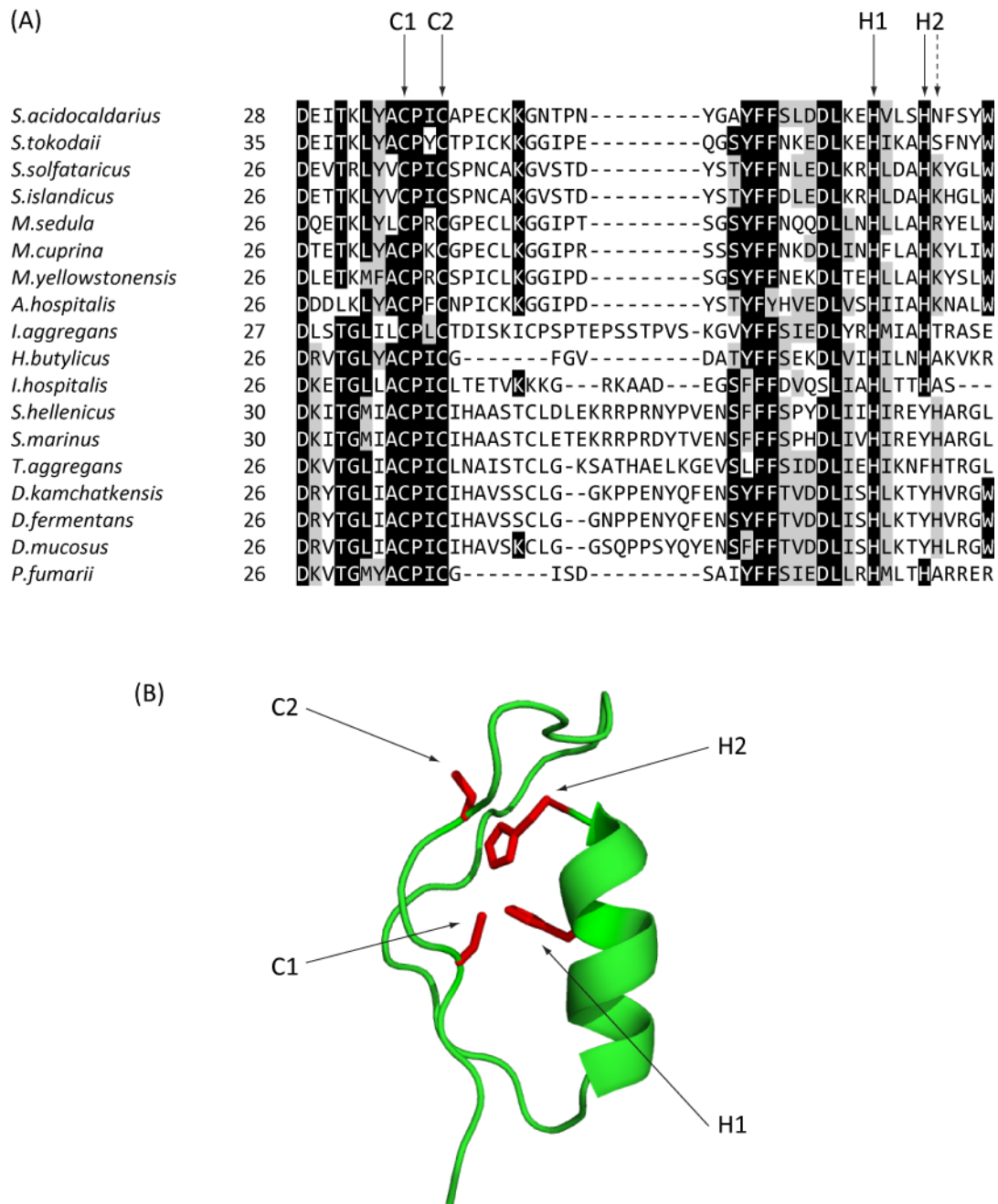


Figure 6.5 – Identification of a predicted zinc finger domain in PBP1

(A) Multiple sequence alignment of PBP1 orthologues from species labelled on the left. Residues with identity to the consensus are highlighted in black, residues with similarity to the consensus are highlighted in grey. The position of conserved cysteine (C) and histidine (H) residues as putative metal coordinating ligands is indicated above the alignment. The dotted line indicates the displacement of the second conserved histidine by one residue in some species. Sequences were aligned along their entire length but only selected regions are shown for clarity; the first residue shown of each protein sequence is numbered on the left

(B) Model of the predicted zinc finger of PBP1 generated by Phyre. Proposed zinc-coordinating cysteine (C) and histidine (H) residues are labelled and highlighted in red. The model is based on PDB file ADR1.

To facilitate further analysis of these proteins *in vitro*, recombinant *S. solfataricus* orthologues of Saci_0746 (Sso0150) and Saci_1566 (Sso6202) were purified and migrated with the expected molecular weights of 11.5 and 8.7 kDa respectively (Figure 6.4B). *S. solfataricus* orthologues were used due to the extensive previous characterisation and development of reagents for *S. solfataricus*, rather than *S. acidocaldarius*, PolB1. The primary sequences of Sso0150 and Sso6202 possess a high degree of conservation with their *S. acidocaldarius* orthologues (48 % identity/68 % similarity and 54 % identity/74 % similarity, respectively) and thus are likely to fulfil the same function in both species. For simplicity, Sso0150 and Sso6202 are henceforth referred to as PolB1-binding protein 1 (PBP1) and PBP2, respectively.

PBP1 and PBP2 directly interact with PolB1 via non-overlapping interfaces

To test whether PBP1 and PBP2 do indeed directly interact with PolB1, they were additionally generated as GST fusion proteins, immobilised on beads, and their ability to pull-down purified recombinant *S. solfataricus* PolB1 was measured (Figure 6.6A). While there was no pull-down of PolB1 by GST alone, robust pull-down was observed with either PBP1 or PBP2 fused to GST, demonstrating a direct physical interaction between these proteins and PolB1. No interaction was observed between PBP1 or PBP2 and the *S. solfataricus* Y-family translesion DNA polymerase Dpo4, suggesting that these are not general DNA polymerase-associated factors. Additionally, neither PBP1 nor PBP2 was capable of pulling down *S. solfataricus* Vps4, a cell division protein with no known or predicted DNA replication function (Samson et al, 2008), further ruling out non-specific promiscuous binding by these two proteins.

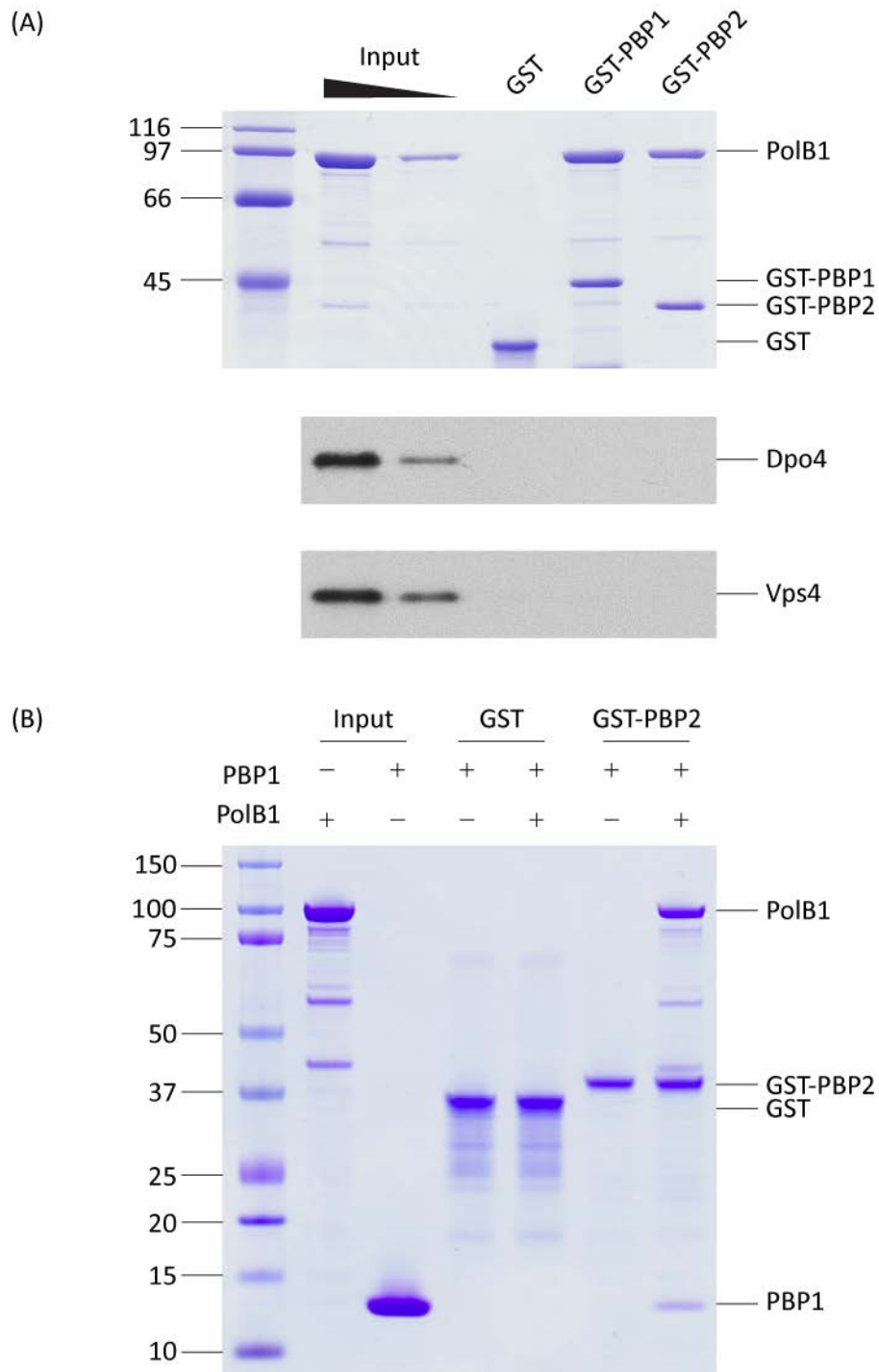


Figure 6.6 – An interaction between PolB1 and PBP1/PBP2 *in vitro*

(A) Pull-down of recombinant PolB1 by GST-tagged PBP1 or PBP2. Samples were analysed by SDS-PAGE and visualised by Coomassie staining (PolB1) or western blotting using Dpo4 or Vps4 antisera. Inputs are 50 % and 10 %. Molecular weight markers are shown on the left, sizes are in kDa.

(B) PolB1-mediated pull-down of recombinant PBP1 by GST-tagged PBP2. Samples were analysed by SDS-PAGE and visualised by Coomassie staining. Inputs are 50 %. Molecular weight markers are shown on the left, sizes are in kDa.

While both PBP1 and PBP2 co-purified with endogenous PolB1, and both individually interacted with PolB1 *in vitro*, it was not clear whether binding of these two factors to PolB1 was mutually exclusive. To test this, GST-tagged PBP2 was tested in its ability to pull-down recombinant untagged PBP1 (Figure 6.6B). No direct interaction was observed between PBP2 and PBP1. However, when PolB1 was present, pull-down of PBP1 by PBP2 was observed, suggesting that PolB1 can act as a bridging factor between PBP1 and PBP2. Importantly, this is only possible if PBP1 and PBP2 interact with different regions of PolB1 in a non-competitive manner. These data therefore suggest that PolB1, PBP1, and PBP2 may form a trimeric complex.

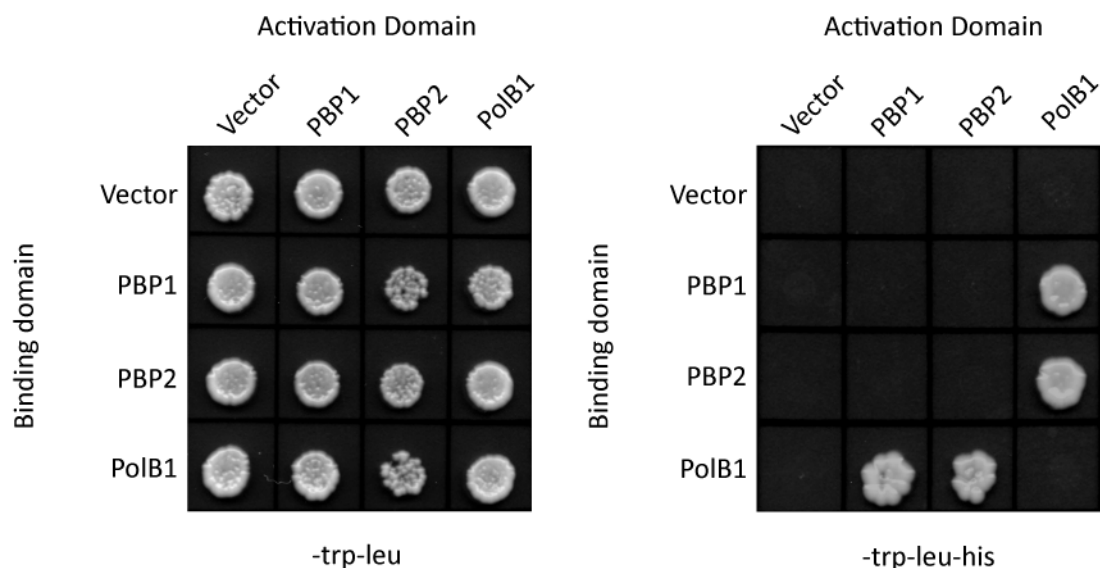


Figure 6.7 – An interaction between PolB1 and PBP1/PBP2 analysed by yeast two-hybrid assay

Successful transformation of yeast with the labelled genes, fused to either the GAL4 DNA binding domain or activation domain as indicated, is demonstrated by growth on media lacking tryptophan and leucine (left hand panel). Interaction between the indicated fusion proteins is demonstrated by growth on media lacking tryptophan, leucine, and histidine (right hand panel). The pattern of interactions shown here was consistent across two independent clones tested for each interaction.

To verify and extend the results of this pull-down assay, a yeast two-hybrid assay was used to analyse all potential pairwise interactions between PolB1, PBP1 and PBP2 (Figure 6.7).

In agreement with the pull-down assay, growth of diploid yeast on selective (-trp-leu-his) media – indicative of a protein-protein interaction – was observed for both PolB1-PBP1 and PolB1-PBP2. No interaction was observed between PBP1 and PBP2, again supportive of these two factors binding PolB1 independently of one another.

PolB1, PBP1 and PBP2 form a stoichiometric trimeric complex

To further determine the molecular composition of a potential complex between PolB1, PBP1 and PBP2, these proteins were analysed by gel filtration on a Superdex 200 column. Each recombinant protein was first analysed individually (Figure 6.8). Relative to standards of known molecular weight, PolB1, PBP1 and PBP2 all eluted as monomeric species; this is in agreement with the yeast two-hybrid data, which failed to detect self-interaction for any of these proteins. PolB1 was then incubated with a three-fold molar excess of either PBP1 or PBP2 before gel filtration analysis (Figure 6.9A and 6.9B). In both cases, a very small decrease in the elution time of the PolB1 peak was observed, which together with an increase in the absorbance value of this peak, was indicative of the formation of a complex of marginally larger molecular weight than PolB1 alone. The formation of a complex between PolB1 and each factor individually was confirmed by SDS-PAGE analysis of the eluted fractions. Incubation of a three-fold excess of both PBP1 and PBP2 simultaneously with PolB1 prior to gel filtration resulted in a further decrease in elution time and increase in absorbance of the PolB1 peak relative to either pairwise combination, suggesting the formation of a complex of greater molecular weight than either PolB1-PBP1 or PolB1-PBP2 alone (Figure 6.9C). SDS-PAGE of eluted fractions confirmed co-elution of PolB1, PBP1 and PBP2. While this co-elution could represent a poorly resolved mixture of PolB1-PBP1 and PolB1-PBP2, the stepwise decrease in elution time observed, together with the preceding GST pull-down data, suggests these three proteins do indeed form a trimeric complex.

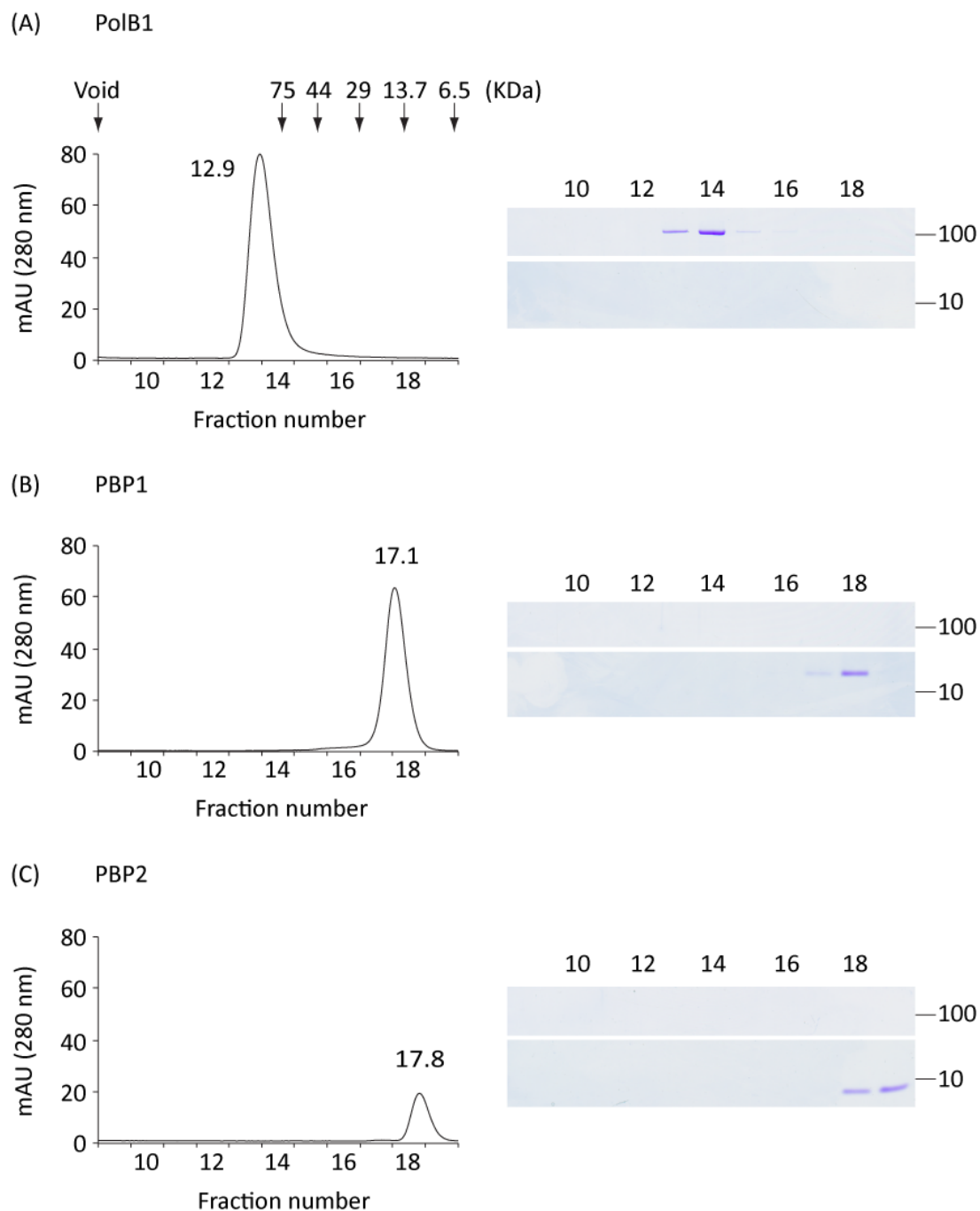


Figure 6.8 – Analytical gel filtration of recombinant PolB1, PBP1 and PBP2

(A) Analysis of 0.8 nmol PolB1 on a Superdex 200 HR 10/30 column. Elution of PolB1, monitored by UV absorbance at 280 nm, is shown in the left hand panel with the peak elution volume in ml indicated next to the peak. The elution volumes of calibration proteins of known molecular weight are indicated above the chromatogram. Analysis of indicated fractions by SDS-PAGE and Coomassie staining is shown in the right hand panel. Molecular weight markers are indicated on the right of SDS-PAGE gels, sizes are in kDa.

