Mammalian DNA Ligase II

Emma Roberts

A thesis submitted for the degree of Ph.D.

at

The University of London,

November 1993.

Imperial Cancer Research Fund,

Clare Hall Laboratories,

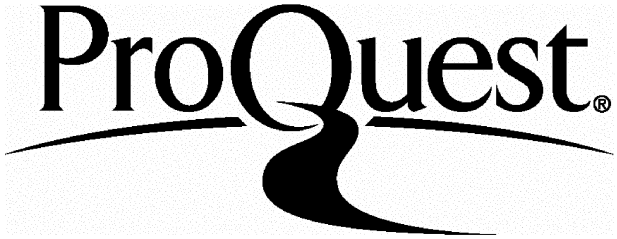
South Mimms, Herts., EN6 3LD

and

University College London,

Gower Street,

London WC1E 6BT



ACKNOWLEDGEMENTS

I am very grateful to Dr. Tomas Lindahl for his advice during the course of this work and to my external supervisor, Prof. Brian Ketterer. I am indebted to the Imperial Cancer Research Fund for supporting this study, John Nicholson for the photographic work presented in my thesis and to Nick Totty for peptide sequencing. Many thanks to Alan Tomkinson, Debbie Barnes, Graham Daly and the Mutagenesis group, Monica Olsson, Brenda Marriott, Frank Fitzjohn, and my friends and colleagues for all their advice and laughter, without whom life at Clare Hall would have been far harder and a lot less fun. Lastly, to my family - I couldn't have done it without you.

ABSTRACT

DNA ligases are essential for DNA replication, repair and recombination. Two DNA ligases have been described previously in mammalian cells and other eukaryotes. DNA ligase I has been purified and the cDNA encoding the enzyme has been cloned. In this study, DNA ligase II has been purified from bovine thymus and it has been confirmed that it is distinct from DNA ligase I in its chromatographic, biochemical and catalytic properties. Internal Lys C peptides have been sequenced from apparently homogeneous DNA ligase II. These show no homology to the predicted amino acid sequence of DNA ligase I. However, both analysis of the peptide sequences and two-dimensional gel electrophoresis showed there to be two proteins present in this DNA ligase II preparation. One of these proteins has been identified as cytovillin, an abundant structural protein. Purification from bovine liver gave a source of DNA ligase II that was not contaminated with cytovillin. A radiolabelled reaction intermediate can be generated for the DNA ligases and in this way it has been possible to compare the active site peptides by electrophoretic techniques. This has confirmed that DNA ligase II is distinct from DNA ligase I. A third DNA ligase, DNA ligase III, was identified in the laboratory during the course of this work, and it has been possible to establish a range of substrate specificities for this enzyme that is different from that of either DNA ligase I or II. Comparison of the active site peptides of DNA ligases II and III has indicated a degree of similarity between these two enzymes and suggests they may be related, although both are unrelated to DNA ligase I.

ABBREVIATIONS

SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
DTT	Dithiothreitol
HPLC	High pressure liquid chromatography
FPLC	Fast protein liquid chromatography
EDTA	Ethylenediaminetetraacetic acid
TLC	Thin layer chromatography

CONTENTS Page

ACKNOWLEDGEMENTS 2
ABSTRACT 3
ABBREVIATIONS 4
CHAPTER 1, Introduction 8
CHAPTER 2, Materials and Methods 38
2.1 Reagents 39
2.2 Tissues 39
2.3 Enzyme Assays 39
2.3.1 Preparation of polynucleotide substrates 39
2.3.2 DNA ligation assays 40
2.3.3 Analysis of ligation products 40
2.3.4 Formation of DNA ligase-adenylate 40
2.3.5 Reactivity of the enzyme-adenylate intermediate 41
2.4 Purification Schemes 41
2.4.1 Purification of DNA ligase II from calf thymus 41
2.4.2 Partial purification of DNA ligase II from bovine liver 43
2.4.3 Purification of DNA ligases I and III from calf thymus 45
2.5 Peptide Sequencing 45
2.5.1 N-terminal amino acid determination 45
2.5.2 Sequencing of internal Lys-C peptides 45
2.6 Two-Dimensional Gel Electrophoresis 45
2.7 Immunisation Procedures 46
2.8 Immunoblots 46
2.9 Analysis of Radiolabelled Active Site Peptides by Two-Dimensional Peptide Mapping 47
2.10 Analysis of Radiolabelled Active Site Peptides by HPLC 48
2.11 Analysis of Radiolabelled DNA Ligases by Proteolytic Digestion 49
2.11.1 Partial digestion of adenylylated DNA ligases I and II using trypsin 49
2.11.2 Complete digestion of adenylylated DNA ligases I and II by trypsin or Lys-C 49
CHAPTER 3, Results - Purification of DNA Ligase II 51
3.1 DNA Ligase II from Calf Thymus 52

3.1.1 Purification 52
3.1.2 Apparently homogeneous DNA ligase II 55
3.1.3 Partial amino acid sequence of purified DNA ligase II 58
3.1.4 Analysis of purified DNA ligase II by two-dimensional gel electrophoresis 60
3.2 DNA Ligase II from Bovine Liver 62
3.2.1 Purification 62
3.2.2 Partially purified DNA ligase II 63
CHAPTER 4, Results - Characterisation of DNA Ligase II 66
4.1 Reactivity of the Adenylylated Forms of DNA Ligases I and II 67
4.2 DNA Ligation Using the ATP Analogue, ATP-αS, as a Cofactor 69
4.3 Polynucleotide Substrate Specificity 69
4.4 Stabilisation of DNA Ligase II Activity by Non-Ionic Detergent 70
4.5 Inhibition of DNA Ligase II Activity by Sodium Chloride 72
4.6 Neutralising Antiserum against DNA Ligase I does not Inhibit DNA Ligases II and III 74
4.7 Immunoblots 74
4.7.1 Polyclonal antiserum against purified DNA ligase I 74
4.7.2 Polyclonal antiserum against a conserved peptide 76
4.7.3 Cross-reactivity of vaccinia virus DNA ligase with antiserum against a conserved peptide 76
4.7.4 Polyclonal antiserum against apparently homogeneous DNA ligase II 76
4.7.5 Polyclonal antiserum against a 'DNA ligase II' peptide 78
CHAPTER 5, Results - Comparison of DNA Ligases I and II by Proteolytic Digests 80
5.1 Comparison of Adenylylated DNA Ligases I and II by Partial Proteolytic Digests 81
5.2 Comparison of the Active Site Peptides of DNA Ligases I and II after Complete Proteolytic Digests 83

5.2.1 Analysis of the radiolabelled active site peptides of DNA ligases I and II by two-dimensional peptide mapping 83
5.2.2 Analysis of the active sites of DNA ligases I and II by HPLC 86
5.2.3 Analysis of the active site peptides of DNA ligases I and II by non-denaturing polyacrylamide gel electrophoresis 90
5.3 Comparison of the Active Site Peptides from DNA Ligases II and III 96
CHAPTER 6, Discussion 97
REFERENCES 110
PUBLICATIONS 119

CHAPTER 1

INTRODUCTION

The continued integrity of the genome is essential for survival in all species. Inaccuracies during replication or deficient repair of exogenous damage can produce mutations which may lead to carcinogenesis and/or cell death. DNA replication, DNA repair and DNA recombination, the processes involved in the maintenance of DNA, are in general, better understood in prokaryotes, although recent advances in *in vitro* systems have produced a greater understanding of these events in eukaryotes. The phenotypes of mutant cell lines of both yeast and higher eukaryotes deficient in these processes have aided in the elucidation of particular enzyme functions. Furthermore, certain human genetic diseases associated with a predisposition to cancer, have been shown to be due to defects in DNA repair, underlining their importance for correct cellular function.

The process of cell division involves the precise duplication of DNA and its genetic information, followed by the controlled separation of the daughter molecules into two new cells. Replication of DNA involves many proteins and must necessarily be accurate and complete before cell division if the genome is to be maintained. Reconstruction *in vitro* of the replication of the Simian virus 40 (SV40) genome has generated a model system for mammalian DNA replication (Li and Kelly, 1984). Detailed analysis of the functions of mammalian enzymes and their accessory proteins has been possible in the context of this system. A possible scheme of events is becoming clearer and this is shown in Fig. 1.1. This scheme involves the separate synthesis of a continuous leading strand and discontinuous synthesis of 30-300 nucleotide Okazaki fragments on the lagging strand (Anderson and DePamphilis, 1979). It may be that the lagging strand is wrapped around, such that a single replication complex can simultaneously synthesise both strands, although this has yet to be firmly established (Hubscher and Thommes, 1992). Recognition of the

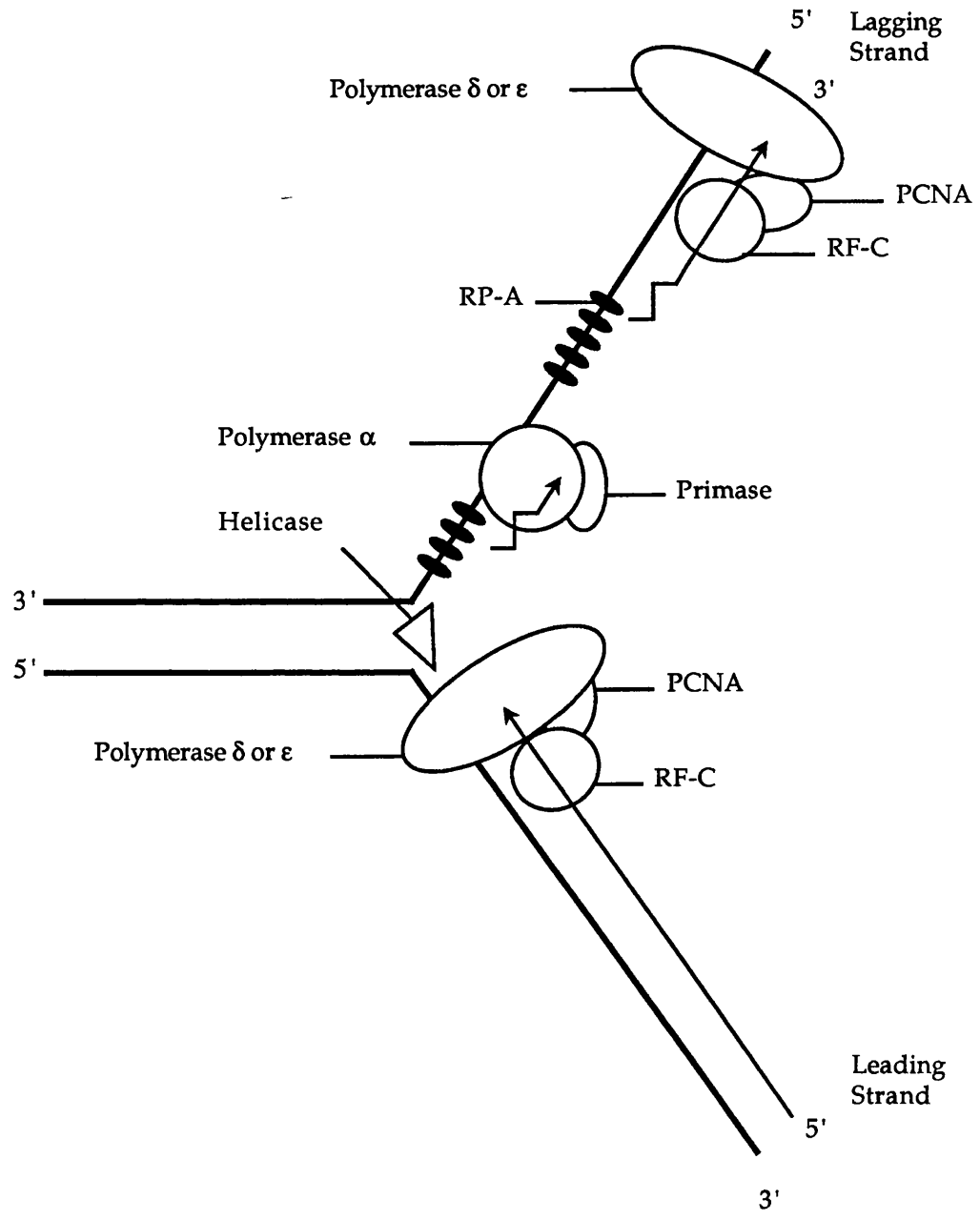


Fig. 1.1. Mammalian DNA Replication (taken from So and Downey, 1992)

origin of replication in the SV40 genome is achieved by a single protein, T antigen (TAg). TAg provides the helicase action necessary to unwind the region of DNA around the origin prior to binding of the proteins that initiate replication (Borowiec and Hurwitz, 1988). Origin recognition and binding appears to be more complex in the mammalian genome. Origins of replication are less well defined and a number of proteins or a complex may be involved (Bell and Stillman, 1992; Diffley and Cocker, 1992).

DNA replication is thought to be initiated by DNA polymerase α and its associated primase acitivity. This complex forms RNA primers, initiating DNA polymerisation on both strands (reviewed in Wang, 1991). Synthesis of the leading strand is taken over by another polymerase, with DNA polymerase δ being the most likely candidate. DNA polymerase δ requires proliferating cell nuclear antigen (PCNA) for significant processivity *in vitro* and is thought to be associated with it at the replication fork *in vivo* (Bravo et al., 1987; Prelich et al., 1987). It is not yet clear whether the lagging strand synthesis is continued by DNA polymerase α or by either DNA polymerase δ or ε (Hubscher and Thommes, 1992). Unlike DNA polymerase δ, DNA polymerase ε does not require PCNA for processivity (Syvaoja and Linn, 1989). DNA polymerases α and δ and their accessory proteins are sufficient for efficient synthesis of both leading and lagging strands in the SV40 *in vitro* system (Tsurimoto and Stillman, 1991a; Tsurimoto and Stillman, 1991b) and so the exact role of DNA polymerase ε at the replication fork is still to be determined. Studies of DNA polymerase ε in yeast have shown it to be an essential enzyme, suggesting that it is indeed necessary for DNA replication (Morrison et al., 1990; Araki et al., 1992). The associated 3'→5' exonuclease activities of both DNA polymerases δ and ε produce a high accuracy of DNA replication (Goscin and Byrnes, 1982; Syvaoja and Linn, 1989). Unlike DNA polymerases α, δ and ε, DNA polymerase β activity is

invariant through the cell cycle, and this polymerase is able to fill small gaps in DNA; features that have implicated it in DNA repair (Mosbaugh and Linn, 1984). However, the property of gap filling may also be important in the completion of Okazaki fragments after removal of the RNA primers. Identification of DNA polymerase β in yeast should allow further analysis of its role. The fifth polymerase, DNA polymerase γ, is associated with the replication and repair of mitochondrial DNA and so is not represented on the scheme in Fig. 1.1 (Hubscher et al., 1979).

Accessory proteins are necessary at particular points in the reactions of initiation and polymerisation. As already mentioned, PCNA is important for correct DNA polymerase δ activity, significantly increasing the processivity of this polymerase. RF-A (also known as human SSB or RP-A) is a multisubunit single strand binding protein that stimulates DNA polymerase α activity and is involved, with a helicase, in the initial stages of priming (Weinberg et al., 1990; Tsurimoto and Stillman, 1991a; Tsurimoto and Stillman, 1991b). RF-C is a DNA-dependent ATPase that enhances primer recognition by DNA polymerase δ (Tsurimoto and Stillman, 1989). Certain accessory factors and their cognate DNA polymerases may also be involved in DNA repair, suggesting a degree of overlap in the enzyme functions required for replication and repair (see later).

Mutations that occur through incorrect DNA replication are rare due to the high fidelity and proof reading activity of the DNA polymerases. The susceptibility of DNA to many kinds of damage, both spontaneous and as a result of the action of exogenous agents, may result in mutation. Cells have evolved many processes to deal with these forms of damage. The pathways of repair in eukaryotes appear to be far more complex than those elucidated in prokaryotes, preventing direct comparisons (Hoeijmakers, 1993a). In contrast, the homology found

between repair pathways in yeast and mammalian cells so far suggests that yeast is a good model for eukaryotic repair systems in general, and this has been used to great advantage in the identification of certain repair functions (Hoeijmakers, 1993b). Yeast mutants have been used to identify the genes encoding repair enzymes, and it has often then been possible to establish their homologues in mammalian cells.

Human repair functions have been identified by cross-species complementation of laboratory generated mutant rodent cell lines, upon transfection with human DNA. Furthermore, certain rare, inherited human syndromes, such as xeroderma pigmentosum (XP), Cockayne's syndrome (CS), ataxia telangiectasia (AT), Fanconi's anaemia (FA), and Bloom's syndrome (BS), show an increased sensitivity to DNA damaging agents and, in some cases, an increased incidence of cancer (Barnes, 1992; Wevrick and Buchwald, 1993). Complementation of cell lines representative of these DNA repair syndromes has allowed direct identification of yet more mammalian proteins that have functional and/or sequence homology to the yeast repair enzymes. Although the phenotypic consequences of the defect in many of these genes are known, the exact biochemical function of the encoded protein has yet to be established. The use of reconstituted *in vitro* repair systems has been of enormous advantage in this respect, in identifying the roles of particular proteins acting within a given repair pathway (Wood et al., 1988; Hoeijmakers et al., 1990).

A major pathway in the removal of DNA damage is that of nucleotide excision repair (NER). This pathway serves to excise bulky lesions that distort the DNA backbone. A scheme for nucleotide excision repair is shown in Fig. 1.2.

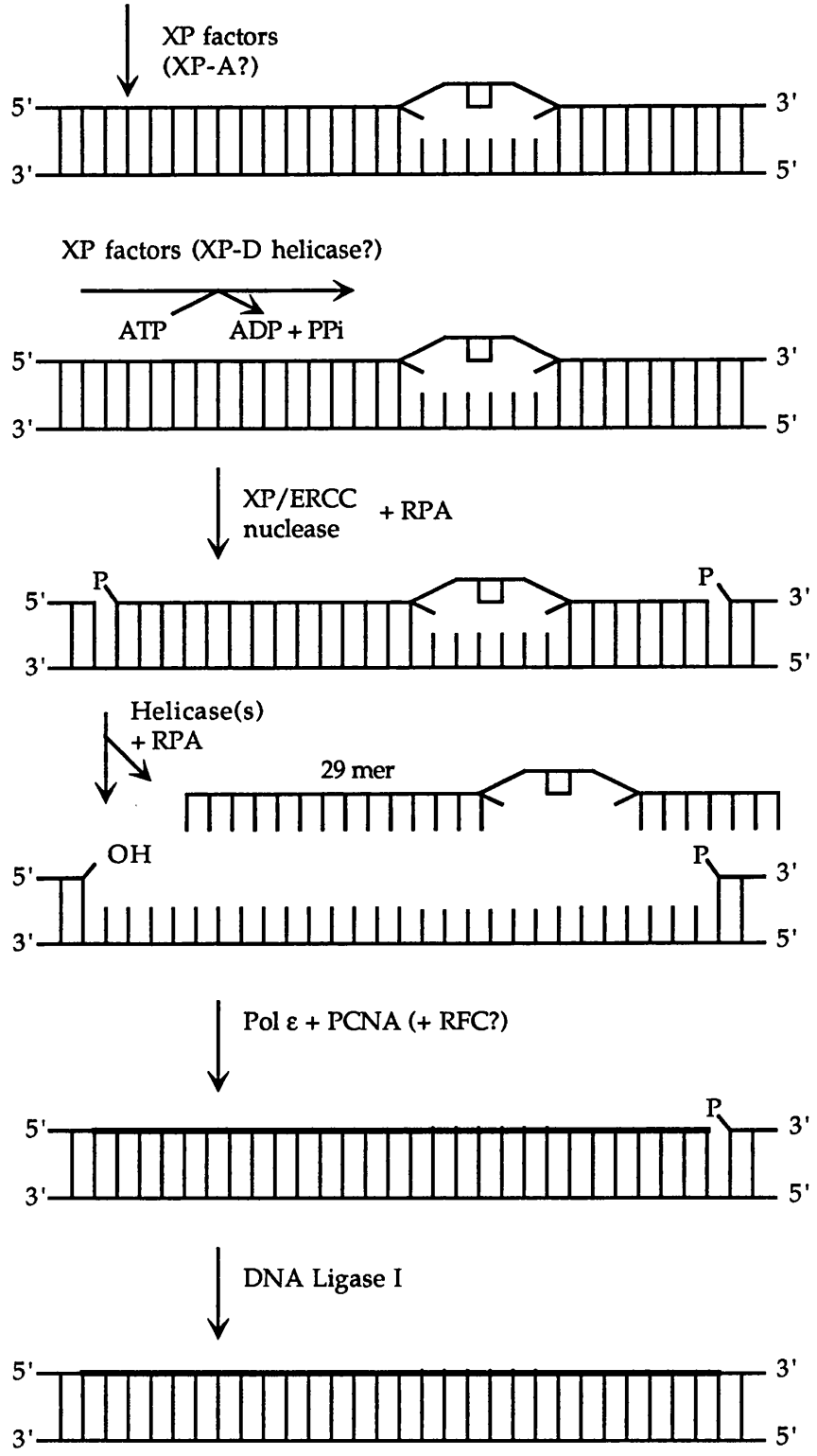


Fig. 1.2. Mammalian Nucleotide Excision Repair (taken from Barnes et al., 1993)

The lesion, possibly a pyrimidine dimer as shown here, is first identified and the necessary proteins located at the site of damage. The DNA is nicked both 5' and 3' to the adduct which is excised in a single stranded oligonucleotide of approximately 30 bp (Huang et al., 1992). Repair synthesis can then take place, possibly using DNA polymerase δ or ε, although other polymerases may also be proficient in this function, and the nick is then sealed by a DNA ligase. A human inherited disease, xeroderma pigmentosum (XP), is due to a defect in this pathway (Wood et al., 1988). A number of genetic complementation groups are known, accounting for the heterogeneity of the clinical syndrome and suggesting that many gene products act in the NER pathway. A number of genes that are defective in XP have been identified from the distinct complementation groups.

Features seen in the cDNA sequences have enabled tentative functional roles to be assigned to certain of these repair proteins. For example, motifs indicative of ATP-dependent DNA helicase activity have been noted in the cDNA of ERCC2, which is equivalent to the XP-D complementing factor (Flejter et al., 1992). The *Saccharomyces cerevisiae* homologue of this protein is RAD3 (Weber et al., 1990), which has been shown to exhibit DNA helicase activity *in vitro* (Harosh et al., 1989). This is in agreement with the proposed role of the XP-D factor in NER (Fig. 1.2), where a DNA helicase would translocate along the DNA, stalling at the site of damage, thereby providing a point of assembly for the repair complex. Although specific functions have not been assigned, many of the proteins required for NER have been shown to be involved in the pre-incision stage of the pathway (Shivji et al., 1992).

Development of an *in vitro* assay system has allowed analysis of repair through the NER pathway (Wood et al., 1988). Extracts from cell lines deficient in NER, such as XP, can be easily identified in this assay.

Using this assay, it has been possible to identify and purify some of the proteins that restore repair capabilities to the defective cell lines. Use of the system to examine the reaction pathway in more detail has identified different stages where enzymes and cofactors are active (Shivji et al., 1992). It has also been possible to look at the repair of damage caused by chemotherapeutic agents such as cis-platin, where the use of single lesion substrates has indicated repair capabilities for specific types of damage (Szymkowski et al., 1992; Szymkowski et al., 1993). Results show that certain lesions are poorly repaired; their retention within the genome possibly accounting for the cytotoxicity of this drug.

Smaller lesions that do not distort the DNA structure, such as modified or damaged bases, are removed by specific and efficient DNA glycosylases. The resulting apurinic/apyrimidinic sites (AP sites), which can also occur through spontaneous deamination or through damage caused by ionizing radiation, are then processed by the base excision repair (BER) pathway, shown in Fig. 1.3. The sequential action of an AP endonuclease, which nicks the DNA 5' to the AP site, and then a deoxyribophosphodiesterase, leaves a single nucleotide gap that can then be filled by a polymerase and sealed by DNA ligase (Price and Lindahl, 1991reviewed in Barnes et al., 1993). This process can also be assayed in cell-free extracts (Dianov et al., 1992; Satoh et al., 1993; Prigent et al., 1994). The abundant nuclear protein, poly(ADP-ribose) polymerase (PARP) is thought to affect the process of BER. The PARP molecule binds to the strand breaks, activating auto-modification through synthesis of long chain poly(ADP-ribose) molecules. This causes release of PARP from the DNA, allowing access of the repair enzymes to the nick (Satoh and Lindahl, 1992; Satoh et al., 1993). It appears that this occurs only during BER and not in NER or during replication of the lagging strand when, presumably, breaks in the DNA are hidden within a protein complex.

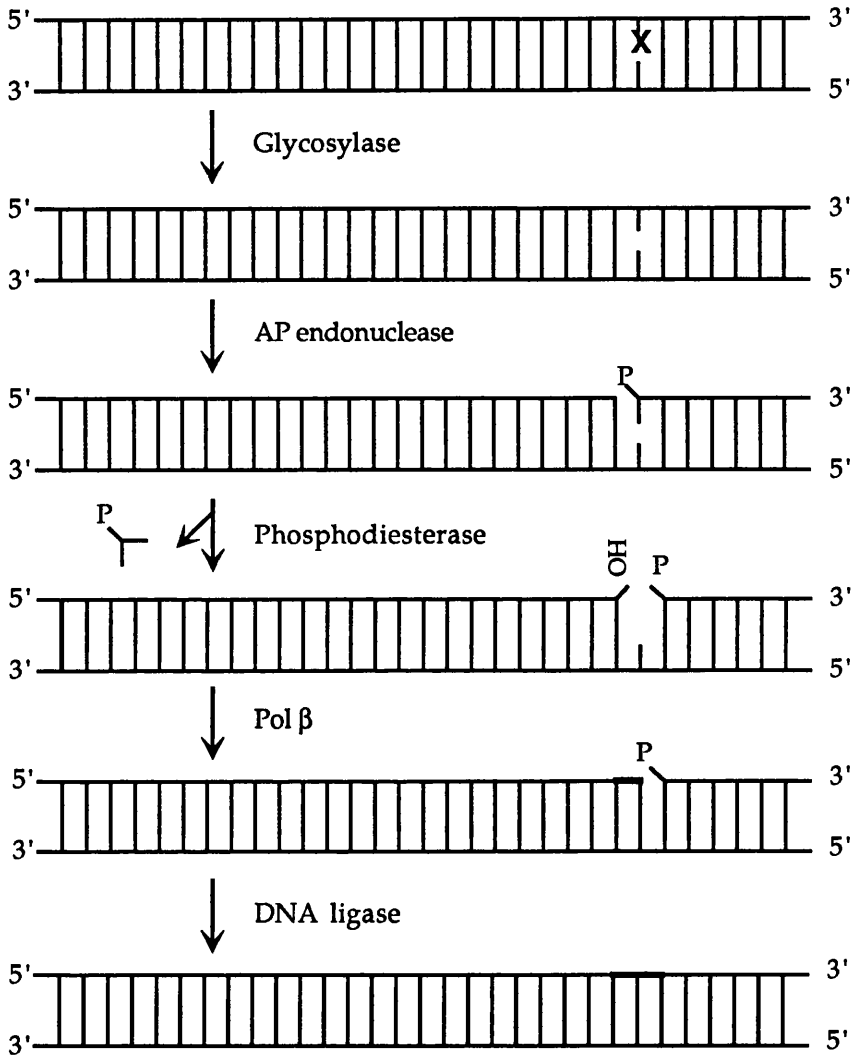


Fig. 1.3. Mammalian Base Excision Repair (taken from Barnes et al., 1993)

A role for PARP in repair may possibly be one of signalling that damage has occurred, producing inhibition of DNA replication where damage would otherwise go unrecognised, although this role has yet to be firmly established.

Single and double strand breaks in the DNA , as well as occurring during excision repair processes, can also occur directly as a consequence of ionizing radiation. *In vitro* systems have shown that the majority of these are not simply religated (Fairman et al., 1992) and may require the involvement of proteins such as that proposed by Thode et al to align non-complementary ends (Thode et al., 1990). Analysis of the joining of nonhomologous ends and the conditions necessary will help in elucidating such functions (Roth and Wilson, 1986). In contrast, some forms of damage, such as O^6-methylguanine (O^6-meG), are repaired without incision of the DNA backbone, giving direct reversal of DNA damage. O^6-meG has both mutagenic and cytotoxic properties, although exactly how this lesion causes cell death is not yet known. A methyltransferase in human cells directly demethylates O^6-meG in response to the alkylation damage, thereby reducing the mutation rate through G:C to A:T transitions. Loss of this enzyme and subsequent inability to avoid these mutations, may have serious consequences for the progression of malignancy (Karran et al., 1993). Certain minor lesions that evade the major repair mechanisms may not be important at the time, but later in the lifespan of the organism, accumulation of such damage may contribute towards malignancy and ageing (Lindahl, 1993).

DNA damage can also be repaired using the recombinational repair pathway, where the damaged DNA is replaced with an intact copy from its sister chromatid. The pathway of recombinational repair is one that appears to have been appropriated by the cell at some point to produce an advantageous genetic diversity. The similarities of the two processes can

be seen in Fig. 1.4, where genetic recombination is detailed alongside an example of recombinational repair, namely that of post-replication recombinational repair. In both, the elements of strand exchange and branch migration, followed by resolution and ligation, are essential (Cox, 1993). The two processes require homology between the DNA molecules involved, and the proteins and reactions necessary for successful recombination are becoming well characterised in *E.coli* (for review see West, 1992). Central to the process of recombination is the RecA protein that forms nucleoprotein filaments within which the strand exchange and branch migration can occur. The RAD51 protein from *S. cerevisiae* is homologous to the RecA protein and is thought to function in a similar way, possibly in conjunction with another protein, RAD52 (Shinohara et al., 1992). Both of these proteins are required in the repair of double strand breaks and mutants in either are sensitive to DNA damaging agents and have aberrant recombination phenotypes. The cDNA encoding the mammalian equivalent of the RecA/RAD51 proteins has recently been cloned and along with the identification of a mammalian resolving protein (Elborough and West, 1990; Shinohara et al., 1993), may now provide the possibility of elucidating the more complex mechanisms of mammalian recombination.

The repair of double strand breaks also appears to involve the process of non-homologous recombination, specifically the functions necessary for V(D)J recombination. In this process, recombination produces antibody diversity in differentiated B cells, and T-cell receptor specificity, through rearrangement of specific gene segments. A mouse mutant that is defective in this processing (Fulop and Phillips, 1990) is extremely immunodeficient, and also shows sensitivity to DNA damaging agents that produce double strand breaks. This implies a common element

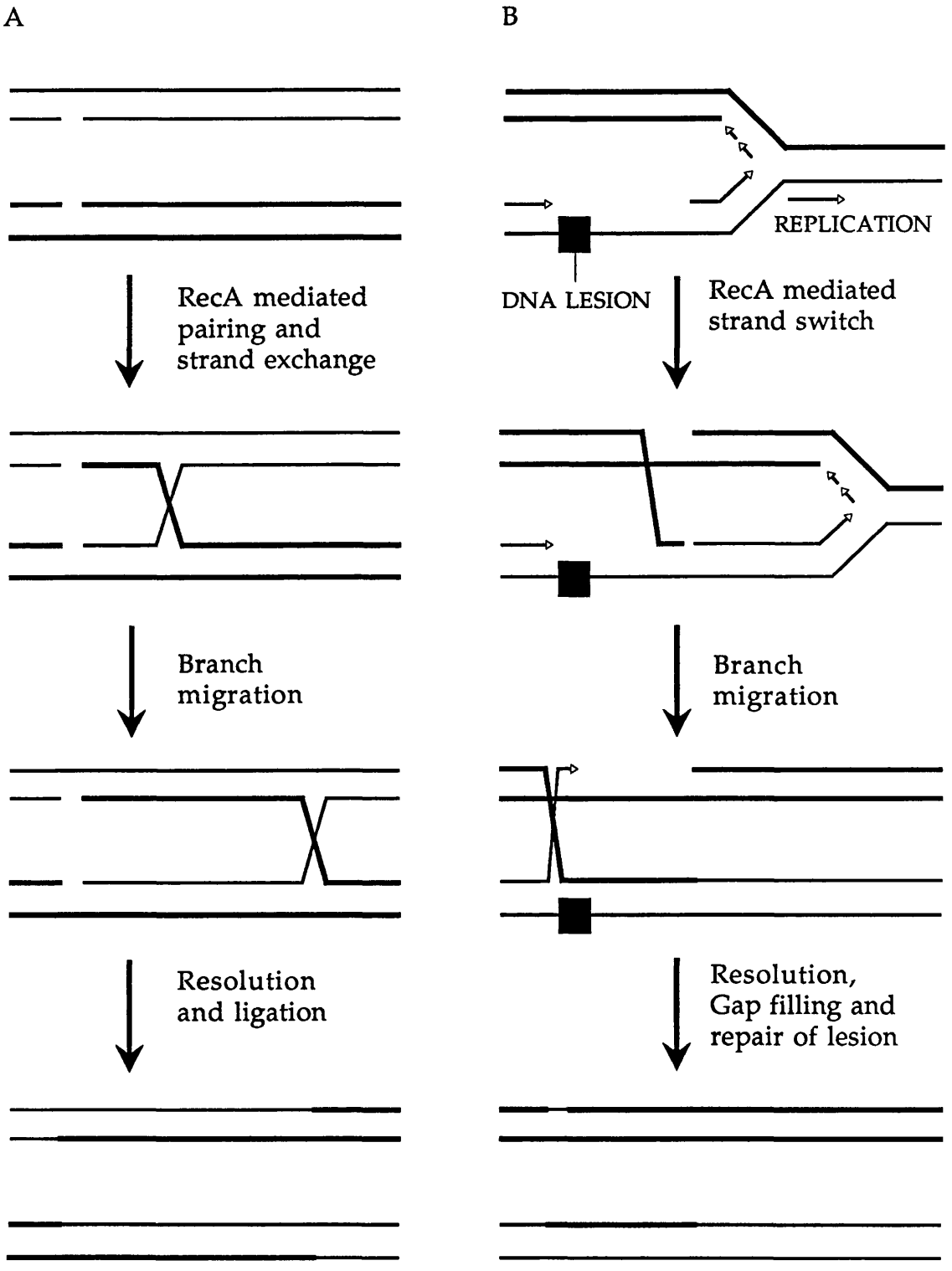


Fig. 1.4. (A) Genetic recombination and (B) post-replication recombinational repair, (taken from Cox, 1993)

in V(D)J joining and the repair of double strand breaks (Biederman et al., 1991; Hendrickson et al., 1991). Indeed, other rodent cell lines isolated as being defective in double strand break repair also show an inability to perform correct V(D)J recombination when stimulated to do so (Pergola et al., 1993).

It is now known that DNA repair is closely coupled to other cellular processes. For example, DNA replication and DNA repair show a common need for DNA synthesis, suggesting that there may be proteins that are conserved between the two processes. This has been shown to be the case for certain accessory proteins required for replication that are also implicated in repair. Using the *in vitro* repair system for NER, Wood and colleagues have shown that both PCNA and human SSB are necessary for correct DNA repair following UV exposure (Coverley et al., 1991; Shivji et al., 1992). Similarly, proteins are also common to DNA repair and other pathways acting on DNA; certain proteins are involved in both DNA repair and transcription. There is a basal level of repair for the overall genome that contrasts with the much faster system of repair for those genes that are being transcribed, particularly for the transcribed strand itself (Hanawalt and Mellon, 1993). Some proteins of the NER repair pathway have been shown to be directly involved in transcription, such as ERCC3, a DNA helicase (Schaeffer et al., 1993), and these may be important in recognition of damage in active genes (Buratowski, 1993). This preferential repair has been shown to be defective in Cockayne's Syndrome, unlike XP where either both modes of repair are defective or solely the overall repair (Venema et al., 1990; Hanawalt, 1991). Not all active genes have been shown to be repaired preferentially and whether this is a process that is commonly used for all types of lesions has yet to be established (Vos, 1992).

A cell is able to recognise and respond to damage in its DNA in a number of ways. As well as initiating repair processes necessary to remove the damage from the genome, certain checkpoints can be induced to arrest the progression of the cell through the cell cycle until repair is complete. In the yeast *S.cerevisiae*, the RAD9 gene product is essential for the damage-arrest response. Mutations in this gene have been shown to affect both the G1 arrest, i.e. the prevention of the initiation of replication on a damaged template (Siede et al., 1993), and the G2 arrest, where the cell is prevented from undergoing mitosis until the genome is fully repaired (Weinert and Hartwell, 1988). These responses are now being assessed in mammalian systems. In AT cells, the normal G1 arrest in response to damage is defective, allowing accumulation of mutations through continued replication (Painter and Young, 1980). The G1 arrest normally seen in response to ionizing radiation is mediated by an increase in p53 (Kuerbitz et al., 1992), which in turn effects the induction of GADD45. In AT cells, there is no increase in p53 protein and therefore no G1 arrest, indicating the key role that the protein plays in this response (Kastan et al., 1992).

As discussed, the processes of DNA metabolism are interlinked, share some similar features, and certain proteins are involved in more than one process. Another characteristic common to replication, excision repair and recombination is the generation of breaks in one or both strands of the DNA double helix. DNA ligases were first discovered in *E. coli* and T4 bacteriophage as enzymes that were able to seal single strand nicks in double stranded DNA (Gefter et al., 1967; Gellert, 1967; Olivera and Lehman, 1967; Weiss and Richardson, 1967). A mammalian DNA ligase, later termed DNA ligase I, was identified soon after (Lindahl and Edelman, 1968). DNA ligase mutants of both *E. coli* and yeast have allowed further characterisation of the role of these enzymes and established them

as essential for correct metabolism and maintenance of DNA, and for viability (Gellert and Bullock, 1970; Modrich and Lehman, 1971; Johnston and Nasmyth, 1978).

In the 1960's, it was established that the T4 DNA ligase required ATP for catalysis whereas the *E. coli* enzyme utilised NAD^+ as its cofactor (Little et al., 1967; Weiss and Richardson, 1967). Since then, eukaryotic DNA ligases have also been shown to be ATP-dependent with only bacterial DNA ligases using NAD^+, although all require Mg^{2+} for activity (Lindahl and Edelman, 1968). Studies on the enzymes from *E. coli* and T4 bacteriophage allowed the elucidation of the mechanism of action and the requirements for active catalysis. Aside of the difference in cofactor requirements, the overall reaction mechanism is the same for all DNA ligases (Lehman, 1974), and is shown schematically in Fig. 1.5.