(B) Analysis of 2.4 nmol PBP1 as (A)

(C) Analysis of 2.4 nmol PBP2 as (A)

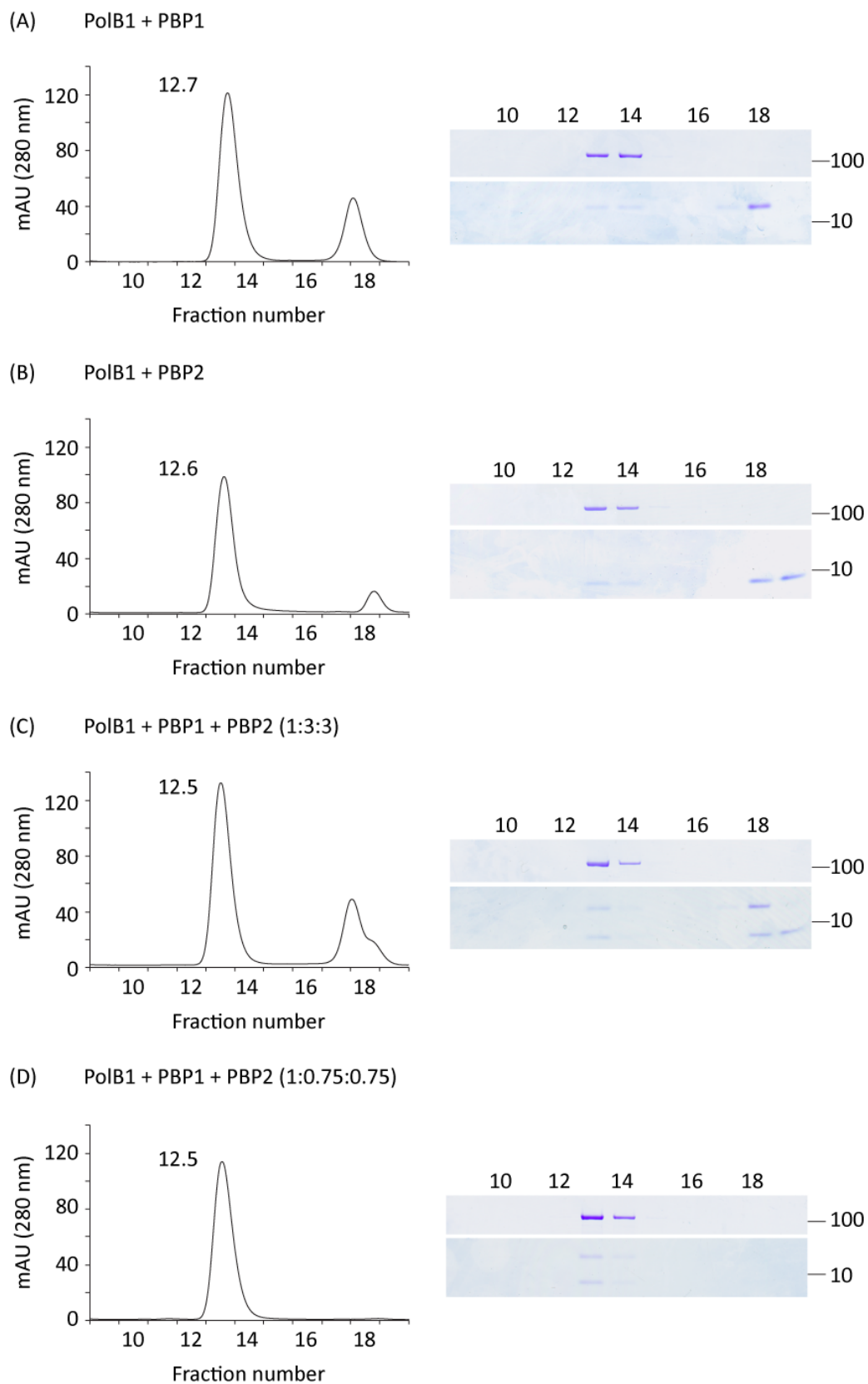


Figure 6.9 – Analytical gel filtration of PolB1-PBP1-PBP2 complexes

(A) Analysis of 0.8 nmol PolB1 + 2.4 nmol PBP1 on a Superdex 200 HR 10/30 column. Elution of proteins monitored by UV absorbance at 280 nm is shown in the left hand panel, with the peak elution volume in

mls indicated next to the peak. Analysis of indicated fractions by SDS-PAGE and Coomassie staining is shown in the right hand panel. Molecular weight markers are indicated on the right of SDS-PAGE gels, sizes are in kDa.

(B) Analysis of 0.8 nmol PolB1 + 2.4 nmol PBP2 as (A)

(C) Analysis of 0.8 nmol PolB1 + 2.4 nmol PBP1 + 2.4 nmol PBP2 as (A)

(D) Analysis of 0.8 nmol PolB1 + 0.6 nmol PBP1 + 0.6 nmol PBP2 as (A)

When PolB1 was incubated with a substoichiometric amount of both PBP1 and PBP2, complex formation between PolB1, PBP1 and PBP2 was still observed, but no free PBP1 or PBP2 was detectable, suggesting that under these conditions there is no appreciable complex dissociation during gel filtration (Figure 6.9D). Free PBP1 and PBP2 observed when these proteins were in molar excess over PolB1 is thus most likely to represent saturation of PolB1 binding sites rather than dissociation from PolB1. PolB1-containing fractions under these conditions were therefore assumed to represent the fully-formed trimeric complex. Two-dimensional densitometry analysis of these fractions run on Coomassie-stained SDS-PAGE gels established a mass ratio of 1:0.11:0.08 for PolB1:PBP1:PBP2, which corresponds to a molar ratio of 1:1:1.

PBP1 and PBP2 confer increased thermostability upon PolB1

S. solfataricus is an obligate hyperthermophile, growing optimally at temperatures of 80°C, and displaying viability up to 92°C (Zaparty et al, 2010), and thus proteins derived from this organism typically display innate thermostability. During preparation of recombinant PolB1, it was previously noted that this protein did not display the extreme thermostability observed with many other *S. solfataricus* proteins. It has previously been demonstrated that the *S. solfataricus* heterotrimeric PCNA complex is significantly more thermostable than any of the PCNA subunits individually (Pascal et al, 2006), supportive of the heterotrimer being the dominant form of *S. solfataricus* PCNA present *in vivo*. It was hypothesised that PolB1 may also only exhibit maximum stability when present in its native complex form found *in vivo*. To

test this, PolB1 or the PolB1-PBP1-PBP2 ternary complex isolated by gel filtration (henceforth referred to as PolB1*) were incubated at a range of temperatures before removal of heat-denatured proteins by centrifugation. The remaining heat-stable soluble protein was compared by SDS-PAGE (Figure 6.10). While PBP1 and PBP2 are both stable up to temperatures of at least 85°C, PolB1 appears to be highly unstable at this temperature, despite it being within the range of physiological growth conditions for *S. solfataricus*. In contrast, PolB1* is stable up to the maximum temperature tested of 85°C. This crude test therefore suggests that the ternary complex of PolB1-PBP1-PBP2 may be the predominant form of PolB1 present in *S. solfataricus*.

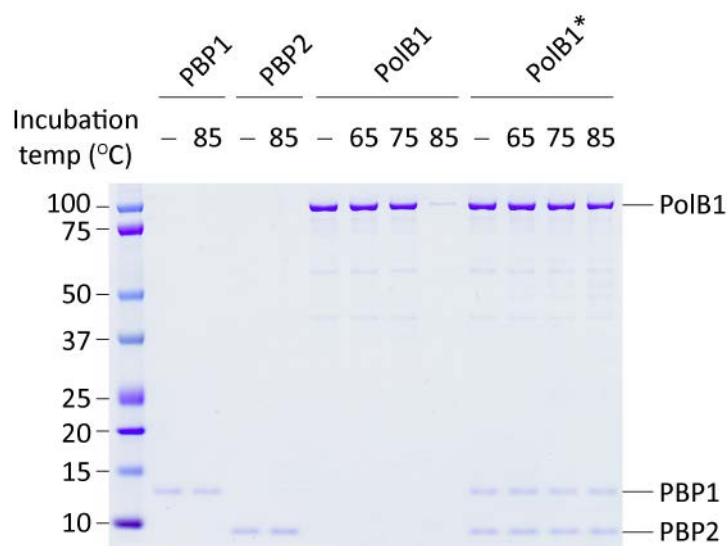


Figure 6.10 – Heat stability of PolB1 and PolB1*

Soluble protein remaining after incubation at the indicated temperatures was analysed by SDS-PAGE and visualised by Coomassie staining. The migration of individual subunits is indicated on the right. Molecular weight markers are shown on the left, sizes are in kDa.

PolB1* is less active in DNA binding and primer extension activities than PolB1

Hypothesising that PolB1* is the main form of PolB1 *in vivo*, the reconstituted ternary complex was compared with PolB1 in biochemical assays, to test whether the formation of such a

complex within the cell could confer specific biological functions upon the PolB1 enzyme. The two forms of PolB1 were first compared in their ability to bind a primer-template substrate by EMSA (Figure 6.11A and 6.11B). Both PolB1 and PolB1* bound this substrate, and consistent with the presence of its additional subunits, PolB1*-DNA complexes migrated more slowly than PolB1-DNA. Notably however, DNA binding by PolB1* is significantly decreased relative to PolB1. A trimeric arrangement of PolB1 on DNA has previously been proposed based on the presence of multiple protein-DNA complexes observed by EMSA (Mikheikin et al, 2009). Although the protein concentrations employed here are significantly lower, and the buffer conditions different, there does appear to be at least one slower migrating form of PolB1 at the highest concentration tested. It will be interesting for future studies to determine if there is any difference in the oligomeric composition of DNA binding species between these two forms of PolB1; the lower DNA binding activity of PolB1* precluded a direct comparison from being performed here.

To further investigate the activity of PolB1*, the two samples were compared in their DNA synthesis activity on the same primer-template substrate (Figure 6.11C). Again, relative to PolB1, PolB1* demonstrated significantly lower activity. Complete inhibition was not observed however, as DNA synthesis products were observed at the highest concentration of protein, and thus PolB1* is still an active polymerase. This residual activity is unlikely to result from contaminating uncomplexed PolB1, as the DNA binding activity of PolB1* observed by EMSA is clearly consistent with a homogeneous complex of higher molecular weight than PolB1. Additionally, the distribution of DNA synthesis products does not suggest any significant difference in processivity between the two species of PolB1. Currently, the biological explanation for these observations is unclear.

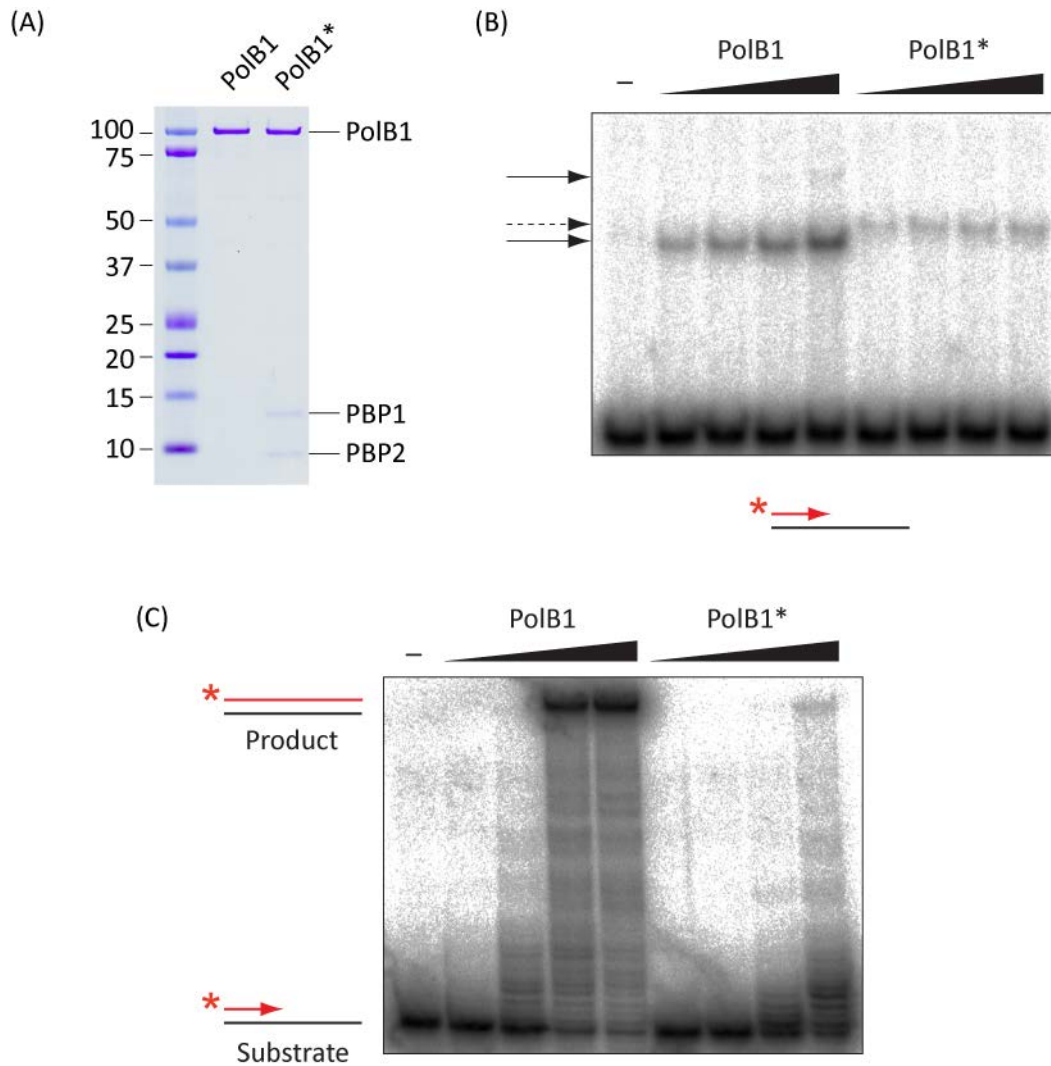


Figure 6.11 – Comparison of the *in vitro* activities of PolB1 and PolB1*

(A) Comparison of PolB1 and PolB1* reconstituted by gel filtration. Samples were analysed by SDS-PAGE and visualised by Coomassie staining. The migration of individual subunits is indicated on the right. Molecular weight markers are shown on the left, sizes are in kDa.

(B) Comparison of the *in vitro* DNA binding activity of PolB1 and PolB1*. Reaction products were separated by native gel electrophoresis. Substrate is shown below the gel, with the labelled strand highlighted in red, and label position indicated by an asterisk. Substrate was incubated with 0, 0.05, 0.1, 0.2, 0.4 pmol indicated protein. The migration of DNA species bound by PolB1 (solid line) and PolB1* (dashed line) are indicated by arrows.

(C) Comparison of *in vitro* primer extension activity of PolB1 and PolB1*. Reaction products were separated by denaturing PAGE. Substrate was incubated with 0, 0.4, 4, 40, 400 fmol protein.

PBP1 and PBP2 do not mediate interactions between PolB1 and the GINS complex

The identification of two new PolB1-interacting proteins raises the possibility that they may function by mediating interactions between PolB1 and other core components of the *S. solfataricus* replisome. One potential candidate for a replisome-organising complex in *S. solfataricus* is that of the GINS complex (GINS15-GINS23-RecJ_{dbh}) which has previously been shown to act as a bridge between the replicative helicase, MCM, on the leading strand and DNA primase on the lagging strand (Marinsek et al, 2006). Recent bioinformatic analysis has demonstrated that RecJ_{dbh} is a distant orthologue of eukaryotic Cdc45, a key component of the replicative CMG helicase (Sanchez-Pulido & Ponting, 2011). Furthermore, the localisation of *S. acidocaldarius* RecJ_{dbh} during S-phase supports a central role for the GINS complex at active replication forks *in vivo* (Tamzin Gristwood, unpublished data). GINS15, GINS23 and RecJ_{dbh} have all been previously well-characterised in a yeast two-hybrid assay (Marinsek et al, 2006), and were therefore suitable candidates to test as potential interaction partners of PBP1, PBP2 and PolB1. However, no interactions were detected between any of these components in a yeast two-hybrid screen, suggesting that neither PolB1 nor PolB1* directly interacts with the GINS complex (Figure 6.12). The fact that GINS15, GINS23 and RecJ_{dbh} (Marinsek et al, 2006), together with PolB1, PBP1 and PBP2 (Figure 6.7) have all previously been demonstrated to be active in the yeast two-hybrid system suggests this is genuinely due to an absence of any interactions rather than lack of expression or appropriate folding of these factors in *S. cerevisiae*.