The first stage of the DNA ligation reaction involves the interaction of the enzyme with the cofactor to produce an enzyme-AMP complex, with the release of either PP_i or NMN. The adenylyl group is covalently linked to the enzyme through a phosphoamide bond to the ε-amino group of a reactive lysine (Gumport and Lehman, 1971). As this is a covalent linkage, the intermediate can be easily isolated and characterised by producing, in the absence of DNA, a radiolabelled enzyme-AMP reaction intermediate using [α-^{32}P]ATP. The complex can be separated using SDS-PAGE and the products of the reaction visualised by autoradiography (for example see Tomkinson et al., 1990). This assay is used frequently in the purification and identification of the different DNA ligases. Enzymes other than the DNA ligases have also been shown to utilise this form of reaction intermediate, including the T4 RNA ligase, the yeast tRNA ligase and an mRNA capping enzyme from vaccinia virus involved in the

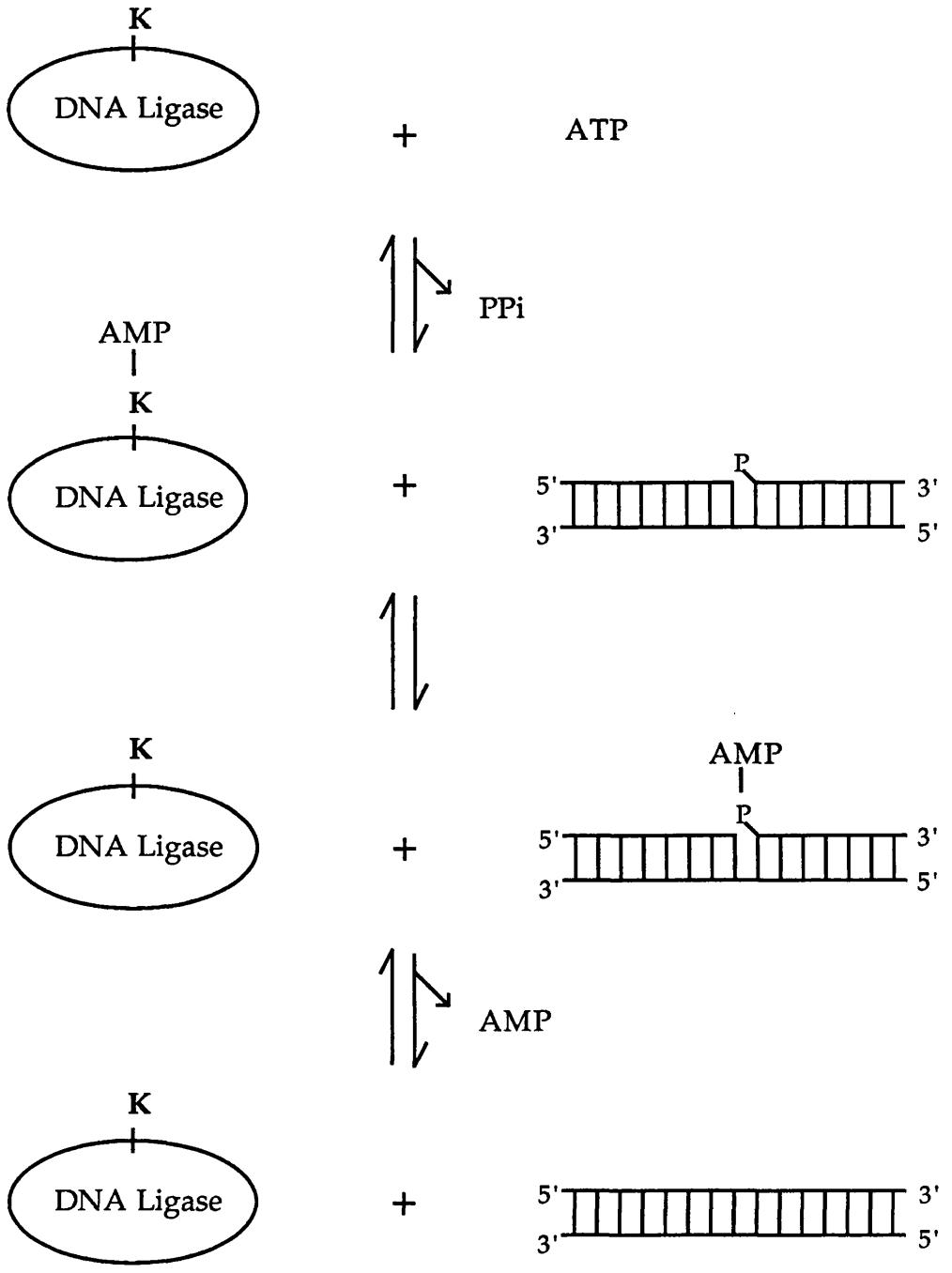


Fig. 1.5. Reaction Mechanism of DNA Ligases

transfer of GMP from GTP to the 5' terminus of RNA via an enzyme-GMP intermediate (Greer et al., 1983; Thorgersen et al., 1985; Cong and Shuman, 1993).

The second step of the reaction involves the transfer of the AMP group from the lysine of the enzyme to the 5'-phosphate of a single strand break in double stranded DNA to generate a short-lived covalent DNA-AMP complex with a 5'-5' phosphoanhydride bond. Using endonuclease I, which digests DNA in a 3'→ 5' direction, it was possible to determine that the AMP group was attached to the 5' end of a radiolabelled DNA-AMP intermediate. Mg^{2+} is still required at this stage of the reaction, as is a 3'-hydroxyl group at the DNA nick (Soderhall, 1975; Yang and Chan, 1992). The final step of the reaction involves attack of the 3'-hydroxyl group on the DNA-AMP intermediate, releasing AMP and forming a new phosphodiester bond (Lehman, 1974; Soderhall, 1975).

The reactive group within the DNA ligase enzyme that allows formation of the AMP complex was identified as a lysine residue (Little et al., 1967). For T4 RNA ligase, which catalyses phosphodiester bond formation between single stranded RNA or DNA not hybridised to a complementary strand, and DNA ligase I from mammalian cells, the particular lysine residue involved in intermediate formation has been determined. The T4 RNA ligase was incubated with [^{14}C]ATP, the RNA ligase-AMP intermediate was digested with chymotrypsin and the radiolabelled active site fragment isolated. The identity of the peptide within the protein sequence was determined by mass spectrometry, giving the labelled lysine at position 99 in the protein (Thorgersen et al., 1985). The active site lysine for mammalian DNA ligase I was elucidated in a similar manner (Tomkinson et al., 1991b), by isolation and amino acid sequencing of a tryptic active site peptide labelled with tritiated AMP. In this situation, trypsin is unable to cleave at the lysine-AMP residue, as

confirmed by the presence in the isolated peptide of an uncleaved lysine residue representing the active site of this enzyme (Tomkinson et al., 1991b). Site-specific mutagenesis of this active site lysine residue results in loss of enzyme activity, confirming its role in the reaction (Kodama et al., 1991). It has been possible to predict active sites in other cloned and sequenced DNA ligases by comparison with the DNA ligase I peptide sequence (Tomkinson et al., 1991b) and to identify a consensus active site motif, KXDG, as can be seen in Fig. 1.6.

All DNA ligases that have so far been cloned and sequenced contain a lysine residue at their active sites, and even the NAD-dependent enzymes retain elements of a wider consensus sequence, indicating a high conservation of function throughout these diverse species. The mRNA capping enzyme from vaccinia virus also contains the KXDG motif, indicating a common mechanism to the nucleotide transfer reactions catalysed by these enzymes. Aligning the peptide sequences of the ligases around the active site lysine indicates a striking conservation of length from the active site to the C-terminus (Tomkinson et al., 1991b). Even the smallest of the ATP-dependent DNA ligases seems to have preserved this feature, suggesting that there may be a minimum size of catalytic region necessary to retain activity. The only DNA ligase not conforming to this pattern is the NAD-dependent DNA ligase from *E. coli*. In contrast, the N-terminal regions of the proteins show a large variation in size and sequence, that may give rise to differences in enzyme regulation via species-specific interactions with other proteins. The presence of the active site motif in DNA ligases from such varied organisms indicates conservation of reaction mechanism during evolution, and the central and essential role that DNA ligases play within the cell.

Human	561	A A F T C E Y **K** Y D G Q R A Q I H	577
Sch. pombe	409	A A F T C E Y **K** Y D G E R A Q V H	425
S. cerevisiae	412	E T F T S E Y **K** Y D G E R A Q V H	428
Vaccinia	224	S G M F A E V **K** Y D G E R V Q V H	240
ASFV	144	T D P I V H G **K** R N G V R A V A C	160
D. ambivalens	254	N I A L V D Y **K** Y D G E R G Q I H	270
T7	27	G Y L I A E I **K** Y D G V R G N I C	43
T3	27	G Y L I A D C **K** Y D G V R G N I V	43
T4	152	F P A F A Q L **K** A D G A R C F A E	168
E.coli	108	V T W C C E L **K** L D G L A V S I L	124
T.thermophilus	111	F A Y T V E H **K** V D G L S V N L Y	127
Z.mobilis	137	V I C T V E P **K** I D G L S C S L R	153
T4 RNA ligase	92	D V D Y I L T **K** E D G S L V S T Y	108
tRNA ligase	107	G P Y D V T I **K** A N G C I I F I S	123

Fig. 1.6. **Comparison of amino acid sequences in the vicinity of the active site lysine residue in DNA and RNA ligases** - The peptide sequences for African swine fever virus (ASFV), the archaeon *Desulfurolobus ambivalens* and *Zymomonas mobilis*, are from (Hammond et al., 1992), (Kletsin, 1992) and (Shark and Conway, 1992) respectively; references to other sequences are in (Lindahl and Barnes, 1992) and (Tomkinson et al., 1991b).

DNA ligase I was the first DNA ligase to be identified in mammalian cells (Lindahl and Edelman, 1968) and since then there has been much confusion over the exact size and number of different forms of this enzyme. Molecular mass estimates range from 85 to 480 kDa, and there have been reports of monomers, dimers and multiple forms of DNA

ligase I (Pedrali Noy et al., 1973; Teraoka et al., 1975; Zimmerman and Levin, 1975; Mezzina et al., 1984). The unusual structure of the enzyme, which is markedly asymmetric, allows it to run anomalously slowly on gel filtration and may account for much of the confusion (Tomkinson et al., 1990). Cloning of the cDNA encoding this enzyme has established that human DNA ligase I is a protein of 919 amino acids with a molecular mass of 102 kDa, although it migrates on SDS-PAGE as a 125 kDa polypeptide due, in part, to a high proline content (Barnes et al., 1990). The full length protein is susceptible to limited proteolysis which removes the N-terminal portion, leaving an 85 kDa C-terminal catalytic domain. This proteolytic fragment retains full activity and may have contributed to the confusion over multiple forms of DNA ligase I (Tomkinson et al., 1990). Immunoblotting analysis of the full length protein in extracts from rapidly lysed cells has confirmed the apparent molecular mass of the protein on SDS-PAGE to be 125 kDa (Lasko et al., 1990).

The cDNA encoding human DNA ligase I was cloned by two independent methods (Barnes et al., 1990). The amino acid sequences of tryptic peptides from the purified bovine protein were used to design oligonucleotides which allowed isolation of clones encoding DNA ligase I from a human cDNA library. Functional complementation of a *cdc9* conditional-lethal DNA ligase mutant of *S. cerevisiae*, also identified the human DNA ligase I cDNA. The 53 kb unique gene encoding human DNA ligase I is located on chromosome 19q13.2-13.3 and produces a single transcript of 3.2 kb (Barnes et al., 1992a; Noguiez et al., 1992). Comparison of the amino acid sequence derived from the DNA ligase I cDNA with the DNA ligases of *S. cerevisiae* and *Schizosaccharomyces pombe* shows homology between the three proteins throughout their C-terminal regions, whereas the N-terminal regions diverge greatly, suggesting a species-specific role for this domain (Tomkinson et al., 1991b).

The 85 kDa C-terminal fragment of DNA ligase I is able to complement a conditional-lethal *E. coli* DNA ligase mutant, *lig251* (Kodama et al., 1991). A slightly smaller fragment of 80 kDa is also active, but deletion of the N-terminus by more than this amount results in complete loss of activity. Similarly, a 78 kDa C-terminal fragment of the *S. cerevisiae* CDC9 DNA ligase remains active (Tomkinson et al., 1992). The very C-terminus of the protein also seems to be essential for activity. Removal of the last 14 amino acids can be tolerated but removing a further 8 residues causes loss of activity. The conservation of this C-terminal 'tail' between human, bovine, yeast and vaccinia DNA ligases reinforces its importance to the enzyme.

The enzyme forms a covalent enzyme-AMP complex that migrates at 125 kDa on SDS-PAGE, with a K_m for ATP of ~ 1 μM (Tomkinson et al., 1990), a property that is retained by the 85 kDa proteolytic fragment which will readily form the AMP complex (Tomkinson et al., 1990). The full-length enzyme-AMP complex has been isolated and shown to ligate nicked DNA in the absence of additional ATP (Soderhall and Lindahl, 1973a). The DNA-AMP intermediate has also been isolated. Addition of unadenylylated enzyme to this complex results in release of AMP and ligation of the nicks (Soderhall, 1975). Thus, the reaction mechanism for the mammalian DNA ligase is identical to that described for the microbial DNA ligases. However, DNA ligase I has a different substrate specificity to that of T4 DNA ligase, being unable to ligate the hybrid substrate oligo(dT)/poly(rA) (Arrand et al., 1986). It is able to ligate blunt ends, an activity that is increased by macromolecular crowding in the presence of polyethylene glycol (Zimmerman and Pheiffer, 1983; Arrand et al., 1986). DNA ligase I can also act as a weak topoisomerase under certain conditions in a reversal of the last step of the ligation reaction (Montecucco et al., 1988).

An inhibitor of the DNA joining activity of DNA ligase I has been identified from HeLa cells (Yang et al., 1992). The inhibitor is a 65 kDa protein, as estimated by gel filtration, that copurifies with DNA ligase I, suggesting some protein:protein interaction. The two proteins can be separated by Mono S chromatography and in this way the inhibitor has been characterised further. Although it affects the DNA joining activity, formation of the enzyme-AMP complex remains normal. This allows DNA ligase I activity to be assayed through adenylylation even in the presence of the inhibitor. An association of the two proteins *in vivo* has not been established, although it is possible that this inhibitor is a means of regulating cellular DNA ligase I activity. Information on the levels of the inhibitor during the cell cycle may help to establish a definite role.

Another possible form of regulation of DNA ligase I activity is through phosphorylation. DNA ligase I enzyme immunoprecipitated from cells labelled with ^{32}P orthophosphate is radiolabelled (Prigent et al., 1992). If this radiolabelled protein is treated with pyrophosphate, a proportion of the radioactivity is lost, indicating that some of the enzyme has incorporated ^{32}P through adenylylation by [α-^{32}P]ATP formed *in vivo*. Further treatment of the enzyme with a phosphatase releases the rest of the radioactivity, confirming that the DNA ligase I is phosphorylated *in vivo*. Casein kinase II is able to phosphorylate DNA ligase I *in vitro* producing an increase in activity. The sites of phosphorylation by casein kinase II have been identified and shown to be on the N-terminal portion of the DNA ligase I protein (C. Prigent, personal communication). This region is thought to regulate the activity of DNA ligase I through phosphorylation and a model has been proposed (see Fig. 1.7). In the inactive form of the enzyme, the N-terminal region is positioned to block the reactive lysine residue at the active site. Phosphorylation of the N-terminal region causes a conformational change that results in exposure of

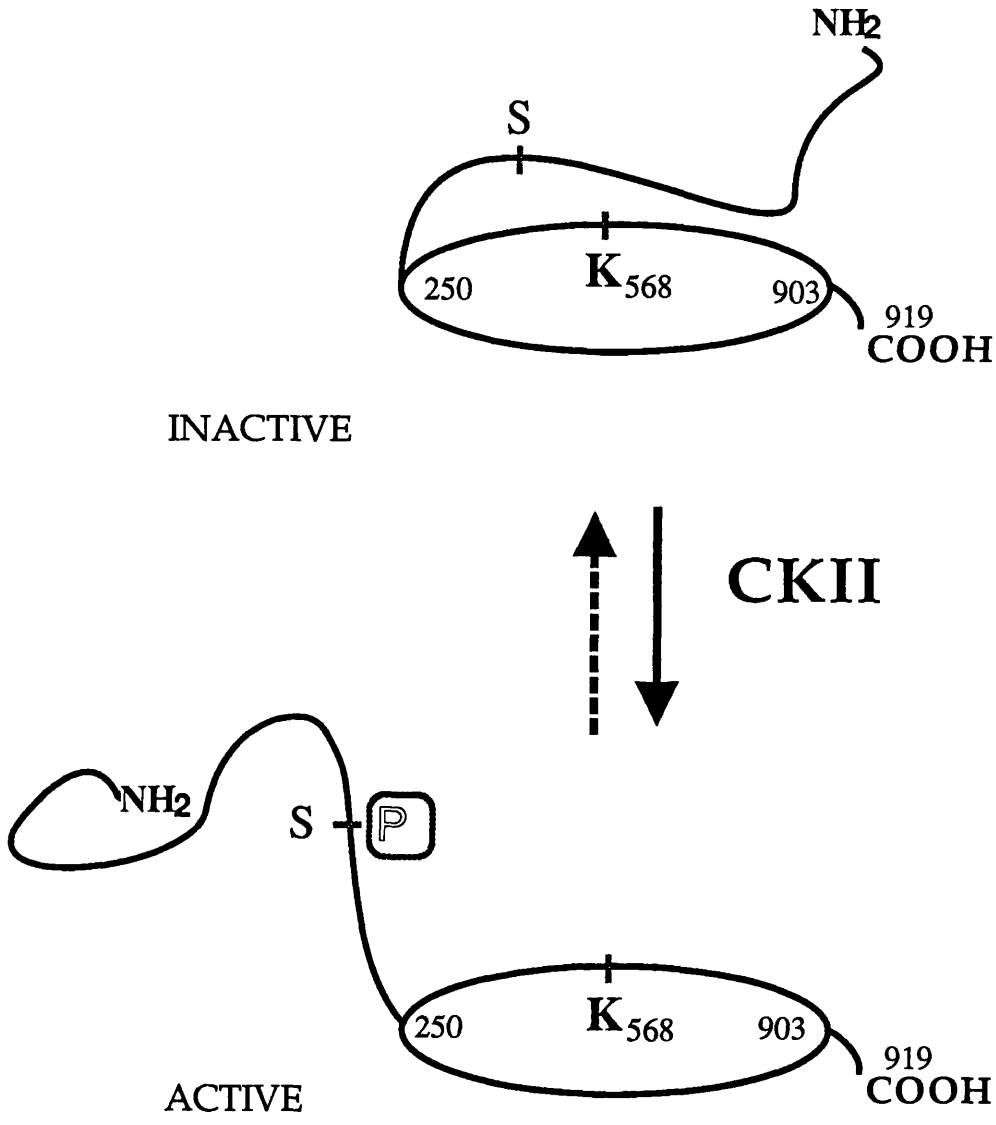


Fig. 1.7. Phosphorylation of DNA Ligase I by Casein Kinase II (taken from Prigent et al., 1992)

the active site and reactivation of the enzyme. Furthermore, expression of full length DNA ligase I does not complement a *lig251* mutant of *E. coli* (Kodama et al., 1991). As *E. coli* does not contain a casein kinase II activity, these data could be explained by the DNA ligase I protein being in the inactive dephosphorylated form. In contrast, the 85 kDa C-terminal fragment is able to complement the *E. coli* mutant as it is constitutively active, having no regulatory domain. Phosphorylation, by casein kinase II, of the inactive, full-length protein expressed in *E. coli* results in enzyme activation, confirming this model (Prigent et al., 1992). Phosphorylation may be an important form of regulatory control of DNA ligase I *in vivo* as the protein has a long half life of 7 h, so that control at the transcriptional level would have little effect during the cell cycle (Lasko et al., 1990).

Cell staining studies using antibodies against the purified bovine DNA ligase I have shown it to be located in the nucleus (Lasko et al., 1990). Further work with virus-infected cells showed that DNA ligase I was recruited to the sites of viral replication (Wilcock and Lane, 1991). These data, together with the observation that both the activity and protein levels increase in rapidly dividing cells (Soderhall, 1976), suggest a role for DNA ligase I in DNA replication. DNA ligase mutants of *S. cerevisiae* and *E. coli* all show a DNA replication defect (Modrich and Lehman, 1971; Johnston and Nasmyth, 1978) and this is effectively complemented by human DNA ligase I, further supporting the idea that this enzyme is involved in replication (Barnes et al., 1990; Kodama et al., 1991). The *cdc9* conditional-lethal DNA ligase mutant of *S. cerevisiae* also shows a defect in excision repair and mitotic recombination, indicating that the enzyme is also involved in these processes (Johnston and Nasmyth, 1978; Fabre and Roman, 1979).

The role of DNA ligase I in mammalian cells has been demonstrated by the identification of a patient with an inherited defect in

DNA ligase I (Barnes et al., 1992b). The patient showed stunted growth, sun sensitivity and severe immunodeficiency, and died aged 19 with lymphoma (Webster et al., 1992). The cell line established from this patient, 46BR, is hypersensitive to a wide range of DNA-damaging agents, such as ionizing radiation, UV-light and alkylating agents (Teo et al., 1983a). The cells are also hypersensitive to killing by 3-aminobenzamide, an inhibitor of poly(ADP-ribose) polymerase, show delayed joining of Okazaki fragments during DNA replication (Henderson et al., 1985; Lonn et al., 1989) and delayed break rejoining after DNA damage (Squires and Johnson, 1983; Teo et al., 1983b; Lehmann et al., 1988). These observations suggested that there may be a defect in DNA ligation. Sequencing of the DNA ligase I cDNA from this cell line revealed two distinct missense mutations, one in each allele of the gene. One allele contains a mutation close to the active site lysine and site-directed mutagenesis analysis has indicated that this would lead to a complete loss of enzyme activity. The second allele carries an Arg→Trp change and appears to produce a partially activate enzyme with only 5 % of normal activity. The defect in the reaction mechanism appears in the last step where the DNA ligase I from 46BR is unable to convert the DNA-AMP intermediate and therefore cannot complete ligation (Prigent et al., 1994). The role of DNA ligase I in lagging strand DNA replication and excision repair has been analysed in 46BR using *in vitro* assays. These have indicated delayed joining of strand breaks, with the resultant accumulation of Okazaki fragments and generation of anomalously long repair patches (Prigent et al., 1994) The severe immunodeficiency seen in 46BR appears not to be due to the defect in DNA ligase I affecting V(D)J recombination and the possiblity remains that another DNA ligase is involved in this process (Hsieh et al., 1993).

The inherited cancer-prone disease, Bloom's syndrome, shows defects in DNA ligation (Chan et al., 1987; Willis and Lindahl, 1987;

Runger and Kraemer, 1989). Cells from these patients show a hyper-recombination and hyper-mutable phenotype, as do the yeast DNA ligase mutants (Nasmyth, 1977; Johnston and Nasmyth, 1978; Langlois et al., 1989). This has not been correlated with a defect in the structural gene for DNA ligase I and the molecular defect in Bloom's syndrome has yet to be established (Petrini et al., 1991; Barnes et al., 1992b).

Mammalian cells contain multiple DNA ligases. This may reflect the increased complexity of DNA replication and excision repair in eukaryotic cells. DNA ligase II was first observed as a separate peak of activity, distinct from DNA ligase I, on hydroxylapatite chromatography (Soderhall and Lindahl, 1973b). The two enzymes were shown to be serologically distinct using antiserum directed against DNA ligase I (Soderhall and Lindahl, 1975) and later against DNA ligase II (Teraoka and Tsukada, 1986). Many groups could not identify a second ligase, possibly due to inappropriate extraction conditions or loss of activity during enzyme assays (Bertazzoni et al., 1972; Teraoka et al., 1979) and this led to some confusion as to whether two enzymes do indeed exist in mammalian cells. The establishment of a specific assay for DNA ligase II confirmed the existence of a second ligase and allowed identification of this enzyme in different species and tissues (Arrand et al., 1986). Both DNA ligase I and II are able to ligate oligo(dT) hybridised to poly(dA), but only DNA ligase II is able to ligate oligo(dT)/poly(rA).

DNA ligase II is a much smaller enzyme than DNA ligase I with a molecular weight of 68-72 kDa (Teraoka et al., 1986; Tomkinson et al., 1991a). During purification, the molecular weight of DNA ligase II remains constant and the enzyme does not seem susceptible to proteolysis in the same way as DNA ligase I, having no catalytically active proteolytic fragment. DNA ligase II shares the same reaction mechanism as other DNA ligases, requiring Mg^{2+} and using ATP as a cofactor, although it has a

much higher K_m for ATP compared to DNA ligase I, of approximately 40 μM (Teraoka et al., 1986). This, together with the differences in substrate specificity, may imply dissimilarity in their active sites and different roles *in vivo*. A proteolytic digest of the enzyme-AMP complex of DNA ligase I results in discrete fragments, none of which are of a similar size to DNA ligase II, implying that DNA ligase II is not a proteolytic product of DNA ligase I (Tomkinson et al., 1990). More recently there has been a report of a common domain between the two DNA ligases (Yang et al., 1990). A V8 protease digest of the two adenylylated DNA ligases apparently produced a 22 kDa polypeptide from both the enzymes. This led the authors to suggest that DNA ligases I and II may be derived from a common ancestral gene and share siginificant homology, contradicting the biochemical evidence that suggests the two enzymes are distinct.

The major DNA ligase activity in other organisms seems most closely to resemble the DNA ligase I from mammalian cells. The DNA ligases from *S. cerevisiae*, *S. pombe* and vaccinia virus have been cloned and show significant homology both to each other and to the human DNA ligase I, particularly in the region of the active site (Barker et al., 1985; Barker et al., 1987; Smith et al., 1989; Tomkinson et al., 1991b). The similarity of function between these enzymes can be seen in the ability of the human DNA ligase I to complement the DNA ligase mutant of *S. cerevisiae* (Barnes et al., 1990). The DNA ligase of *S. cerevisiae*, is similar to the bovine DNA ligase I in that it readily produces a catalytically active proteolytic fragment (Tomkinson et al., 1992). However, a second DNA ligase has been described in *S. cerevisiae*, *Xenopus laevis* and *Drosophila melanogaster* (Takahashi and Senshu, 1987; Aoufouchi et al., 1992; Tomkinson et al., 1992). These enzymes closely resemble mammalian DNA ligase II in their biochemical properties.

Unlike DNA ligase I, the level of DNA ligase II activity remains constant irrespective of DNA replication in both *D. melanogaster* and mammalian cells (Soderhall, 1976; Takahashi and Senshu, 1987). Although the cellular function of DNA ligase II is unknown, this observation has led some groups to postulate that DNA ligase II may be involved in DNA repair (Teraoka et al., 1986). The involvement of poly(ADP-ribose) in activation of DNA ligase II in response to DNA damage has been reported but has yet to be confirmed (Creissen and Shall, 1982; Ohashi et al., 1983). In addition DNA ligase II has been proposed to have a role in meiotic recombination. A mutant of the DNA ligase I homologue in *S. pombe, cdc17*$^{+}$, has impaired DNA replication, repair and mitotic recombination but normal levels of meiotic recombination, suggesting that a second DNA ligase is capable of fulfilling this role (Nasmyth, 1977; Sipiczki et al., 1990). In agreement with this, Higashitani and coworkers have observed an increase in DNA ligase II activity during meiosis in mouse testis, in line with other proteins known to be involved in meiotic recombination (Higashitani et al., 1990). Further characterisation of DNA ligase II is necessary to elucidate the possible role of multiple DNA ligases in mammalian cells.

Aims

Several groups had failed to detect DNA ligase II in mammalian cells and had dismissed it as a proteolytic product of DNA ligase I. Confusion existed as to the exact relationship of the two enzymes and the degree of similarity between them. My aim was to establish that DNA ligase II is indeed a distinct protein, to further characterise this enzyme, and to examine its relationship to DNA ligase I at the biochemical, immunological and catalytic levels. During the course of this work on DNA ligase II, a third DNA ligase, DNA ligase III, was identified in the

laboratory. Some characterisation experiments have included this enzyme and its relationship to DNA ligase II is discussed.

CHAPTER 2

MATERIALS AND METHODS

2.1 Reagents

Laboratory reagents and solvents, including protease inhibitors, were from Sigma, BDH, and Fisons, or as specified in the text. Commercial enzyme preparations were from New England Biolabs Inc. All radiolabelled reagents, including ^{14}C-radiolabelled protein molecular mass markers, were from Amersham. Unlabelled protein markers were from Bio-Rad Laboratories Ltd. X-ray film was from Fuji (RX) and Kodak (X-OMAT).

2.2 Tissues

Calf thymus glands (3-6 month old calves), and bovine liver (18 months to 2 yr old steers), were obtained at the local abattoir. The tissues were packed in ice and used for initial enzyme preparation within 3 h.

2.3 Enzyme Assays

2.3.1 Preparation of polynucleotide substrates

Oligo$(dT)_{30}$ was synthesized on an Applied Biosystems model 380B DNA synthesizer and purified by HPLC. Oligonucleotides $(dT)_{16}$ and $(rA)_{12\text{-}18}$, and polynucleotides (dA), (dT) and (rA) were from Pharmacia LKB Biotechnology Inc. Oligo$(rA)_{16}$ was purified from oligo$(rA)_{12\text{-}18}$ by gel purification on a denaturing 20% polyacrylamide/8M urea gel. Oligonucleotides (10 µg) were radioactively labelled with 100 µCi [γ-^{32}P] ATP (>5,000 Ci/mmol, Amersham) and 10 µmol unlabelled ATP using T4 polynucleotide kinase (New England Biolabs Inc.). The labelled oligonucleotide was mixed with an equimolar amount of polynucleotide, incubated at 90°C for 10 min and then slowly cooled to room temperature. Non-radioactively labelled substrates were prepared as described above with the omission of [γ-^{32}P]ATP.

2.3.2 DNA ligation assays

Reaction mixtures (60 µl) contained 60 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM DTT, 1 mM ATP, 50 µg/ml nuclease-free bovine serum albumin, polynucleotide substrate (20,000 cpm) and a limiting amount of DNA ligase. Incubations were at 16°C for 15 min. The conversion of [5'-^{32}P]-labelled phosphomonoesters to alkaline phosphatase-resistant diesters was measured (Lasko et al., 1990; Prigent et al., 1990). One unit of DNA ligase activity catalyses the conversion of 1 pmol terminal phosphate residues to a phosphatase resistant form in 15 min at 16°C.

2.3.3 Analysis of ligation products

DNA ligation assays were carried out as described above. The reactions were stopped by heating at 90°C for 10 min. An aliquot (5 µl) was heated for 5 min at 100°C in 65% formamide prior to loading onto a denaturing, 20% acrylamide/8M urea gel. After electrophoresis for 3 h at 300 V (when monitoring purification procedures, smaller gels were used and run for 30 min), gels were fixed in 40% methanol, 10% acetic acid for 30 min and then dried. Oligonucleotides were visualised by autoradiography. Using this system, the DNA ligase activity is expressed as units based on a T4 DNA ligase control (New England Biolabs Inc.) run on the same gel.

2.3.4 Formation of DNA ligase-adenylate

Reaction mixtures (10 µl) contained 60 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM DTT, 50 µg/ml bovine serum albumin, 0.5 µCi [α-^{32}P] ATP (3000 Ci/mmol, Amersham) and DNA ligase. Incubations were at room temperature for 15 min. After the addition of SDS sample buffer, reaction mixtures were heated at 90°C for 10 min. Proteins were separated by SDS-PAGE (Laemmli, 1970) on a 10 % gel (unless otherwise stated) at 200 V using the Bio-Rad Mini Protean II apparatus. Gels were fixed for 10

min in 10% acetic acid, dried and adenylylated proteins detected by autoradiography.

2.3.5 Reactivity of the enzyme-adenylate intermediate

To examine the reactivity of this intermediate, the adenylylation reaction was performed as described above in a final volume of 40 µl. Aliquots (10 µl) were incubated with; 0.8 µg unlabelled oligo(dT)/poly(dA), 0.8 µg unlabelled oligo(dT)/poly(rA), 1 mM sodium pyrophosphate or TE for 45 min at 37°C. The reactions were stopped by the addition of SDS sample buffer and the reaction mixtures were then heated at 90°C for 10 min. Adenylylated polypeptides were separated by 7.5 % SDS-PAGE and visualised by autoradiography (Laemmli, 1970)

2.4 Purification Schemes

2.4.1 Purification of DNA ligase II from calf thymus

All procedures were carried out at 0-4°C and centrifugations were at 10,000 x *g* for 30 min unless otherwise stated. Calf thymus glands (3 kg) were disrupted in aliquots by homogenisation for 3 x 30 sec in a Waring blendor with 3 litres of a buffer containing 0.2 M NaCl, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 10 mM β-mercaptoethanol, 1 mM phenylmethylsulphonylfluoride (PMSF), 1.9 µg/ml aprotinin and 0.5 µg/ml each of leupeptin, pepstatin, chymostatin and Nα-tosyl-L-lysine chloromethyl ketone (TLCK). After gentle stirring for 1 h, cellular debris was removed by centrifugation. The supernatant (Fraction I, crude extract) was diluted to a final NaCl concentration of 20 mM by the addition of 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 10 mM β-mercaptoethanol. This was batch adsorbed to 3.5 litres of a thick slurry (500g dry weight) of P_{11} phosphocellulose (Whatman), which had been pre-equilibrated with 20 mM NaCl, 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.5 mM DTT (buffer A). After washing the phosphocellulose with 10 litres of buffer A, adsorbed

proteins were eluted with 4 litres of buffer A containing 0.7 M NaCl. Ammonium sulphate was added to the eluate (Fraction II) at 231 g/l and 1 M Tris base was added as necessary to maintain the pH at 7.1-7.5. After 30 min of gentle stirring, the precipitate was removed by centrifugation. Additional ammonium sulphate (160 g/l) was added to the supernatant and the resulting suspension was again neutralised and stirred gently for 30 min. The precipitate was collected by centrifugation, divided into six aliquots and quickly frozen on dry ice. These precipitates could be stored at -70°C with no significant loss of DNA ligase activity after several weeks. This procedure avoided the use of frozen thymus glands which yield extracts containing higher levels of protease activity than fresh glands. One aliquot of the ammonium sulphate precipitate (material from 500 g of tissue) was resuspended in 15 ml and dialysed for 3 h against 50 mM potassium phosphate (pH 7.5), 0.5 mM DTT (buffer B). Insoluble material was removed by centrifugation (Fraction III, 25ml). Fraction III was loaded onto a 6 x 7.5 cm hydroxylapatite (Bio Rad, HT) column which had been equilibrated with buffer B. Proteins were eluted by steps of 250 mM potassium phosphate (pH 7.5), 0.5 mM DTT and 400 mM potassium phosphate (pH 7.5), 0.5 mM DTT and assayed for DNA ligation activity and for enzyme-adenylate complex formation. The 250 mM potassium phosphate (pH 7.5) eluate contained the majority of the DNA ligase I and DNA ligase III activities, whereas the 400 mM potassium phosphate (pH 7.5) eluate contained the DNA ligase II activity. This pool was concentrated by ultrafiltration using a YM-10 membrane in a 50 ml filtration unit (Amicon) and a pressure of 50 psi. The concentrate (Fraction IV, 6 ml) was then dialysed into 50 mM Tris-HCl (pH 7.5), 1 M NaCl, 0.5 mM DTT, 1 mM EDTA, 0.2% Tween 20 (buffer C). The dialysed Fraction IV was centrifuged for 15 min and loaded onto a 0.9 x 97.5 cm Ultrogel AcA 44 column (Pharmacia LKB Biotechnology Inc.), pre-equilibrated with buffer C. After

elution with buffer C, fractions were assayed for DNA ligation activity and for enzyme-adenylate complex formation. Active fractions were pooled and dialysed against 50 mM Tris-HCl (pH 7.5), 0.9 M $(NH_4)_2SO_4$, 0.5 mM DTT, 1 mM EDTA, 0.2% Tween 20 (buffer D). This pool (Fraction V, 8 ml) contained two major polypeptides of approximately 72 kDa and 74 kDa. After centrifugation for 15 min, Fraction V was loaded onto a FPLC Phenyl Superose column (Pharmacia), pre-equilibrated with buffer D. DNA ligase II activity and the 72 kDa polypeptide passed through the column (Fraction VI, 10ml), whereas the 74 kDa polypeptide was retained. Fraction VI was dialysed against 50 mM Tris-HCl (pH 7.5), 30 mM NaCl, 0.5 mM DTT, 1 mM EDTA, 0.2% Tween 20 (buffer E) and then loaded onto a FPLC Mono Q column (Pharmacia), pre-equilibrated in buffer E. DNA ligase II activity passed through the column whereas the remaining traces of DNA ligases I and III were retained by the column. The pass-through fractions were pooled (Fraction VII, 20ml), and loaded onto a FPLC Mono S column (Pharmacia) pre-equilibrated in buffer E. Proteins were eluted with a linear gradient of 30 - 500 mM NaCl in buffer E. Fractions containing DNA ligase II activity, which eluted at 0.22 - 0.25 M NaCl (Fraction VIII), were dialysed against buffer E containing 50 % glycerol and stored at -20°C.

2.4.2 Partial purification of DNA ligase II from bovine liver

Bovine liver (2.5 kg) was disrupted in aliquots by homogenisation for 2 x 20 sec in a Waring blendor with 3 litres of a buffer containing 0.2 M NaCl, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 10 mM β-mercaptoethanol, 1 mM PMSF, 1.9 µg/ml aprotinin and 0.5 µg/ml each of leupeptin and pepstatin, 1 µg/ml of chymostatin and 5 µg/ml of TLCK. After gentle stirring for 1 h, cellular debris was removed by filtration through two layers of gauze, followed by centrifugation at 10,000 x *g* for 90 min. The supernatant was filtered through one layer of Miracloth (Calbiochem) then diluted to a final NaCl concentration of 20 mM by the addition of 50

mM Tris-HCl (pH 7.5), 1 mM EDTA, 10 mM β-mercaptoethanol. Proteins were batch adsorbed onto phosphocellulose, washed, eluted and precipitated with ammonium sulphate as described for calf thymus. One aliquot of the ammonium sulphate precipitate was resuspended in 10 ml of 50 mM Tris-HCl (pH 7.5), 1 M NaCl, 1 mM EDTA, 2 mM DTT (buffer B), centrifuged at 8,000 x *g* for 15 min. and loaded onto a 2.5 cm x 110 cm Ultrogel AcA 34 (Pharmacia) gel filtration column equilibrated in buffer B. Proteins were eluted with buffer B and fractions were assayed for DNA ligase activity. Active fractions were pooled and adjusted to 1 mM potassium phosphate (pH 7.5) and loaded on to a 6 cm x 4.5 cm hydroxylapatite (Bio-Rad, HT) column which had been equilibrated with 50 mM potassium phosphate (pH 7.5), 2 mM DTT. Proteins were eluted by steps of; 100 mM potassium phosphate (pH 7.5), 2 mM DTT; 250 mM potassium phosphate (pH 7.5), 2 mM DTT and 500 mM potassium phosphate (pH 7.5), 2 mM DTT. The 100 mM and 250 mM potassium phosphate eluates contained the majority of the DNA ligase I and DNA ligase III activities, whereas the 500 mM potassium phosphate eluate contained the DNA ligase II activity. Active fractions were pooled and dialysed against 50 mM Tris-HCl (pH 7.5), 1.1 M $(NH_4)_2SO_4$, 2 mM DTT, 1 mM EDTA (buffer C). After centrifugation for 15 min, this fraction was loaded onto a FPLC Phenyl Superose column (Pharmacia), pre-equilibrated with buffer C. Proteins were eluted in a linear gradient from 1.1 M $(NH_4)_2SO_4$ to 0 M $(NH_4)_2SO_4$, 10 % glycerol in buffer C. DNA ligase II activity from this column was dialysed against 50 mM Tris-HCl (pH 7.5), 30 mM NaCl, 2 mM DTT, 1 mM EDTA, 10 % glycerol (buffer D), and then loaded onto a FPLC Mono S column (Pharmacia) pre-equilibrated in buffer D. Proteins were eluted with a linear gradient of 30 - 500 mM NaCl in buffer D. Fractions containing DNA ligase II activity, which eluted at 150

mM NaCl, were dialysed against buffer D containing 50 % glycerol and stored at -20°C.

2.4.3 Purification of DNA ligases I and III from calf thymus

DNA ligases I and III were kindly supplied by A.E. Tomkinson, C. Prigent, P. Robins and G. Daly, purified as described (Tomkinson et al., 1990; Tomkinson et al., 1991a).

2.5 Peptide Sequencing (Carried out by N.F. Totty, Ludwig Institute, Middlesex Hospital, London)

2.5.1 N-terminal amino acid determination

DNA ligase II (Fraction VIII from calf thymus) was dialysed into 100 mM Tris-HCl, pH 8.2, 0.1% SDS. Automated N-terminal sequencing of approximately 5 µg of protein was performed on an Applied Biosystems model 473A sequenator.

2.5.2 Sequencing of internal Lys-C peptides

DNA ligase II (Fraction VIII from calf thymus) was concentrated and the buffer exchanged into 100 mM Tris-HCl pH 8.2, 0.1% SDS using a Centricon-10 micro-concentrator (Amicon). The addition of SDS prevented precipitation of the intact protein and proteolytic fragments. Approximately 30 µg of the concentrated protein (1mg/ml), was digested using Lysylendopeptidase (WAKO) at a concentration of 1:30, enzyme/substrate (w/w), for 72 h at 30°C. Following peptide purification by ion-exchange/reversed phase HPLC (Kawasaki and Suzuki, 1990) the sequences of 14 peptides were obtained by microsequencing (Downward et al., 1984).

2.6 Two-Dimensional Gel Electrophoresis

Proteins were separated by the method of O'Farrell (O'Farrell, 1975). In the first dimension, proteins were resolved by isoelectric focusing using

5% polyacrylamide rod gels containing 2% ampholines (pH range 3.5-10, Pharmacia). The gels were then equilibrated in SDS sample buffer and proteins separated in the second dimension by electrophoresis through a 10% SDS-polyacrylamide gel (Laemmli, 1970).

2.7 Immunisation Procedures

Peptides were made on an Applied Biosystems 431A peptide synthesizer and were coupled to either keyhole limpet haemocyanin or thyroglobulin using the glutaraldehyde method (Harlow and Lane, 1988). Rabbit antibodies were raised against the coupled peptides by emulsifying 200 µg peptide conjugate with Freund's complete adjuvant and injecting subcutaneously. Subsequent injections were emulsified with Freunds incomplete adjuvant. The rabbit received 6-8 injections and was bled 10 days after the last injection. Antibodies against the whole protein were raised by injecting aliquots of 50 µg following the same protocol as that used for the peptides. The rabbit received 9 injections.

2.8 Immunoblots

Proteins were separated by SDS-PAGE (Laemmli, 1970) and then transferred to nitrocellulose filters (Schleicher and Schuell) in 25 mM Tris, 192 mM glycine, 20% methanol (pH 8.3). After transfer, membranes were incubated for 1 h at 37°C in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.05% Tween 20 (TBST) with 5% non-fat dry milk (Marvel) and then washed twice with TBST. Membranes were incubated for 12-16 h at 4°C with antibody diluted in TBST plus 2% non-fat dry milk. All subsequent steps were carried out at room temperature. Membranes were washed four times with TBST. [^{125}I]-linked goat anti-rabbit IgG secondary antibody (Amersham) was added (0.1 µCi/ml) and incubation continued for 3-4 h. Membranes were then washed with TBST, dried and exposed to film.

Alternatively, antigen-antibody complexes were detected using a secondary antibody coupled to alkaline phosphatase. After incubation in TBST plus 2% non-fat dry milk, membranes were incubated with goat anti-rabbit IgG antibody linked to alkaline phosphatase (1:500 dilution) (DAKO Immunoglobulins) for 2 h, washed with TBST, rinsed with water and blotted dry. Alkaline phosphatase activity was visualized by washing the membrane in 100 mM Tris-HCl (pH 9.5), 100 mM NaCl and 5 mM $MgCl_2$ followed by incubation in 10 ml of the same buffer containing 44 µl of 75 mg/ml p-nitrotetrazolium blue chloride (Life Technologies) and 33 µl of 50 mg/ml 5-bromo-4-chloro-3-indolylphosphate p-toluidine salt (Life Technologies). The reaction was stopped by the addition of 20 mM Tris-HCl (pH 8.0), 5 mM EDTA.

2.9 Analysis of Radiolabelled Active Site Peptides by Two-Dimensional Peptide Mapping

Radiolabelled peptides were isolated and analysed in a similar manner to that of Boyle et al (Boyle et al., 1991). Starting with 5-10 ng of DNA ligase II (from bovine liver), enzyme-AMP complexes were formed using [α-^{32}P]ATP and separated by SDS-PAGE. These were electrophoretically transferred onto Immobilon-P membrane (Millipore) and the protein visualised by autoradiography. The bands were cut out and the piece of membrane blocked by incubation in either 0.5 % polyvinylpyrrolidone in 100 mM acetic acid, or 0.2 % Tween 20 in PBS for 30 min at 37 °C, then washed with water followed by 50 mM ammonium bicarbonate. The protein was digested in 200 µl of 50 mM ammonium bicarbonate containing 10 µg of trypsin (sequencing grade, Boehringer Mannheim), at 37 °C for 2 h and then for 16 h at 37 °C with an additional 10 µg of trypsin. Peptides released from the membrane were lyophilised and washed twice with 100 µl of water. Initially peptides were oxidised by

incubating in 100 µl of fresh performic acid at 0 °C for 60 min, though in later experiments this step was omitted. Peptides were again lyophilised, washed twice with water and resuspended in the appropriate electrophoresis buffer. The peptides were run in two dimensions on silica TLC plates (Polygram Sil G, 805 013, Macherey-Nagel, Camlab) or in one dimension on cellulose TLC plates (PEI-cellulose F, 5579, Merck). The first dimension on the silica plates was electrophoresis at 750 V for 40 min or 1 kV for 1 h using the Hunter Thin Layer Electrophoresis System (C.B.S. Scientific Company, Inc.), in either 2.2 % formic acid, 7.8 % acetic acid (pH 1.9 electrophoresis buffer), 1 % ammonium carbonate (pH 8.9 electrophoresis buffer) or 10 % pyridine, 0.4 % acetic acid (pH 6.5 electrophoresis buffer). The second dimension was ascendant chromatography in 39.25 % n-butanol, 30.35 % pyridine, 6.1 % acetic acid for approximately 6 h. Cellulose TLC plates were used for chromatography with a buffer of 0.85 M KH_2PO_4, pH 3.4 for approximately 2 h. Radiolabelled products were visualised by autoradiography.

2.10 Analysis of Radiolabelled Active Site Peptides by HPLC

DNA ligases (approximately 5 ng) were adenylylated as described above and the reaction stopped by the addition of EDTA to a final concentration of 20 mM. To remove unincorporated ATP and to exchange the buffer, reactions were spun through Centricon-10's (Amicon), preadsorbed with 1 % dried milk overnight at room temperature, washed twice with 2 ml of 50 mM ammonium bicarbonate, and the retentate collected. This was digested with 10 µg of trypsin (sequencing grade, Boehringer Mannheim) for 16 h at 37 °C. The peptides were lyophilised and taken up in the start buffer (buffer A) for the HPLC. The resulting peptides were fractionated by reversed phase HPLC on a C18 column (Beckman Ultrasphere ODS, 4.6 mm x 25 cm) using a linear gradient from

15 mM ammonium bicarbonate, 2 % acetonitrile (Acetonitrile 190 (far UV), Romil Chemicals) (buffer A) to 15 mM ammonium bicarbonate, 80 % acetonitrile (buffer B), over 60 min at a flow rate of 1 ml/min. Absorbance was monitored at 220 nm. Fractions were collected and radioactivity determined by scintillation counting. Peptides were treated with iodoacetamide to reduce the -SH groups (Lane, 1978). After digestion, peptides were dried and resuspended in 50 mM DTT, 50 mM Tris-HCl (pH 7.5) and incubated for 30 min at 37°C. Iodoacetamide was added to 100 mM and samples reincubated for 30 min at 37°C. These were then dried and resuspended in the appropriate buffer for separation by HPLC.

2.11 Analysis of Radiolabelled DNA Ligases by Proteolytic Digestion

2.11.1 Partial digestion of adenylylated DNA ligases I and II using trypsin

Partially purified calf thymus DNA ligase I and bovine liver DNA ligase II were adenylylated as above and digested with limiting amounts (0-10 ng) of trypsin (sequencing grade, Boehringer Mannheim) at 37°C for 15 min to produce partial digestion products. The reaction was stopped by the addition of SDS sample buffer and heating at 90°C for 10 min. The products were separated by electrophoresis on 10-20 % gradient SDS-PAGE at a constant current of 12 mA. The dimensions of the gel (45 x 180 x 5 mm) allowed maximum separation of the products. Gels were fixed in 5 % MeOH, 7 % acetic acid, 5 % glycerol for 1 h before drying and autoradiography.

2.11.2 Complete digestion of adenylylated DNA ligases I and II by trypsin or Lys-C

Aliquots of the adenylylated DNA ligases I and II were digested with a 2.5 fold excess (w/w) of either trypsin or Lys-C (endoproteinase Lys-C, sequencing grade, Boehringer Mannheim) by incubation at 37°C for 4 h and then dried. Before electrophoresis, peptides were treated with

iodoacetamide to block cysteine residues (Lane 1978). Samples were dried, taken up in 20 µl of loading buffer (8 M urea, 0.125 M Tris-HCl (pH6.8), 12.5 mM DTT and 0.1 % bromophenol blue), and incubated at 37°C for 30 min. Each sample was split into two and iodoacetamide, in 50 mM Tris pH 7.0, was added to one of the aliquots, to a final concentration of 23 mM. Both aliquots were then incubated at 37°C for 30 min prior to electrophoresis. The products of these digestions were separated on gels prepared in a manner similar to that of Pantazis and Bonner (Pantazis and Bonner, 1981). Briefly, for products of Lys-C digestion, the gels contained 25 % acrylamide, 0.038 % bis-acrylamide and 0.75 M Tris-HCl (pH 8.8). For the separation of trypsin digestion products the acrylamide and bis-acrylamide concentrations were increased to 30 % and 0.045 % respectively. The stacking gel in each case was 3.3 % acrylamide, 0.16 % bis-acrylamide, 0.125 M Tris-HCl (pH 6.8) and 8 M urea. Electrophoresis was carried out at 100 V for 1 h at 0°C and then at 200 V at room temperature. The gels were dried without fixing and radioactive products visualised by autoradiography.

CHAPTER 3

RESULTS - PURIFICATION OF DNA LIGASE II

3.1 DNA Ligase II from Calf Thymus

3.1.1 Purification

Details of the DNA ligase II purification scheme are described in the Materials and Methods and are summarised below in Table 3.1.

Table 3.1. **Purification of DNA ligase II from Calf Thymus** - Purification of DNA ligase II from 500 g of tissue. Activities of fractions were assayed by joining of the oligo(dT)/poly(rA) substrate. Protein concentrations were determined by the method of Bradford (Bradford, 1976).

Fraction		Protein(mg)	oligo(dT)/poly(rA)	
			Total Activity (Units)	Specific Activity (Units/mg)
I.	Crude extract	19000	-	-
II.	Phosphocellulose	1600	-	-
III.	Ammonium sulphate	490	-	-
IV.	Hydroxylapatite	56	5	0.09
V.	Gel filtration	28	1.3[a]	0.05
VI.	FPLC Phenyl Superose	5	0.83	0.17
VII.	FPLC Mono Q	2	1.4[b]	0.72
VIII.	FPLC Mono S	0.3	0.41	1.4

[a] These fractions are assayed at a high salt concentration which partly inhibits the enzyme activity (see Chapter 4, section 4.5). The values have been adjusted accordingly.
[b] An apparent increase in total activity has been reproducibly observed at this stage.

Previous work had established that DNA ligase II is extracted at 0.2 M NaCl, binds to phosphocellulose and precipitates between 42-66 % ammonium sulphate (Soderhall and Lindahl, 1973b; A.E. Tomkinson, personal communication). The amount of activity present during the first three steps of the purification is difficult to assess due to the apparent lability of the enzyme and contaminating RNase H which degrades the substrate. The first three steps of the purification were routinely carried out at four-fold larger scale so that, following adsorption to phosphocellulose and ammonium sulphate precipitation, the extract

could be conveniently stored at -70°C with no loss of activity. A single ammonium sulphate pellet, equivalent to 500 g of tissue, was then processed through the purification scheme.

Hydroxylapatite provides a key step in the purification of DNA ligase II, allowing the removal of much of the contaminating DNA ligase I. DNA ligase II adsorbs very strongly to this matrix, unlike DNA ligase I, and requires high phosphate concentrations for its elution. The large purification gained from this column can be seen from the protein profile in Fig. 3.1. Fractions from the 400 mM potassium phosphate elution were pooled and concentrated before applying to a gel filtration column.

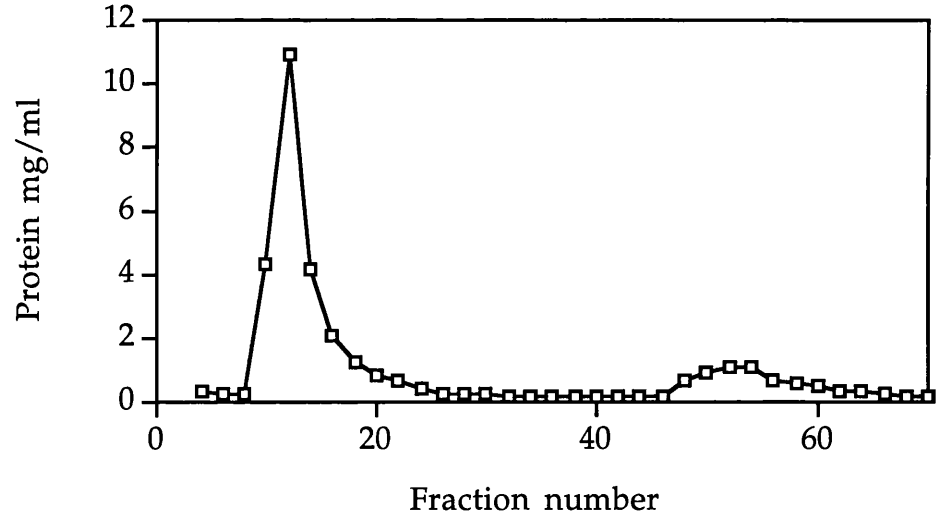


Fig. 3.1. **Fractionation of DNA Ligase II from Calf Thymus by Hydroxylapatite Chromatography** - DNA ligase II activity elutes at 400 mM potassium phosphate, fractions 40-60; contaminating DNA ligase I occurs mainly in the 250 mM potassium phosphate protein peak, fractions 2-30.

The AcA 44 gel filtration matrix has a separation range of 10,000-130,000 Da which allows removal of larger contaminating proteins, including any remaining full length DNA ligase I. Fig. 3.2 shows the size fractionation of the DNA ligase activities identified by adenylylation.

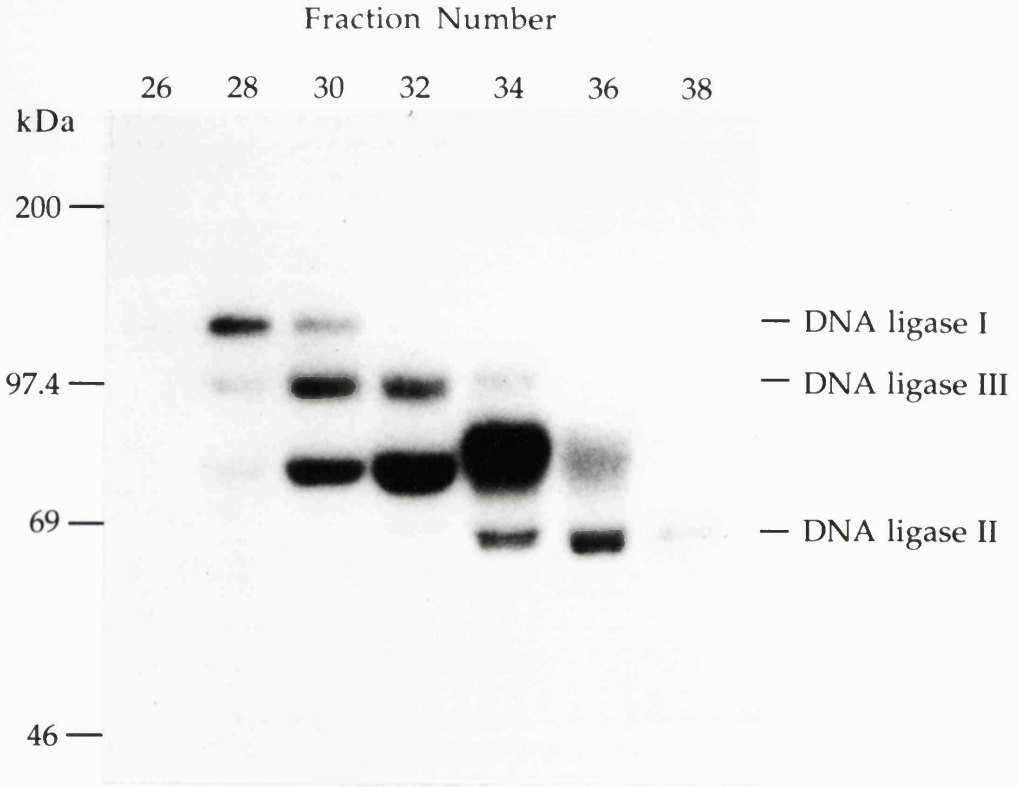


Fig. 3.2. **Adenylylation of DNA Ligases From Calf Thymus Following Separation by Gel Filtration on AcA 44** - Following hydroxylapatite chromatography, fractions were concentrated and run on gel filtration. DNA ligases were detected by formation of a radiolabelled enzyme-AMP complex, separated by 7.5 % SDS-PAGE, and visualised by autoradiography. The positions of DNA ligases I, II and III are indicated, the remaining bands are proteolytic products.

Although this step results in little purification, and a small decrease rather than increase in specific activity, it is useful to remove contaminating DNA ligases I and III.

At this point in the purification two major bands of approximately 72 and 74 kDa were seen by Coomassie Blue staining after SDS-PAGE. The DNA ligase II activity was found to be associated with the lower of the two bands by radiolabelling the enzyme-AMP complex and counting the radioactivity present in the Coomassie Blue stained protein bands. To separate these two proteins, hydrophobic chromatography was used in the form of a Phenyl Superose FPLC column. Mono Q chromatography removed any of the 85 kDa proteolytic product of DNA ligase I that still remained and the DNA ligase II was concentrated using a Mono S column. Final fractions are shown in Fig. 3.3, with corresponding DNA ligase activity assayed on the oligo(dT)/poly(rA) substrate.

A relatively low yield of DNA ligase II activity in the latter purification steps reflects the lability of this enzyme during purification. Stability was improved by the addition of 0.2% Tween 20 to all buffers employed after the hydroxylapatite chromatography step (see Chapter 4, section 4.4).

3.1.2 Apparently homogeneous DNA ligase II

The peak fractions of DNA ligase II activity (Table 3.1, Fraction VIII) eluted from the final FPLC Mono S column at a salt concentration of 220-250 mM. These apparently contained a single band of protein at approximately 70 kDa detected by Coomassie Blue staining following SDS-PAGE (Fig. 3.4A). This observation was confirmed by silver staining of the same gel which revealed very few contaminating proteins (Fig. 3.4B). After incubation of Fraction VIII with [α-^{32}P]ATP, a radioactively labelled 70 kDa polypeptide was detected by autoradiography following SDS-PAGE.

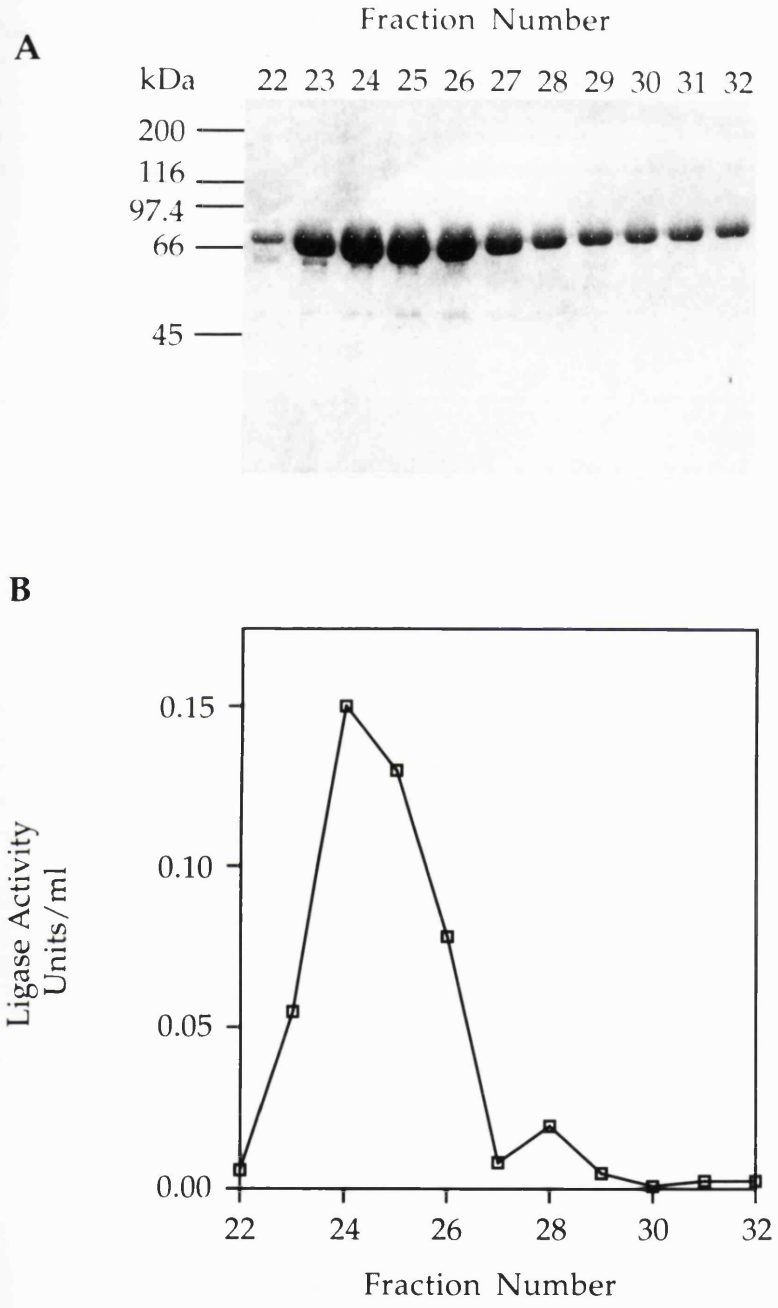


Fig. 3.3. **Fractionation of DNA Ligase II from Calf Thymus by Mono S Chromatography** - Fractions containing DNA ligase II activity following Mono Q chromatography were separated by gradient elution on a Mono S column. (A) Proteins were detected by staining with Coomassie Brilliant Blue. (B) Activity was assayed by joining of the oligo(dT)/poly(rA) substrate, showing a peak of activity in fraction 24.

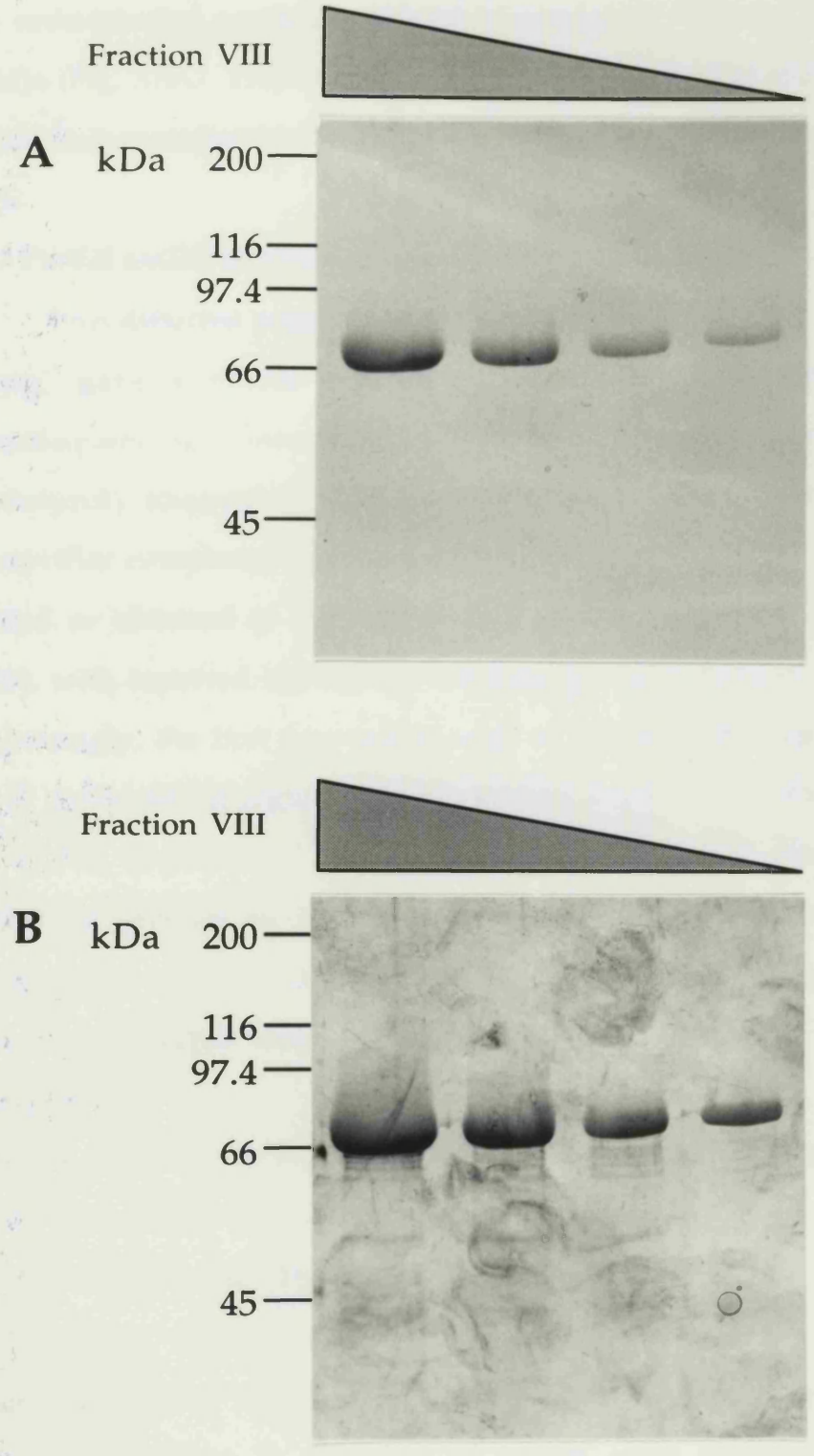


Fig. 3.4. **Apparently Homogeneous DNA Ligase II from Calf Thymus** - DNA ligase II protein (decreasing amounts, from 1 µg to 250 ng) was resolved through 7.5%SDS-PAGE. Proteins were detected by (A) Coomassie Brilliant Blue or (B) silver staining (Bio-Rad).

The radiolabelled protein appeared to comigrate with the protein band at 70 kDa (Fig. 3.5A). This was confirmed by slicing the gel and measuring the radioactive material in the gel slices by liquid scintillation counting (Fig. 3.5B).

3.1.3 Partial amino acid sequence of purified DNA ligase II

Two different preparations of DNA ligase II (Table 3.1, Fraction VIII, 20 μg) gave a single N-terminal sequence of 20 amino acids by microsequencing. Comparison of this sequence with a protein database (Swissprot) showed it was homologous to the N-terminus of the microvillar cytoplasmic protein, ezrin (Gould et al., 1989). Ezrin is closely related or identical to the mammalian protein, cytovillin (Turunen et al., 1989), with reported differences confined to the N-termini of the proteins. Surprisingly, the first few amino acids of our sequence appeared to be in the 5' untranslated region of the cytovillin sequence. Further inspection of the cDNA sequence of cytovillin 5' to the predicted site of translation initiation reveals an in-frame amino acid sequence starting with a methionine, which is identical to the N-terminus of ezrin. This suggests that the tentative translational start site of cytovillin was originally misassigned.

Internal peptides from a DNA ligase II preparation (30 μg, Fraction VIII) were generated by digestion with the reagent proteolytic enzyme, Lysylendopeptidase. These were separated by HPLC and the N-terminal amino acid sequences of 14 different peptides determined by microsequencing. Analysis of these peptide sequences showed that ten were homologous to stretches of amino acids distributed throughout ezrin/cytovillin. The remaining four peptides, which were present in similar molar amounts to the other peptides, were not detectably homologous to ezrin/cytovillin or to any other mammalian protein

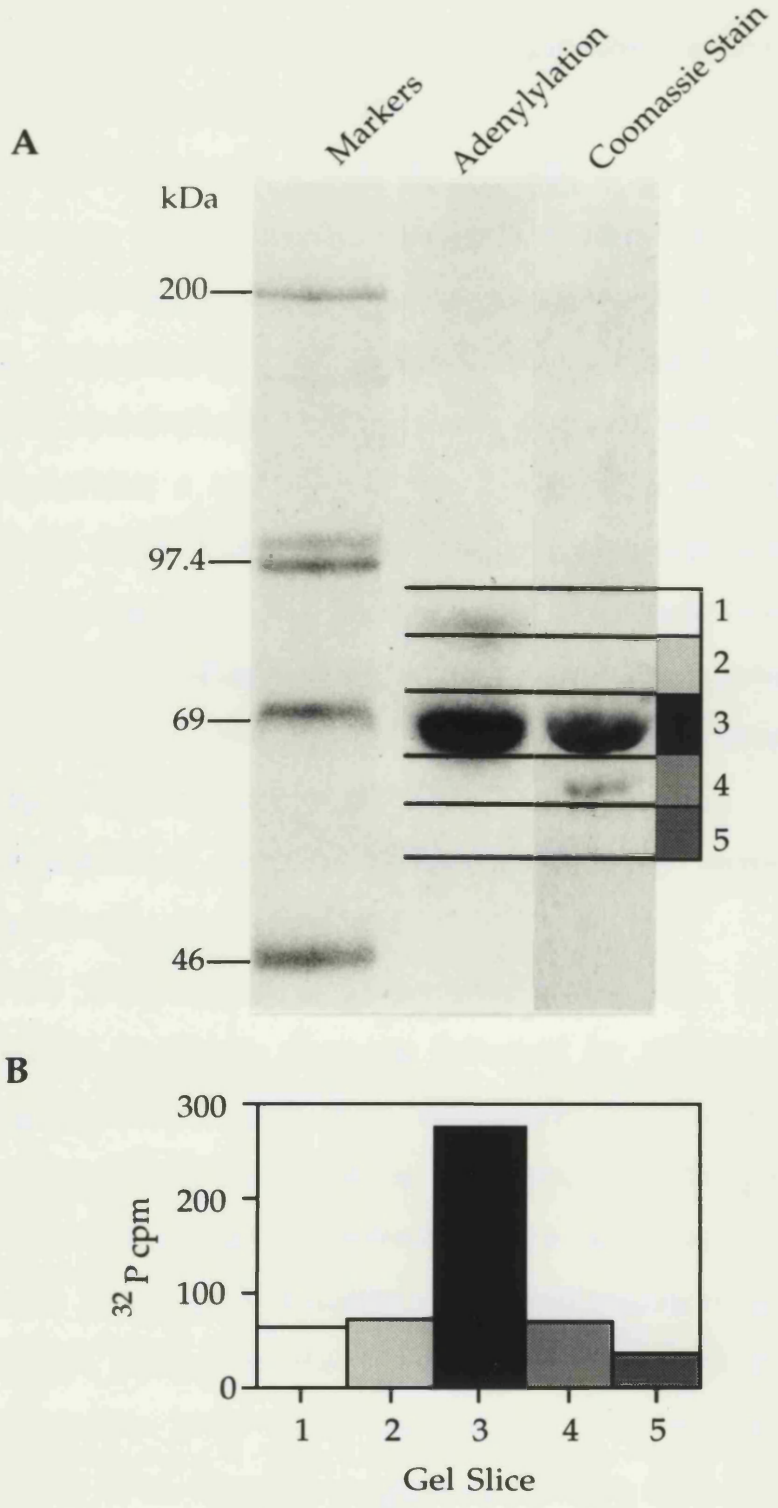


Fig. 3.5. **Comigration of Protein Band and Adenylylated Band of Calf Thymus DNA Ligase II** - (A) DNA ligase II (Table 3.1, Fraction VIII) was separated by 7.5 % SDS-PAGE and stained with Coomassie Brilliant Blue. A sample was also adenylylated and loaded on the same gel.
(B) Radioactivity was quantitated by scintillation counting of gel slices taken from the adenylylated track; the peak of radioactivity comigrated with the protein band at approximately 70 kDa.

sequence in the database, including the predicted amino acid sequence of DNA ligase I (Barnes et al., 1990).

The results indicate that the material in the 70 kDa band (Fig. 3.4) represents an approximately equimolar mixture of two proteins, the 69 kDa (as predicted from the cDNA sequence) ezrin/cytovillin and a second protein, presumably DNA ligase II. Ezrin/cytovillin binds unusually tightly to hydroxylapatite, a property it shares with DNA ligase II and which provides a key step in the purification of both these proteins (Bretscher, 1983). Consequently the complete separation of these two proteins becomes difficult. The presence of a single N-terminal sequence in Fraction VIII, corresponding to ezrin/cytovillin, strongly indicates that the N-terminus of DNA ligase II is blocked, as was found previously for DNA ligase I (Tomkinson et al., 1990).

3.1.4 Analysis of purified DNA ligase II by two-dimensional gel electrophoresis

The DNA ligase II preparation (Table 3.1, Fraction VIII) was resolved by isoelectric focusing followed by SDS-PAGE. Two 70 kDa polypeptides with different isoelectric points were detected in approximately equal amounts, as predicted from the peptide sequencing experiments (Fig. 3.6A). Antibodies were raised against the N-terminal sequence of ezrin/cytovillin and immunoblotting identified the polypeptide closest to the bovine serum albumin marker as bovine ezrin/cytovillin (Fig. 3.6B). The second polypeptide was tentatively assigned as DNA ligase II, suggesting that the final preparation of DNA ligase II was about 50% pure, the residual contaminating protein being ezrin/cytovillin, an abundant structural cellular protein.