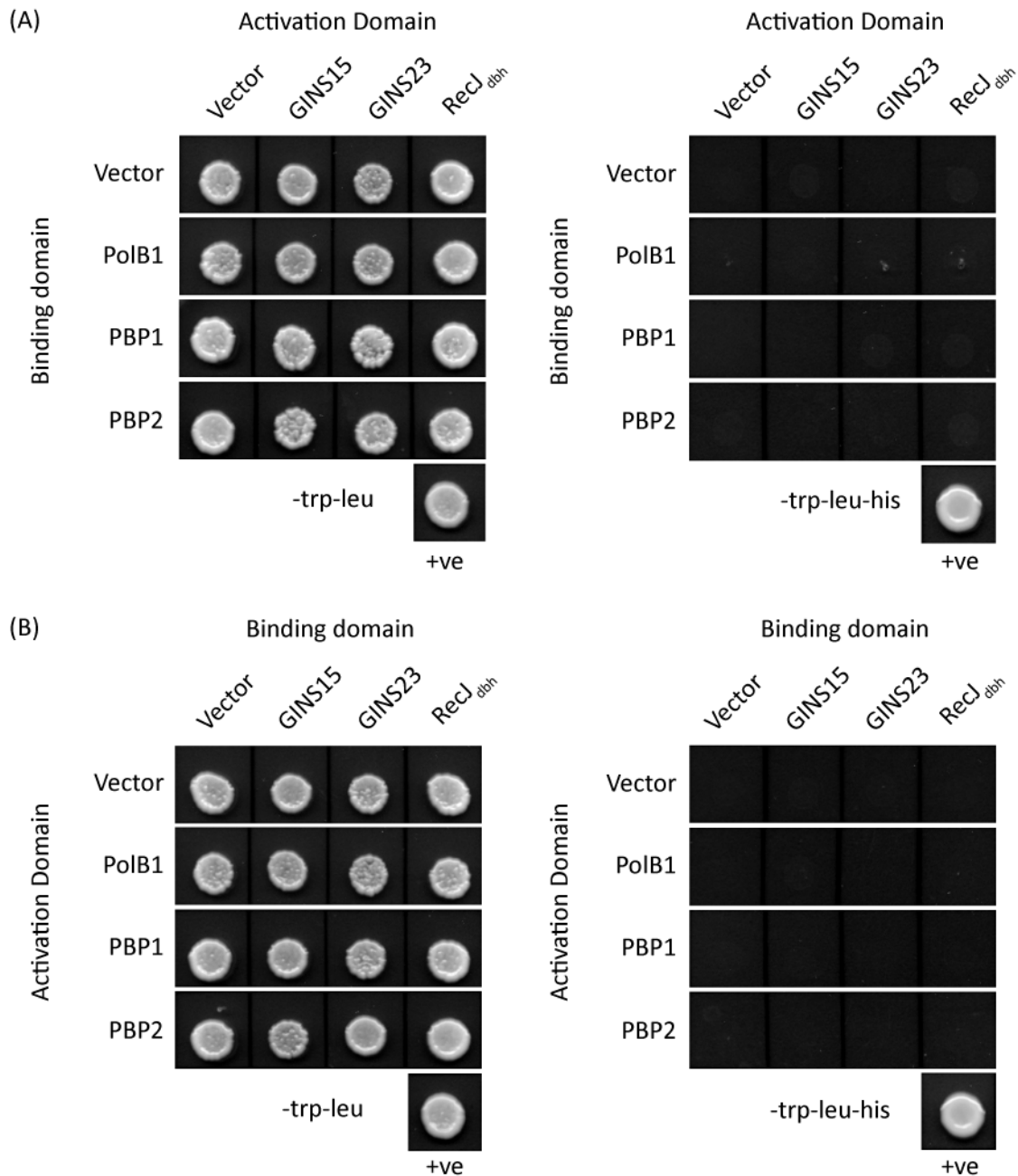


Figure 6.12 – Interactions between PolB1* and the GINS-RecJ_{dbh} complex

(A) Yeast two-hybrid analysis of interactions between PolB1, PBP1 or PBP2 fused to GAL4 DNA binding domain and GINS15, GINS23 or RecJ_{dbh} fused to GAL4 activation domain. Successful transformation of yeast with the indicated genes is demonstrated by growth on media lacking tryptophan and leucine (left hand panel). Interaction between the indicated fusion proteins is demonstrated by growth on media lacking tryptophan, leucine, and histidine (right hand panel). '+ve' indicates interaction between p53 and large T antigen. The pattern of interactions shown here was consistent across two independent clones tested for each interaction.

(B) Yeast two-hybrid analysis of interactions as in (A) between PolB1, PBP1 or PBP2 fused to GAL4 activation domain and GINS15, GINS23 or RecJ_{dbh} fused to GAL4 DNA binding domain.

DISCUSSION

The data presented here suggest that in many crenarchaeal species the replicative polymerase, PolB1, is associated with two previously unidentified accessory subunits, in a stoichiometric ternary complex. Several lines of evidence support these interactions existing *in vivo*. Firstly, an association between these proteins was detected in extracts from *S. acidocaldarius* cells which appear to undergo normal DNA replication. Genuine protein-protein interactions were confirmed by yeast two-hybrid and *in vitro* pull-down assays, together with gel filtration of recombinant purified proteins. Furthermore, the PolB1 ternary complex was found to exhibit greater thermostability than PolB1 alone, demonstrating a physiological relevance for complex formation in this hyperthermophilic species.

It is interesting to note that both *S. acidocaldarius* and *S. solfataricus* PolB1 were first identified by native purification from *Sulfolobus* extracts but no additional subunits were observed (Klimczak et al, 1985; Rella et al, 1990). Without the advantage of a protein affinity tag, both these studies employed numerous chromatographic steps to achieve enzyme purity, and therefore it is possible that subunit interactions within PolB1 were lost throughout these multiple treatments. Furthermore, as purification progress in these studies was monitored by DNA synthesis activity, it is possible that the lower activity of PolB1*, as observed here, may have led to the biased selection of fractions containing PolB1 alone.

It is currently unclear why PolB1* possesses a decreased ability to bind to a DNA substrate and perform primer extension relative to PolB1. One possibility is that these two additional subunits do indeed function as negative regulators of PolB1 *in vivo*, acting to prevent wasteful polymerisation and exonuclease activities outside of S-phase. However, an optimal 1:1:1 stoichiometry of PolB1-PBP1-PBP2 determined *in vitro* was also broadly observed with PolB1 isolated from *S. acidocaldarius* cells (1:1.4:1.2 taking into account degradation products identified as PolB1, and with the caveat that complex stoichiometry may

be disrupted during chromatography). This suggests that the vast majority of PolB1, if not all, was associated with these accessory subunits *in vivo*, and it is difficult to reconcile a potential negative regulatory role with the requirement for ongoing DNA synthesis in these exponentially growing cells. Characterisation of these interactions over the course of the *S. acidocaldarius* cell cycle will enable this hypothesis to be tested directly.

It is worth noting that PolB1* still possesses residual DNA binding and primer extension activities, and therefore it can be hypothesised that this decreased function due to association with PBP1 and PBP2 may still be sufficient *in vivo* while allowing other essential functions to be conferred upon PolB1 by these two subunits. For example, the addition of these subunits clearly enhances the thermostability of PolB1, an essential property for optimal function in the hyperthermophilic environment of *Sulfolobus*. In this respect it is worth noting that all the species identified as possessing PBP1 and PBP2 orthologues are hyperthermophiles and so it may be that these additional subunits have evolved simply to stabilise the catalytic PolB1 subunit at elevated temperatures. However, due to bioinformatic limitations, it cannot currently be ruled out that orthologues of these proteins are cofactors of PolB1 in non-hyperthermophilic species. It could therefore also be hypothesised that these subunits evolved in an ancestral species to confer other functions upon PolB1, with PolB1* evolving as a thermostable complex during the adaptation of certain species to hyperthermophilic environments.

There are many other indices of PolB1 function not measured in this study that may be enhanced by the formation of PolB1*, thus allowing the protein to operate optimally *in vivo* despite an apparent decrease in some of its activities *in vitro*. It has been shown that strand displacement by PolB1 is a key activity in the process of Okazaki fragment maturation (see Chapter 3), and therefore it is possible that PolB1* may display different activity in this respect, optimising its coordination with Fen1. Furthermore, PCNA is clearly an important cofactor for PolB1 activity, and therefore it will be interesting to test the function of PolB1* in

the presence of PCNA as this may reveal the functional importance of the two additional subunits in a more physiological context.

The fidelity of DNA synthesis is another key DNA polymerase property not measured in this study, and it may be that the association of PolB1 with PBP1 and PBP2 sacrifices basal polymerase activity for increased fidelity essential for genome maintenance. Interestingly, kinetic analysis of *S. solfataricus* PolB1 *in vitro* suggested a fidelity of 10^{-6} – 10^{-8} , which is estimated to be approximately 100-fold lower than required to give rise to the actual base substitution rate observed in *S. solfataricus* (Zhang et al, 2009a). This raises the intriguing possibility that additional mechanisms may indeed be required to enhance the fidelity of PolB1, although it is likely that post-replicative repair pathways operating *in vivo* contribute the majority of this additional fidelity.

Alternatively, PBP1 and PBP2 could play a key role in mediating interactions between PolB1 and other DNA replication proteins, and thus the true functional effect of PolB1* assembly may be difficult to measure outside of the context of an active replication fork. Native purification of PolB1 was undertaken partly to identify any interactions between PolB1 and known DNA replication proteins. No such interactions were observed, although it is likely that only high affinity interactions would persist throughout the protocol employed. Furthermore, no interactions were observed between PolB1, PBP1, or PBP2 and components of the GINS complex by yeast two-hybrid assay. However, these tests were directed by an established role for GINS at the core of the *S. solfataricus* replisome (Marinsek et al, 2006) together with the availability of existing reagents, and therefore do not represent an exhaustive screen. To further test the hypothesis that these newly identified subunits facilitate physical coordination of PolB1 with other components of the replisome, future work could use an *S. solfataricus* yeast two-hybrid genomic library (Dionne & Bell, 2005) to perform an unbiased screen for potential interaction partners of PBP1 and PBP2.

One feature of PBP1 that is apparent from sequence analysis is the presence of a highly acidic C-terminal tail. *S. solfataricus* SSB has also been shown to possess an acidic C-terminal tail, and furthermore this region is not required for DNA binding but instead appears to mediate protein-protein interactions, for example with RNA polymerase (Wadsworth & White, 2001; Richard et al, 2004). Interestingly, the magnitude of the negative charge of this motif varies between PBP1 orthologues, and this variation between species is broadly similar among SSB orthologues (Figure 6.13). This potential co-evolution raises the intriguing possibility that this motif facilitates protein-protein interactions of PBP1 and SSB with a shared binding partner, potentially involved in DNA replication. The absence of SSB from the initial native purification suggests this common partner is not simply PolB1.

Another interesting feature of PBP1 is the predicted existence of a zinc finger motif in this protein, potentially extending the use of metal-binding clusters to all archaeal-eukaryotic replicative polymerases. Given that the predicted motif in PBP1 is a C2H2 zinc finger, in contrast to the four cysteine ligands present in all other polymerase metal-binding clusters, it is unlikely that these motifs are evolutionarily related. Nonetheless, a metal binding motif may have evolved multiple times independently to fulfil a common function in different DNA polymerases. The iron-sulphur cluster of eukaryotic polymerases has been shown to be important for assembly of the catalytic subunits with their accessory subunits, and the zinc finger appears to play a functional role in supporting polymerase interaction with PCNA (Netz et al, 2011). Furthermore, the zinc finger of euryarchaeal PolD is important in supporting the interaction between large and small subunits (Shen et al, 2004). Thus it can be speculated that a zinc finger in PBP1 may be important for mediating its interaction with PolB1, or perhaps more intriguingly may facilitate an interaction of PolB1* with additional proteins.

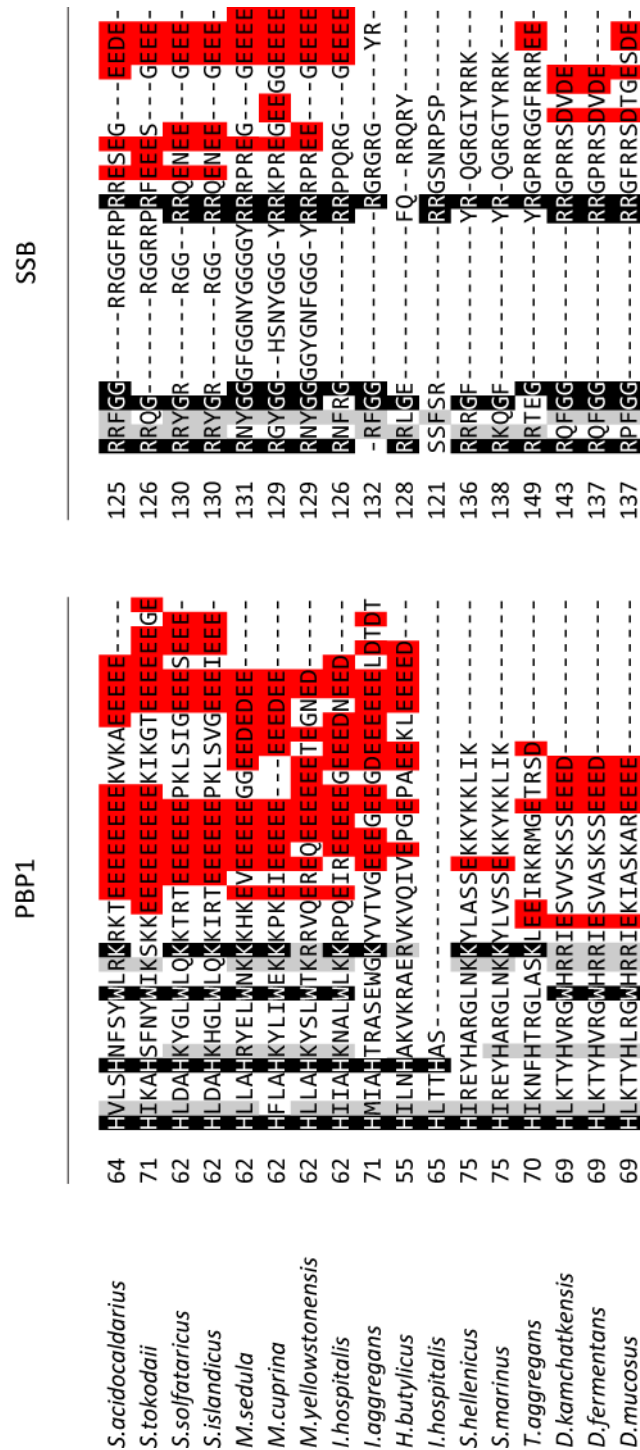


Figure 6.13 – Comparison of the C-terminal tails of PBP1 and SSB orthologues

Multiple sequence alignments of PBP1 (left) and SSB (right) orthologues across the same species, labelled on the left. Note that proteins were aligned along their full length but only C-terminal regions are shown for clarity; the position in each protein sequence of the N-terminal residue displayed is numbered on the left. Residues with identity to consensus are highlighted in black, residues with similarity to consensus are highlighted in grey. Negatively charged residues are highlighted in red.

The presence of multiple B-type polymerases in crenarchaea raises the question of whether these are also able to form complexes with PBP1 and PBP2; *S. solfataricus* PolB2 and PolB3 for example share a reasonable degree of overall homology to PolB1 despite key catalytic amino acid substitutions. In this respect it is interesting that a fourth B-type DNA polymerase present in eukaryotes, the lesion bypass Pol ζ , has recently been shown to require an interaction with the two accessory subunits of Pol δ , Pol31 and Pol32 (Johnson et al, 2012). Thus there is a precedent for distinct but related polymerases using the same accessory factors, and this is an interesting area for future studies in crenarchaea.

Further studies will be required to elucidate the functional role of the novel PolB1-PBP1-PBP2 interaction reported here. In addition to further *in vitro* assays discussed, it will be important to characterise the properties of PBP1 and PBP2 *in vivo*; in this respect improving genetic and microscopy tools in *Sulfolobus* may be particularly useful. Furthermore, the amenability of PolB1* to reconstitution *in vitro* suggests it could be a good candidate for structural studies, potentially revealing a wealth of information on specific interaction interfaces and the influence of these accessory subunits on the global architecture of PolB1.

Chapter 7 – Discussion

MAIN CONCLUSIONS

This thesis describes the biochemical characterisation of DNA replication proteins in the model crenarchaeon *Sulfolobus solfataricus*. In particular, the molecular mechanisms of Okazaki fragment maturation and the molecular architecture of the archaeal replisome were investigated. The resulting conclusions of this work have broad implications not only for DNA replication in the archaea but also for the more complex DNA replication machinery of eukaryotes.

Chapter 3 described the characterisation of *S. solfataricus* Okazaki fragment maturation, which is an essential mechanism for the completion of DNA replication in all cellular organisms. By reconstituting this process *in vitro* from purified components, it was demonstrated that the proteins PolB1, Fen1, Lig1 and heterotrimeric PCNA are necessary and sufficient for fully coupled DNA synthesis, complete RNA primer removal and Okazaki fragment ligation in *S. solfataricus*.

In chapter 4, the role of PCNA in coordinating Okazaki fragment maturation in *S. solfataricus* was investigated. It was demonstrated that the coordination of PolB1, Fen1 and Lig1 enzyme activities was most efficient when all three proteins could simultaneously interact with a single PCNA ring, thus supporting a model whereby PCNA acts as a platform to organise PolB1, Fen1 and Lig1 into a multi-protein maturation complex.

Chapter 5 investigated the structural mechanism by which multiple different enzymes are coordinated around a single PCNA ring. A FRET assay was designed to test the functional relevance of Lig1 conformational flexibility to its role in Okazaki fragment maturation. Protocols and reagents were developed to enable site-specific labelling of Lig1 with two

distinct fluorophores. However, fluorophore labelling was shown to impair Lig1 activity *in vitro*, and thus FRET measurements could not be undertaken.

In Chapter 6, interaction partners of the *Sulfolobus* replicative polymerase, PolB1, were investigated. While no stable interactions between PolB1 and any known components of the DNA replication machinery could be detected, two novel proteins were demonstrated to form a stoichiometric heterotrimeric complex with PolB1. These additional subunits stabilised PolB1 under physiological conditions of elevated temperature but imposed decreased DNA binding and primer extension activities upon PolB1 *in vitro*. It is proposed that the formation of this ternary complex may confer essential activities upon PolB1 in the context of an active replisome *in vivo*.