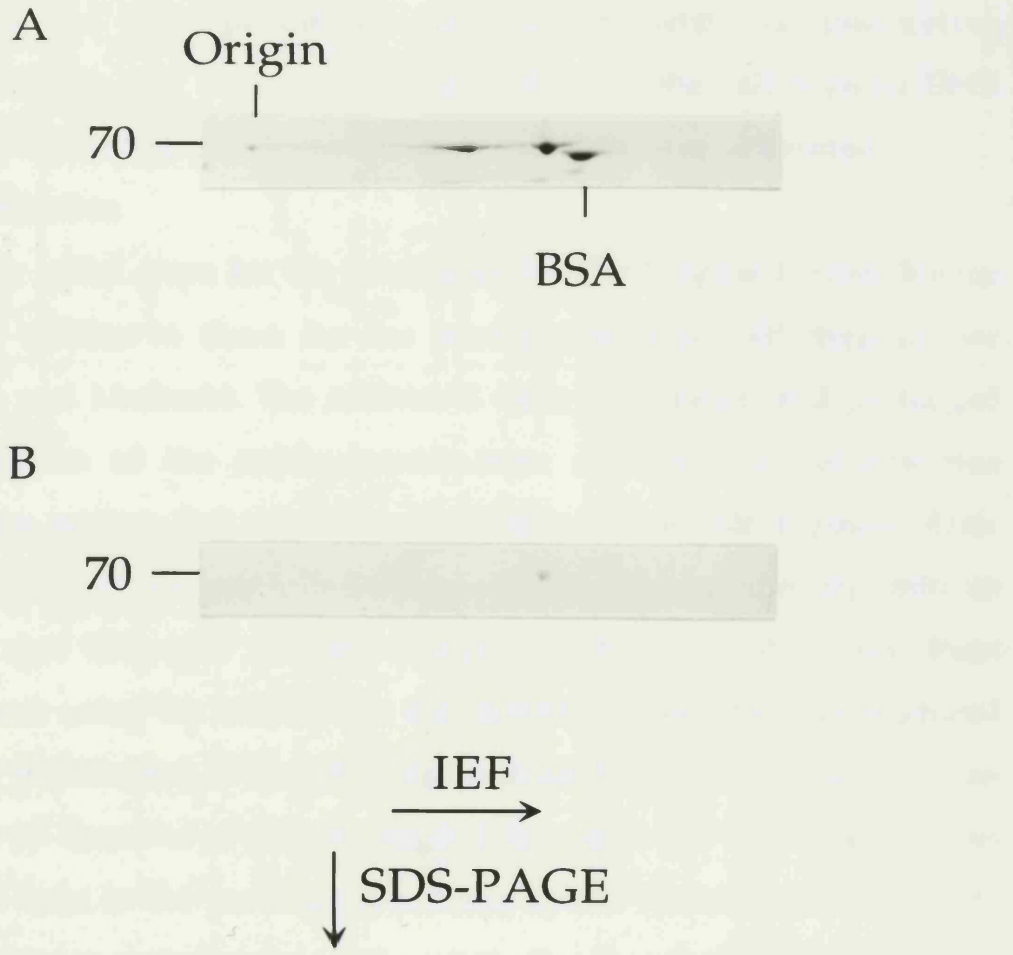


Fig. 3.6. **Two-Dimensional Analysis of DNA Ligase II from Calf Thymus** - DNA ligase II (Table 3.1, Fraction VIII) was electrophoresed through a 5% polyacrylamide isoelectric focussing rod gel and then separated in the second dimension by electrophoresis through 10% SDS-PAGE. Supplementing the fraction with 300 ng of BSA provided an internal marker. (A) Proteins stained with Coomassie Brilliant Blue, the BSA marker was identified by analysis alone on a separate gel.
(B) Immunological detection of ezrin/cytovillin after transfer of the proteins onto nitrocellulose. An antibody against the N-terminal sequence of ezrin/cytovillin was used and immunocomplexes detected by a secondary antibody coupled to alkaline phosphatase. The portion of the gel corresponding to 70 kDa by SDS-PAGE is shown.

3.2 DNA Ligase II from Bovine Liver

Due to the problems encountered with contaminating ezrin/cytovillin and losses of enzyme activity in the calf thymus DNA ligase II preparation, purification from bovine liver was attempted.

3.2.1 Purification

The initial steps for the purification of DNA ligase II from bovine liver are similar to those for the purification from calf thymus (see Materials and Methods). The additional steps of filtration and prolonged centrifugation of the crude extract were necessary to remove fine particulate matter that was not encountered with calf thymus. After ammonium sulphate precipitation, protein was loaded directly onto an AcA 34 gel filtration column (range 20,000 to 350,000 Da). Pilot experiments using the calf thymus enzyme had shown that this produced a greater purification of the DNA ligase II away from DNA ligase I than did AcA 44. Removal of the full length DNA ligase I at this stage reduces problems later in the purification scheme due to accumulation of the 85 kDa catalytic domain by proteolytic action, that then is difficult to separate from DNA ligase II. Mono Q chromatography could therefore be omitted. A decrease in the amount of total protein produced by using gel filtration as an initial step was also of advantage as a smaller hydroxylapatite column could then be used and still give extensive purification (Fig. 3.7). Phenyl Superose and Mono S FPLC further increased the purification and allowed concentration of the final fractions (Fig. 3.8).

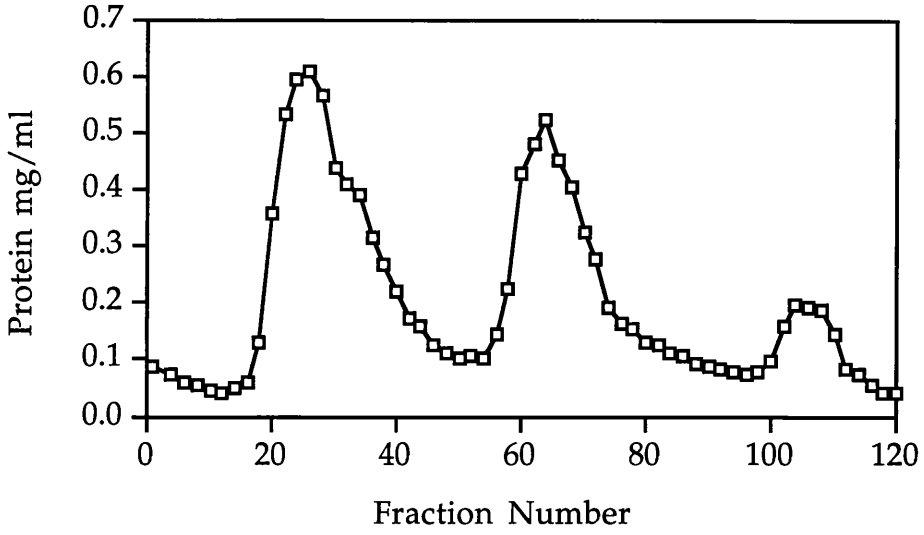


Fig. 3.7. **Fractionation of DNA Ligase II from bovine liver by Hydroxylapatite Chromatography** - Fractions 10-45, 100 mM potassium phosphate; fractions 46-85, 250 mM potassium phosphate; fractions 86-120, 500 mM potassium phosphate, containing the DNA ligase II activity.

3.2.2 Partially purified DNA ligase II

The most active fraction of this preparation was adenylylated and resolved by SDS-PAGE. The gel was silver stained and is shown in Fig. 3.9A. Individual bands were cut from the gel as indicated and the radioactivity in each band determined by scintillation counting (Fig. 3.9B). The radioactivity was detected in a single gel slice, corresponding to a discrete protein band of approximately 70 kDa. The DNA ligase II from bovine liver is only partially purified but is free from other contaminating DNA ligase activities as no other adenylylated bands were detected. Immunoblotting has established that there is no contaminating ezrin/cytovillin in DNA ligase II prepared in this way (R.A. Nash, personal communication).

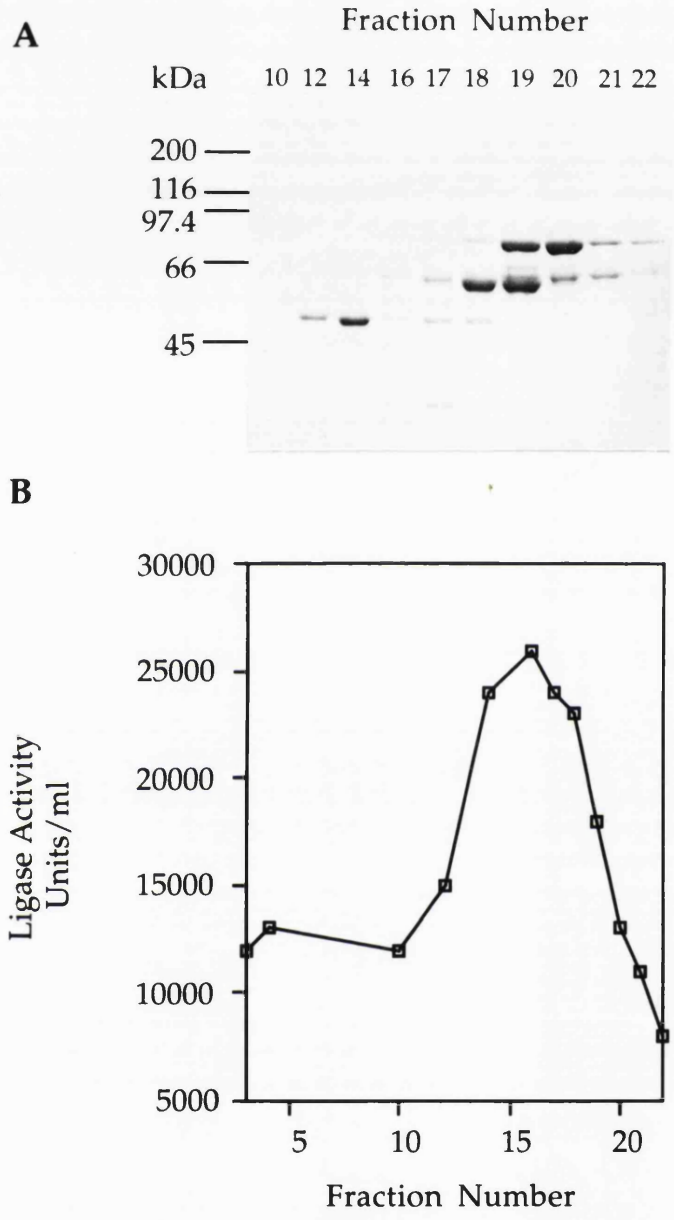


Fig. 3.8. **Fractionation of DNA Ligase II from Bovine Liver by Mono S Chromatography** - Fractions containing DNA ligase II activity were pooled following Phenyl Superose chromatography and separated by gradient elution on a Mono S column. (A) Proteins were detected by staining with Coomassie Brilliant Blue. (B) Activity was assayed by joning of the oligo(dT)/poly(rA) substrate.

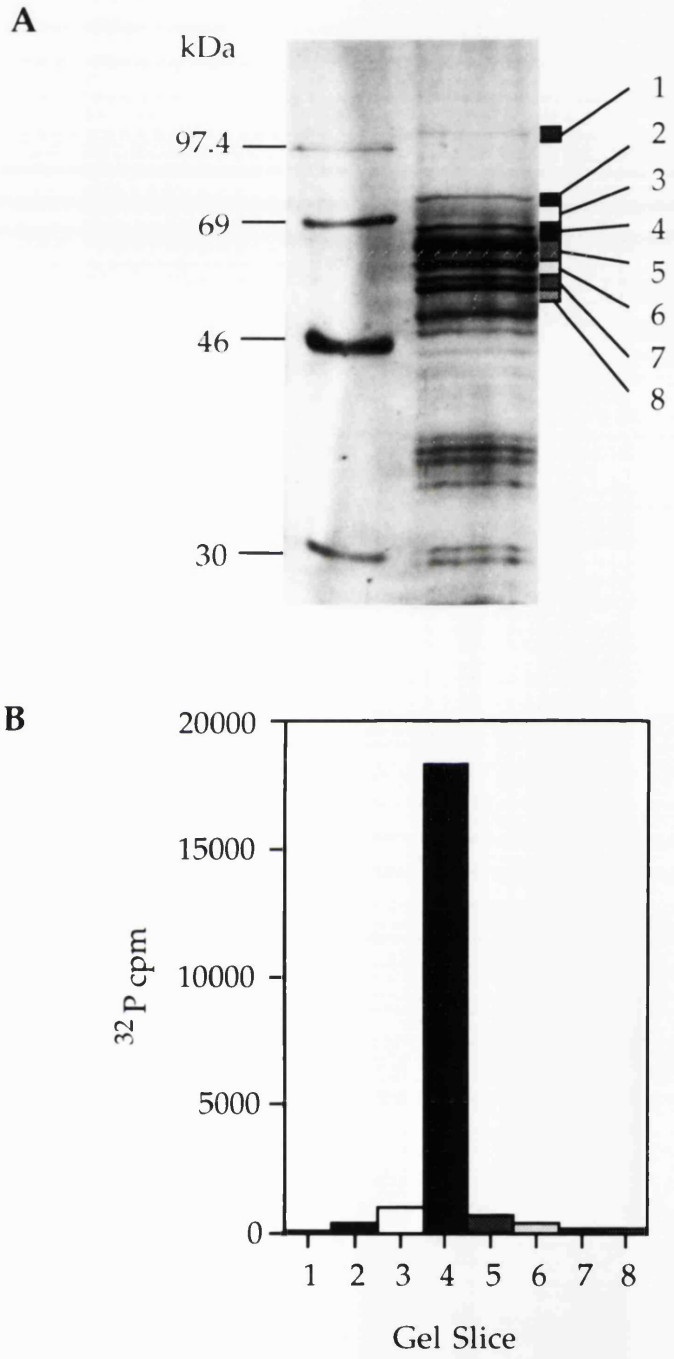


Fig. 3.9. **Comigration of Protein Band and Adenylylated Band of DNA Ligase II from Bovine Liver** - (A) Adenylylated DNA ligase II was separated by 7.5 % SDS-PAGE and the gel silver stained (Bio-Rad). (B) Radioactivity was assessed by scintillation counting of gel slices.

CHAPTER 4

RESULTS - CHARACTERISATION OF DNA LIGASE II

All the following characterisation experiments used DNA ligase II from calf thymus.

4.1 Reactivity of the Adenylylated Forms of DNA Ligases I and II

Incubation of DNA ligases from mammalian cells with [α-^{32}P]ATP allows formation of a radiolabelled enzyme-AMP complex which can be visualised by autoradiography following SDS-PAGE. The DNA ligases can be distinguished by the different sizes of their radiolabelled intermediates (Fig. 4.1, lanes 1 and 5). Bovine DNA ligase I-AMP migrates during SDS-PAGE with an apparent molecular mass of 125 kDa (Tomkinson et al., 1990), while DNA ligase II has an apparent molecular mass of approximately 70 kDa. In the second step of the ligation reaction, the AMP group is transferred from the enzyme to the 5'-phosphate of the DNA strand break. This reaction was mimicked *in vitro* by incubating radiolabelled enzyme-AMP with a molar excess of unlabelled homopolymer substrate, either oligo(dT)/poly(dA) or oligo(dT)/poly(rA) and loss of radioactivity from the enzyme-AMP was monitored. Addition of pyrophosphate reverses the initial step of the reaction, also resulting in loss of radiolabel from the enzyme. The DNA ligase I AMP-intermediate transferred the radiolabelled AMP group to the oligo(dT)/poly(dA) substrate but did not react with the oligo(dT)/poly(rA) substrate (Fig. 4.1, lanes 6 and 7). Unlike DNA ligase I, DNA ligase II was able to transfer the AMP group to both substrates (Fig. 4.1, lanes 2 and 3). In both cases the adenylylation was reversed by incubation with pyrophosphate (Fig. 4.1, lanes 4 and 8). This agrees with previous observations on the substrate specificities of the two DNA ligases (Arrand et al., 1986).

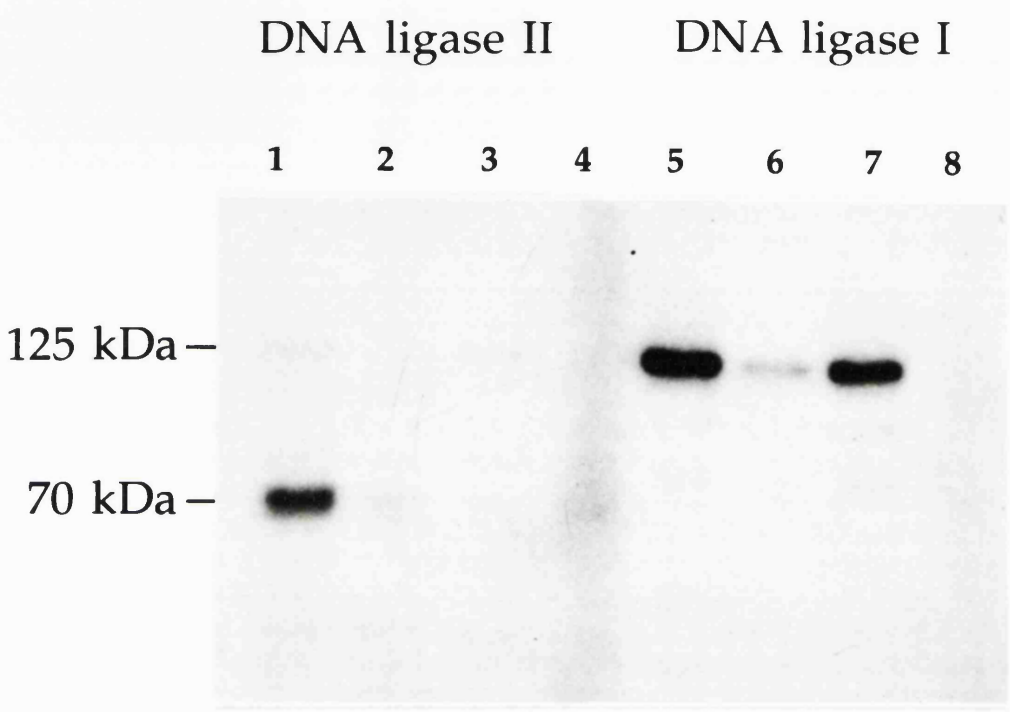


Fig. 4.1. **Reactivity of Enzyme-AMP Complexes of DNA Ligases I and II** - DNA ligase-AMP complexes were formed using [α-^{32}P]ATP and then incubated with unlabelled substrate or sodium pyrophosphate as detailed in Materials and Methods. Lanes 1-4, DNA ligase II; lanes 5-8, DNA ligase I; with: lanes 1 and 5, no addition; lanes 2 and 6, oligo(dT)/poly(dA); lanes 3 and 7, oligo(dT)/poly(rA); and lanes 4 and 8, sodium pyrophosphate. Reactions were stopped by the addition of SDS sample buffer, and analysed by 7.5 % SDS-PAGE and autoradiography.

4.2 DNA Ligation Using the ATP Analogue, ATP-αS, as a Cofactor

DNA ligases I, II and III were analysed for their ability to use ATP-αS as a substitute for ATP. Using the DNA ligase assay detailed in Materials and Methods (section 2.3.2), ATP was substituted with ATP-αS (200 μM) and reactions incubated for 30 min. The results using the oligo(dT)/poly(dA) substrate are shown in Fig. 4.2. DNA ligase I can utilise this analogue to the same extent as ATP, whereas DNA ligases II and III can only use it to 67 % and 20 % respectively compared to ATP. This is in contrast to the results of Montecucco et al. (Montecucco et al., 1990) and Elder and Rossignol (Elder and Rossignol, 1990) who observed that DNA ligase II is better able than DNA ligase I to use ATP-αS with the same efficiency as ATP.

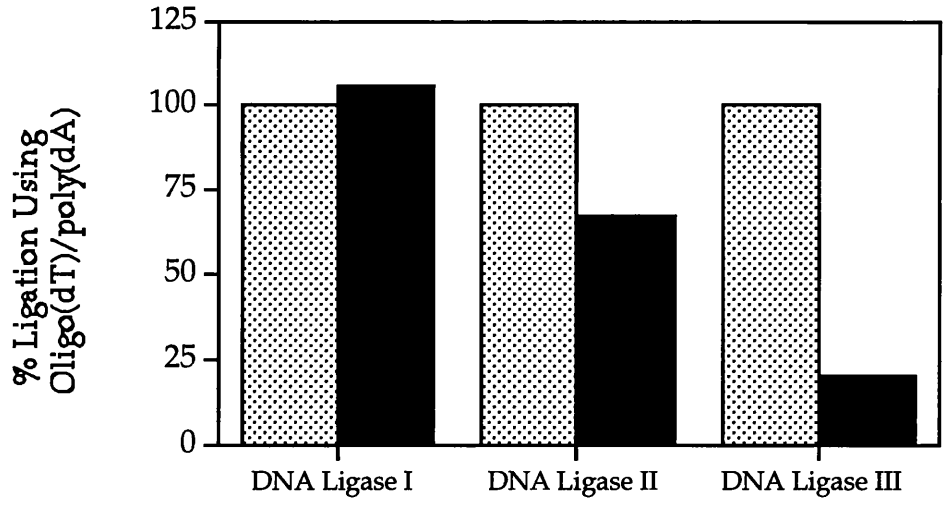


Fig. 4.2. **Adenylylation Using ATP-αS** - DNA ligases I, II and III were assayed using the oligo(dT)/poly(dA) substrate with either ATP (▒) or ATP-αS (■) as a cofactor. Results with ATP-αS are given as a percentage of the activity observed using ATP.

4.3 Polynucleotide Substrate Specificity

The abilities of bovine DNA ligases I, II and III to catalyse ligation of several different homopolymer polynucleotide substrates were investigated. Reaction products were separated by denaturing

polyacrylamide gel electrophoresis to determine whether authentic ligation products were being generated. In agreement with previous results, all three DNA ligases could join oligo(dT) molecules hybridised to a poly(dA) template (Fig. 4.3, lanes 1-4), whereas only DNA ligases II and III could ligate oligo(dT) molecules hybridised to a poly(rA) template (Fig. 4.3, lanes 5 and 8). A novel substrate of 5'-^{32}P labelled oligo(rA) hybridised to poly(dT) was constructed and used in these assays. Surprisingly, DNA ligase I and DNA ligase III, but not DNA ligase II, catalysed ligation of this oligo(rA)/poly(dT) substrate (Fig. 4.3, lanes 9-12). This means that the three mammalian DNA ligases can be distinguished by their polynucleotide substrate specificity (see Table 4.1).

Table 4.1. **Substrate Specificities of the Three Mammalian DNA Ligases**

	I	II	III
Ligation of oligo(dT)/poly(dA)	+	+	+
Ligation of oligo(dT)/poly(rA)	-	+	+
Ligation of oligo(rA)/poly(dT)	+	-	+

4.4 Stabilisation of DNA Ligase II Activity by Non-Ionic Detergent

Fractions from hydroxylapatite chromatography containing DNA ligase II activity (Table 3.1, Fraction IV) were incubated for five days at 0°C in different buffers, to assess the possibility of reducing losses of activity observed during purification and subsequent storage. A control sample remained in the buffer from the column, namely 400 mM potassium phosphate, pH 7.5. Further samples were either dialysed into alternative buffers or remained in the column buffer, to which additions were made. After five days the activity of the samples was compared to the control sample by means of the DNA ligation assay.

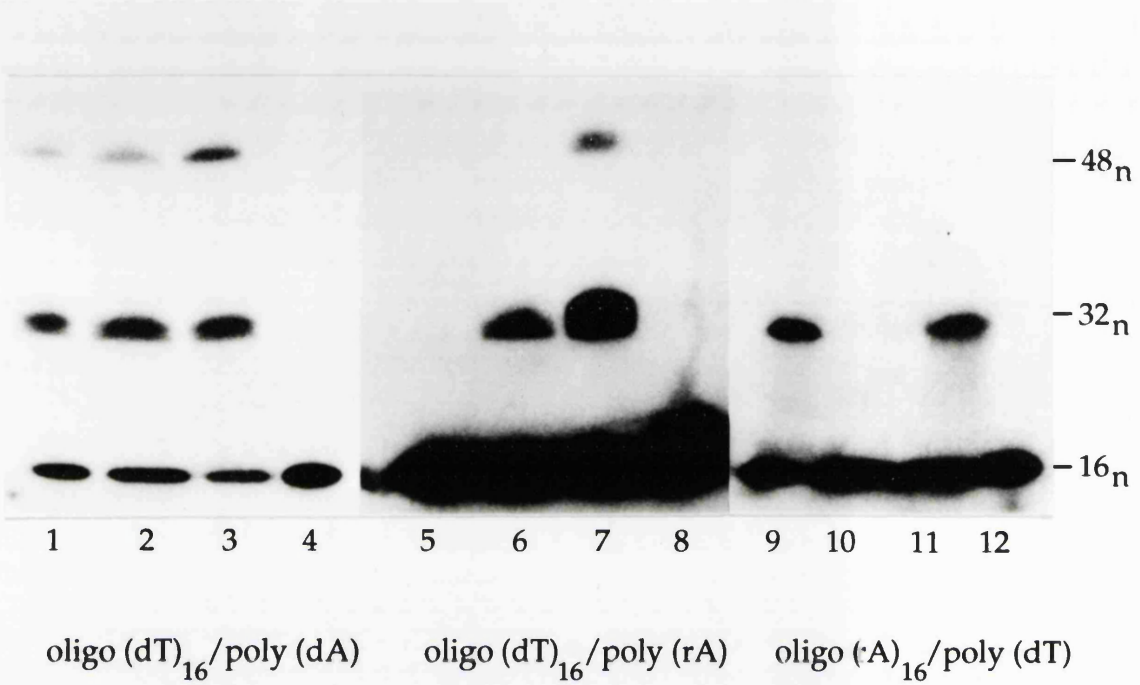


Fig. 4.3. **Analysis of the Ligation Products Formed by DNA Lgases I, II and III** - Equivalent amounts of the three DNA ligases, as measured by activity on the oligo(dT)/poly(dA) substrate, were incubated with dfferent labelled polynucleotide substrates as indicated. Lanes 1, 5 and 9, DNA ligase I; lanes 2, 6 and 10, DNA ligase II; lanes 3, 7 and 11, DNA ligase III; lanes 4, 8 and 12, no enzyme. Ligation products were resolved through a 20 % denaturing polyacrylamide gel and detected by autoradiography.

The results are shown in Fig. 4.4. The only stabilisation of DNA ligase II activity was observed with the addition of the non-ionic detergent, Tween 20 (Fig. 4.4, buffer 3) and in subsequent purifications of DNA ligase II from calf thymus, 0.2 % Tween 20 was added from Fraction IV onwards (see Table 3.1). The other buffers adversely affected enzyme stability. Buffers 5 and 6 (see Fig. 4.4) contained a high concentration of sodium chloride that may have caused some inhibition of activity in the assay (see section 4.5 below). Even when results were adjusted to take into account the % inhibition of activity known to occur at these salt concentrations in buffers 5 and 6, no stabilisation was observed.

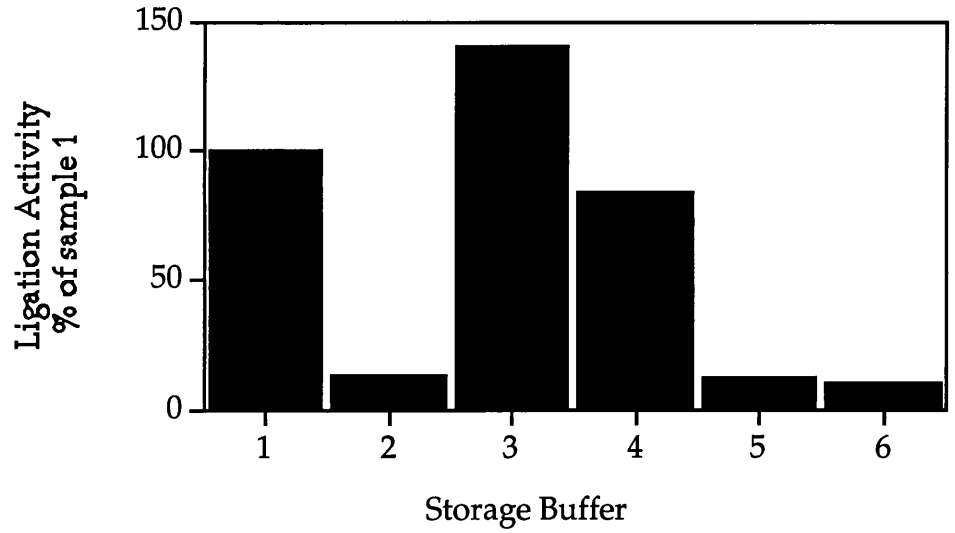


Fig. 4.4. **Stabilisation of DNA ligase II Activity** - Samples were stored for 5 days at 0°C and then assayed using the oligo(dT)/poly(rA) substrate. Storage buffers were: (1) hydroxylapatite column buffer (400 mM potassium phosphate); (2) 50 mM potassium phosphate, 50 % ethylene glycol, 0.5 mM DTT; (3) hydroxylapatite column buffer + 0.2 % Tween 20; (4) hydroxylapatite column buffer + 5 mM DTT; (5) 1 M NaCl, 50 mM Tris-HCl pH 7.5, 0.5 mM DTT; (6) 1 M NaCl, 50 mM Tris-HCl pH 7.5, 0.5 mM DTT, 1 mM EDTA.

4.5 Inhibition of DNA Ligase II Activity by Sodium Chloride

An important consideration in assaying fractions directly from chromatography or gel filtration columns during enzyme purification is

the possibility of inhibition of activity due to high salt concentration present in the column buffer. The sensitivity of DNA ligases I and II to increasing salt concentrations was compared using two partially purified fractions obtained during the DNA ligase II purification from calf thymus, that contained either DNA ligase I or DNA ligase II as assessed by adenylylation. The DNA ligases were assayed using both the oligo(dT)/poly(dA) and the oligo(dT)/poly(rA) substrates. Final sodium chloride concentrations were adjusted to 50, 100 and 200 mM NaCl in the reaction mixture. DNA ligase II appears to be far more sensitive than DNA ligase I with 50 % inhibition of activity occurring at a salt concentration of 30 mM, compared to 130 mM for DNA ligase I (Fig. 4.5).

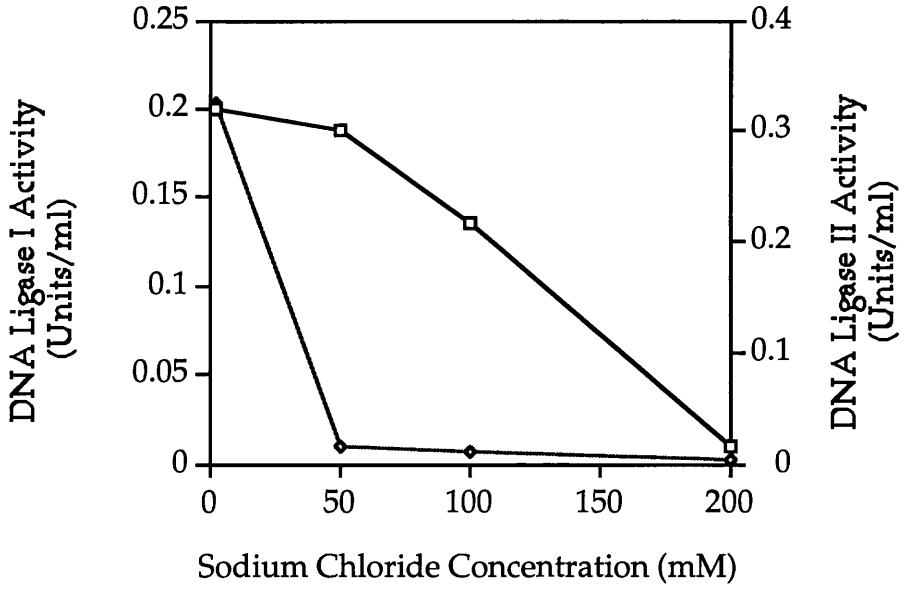


Fig. 4.5. **Inhibition of DNA Ligases I and II by Increasing NaCl Concentration** - DNA ligases I (—□—) and II (—◇—) were assayed using the oligo(dT)/poly(dA) substrate in reaction mixture containing sodium chloride as indicated. Assays and units of activity were as described in (section 2.3.2., Materials and Methods).

4.6 Neutralising Antiserum against DNA Ligase I does not Inhibit DNA Ligases II and III

A polyclonal antiserum against DNA ligase I had been shown to inhibit the ligation activity of this enzyme to >95 % (Soderhall and Lindahl, 1975). This was tested against the calf thymus DNA ligases II and III. Differing concentrations of antiserum were preincubated with the DNA ligase for 5 min at 0°C, DNA joining activity was assessed by analysis of the ligation products using the oligo(dT)/poly(dA) and oligo(dT)/poly(rA) substrates (section 2.3.3.), and % inhibition of the reaction calculated by scanning densitometry of the autoradiograph. Under conditions that gave 84 % inhibition of DNA ligase I, there was no inhibition of DNA ligases II and III. No effect of the DNA ligase I antiserum was seen with either DNA ligase II or III, even at higher concentrations of antiserum. The activities of both these enzymes incubated with the antiserum remained identical to that seen with similar concentrations of preimmune serum (Fig. 4.6).

4.7 Immunoblots

4.7.1 Polyclonal antiserum against purified DNA ligase I

A polyclonal antiserum raised against homogeneous bovine DNA ligase I (Lasko et al., 1990) was used in immunoblotting experiments with DNA ligases II and III. This antiserum apparently does not detect bovine DNA ligase II on immunoblotting (Fig. 4.7A, lane 3). Neither does this antiserum cross-react with the bovine DNA ligase III (Fig. 4.7A, lane 5), indicating that both enzymes are distinct from DNA ligase I.

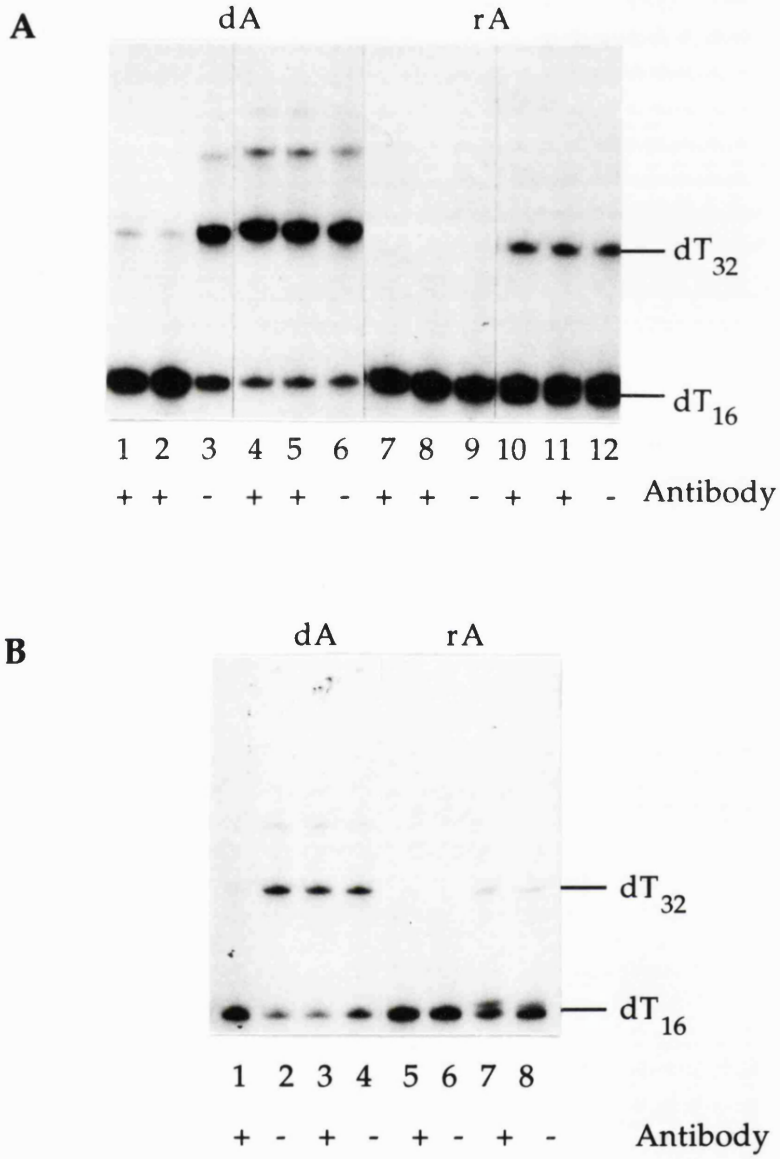


Fig. 4.6. **Inhibition of DNA Ligase Activity by Antibodies Against DNA Ligase I** - Equivalent amounts of DNA ligase I or II were used, as assayed on the oligo(dT)/poly(dA) substrate. DNA ligases were preincubated with either the immune (+) or preimmune serum (-) and the activity assayed with the labelled oligo(dT)/poly(dA) and oligo(dT)/poly(rA) substrates as indicated. Total protein in the antiserum was assessed by the method of Bradford (Bradford, 1976) and concentrations are given as amount of antiserum per ml of reaction mixture. (A) Lanes 1-3 and 7-9, DNA ligase I; lanes 4-6 and 10-12, DNA ligase II. Lanes 1, 4, 7, and 10, 0.6 mg/ml serum; lanes 2, 5, 8, and 11, 2mg/ml serum; lanes 3, 6, 9 and 12, 4 mg/ml serum. (B) Lanes 1, 2, 5, and 6, DNA ligase I; lanes 3, 4, 7, and 8, DNA ligase III; incubated with 1mg/ml serum.

4.7.2 Polyclonal antiserum against a conserved peptide

A highly conserved peptide sequence is located close to the C-termini of mammalian DNA ligase I and the DNA ligases of vaccinia virus, *S. cerevisiae* and *S. pombe* (Barker et al., 1985; Barker et al., 1987; Kerr and Smith, 1989; Barnes et al., 1990). A polyclonal antipeptide antiserum has been raised against a synthetic peptide of this sequence (Lasko et al., 1990) and was used in immunoblotting experiments with partially purified DNA ligases II and III. Neither of these two enzymes cross-reacts detectably with this antiserum (Fig. 4.7B, lanes 3 and 5), indicating that DNA ligases II and III apparently did not contain this epitope.

4.7.3 Cross-reactivity of vaccinia virus DNA ligase with antiserum against a conserved peptide

Crude cell extracts of bacteria containing a plasmid expressing the vaccinia virus DNA ligase cDNA or a vector-only control were kindly provided by S. Kerr and G. Smith (Sir William Dunn School of Pathology, University of Oxford). These extracts were tested with various antisera to DNA ligase I, the conserved peptide (see section 4.7.2 above) or DNA ligase II (see section 4.7.4). A positive reaction with the over-expressing extract and not the vector control was only seen with the antiserum against the conserved peptide (Fig. 4.7C).

4.7.4 Polyclonal antiserum against apparently homogeneous DNA ligase II

DNA ligase II was purified to apparent homogeneity from calf thymus and the final fractions analysed by two-dimensional gel electrophoresis (see section 3.1.4). This showed that the DNA ligase II preparation was not homogeneous but contained approximately 50 % of a second protein, ezrin/cytovillin. As a proportion of this preparation (Table 3.1, Fraction VIII) was assumed to be DNA ligase II, antibodies were raised

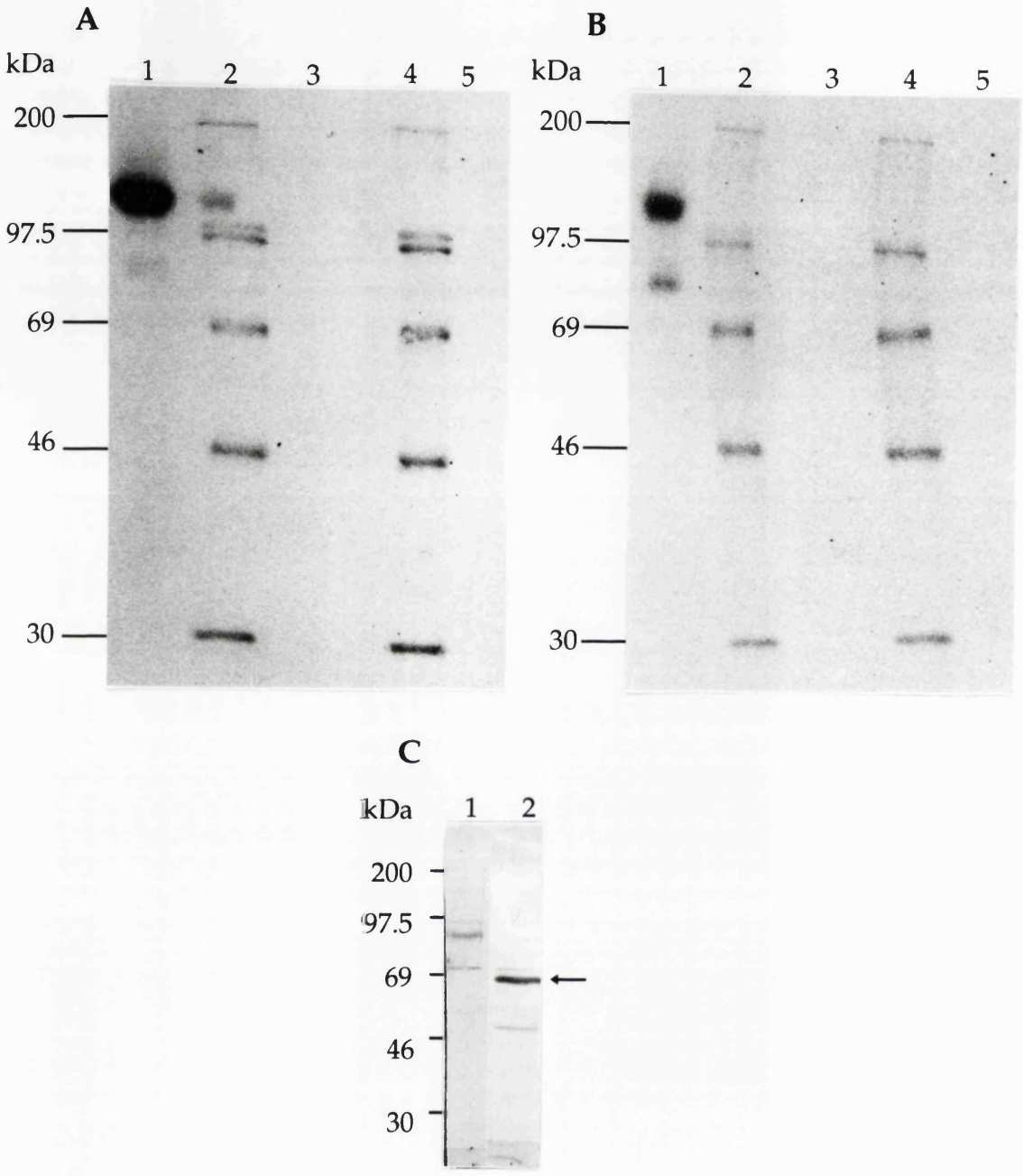


Fig. 4.7. **Immunoblots of DNA Ligases I, II and III and Vaccinia Virus DNA Ligase** - Proteins were separated by 10 % SDS-PAGE and transferred to nitrocellulose. Membranes were incubated with the primary antibody for 16 h. (A) Incubation with rabbit polyclonal antiserum against purified bovine DNA ligase I, 1:200 dilution. (B) Incubation with rabbit polyclonal antiserum against a conserved peptide sequence (Lasko et al., 1990)), 1:100 dilution. Immunocomplexes were detected with a ^{125}I-linked secondary antibody. Lane 1, DNA ligase I; lanes 2 and 4, ^{14}C-methylated molecular weight markers (Amersham); lane 3, DNA ligase II; lane 5, DNA ligase III. (C) Incubation with the antiserum against a conserved peptide, 1:50 dilution; lane 1, control extract; lane 2, extract from bacteria expressing vaccinia virus DNA ligase. Immunocomplexes were detected using a secondary antibody coupled to alkaline phosphatase. The vaccinia virus DNA ligase-specific band is marked with an arrow at the expected molecular mass of 63 kDa (Kerr and Smith, 1989).

against this partially purified fraction and used in immunblotting experiments against DNA ligases I and III. Although these antibodies reacted with the DNA ligase II preparation (Fig. 4.8A, lane 3), no cross-reactivity was observed with either DNA ligase I or III (Fig. 4.8A, lanes 1 and 5).

4.7.5 Polyclonal antiserum against a 'DNA ligase II' peptide

Peptide sequences were obtained from the apparently homogeneous preparation of DNA ligase II from calf thymus. A number of these sequences corresponded to the published sequence of ezrin/cytovillin, while others showed no homology. A polyclonal antiserum was raised against a synthetic peptide (-CTPHVAEPAENEQD-) corresponding to one of the unknown sequences. This reacted with the same 70 kDa protein band identified with the polyclonal serum against the protein mixture (Fig. 4.8B). As this peptide was known not to be from ezrin/cytovillin, it was assumed to come from DNA ligase II. To attempt to prove more conclusively that this might be the case, the antipeptide antiserum was assessed in its ability to inhibit DNA ligase II activity, using both the DNA ligation and adenylylation assays. These experiments were unsuccessful. Furthermore, the antiserum failed to immunoprecipitate the adenylylated form of DNA ligase II. Immunoprecipitation of the enzyme followed by adenylylation was also unsuccessful in establishing any interaction of DNA ligase II with the antiserum. Therefore, it remains uncertain whether this peptide was from DNA ligase II.

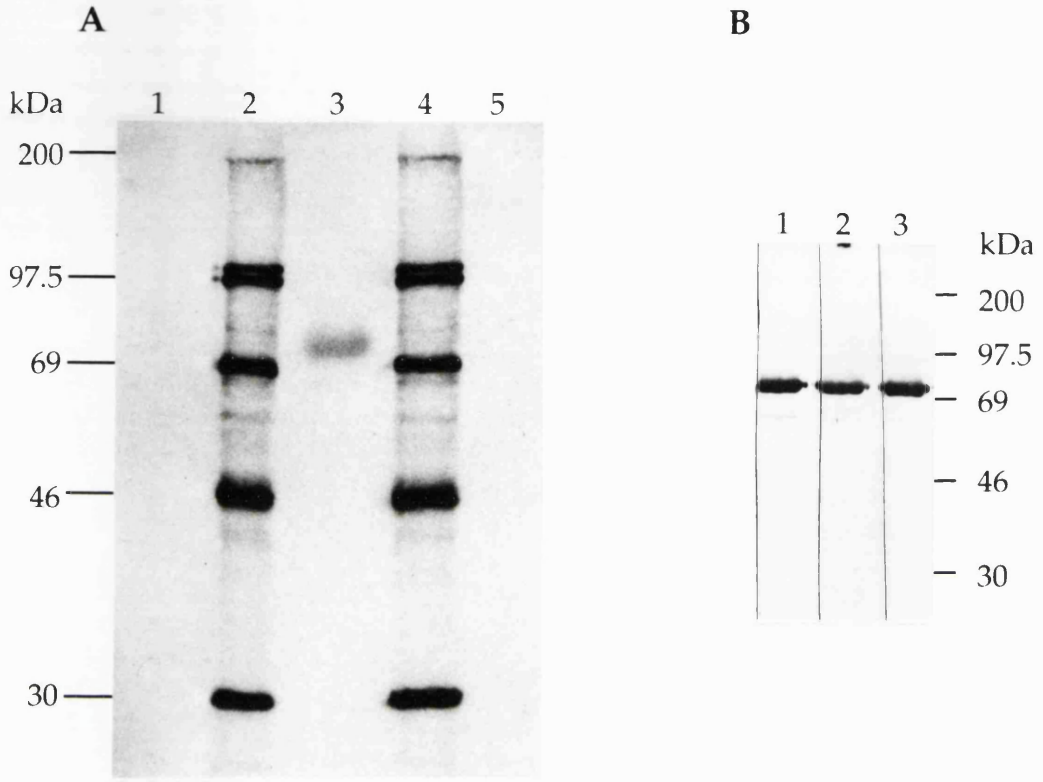


Fig. 4.8. **Immunoblots of DNA Ligases I, II and III** - Proteins were electrophoresed through 10 % SDS-PAGE, transferred to nitrocellulose, and incubated with the primary antibody for 16 h. (A) Incubation with rabbit polyclonal antiserum against bovine DNA ligase II (Table 3.1, Fraction VIII), 1:20 dilution. Lane 1, DNA ligase I; lanes 2 and 4, ^{14}C-methylated molecular weight markers (Amersham); lane 3, DNA ligase II; lane 5, DNA ligase III. Immunocomplexes were detected with a ^{125}I-linked secondary antibody. (B) Bovine DNA ligase II (Table 3.1, Fraction VIII), detected with; lane 1, rabbit polyclonal antiserum against Fraction VIII, 1:5 dilution; lanes 2 and 3, rabbit polyclonal antiserum against a possible DNA ligase II peptide, 1:5 and 1:10 dilutions. Immunocomplexes were detected using a secondary antibody coupled to alkaline phosphatase.

CHAPTER 5

RESULTS - COMPARISON OF DNA LIGASES I AND II BY PROTEOLYTIC DIGESTS

5.1 Comparison of Adenylylated DNA Ligases I and II by Partial Proteolytic Digests

Digestion of the adenylylated forms of DNA ligases I and II by trypsin was carried out using a range of protease concentrations. The individual arrays of radioactive peptides characteristic of each protein were compared by SDS-PAGE and autoradiography (Cleveland et al., 1977). DNA ligase I is a 102 kDa protein that migrates anomalously slowly as an 125 kDa polypeptide during SDS-PAGE on 7.5% gels (Tomkinson et al., 1990). This discrepancy is due to the presence of phosphoserine residues and a relatively high proportion of charged amino acids and prolines (Lindahl and Barnes, 1992). Trypsin digestion yielded radioactively labelled major degradation products of DNA ligase I of approximately 85, 80, 68, 42 and 30 kDa (Fig. 5.1). The 85 kDa fragment apparently corresponds to the previously characterized 85 kDa catalytic domain of the enzyme that is generated by endogenous cellular proteases and has been observed after limited treatment of DNA ligase I with subtilisin, trypsin, chymotrypsin, V8 protease (Tomkinson et al., 1990). Digestion of a DNA ligase II-AMP complex with trypsin yielded a completely different set of peptides (Fig. 5.1). DNA ligase II migrates with an apparent molecular mass of approximately 70 kDa on SDS-PAGE. The major tryptic peptides in this case were 41, 40 and 10 kDa. Similarly, further experiments within the laboratory using V8 protease clearly distinguished between the two DNA ligases (Roberts et al., 1994). A 22 kDa common active site domain generated after partial V8 digestion and reported to be common to both DNA ligases I and II (Yang et al., 1990) was not observed in the present experiments, and there was no indication of any shared active site fragment above a molecular mass of approximately 10 kDa.

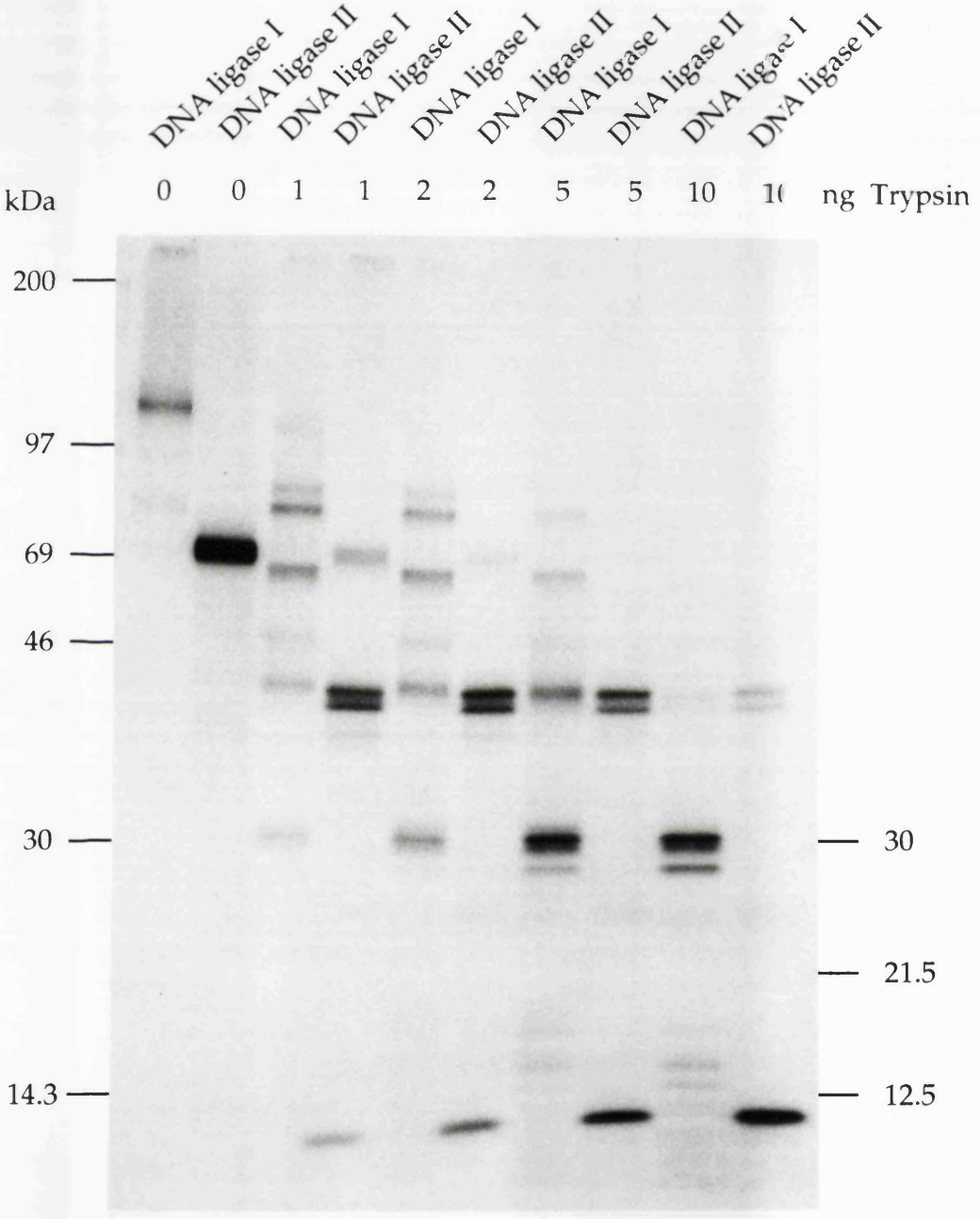


Fig. 5.1. **Partial Trypsin Digestion of Adenylylated DNA Ligases I and II** - Radiolabelled DNA ligase-AMP complexes were incubated for 15 min at 37°C with trypsin as indicated. Products were separated on 10-20 % gradient SDS-PAGE and detected by autoradiography.

5.2 Comparison of the Active Site Peptides of DNA Ligases I and II after Complete Proteolytic Digests

Since partial digests of DNA ligase I had not shown any similarity to DNA ligase II, it was of interest to establish the relationship of the two enzymes within a smaller region immediately flanking the active site by complete digestion of the radiolabelled DNA ligase-AMP complex.

5.2.1 Analysis of the radiolabelled active site peptides of DNA ligases I and II by two-dimensional peptide mapping

Partially purified DNA ligases I and II were adenylylated, separated by 15 % SDS-PAGE and transferred onto Immobilon-P membrane. The two DNA ligases could be readily distinguished following autoradiography (Fig. 5.2). A radiolabelled band could then be excised from the blot, free from contamination by other DNA ligases. Adenylylated DNA ligases were digested, released from the membrane, and samples of the radiolabelled peptides from DNA ligases I and II were analysed in two dimensions by electrophoresis/chromatography on TLC plates. Initial results indicated a close similarity between the two enzymes and for direct comparison both samples were loaded at the same origin.

Following electrophoresis and chromatography, autoradiography of the TLC plate showed only a single radioactive spot, indicating that the products from the two enzymes were the same (Fig. 5.3). However, certain steps in the preparation of the peptides for two-dimensional mapping are executed at low pH and the DNA ligase-AMP bond is known to be somewhat sensitive to acidic pH (Tomkinson et al., 1991b). Thus, it was possible that cleavage of the bond had occurred during these procedures, and this generated the same radioactive product from the two DNA ligases.

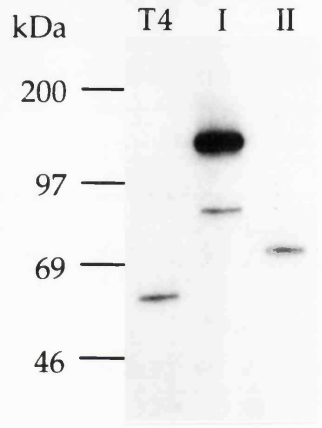


Fig. 5.2. **Adenylylated DNA Ligases prior to Analysis by Tryptic Peptide Mapping** - T4 DNA ligase and DNA ligases I and II were adenylylated as described in Materials and Methods, resolved by 7.5 % SDS-PAGE and transferred to Immobilon-P membrane. Radiolabelled bands were visualised by autoradiography. As well as the full length 125 kDa DNA ligase I, this fraction also contained some of the 85 kDa proteolytic product. Bands were cut from the membrane and used for two-dimensional maps following digestion with trypsin.

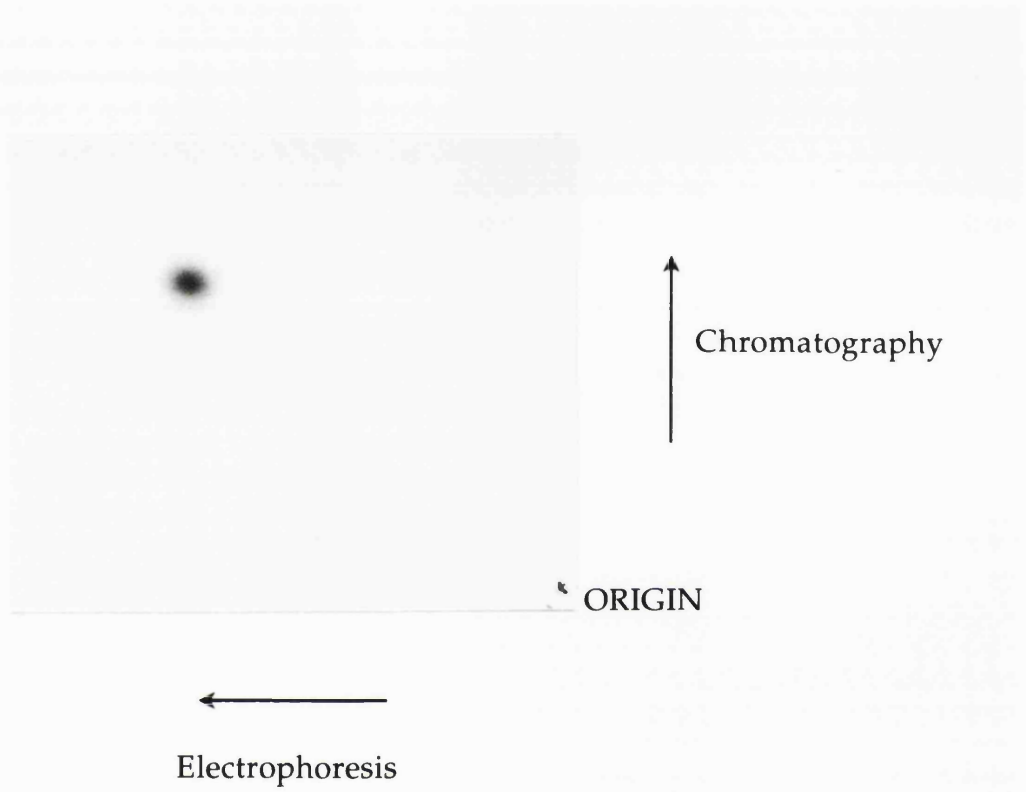


Fig. 5.3. **Two-Dimensional Peptide Mapping of the Adenylylated Active Site Peptides of DNA Ligases I and II** - DNA ligases I and II were adenylylated and radiolabelled active site peptides generated by digestion with trypsin. Samples were loaded on a TLC plate at a single origin, as marked. Electrophoresis at pH 4.72, 1 kV for 1 h, was carried out in the direction indicated followed by ascendant chromatography. Peptides were visualised by autoradiography.

To establish whether the final product was in fact released AMP, samples were run on a cellulose plate, allowing migration of nucleotides, alongside [^{14}C]AMP. The product of trypsin digestion of radiolabelled T4 DNA ligase and radiolabelled AMP comigrated using this technique, suggesting that AMP had been released during preparation for two-dimensional mapping (Fig. 5.4A). Comparison of the DNA ligase I sample and [^{14}C]AMP confirmed this observation (Fig. 5.4B). In an attempt to retain the [^{32}P]AMP radiolabel, the acidic steps of the procedure were omitted or adjusted to a higher pH (Materials and Methods, section 2.9). Electrophoresis of samples was then attempted at higher pH to avoid dissociation of the label from the peptide. However, poor migration of the DNA ligase peptides in this modified buffer system was insufficient to allow any distinction between them to be visible.

5.2.2 Analysis of the active sites of DNA ligases I and II by HPLC

Reversed phase HPLC of peptides is usually performed in a buffer containing trifluoroacetic acid but due to the lability of the DNA ligase-AMP bond in acidic conditions, as encountered with the two-dimensional peptide mapping, it was necessary to establish a buffer system that allowed separation of peptides at a higher pH. The use of ammonium bicarbonate with acetonitrile (suggested by D. Pappin, Protein Sequencing Laboratory, ICRF) allowed good separation of peptides in this system at approximately neutral pH, as shown by the analysis of the two synthetic peptides in Fig. 5.5.

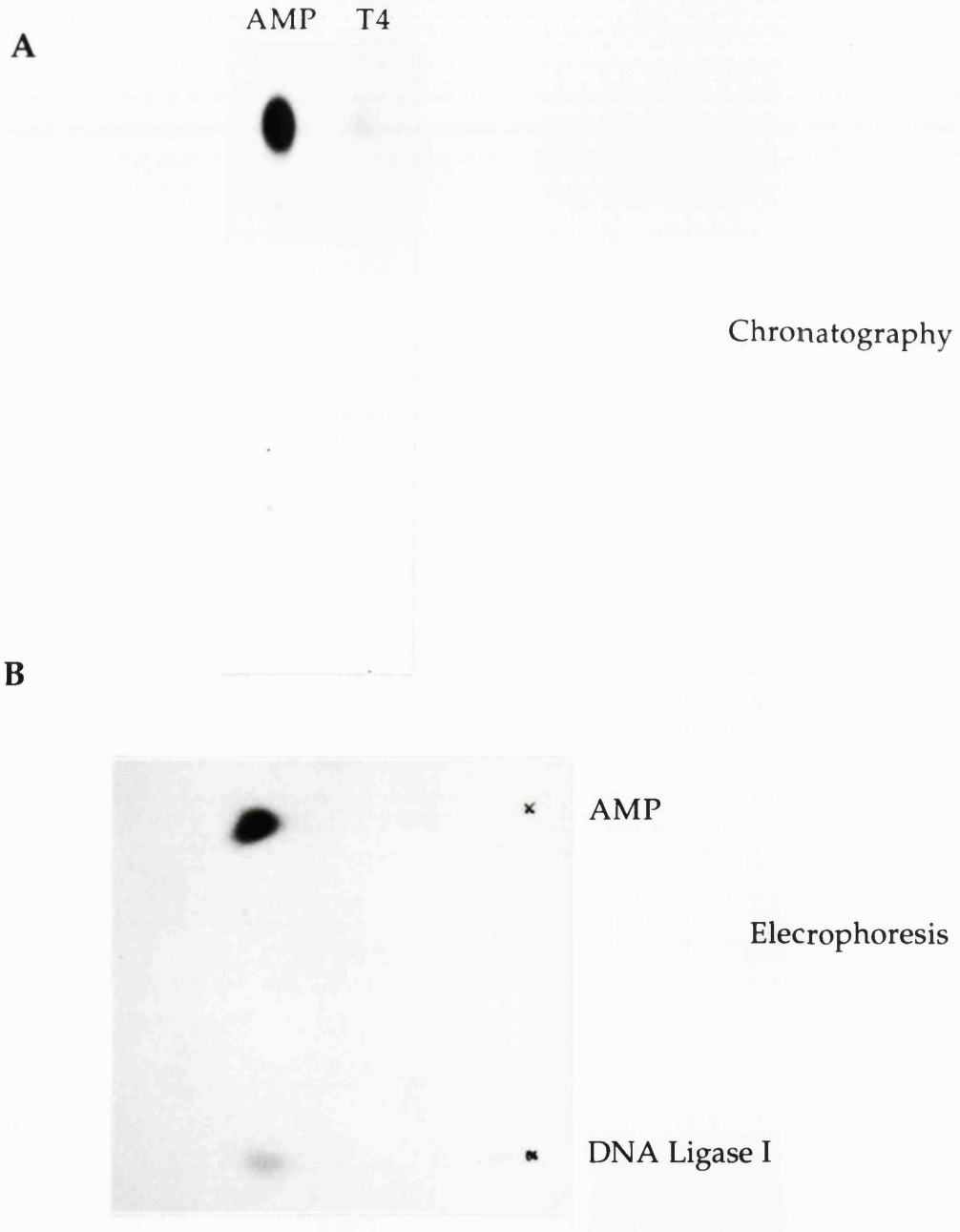


Fig. 5.4. **Comparison of Possible Active Site Peptides with AMP using Electrophoresis and Chromatography** - Trypsin digests of T4 DNA ligase and bovine DNA ligase I were compared to [^{14}C]AMP by: (A) chromatography on a PEI cellulose plate; or (B) electrophoresis at pH 6.5, 1 kV for 1 h.

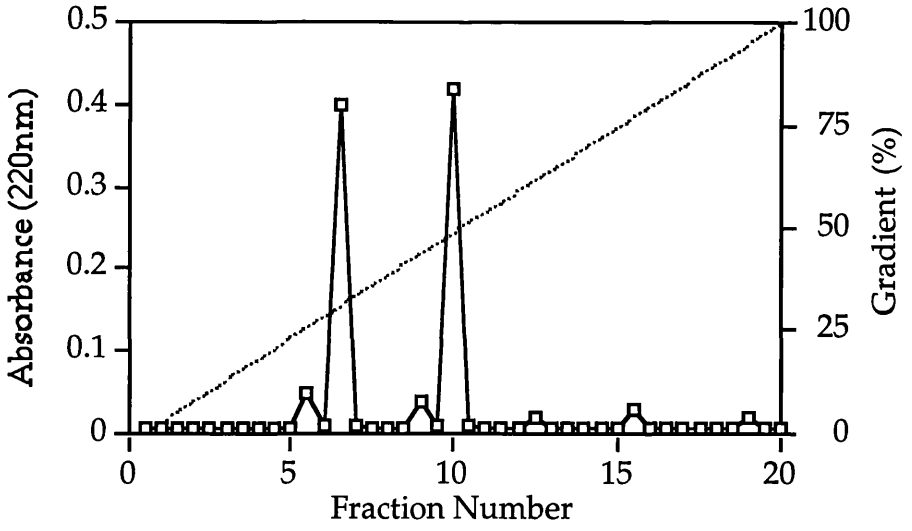


Fig. 5.5. **Separation of Synthetic Peptides by Reversed Phase HPLC** - Peptides were separated using a linear gradient to 100 % buffer B (.....) as detailed in Materials and Methods. Absorbance was monitored at 220 nm (—□—).

To further eliminate the possibility of confusion due to free AMP or to unreacted ATP, elution profiles were established for the nucleotides using these buffer conditions. As is shown in Fig. 5.6, AMP eluted in the column flow-through giving no interference with peptide fractions. The elution profile for ATP was the same as that for AMP (data not shown).

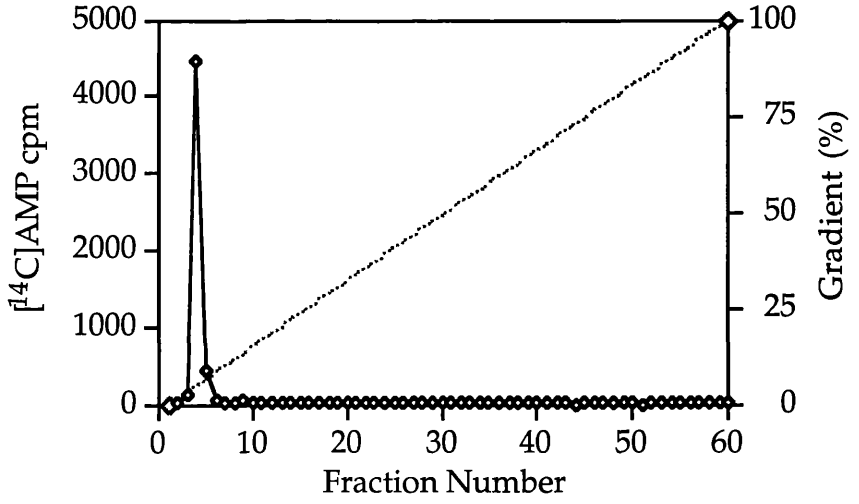


Fig. 5.6. **Elution Profile of ^{14}C-AMP by Reversed Phase HPLC** - Radiolabelled AMP was eluted using a linear gradient to 100 % buffer B (.....) and radioactivity assessed by scintillation counting (—◇—).

Initial analysis of the tryptic DNA ligase I active site peptide had produced multiple peaks of radioactivity (Fig. 5.7). As these were unlikely to represent partial digestion products under the conditions used, it was thought that they may be due to interactions through a cysteine residue, known from the DNA ligase I cDNA sequence to be present in this peptide. These interactions were eliminated by irreversible reaction with iodoacetamide (Lane, 1978). Separation on HPLC after this treatment gave a single peak of radioactivity that comigrated with one of the peaks from the non-treated sample (Fig.5.7). Subsequently, this step was included in the preparation of peptides analysed in this way.

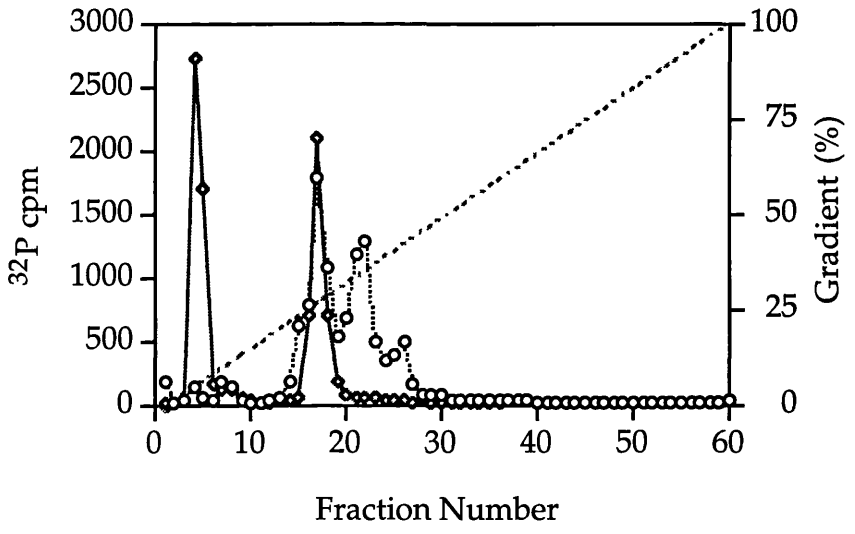


Fig. 5.7. **Analysis of Iodoacetamide Treated DNA Ligase I Active Site Tryptic Peptide by Reversed Phase HPLC** - DNA ligase I peptides before treatment (·····o·····) and after treatment (—◆—) separated by a linear gradient to 100 % buffer B (·····).

The radiolabelled active site peptides from DNA ligases I and II were treated with iodoacetamide and loaded as a single injection onto the HPLC column. Using a shallow gradient of 15 to 30 % buffer B, to obtain maximum separation, resolved only a single peak of radioactivity (Fig. 5.8). As gel electrophoresis (see below) indicated that the active site peptides from these two enzymes can be separated the inability to

distinguish between them here may reflect the limits of this technique. Later work using these partially purified fractions of DNA ligases I and II suggested that the digestion conditions may also not have been optimal here; a certain amount of non-specific degradation was seen with digestion times longer than a few hours (see below).

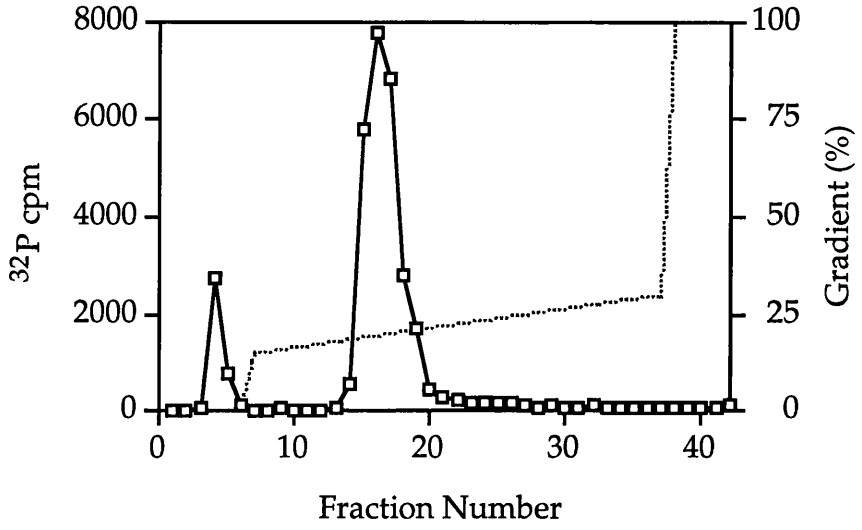


Fig. 5.8. **Comigration of Active Site Tryptic Peptides of DNA Ligases I and II on Reversed Phase HPLC** - DNA ligases I and II samples were loaded together and eluted with a shallow gradient as indicated (·····). Elution of the peptides was monitored by scintillation counting (—□—).

5.2.3 Analysis of the active site peptides of DNA ligases I and II by non-denaturing polyacrylamide gel electrophoresis

Partially purified DNA ligases from calf thymus or liver were radiolabelled with [α-^{32}P]ATP and active site peptides characterised following complete digestion with trypsin, Lys C or V8 protease.

Titrating the protease concentrations and varying digestion times (Fig. 5.9), suggested that high protease concentrations for short periods of time were preferable to prolonged incubation periods with lesser amounts of enzyme to minimize slow non-enzymatic hydrolysis of the Lys-AMP phosphoamide bond. Incubation with protease for 16 h (as recommended

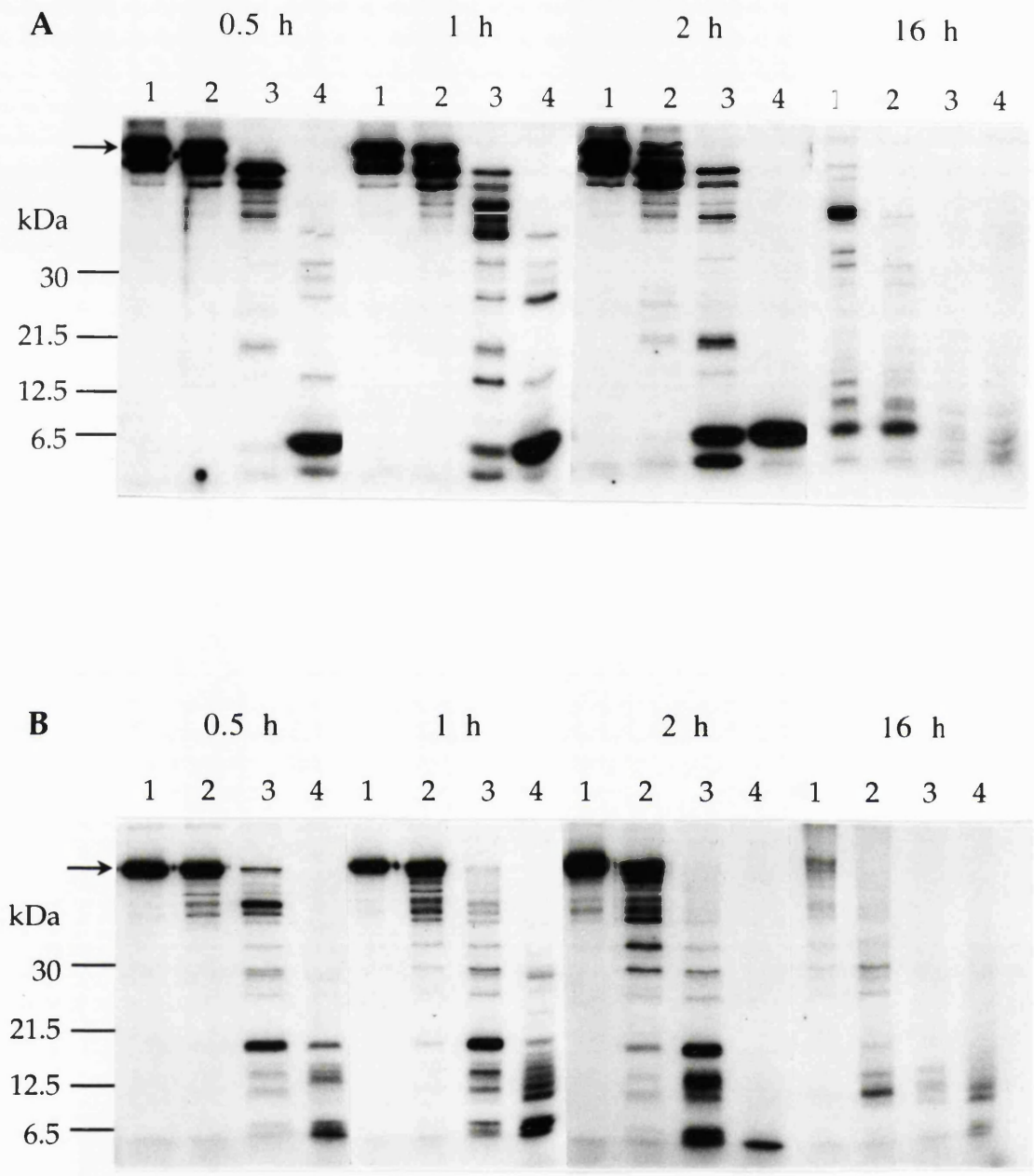


Fig. 5.9. **V8 Digestion of Adenylylated DNA Ligases I and II** - Adenylylated proteins were digested with a range of V8 concentrations for various times, as indicated, and products were separated on 15 % SDS-PAGE. (A) DNA ligase I; (B) DNA ligase II. Full length adenylylated enzymes are indicated with arrows. Lane 1, 0.5 % protease; lane 2, 5 % protease; lane 3, 50 % protease; lane 4, 500 % protease (w/w).

by the manufacturer) produced significant losses of radiolabelled products (Fig. 5.9).