IMPLICATIONS IN A BROADER CONTEXT

Okazaki fragment maturation

The work presented here demonstrates that *S. solfataricus* employs a simplified machinery compared to eukaryotes for processing Okazaki fragments. It is likely that this can be extended to all archaeal species as DNA polymerase, Fen1, Lig1 and PCNA orthologues are conserved throughout archaea while Dna2 and Pif1, which have been proposed to play a role in eukaryotic Okazaki fragment maturation, are not. Okazaki fragment maturation in euryarchaea is anticipated to exhibit some key differences compared to the crenarchaea however. A D-family DNA polymerase, rather than a PolB1 orthologue, is likely to be employed in maturation, given its characterised ability to extend RNA primers and perform strand displacement synthesis (Henneke et al, 2005). Furthermore, euryarchaeal PCNA is homotrimeric rather than heterotrimeric, but nonetheless has still been shown to interact with Fen1 and Lig1 orthologues, together with PolD, and is therefore likely to play a similar role in

coordinating these enzyme activities (Chapados et al, 2004; Henneke et al, 2005; Kiyonari et al, 2006).

The existence of additional back-up Okazaki fragment maturation pathways in eukaryotes may reflect an increased workload imposed on the maturation machinery; one round of chromosome replication in *S. solfataricus* is calculated to give rise to approximately 30000 Okazaki fragments, while in human cells as many as 50 million Okazaki fragments may be generated (Burgers, 2009). Additionally, redundancy in eukaryotic Okazaki fragment maturation is consistent with the greater complexity of Okazaki fragment structure in eukaryotes: while archaeal Okazaki fragment primers consist only of RNA (Lao-Sirieix & Bell, 2004), eukaryotic primers also contain low-fidelity DNA synthesised by Pol α , which must be removed by the maturation machinery in addition to RNA.

The toolbelt

The work presented in this thesis strongly supports a toolbelt mechanism for PCNA function, in which coordination of PolB1, Fen1 and Lig1 activities is achieved through simultaneous binding to a single PCNA ring. In *S. solfataricus* this is achieved through a specific interaction of each enzyme with a distinct subunit of heterotrimeric PCNA. A key question is whether it is possible to extend such a toolbelt mechanism beyond the crenarchaea to bacterial, euryarchaeal and eukaryotic organisms that possess homo-oligomeric DNA sliding clamps. Functional and biophysical studies have demonstrated that simultaneous binding of Pol III and Pol IV to the β clamp in bacteria facilitates transient polymerase switching during bypass of template lesions (Indiani et al, 2005). An interaction of two different factors with homotrimeric PCNA has also been observed *in vitro* in both eukaryotes (Doré et al, 2009) and euryarchaea (Kiyonari et al, 2009).

The mechanisms are unclear however, by which appropriate combinations of different enzymes may be assembled on a single sliding clamp with multiple identical binding sites. It is

possible that this is mediated by interactions between PCNA-binding proteins, although given the proposed conformational flexibility of proteins upon PCNA, it is difficult to envisage how such protein-protein interactions could be stably maintained. Alternatively, assembly may be a stochastic process in which extensive exchange of partner proteins transiently generates appropriate toolbelt complexes. Potential combinations of different factors could be limited by active regulation of interactions with PCNA; a number of post-translational modifications have been identified on both PCNA and PCNA-binding proteins which can modulate the interaction between the two (Moldovan et al, 2007). This pool of exchangeable factors may be further regulated by local protein concentrations at sites of DNA replication *in vivo*.

In contrast, the specificity of PCNA binding in crenarchaea presumably facilitates the formation of more stable toolbelt complexes. Okazaki fragment maturation may therefore be intrinsically more efficient in crenarchaeal species and this could be biologically relevant; in the hyperthermophilic acidophilic environment of *S. solfataricus* for example, DNA damage is likely to be more prevalent, and thus it may be speculated that efficient Okazaki fragment maturation minimises the life-time of vulnerable single-stranded and nicked DNA species. Additionally, given that archaea only require RNA to be removed from the 5' end of Okazaki fragments, rapid ligation of nicked DNA by PCNA-localised Lig1 is clearly advantageous. In contrast, it may be beneficial for eukaryotic Lig1 to be relatively less well coordinated with other components of the Okazaki fragment maturation machinery, to enable the additional removal of Pol α - synthesised DNA prior to Okazaki fragment ligation. Similarly the much greater idling activity observed for *S. cerevisiae* Pol δ compared to *S. solfataricus* PolB1 could reflect a greater dependence on a mechanism to limit strand displacement in eukaryotes, where polymerase and Fen1 may be less robustly coordinated around PCNA.

One possibility suggested by this work is that multistep DNA transactions other than Okazaki fragment maturation could also be mediated by the organisation of multiple factors on PCNA. Several protein complexes are required to facilitate DNA mismatch repair (MMR).

Many of these complexes, such as the mismatch sensors MutS α and MutS β and the accessory complex MutL α , have been shown to interact with PCNA in eukaryotes (Umar et al, 1996). Similarly, the bacterial equivalents of these complexes, MutS and MutL, have also been shown to interact with the β clamp (López de Saro et al, 2006). Multiple base excision repair (BER) factors have also been shown to interact with PCNA in both eukaryotes and archaea (Moldovan et al, 2007). Indeed, in the euryarchaeon *Pyrococcus furiosus* a uracil DNA glycosylase – which recognises and excises damaged bases – and an apurinic/apyrimidinic (AP) endonuclease enzyme – which cleaves the phosphodiester backbone of DNA at the resulting abasic site – have both been shown to interact simultaneously with PCNA (Kiyonari et al, 2009). A uracil DNA glycosylase, UDG1, has also been shown to specifically interact with PCNA3 in *S. solfataricus* (Dionne & Bell, 2005), while *S. solfataricus* also possesses an orthologue of an AP endonuclease characterised in the crenarchaeon *Pyrobaculum aerophilum* (Sartori & Jiricny, 2003). Thus, as with Okazaki fragment maturation, *S. solfataricus* heterotrimeric PCNA could potentially provide a unique opportunity to test the functional relevance of multiple enzyme coordination around PCNA during BER. Given the conservation of factors with eukaryotes, archaea are proposed to solely employ ‘long-patch’ repair – mediated by PolB1, Fen1 and Lig1 – to replace damaged bases following AP endonuclease cleavage (Sartori & Jiricny, 2003). This raises the interesting possibility that BER involves sequential activity by two distinct toolbelt complexes.

The toolbelt model demonstrated here for PCNA may also have implications for the Rad9-Rad1-Hus1 (9-1-1) complex which functions as an essential component of the DNA damage response in eukaryotes. 9-1-1 is a key checkpoint activator and effector of DNA repair pathways; interactions have been demonstrated between 9-1-1 and checkpoint mediators, translesion DNA polymerases, BER and MMR factors (Bai et al, 2010; Eichinger & Jentsch, 2011). Despite only possessing a weak sequence relationship to PCNA, 9-1-1 has been shown

to adopt a similar ring-shaped architecture to the DNA sliding clamps (Figure 7.1A) (Doré et al, 2009; Sohn & Cho, 2009; Xu et al, 2009).

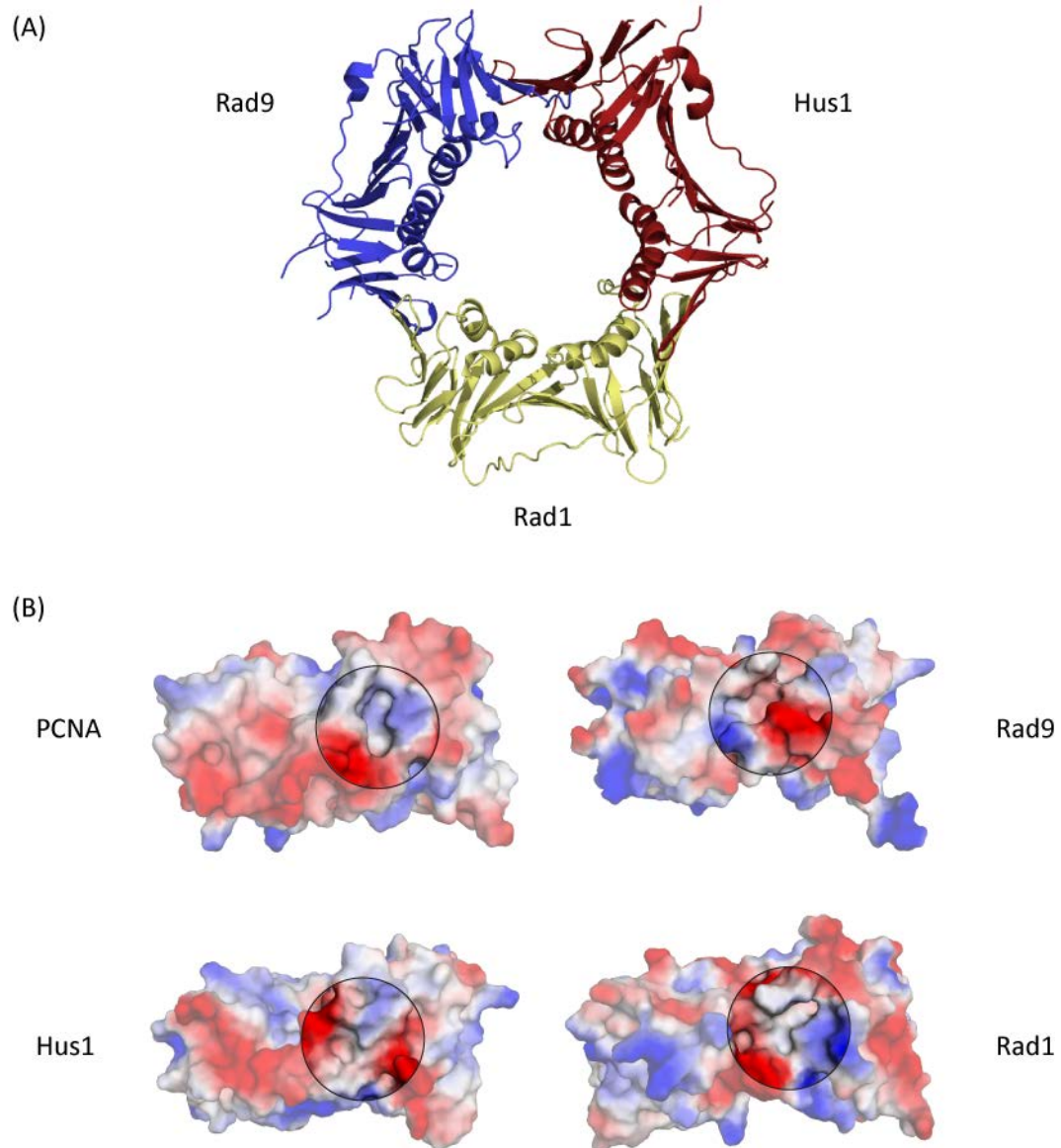


Figure 7.1 – The structure of human Rad9-Rad1-Hus1 (9-1-1)

(A) The complex structure of Rad9 (blue), Rad1 (yellow) and Hus1 (red), generated from PDB file 3G65.

(B) A comparison of the putative protein binding pockets of Rad9, Rad1 and Hus1 with human PCNA (PDB file 1AXC), highlighted by circles. Structures are displayed with electrostatic surface rendering.

Furthermore, this complex has been shown to be loaded onto dsDNA at recessed 5' ends by the specialised Rad17-RFC clamp loader (Bermudez et al, 2003; Majka et al, 2006),

suggesting a similar mechanism of action to PCNA as a protein recruitment platform on DNA. Similarly to PCNA, Rad9, Rad1 and Hus1 all possess a hydrophobic surface pocket formed by an IDCL, but interestingly each site possesses distinct shape and charge characteristics, reminiscent of the three distinct protein binding sites present in *S. solfataricus* PCNA (Figure 7.1B). Thus, although the molecular basis of interactions between 9-1-1 and its binding partners is still poorly defined, it has been hypothesised that different proteins may exhibit preferential binding to different subunits of the 9-1-1 ring; a peptide encompassing the PIP motif of human Fen1 for example has been demonstrated to possess different binding affinities for Rad9, Rad1 and Hus1 (Xu et al, 2009). It is possible therefore that similarly to *S. solfataricus* PCNA, 9-1-1 organises multiple distinct enzymes into productive complexes to facilitate coordinated DNA damage responses. Indeed, it has been proposed from biochemical analysis of BER that the necessary enzymes Pol β , Fen1 and Lig1 are coordinated by a single 9-1-1 ring (Balakrishnan et al, 2009).

DNA polymerases and replisome architecture

A second key finding of this thesis was the identification of two previously uncharacterised PolB1-interacting proteins, PBP1 and PBP2. As a relatively simple, biochemically amenable single-subunit DNA polymerase, *S. solfataricus* PolB1 has been used as a model for archaeal B-type replicative polymerases in numerous *in vitro* studies. The kinetic mechanism (Brown & Suo, 2009), fidelity (Zhang et al, 2009a), uracil-sensing activity (Gruz et al, 2003), structure (Savino et al, 2004) and assembly on DNA (Mikheikin et al, 2009; Lin et al, 2012) of PolB1 have all been characterised in detail. The data here support the formation of a constitutive stoichiometric complex *in vivo* between PolB1, PBP1 and PBP2, termed PolB1*. These previous studies may therefore have to be reassessed to account for this physiologically relevant form of the enzyme, particularly given that the additional subunits appear inhibitory to some PolB1 activities *in vitro*. Similarly, it will be interesting to determine the role of PolB1*

in Okazaki fragment maturation given the essential role for PolB1 demonstrated here. Subunit conservation suggests the PolB1* architecture is common to the majority of crenarchaeal species; whether this can be extended to the thermoproteales and euryarchaeal species may have to be determined empirically.

This study failed to detect any interactions between PolB1 and known components of the *S. solfataricus* DNA replication machinery, suggesting that this replicative polymerase does not stably interact with the core replisome, in agreement with studies in eukaryotes (Gambus et al, 2006). However, it cannot be concluded from the current evidence whether replicative polymerases act in isolation from other components of the replisome or simply associate through transient interactions; an interaction between PolB1 and PCNA2 was not detected during native purification for example, despite a physical and functional interaction between these proteins being well characterised. Given the complexities of coordinating simultaneous leading and lagging strand polymerase activities – in the case of eukaryotes mediated by three distinct multimeric assemblies – it seems likely that some degree of physical coordination with other key activities at the replication fork exists, as observed in bacteria. This remains to be demonstrated experimentally however. In this respect the DNA replication machinery of *S. solfataricus*, with its simplified complement of proteins, may serve as a useful model system in future studies for determining functionally important protein-protein interactions at the replication fork.

LIMITATIONS AND FUTURE WORK

The functional data presented here are strongly supportive of a stoichiometric complex between PCNA, PolB1, Fen1 and Lig1 being required for optimal Okazaki fragment maturation, and this is in agreement with previous evidence demonstrating that PCNA can indeed act as a bridge between these three factors (Dionne et al, 2003). However, such a complex has not

been directly visualised. Attempts to isolate and characterise this complex by DNA pull-down or gel filtration were unsuccessful. However, it is unlikely that these proteins form a highly stable complex – as required for isolation by these techniques – given that the exchange of PolB1, Fen1 or Lig1 with additional factors is likely to be important for PCNA function *in vivo*.

In collaboration with Laura Spagnolo (University of Edinburgh), complexes of PCNA, PolB1, Fen1 and Lig1 in the presence of a DNA flap substrate have now been isolated by gel filtration following chemical crosslinking, and visualised by negative stain electron microscopy and single particle analysis (Figure 7.2). Preliminary analysis suggests that PolB1, Fen1 and Lig1 can indeed all simultaneously associate with a single heterotrimeric PCNA ring. Fen1 appears to engage the substrate while PolB1 and Lig1 are positioned away from DNA passing through the central PCNA pore; interestingly, Lig1 adopts a position on the ‘back’ face of PCNA, away from PolB1 and Fen1. It should be noted that this structure is in the early stages of refinement but nonetheless currently supports the toolbelt mechanism for PCNA function. It is hoped that this structure will provide a useful structural basis for understanding multiple factor coordination by PCNA in the future.

The *in vitro* system established here is a powerful tool for establishing the molecular mechanisms underlying Okazaki fragment maturation. Notably however, the factors tested in this assay were all chosen on the basis of homology to proteins previously identified in eukaryotic studies. Thus it is not possible to rule out the involvement in this process of additional archaeal-specific factors. While considerable progress has been made in recent years, genetic tools in *Sulfolobus* are still relatively primitive compared to those employed in bacteria or model eukaryotic species such as *S. cerevisiae* (Wagner et al, 2009b), and therefore it is still not possible to perform genetic assays in *Sulfolobus* similar to those which enabled the identification of eukaryotic Okazaki fragment maturation proteins. In this respect, the *in vivo* tagging methodology employed in Chapter 6 to identify PBP1 and PBP2 could be important in future studies for identifying additional interaction partners of Okazaki fragment maturation

proteins, and indeed other components of the archaeal replisome. One possibility for improving the sensitivity of such studies in future could be to combine native purification with chemical crosslinking agents to preserve weak or transient interactions.

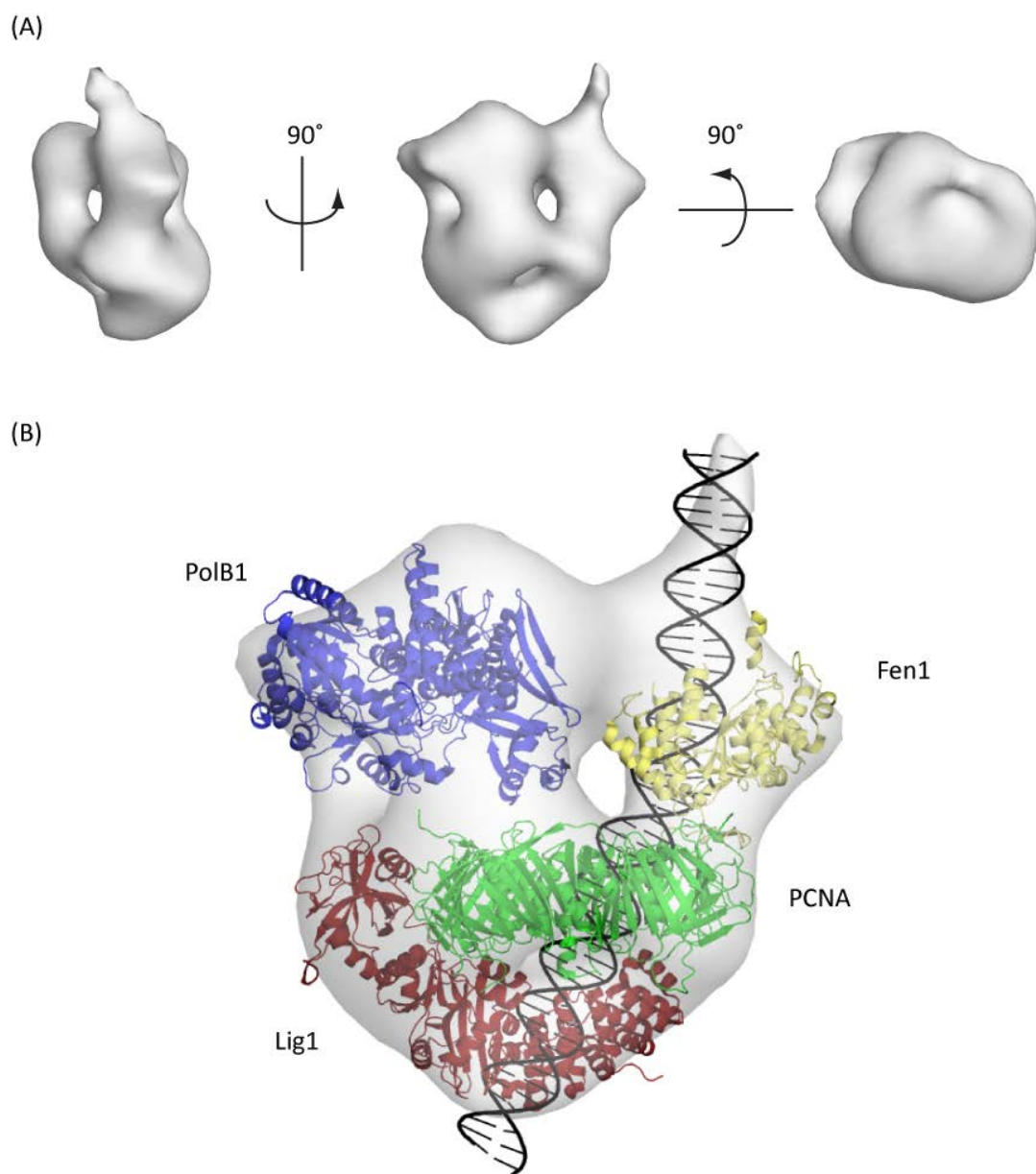


Figure 7.2 – The structure of the *S. solfataricus* DNA-PCNA-PolB1-Fen1-Lig1 complex

(A) 3D reconstructions from single particle analysis of purified DNA-PCNA-PolB1-Fen1-Lig1, displayed with surface rendering. Complex preparation, image acquisition and image processing were performed by Giuseppe Cannone and Laura Spagnolo.

(B) An atomic model of DNA-PCNA-PolB1-Fen1-Lig1. Fitting of crystal structures was performed by Laura Spagnolo using PDB files 1UL1, 2VWJ and 2HIV.

Furthermore, with only limited genetic tools available it is not possible to unambiguously conclude that the importance of PolB1, Fen1 and Lig1 determined *in vitro* is the same *in vivo*. With improving genetic tools, such as the increasing efficiency with which gene deletion or replacement can be performed in *Sulfolobus* species, it is hoped that future studies may be directed towards confirming *in vivo* the molecular mechanisms determined here. Nonetheless the complete maturation activity displayed by PolB1, Fen1 and Lig1, together with their evolutionary relationship to the core eukaryotic maturation factors, does suggest these proteins are the primary enzymes responsible for Okazaki fragment maturation in *S. solfataricus*. Similarly, genetic analysis will be required in future studies to investigate the *in vivo* function of the PBP1 and PBP2 proteins identified in this study, in particular the *in vivo* relevance of their observed interaction with PolB1.

The linear substrate employed here, while providing an excellent model on which to analyse the final stages of Okazaki fragment maturation, is not able to provide information on many other aspects of lagging strand DNA replication. By adapting the system developed here, future studies could potentially investigate the full complement of activities required for the complete synthesis and maturation of Okazaki fragments. For example, use of more physiological closed circular DNA substrates would introduce a dependence on PCNA loading by the RFC clamp loader (Dionne et al, 2008). Given that the mechanism and timing of PCNA unloading and recycling in archaea and eukaryotes is still poorly understood, such an *in vitro* system could enable high resolution analysis of this process in relation to the completion of Okazaki fragment maturation. Additionally, previous studies have demonstrated a significant influence of SSB (Choi et al, 2011) and the chromatin protein Sul7 (Lou et al, 2004) on PolB1 activity, while the effect of other chromatin proteins, such as Alba, has not been determined. Thus on a larger substrate the full influence of a more physiological chromatin environment on the Okazaki fragment maturation machinery could be tested.

A key advance in bacterial and bacteriophage DNA replication studies has been the reconstitution of chromosomal DNA replication *in vitro*; thus, using a minicircle template, fully coupled leading and lagging strand synthesis has been demonstrated *in vitro* for proteins from *E. coli* (Wu et al, 1992), *B. subtilis* (Sanders et al, 2010) and bacteriophages T4 (Salinas & Benkovic, 2000) and T7 (Lee et al, 1998). The establishment of such systems has enabled powerful biochemical analysis of all aspects of the DNA replication mechanism. In particular, the development of single molecule techniques has facilitated analysis of these machineries at previously unprecedented resolution, allowing determination of the mechanism of polymerase and β clamp recycling between Okazaki fragments (Hamdan et al, 2009; Tanner et al, 2011), the identification of a previously uncharacterised priming loop formed during the initiation of each Okazaki fragment (Manosas et al, 2009; Pandey et al, 2009), and the determination of the role of a third polymerase present in the *E. coli* replisome (Georgescu et al, 2011). Despite the huge potential of such *in vitro* systems, the much greater complexity of the eukaryotic DNA replication machinery has thus far prevented equivalent studies from being developed. In this respect *S. solfataricus*, with its simplified complement of eukaryotic-like DNA replication factors, is an excellent potential model system in which to develop reconstituted systems, as demonstrated by the Okazaki fragment maturation analysis presented here.

References

- Acharya N, Klassen R, Johnson RE, Prakash L, Prakash S (2011) PCNA binding domains in all three subunits of yeast DNA polymerase delta modulate its function in DNA replication. *Proc Natl Acad Sci USA* **108**: 17927-17932
- Alabert C, Groth A (2012) Chromatin replication and epigenome maintenance. *Nat Rev Mol Cell Biol* **13**: 153-167
- Aravind L, Koonin EV (1998) Phosphoesterase domains associated with DNA polymerases of diverse origins. *Nucleic Acids Res* **26**: 3746-3752
- Avery OT, Macleod CM, McCarty M (1944) Studies on the chemical nature of the substance inducing transformation of Pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus type iii. *J Exp Med* **79**: 137-158
- Ayyagari R, Gomes XV, Gordenin DA, Burgers PM (2003) Okazaki fragment maturation in yeast. I. Distribution of functions between FEN1 AND DNA2. *J Biol Chem* **278**: 1618-1625
- Bae SH, Bae KH, Kim JA, Seo YS (2001) RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes. *Nature* **412**: 456-461
- Bae SH, Seo YS (2000) Characterization of the enzymatic properties of the yeast dna2 helicase/endonuclease suggests a new model for Okazaki fragment processing. *J Biol Chem* **275**: 38022-38031
- Bai H, Madabushi A, Guan X, Lu AL (2010) Interaction between human mismatch repair recognition proteins and checkpoint sensor Rad9-Rad1-Hus1. *DNA Repair* **9**: 478-487
- Balakrishnan L, Brandt PD, Lindsey-Boltz LA, Sancar A, Bambara RA (2009) Long patch base excision repair proceeds via coordinated stimulation of the multienzyme DNA repair complex. *J Biol Chem* **284**: 15158-15172
- Balakrishnan L, Stewart JA, Polaczek P, Campbell JL, Bambara RA (2010) Acetylation of Dna2 endonuclease/helicase and flap endonuclease 1 by p300 promotes DNA stability by creating long flap intermediates. *J Biol Chem* **285**: 4398-4404
- Barns SM, Delwiche CF, Palmer JD, Pace NR (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc Natl Acad Sci USA* **93**: 9188-9193
- Barry ER, Bell SD (2006) DNA replication in the archaea. *Microbiol Mol Biol Rev* **70**: 876-887
- Bauer RJ, Begley MT, Trakselis MA (2012) Kinetics and fidelity of polymerization by DNA polymerase III from *Sulfolobus solfataricus*. *Biochemistry* **51**: 1996-2007
- Bebenek K, Kunkel TA (2004) Functions of DNA polymerases. *Adv Protein Chem* **69**: 137-165

- Bell SP, Dutta A (2002) DNA replication in eukaryotic cells. *Annu Rev Biochem* **71**: 333-374
- Benkovic SJ, Valentine AM, Salinas F (2001) Replisome-mediated DNA replication. *Annu Rev Biochem* **70**: 181-208
- Bermudez VP, Lindsey-Boltz LA, Cesare AJ, Maniwa Y, Griffith JD, Hurwitz J, Sancar A (2003) Loading of the human 9-1-1 checkpoint complex onto DNA by the checkpoint clamp loader hRad17-replication factor C complex in vitro. *Proc Natl Acad Sci USA* **100**: 1633-1638
- Berquist BR, DasSarma P, DasSarma S (2007) Essential and non-essential DNA replication genes in the model halophilic Archaeon, *Halobacterium* sp. NRC-1. *BMC genet* **8**: 31
- Bessman MJ, Lehman IR, Simms ES, Kornberg A (1958) Enzymatic synthesis of deoxyribonucleic acid. II. General properties of the reaction. *J Biol Chem* **233**: 171-177
- Bienko M, Green CM, Crosetto N, Rudolf F, Zapart G, Coull B, Kannouche P, Wider G, Peter M, Lehmann AR, Hofmann K, Dikic I (2005) Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science* **310**: 1821-1824
- Bochman ML, Sabouri N, Zakian VA (2010) Unwinding the functions of the Pif1 family helicases. *DNA Repair* **9**: 237-249
- Boudsocq F, Iwai S, Hanaoka F, Woodgate R (2001) *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol eta. *Nucleic Acids Res* **29**: 4607-4616
- Bowman GD, O'Donnell M, Kuriyan J (2004) Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature* **429**: 724-730
- Branzei D, Foiani M (2010) Maintaining genome stability at the replication fork. *Nat Rev Mol Cell Biol* **11**: 208-219
- Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol* **6**: 245-252
- Brochier-Armanet C, Forterre P, Gribaldo S (2011) Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr Opin Microbiol* **14**: 274-281
- Brochier C, Gribaldo S, Zivanovic Y, Confalonieri F, Forterre P (2005) Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol* **6**: R42
- Brown JA, Suo Z (2009) Elucidating the kinetic mechanism of DNA polymerization catalyzed by *Sulfolobus solfataricus* P2 DNA polymerase B1. *Biochemistry* **48**: 7502-7511
- Bubeck D, Reijns MAM, Graham SC, Astell KR, Jones EY, Jackson AP (2011) PCNA directs type 2 RNase H activity on DNA replication and repair substrates. *Nucleic Acids Res* **39**: 3652-3666

- Budd ME, Antoshechkin IA, Reis C, Wold BJ, Campbell JL (2011) Inviability of a DNA2 deletion mutant is due to the DNA damage checkpoint. *Cell cycle* **10**: 1690-1698
- Budd ME, Campbell JL (1995) A yeast gene required for DNA replication encodes a protein with homology to DNA helicases. *Proc Natl Acad Sci USA* **92**: 7642-7646
- Budd ME, Campbell JL (1997) A yeast replicative helicase, Dna2 helicase, interacts with yeast FEN-1 nuclease in carrying out its essential function. *Mol Cell Biol* **17**: 2136-2142
- Budd ME, Reis CC, Smith S, Myung K, Campbell JL (2006) Evidence suggesting that Pif1 helicase functions in DNA replication with the Dna2 helicase/nuclease and DNA polymerase delta. *Mol Cell Biol* **26**: 2490-2500
- Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NS, Venter JC (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**: 1058-1073
- Bunting KA, Roe SM, Pearl LH (2003) Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the beta-clamp. *EMBO J* **22**: 5883-5892
- Burgers PM (2009) Polymerase dynamics at the eukaryotic DNA replication fork. *J Biol Chem* **284**: 4041-4045
- Bylund GO, Burgers PM (2005) Replication protein A-directed unloading of PCNA by the Ctf18 cohesion establishment complex. *Mol Cell Biol* **25**: 5445-5455
- Cann IK, Ishino S, Nomura N, Sako Y, Ishino Y (1999) Two family B DNA polymerases from *Aeropyrum pernix*, an aerobic hyperthermophilic crenarchaeote. *J Bacteriol* **181**: 5984-5992
- Cann IK, Komori K, Toh H, Kanai S, Ishino Y (1998) A heterodimeric DNA polymerase: evidence that members of Euryarchaeota possess a distinct DNA polymerase. *Proc Natl Acad Sci USA* **95**: 14250-14255
- Cejka P, Cannavo E, Polaczek P, Masuda-Sasa T, Pokharel S, Campbell JL, Kowalczykowski SC (2010) DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2. *Nature* **467**: 112-116
- Chapados BR, Hosfield DJ, Han S, Qiu J, Yelent B, Shen B, Tainer JA (2004) Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell* **116**: 39-50
- Chen L, Brugger K, Skovgaard M, Redder P, She Q, Torarinsson E, Greve B, Awayez M, Zibat A, Klenk HP, Garrett RA (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. *J Bacteriol* **187**: 4992-4999
- Chia N, Cann IKO, Olsen GJ (2010) Evolution of DNA replication protein complexes in eukaryotes and archaea. *PLoS One* **5**: e10866

- Chilkova O, Stenlund P, Isoz I, Stith CM, Grabowski P, Lundström E-B, Burgers PM, Johansson E (2007) The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. *Nucleic Acids Res* **35**: 6588-6597
- Choi J-Y, Eoff RL, Pence MG, Wang J, Martin MV, Kim E-J, Folkmann LM, Guengerich FP (2011) Roles of the four DNA polymerases of the crenarchaeon *Sulfolobus solfataricus* and accessory proteins in DNA replication. *J Biol Chem* **286**: 31180-31193
- Chon H, Vassilev A, DePamphilis ML, Zhao Y, Zhang J, Burgers PM, Crouch RJ, Cerritelli SM (2009) Contributions of the two accessory subunits, RNASEH2B and RNASEH2C, to the activity and properties of the human RNase H2 complex. *Nucleic Acids Res* **37**: 96-110
- Ciccia A, Nimonkar AV, Hu Y, Hajdu I, Achar YJ, Izhar L, Petit SA, Adamson B, Yoon JC, Kowalczykowski SC, Livingston DM, Haracska L, Elledge SJ (2012) Polyubiquitinated PCNA recruits the ZRANB3 translocase to maintain genomic integrity after replication stress. *Mol Cell* **47**: 396-409
- Cotner-Gohara E, Kim I-K, Hammel M, Tainer JA, Tomkinson AE, Ellenberger T (2010) Human DNA ligase III recognizes DNA ends by dynamic switching between two DNA bound states. *Biochemistry* **49**: 6165-6176
- Creze C, Ligabue A, Laurent S, Lestini R, Laptenok SP, Kuhn J, Vos MH, Czjzek M, Myllykallio H, Flament D (2012) Modulation of the *Pyrococcus abyssi* NucS endonuclease activity by the replication clamp PCNA at functional and structural levels. *J Biol Chem* **287**: 15648-15660
- Crick F (1970) Central dogma of molecular biology. *Nature* **227**: 561-563
- Cubeddu L, White MF (2005) DNA damage detection by an archaeal single-stranded DNA-binding protein. *J Mol Biol* **353**: 507-516
- Dalrymple BP, Kongsuwan K, Wijffels G, Dixon NE, Jennings PA (2001) A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc Natl Acad Sci USA* **98**: 11627-11632
- Demczuk S, Harbers M, Vennstrom B (1993) Identification and analysis of all components of a gel retardation assay by combination with immunoblotting. *Proc Natl Acad Sci USA* **90**: 2574-2578
- Dionne I, Bell SD (2005) Characterization of an archaeal family 4 uracil DNA glycosylase and its interaction with PCNA and chromatin proteins. *Biochem J* **387**: 859-863
- Dionne I, Brown NJ, Woodgate R, Bell SD (2008) On the mechanism of loading the PCNA sliding clamp by RFC. *Mol Microbiol* **68**: 216-222
- Dionne I, Nookala RK, Jackson SP, Doherty AJ, Bell SD (2003) A heterotrimeric PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol Cell* **11**: 275-282
- Dorazi R, Parker JL, White MF (2006) PCNA activates the Holliday junction endonuclease Hjc. *J Mol Biol* **364**: 243-247

- Doré AS, Kilkenny ML, Jones SA, Oliver AW, Roe SM, Bell SD, Pearl LH (2006) Structure of an archaeal PCNA1-PCNA2-FEN1 complex: elucidating PCNA subunit and client enzyme specificity. *Nucleic Acids Res* **34**: 4515-4526
- Doré AS, Kilkenny ML, Rzechorzek NJ, Pearl LH (2009) Crystal structure of the rad9-rad1-hus1 DNA damage checkpoint complex-implications for clamp loading and regulation. *Mol Cell* **34**: 735-745
- Duggin IG, McCallum Sa, Bell SD (2008) Chromosome replication dynamics in the archaeon *Sulfolobus acidocaldarius*. *Proc Natl Acad Sci USA* **105**: 16737-16742
- Duxin JP, Dao B, Martinsson P, Rajala N, Guittat L, Campbell JL, Spelbrink JN, Stewart SA (2009) Human Dna2 is a nuclear and mitochondrial DNA maintenance protein. *Mol Cell Biol* **29**: 4274-4282
- Duxin JP, Moore HR, Sidorova J, Karanja K, Honaker Y, Dao B, Piwnica-Worms H, Campbell JL, Monnat RJ, Stewart SA (2012) An Okazaki fragment processing-independent role for human Dna2 during DNA replication. *J Biol Chem* **287**: 21980-21991
- Edenberg HJ, Huberman JA (1975) Eukaryotic chromosome replication. *Annu Rev Genet* **9**: 245-284
- Edgell DR, Doolittle WF (1997) Archaea and the origin(s) of DNA replication proteins. *Cell* **89**: 995-998
- Edgell DR, Klenk HP, Doolittle WF (1997) Gene duplications in evolution of archaeal family B DNA polymerases. *J Bacteriol* **179**: 2632-2640
- Eichinger CS, Jentsch S (2011) 9-1-1: PCNA's specialized cousin. *Trends Biochem Sci* **36**: 563-568
- Elkins JG, Podar M, Graham DE, Makarova KS, Wolf Y, Randau L, Hedlund BP, Brochier-Armanet C, Kunin V, Anderson I, Lapidus A, Goltsman E, Barry K, Koonin EV, Hugenholtz P, Kyrpides N, Wanner G, Richardson P, Keller M, Stetter KO (2008) A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc Natl Acad Sci USA* **105**: 8102-8107
- Fox GE, Magrum LJ, Balch WE, Wolfe RS, Woese CR (1977) Classification of methanogenic bacteria by 16S ribosomal RNA characterization. *Proc Natl Acad Sci USA* **74**: 4537-4541
- Fridman Y, Palgi N, Dovrat D, Ben-Aroya S, Hieter P, Aharoni A (2010) Subtle alterations in PCNA-partner interactions severely impair DNA replication and repair. *PLoS Biol* **8**: e1000507
- Funnell BE, Baker TA, Kornberg A (1986) Complete enzymatic replication of plasmids containing the origin of the *Escherichia coli* chromosome. *J Biol Chem* **261**: 5616-5624
- Gambus A, Jones RC, Sanchez-Diaz A, Kanemaki M, van Deursen F, Edmondson RD, Labib K (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol* **8**: 358-366