In preliminary studies, several different methods of polyacrylamide gel electrophoresis were employed in attempts to resolve the peptide products of these digests. 12.5 % SDS-urea gels did not resolve these products adequately to be able to distinguish individual peptides (Swank and Munkres, 1971). The tricine-SDS system of Schagger and von Jagow (Schagger and von Jagow, 1987) gave better resolution, although the bands produced were somewhat indistinct (data not shown). However, the native gel system of Pantazis and Bonner (Pantazis and Bonner, 1981) with high polyacrylamide concentrations and no SDS consistently yielded clear results. With this gel system, it was possible to resolve small proteolytic products according to both the molecular mass and charge of the peptide.

The tryptic peptide containing the Lys-AMP has been sequenced for bovine DNA ligase I and is identical to a 16-amino acid (1.9 kDa) peptide defined by Arg residues in the human cDNA (Tomkinson et al., 1991b). The probable molecular masses of other active site peptides of bovine DNA ligase I could be estimated from the predicted amino acid sequence of the human cDNA (Fig. 5.10).

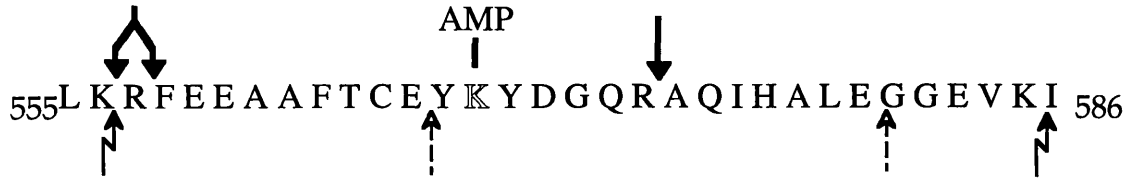


Fig. 5.10. **Protease Cleavage Sites in the DNA Ligase I Active Site Sequence** - Amino acid residues 555 to 586 in mammalian DNA ligase I, derived from the human DNA ligase I cDNA (Barnes et al., 1990). Residues 558 to 573 have also been determined by amino acid sequencing of a bovine DNA ligase I tryptic peptide and shown to be identical with the human sequence (Tomkinson et al., 1991b). The active site Lys 568, and cleavage sites around this residue for trypsin (↓), Lys-C (↑) and V8 protease (⇡).

Lys-C cuts on the C-terminal side of lysine residues but does not recognise a Lys-AMP moiety, digestion of DNA ligase I-AMP would therefore yield a peptide of 3.3 kDa (29 amino acids) (Fig. 5.11A). Two peptides derived from DNA ligase I were different from the single, more slowly migrating DNA ligase II peptide (Fig. 5.11A). A three-fold increase or decrease in the amount of Lys-C protease used for digestion produced no change in migration of the DNA ligase-derived peptides. However, blocking of cysteine residues using iodoacetamide converted the two radioactively labelled DNA ligase I peptides to a single band (Fig. 5.11A, lanes 1 and 2). This is in agreement with the DNA ligase I sequence which shows that the active site peptide contains a Cys residue (Fig. 5.10.).

The tryptic peptides of DNA ligase I were obtained consistently after 4 h incubation with a range of high trypsin concentrations (0.5-2.5 times the amount of substrate). Four different peptides were detected, which were reduced to two by iodoacetamide treatment (Fig. 5.11B, lanes 1 and 2), consistent with the presence of a Cys residue in the active site peptide. These migrated more rapidly than the Lys-C peptides of DNA ligase I under identical conditions (data not shown), which would be expected for smaller peptides of similar charge (Fig. 5.10). DNA ligase II (iodoacetamide-treated) also yielded two tryptic peptides (Fig. 5.11B, lane 4), but these migrated significantly more slowly than the peptides from DNA ligase I (Fig. 5.11B, lanes 2 and 4). The occurrence of two different tryptic peptides of DNA ligase I might be explained by the presence of two adjacent basic amino acids at the N-terminal cleavage site (Fig. 5.10). Cleavage could occur at either the lysine or arginine residues but cleavage at the lysine residue would preclude the subsequent removal of the N-terminal arginine in any one peptide (D. Pappin, personal communication). This would generate two peptides; one with an N-terminal phenylalanine residue, which is the DNA ligase I active site

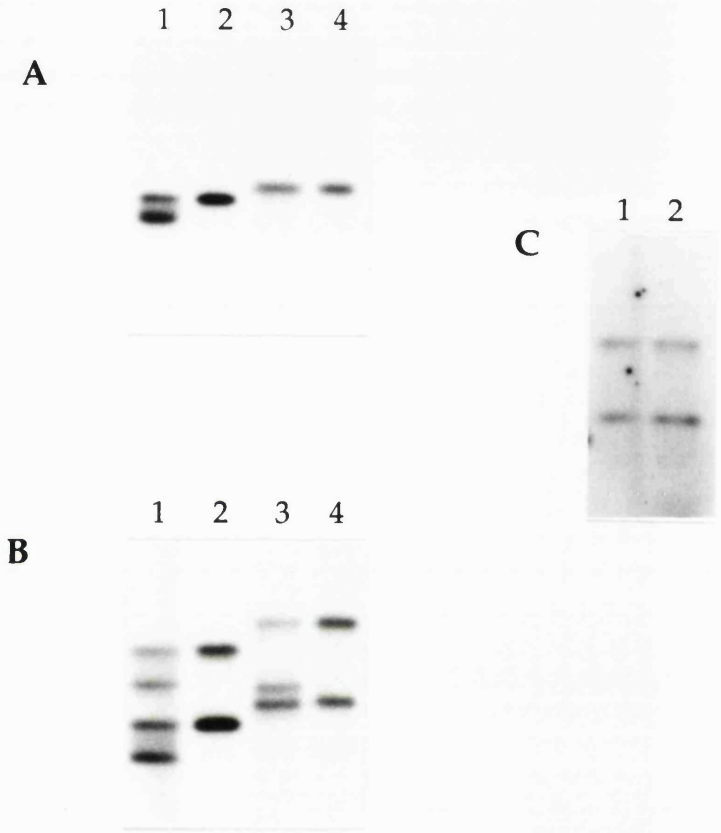


Fig. 5.11. **Analysis of Radiolabelled Active Site Peptides from DNA Ligases I, II and III by Non-denaturing Gel Electrophoresis** - Radiolabelled enzyme-AMP was digested with 250 % (w/w) protease for 4 h at 37 °C. (A) Digestion with Lys C, products separated on a 25 % gel; lane 1, DNA ligase I; lane 2, DNA ligase I treated with iodoacetamide; lane 3, DNA ligase II; lane 4, DNA ligase II treated with iodoacetamide. (B) Digestion with trypsin, products separated on a 30 % gel; lane 1, DNA ligase I; lane 2, DNA ligase I treated with iodoacetamide; lane 3, DNA ligase II; lane 4, DNA ligase II treated with iodoacetamide. (C) Digestion with trypsin, products separated on a 30 % gel; lane 1, DNA ligase II; lane 2, DNA ligase III; both treated with iodoacetamide.

tryptic peptide previously sequenced (Tomkinson et al., 1991b), the other would be a slightly longer peptide with an N-terminal arginine residue, which would be less negatively charged and consequently would migrate more slowly in the gel electrophoresis system employed here.

Complete digestion of DNA ligase II with trypsin showed an extra band prior to treatment with iodoacetamide, indicating the presence of a Cys residue, however, two bands still remained after treatment. Lys-C digestion of DNA ligase II yielded a single radioactive peptide. These data could be explained by potential Lys-C and trypsin sites within a sequence context such as -*X*RCK*X*- at the N-termini of the active site peptides.

Digestion of DNA ligase I-AMP to completion with V8 protease yielded a single radioactively labelled peptide which migrated slightly more rapidly than the Lys-C-derived peptide shown in Fig. 5.11A, (lane 2) under identical conditions. The peptide does not contain a cysteine residue (Fig. 5.10) and, as expected, there was no effect of iodoacetamide treatment. In contrast, no active site peptide of V8-digested DNA ligase II was detected in repeated attempts. It seems likely that extensive V8 protease treatment of DNA ligase II generated a peptide that is either very small (or highly charged) such that it migrates near the unincorporated ATP on these gels and may be obscured by this radiolabel. In this connection, the Gln residue at position 572 in the mammalian DNA ligase I is unusual in that several other DNA ligases with evolutionarily conserved active site regions have a Glu residue at this position (Tomkinson et al., 1991b). Thus, the sequence at the active site of *Saccharomyces cerevisiae* and *Schizosaccharmyces pombe* is -EY**K**YDG*E*-, whereas that of mammalian DNA ligase I is -EY**K**YDG*Q*-. A DNA ligase containing the former sequence would yield a very short V8 peptide comprising the active site Lys residue. A second possibility is that the peptide is not negatively charged in the buffer system used. Despite the problem of identifying the DNA ligase II V8 peptide, it is

clearly not the same as the DNA ligase I V8 peptide that could be resolved in this system, confirming the difference between the active site peptides of DNA ligases I and II generated with the other proteases.

5.3. Comparison of the Active Site Peptides from DNA Ligases II and III

Having successfully resolved the active site peptides of DNA ligases I and II in a non-denaturing gel electrophoresis system following complete proteolytic digestion of a radiolabelled DNA ligase-AMP intermediate, the method was extended to a comparison of DNA ligases II and III. Digestion to completion of DNA ligase III-AMP with trypsin yielded active site peptides indistinguishable from those of DNA ligase II (Fig. 5.11C), and thus different from DNA ligase I. Moreover, parallel work in the laboratory comparing radiolabelled DNA ligases II and III after partial proteolytic digests (as carried out here for DNA ligases I and II, see section 5.1) showed that several small peptides (< 35 kDa) were common, whereas larger fragments differed (R. Nash, personal communication, Roberts et al., 1994). These data indicate close similarity or identity of the region around the active site Lys residue between DNA ligases II and III.

CHAPTER 6

DISCUSSION

Two DNA ligases have been previously identified in mammalian cells and in other eukaryotes, including *Drosophila, Xenopus* and the budding yeast, *S. cerevisiae* (Teraoka et al., 1986; Takahashi and Senshu, 1987; Tomkinson et al., 1991a; Aoufouchi et al., 1992; Tomkinson et al., 1992). Mammalian DNA ligase I has been well characterised biochemically, and the cDNA encoding this enzyme has been cloned and sequenced (Barnes et al., 1990), allowing comparison to DNA ligases from other species and the identification of certain conserved features (Barnes et al., 1990; Tomkinson et al., 1991b). Mammalian DNA ligase II appears to be a minor activity that has been purified from calf thymus and shown to be distinct from DNA ligase I, both in its chromatographic properties and in its *in vitro* specificities for homopolymer substrates (Arrand et al., 1986; Teraoka et al., 1986); both enzymes are able to ligate oligo(dT) molecules hybridised to poly(dA), but only DNA ligase II is also able to ligate oligo(dT) hybridised to poly(rA). Immunological studies using antisera against purified DNA ligase I or DNA ligase II have shown that these enzymes are also serologically unrelated (Soderhall and Lindahl, 1975; Teraoka et al., 1986). An inhibitory antiserum against DNA ligase I did not affect the activity of DNA ligase II in *in vitro* assays (Soderhall and Lindahl, 1975) and an inhibitory antiserum against DNA ligase II had no affect on DNA ligase I acitivity (Teraoka et al., 1986). These data suggest that DNA ligase II is a distinct enzyme from DNA ligase I. However, some confusion still remained, as certain groups had failed to detect DNA ligase II in mammalian cells, while others had reported similarities between the two enzymes. The object of this study was to establish the relationship between the two proteins.

DNA ligase II was initially purified to apparent homogeneity from calf thymus glands. Further analysis of the final purified fraction, by peptide sequencing and two-dimensional gel electrophoresis, showed that

it contained at least two distinct 70 kDa proteins. One of these proteins was identified by sequence comparisons to be the abundant structural protein, ezrin/cytovillin (Gould et al., 1989; Turunen et al., 1989). The contamination of DNA ligase II by this protein is probably accounted for by their common ability to adsorb very strongly to hydroxylapatite. This matrix was used here as a key step in the purification of DNA ligase II and is similarly used in the purification of ezrin/cytovillin (Bretscher, 1983), therefore the DNA ligase II purification was biased to include a certain amount of ezrin/cytovillin. A single N-terminal sequence was obtained from this preparation and corresponded to that of ezrin/cytovillin, suggesting that DNA ligase II has a blocked N-terminus and is a primary translation product. No obvious degradation was seen during the purification and immunoblotting analysis of crude extracts by Teraoka et al., suggesting that this is the full length product (Teraoka et al., 1986).

In an attempt to avoid contamination by ezrin/cytovillin, DNA ligase II was purified from bovine liver. This was successful, as analysis of the bovine liver preparation using an antiserum against the N-terminus of ezrin/cytovillin has indicated that none of this contaminating protein was present (R. Nash, personal communication). In addition, the DNA ligase II activity, assayed by adenylylation, appeared to be greater than that purified from calf thymus, although the amounts of DNA ligase II protein, identified by silver staining, appeared to be much less. However, accurate comparisons can not be made due to the difficulties in assessing the quantity of DNA ligase II protein in the calf thymus preparation. The two-dimensional gels of calf thymus DNA ligase II showed that a second spot of protein was present in similar amounts to ezrin/cytovillin, but the possibility remains that this spot represents a second contaminating protein. Indeed, if levels of DNA ligase II in other tissues are similar to those seen in liver, the protein would not have been visible with the

staining procedure used for these gels. A specific antiserum against DNA ligase II would be able to resolve these possibilities. It is now hoped that peptide sequences can be obtained from the bovine liver DNA ligase II preparation by isolation of a specific band after electrophoretic transfer of the protein preparation to a membrane (R. Nash, personal communication).

Purification of DNA ligase II from both calf thymus and liver had confirmed that this enzyme was chromatographically distinct from DNA ligase I and resembled that previously characterised (Soderhall and Lindahl, 1973b; Teraoka et al., 1986). DNA ligases I and II are also distinct in their substrate specificities, with DNA ligase II being able to ligate oligo(dT) hybridised to poly(rA), whereas DNA ligase I can not. These data were confirmed here and the analysis extended to include the substrate, oligo(rA)/poly(dT). DNA ligase II showed no ability to ligate this substrate, whereas DNA ligase I was able to do so efficiently. Differences were also observed between the two enzyme activities in their response to increasing levels of sodium chloride, with DNA ligase II being extremely sensitive to concentrations above 30 mM that did not significantly affect DNA ligase I activity.

The lack of serological cross-reactivity between DNA ligases I and II was confirmed using antisera directed against purified DNA ligase I. An antiserum against a peptide conserved between the vaccinia virus and yeast DNA ligases, and mammalian DNA ligase I (Lasko et al., 1990), did not cross-react detectably with calf thymus DNA ligase II, suggesting that the relevant sequence is not precisely conserved in this protein. As the quantity of DNA ligase II protein in the calf thymus fraction is uncertain, the possibility of error arises in assessing these immunoblotting experiments, as insufficient protein would also give a negative result. This problem is circumvented by using an inhibitory DNA ligase I antiserum

and measuring enzyme activity. No inhibition of DNA ligase II activity was seen using this antiserum, in apparent agreement with the immunoblotting data and suggesting that the two proteins are unrelated. As more accurate estimates of DNA ligase II protein quantities can be made with the liver preparation, a reassessment of the immunoblotting data has been possible. More recent immunoblotting analysis using greater quantities of the protein from bovine liver has shown cross-reactivity of the antiserum against the conserved peptide with DNA ligase II, implying that DNA ligase II does in fact contain a version of this epitope in common with a number of the eukaryotic DNA ligases (P. Robins, personal communication). The vaccinia virus DNA ligase showed a positive reaction with this antibody, as expected from its amino acid sequence.

During the course of the studies on DNA ligases I and II, a further high molecular mass species, with the ability to ligate oligo(dT)/poly(rA), had been identified from calf thymus and termed DNA ligase III (Tomkinson et al., 1991a). This enzyme has distinct chromatographic and biochemical properties from both the previously characterised calf thymus DNA ligases. Many properties are shared with the second high molecular mass DNA ligase from rat liver observed by Elder and Rossignol and designated "DNA ligase II" (Elder and Rossignol, 1990), and the two may be the same enzyme or closely related. However, unlike the rat liver DNA ligase II, the bovine DNA ligase III here was not able to efficiently utilise ATP-αS as a cofactor in place of ATP. Studies were carried out on the substrate specificity of this enzyme to establish its relationship to the other DNA ligases. Initially, DNA ligase III was thought to have the same substrate specificity as DNA ligase II, being able to ligate both the oligo(dT)/poly(dA) and oligo(dT)/poly(rA) substrates. Use of a third substrate, oligo(rA)/poly(dT), established a range of substrate specificities

distinct from that of either DNA ligase I or II, where DNA ligase III can ligate all three substrates so far analysed (see Table 6.1).

Immunoblotting analysis of DNA ligase III using an antiserum against DNA ligase I supported the catalytic data and suggested that this protein is not related to DNA ligase I. Neither did an inhibitory antiserum against DNA ligase I affect the ligation activity of DNA ligase III *in vitro*. As already described for DNA ligase II, work here with partially purified DNA ligase III indicated that it did not contain the conserved DNA ligase peptide, while more recent work with larger available quantities of the enzyme has shown that it too contains this or a similar epitope, implying that the peptide is an important element of eukaryotic DNA ligases. An antiserum against a possible DNA ligase II peptide did not recognise either DNA ligases I or III, but the exact origin of this peptide has not been conclusively determined. Specific antisera against DNA ligase II or III will allow further analysis of the relationship between these two enzymes. Immunoblotting of crude extracts from rapidly lysed cells would establish whether these are both present *in vivo* or whether DNA ligase II is produced by proteolytic processing.

The properties of the three mammalian DNA ligases are summarised in the table below.

Table 6.1. **Properties of Mammalian DNA Ligases**

	I	II	III
Size by SDS-PAGE	125 kDa	72 kDa	100 kDa
Size by cDNA sequence	102 kDa	-	-
Adsorption to hydroxylapatite	Weak	Strong	Moderate
Ligation of oligo(dT)/poly(dA)	+	+	+
Ligation of oligo(dT)/poly(rA)	-	+	+
Ligation of oligo(rA)/poly(dT)	+	-	+
K_m for ATP	Low	High	Low
Inhibition by Ligase I Antiserum	+	-	-

DNA ligases II and III are smaller than DNA ligase I, as estimated by SDS-PAGE, and despite the biochemical and catalytic evidence to the contrary, it remained a possiblity that they were active fragments of the full length DNA ligase I, produced by proteolysis during the purification procedure as is the 85 kDa C-terminal catalytic domain. Yang and coworkers had reported that DNA ligases I and II from mammalian cells had a common 22 kDa AMP-binding domain noted after digestion with the protease V8, suggesting that DNA ligases I and II were isozymes, possibly sharing a common ancestral gene (Yang et al., 1990). Initial studies of partial digests of the adenylylated DNA ligase I using subtilisin had not revealed any fragments the size of full length DNA ligase II or III (Tomkinson et al., 1990). In the present work, a more detailed study was undertaken.

A comparison of the products of partial proteolytic digestion with trypsin indicated that DNA ligases I and II are not directly related, as common proteolytic fragments were not observed. Analysis of the partial products from V8 digestion has confirmed this observation; no common 22 kDa domain was detected, and the two proteins did not share any common fragments above a molecular mass of 10 kDa (R. Nash, personal communication, Roberts et al., 1994). Deletion mutagenesis at both termini of the DNA ligase I cDNA has shown that it is possible to generate an active fragment of 80 kDa from DNA ligase I but further deletions remove evolutionarily conserved sequences and result in enzyme inactivation. It is not possible to generate an active fragment the size of DNA ligase II from DNA ligase I. Any sequence identity between the two enzymes was likely to be within the region of the active site as they share a common reaction mechanism. Analysis of the peptides produced by complete digestion of DNA ligases I and II with trypsin or Lys-C confirmed

that DNA ligases I and II are distinct enzymes, and showed that the sequence of the active site of DNA ligase II is different from that of DNA ligase I.

Having compared the active site peptides from DNA ligases I and II and established that they are different, the relationship between DNA ligases II and III was analysed. Following complete digestion with trypsin or Lys-C, the active site peptide of DNA ligase III was indistinguishable from that of DNA ligase II, but clearly different from DNA ligase I. This implies a close similarity or identity in the region around the active site Lys residue in DNA ligases II and III. The larger fragments produced from partial digests of DNA ligase III with trypsin or V8 were not similar to the initial products of DNA ligase II digestion. However, the smaller products of these digests (< 35 kDa) were similar or identical between DNA ligases II and III (R. Nash, personal communication, Roberts et al., 1994). The relationship between DNA ligases II and III is still not clear, although they are obviously closely related to one another and unrelated to DNA ligase I.

The simplest conclusion from the data presented here might be that DNA ligase II is a proteolytic fragment of DNA ligase III generated *in vivo* or during the purification procedure. DNA ligase II is more easily purified than full length DNA ligase III, suggesting that it may be a stable fragment produced *in vivo.* DNA ligase III may be a precursor of DNA ligase II, and a specific protease in mammalian cells may generate the smaller active form, DNA ligase II. There is also evidence to the contrary as it is not possible to generate a product *in vitro* from DNA ligase III that corresponds to the full length DNA ligase II. Thus, the 100 kDa full length DNA ligase III is partly converted to an enzymatically active 87 kDa form by an endogenous protease during enzyme purification, but no conversion to an active 70 kDa form has been detected. Moreover, attempts to obtain an N-terminal sequence of bovine liver DNA ligase II have failed,

indicating that DNA ligase II has a blocked N-terminus (R. Nash and D. Pappin, personal communication); this would be expected for a full length protein but not for an active fragment. Takahashi and Tomizawa have described two forms of DNA ligase II from *Drosophila*, with molecular masses of 90 kDa and 70 kDa (Takahashi and Tomizawa, 1990), suggesting that the higher molecular mass form may in fact be DNA ligase III. They also observed the similarities in digestion patterns identified here between DNA ligases II and III, while both enzymes were again different from DNA ligase I. Alternatively to proteolytic processing of a single gene product, DNA ligases II and III may be distinct enzymes within the cell and have particular roles to play that do not overlap. It is possible that they are derived from two closely related genes or from a single gene transcript by alternative splicing, retaining some common features, while gaining specific functions, as manifested by the differences in substrate specificity, from other domains.

The work presented here has shown that mammalian DNA ligases I and II are distinct enzymes. This implies that they may have different roles within the cell. The differences in substrate specificity, although observed *in vitro* using homopolymer substrates, may give some indication as to the possible roles of the DNA ligases *in vivo*, e.g. the ability of DNA ligase II to ligate single strand breaks in a DNA/RNA hybrid may be important during transcription; precise roles have yet to be defined. The biological role of DNA ligase I in mammalian cells has been most closely examined and data have indicated that it is involved in DNA replication. Cell staining studies have shown that DNA ligase I is localised in the nucleus and is recruited to replication factories, along with other proteins known to be involved in DNA replication, upon Herpes virus infection (Lasko et al., 1990; Wilcock and Lane, 1991). DNA ligase I activity is increased upon cellular proliferation and mutations affecting this

enzyme in human cells, or its homologue in yeast, produce defects in DNA replication (Soderhall, 1976; Nasmyth, 1977; Johnston and Nasmyth, 1978; Barnes et al., 1992b). The human DNA ligase I is able to complement these defects in *S. cerevisiae* and *E. coli* conditional lethal mutants, restoring competent DNA replication and allowing growth at the restrictive temperature (Barnes et al., 1990; Kodama et al., 1991), implying that DNA ligase I is the replicative DNA ligase. Data on the defect in the human cell line, 46BR, has confirmed this, as the DNA ligase I gene from these cells contains mutations and produces an enzyme that is only partially active and can not maintain normal DNA replication (Prigent et al. 1994).

In contrast to DNA ligase I, the level of mammalian DNA ligase II activity is constant throughout the cell cycle and this enzyme was therefore implicated in DNA repair (Soderhall and Lindahl, 1973b; Creissen and Shall, 1982; Takahashi and Senshu, 1987; Li and Rossman, 1989). Poly(ADP-ribose) polymerase has been shown to have a role in base excision repair (Satoh and Lindahl, 1992) and can apparently modulate DNA ligase II activity in the cell, supporting the possiblity of a role for DNA ligase II in DNA repair. Creissen and Shall reported that poly(ADP-ribosyl)ation, possibly of DNA ligase II itself, produced an increase in its activity (Creissen and Shall, 1982). Ohashi et al. also observed an increase in DNA ligase activity in response to poly(ADP-ribosyl)ation, although they could not ascribe this specfically to DNA ligase II (Ohashi et al., 1983). In line with these results, benzamide, an inhibitor of PARP, appeared to increase the sensitivity of resistant cells to the cytotoxic effects of a DNA cross-linking agent, through inhibition of DNA ligase II activity and the consequent defect in the repair pathway (Institoris et al., 1992). In direct contrast, Yoshihara et al. observed that poly(ADP-ribosyl)ation of either DNA ligase II itself or a regulatory protein inhibited DNA ligase II activity

(Yoshihara et al., 1985); these two lines of evidence have yet to be resolved. The compound, arsenite, is able to inhibit DNA ligase II activity. Arsenite is not mutagenic itself but is thought to increase the mutagenicity of compounds such as DNA methylating agents by inhibiting or modifying the repair process. The inhibition of DNA ligase II by arsenite suggests that this is the target enzyme in the repair pathway (Li and Rossman, 1989).

Although previous work has supported a proposed role for DNA ligase II in DNA repair, studies on the DNA ligase mutants of budding and fission yeast, and more recent identification of a DNA ligase I defect in the mutant human cell line, 46BR, revealed that the situation may not be as simple as DNA ligase I being the replicative ligase and DNA ligase II being involved in DNA repair. *S. cerevisiae cdc9* mutants have, in addition to defective DNA replication, an inability to repair UV-induced DNA damage and altered mitotic recombination, implicating a single DNA ligase, the product of the CDC9 gene, in these processes (Fabre and Roman, 1979; Game et al., 1979; Johnston, 1979). Complementation of a *cdc9* mutant with the human DNA ligase I restores all three functions, suggesting that DNA ligase I, which shows extensive sequence homology with the CDC9 gene product, is capable of functioning in these processes *in vivo* (Barnes et al., 1990; Lindahl and Barnes, 1992). A slow increase in DNA ligase I mRNA and an increase in DNA ligase I activity through *de novo* synthesis of the protein have been observed in mammalian cells in response to UV, and may be important within the cell as a means of regulating DNA ligase I activity in response to damage (Mezzina and Nocentini, 1978; Montecucco et al., 1992). The human cell line deficient in DNA ligase I, 46BR, as well as having a defect in DNA replication, shows sensitivity to a wide range of DNA damaging agents, and *in vitro* studies directly implicate DNA ligase I in excision repair pathways (Prigent et al., 1994). Therefore, the proposed role of DNA ligase II as a repair enzyme is

less clear, as it appears that mammalian DNA ligase I and the *S. cerevisiae* CDC9 homologue are capable of functioning in excision repair processes, and in mutant cell lines, mammalian DNA ligase II or a second DNA ligase activity described in *S. cerevisiae* (Tomkinson et al., 1992), do not appear able to replace the functions of DNA ligase I/CDC9.

Not all processes involving the need for ligation of strand breaks in DNA are defective in a DNA ligase I mutant, as demonstrated by a mutant of the *S. pombe* DNA ligase I homologue, *cdc17*$^+$. This shows similar defects to the *S. cerevisiae cdc9* DNA ligase mutant in defective DNA replication, DNA repair and mitotic recombination. However, examination of meiotic recombination in this mutant has shown it to be normal (Sipiczki et al., 1990). This implies that the *cdc17*$^+$ DNA ligase is not involved in this process, but that this may be the role of a second DNA ligase. In addition, Higashitani and coworkers have described an apparent increase in DNA ligase II activity during mouse spermatogenesis in parallel with other proteins presumed to be involved in meiotic recombination, such as DNA polymerase β (Higashitani et al., 1990). This implicates DNA ligase II in the ligation of strand breaks occuring during meiotic recombination.

It may be that DNA ligases II and III are involved in specific forms of recombination, as work by Hsieh et al. has indicated (Hsieh et al., 1993). They observed that V(D)J recombination is normal in the human DNA ligase I mutant, 46BR, implying that DNA ligase I is not involved in DNA end-joining in this process. As DNA ligases II and III are normal in 46BR (Lehmann et al., 1988; Barnes et al., 1992b), it is possible that one of these may be involved in V(D)J joining. DNA ligase III may also be involved in recombinational repair, as it has been shown to be associated with a complex capable of double strand break repair by recombination *in vitro* (Jessberger, et al., 1993). A defect in DNA ligation has also been seen in

cells from Bloom's syndrome (BS) patients but this is not correlated with a mutation in the structural gene for DNA ligase I (Chan et al., 1987; Willis and Lindahl, 1987; Runger and Kraemer, 1989; Petrini et al., 1991; Barnes et al., 1992b). BS cells exhibit a high level of spontaneous sister chromatid exchange suggestive of altered homologous recombination. Therefore, the possibility remains that there is a defect in a different DNA ligase in BS and that DNA ligase I is unable to substitute for this enzyme. DNA ligase II activity was originally observed to be normal in BS lines (Willis and Lindahl, 1987) but the DNA ligase III activity has not been analysed separately from DNA ligase I and this remains to be studied.

The degree of similarity between the three DNA ligases and their relative roles in mammalian cells will ultimately be determined by cloning of the cDNAs for DNA ligases II and III and comparison with that already obtained for DNA ligase I. Certain features may be found to be common between them, as suggested by the immunoblotting data and their active site peptides, and yet they clearly retain differences within this framework that allow specific functions. Analysis of the expression and regulation of these enzymes would then be possible. Specific antisera against DNA ligases II or III would allow localisation of these enzymes within the cell. Depletion of the DNA ligases from crude cell extracts and addition of the purified enzymes to *in vitro* assay systems for DNA repair and replication may provide further clues to their individual functions, as well as the need for a multiplicity of DNA ligases and their lack of complementarity. Analysis of the DNA ligase I defect in the 46BR individual has indicated the importance of DNA ligases to the mammalian cell. The identification of a second DNA ligase activity in *S. cerevisiae*, an organism that is amenable to genetic manipulation, should allow analysis of mutants for the other DNA ligases, and may provide valuable clues to their *in vivo* roles.

REFERENCES

Anderson, S. and DePamphilis, M.L. (1979). J. Biol. Chem. **254**, 11495-11504.

Aoufouchi, S., Prigent, C., Theze, N., Philippe, M. and Thiebaud, P. (1992). Dev. Biol. **152**, 199-202.

Araki, H., Ropp, P.A., Johnson, A.L., Johnston, L.H., Morrison, A. and Sugino, A. (1992). EMBO J. **11**, 733-740.

Arrand, J.E., Willis, A.E., Goldsmith, I. and Lindahl, T. (1986). J. Biol. Chem. **261**, 9079-9082.

Barker, D.G., White, J.H.M. and Johnston, L.H. (1985). Nucleic Acids Res. **13**, 8323-8337.

Barker, D.G., White, J.H. and Johnston, L.H. (1987). Eur. J. Biochem. **162**, 659-667.

Barnes, D.E., Johnston, L.H., Kodama, K., Tomkinson, A.E., Lasko, D.D. and Lindahl, T. (1990). Proc. Natl. Acad. Sci. USA **87**, 6679-6683.

Barnes, D.E. (1992). Nature **359**, 12-13.

Barnes, D.E., Kodama, K., Tynan, K., Trask, B.J., Christensen, M., De Jong, P.J., Spurr, N.K., Lindahl, T. and Mohrenweiser, H.W. (1992a). Genomics **12**, 164-166.

Barnes, D.E., Tomkinson, A.E., Lehmann, A.R., Webster, A.D.B. and Lindahl, T. (1992b). Cell **69**, 495-503.

Barnes, D.E., Lindahl, T. and Sedgwick, B. (1993). Curr. Opin. Cell Biol. **5**, 424-433.

Bell, S.P. and Stillman, B. (1992). Nature **357**, 128-134.

Bertazzoni, U., Mathelet, M. and Campagnari, F. (1972). Biochim. Biophys. Acta **287**, 404-414.

Biederman, K.A., Sun, J., Giaccia, A.J., Tosto, L.M. and Brown, M. (1991). Proc. Natl. Acad. Sci. USA **88**, 1394-1397.

Borowiec, J.A. and Hurwitz, J. (1988). EMBO J. **7**, 3149-3158.

Boyle, W.J., Van Der Geer, P. and Hunter, T. (1991). Methods Enzymol. **201B**, 110-149.

Bradford, M.M. (1976). Anal. Biochem. **76**, 248-254.

Bravo, R., Frank, R., Blundell, P.A. and Macdonald-Bravo, H. (1987). Nature **326**, 515-517.

Bretscher, A. (1983). J. Cell Biol. **97**, 425-432.

Buratowski, S. (1993). Science **260**, 37-38.

Chan, J.Y.H., Becker, F.F., German, J. and Ray, J.H. (1987). Nature **325**, 357-359.

Cleveland, D.W., Fischer, S.G., Kirschner, M.W. and Laemmli, U.K. (1977). J. Biol. Chem. **252**, 1102-1106.

Cong, P. and Shuman, S. (1993). J. Biol. Chem. **268**, 7256-7260.

Coverley, D., Kenny, M.K., Munn, M., Rupp, W.D., Lane, D.P. and Wood, R.D. (1991). Nature **349**, 538-541.

Cox, M.M. (1993). BioEssays **15**, 617-623.

Creissen, D. and Shall, S. (1982). Nature **296**, 271-272.

Dianov, G., Price, A. and Lindahl, T. (1992). Mol. Cell. Biol. **12**, 1605-1612.

Diffley, J.F.X. and Cocker, J.H. (1992). Nature **357**, 169-172.

Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J. and Waterfield, M.D. (1984). Nature **307**, 521-527.

Elborough, K.M. and West, S.C. (1990). EMBO J. **9**, 2931-2936.

Elder, R.H. and Rossignol, J.-M. (1990). Biochemistry **29**, 6009-6017.

Fabre, F. and Roman, H. (1979). Proc. Natl. Acad. Sci. USA **76**, 4586-4588.

Fairman, M.P., Johnson, A.P. and Thacker, J. (1992). Nucleic Acids Res. **20**, 4145-4152.

Flejter, W.L., McDaniel, L.D., Johns, D., Friedberg, E.C. and Schultz, R.A. (1992). Proc. Natl. Acad. Sci. USA **89**, 261-265.

Fulop, G.M. and Phillips, R.A. (1990). Nature **347**, 479-482.

Game, J.C., Johnston, L.H. and von Borstel, R.C. (1979). Proc. Natl. Acad. Sci. USA **76**, 4589-4592.

Gefter, M.L., Becker, A. and Hurwitz, J. (1967). Proc. Natl. Acad. Sci. USA **58**, 240-247.

Gellert, M. (1967). Proc. Natl. Acad. Sci. USA **57**, 148-155.

Gellert, M. and Bullock, M.L. (1970). Proc. Natl. Acad. Sci. USA **67**, 1580-1587.

Goscin, L.P. and Byrnes, J.J. (1982). Biochemistry **21**, 2513-2518.

Gould, K.L., Bretscher, A., Esch, F.S. and Hunter, T. (1989). EMBO J. **8**, 4133-4142.

Greer, C.L., Peebles, C.L., Gegenheimer, P. and Abelson, J. (1983). Cell **32**, 537-546.

Gumport, R.I. and Lehman, I.R. (1971). Proc. Nat. Acad. Sci. USA **68**, 2559-2563.

Hammond, J.M., Kerr, S.M., Smith, G.L. and Dixon, L.K. (1992). Nucleic Acids Res. **20**, 2667-2671.

Hanawalt, P.C. (1991). Mutat. Res. **247**, 203-211.

Hanawalt, P. and Mellon, I. (1993). Curr. Biol. **3**, 67-69.

Harlow, E. and Lane, D. (1988). Antibodies. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY,

Harosh, I., Naumovski, L. and Friedberg, E.C. (1989). J. Biol. Chem. **264**, 20532-20539.

Henderson, L.M., Arlett, C.F., Harcourt, S.A., Lehmann, A.R. and Broughton, B. (1985). Proc. Natl. Acad. Sci. USA **82**, 2044-2048.

Hendrickson, E.A., Qin, X.-Q., Bump, E.A., Schatz, D.G., Oettinger, M. and Weaver, D.T. (1991). Proc. Natl. Acad. Sci. USA **88**, 4061-4065.

Higashitani, A., Tabata, S., Endo, H. and Hotta, Y. (1990). Cell Struct. Funct. **15**, 67-72.

Hoeijmakers, J.H.J., Eker, A.P.M., Wood, R.D. and Robins, P. (1990). Mutat. Res. **236**, 223-238.

Hoeijmakers, J.H.J. (1993a). TIGS **9**, 173-177.

Hoeijmakers, J.H.J. (1993b). TIGS **9**, 211-217.

Hsieh, C.-L., Arlett, C.F. and Lieber, M.R. (1993). J. Biol. Chem. **268**, 20105-20109.

Huang, J.-C., Svoboda, D.L., Reardon, J.T. and Sancar, A. (1992). Proc. Natl. Acad. Sci. USA **89**, 3664-3668.

Hubscher, U., Kuenzle, C.C. and Spadari, S. (1979). Proc. Natl. Acad. Sci. USA **76**, 2316-2320.

Hubscher, U. and Thommes, P. (1992). TIBS **17**, 55-58.

Institoris, E., Fox, B.W. and Palyi, I. (1992). Cancer Chemother. Pharmacol. **30**, 325-329.

Jessberger, R., Podust, V., Hubscher, U. and Berg, P. (1993). J. Biol. Chem. **268**, 15070-15079.

Johnston, L.H. and Nasmyth, K.A. (1978). Nature **247**, 891-893.

Johnston, L.H. (1979). Mol. Gen. Genet. **170**, 89-92.

Karran, P., Macpherson, P., Ceccotti, S., Dogliotti, E., Griffin, S. and Bignami, M. (1993). J. Biol. Chem. **268**, 15878-15886.