- Gambus A, van Deursen F, Polychronopoulos D, Foltman M, Jones RC, Edmondson RD, Calzada A, Labib K (2009) A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. *EMBO J* **28**: 2992-3004
- Garg P, Stith CM, Sabouri N, Johansson E, Burgers PM (2004) Idling by DNA polymerase delta maintains a ligatable nick during lagging-strand DNA replication. *Genes Dev* **18**: 2764-2773
- Georgescu RE, Kim S-S, Yurieva O, Kuriyan J, Kong X-P, O'Donnell M (2008) Structure of a sliding clamp on DNA. *Cell* **132**: 43-54
- Georgescu RE, Kurth I, O'Donnell M (2011) Single-molecule studies reveal the function of a third polymerase in the replisome. *Nat Struct Mol Biol* **19**: 113-116
- Georgescu RE, Kurth I, Yao NY, Stewart J, Yurieva O, O'Donnell M (2009) Mechanism of polymerase collision release from sliding clamps on the lagging strand. *EMBO J* **28**: 2981-2991
- Gomes XV, Burgers PM (2000) Two modes of FEN1 binding to PCNA regulated by DNA. *EMBO J* **19**: 3811-3821
- Gristwood T, Duggin IG, Wagner M, Albers SV, Bell SD (2012) The sub-cellular localization of *Sulfolobus* DNA replication. *Nucleic Acids Res* **40**: 5487-5496
- Gruz P, Shimizu M, Pisani FM, De Felice M, Kanke Y, Nohmi T (2003) Processing of DNA lesions by archaeal DNA polymerases from *Sulfolobus solfataricus*. *Nucleic Acids Res* **31**: 4024-4030
- Gulbis JM, 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**: 297-306
- Guo C, Sonoda E, Tang TS, Parker JL, Bielen AB, Takeda S, Ulrich HD, Friedberg EC (2006) REV1 protein interacts with PCNA: significance of the REV1 BRCT domain *in vitro* and *in vivo*. *Mol Cell* **23**: 265-271
- Guo Z, Zheng L, Xu H, Dai H, Zhou M, Pascua MR, Chen QM, Shen B (2010) Methylation of FEN1 suppresses nearby phosphorylation and facilitates PCNA binding. *Nat Chem Biol* **6**: 766-773
- Hamdan SM, Johnson DE, Tanner NA, Lee JB, Qimron U, Tabor S, van Oijen AM, Richardson CC (2007) Dynamic DNA helicase-DNA polymerase interactions assure processive replication fork movement. *Mol Cell* **27**: 539-549
- Hamdan SM, Loparo JJ, Takahashi M, Richardson CC, van Oijen AM (2009) Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature* **457**: 336-339
- Hamdan SM, Richardson CC (2009) Motors, switches, and contacts in the replisome. *Annu Rev Biochem* **78**: 205-243

- Haseltine CA, Kowalczykowski SC (2002) A distinctive single-strand DNA-binding protein from the Archaeon *Sulfolobus solfataricus*. *Mol Microbiol* **43**: 1505-1515
- Heltzel JMH, Maul RW, Scouten Ponticelli SK, Sutton MD (2009) A model for DNA polymerase switching involving a single cleft and the rim of the sliding clamp. *Proc Natl Acad Sci USA* **106**: 12664-12669
- Henneke G (2012) *In vitro* reconstitution of RNA primer removal in archaea reveals the existence of two pathways. *Biochem J* **447**: 271-280
- Henneke G, Flament D, Hübscher U, Querellou J, Raffin J-P (2005) The hyperthermophilic euryarchaeota *Pyrococcus abyssi* likely requires the two DNA polymerases D and B for DNA replication. *J Mol Biol* **350**: 53-64
- Henneke G, Koundrioukoff S, Hübscher U (2003) Phosphorylation of human Fen1 by cyclin-dependent kinase modulates its role in replication fork regulation. *Oncogene* **22**: 4301-4313
- Hershey AD, Chase M (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* **36**: 39-56
- Higashibata H, Kikuchi H, Kawarabayasi Y, Matsui I (2003) Helicase and nuclease activities of hyperthermophile *Pyrococcus horikoshii* Dna2 inhibited by substrates with RNA segments at 5'-end. *J Biol Chem* **278**: 15983-15990
- Hlinkova V, Xing G, Bauer J, Shin YJ, Dionne I, Rajashankar KR, Bell SD, Ling H (2008) Structures of monomeric, dimeric and trimeric PCNA: PCNA-ring assembly and opening. *Acta Crystallogr Sect D Biol Crystallogr* **64**: 941-949
- Hoege C, Pfander B, Moldovan G-L, Pyrowolakis G, Jentsch S (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* **419**: 135-141
- Hu J, Sun L, Shen F, Chen Y, Hua Y, Liu Y, Zhang M, Hu Y, Wang Q, Xu W, Sun F, Ji J, Murray JM, Carr AM, Kong D (2012) The intra-S phase checkpoint targets Dna2 to prevent stalled replication forks from reversing. *Cell* **149**: 1221-1232
- Humbard MA, Miranda HV, Lim J-m, Krause DJ, Pritz JR, Zhou G, Chen S, Wells L, Maupin-Furlow JA (2010) Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. *Nature* **463**: 54-60
- Hutton RD, Craggs TD, White MF, Penedo JC (2010) PCNA and XPF cooperate to distort DNA substrates. *Nucleic Acids Res* **38**: 1664-1675
- Hutton RD, Roberts JA, Penedo JC, White MF (2008) PCNA stimulates catalysis by structure-specific nucleases using two distinct mechanisms: substrate targeting and catalytic step. *Nucleic Acids Res* **36**: 6720-6727
- Imamura K, Fukunaga K, Kawarabayasi Y, Ishino Y (2007) Specific interactions of three proliferating cell nuclear antigens with replication-related proteins in *Aeropyrum pernix*. *Mol Microbiol* **64**: 308-318

- Indiani C, McInerney P, Georgescu RE, Goodman MF, O'Donnell M (2005) A sliding-clamp toolbelt binds high- and low-fidelity DNA polymerases simultaneously. *Mol Cell* **19**: 805-815
- Indiani C, O'Donnell M (2006) The replication clamp-loading machine at work in the three domains of life. *Nat Rev Mol Cell Biol* **7**: 751-761
- Ishmael FT, Trakselis MA, Benkovic SJ (2003) Protein-protein interactions in the bacteriophage T4 replisome. The leading strand holoenzyme is physically linked to the lagging strand holoenzyme and the primosome. *J Biol Chem* **278**: 3145-3152
- Iwai T, Kurosawa N, Itoh YH, Kimura N, Horiuchi T (2000) Sequence analysis of three family B DNA polymerases from the thermoacidophilic crenarchaeon *Sulfolobus solfataricus*. *DNA Res* **7**: 243-251
- Jeruzalmi D, Yurieva O, Zhao Y, Young M, Stewart J, Hingorani MM, O'Donnell M, Kuriyan J (2001) Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III. *Cell* **106**: 417-428
- Johansson E, Macneill SA (2010) The eukaryotic replicative DNA polymerases take shape. *Trends Biochem Sci* **35**: 339-347
- Johnson A, O'Donnell M (2005) Cellular DNA replicases: components and dynamics at the replication fork. *Annu Rev Biochem* **74**: 283-315
- Johnson RE, Prakash L, Prakash S (2012) Pol31 and Pol32 subunits of yeast DNA polymerase delta are also essential subunits of DNA polymerase zeta. *Proc Natl Acad Sci USA* **109**: 12455-12460
- Jokela M, Eskelinen A, Pospiech H, Rouvinen J, Syvaioja JE (2004) Characterization of the 3' exonuclease subunit DP1 of *Methanococcus jannaschii* replicative DNA polymerase D. *Nucleic Acids Res* **32**: 2430-2440
- Kao H-I, Henricksen LA, Liu Y, Bambara RA (2002) Cleavage specificity of *Saccharomyces cerevisiae* flap endonuclease 1 suggests a double-flap structure as the cellular substrate. *J Biol Chem* **277**: 14379-14389
- Kapanidis AN, Laurence TA, Lee NK, Margeat E, Kong X, Weiss S (2005) Alternating-laser excitation of single molecules. *Acc Chem Res* **38**: 523-533
- Kapanidis AN, Weiss S (2002) Fluorescent probes and bioconjugation chemistries for single-molecule fluorescence analysis of biomolecules. *J Chem Phys* **117**: 10953
- Kazmirski SL, Zhao Y, Bowman GD, O'Donnell M, Kuriyan J (2005) Out-of-plane motions in open sliding clamps: molecular dynamics simulations of eukaryotic and archaeal proliferating cell nuclear antigen. *Proc Natl Acad Sci USA* **102**: 13801-13806
- Kelch BA, Makino DL, O'Donnell M, Kuriyan J (2011) How a DNA polymerase clamp loader opens a sliding clamp. *Science* **334**: 1675-1680

- Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**: 363-371
- Kelly RB, Cozzarelli NR, Deutscher MP, Lehman IR, Kornberg A (1970) Enzymatic synthesis of deoxyribonucleic acid. XXXII. Replication of duplex deoxyribonucleic acid by polymerase at a single strand break. *J Biol Chem* **245**: 39-45
- Kelman Z, Yuzhakov A, Andjelkovic J, O'Donnell M (1998) Devoted to the lagging strand-the subunit of DNA polymerase III holoenzyme contacts SSB to promote processive elongation and sliding clamp assembly. *EMBO J* **17**: 2436-2449
- Kerr ID, Wadsworth RIM, Cubeddu L, Blankenfeldt W, Naismith JH, White MF (2003) Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of *Sulfolobus* SSB protein. *EMBO J* **22**: 2561-2570
- Kilkenny ML, De Piccoli G, Perera RL, Labib K, Pellegrini L (2012) A conserved motif in the C-terminal tail of DNA polymerase alpha tethers primase to the eukaryotic replisome. *J Biol Chem* **287**: 23740-23747
- Kim DJ, Kim O, Kim HW, Kim HS, Lee SJ, Suh SW (2009) ATP-dependent DNA ligase from *Archaeoglobus fulgidus* displays a tightly closed conformation. *Acta Crystallogr F Struct Biol Cryst Commun* **65**: 544-550
- Kiyonari S, Tahara S, Shirai T, Iwai S, Ishino S, Ishino Y (2009) Biochemical properties and base excision repair complex formation of apurinic/apyrimidinic endonuclease from *Pyrococcus furiosus*. *Nucleic Acids Res* **37**: 6439-6453
- Kiyonari S, Takayama K, Nishida H, Ishino Y (2006) Identification of a novel binding motif in *Pyrococcus furiosus* DNA ligase for the functional interaction with proliferating cell nuclear antigen. *J Biol Chem* **281**: 28023-28032
- Kletzin A (2007) General characteristics and important model organisms. In *Archaea: Molecular and Cellular Biology*, Cavicchioli R (ed), pp 14-92. Washington, DC: ASM Press
- Klimczak LJ, Grummt F, Burger KJ (1985) Purification and characterization of DNA polymerase from the archaebacterium *Sulfolobus acidocaldarius*. *Nucleic Acids Res* **13**: 5269-5282
- Klinge S, Núñez-Ramírez R, Llorca O, Pellegrini L (2009) 3D architecture of DNA Pol alpha reveals the functional core of multi-subunit replicative polymerases. *EMBO J* **28**: 1978-1987
- Komori K, Ishino Y (2001) Replication protein A in *Pyrococcus furiosus* is involved in homologous DNA recombination. *J Biol Chem* **276**: 25654-25660
- Kong XP, Onrust R, O'Donnell M, Kuriyan J (1992) Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* **69**: 425-437
- Kornberg A, Baker TA (1992) *DNA replication*, 2nd edn. Sausalito, California: University Science Books.

- Krishna SS, Majumdar I, Grishin NV (2003) Structural classification of zinc fingers: survey and summary. *Nucleic Acids Res* **31**: 532-550
- Krishna TS, Kong X-P, Gary S, Burgers PM, Kuriyan J (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79**: 1233-1243
- Kucherlapati M, Yang K, Kuraguchi M, Zhao J, Lia M, Heyer J, Kane MF, Fan K, Russell R, Brown AMC, Kneitz B, Edelmann W, Kolodner RD, Lipkin M, Kucherlapati R (2002) Haploinsufficiency of Flap endonuclease (Fen1) leads to rapid tumor progression. *Proc Natl Acad Sci USA* **99**: 9924-9929
- Kunkel TA, Burgers PM (2008) Dividing the workload at a eukaryotic replication fork. *Trends Cell Biol* **18**: 521-527
- Ladner JE, Pan M, Hurwitz J, Kelman Z (2011) Crystal structures of two active proliferating cell nuclear antigens (PCNAs) encoded by *Thermococcus kodakaraensis*. *Proc Natl Acad Sci USA* **108**: 2711-2716
- Langston LD, O'Donnell M (2008) DNA polymerase delta is highly processive with proliferating cell nuclear antigen and undergoes collision release upon completing DNA. *J Biol Chem* **283**: 29522-29531
- Lao-Sirieix S-H, Bell SD (2004) The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. *J Mol Biol* **344**: 1251-1263
- Lazzaro F, Novarina D, Amara F, Watt DL, Stone JE, Costanzo V, Burgers PM, Kunkel TA, Plevani P, Muzi-Falconi M (2012) RNase H and postreplication repair protect cells from ribonucleotides incorporated in DNA. *Mol Cell* **45**: 99-110
- Le Laz S, Le Goaziou A, Henneke G (2010) Structure-specific nuclease activities of *Pyrococcus abyssi* RNase HII. *J Bacteriol* **192**: 3689-3698
- Lee J, Chastain PD, Kusakabe T, Griffith JD, Richardson CC (1998) Coordinated leading and lagging strand DNA synthesis on a minicircular template. *Mol Cell* **1**: 1001-1010
- Lee JY, Chang C, Song HK, Moon J, Yang JK, Kim HK, Kwon ST, Suh SW (2000) Crystal structure of NAD(+)-dependent DNA ligase: modular architecture and functional implications. *EMBO J* **19**: 1119-1129
- Lehman IR, Bessman MJ, Simms ES, Kornberg A (1958) Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from *Escherichia coli*. *J Biol Chem* **233**: 163-170
- Leu FP, Hingorani MM, Turner J, O'Donnell M (2000) The delta subunit of DNA polymerase III holoenzyme serves as a sliding clamp unloader in *Escherichia coli*. *J Biol Chem* **275**: 34609-34618
- Levin DS, Bai W, Yao N, O'Donnell M, Tomkinson AE (1997) An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining. *Proc Natl Acad Sci USA* **94**: 12863-12868

- Levin DS, McKenna AE, Motycka TA, Matsumoto Y, Tomkinson AE (2000) Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. *Curr Biol* **10**: 919-922
- Li X, Li J, Harrington J, Lieber MR, Burgers PM (1995) Lagging strand DNA synthesis at the eukaryotic replication fork involves binding and stimulation of FEN-1 by proliferating cell nuclear antigen. *J Biol Chem* **270**: 22109-22112
- Li X, Marians KJ (2000) Two distinct triggers for cycling of the lagging strand polymerase at the replication fork. *J Biol Chem* **275**: 34757-34765
- Li Z, Santangelo TJ, Cuboňová L, Reeve JN, Kelman Z (2010) Affinity purification of an archaeal DNA replication protein network. *mBio* **1**: e00221
- Lia G, Michel B, Allemand J-F (2011) Polymerase exchange during Okazaki fragment synthesis observed in living cells. *Science* **335**: 328-331
- Lin HK, Chase SF, Laue TM, Jen-Jacobson L, Trakselis MA (2012) Differential temperature dependent multimeric assemblies of replication and repair polymerases on DNA increase processivity. *Biochemistry* **51**: 7367-7382
- Liu L, Komori K, Ishino S, Bocquier AA, Cann IK, Kohda D, Ishino Y (2001) The archaeal DNA primase: biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. *J Biol Chem* **276**: 45484-45490
- López de Saro FJ, Marinus MG, Modrich P, O'Donnell M (2006) The beta sliding clamp binds to multiple sites within MutL and MutS. *J Biol Chem* **281**: 14340-14349
- López de Saro FJ, O'Donnell M (2001) Interaction of the beta sliding clamp with MutS, ligase, and DNA polymerase I. *Proc Natl Acad Sci USA* **98**: 8376-8380
- Lou H, Duan Z, Huo X, Huang L (2004) Modulation of hyperthermophilic DNA polymerase activity by archaeal chromatin proteins. *J Biol Chem* **279**: 127-132
- Lou H, Komata M, Katou Y, Guan Z, Reis CC, Budd M, Shirahige K, Campbell JL (2008) Mrc1 and DNA polymerase epsilon function together in linking DNA replication and the S phase checkpoint. *Mol Cell* **32**: 106-117
- Lundgren M, Andersson A, Chen L, Nilsson P, Bernander R (2004) Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination. *Proc Natl Acad Sci USA* **101**: 7046-7051
- Lundgren M, Malandrin L, Eriksson S, Huber H, Bernander R (2008) Cell cycle characteristics of crenarchaeota: unity among diversity. *J Bacteriol* **190**: 5362-5367
- Majka J, Binz SK, Wold MS, Burgers PM (2006) Replication protein A directs loading of the DNA damage checkpoint clamp to 5'-DNA junctions. *J Biol Chem* **281**: 27855-27861
- Makiniemi M, Pospiech H, Kilpelainen S, Jokela M, Vihinen M, Syvaoja JE (1999) A novel family of DNA-polymerase-associated B subunits. *Trends Biochem Sci* **24**: 14-16

- Manosas M, Spiering MM, Zhuang Z, Benkovic SJ, Croquette V (2009) Coupling DNA unwinding activity with primer synthesis in the bacteriophage T4 primosome. *Nat Chem Biol* **5**: 904-912
- Marinsek N, Barry ER, Makarova KS, Dionne I, Koonin EV, Bell SD (2006) GINS, a central nexus in the archaeal DNA replication fork. *EMBO rep* **7**: 539-545
- Marsh VL, McGeoch AT, Bell SD (2006) Influence of chromatin and single strand binding proteins on the activity of an archaeal MCM. *J Mol Biol* **357**: 1345-1350
- Matsunaga F, Norais C, Forterre P, Myllykallio H (2003) Identification of short 'eukaryotic' Okazaki fragments synthesized from a prokaryotic replication origin. *EMBO rep* **4**: 154-158
- Mayanagi K, Kiyonari S, Nishida H, Saito M, Kohda D, Ishino Y, Shirai T, Morikawa K (2011) Architecture of the DNA polymerase B-proliferating cell nuclear antigen (PCNA)-DNA ternary complex. *Proc Natl Acad Sci USA* **108**: 1845-1849
- Mayanagi K, Kiyonari S, Saito M, Shirai T, Ishino Y, Morikawa K (2009) Mechanism of replication machinery assembly as revealed by the DNA ligase-PCNA-DNA complex architecture. *Proc Natl Acad Sci USA* **106**: 4647-4652
- McGeoch AT, Bell SD (2008) Extra-chromosomal elements and the evolution of cellular DNA replication machineries. *Nat Rev Mol Cell Biol* **9**: 569-574
- McInerney P, Johnson A, Katz F, O'Donnell M (2007) Characterization of a triple DNA polymerase replisome. *Mol Cell* **27**: 527-538
- McNally R, Bowman GD, Goedken ER, O'Donnell M, Kuriyan J (2010) Analysis of the role of PCNA-DNA contacts during clamp loading. *BMC Struct Biol* **10**: 3
- Mendel G (1865) Versuche uber Pflanzen-Hybriden. *Verh Naturforsch Ver Brunn* **4**: 3-47
- Meselson M, Stahl FW (1958) The replication of DNA in *Escherichia coli*. *Proc Natl Acad Sci USA* **44**: 671-682
- Meslet-Cladière L, Norais C, Kuhn J, Briffotiaux J, Sloostra JW, Ferrari E, Hübscher U, Flament D, Myllykallio H (2007) A novel proteomic approach identifies new interaction partners for Proliferating Cell Nuclear Antigen. *J Mol Biol* **372**: 1137-1148
- Miescher F (1871) Ueber die chemische Zusammensetzung der Eiterzellen. *Medicinischemische Untersuchungen* **4**: 441-460
- Mikheikin AL, Lin H-K, Mehta P, Jen-Jacobson L, Trakselis MA (2009) A trimeric DNA polymerase complex increases the native replication processivity. *Nucleic Acids Res* **37**: 7194-7205
- Miyabe I, Kunkel TA, Carr AM (2011) The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genet* **7**: e1002407

- Miyachi K, Fritzler MJ, Tan EM (1978) Autoantibody to a nuclear antigen in proliferating cells. *J Immunol* **121**: 2228-2234
- Miyata T, Suzuki H, Oyama T, Mayanagi K, Ishino Y, Morikawa K (2005) Open clamp structure in the clamp-loading complex visualized by electron microscopic image analysis. *Proc Natl Acad Sci USA* **102**: 13795-13800
- Moarefi I, Jeruzalmi D, Turner J, O'Donnell M, Kuriyan J (2000) Crystal structure of the DNA polymerase processivity factor of T4 bacteriophage. *J Mol Biol* **296**: 1215-1223
- Moldovan GL, Pfander B, Jentsch S (2006) PCNA controls establishment of sister chromatid cohesion during S phase. *Mol Cell* **23**: 723-732
- Moldovan GL, Pfander B, Jentsch S (2007) PCNA, the maestro of the replication fork. *Cell* **129**: 665-679
- Montecucco A, Rossi R, Levin DS, Gary R, Park MS, Motycka TA, Ciarrocchi G, Villa A, Biamonti G, Tomkinson AE (1998) DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: identification of a common targeting mechanism for the assembly of replication factories. *EMBO J* **17**: 3786-3795
- Morris GF, Mathews MB (1989) Regulation of proliferating cell nuclear antigen during the cell cycle. *J Biol Chem* **264**: 13856-13864
- Motz M, Kober I, Girardot C, Loeser E, Bauer U, Albers M, Moeckel G, Minch E, Voss H, Kilger C, Koegl M (2002) Elucidation of an archaeal replication protein network to generate enhanced PCR enzymes. *J Biol Chem* **277**: 16179-16188
- Muramatsu S, Hirai K, Tak Y-S, Kamimura Y, Araki H (2010) CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol epsilon, and GINS in budding yeast. *Genes Dev* **24**: 602-612
- Myllykallio H, Lopez P, Lopez-Garcia P, Heilig R, Saurin W, Zivanovic Y, Philippe H, Forterre P (2000) Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science* **288**: 2212-2215
- Nair PA, Nandakumar J, Smith P, Odell M, Lima CD, Shuman S (2007) Structural basis for nick recognition by a minimal pluripotent DNA ligase. *Nat Struct Mol Biol* **14**: 770-778
- Nandakumar J, Nair PA, Shuman S (2007) Last stop on the road to repair: structure of *E. coli* DNA ligase bound to nicked DNA-adenylate. *Mol Cell* **26**: 257-271
- Napoli A, Valenti A, Salerno V, Nadal M, Garnier F, Rossi M, Ciaramella M (2005) Functional interaction of reverse gyrase with single-strand binding protein of the archaeon *Sulfolobus*. *Nucleic Acids Res* **33**: 564-576
- Nethanel T, Reisfeld S, Dinter-Gottlieb G, Kaufmann G (1988) An Okazaki piece of simian virus 40 may be synthesized by ligation of shorter precursor chains. *J Virol* **62**: 2867-2873

- Netz DJ, Stith CM, Stümpfig M, Köpf G, Vogel D, Genau HM, Stodola JL, Lill R, Burgers PMJ, Pierik AJ (2011) Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nat Chem Biol* **8**: 125-132
- Nick McElhinny SA, Gordenin DA, Stith CM, Burgers PM, Kunkel TA (2008) Division of labor at the eukaryotic replication fork. *Mol Cell* **30**: 137-144
- Nick McElhinny SA, Watts BE, Kumar D, Watt DL, Lundström E-B, Burgers PM, Johansson E, Chabes A, Kunkel TA (2010) Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc Natl Acad Sci USA* **107**: 4949-4954
- Nishida H, Kiyonari S, Ishino Y, Morikawa K (2006) The closed structure of an archaeal DNA ligase from *Pyrococcus furiosus*. *J Mol Biol* **360**: 956-967
- Nishida H, Mayanagi K, Kiyonari S, Sato Y, Oyama T, Ishino Y, Morikawa K (2009) Structural determinant for switching between the polymerase and exonuclease modes in the PCNA-replicative DNA polymerase complex. *Proc Natl Acad Sci USA* **106**: 20693-20698
- Niu H, Chung WH, Zhu Z, Kwon Y, Zhao W, Chi P, Prakash R, Seong C, Liu D, Lu L, Ira G, Sung P (2010) Mechanism of the ATP-dependent DNA end-resection machinery from *Saccharomyces cerevisiae*. *Nature* **467**: 108-111
- Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H, Chee G-J, Hattori M, Kanai A, Atomi H, Takai K, Takami H (2011) Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res* **39**: 3204-3223
- Odell M, Shuman S (1999) Footprinting of *Chlorella* virus DNA ligase bound at a nick in duplex DNA. *J Biol Chem* **274**: 14032-14039
- Odell M, Sriskanda V, Shuman S, Nikolov DB (2000) Crystal structure of eukaryotic DNA ligase-adenylate illuminates the mechanism of nick sensing and strand joining. *Mol Cell* **6**: 1183-1193
- Ogawa T, Okazaki T (1984) Function of RNase H in DNA replication revealed by RNase H defective mutants of *Escherichia coli*. *Mol Gen Genet* **193**: 231-237
- Paeschke K, Capra JA, Zakian VA (2011) DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. *Cell* **145**: 678-691
- Pandey M, Syed S, Donmez I, Patel G, Ha T, Patel SS (2009) Coordinating DNA replication by means of priming loop and differential synthesis rate. *Nature* **462**: 940-943
- Park K, Debyser Z, Tabor S, Richardson CC, Griffith JD (1998) Formation of a DNA loop at the replication fork generated by bacteriophage T7 replication proteins. *J Biol Chem* **273**: 5260-5270
- Pascal JM (2008) DNA and RNA ligases: structural variations and shared mechanisms. *Curr Opin Struct Biol* **18**: 96-105

- Pascal JM, O'Brien PJ, Tomkinson AE, Ellenberger T (2004) Human DNA ligase I completely encircles and partially unwinds nicked DNA. *Nature* **432**: 473-478
- Pascal JM, Tsodikov OV, Hura GL, Song W, Cotner EA, Classen S, Tomkinson AE, Tainer JA, Ellenberger T (2006) A flexible interface between DNA ligase and PCNA supports conformational switching and efficient ligation of DNA. *Mol Cell* **24**: 279-291
- Pavlov YI, Frahm C, Nick McElhinny SA, Niimi A, Suzuki M, Kunkel TA (2006) Evidence that errors made by DNA polymerase alpha are corrected by DNA polymerase delta. *Curr Biol* **16**: 202-207
- Paytubi S, McMahon Sa, Graham S, Liu H, Botting CH, Makarova KS, Koonin EV, Naismith JH, White MF (2011) Displacement of the canonical single-stranded DNA-binding protein in the Thermoproteales. *Proc Natl Acad Sci USA* **109**: 398-405
- Pfander B, Moldovan G-L, Sacher M, Hoege C, Jentsch S (2005) SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. *Nature* **436**: 428-433
- Pike JE, Burgers PM, Campbell JL, Bambara RA (2009) Pif1 helicase lengthens some Okazaki fragment flaps necessitating Dna2 nuclease/helicase action in the two-nuclease processing pathway. *J Biol Chem* **284**: 25170-25180
- Pike JE, Henry RA, Burgers PM, Campbell JL, Bambara RA (2010) An alternative pathway for Okazaki fragment processing: resolution of fold-back flaps by Pif1 helicase. *J Biol Chem* **285**: 41712-41723
- Pisani FM, De Martino C, Rossi M (1992) A DNA polymerase from the archaeon *Sulfolobus solfataricus* shows sequence similarity to family B DNA polymerases. *Nucleic Acids Res* **20**: 2711-2716
- Piserchio A, Nair PA, Shuman S, Ghose R (2010) Solution NMR studies of *Chlorella* virus DNA ligase-adenylate. *J Mol Biol* **395**: 291-308
- Prangishvili D, Klenk HP (1993) Nucleotide sequence of the gene for a 74 kDa DNA polymerase from the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res* **21**: 2768
- Prelich G, Kostura M, Marshak DR, Mathews MB, Stillman B (1987a) The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication *in vitro*. *Nature* **326**: 471-475
- Prelich G, Tan CK, Kostura M, Mathews MB, So AG, Downey KM, Stillman B (1987b) Functional identity of proliferating cell nuclear antigen and a DNA polymerase-delta auxiliary protein. *Nature* **326**: 517-520
- Pursell ZF, Isoz I, Lundström E-B, Johansson E, Kunkel TA (2007) Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* **317**: 127-130
- Qiu J, Qian Y, Frank P, Wintersberger U, Shen B (1999) *Saccharomyces cerevisiae* RNase H(35) functions in RNA primer removal during lagging-strand DNA synthesis, most efficiently in cooperation with Rad27 nuclease. *Mol Cell Biol* **19**: 8361-8371

- Querol-Audí J, Yan C, Xu X, Tsutakawa SE, Tsai M-S, Tainer JA, Cooper PK, Nogales E, Ivanov I (2012) Repair complexes of FEN1 endonuclease, DNA, and Rad9-Hus1-Rad1 are distinguished from their PCNA counterparts by functionally important stability. *Proc Natl Acad Sci USA* **109**: 8528-8533
- Rahmeh AA, Zhou Y, Xie B, Li H, Lee EY, Lee MY (2012) Phosphorylation of the p68 subunit of Pol delta acts as a molecular switch to regulate its interaction with PCNA. *Biochemistry* **51**: 416-424
- Reagan MS, Pittenger C, Siede W, Friedberg EC (1995) Characterization of a mutant strain of *Saccharomyces cerevisiae* with a deletion of the RAD27 gene, a structural homolog of the RAD2 nucleotide excision repair gene. *J Bacteriol* **177**: 364-371
- Reijns MA, Rabe B, Rigby RE, Mill P, Astell KR, Lettice LA, Boyle S, Leitch A, Keighren M, Kilanowski F, Devenney PS, Sexton D, Grimes G, Holt IJ, Hill RE, Taylor MS, Lawson KA, Dorin JR, Jackson AP (2012) Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* **149**: 1008-1022
- Rella R, Raia CA, Pisani FM, D'Auria S, Nucci R, Gambacorta A, De Rosa M, Rossi M (1990) Purification and properties of a thermophilic and thermostable DNA polymerase from the archaebacterium *Sulfolobus solfataricus*. *Ital J Biochem* **39**: 83-99
- Reyes-Lamothe R, Sherratt DJ, Leake MC (2010) Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science* **328**: 498-501
- Richard DJ, Bell SD, White MF (2004) Physical and functional interaction of the archaeal single-stranded DNA-binding protein SSB with RNA polymerase. *Nucleic Acids Res* **32**: 1065-1074
- Rivera MC, Lake JA (2004) The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* **431**: 152-155
- Robbins JB, Murphy MC, White BA, Mackie RI, Ha T, Cann IKO (2004) Functional analysis of multiple single-stranded DNA-binding proteins from *Methanosarcina acetivorans* and their effects on DNA synthesis by DNA polymerase β . *J Biol Chem* **279**: 6315-6326
- Roberts JA, Bell SD, White MF (2003) An archaeal XPF repair endonuclease dependent on a heterotrimeric PCNA. *Mol Microbiol* **48**: 361-371
- Robertson CE, Harris JK, Spear JR, Pace NR (2005) Phylogenetic diversity and ecology of environmental Archaea. *Curr Opin Microbiol* **8**: 638-642
- Robinson NP, Bell SD (2007) Extrachromosomal element capture and the evolution of multiple replication origins in archaeal chromosomes. *Proc Natl Acad Sci USA* **104**: 5806-5811
- Robinson NP, Blood KA, McCallum SA, Edwards PAW, Bell SD (2007) Sister chromatid junctions in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *EMBO J* **26**: 816-824
- Robinson NP, Dionne I, Lundgren M, Marsh VL, Bernander R, Bell SD (2004) Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. *Cell* **116**: 25-38

- Rogozin IB, Makarova KS, Pavlov YI, Koonin EV (2008) A highly conserved family of inactivated archaeal B family DNA polymerases. *Biol Direct* **3**: 32
- Rolfsmeier ML, Haseltine CA (2010) The single-stranded DNA binding protein of *Sulfolobus solfataricus* acts in the presynaptic step of homologous recombination. *J Mol Biol* **397**: 31-45
- Rossi ML, Bambara RA (2006) Reconstituted Okazaki fragment processing indicates two pathways of primer removal. *J Biol Chem* **281**: 26051-26061
- Rossi ML, Pike JE, Wang W, Burgers PM, Campbell JL, Bambara RA (2008) Pif1 helicase directs eukaryotic Okazaki fragments toward the two-nuclease cleavage pathway for primer removal. *J Biol Chem* **283**: 27483-27493
- Sabouri N, McDonald KR, Webb CJ, Cristea IM, Zakian VA (2012) DNA replication through hard-to-replicate sites, including both highly transcribed RNA Pol II and Pol III genes, requires the *S. pombe* Pfh1 helicase. *Genes Dev* **26**: 581-593
- Sakabe K, Okazaki R (1966) A unique property of the replicating region of chromosomal DNA. *Biochim Biophys Acta* **129**: 651-654
- Sakurai S, Kitano K, Yamaguchi H, Hamada K, Okada K, Fukuda K, Uchida M, Ohtsuka E, Morioka H, Hakoshima T (2005) Structural basis for recruitment of human flap endonuclease 1 to PCNA. *EMBO J* **24**: 683-693
- Salinas F, Benkovic SJ (2000) Characterization of bacteriophage T4-coordinated leading- and lagging-strand synthesis on a minicircle substrate. *Proc Natl Acad Sci USA* **97**: 7196-7201
- Samson RY, Bell SD (2011) Cell cycles and cell division in the archaea. *Curr Opin Microbiol* **14**: 350-356
- Samson RY, Obita T, Freund SM, Williams RL, Bell SD (2008) A role for the ESCRT system in cell division in archaea. *Science* **322**: 1710-1713
- Sanchez-Pulido L, Ponting CP (2011) Cdc45: the missing RecJ ortholog in eukaryotes? *Bioinformatics* **27**: 1885-1888
- Sanders GM, Dallmann HG, McHenry CS (2010) Reconstitution of the *B. subtilis* replisome with 13 proteins including two distinct replicases. *Mol Cell* **37**: 273-281
- Sartori AA, Jiricny J (2003) Enzymology of base excision repair in the hyperthermophilic archaeon *Pyrobaculum aerophilum*. *J Biol Chem* **278**: 24563-24576
- Savino C, Federici L, Johnson KA, Vallone B, Nastopoulos V, Rossi M, Pisani FM, Tsernoglou D (2004) Insights into DNA replication: the crystal structure of DNA polymerase B1 from the archaeon *Sulfolobus solfataricus*. *Structure* **12**: 2001-2008