Kastan, M.B., Zhan, Q., El-Deiry, W.S., Carrier, F., Jacks, T., Walsh, W.V., Plunkett, B.S., Vogelstein, B. and Fornace, A.J. (1992). Cell **71**, 587-597.

Kawasaki, H. and Suzuki, K. (1990). Anal. Biochem. **186**, 264-268.

Kerr, S.M. and Smith, G.L. (1989). Nucleic Acids Res. **17**, 9039-9050.

Kletsin, A. (1992). Nucleic Acids Res. **20**, 5389-5396.

Kodama, K., Barnes, D.E. and Lindahl, T. (1991). Nucleic Acids Res. **19**, 6093-6099.

Kuerbitz, S.J., Plunkett, B.S., Walsh, W.V. and Kastan, M.B. (1992). Proc. Natl. Acad. Sci. USA **89**, 7491-7495.

Laemmli, U.K. (1970). Nature **227**, 680-685.

Lane, L.C. (1978). Anal. Biochem. **86**, 655-664.

Langlois, R.G., Bigbee, W.L., Jensen, R.H. and German, J. (1989). Proc. Natl. Acad. Sci. USA **86**, 670-674.

Lasko, D.D., Tomkinson, A.E. and Lindahl, T. (1990). J. Biol. Chem. **265**, 12618-12622.

Lehman, I.R. (1974). Science **186**, 790-797.

Lehmann, A.R., Willis, A.E., Broughton, B.C., James, M.R., Steingrimsdottir, H., Harcourt, S.A., Arlett, C.F. and Lindahl, T. (1988). Cancer Res. **48**, 6343-6347.

Li, J.J. and Kelly, T.J. (1984). Proc. Natl. Acad. Sci. USA **81**, 6973-6977.

Li, J.-H. and Rossman, T.G. (1989). Mol. Toxicol. **2**, 1-9.

Lindahl, T. and Edelman, G.M. (1968). Proc. Natl. Acad. Sci. USA **61**, 680-687.

Lindahl, T. and Barnes, D.E. (1992). Annu. Rev. Biochem. **61**, 251-281.

Lindahl, T. (1993). Nature **362**, 709-715.

Little, J.W., Zimmerman, S.B., Oshinsky, C.K. and Gellert, M. (1967). Proc. Natl. Acad. Sci. USA **58**, 2004-2011.

Lonn, U., Lonn, S., Nylen, U. and Winblad, G. (1989). Carcinogenesis **10**, 981-985.

Mezzina, M. and Nocentini, S. (1978). Nucleic Acids Res. **5**, 4317-4328.

Mezzina, M., Sarasin, A., Politi, N. and Bertazzoni, U. (1984). Nucleic Acids Res. **12**, 5109-5122.

Modrich, P. and Lehman, I.R. (1971). Proc. Nat. Acad. Sci. USA **68**, 1002-1005.

Montecucco, A., Pedrali-Noy, G., Spadari, S. and Ciarrocchi, G. (1988). Biochim. Biophys. Acta **951**, 330-334.

Montecucco, A., Lestingi, M., Pedrali-Noy, G., Spadari, S. and Ciarrocchi, G. (1990). Biochem. J. **271**, 265-268.

Montecucco, A., Biamonti, G., Savini, E., Focher, F., Spadari, S. and Ciarrocchi, G. (1992). Nucleic Acids Res. **20**, 6209-6214.

Morrison, A., Araki, H., Clark, A.B., Hamatake, R.K. and Sugino, A. (1990). Cell **62**, 1143-1151.

Mosbaugh, D.W. and Linn, S. (1984). J. Biol. Chem. **259**, 10247-10251.

Nasmyth, K.A. (1977). Cell **12**, 1109-1120.

Noguiez, P., Barnes, D.E., Mohrenweiser, H.W. and Lindahl, T. (1992). Nucleic Acids Res. **20**, 3845-3850.

O'Farrell, P.H. (1975). J. Biol. Chem. **250**, 4007-4021.

Ohashi, Y., Ueda, K., Kawaichi, M. and Hayaishi, O. (1983). Proc. Natl. Acad. Sci. USA **80**, 3604-3607.

Olivera, B.M. and Lehman, I.R. (1967). Proc. Natl. Acad. Sci. USA **57**, 1426-1433.

Painter, R.B. and Young, B.R. (1980). Proc. Natl. Acad. Sci. USA **77**, 7315-7317.

Pantazis, P. and Bonner, W.M. (1981). J. Biol. Chem. **256**, 4669-4675.

Pedrali Noy, G.C.F., Spadari, S., Ciarrocchi, G., Pedrini, A.M. and Falaschi, A. (1973). Eur. J. Biochem. **39**, 343-351.

Pergola, F., Zdzienicka, M.Z. and Lieber, M.R. (1993). Mol. Cell. Biol. **13**, 3464-3471.

Petrini, J.H.J., Huwiler, K.G. and Weaver, D.T. (1991). Proc. Natl. Acad. Sci. USA **88**, 7615-7619.

Prelich, G., Tan, C.-K., Kostura, M., Mathews, M.B., So, A.G., Downey, K.M. and Stillman, B. (1987). Nature **326**, 517-520.

Price, A. and Lindahl, T. (1991). Biochemistry **30**, 8631-8637.

Prigent, C., Aoufouchi, S. and Phillipe, M. (1990). Biochem. Biophys. Res. Commun. **169**, 888-895.

Prigent, C., Lasko, D.D., Kodama, K., Woodgett, J.R. and Lindahl, T. (1992). EMBO J. **11**, 2925-2933.

Prigent, C., Satoh, M.S., Daly, G., Barnes, D.E. and Lindahl, T. (1994). Mol. Cell. Biol. **14**, in press.

Roberts, E., Nash, R.A., Robins, P. and Lindahl, T. (1994). J. Biol. Chem. **269**, in press.

Roth, D.B. and Wilson, J.H. (1986). Mol. Cell. Biol. **6**, 4295-4304.

Runger, T.M. and Kraemer, K.H. (1989). EMBO J. **8**, 1419-1425.

Satoh, M.S. and Lindahl, T. (1992). Nature **356**, 356-358.

Satoh, M.S., Poirier, G.G. and Lindahl, T. (1993). J. Biol. Chem. **268**, 5480-5487.

Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J.H.J., Chambon, P. and Egly, J.-M. (1993). Science **260**, 58-63.

Schagger, H. and von Jagow, G. (1987). Anal. Biochem. **166**, 368-379.

Shark, K.B. and Conway, T. (1992). FEMS Microbiol. Lett. **96**, 19-26.
Shinohara, A., Ogawa, H. and Ogawa, T. (1992). Cell **69**, 457-470.

Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K. and Ogawa, T. (1993). Nature Genet. **4**, 239-243.

Shivji, M.K.K., Kenny, M.K. and Wood, R.D. (1992). Cell **69**, 367-374.

Siede, W., Friedberg, A.S. and Friedberg, E.C. (1993). Proc. Natl. Acad. Sci. USA **90**, 7985-7989.

Sipiczki, M., Grossenbacher-Grunder, A.-M. and Bodi, Z. (1990). Mol. Gen. Genet. **220**, 307-313.

Smith, G.L., Chan, Y.S. and Kerr, S.M. (1989). Nucleic Acids Res. **17**, 9051-9062.

So, A.G. and Downey, K.M. (1992). Crit. Rev. Biochem. Mol. Biol. **27**, 129-155.

Soderhall, S. and Lindahl, T. (1973a). J. Biol. Chem. **248**, 672-675.

Soderhall, S. and Lindahl, T. (1973b). Biochem. Biophys. Res. Commun. **53**, 910-916.

Soderhall, S. (1975). Eur. J. Biochem. **51**, 129-136.

Soderhall, S. and Lindahl, T. (1975). J. Biol. Chem. **250**, 8438-8444.

Soderhall, S. (1976). Nature **260**, 640-642.

Squires, S. and Johnson, R.T. (1983). Carcinogenesis **4**, 565-572.

Swank, R.T. and Munkres, K.D. (1971). Anal. Biochem. **39**, 462-477.

Syvaoja, J. and Linn, S. (1989). J. Biol. Chem. **264**, 2489-2497.

Szymkowski, D.E., Yarema, K., Essigman, J.M., Lippard, S.J. and Wood, R.D. (1992). Proc. Natl. Acad. Sci. USA **89**, 10772-10776.

Szymkowski, D.E., Hajibagheri, M.A.N. and Wood, R.D. (1993). J. Mol. Biol. **231**, 251-260.

Takahashi, M. and Senshu, M. (1987). FEBS Lett. **213**, 345-352.

Takahashi, M. and Tomizawa, K. (1990). Eur. J. Biocem. **192**, 735-740.

Teo, I.A., Arlett, C.F., Harcourt, S.A., Priestley, A. and Broughton, B.C. (1983a). Mutat. Res. **107**, 371-386.

Teo, I.A., Broughton, B.C., Day, R.S., James, M.R., Karran, P., Mayne, L.V. and Lehmann, A.R. (1983b). Carcinogenesis **4**, 559-564.

Teraoka, H., Shimoyachi, M. and Tsukada, K. (1975). FEBS Lett. **54**, 217-220.

Teraoka, H., Tamura, S. and Tsukada, K. (1979). Biochim. Biophys. Acta **563**, 535-539.

Teraoka, H., Sumikawa, T. and Tsukada, K. (1986). J. Biol. Chem. **261**, 6888-6892.

Teraoka, H. and Tsukada, K. (1986). Biochim. Biophys. Acta **873**, 297-303.

Thode, S., Schafer, A., Pfeiffer, P. and Vielmetter, W. (1990). Cell **60**, 921-928.

Thorgersen, H.C., Morris, H.R., Rand, K.N. and Gait, M.J. (1985). Eur. J. Biochem. **147**, 325-329.

Tomkinson, A.E., Lasko, D.D., Daly, G. and Lindahl, T. (1990). J. Biol. Chem. **265**, 12611-12617.

Tomkinson, A.E., Roberts, E., Daly, G., Totty, N.F. and Lindahl, T. (1991a). J. Biol. Chem. **266**, 21728-21735.

Tomkinson, A.E., Totty, N.F., Ginsburg, M. and Lindahl, T. (1991b). Proc. Natl. Acad. Sci. USA **88**, 400-404.

Tomkinson, A.E., Tappe, N.J. and Friedberg, E.C. (1992). Biochemistry **31**, 11762-11771.

Tsurimoto, T. and Stillman, B. (1989). Mol. Cell. Biol. **9**, 609-619.

Tsurimoto, T. and Stillman, B. (1991a). J. Biol. Chem. **266**, 1950-1960.

Tsurimoto, T. and Stillman, B. (1991b). J. Biol. Chem. **266**, 1961-1968.

Turunen, O., Winqvist, R., Pakkanen, R., Grzeschik, K.-H., Wahlstrom, T. and Vaheri, A. (1989). J. Biol. Chem. **264**, 16727-16732.

Venema, J., Mullenders, L.H.F., Natarajan, A.T., van Zeeland, A.A. and Mayne, L.V. (1990). Proc. Natl. Acad. Sci. USA **87**, 4707-4711.

Vos, J.-M.H. (1992). Curr. Opin. Cell Biol. **4**, 385-395.

Wang, T.S.-F. (1991). Annu. Rev. Biochem. **60**, 513-552.

Weber, C.A., Salazar, E.P., Stewart, S.A. and Thompson, L.H. (1990). EMBO J. **9**, 1437-1447.

Webster, A.D.B., Barnes, D.E., Arlett, C.F., Lehmann, A.R. and Lindahl, T. (1992). Lancet **339**, 1508-1509.

Weinberg, D.H., Collins, K.L., Simaneck, P., Russo, A., Wold, M.S., Virshup, D.M. and Kelly, T.J. (1990). Proc. Natl. Acad. Sci. USA **87**, 8692-8696.

Weinert, T.A. and Hartwell, L.H. (1988). Science **241**, 317-322.

Weiss, B. and Richardson, C.C. (1967). Proc. Natl. Acad. Sci. USA **57**, 1021-1028.

West, S.C. (1992). Annu. Rev. Biochem. **61**, 603-640.

Wevrick, R. and Buchwald, M. (1993). Curr. Opin. Genet. Dev. **3**, 470-474.

Wilcock, D. and Lane, D.P. (1991). Nature **349**, 429-431.

Willis, A.E. and Lindahl, T. (1987). Nature **325**, 355-357.

Wood, R.D., Robins, P. and Lindahl, T. (1988). Cell **53**, 97-106.

Yang, S.-W., Becker, F.F. and Chan, J.Y. (1990). J. Biol. Chem. **265**, 18130-18134.

Yang, S.-W., Becker, F.F. and Chan, J.Y.H. (1992). Proc. Natl. Acad. Sci. USA **89**, 2227-2231.

Yang, S.-W. and Chan, J.Y.-H. (1992). J. Biol. Chem. **267**, 8117-8122.

Yoshihara, K., Itaya, A., Tanaka, Y., Ohashi, Y., Ito, K., Teraoka, H., Tsukada, K., Matsukage, A. and Kamiya, T. (1985). Biochem. Biophys. Res. Comm. **128**, 61-67.

Zimmerman, S.B. and Levin, C.J. (1975). J. Biol. Chem. **250**, 149-155.

Zimmerman, S.B. and Pheiffer, B.H. (1983). Proc. Natl. Acad. Sci. USA **80**, 5852-5856.

PUBLICATIONS

Tomkinson, A.E., Roberts, E., Daly, G., Totty, N.F. and Lindahl, T. (1991) Three Distinct DNA Ligases in Mammalian Cells. J. Biol. Chem. **266**, 21728-21735.

Roberts, E., Nash, R.A., Robins, P. and Lindahl, T. (1994) Different Active Sites of Mammalian DNA Ligases I and II. J. Biol. Chem. **269**, in press.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 266, No. 32, Issue of November 15, pp. 21728-21735, 1991
Printed in U.S.A.

Three Distinct DNA Ligases in Mammalian Cells*

(Received for publication, July 8, 1991)

Alan E. Tomkinson‡§, Emma Roberts‡, Graham Daly‡, Nicholas F. Totty¶, and Tomas Lindahl‡‖

From the ‡Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Hertfordshire, EN6 3LD, United Kingdom and the ¶Ludwig Institute for Cancer Research, Middlesex Hospital, University College Branch, London, W1P 8BT, United Kingdom

The major DNA ligase of proliferating mammalian cells, DNA ligase I, catalyzes the joining of single strand breaks in double stranded DNA and is active on a synthetic substrate of oligo(dT) hybridized to poly(dA). DNA ligase I does not catalyze the joining of an oligo(dT)·poly(rA) substrate. Two additional DNA ligases, II and III, which can act on the latter substrate have been purified from calf thymus. DNA ligase II, which has been described previously, is a 72-kDa protein. DNA ligase III migrates as a 100-kDa protein in denaturing gel electrophoresis. Structural, immunochemical, and catalytic studies on the three DNA ligase activities strongly indicate that they are the products of three different genes.

DNA ligases catalyze the formation of phosphodiester bonds at single strand breaks in double stranded DNA. Genetic studies in *Escherichia coli* and the yeasts, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, have demonstrated that this enzyme is essential for DNA replication and is also required for DNA repair and genetic recombination (1–4). Microorganisms appear to have only one species of DNA ligase, although this has been more firmly established for bacteria than for yeasts. In contrast, two distinct DNA ligases, designated DNA ligase I and DNA ligase II, have been found in extracts from mammalian tissues and cells (5–9) and from *Drosophila melanogaster* cells (10, 11). A human cDNA encoding the 102-kDa DNA ligase I has been cloned and sequenced (12), and the biochemical properties of mammalian DNA ligase I have been extensively characterized (5, 13–17). DNA ligase II has been investigated to a lesser extent. The latter enzyme was initially detected as a minor activity that adsorbed more firmly to hydroxylapatite than DNA ligase I (5, 7). Subsequent studies have demonstrated that these two nuclear enzymes are of different size, are immunologically unrelated, and have somewhat different catalytic properties, including the unique ability of DNA ligase II to ligate oligo(dT) molecules hybridized to poly(rA) (6, 7, 9, 18).

Rapidly dividing cells contain higher levels of DNA ligase I activity than nonproliferating cells, indicating a role for this enzyme in DNA replication (14, 19, 20). Complementation of the conditional lethal phenotype of a *S. cerevisiae cdc9* DNA ligase mutant by human DNA ligase I cDNA demonstrates that this enzyme can function in DNA replication (12). In contrast, the level of DNA ligase II activity is independent of the proliferative state of the cell (19, 20). It has been reported that the level of DNA ligase II is induced by DNA damage (21, 22), suggesting that DNA ligase II may be involved in at least some forms of DNA repair. In addition, the regulation of DNA ligase II activity during spermatogenesis is consistent with this enzyme being involved in meiotic recombination (9).

Since DNA ligase I does not ligate an oligo(dT)·poly(rA) substrate (7), we have used this specific assay to purify DNA ligase II from calf thymus extracts. During the course of these studies, we detected two minor DNA ligase activities with different chromatographic properties that catalyze joining of the oligo(dT)·poly(rA) substrate. The biochemical properties of one of these enzymes were identical to those described previously for DNA ligase II (5, 6, 9, 18). However, the chromatographic, physical, and catalytic properties of the other activity, designated DNA ligase III, were clearly different from those of either DNA ligase I or II.

EXPERIMENTAL PROCEDURES[1]

RESULTS

Purification and Size of DNA Ligase II—Details of the DNA ligase II purification scheme are described in the Miniprint Section. The peak fractions of DNA ligase II activity (Table IS, Fraction VIII) eluting from the final FPLC[2] Mono S column apparently contained a single 72-kDa polypeptide detected by Coomassie Blue staining (Fig. 1) or by silver staining (data not shown) following SDS-polyacrylamide gel electrophoresis (SDS-PAGE). After incubation of DNA ligase II (Table 1S, Fraction VIII) with [α-^{32}P]ATP, a radioactively labeled 72-kDa polypeptide was detected by autoradiography following denaturing gel electrophoresis. The co-migration of the 72-kDa polypeptides, detected by protein staining (Fig. 1) and by autoradiography (see below, Fig. 4), was confirmed by staining the adenylylated polypeptide with Coomassie Blue after SDS-PAGE before slicing the gel and measuring the radioactive material in the gel slices by liquid scintillation counting. A relatively low yield of DNA ligase II activity in the latter purification steps reflects the lability of this enzyme during purification. Stability was improved by the addition of 0.2% Tween 20 to all buffers employed after the hydroxylapatite chromatography step.

Partial Amino Acid Sequencing of Purified DNA Ligase II—A single N-terminal sequence of 20 amino acids from 20 μg

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: Dept. of Pathology, University of Texas, Southwestern Medical Center, Dallas, TX 75235-9072.

‖ To whom correspondence should be addressed. Fax: 44-707-46332.

[1] Portions of this paper (including "Experimental Procedures," Tables 1S and 2S, and Figs 1S and 2S) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

[2] The abbreviations used are: FPLC, fast protein liquid chromatography; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

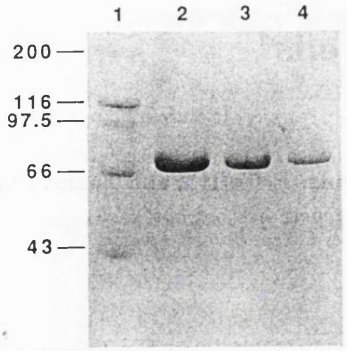


FIG. 1. **Size of DNA ligase II from calf thymus as measured by SDS-polyacrylamide gel electrophoresis.** Purified DNA ligase II protein (Table 1S, Fraction VIII) was electrophoresed through a 10% SDS-polyacrylamide gel. Proteins were stained using Coomassie Brilliant Blue (Sigma). *Lane 1,* molecular mass standards (Bio-Rad); *lane 2,* 1 μg of protein (Fraction VIII); *lane 3,* 500 ng of protein; *lane 4,* 250 ng of protein.

FIG. 2. **Two-dimensional gel analysis of purified bovine DNA ligase II.** DNA ligase II (Fraction VIII, 500 ng) was supplemented with 300 ng of bovine serum albumin as an internal marker and electrophoresed through a 5% polyacrylamide isoelectric focusing rod gel and then separated in the second dimension by electrophoresis through a 10% SDS-polyacrylamide slab gel. *A,* proteins detected by staining with Coomassie Brilliant Blue. The position of bovine serum albumin was confirmed in a separate experiment. *B,* immunological detection of ezrin/cytovillin. Proteins were transferred from the SDS-polyacrylamide gel onto nitrocellulose. The antigen recognized by a polyclonal antiserum directed against the N-terminal amino acid sequence of ezrin/cytovillin was detected using a secondary antibody coupled to alkaline phosphatase. *BSA,* bovine serum albumin.

of two different preparations of DNA ligase II (Table 1S, Fraction VIII) was obtained by microsequencing. Unexpectedly, this sequence was homologous to the N terminus of the microvillar cytoplasmic protein, ezrin (23). Ezrin is closely related or identical to the mammalian protein, cytovillin (24), with reported differences confined to the N termini of the proteins.[3] The DNA ligase II preparation (30 μg, Fraction VIII) was digested with the reagent proteolytic enzyme, lysylendopeptidase. After separation by high pressure liquid chromatography, the N-terminal amino acid sequences of 14 different peptides were determined by microsequencing (25). Ten of these peptide sequences were homologous to stretches of amino acid sequences distributed throughout ezrin/cytovillin. The amino acid sequences of the remaining four peptides, which were present in similar molar amounts to the other peptides, were not detectably homologous to ezrin/cytovillin or to any other mammalian protein sequence in the data base. In particular, none of these peptides showed detectable homology with the predicted amino acid sequence of DNA ligase I (12). The results indicate that the material in the 72-kDa band (Fig. 1) represents an approximately equimolar mixture of two proteins, the 69-kDa (as predicted from the cDNA sequence) ezrin/cytovillin and the 72-kDa DNA ligase II. The major purification step for ezrin is hydroxylapatite chromatography since this protein adsorbs unusually strongly (26), but this property is shared with DNA ligase II and consequently the complete separation of these two proteins becomes difficult. The presence of a single N-terminal sequence in Fraction VIII, corresponding to ezrin/cytovillin, strongly indicates that the N terminus of DNA ligase II is blocked, as was found previously for DNA ligase I (15).

Analysis of Purified DNA Ligase II by Two-dimensional Gel Electrophoresis—After fractionation of the DNA ligase II preparation (Table 1S, Fraction VIII) by isoelectric focusing followed by SDS-PAGE, two 72-kDa polypeptides with different isoelectric points were detected in approximately equal amounts, as predicted from the peptide sequencing experiments (Fig. 2*A*). The polypeptide closest to the bovine serum albumin marker was identified as bovine ezrin/cytovillin by immunoblotting with antibodies against the N-terminal sequence of ezrin/cytovillin (Fig. 2*B*). The second polypeptide was tentatively assigned as DNA ligase II. This confirms that the final preparation of DNA ligase II was about 50% pure, the residual contaminating protein being ezrin/cytovillin, an abundant structural cellular protein.

Purification and Size of DNA Ligase III—During the isolation of DNA ligase I from calf thymus (15), an enzyme activity which could join an oligo(dT)·poly(rA) substrate was observed to elute from the AcA 34 gel filtration column immediately before DNA ligase I (Fig. 1S, Miniprint Section). This high molecular weight activity, exhibiting a Stokes radius of approximately 58 Å, might represent DNA ligase II complexed to another protein or a separate DNA ligase activity. Purification and characterization of the activity showed that it was clearly distinct from DNA ligase II. Details of the purification scheme for this high molecular weight activity, DNA ligase III, are described in the Miniprint Section (Table 2S). After following the same initial purification steps used for DNA ligases I and II, it was observed that DNA ligase III precipitated at lower ammonium sulfate concentrations compared to DNA ligases I and II. Using such a fraction, DNA ligase III activity was found to co-elute on gel filtration (Fig. 1S, Miniprint Section) with a 100-kDa polypeptide that formed a covalent enzyme-adenylate intermediate (data not shown). During the later stages of the purification, DNA ligase III fractions contained 100- and 87-kDa polypeptides that formed enzyme-adenylate complexes (Fig. 4). The relative amounts of the 87-kDa polypeptide increased with time, suggesting that it was a proteolytic fragment of the 100-kDa polypeptide (Fig. 5). Active fractions were pooled to specifically purify the larger polypeptide. In the final purification step, DNA ligase III activity (Table 2S, Fraction VIII) was eluted from a Mono S column at 0.3 M NaCl (Fig. 3*A*). Analysis of the active fractions by SDS-PAGE demonstrated that a polypeptide of 100 kDa co-purified with enzyme activity (Fig. 3*B*). In different preparations of DNA ligase III, a 46-kDa polypeptide was also present in the active fractions (Fig. 3B). When DNA ligase III (Table 2S, Fraction VIII) was incubated with [α-^{32}P]ATP, a 100-kDa adenylylated polypeptide was detected (data not shown). The data indicate that DNA ligase III has a molecular mass of 100 kDa as determined by SDS-PAGE and that the preparation was at least 50–70% pure. This enzyme was also partially purified from the human B cell line GM1953, so DNA ligase III is not a thymus- or bovine-specific enzyme. Sucrose gradient centrifugation of the DNA ligase III activity yielded a sedimentation coefficient of 5.5 S. DNA ligase III has a native molecular mass of 147

[3] Inspection of the DNA sequence of cytovillin 5′ to the predicted site of translation initiation reveals an in-frame amino acid sequence starting with a methionine, which is identical to the N terminus of ezrin. This suggests that the tentative translational start site of cytovillin was originally misassigned.

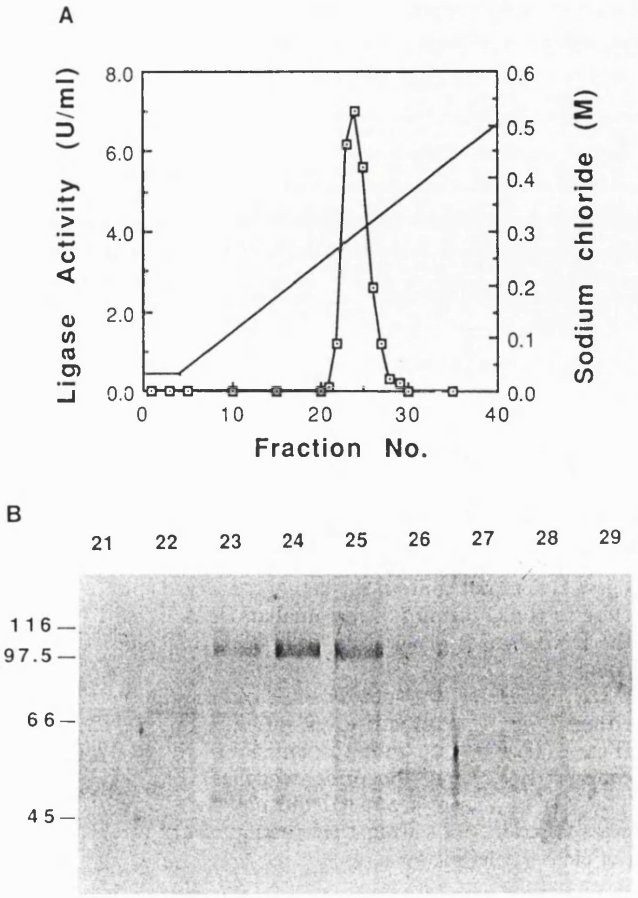


FIG. 3. **Purification of bovine DNA ligase III by FPLC Mono-S chromatography.** Fractions containing DNA ligase activity following phenyl-Sepharose chromatography (Table 2S, Fraction VII) were loaded by a single injection of 4 ml onto a FPLC Mono-S column, and bound proteins were gradient eluted. *A*, ligation activity with an oligo(dT)·poly(rA) substrate (⊡). *B*, aliquots (20 μl) of fractions that contained ligase activity were separated by SDS-polyacrylamide gel electrophoresis and proteins were detected by silver staining (Bio-Rad). The molecular weight standards were as for Fig. 1.

kDa, calculated from the sedimentation coefficient and Stokes radius values (27). This result indicates that the 100- and 46-kDa polypeptides seen on silver staining (Fig. 3*B*) may be associated with the 100-kDa polypeptide containing the AMP-binding domain. Further studies are required to establish whether the 46-kDa polypeptide is a functionally significant component of the enzyme.

Reactivity of the Enzyme-Adenylate Intermediates of Bovine DNA Ligases I, II, and III—In the first step of the ligation reaction, eukaryotic and viral DNA ligases interact with ATP to form a covalent enzyme-AMP intermediate (4). The sizes of the labeled reaction intermediates can be estimated by denaturing gel electrophoresis after incubation with [α-^{32}P] ATP. The adenylate group can be specifically removed from the complex either by transfer to a 5′-phosphate acceptor of a polynucleotide substrate or by incubation with pyrophosphate, reversing the initial reaction to release ATP.

The 102-kDa DNA ligase I migrates anomalously slowly during SDS-PAGE, probably due to the high proline content of the protein (12). Under these conditions it has an apparent molecular mass of 125 kDa (15). The DNA ligase I-adenylate intermediate transferred the AMP group to an oligo(dT)·poly(dA) substrate but failed to react with an oligo(dT)·poly(rA) substrate (Fig. 4*A*, *lanes 5–7*), whereas the 72-kDa DNA ligase II-adenylate intermediate reacted with both substrates (Fig. 4*A*, *lanes 1–3*). The adenylylation reactions of DNA ligases I and II were reversed by incubation with an excess of pyrophosphate (Fig. 4*A*, *lanes 4* and *8*). These results are in agreement with our previous observation that DNA ligase II, but not DNA ligase I, can ligate oligo(dT) molecules hybridized to poly(rA) (7). A DNA ligase III·AMP complex, when assayed in the same way, was dissociated by both oligo(dT)·poly(dA) and oligo(dT)·poly(rA) (Fig. 4*B*, *lanes 7–9*). Thus, DNA ligase III was similar to DNA ligase II in this regard, but clearly different from DNA ligase I.

During the purification of DNA ligase I, the intact form of the enzyme was partially proteolytically cleaved to generate a catalytically active 85-kDa fragment (15). The 85-kDa poly-

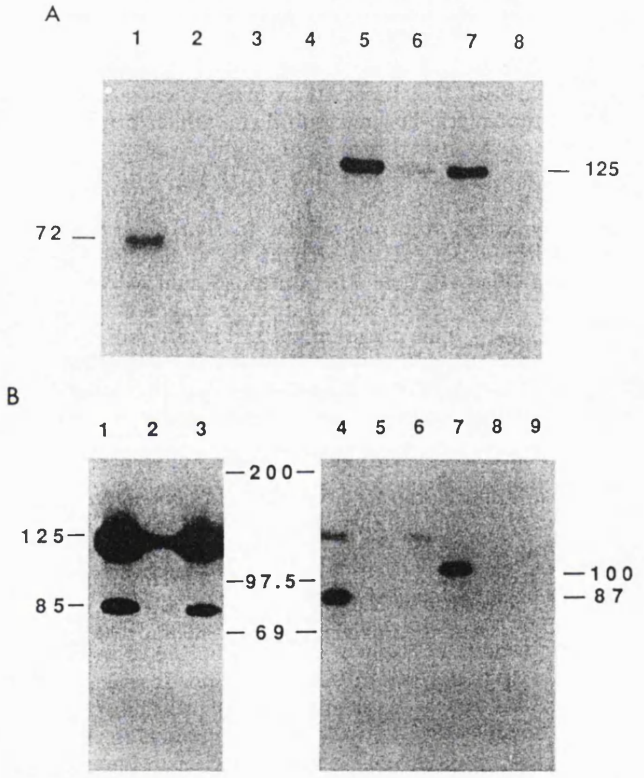


FIG. 4. **Reactivity of the adenylylated forms of bovine DNA ligases I, II, and III.** DNA ligase-adenylate intermediates were formed and then further incubated with unlabeled polynucleotide substrate or sodium pyrophosphate as described under "Experimental Procedures." *A*, *lanes 1–4*, DNA ligase II-adenylate complex (Table 1S, Fraction V, 500 ng) and *lanes 5–8*, DNA ligase I-adenylate complex (Fraction VII (15), 20 ng) incubated with: *lanes 1* and *5*, no addition; *lanes 2* and *6*, poly(dA)·oligo(dT); *lanes 3* and *7*, poly(rA)·oligo(dT); and *lanes 4* and *8*, sodium pyrophosphate. Reactions were stopped by the addition of SDS sample buffer. Polypeptide-adenylate complexes were separated by electrophoresis through a 10% SDS-polyacrylamide gel, which was fixed in 10% acetic acid, dried, and exposed to x-ray film. *B*, polypeptide-adenylate complexes formed by fractions eluting from the phosphocellulose column (Fig. 2S, Miniprint Section) during purification of DNA ligase III. *Lanes 1–3*, Fraction 21 (DNA ligase I, intact enzyme and 85-kDa active fragment); *lanes 4–6*, Fraction 30 (containing both DNA ligases I and III); *lanes 7–9*, Fraction 40 (DNA ligase III, Fraction V, Table 2S); *lanes 1, 4,* and *7*, no addition; *lanes 2, 5,* and *8*, oligo(dT)·poly(dA), and *lanes 3, 6,* and *9*, oligo(dT)·poly(rA). Reactions were stopped by the addition of SDS sample buffer and, after electrophoresis through a 7.5% SDS-polyacrylamide gel, polypeptide-adenylate complexes were detected by autoradiography as described above. ^{14}C-Labeled size markers (Amersham Corp.) were electrophoresed on the same gels.

peptide-adenylate complex retained the same polynucleotide substrate specificity as the reaction intermediate formed by the intact enzyme (Fig. 4*B*, *lanes 1–3*). A similar partial degradation by endogenous proteases occurred during the purification of DNA ligase III, *i.e.* the 100-kDa polypeptide was cleaved to an 87-kDa fragment which retained catalytic activity (data not shown). The 87-kDa adenylylated polypeptide retained the same polynucleotide substrate specificity as the reaction intermediate formed by the 100-kDa polypeptide in that both transferred the AMP group to the oligo(dT)·poly(rA) and the oligo(dT)·poly(dA) substrates (Fig. 4*B*, *lanes 4–9*). No detectable cleavage of the 72-kDa DNA ligase II to active fragments occurred during purification.

DNA ligase I and DNA ligase III exhibited similar K_m values for ATP in the enzyme adenylylation reaction. DNA ligase III has a K_m of 1–2 μM (data not shown) compared to 0.6 μM for DNA ligase I (15). In contrast, the K_m for DNA ligase II was considerably higher, about 10–20 μM.

Partial Proteolysis of Adenylylated DNA Ligase III—Digestion of purified DNA ligase III by low concentrations of the reagent proteolytic enzyme, subtilisin, which preferentially cleaves non-structured regions of proteins separating folded domains was performed. On SDS-PAGE (28) a discontinuous conversion of the 100-kDa adenylylated polypeptide to an 87-kDa polypeptide occurred, similar to that observed during partial degradation of the enzyme by endogenous cellular proteases (Fig. 4*B*, *lane 7*). Continued subtilisin digestion subsequently generated smaller discrete fragments of 60 and 38 kDa (Fig. 5). This digestion pattern is different from that observed for the adenylylated form of DNA ligase I (15), indicating that DNA ligase I and III are two different proteins. Furthermore, digestion of adenylylated DNA ligase III did not generate a fragment of 72 kDa, similar in size to DNA ligase II.

Polynucleotide Substrate Specificity of Bovine DNA Ligases I, II, and III—The abilities of bovine DNA ligases I, II, and III to catalyze ligation of several different homopolymer polynucleotide substrates were investigated in preliminary experiments. The results with 5′-^{32}P-labeled oligo(rA) hybridized to poly(dT) seemed of particular interest, and the reaction was studied in more detail. Reaction products were separated by denaturing polyacrylamide gel electrophoresis to determine whether authentic ligation products were being generated. In agreement with previous results, all three DNA ligases could join oligo(dT) molecules hybridized to a poly(dA) template (Fig. 6, *lanes 1–3*), whereas only DNA ligases II and III could ligate oligo(dT) molecules hybridized to a poly(rA) template. Surprisingly, DNA ligase I and III, but not DNA ligase II, catalyzed ligation of an oligo(rA)·poly(dT) substrate (Fig. 6, *lanes 9–11*). Therefore, the three ligases can be distinguished by their polynucleotide substrate specificity. In addition to acting on these synthetic substrates, all three DNA ligases could rejoin *Eco*RI restriction enzyme fragments of DNA (data not shown).

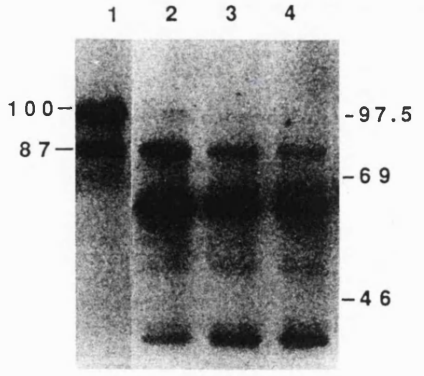


FIG. 5. **Partial subtilisin digestion of bovine DNA ligase III-AMP complex.** DNA ligase III (Table 2S, Fraction V, 800 ng) was incubated with 10 μCi of [α-^{32}P]ATP in the presence of 60 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM dithiothreitol, 50 μg/ml of bovine serum albumin in a final volume of 40 μl for 15 min at room temperature. The reaction was stopped by the addition of 2 μl of 0.5 M EDTA. Protein-adenylate (200 ng) was incubated with subtilisin for 15 min as follows; *lane 1*, no addition; *lane 2*, 2 ng of subtilisin; *lane 3*, 5 ng; *lane 4*, 10 ng. Reactions were stopped by the addition of 3.8 μg of aprotinin and 5 μl of SDS sample buffer. Polypeptide-adenylate complexes were detected by autoradiography after separation through a 12% SDS-polyacrylamide gel. Size markers were as for Fig. 4.

DNA Ligases II and III Are Not Immunologically Related to DNA Ligase I—It has been shown previously that bovine DNA ligase I and II are serologically distinct proteins (6, 18). This observation was confirmed by demonstrating that a polyclonal antiserum raised against homogeneous bovine DNA ligase I (15) does not detect homogeneous bovine DNA ligase II on immunoblotting (Fig. 7*A*, *lane 2*). Moreover, DNA ligase II does not contain the highly conserved epitope located close to the C termini of mammalian DNA ligase I (15) and the DNA ligases of vaccinia virus (29), *S. cerevisiae* and *S. pombe* (Fig. 7*B*, *lane 2*).