- Scott MT, Morrice N, Ball KL (2000) Reversible phosphorylation at the C-terminal regulatory domain of p21(Waf1/Cip1) modulates proliferating cell nuclear antigen binding. *J Biol Chem* **275**: 11529-11537
- Shamoo Y, Steitz TA (1999) Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* **99**: 155-166
- She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CC, Clausen IG, Curtis BA, De Moors A, Erauso G, Fletcher C, Gordon PM, Heikamp-de Jong I, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW, Van der Oost J (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc Natl Acad Sci USA* **98**: 7835-7840
- Shen Y, Tang XF, Yokoyama H, Matsui E, Matsui I (2004) A 21-amino acid peptide from the cysteine cluster II of the family D DNA polymerase from *Pyrococcus horikoshii* stimulates its nuclease activity which is Mre11-like and prefers manganese ion as the cofactor. *Nucleic Acids Res* **32**: 158-168
- Shi X, Jung Y, Lin L-J, Liu C, Wu C, Cann IKO, Ha T (2012) Quantitative fluorescence labeling of aldehyde-tagged proteins for single-molecule imaging. *Nat Methods* **9**: 499-503
- Simonetta KR, Kazmirski SL, Goedken ER, Cantor AJ, Kelch BA, McNally R, Seyedin SN, Makino DL, O'Donnell M, Kuriyan J (2009) The mechanism of ATP-dependent primer-template recognition by a clamp loader complex. *Cell* **137**: 659-671
- Sinha NK, Morris CF, Alberts BM (1980) Efficient *in vitro* replication of double-stranded DNA templates by a purified T4 bacteriophage replication system. *J Biol Chem* **255**: 4290-4293
- Smith DJ, Whitehouse I (2012) Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature* **483**: 434-438
- Sohn SY, Cho Y (2009) Crystal structure of the human rad9-hus1-rad1 clamp. *J Mol Biol* **390**: 490-502
- Sommers CH, Miller EJ, Dujon B, Prakash S, Prakash L (1995) Conditional lethality of null mutations in RTH1 that encodes the yeast counterpart of a mammalian 5'- to 3'-exonuclease required for lagging strand DNA synthesis in reconstituted systems. *J Biol Chem* **270**: 4193-4196
- Sparks JL, Chon H, Cerritelli SM, Kunkel TA, Johansson E, Crouch RJ, Burgers PM (2012) RNase H2-Initiated Ribonucleotide Excision Repair. *Mol Cell* **47**: 980-986
- Sporbert A, Domaing P, Leonhardt H, Cardoso MC (2005) PCNA acts as a stationary loading platform for transiently interacting Okazaki fragment maturation proteins. *Nucleic Acids Res* **33**: 3521-3528

- Sporbert A, Gahl A, Ankerhold R, Leonhardt H, Cardoso MC (2002) DNA polymerase clamp shows little turnover at established replication sites but sequential *de novo* assembly at adjacent origin clusters. *Mol Cell* **10**: 1355-1365
- Sriskanda V, Kelman Z, Hurwitz J, Shuman S (2000) Characterization of an ATP-dependent DNA ligase from the thermophilic archaeon *Methanobacterium thermoautotrophicum*. *Nucleic Acids Res* **28**: 2221-2228
- Stelter P, Ulrich HD (2003) Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature* **425**: 188-191
- Stewart JA, Campbell JL, Bambara RA (2009) Significance of the dissociation of Dna2 by flap endonuclease 1 to Okazaki fragment processing in *Saccharomyces cerevisiae*. *J Biol Chem* **284**: 8283-8291
- Stewart JA, Campbell JL, Bambara RA (2010) Dna2 is a structure-specific nuclease, with affinity for 5'-flap intermediates. *Nucleic Acids Res* **38**: 920-930
- Stewart JA, Miller AS, Campbell JL, Bambara RA (2008) Dynamic removal of replication protein A by Dna2 facilitates primer cleavage during Okazaki fragment processing in *Saccharomyces cerevisiae*. *J Biol Chem* **283**: 31356-31365
- Stith CM, Sterling J, Resnick MA, Gordenin DA, Burgers PM (2008) Flexibility of eukaryotic Okazaki fragment maturation through regulated strand displacement synthesis. *J Biol Chem* **283**: 34129-34140
- Stroud A, Liddell S, Allers T (2012) Genetic and biochemical identification of a novel single-stranded DNA-binding complex in *Haloferax volcanii*. *Front Microbiol*: DOI:10.3389/fmicb.2012.00224
- Stukenberg PT, Studwell-Vaughan PS, O'Donnell M (1991) Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme. *J Biol Chem* **266**: 11328-11334
- Stukenberg PT, Turner J, O'Donnell M (1994) An explanation for lagging strand replication: polymerase hopping among DNA sliding clamps. *Cell* **78**: 877-887
- Su'etsugu M, Errington J (2011) The replicase sliding clamp dynamically accumulates behind progressing replication forks in *Bacillus subtilis* cells. *Mol Cell* **41**: 720-732
- Subramanian J, Vijayakumar S, Tomkinson AE, Arnheim N (2005) Genetic instability induced by overexpression of DNA ligase I in budding yeast. *Genetics* **171**: 427-441
- Sutton MD, Duzen JM, Scouten Ponticelli SK (2010) A single hydrophobic cleft in the *Escherichia coli* processivity clamp is sufficient to support cell viability and DNA damage-induced mutagenesis in vivo. *BMC Mol Biol* **11**: 102
- Tahirov TH, Makarova KS, Rogozin IB, Pavlov YI, Koonin EV (2009) Evolution of DNA polymerases: an inactivated polymerase-exonuclease module in Pol epsilon and a chimeric origin of eukaryotic polymerases from two classes of archaeal ancestors. *Biol Direct* **4**: 11

- Tan CK, Castillo C, So AG, Downey KM (1986) An auxiliary protein for DNA polymerase-delta from fetal calf thymus. *J Biol Chem* **261**: 12310-12316
- Tanner NA, Tolun G, Loparo JJ, Jergic S, Griffith JD, Dixon NE, van Oijen AM (2011) *E. coli* DNA replication in the absence of free β clamps. *EMBO J* **30**: 1830-1840
- Tishkoff DX, Boerger AL, Bertrand P, Filosi N, Gaida GM, Kane MF, Kolodner RD (1997a) Identification and characterization of *Saccharomyces cerevisiae* EXO1, a gene encoding an exonuclease that interacts with MSH2. *Proc Natl Acad Sci USA* **94**: 7487-7492
- Tishkoff DX, Filosi N, Gaida GM, Kolodner RD (1997b) A novel mutation avoidance mechanism dependent on *S. cerevisiae* RAD27 is distinct from DNA mismatch repair. *Cell* **88**: 253-263
- Tom S, Henricksen LA, Bambara RA (2000) Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1. *J Biol Chem* **275**: 10498-10505
- Tom S, Henricksen LA, Park MS, Bambara RA (2001) DNA ligase I and proliferating cell nuclear antigen form a functional complex. *J Biol Chem* **276**: 24817-24825
- Tsutakawa SE, Classen S, Chapados BR, Arvai AS, Finger LD, Guenther G, Tomlinson CG, Thompson P, Sarker AH, Shen B, Cooper PK, Grasby JA, Tainer JA (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* **145**: 198-211
- Turchi JJ, Huang L, Murante RS, Kim Y, Bambara RA (1994) Enzymatic completion of mammalian lagging-strand DNA replication. *Proc Natl Acad Sci USA* **91**: 9803-9807
- Uemori T, Ishino Y, Doi H, Kato I (1995) The hyperthermophilic archaeon *Pyrodicticum occultum* has two alpha-like DNA polymerases. *J Bacteriol* **177**: 2164-2177
- Umar A, Buermeyer AB, Simon JA, Thomas DC, Clark AB, Liskay RM, Kunkel TA (1996) Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell* **87**: 65-73
- Vos SM, Tretter EM, Schmidt BH, Berger JM (2011) All tangled up: how cells direct, manage and exploit topoisomerase function. *Nat Rev Mol Cell Biol* **12**: 827-841
- Wadsworth RIM, White MF (2001) Identification and properties of the crenarchaeal single-stranded DNA binding protein from *Sulfolobus solfataricus*. *Nucleic Acids Res* **29**: 914-920
- Wagner J, Etienne H, Fuchs RP, Cordonnier A, Burnouf D (2009a) Distinct beta-clamp interactions govern the activities of the Y family PolIV DNA polymerase. *Mol Microbiol* **74**: 1143-1151
- Wagner M, Berkner S, Ajon M, Driessen AJ, Lipps G, Albers SV (2009b) Expanding and understanding the genetic toolbox of the hyperthermophilic genus *Sulfolobus*. *Biochem Soc Trans* **37**: 97-101

- Wardle J, Burgers PM, Cann IK, Darley K, Heslop P, Johansson E, Lin LJ, McGlynn P, Sanvoisin J, Stith CM, Connolly BA (2008) Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya. *Nucleic Acids Res* **36**: 705-711
- Watson JD, Crick FH (1953a) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**: 737-738
- Watson JD, Crick FH (1953b) Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171**: 964-967
- White MF, Dillingham MS (2011) Iron-sulphur clusters in nucleic acid processing enzymes. *Curr Opin Struct Biol* **22**: 92-100
- Wickner S (1976) Mechanism of DNA elongation catalyzed by *Escherichia coli* DNA polymerase III, dnaZ protein, and DNA elongation factors I and III. *Proc Natl Acad Sci USA* **73**: 3511-3515
- Williams GJ, Johnson K, Rudolf J, McMahon SA, Carter L, Oke M, Liu H, Taylor GL, White MF, Naismith JH (2006) Structure of the heterotrimeric PCNA from *Sulfolobus solfataricus*. *Acta Crystallogr F Struct Biol Cryst Commun* **62**: 944-948
- Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci USA* **74**: 5088-5090
- Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* **87**: 4576-4579
- Wold MS (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* **66**: 61-92
- Wu CA, Zechner EL, Marians KJ (1992) Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. I. Multiple effectors act to modulate Okazaki fragment size. *J Biol Chem* **267**: 4030-4044
- Xing G, Kirouac K, Shin YJ, Bell SD, Ling H (2009) Structural insight into recruitment of translesion DNA polymerase Dpo4 to sliding clamp PCNA. *Mol Microbiol* **71**: 678-691
- Xu M, Bai L, Gong Y, Xie W, Hang H, Jiang T (2009) Structure and functional implications of the human rad9-hus1-rad1 cell cycle checkpoint complex. *J Biol Chem* **284**: 20457-20461
- Yamasaki K, Urushibata Y, Yamasaki T, Arisaka F, Matsui I (2010) Solution structure of the N-terminal domain of the archaeal D-family DNA polymerase small subunit reveals evolutionary relationship to eukaryotic B-family polymerases. *FEBS Lett* **584**: 3370-3375
- Yao NY, Johnson A, Bowman GD, Kuriyan J, O'Donnell M (2006) Mechanism of proliferating cell nuclear antigen clamp opening by replication factor C. *J Biol Chem* **281**: 17528-17539

- Yao NY, Turner J, Kelman Z, Stukenberg PT, Dean F, Shechter D, Pan ZQ, Hurwitz J, O'Donnell M (1996) Clamp loading, unloading and intrinsic stability of the PCNA, beta and gp45 sliding clamps of human, *E. coli* and T4 replicases. *Genes Cells* **1**: 101-113
- Yarza P, Ludwig W, Euzéby J, Amann R, Schleifer KH, Glockner FO, Rossello-Mora R (2010) Update of the All-Species Living Tree Project based on 16S and 23S rRNA sequence analyses. *Syst Appl Microbiol* **33**: 291-299
- Yuan J, Ghosal G, Chen J (2012) The HARP-like domain-containing protein AH2/ZRANB3 binds to PCNA and participates in cellular response to replication stress. *Mol Cell* **47**: 410-421
- Yuzhakov A, Turner J, O'Donnell M (1996) Replisome assembly reveals the basis for asymmetric function in leading and lagging strand replication. *Cell* **86**: 877-886
- Zaparty M, Esser D, Gertig S, Haferkamp P, Kouril T, Manica A, Pham TK, Reimann J, Schreiber K, Sierocinski P, Teichmann D, van Wolferen M, von Jan M, Wieloch P, Albers SV, Driessen AJ, Klenk HP, Schleper C, Schomburg D, van der Oost J, Wright PC, Siebers B (2010) "Hot standards" for the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Extremophiles* **14**: 119-142
- Zhang C, Guo L, Deng L, Wu Y, Liang Y, Huang L, She Q (2010) Revealing the essentiality of multiple archaeal pcna genes using a mutant propagation assay based on an improved knockout method. *Microbiology* **156**: 3386-3397
- Zhang L, Brown JA, Newmister SA, Suo Z (2009a) Polymerization fidelity of a replicative DNA polymerase from the hyperthermophilic archaeon *Sulfolobus solfataricus* P2. *Biochemistry* **48**: 7492-7501
- Zhang L, Zhang L, Liu Y, Yang S, Gao C, Gong H, Feng Y, He Z-G (2009b) Archaeal eukaryote-like Orc1/Cdc6 initiators physically interact with DNA polymerase B1 and regulate its functions. *Proc Natl Acad Sci USA* **106**: 7792-7797
- Zhang SJ, Zeng XR, Zhang P, Toomey NL, Chuang RY, Chang LS, Lee MY (1995) A conserved region in the amino terminus of DNA polymerase delta is involved in proliferating cell nuclear antigen binding. *J Biol Chem* **270**: 7988-7992
- Zheng L, Dai H, Qiu J, Huang Q, Shen B (2007) Disruption of the FEN-1/PCNA interaction results in DNA replication defects, pulmonary hypoplasia, pancytopenia, and newborn lethality in mice. *Mol Cell Biol* **27**: 3176-3186
- Zheng L, Shen B (2011) Okazaki fragment maturation: nucleases take centre stage. *J Mol Cell Biol* **3**: 23-30
- Zheng L, Zhou M, Guo Z, Lu H, Qian L, Dai H, Qiu J, Yakubovskaya E, Bogenhagen DF, Demple B, Shen B (2008) Human DNA2 is a mitochondrial nuclease/helicase for efficient processing of DNA replication and repair intermediates. *Mol Cell* **32**: 325-336
- Zhu Z, Chung W-H, Shim EY, Lee SE, Ira G (2008) Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. *Cell* **134**: 981-994

- Zhuang Z, Yoder BL, Burgers PM, Benkovic SJ (2006) The structure of a ring-opened proliferating cell nuclear antigen-replication factor C complex revealed by fluorescence energy transfer. *Proc Natl Acad Sci USA* **103**: 2546-2551
- Zillig W, Stetter KO, Wunderl S, Schulz W, Priess H, Scholz I (1980) The *Sulfolobus-Caldariella* group - taxonomy on the basis of the structure of DNA-dependent RNA polymerases. *Arch Microbiol* **125**: 259-269

Appendix I – Oligonucleotide sequences

PRIMERS USED FOR CLONING

Name	Sequence (5'-3')
ΔP1	ATGATCTTTTGGATAGAGCCAAGGTTACTCGAG
ΔP2	GGGACTTTTACATTGAACCAAGGGCAGACGTC
ΔP2	GTTACTTATTTATTAGAACCAAAAGTTGTCGAG
PolB1_stop	CCCAAGTAAAGGAAATAGTTAACTCGAGCACCACCACCACCACCTG AGAT
PolB1_PIP	GCTAAGCAACTTACCGCAGCAGATATTCCTTCATC
Fen1_PIP	CAGGATTGGACAGAGCAGCATAGCTCGAGCACCAC
Lig1_PIP	CAGGAATTTTAGGAGCTGCGGGTACCACAAGTAAAG
CC544SS	TCTCCTTTACATACTTCTTCTCAAGACGTAGTAG
C483S	GATTCTTTTGAATCCGTTTCCAAGGTAGCTAGTGTTTTC
I464C	GGGTAAAAGAGGCGGTAAGTGTAAGCTCGTTATTAATGGC
K463C	GGGTAAAAGAGGCGGTTGTATAAGCTCGTTATTAATGGC
V507C	ATGGAAATAAAGAGGGACTGTAAGCATCCTAGGGTTAATTC
K515C	CATCCTAGGGTTAATTCTTGTATGGAACCAGATATATGG
M516C	CCTAGGGTTAATTCTAAGTGTGAACCAGATATATGGG
E2C_for	GGGATCCATATGTGTTTTAAAGTTATTGCCGAGTAC
T7R	GCTAGTTATTGCTCAGCGG
Sso0150_for	CATGACCATATGTCAACGAGATGGCTACC
Sso0150_for2	CATGACCCATGGCAACGAGATGGCTACC
Sso0150_rev	CATGACCTCGAGCGTAAATCCTCCTACTATTAATATCTTAC
Sso6202_for	CATGACCATATGTCTGTTAATCAGAAGGAAATTG
Sso6202_for2	CATGACGAATTCATGTCTGTTAATCAGAAGGAAATTG

Sso6202_rev	CATGACCTCGAGCTTCTTGTCTAGTAGATTTCTCAC
Sso6202_rev2	CATGACCTCGAGTCACTTCTTGTCTAGTAGATTTCC
RNase_for	CATGACCCATGGGAGTAGGCATAGATGAAG
RNase_rev	GGATCCCTCGAGTCACCATAACCTTCTACCCTCT
PolB1_for	CATGACCCATGGCTAAGCAACTTACCTTATTTG
PolB1_rev	CATGACCTCGAGTTAACTATTTCTTTACTTGGGTATG

OLIGONUCLEOTIDES USED FOR SUBSTRATES

Ribonucleotides are underlined. Dideoxynucleotides are indicated in parentheses. Sites of Cy3B attachment are indicated in red. Sites of biotin incorporation are indicated by an asterisk.

Name	Sequence (5'-3')
Template	GTATAATGTTTATCTTATTTATCTCAATTAGCTAGATTAAGTAATTTCCAGAGGAAATAGATGGGTCCCACTTCGCTTTTAAATGGTATTTTGTGGACTT
Upstream	AAGTCCACAAAATACCATTTAAAAGCGAAG
Downstream_DNA	ATTTCTCTGGAAATTACTTAATCTAGCTAATTGAGATAAATAAGATAAACATTATAC
Downstream_RNA	<u>AUUUCCUCUGGAAATTACTTAATCTAGCTAATTGAGATAAATAAGATAAACATTATAC</u>
Upstream_1nt	AAGTCCACAAAATACCATTTAAAAGCGAAGTGGGACCCATCTATTTCTCTGGAAATTAC
Downstream_flap	TAAAGGAGACCTTTAATCTTAATCTAGCTAATTGAGATAAATAAGATAAACATTATAC
Nick1	GATCGTGGCTATTGTGCGCCCTTATTCCGAC
Nick1_ddC	GATCGTGGCTATTGTGCGCCCTTATTCCGA(C)
Nick2	AGTGACTACATTTTTGTAGTCACTCTCGGAATAAGGGCGACAATAGCCACGATC

Nick2_internal	AGTGACTACATTTT T GTAGTCACTCTCGGAATAAGGGCGACAATAGCC ACGATC
Nick2_3'	AGTGACTACATTTT T GTAGTCACTCTCGGAATAAGGGCGACAATAGCC ACGAT C
Upstream_nick	AAGTCCACAAAATACCATTTAAAGCGAAGTGGGACCCATCT
Template_I	GTTGTTGGGTTGGTTTTGTTTGGTGGAACCTTTTTTTTTTTTTTTTTT TTTTTTTCTCCCTTCTTCTCCTCCCTCTCCCTTCCCT
Upstream_I	AGGGAAGGGAGAGGGAGGAGAAGAAGGGAG
Downstream_I	<u>GGU</u> UCCACCAAACAAAACCAACCCAACAAC
B-Template	*GTATAATGTTTATCTTATTTATCTCAATTAGCTAGATTAAGTAATTT CCAGAGGAAATAGATGGGTCCCACTTCGCTTTTAAATGGTATTTTGTG GACTT