DNA ligase III was not detected by immunoblotting with either the polyclonal antiserum raised against DNA ligase I (Fig. 7*A*, *lane 3*) or with the antiserum raised against the conserved C-terminal peptide sequence (Fig. 7*B*, *lane 3*). Thus, DNA ligases I and III did not show any serological cross-reaction, indicating that their primary structures are not closely related.

DISCUSSION

Two distinct DNA ligases which differ from the major enzyme of this type present in proliferating mammalian cells, DNA ligase I, have been found in extracts of calf thymus and human lymphoid B-cell lines. The identification of these enzymes was greatly facilitated by their ability to join oligo(dT) molecules hybridized to poly(rA), a reaction not catalyzed by DNA ligase I (7, 30). One of the enzymes is identical with the previously described DNA ligase II, an enzyme characterized in several laboratories (5–9, 18, 20, 31). The other minor activity, which is clearly separate from DNA ligase II, has been termed DNA ligase III.

DNA ligase II is a labile enzyme present in cell nuclei which has been purified from calf thymus to apparent homogeneity, as estimated by SDS-PAGE, by three different groups using similar methods. The molecular mass of the protein has been observed to be about 68 (18), 65 (9), or 72 kDa (this work). Since different size markers were employed in each case, these numbers are probably identical within experimental error. As reported here, such a purified preparation exhibited a unique N-terminal sequence. Nevertheless, these results appear somewhat misleading because a more detailed investigation by isoelectric focusing, and sequencing of Lys-C peptides, revealed that our most purified preparation of DNA ligase II was contaminated with a similar amount of the abundant structural protein, ezrin/cytovillin, which is the same size as DNA ligase II. In spite of this complication, sequences of several peptides of DNA ligase II have been obtained and are being used as probes to isolate cDNAs by hybridization.

The N-terminal sequence of DNA ligase II is apparently blocked to Edman degradation, in common with about 80% of the soluble proteins in mammalian cell extracts (32). The N-terminal sequence detected was that reported for ezrin (23).

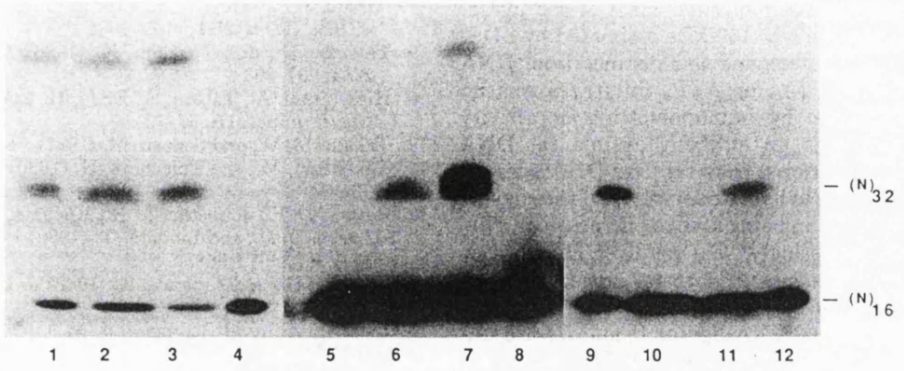


FIG. 6. **Analysis of ligation products formed by bovine DNA ligases I, II, and III.** Equal amounts of bovine DNA ligases I, II, and III as measured by activity on the oligo(dT)·poly(dA) substrate were incubated for 15 min at 16 °C with different labeled polynucleotide substrates. *Lanes 1–4,* oligo(dT)·poly(dA); *lanes 5–8,* oligo(dT)·poly(rA); *lanes 9–12,* oligo(rA)·poly(dT); incubated with, *lanes 1, 5,* and *9,* DNA ligase I (Fraction VII (15)); *lanes 2, 6,* and *10,* DNA ligase II (Table 1S, Fraction VIII); *lanes 3, 7,* and *11,* DNA ligase III (Table 2S, Fraction VI); *lanes 4, 8,* and *12,* no enzyme. Oligonucleotides were detected by autoradiography after electrophoresis through a 20% denaturing polyacrylamide gel.

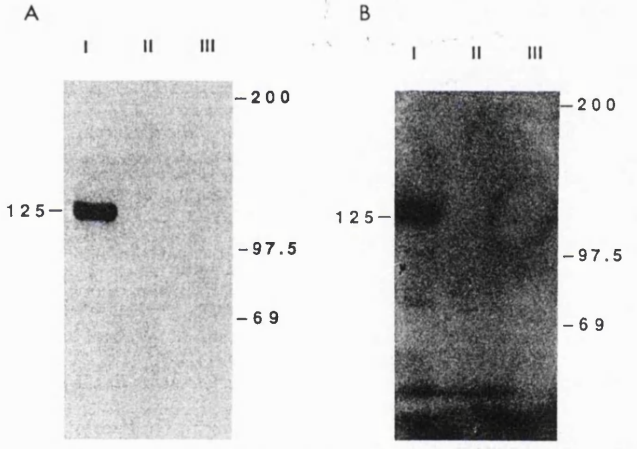


FIG. 7. **Antisera raised against bovine DNA ligase I or against a conserved peptide common to several eukaryotic DNA ligases do not cross-react with bovine DNA ligase II or DNA ligase III.** Fractions containing DNA ligase I (200 ng, Fraction VII, (15), DNA ligase II (Table 1S, 200 ng, Fraction VIII), and DNA ligase III (Table 2S, 400 ng, Fraction VIII) were separated by SDS-PAGE with a 4–15% gradient of acrylamide (Bio-Rad). Under these conditions, bands of approximately equal intensity are detected after Coomassie Blue staining with relative molecular masses of: *lane I,* 125 kDa, DNA ligase I; *lane II,* 72 kDa, DNA ligase II; and *lane III,* 100 kDa, DNA ligase III. Proteins were separated under identical conditions on two more gels and transferred to nitrocellulose membranes. *A,* nitrocellulose membrane incubated with antiserum (diluted 1 in 100) raised against bovine DNA ligase I. *B,* nitrocellulose membrane incubated with antiserum (diluted 1 in 400) raised against a peptide common to several eukaryotic DNA ligases. Immune complexes were detected by incubation of the membranes with ^{125}I-labeled protein A.

The presence of a blocked N terminus in DNA ligase II together with the lack of immunological cross-reactivity between DNA ligases I and II, strongly indicate that the 65–72-kDa form of DNA ligase II is a primary translation product and not a proteolytic fragment of DNA ligase I, or a product of alternative splicing of the transcript of the DNA ligase I gene. Teraoka *et al.* (18) arrived at a similar conclusion in an immunochemical study by using a neutralizing antibody against DNA ligase II which did not recognize DNA ligase I, and showing that DNA ligase II appears to remain the same size in crude cell extracts and throughout enzyme purification. Other properties of DNA ligase II which distinguish this enzyme from DNA ligase I include a firm association with cell nuclei (DNA ligase I leaches out rapidly in isotonic sucrose solutions), a high K_m for ATP, and a lack of induction on cell proliferation (9, 18, 33).

DNA ligase III is clearly dissimilar from either DNA ligases I or II with regard to its physical and catalytic properties, and we conclude that it is a third mammalian enzyme of this type. Polyclonal antibodies directed against either DNA ligase I, or an evolutionarily conserved peptide within the DNA ligase I sequence, failed to detect DNA ligase III. Moreover, partial degradation of DNA ligases I and III with subtilisin yielded different polypeptide patterns, so it seems unlikely that the 100-kDa DNA ligase III detected by SDS-gel electrophoresis could be a derivative or active fragment of DNA ligase I (28). As mentioned above, there is strong evidence that the 65–72-kDa DNA ligase II is a primary translation product rather than being a fragment derived from a larger precursor protein and there is no indication of a direct relationship between DNA ligases II and III. Partial proteolysis *in vitro* of DNA ligase III did not yield a fragment the size of DNA ligase II. However, it is unknown whether DNA ligases II and III would cross-react immunologically, because suitable antisera are not yet available. DNA ligase III has a low K_m for ATP and does not bind strongly to hydroxylapatite; in these respects it resembles DNA ligase I but is clearly different from DNA ligase II.

A DNA ligase activity with the properties of DNA ligase III has been observed independently by Elder and Rossignol (14) in rat liver extracts. Since a number of reports had appeared claiming that only a single DNA ligase might be present in mammalian cells (13, 34), these workers used the oligo(dT)·poly(rA) assay to investigate whether an activity distinct from DNA ligase I could be detected in postmicrosomal supernatants of rat liver. It is noteworthy that the isotonic sucrose buffer employed for extraction would not have released DNA ligase II in soluble form from cell nuclei (9, 20, 33). Nevertheless, Elder and Rossignol (14) observed and partially purified two distinct DNA ligases, one of which was identical with DNA ligase I. The other activity was proposed to be "a larger more native form of DNA ligase II" than the enzyme studied in other laboratories. We believe this interpretation to be in error, but the rat liver enzyme detected is similar to the calf thymus DNA ligase III described here in

the following respects: (i) the molecular mass of the ligase-AMP reaction intermediate was 100 kDa estimated by SDS-gel electrophoresis, (ii) the enzyme was distinct from DNA ligase II in having a low K_m for ATP, (iii) there was no recognition of the protein by immunoblotting or activity neutralization assays using an antibody against rat DNA ligase I, (iv) the enzyme differed from rat liver DNA ligase II (19) in binding only weakly to hydroxylapatite, (v) the enzyme differed from DNA ligase I in being able to join an oligo(dT)·poly(rA) substrate. Interestingly, this rat liver DNA ligase was present in similar amounts in normal and regenerating liver and differed from DNA ligase I in that regard (14).

The physiological roles of DNA ligases II and III are unknown. Several authors have suggested that DNA ligase II could be involved in DNA repair, but there is no direct experimental evidence for this notion. Mammalian DNA ligase I shows strong sequence homology to the *cdc9* gene product of *S. cerevisiae* (12), which is a DNA ligase active in replication, constitutive excision repair, and mitotic recombination (2, 3). The human DNA ligase I cDNA can complement a *S. cerevisiae cdc9* mutant with regard to all these phenotypes,[4] indicating that mammalian DNA ligase I performs similar functions as the *cdc9* gene product. This would relegate DNA ligases II and III to more specialized cellular functions. Neither enzyme is exclusively a mitochondrial activity (9, 14).[5] Higashitani *et al.* (9) followed DNA ligase II levels during spermatogenesis and speculated that the enzyme might catalyze the final step in meiotic recombination. DNA ligases II or III might be involved in the joining of three- or four-stranded recombination intermediates (35), in inducible repair, or in illegitimate reactions such as the joining of protruding single strands in DNA (36, 37). The ability of DNA ligases II and III to act on the oligo(dT)·poly(rA) substrate might point to a role in ligation of strand breaks during transcription or in other processes where DNA/RNA hybrids occur. Cloning and analysis of cDNAs encoding DNA ligases II and III, and detection of analogous enzymes in organisms that allow detailed genetic analysis, *e.g.* yeast, should facilitate functional investigations.

Acknowledgments—We thank Michelle Ginsburg for performing some of the database searches and Deborah Barnes for valuable discussions.

REFERENCES

1. Johnston, L. H., and Nasmyth, K. A. (1978) *Nature* **247,** 891–893
2. Fabre, F., and Roman, H. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76,** 4586–4588
3. Johnston, L. H. (1979) *Mol. Gen. Genet.* **170,** 89–92
4. Engler, M. J., and Richardson, C. C. (1982) in *The Enzymes* (Boyer, P. D., ed) 3rd Ed., 15 B, pp. 1–29, Academic Press, New York
5. Soderhall, S., and Lindahl, T. (1973) *Biochem. Biophys. Res. Commun.* **53,** 910–916
6. Soderhall, S., and Lindahl, T. (1975) *J. Biol. Chem.* **250,** 8438–8444
7. Arrand, J. E., Willis, A. E., Goldsmith, I., and Lindahl, T. (1986) *J. Biol. Chem.* **261,** 9079–9082
8. Teraoka, H., and Tsukada, K. (1986) *Biochim. Biophys. Acta* **873,** 297–303
9. Higashitani, A., Tabata, S., Endo, H., and Hotta, Y. (1990) *Cell Struct. Funct.* **15,** 67–72
10. Takahashi, M., and Senshu, M. (1987) *FEBS Lett.* **213,** 345–352
11. Takahashi, M., and Tomizawa, K. (1990) *Eur. J. Biochem.* **192,** 735–740
12. Barnes, D. E., Johnston, L. H., Kodama, K., Tomkinson, A. E., Lasko, D. D., and Lindahl, T. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 6679–6683
13. Teraoka, H., and Tsukada, K. (1982) *J. Biol. Chem.* **257,** 4758–4763
14. Elder, R. H., and Rossignol, J.-M. (1990) *Biochem.* **29,** 6009–6017
15. Tomkinson, A. E., Lasko, D. D., Daly, G., and Lindahl, T. (1990) *J. Biol. Chem.* **265,** 12611–12617
16. Lasko, D. D., Tomkinson, A. E., and Lindahl, T. (1990) *J. Biol. Chem.* **265,** 12618–12622
17. Tomkinson, A. E., Totty, N. F., Ginsburg, M., and Lindahl, T. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 400–404
18. Teraoka, H., Sumikawa, T., and Tsukada, K. (1986) *J. Biol. Chem.* **261,** 6888–6892
19. Soderhall, S. (1976) *Nature* **260,** 640–642
20. Chan, J. Y. H., and Becker, F. F. (1985) *Carcinogenesis* **6,** 1275–1277
21. Creissen, D., and Shall, S. (1982) *Nature* **296,** 271–272
22. Li, J.-H., and Rossman, T. G. (1989) *Mol. Toxicol.* **2,** 1–9
23. Gould, K. L., Bretscher, A., Esch, F. S., and Hunter, T. (1989) *EMBO J.* **8,** 4133–4142
24. Turunen, O., Winqvist, R., Pakkanen, R., Grzeschik, K-H., Wahlstrom, T., and Vaheri, A. (1989) *J. Biol. Chem.* **264,** 16727–16732
25. Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N. F., Stockwell, P., Ullrich, A., Schlessinger, J., and Waterfield, M. D. (1984) *Nature* **307,** 521–527
26. Bretscher, A. (1983) *J. Cell Biol.* **97,** 425–432
27. Siegel, L. M., and Monty, K. J. (1966) *Biochim. Biophys. Acta* **112,** 346–362
28. Cleveland, D. W., Fischer, S. G., Kirschner, M. W., and Laemmli, U. K. (1977) *J. Biol. Chem.* **252,** 1102–1106
29. Kerr, S. M., and Smith, G. L. (1989) *Nucleic Acids Res.* **17,** 9039–9050
30. Pedrali-Noy, G. C., Spadari, S., Ciarrocchi, G., Pedrini, A. M., and Falaschi, A. (1973) *Eur. J. Biochem.* **39,** 343–351
31. Chan, J. Y., Thompson, L. H., and Becker, F. F. (1984) *Mutat. Res.* **131,** 209–214
32. Brown, J. L., and Roberts, W. K. (1976) *J. Biol. Chem.* **251,** 1009–1014
33. Soderhall, S., and Lindahl, T. (1976) *FEBS Lett.* **67,** 1–8
34. Mezzina, M., Rossignol, J. M., Philippe, M., Izzo, R., Bertazzoni, U., and Sarasin, A. (1987) *Eur. J. Biochem.* **162,** 325–332
35. Rao, B. J., Dutreix, M., and Radding, C. M. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 2984-2988
36. Roth, D. B., and Wilson, J. H. (1986) *Mol. Cell. Biol.* **6,** 4295–4304
37. Thode, S., Schafer, A., Pfeiffer, P., and Vielmetter, W. (1990) *Cell* **60,** 921–928
38. Prigent, C., Aoufouchi, S., and Philippe, M. (1990) *Biochem. Biophys. Res. Commun.* **169,** 888–895
39. Laemmli, U. K. (1970) *Nature* **227,** 680–685
40. Kawasaki, H., and Suzuki, K. (1990) *Anal. Biochem.* **186,** 264–268
41. Harlow, E., and Lane, D. (1988) *Antibodies,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
42. O'Farrell, P. H. (1975) *J. Biol. Chem.* **250,** 4007–4021
43. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254

[4] L. Johnston, personal communication.
[5] A. E. Tomkinson and T. Lindahl, unpublished data.

Supplementary Material to:
Three distinct DNA ligases in mammalian cells

Alan E. Tomkinson, Emma Roberts, Graham Daly, Nicholas F. Totty and Tomas Lindahl

EXPERIMENTAL PROCEDURES

Cells and Tissues - Calf thymus glands were obtained at the local abattoir. The tissues were packed in ice and used for initial enzyme preparation within 3 h.

A human lymphoblastoid cell line derived from a healthy individual, GM1953, was obtained from the Human Genetic Mutant Cell Repository (Camden, New Jersey). These cells were grown in suspension culture at 37°C in RPMI 1640 medium supplemented with 15% fetal bovine serum.

Preparation of polynucleotide substrates- Oligo $(dT)_{30}$ was synthesized on a commercial DNA synthesizer. Oligonucleotides $(dT)_{16}$ and $(rA)_{12-18}$, and polynucleotides (dA), (dT) and (rA) were from P-L Biochemicals. Oligo $(rA)_{16}$ was purified from oligo $(rA)_{12-18}$ by gel purification on a denaturing 20% polyacrylamide gel. If necessary, oligonucleotides (10 µg) were dephosphorylated by calf intestinal phosphatase (Boehringer Mannheim) prior to radioactive labelling with 100 µCi [γ-^{32}P] ATP (>5,000 Ci/mmol, Amersham) and 10 µmol unlabelled ATP using T4 polynucleotide kinase (P-L Biochemicals). The labelled oligonucleotide was mixed with an equimolar amount of polynucleotide, incubated at 90°C for 10 min and then slowly cooled to room temperature. Non-radioactively labelled substrates were prepared as described above with the omission of the [γ-^{32}P]ATP.

DNA ligase assays - Reaction mixtures (60 µl) contained 60 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM DTT, 1 mM ATP, 50 µg/ml nuclease-free bovine serum albumin, polynucleotide substrate (20,000 cpm) and a limiting amount of DNA ligase. Incubations were at 16°C for 15 min.The conversion of 5'-^{32}P-labelled phosphomonoesters to alkaline phosphatase-resistant diesters was measured (16, 38). One unit of DNA ligase activity catalyses the conversion of 1 pmol terminal phosphate residues to a phosphatase resistant form in 15 min at 16°C.

Analysis of Ligation Products- DNA ligase assays were carried out as described previously. The reactions were stopped by heating at 90°C for 10 min. An aliquot (5 µl) was heated for 5 min at 100°C in 65% formamide prior to loading onto a denaturing, 20% acrylamide gel. After electrophoresis for 3 h at 300 V, the gel was fixed in 40% methanol, 10% acetic acid for 30 min and then dried. Oligonucleotides were visualised by autoradiography.

Formation of DNA ligase-adenylate - Reaction mixtures (10 µl) contained 60 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM DTT, 50 µg/ml bovine serum albumin, 0.5 µCi [α-^{32}P] ATP (3000 Ci/mmol, Amersham) and DNA ligase. Incubations were at room temperature for 15 min. After the addition of 5 µl of SDS sample buffer, reaction mixtures were heated at 90°C for 10 min. Proteins were separated by SDS-PAGE (39). Gels were fixed for 10 min in 10% acetic acid, dried and adenylylated proteins detected by autoradiography.

To examine the reactivity of the enzyme-adenylate intermediate, the adenylylation reaction was performed as described above in a final volume of 40 µl. Aliquots (10 µl) were incubated with; 0.8 µg unlabelled oligo(dT)-poly(dA); 0.8 µg unlabelled oligo(dT)-poly(rA) or 10 nmoles of sodium pyrophosphate for 1 h at 37°C. The reactions were stopped by the addition of SDS sample buffer and adenylylated polypeptides detected as described above.

N-Terminal Amino Acid Determination - Purified DNA ligase II was dialysed into 100 mM Tris-HCl, pH 8.2, 0.1% SDS. Automated NH_2-terminal sequencing was performed on an Applied Biosystems model 473A sequenator.

Amino Acid Sequencing- DNA ligase II (Table 1S, Fraction VIII) was concentrated and the buffer exchanged into 100 mM Tris-HCl pH 8.2, 0.1% SDS using a Centricon-10, micro-concentrator (Amicon). The addition of SDS prevented precipitation of the intact protein and proteolytic fragments. 30 µg of the concentrated protein (1mg/ml), was digested using Lysylendopeptidase (WAKO) at a concentration of 1:30, enzyme/substrate (w/w), for 72 hrs at 30°C. Following peptide purification by ion-exchange/reverse phase HPLC (40) the sequences of 14 peptides were obtained by microsequencing (25). The peptide sequences were run against the OWL9 protein database.

Immunization Procedures- A 15-mer peptide (sequence; PKTINVVVTTMDAEC) corresponding to the N terminal sequence of ezrin (except for the C terminal cysteine linker) was made with a commercial peptide synthesizer. The peptide was coupled to keyhole limpet hemocyanin using the glutaraldehyde method (41). Rabbit antibodies were raised against the peptide coupled to keyhole limpet hemocyanin by emulsifying 200 µg peptide conjugate with Freund's complete adjuvant and injecting subcutaneously. Subsequent injections were emulsified with Freunds incomplete adjuvant. The rabbit received 8 injections and was bled 10 days after the last injection.

Immunoblots - Proteins were separated by SDS-PAGE (39) and then transferred to nitrocellulose filters (Schleicher and Schuell) in 25 mM Tris, 192 mM glycine, 20% methanol (pH 8.3). After transfer, membranes were incubated for 1 h at 37°C in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.05% Tween 20 (TBST) with 5% non-fat dry milk and then washed twice with TBST. Nitrocellulose membranes were incubated for 12-16 h at 4°C with antibody diluted in TBST plus 2% non-fat dry milk. All subsequent steps were carried out at room temperature. Membranes were washed four times with TBST. After incubation for 5 min in TBST plus 2% non-fat dry milk, [^{125}I] protein A (0.1 µCi/ml, Amersham) was added and incubation continued for 1 h. Membranes were then washed with TBST, dried and exposed to preflashed X-ray film (Kodak) at -70°C with an intensifying screen. Alternatively, antigen-antibody complexes were detected using a second antibody coupled to alkaline phosphatase. After incubation in TBST plus 2% non-fat dry milk, membranes were incubated with goat anti-rabbit antibody linked to alkaline phosphatase for 2 h, washed with TBST, rinsed with water and blotted dry. Alkaline phosphatase activity was visualized by washing the membrane in 100 mM Tris-HCl (pH 9.5), 100 mM NaCl and 5 mM $MgCl_2$ followed by incubation in 15 ml of the same buffer containing 44 µl of 75 mg/ml p-nitrotetrazolium blue chloride and 33 µl of 50 mg/ml 5-bromo-4-chloro-3-indolylphosphate p-toluidine salt, both in dimethylformamide. The reaction was stopped by the addition of 20 mM Tris-HCl (pH 8.0), 5 mM EDTA.

Two-dimensional gel electrophoresis- Proteins were separated by the method of O'Farrell (42). In the first dimension, proteins were resolved by isoelectric focusing using 5% polyacrylamide rod gels containing 2% ampholines (pH range 3.5-10, Pharmacia). The gels were then equilibrated in SDS sample buffer and proteins separated in the second dimension by electrophoresis through a 10% SDS-polyacrylamide gel (39).

Sucrose Gradient Density Centrifugation - This was carried out as described previously (15).

Purification of DNA Ligase I from Calf Thymus- The isolation of a >95% homogeneous enzyme was performed as described (15).

Table 1S. Purification of DNA Ligase II from Calf Thymus

Purification of DNA ligase II from 500 g of tissue. Activities of fractions were assessed by joining a oligo(dT)-poly(rA) substrate. Protein concentrations were determined by the Coomassie Blue method (43).

Fraction	Protein (mg)	oligo(dT)-poly(rA)	
		Total activity (units)	Specific activity (units/mg)
I. Crude extract	19000	-	-
II. Phosphocellulose	1600	-	-
III. Ammonium sulphate (42-66%)	490	-	-
IV. Hydroxylapatite	56	5	0.09
V. Gel filtration	28	1.3[a]	0.05
VI. FPLC Phenyl Superose	5	0.83	0.17
VII. FPLC Mono Q	2	1.4[b]	0.72
VIII. FPLC Mono S	0.3	0.41	1.4

Activity numbers for the first three steps are difficult to estimate given the lability of this enzyme and contamination by RNase H which degrades the substrate. In agreement with previous observations (6), pilot experiments established that DNA ligase II is largely released from cell nuclei with 0.2 M NaCl, adsorbs to phosphocellulose, and is largely recovered in the ammonium sulphate 42-66% saturation fraction.

[a] These fractions are assayed at a high salt concentrations which partly inhibits the enzyme activity (approximately 50% inhibition at 30 mM NaCl). The values have been adjusted accordingly. Although this step results in little purification and a small decrease rather than increase in specific activity, it is useful to remove contaminating DNA ligases I and III.

[b] An apparent increase in total activity has been reproducibly observed at this stage.

Purification of DNA Ligase II from Calf Thymus - All procedures were carried out at 0-4°C and centrifugations were at 10,000 x *g* for 30 min unless otherwise stated. Calf thymus glands (3 kg) were disrupted in aliquots by homogenisation for 3 x 30 sec in a Waring blendor with 3 litres of a buffer containing 0.2 M NaCl, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 10 mM mercaptoethanol, 1 mM phenylmethylsulphonylfluoride, 1.9 µg/ml aprotinin and 0.5 µg/ml each of leupeptin, pepstatin, chymostatin and TLCK. After gentle stirring for 1 h, cellular debris was removed by centrifugation. The supernatant (Fraction I, crude extract) was diluted to a final NaCl concentration of 20 mM by the addition of 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 10 mM mercaptoethanol. This was batch adsorbed to 3.5 litres of a thick slurry (500g dry weight) of P11 phosphocellulose (Whatman), which had been pre-equilibrated with 20 mM NaCl, 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.5 mM DTT (buffer A). After washing the phosphocellulose with 10 litres of buffer A, adsorbed proteins were eluted with 4 litres of buffer A containing 0.7 M NaCl. Ammonium sulphate was added to the eluate (Fraction II) at 231 g/l and 1M Tris base was added as necessary to maintain the pH at 7.1-7.5. After 30 min of gentle stirring the precipitate was removed by centrifugation. Additional ammonium sulphate (160 g/l) was added to the supernatant and the resulting suspension was again neutralised and stirred gently for 30 min. The precipitate was collected by centrifugation, divided into six aliquots and quickly frozen. These precipitates could be stored at -70°C with no significant loss of DNA ligase activity after several weeks. By this procedure, we avoided the use of frozen thymus glands which yield extracts containing higher levels of protease activity than fresh glands.

One aliquot of the ammonium sulphate precipitate (material from 500 g of tissue) was resuspended in 15 ml and dialysed for 3 h against 50 mM potassium phosphate (pH 7.5), 0.5 mM DTT (buffer B). Insoluble material was removed by centrifugation (Fraction III, 25ml). Fraction III was loaded onto a 5.9 x 7.5 cm hydroxylapatite (Bio Rad, HT) column which had been equilibrated with buffer B. Proteins were eluted by steps of 250 mM potassium phosphate (pH 7.5), 0.5 mM DTT and 400 mM potassium phosphate (pH 7.5), 0.5 mM DTT and assayed for DNA ligation activity and for enzyme-adenylate complex formation. The 250 mM potassium phosphate (pH 7.5) eluate contained the majority of the DNA ligase I and DNA ligase III activities, whereas the 400 mM potassium phosphate (pH 7.5) eluate contained the DNA ligase II activity. This pool was concentrated by ultrafiltration using a 50 ml Amicon filtration unit with a YM-10 membrane and a pressure of 50 psi. The concentrate (Fraction IV, 6 ml) was then dialysed into 50 mM Tris-HCl (pH 7.5), 1 M NaCl, 0.5 mM DTT, 1 mM EDTA, 0.2% Tween 20 (buffer C). The dialysed Fraction IV was centrifuged for 15 min and loaded onto a 0.9 x 97.5 cm Ultrogel AcA 44 column (Pharmacia LKB Biotechnology Inc.), pre-equilibrated with buffer C. After elution with buffer C, fractions were assayed for DNA ligation activity and for enzyme-adenylate complex formation. Active fractions were pooled and dialysed against 50 mM Tris-HCl (pH 7.5), 0.9 M $(NH_4)_2SO_4$, 0.5 mM DTT, 1 mM EDTA, 0.2% Tween 20 (buffer D). This pool (Fraction V, 8 ml) contained two major polypeptides of 72 kDa and 74 kDa. After centrifugation for 15 min, Fraction V was loaded onto a FPLC Phenyl Superose column (Pharmacia), pre-equilibrated with buffer D. DNA ligase II activity and the 72 kDa polypeptide passed through the column (Fraction VI, 10ml), whereas the 74 kDa polypeptide was retained. Fraction VI was dialysed against 50 mM Tris-HCl (pH 7.5), 30 mM NaCl, 0.5 mM DTT, 1 mM EDTA, 0.2% Tween 20 (buffer E) and then loaded onto a FPLC Mono-Q column, pre-equilibrated in buffer E. DNA ligase II activity passed through the column whereas the remaining traces of DNA ligases I and III were retained by the column. The pass-through fractions were pooled (Fraction VII, 20ml), and loaded onto a FPLC Mono-S column pre-equilibrated in buffer E. Proteins were eluted with a linear gradient of 30 - 500 mM NaCl in buffer E. Fractions containing DNA ligase II activity, which eluted at 0.22 - 0.25 M NaCl (Fraction VIII), were dialysed against buffer E containing 50 % glycerol and stored at -20°C. The purified enzyme lost 50% of its activity in 2-3 weeks under these conditions.

Table 2S. Purification of DNA Ligase III from Calf Thymus
The purification of DNA ligase III from 500 g of calf thymus is shown. The first three steps were usually performed at a six times larger scale, as described in Experimental Procedures. Protein concentrations were determined by the Coomassie Blue method (43).

Fraction	Protein (mg)	oligo(dT)-poly(rA)	
		Total activity (units)	Specific activity (units/mg)
I. Crude extract	19000	-	-
II. Phosphocellulose I	1600	-	-
III. Ammonium sulphate (0-50%)	1400	-	-
IV. Gel filtration	150	120[a]	0.8
V. Phosphocellulose II	20	114	5.7
VI. Q Sepharose	5	86	17.2
VII. FPLC Phenyl Superose	0.3	30[a]	100
VIII. FPLC Mono S	0.06	24	400

[a] This fraction is assayed at a salt concentration which partially inhibits the ligation reaction. These numbers have been adjusted accordingly.

Purification of DNA Ligase III from Calf Thymus - The initial fractionation steps for DNA ligase III were the same as for DNA ligase II. After the phosphocellulose batch step, the eluate (Fraction II) was supplemented with 291 g/liter ammonium sulphate, and 1 M Tris base was added intermittently to keep the pH at 7.1-7.5. After 30 min of gentle stirring, the precipitate was collected by centrifugation, divided into six equal aliquots, and quickly frozen. The ammonium sulphate precipitates could be stored at -70°C with no significant loss of DNA ligase III activity after 3 months.

One sixth of the ammonium sulphate precipitate (material from 500 g tissue) was diluted to a thick slurry and dialyzed for 4 h against 1 M NaCl, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.5 mM DTT (buffer B). Insoluble material was removed by centrifugation at 30,000 rpm for 1 h in a Beckman 45Ti rotor (Fraction III, 30 ml). Fraction III was loaded onto a 2.5 x 125 cm Ultrogel AcA 34 (Pharmacia) column, equilibrated with buffer B. Proteins were eluted with buffer B and fractions were assayed for ligation activity on both the oligo(dT)-poly(dA) and oligo(dT)-poly(rA) substrates (Figure 1S) and also for formation of enzyme-adenylate complex.

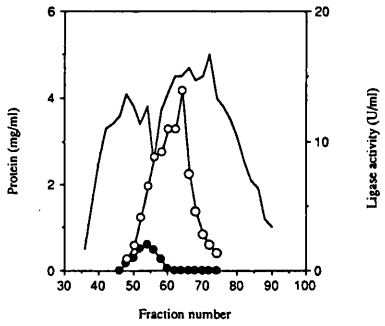


Figure 1S; Fractionation of calf thymus DNA ligase activity by gel filtration - (——) protein (mg/ml), (—●—) ligase activity with a oligo(dT)-poly(rA) substrate, (—○—) ligase activity with a oligo(dT)-poly(dA) substrate. The activity peak at fraction 54 represents DNA ligase III, whereas DNA ligase I is mainly eluted in fractions 60-65. The peak of DNA ligase II activity occurs around fraction 75, but is masked here by a contaminating RNase H activity which is subsequently removed on hydroxylapatite chromatography.

Fractions acting on the oligo(dT)-poly(rA) substrate (Fraction IV, 40 ml) were pooled, dialyzed against 30 mM NaCl, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.5 mM DTT and 10% glycerol (buffer C) and then loaded onto a 1.6 x 20 cm phosphocellulose (P11, Whatman) column, which had been equilibrated with buffer C. Proteins were eluted with a 400 ml linear gradient of 0-1 M NaCl in buffer C. Fractions were assayed for ligation activity on both the poly(dA)-oligo(dT) and the poly(rA)-oligo(dT) substrates (Figure 2S) and also for formation of the enzyme-adenylate complex.

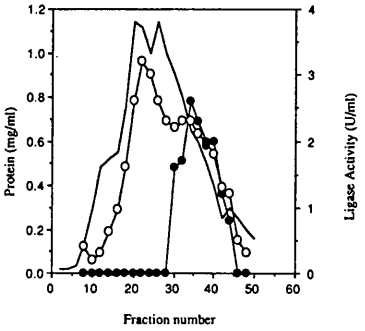


Figure 2S; Separation of DNA ligases I and III by phosphocellulose chromatography - (——) protein (mg/ml), (—●—) oligo(dT)-poly(rA) substrate, (—○—) oligo(dT)-poly(dA) substrate. The peak of DNA ligase III activity occurs in fraction 34, whereas the DNA ligase I activity is highest in fraction 22.

The peak of DNA ligase I activity eluted at 0.2 M NaCl whereas the peak of DNA ligase III activity, joining the oligo(dT)-poly(rA) substrate, eluted at 0.35 M NaCl. Adenylylated polypeptides of 100 kDa and 87 kDa were detected in fractions containing DNA ligase III activity. Fractions were pooled (Fraction V, 20 ml) to enhance purification of the larger polypeptide. After dialysis against buffer C, Fraction V was loaded onto a 1.0 x 20 cm Q Sepharose (Pharmacia) column, which had been equilibrated with buffer C. Proteins were eluted with a 200 ml linear gradient of 0-1 M NaCl in buffer C. Active fractions (Fraction VI, 10 ml), eluting at 0.25 m NaCl, were pooled and dialyzed against buffer C containing 0.9 M ammonium sulfate.
The dialyzed Fraction VI was loaded onto a FPLC Phenyl Superose column (Pharmacia), pre-equilibrated with buffer C containing 0.9 M ammonium sulphate. Proteins were eluted with a 20 ml linear gradient of 0.9-0 M ammonium sulphate. Active fractions, which eluted at 0.45 M ammonium sulfate, were pooled (Fraction VII, 5 ml) and dialysed against buffer C. Although this hydrophobic chromatography step results in a relatively poor recovery of DNA ligase III, it is the most effective procedure, of several investigated, for removing contaminating DNA ligase I. Fraction VII was loaded onto a FPLC Mono S column (Pharmacia), pre-equilibrated with buffer C. Proteins were eluted with a 30 ml linear gradient of 0-1 M NaCl in buffer C. Fractions containing activity (Fraction VIII, 1.0 ml) eluted at 0.3 M and did not significantly lose activity after storage for up to two weeks on ice. For longer term storage these fractions were pooled, dialysed against buffer C containing 50% glycerol and stored at -20°C. DNA ligase III is most active in 5-10 mM NaCl and is inhibited sharply by higher salt concentrations, with 70 % inhibition at 50 mM NaCl. The activity in fractions kept at -20°C was stable for 2-3 months.

DATE:

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 269, No. 0, Issue of xxx 0, pp. 1-4, 1994
Printed in U.S.A.

Different Active Sites of Mammalian DNA Ligases I and II*

(Received for publication, September 24, 1993)

Emma Roberts, Rachel A. Nash, Peter Robins, and Tomas Lindahl‡

From the Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, United Kingdom

Bovine DNA ligases I and II were adenylylated in the presence of [α-^{32}P]ATP and digested with limiting amounts of trypsin or V8 protease. The generation of radioactive peptides of decreasing size was monitored by polyacrylamide gel electrophoresis and autoradiography. Active site peptides obtained by complete proteolytic digestions with trypsin, V8, or Lys-C protease were also compared. The partial digestion products of DNA ligases I and II were entirely different, with no indication of extensive sequence homology. Furthermore, the sequence of the active site region of DNA ligase I is clearly different from that of DNA ligase II. Similar analysis of a third chromatographically distinct mammalian DNA ligase indicated that it is different from DNA ligase I but related to DNA ligase II.

DNA ligases catalyze the formation of phosphodiester bonds at single-strand breaks in double-strand DNA (1, 2). In the first step of the ligation reaction, a covalent DNA ligase-AMP intermediate is formed with a phosphoamide bond between a lysine residue and the nucleotide (3). The active site of mammalian DNA ligase I has been identified by isolation and amino acid sequencing of a tryptic peptide containing a lysine-[^{3}H]AMP moiety (4) and has been localized by comparison with the complete peptide sequence deduced from the cDNA (5). The sequence and position of the active site of the bacteriophage T4 RNA ligase has also been established in this way (6). Sequence comparisons with DNA ligases from other species have identified a ubiquitous, KXDG active site motif (4). This motif has also been found at the active site of the vaccinia virus mRNA capping enzyme that catalyzes transfer of GMP from GTP to the 5′ terminus of RNA and forms an enzyme-guanylate reaction intermediate with a lysine-GMP phosphoamide bond. The observed sequence conservation indicates a common mechanism of covalent catalysis in these nucleotidyl transfer reactions (7). Outside the short KXDG motif, the amino acid sequences of DNA ligases from different sources diverge.

A second distinct enzyme, DNA ligase II, has been described in mammalian cell nuclei (8, 9). This enzyme differs from DNA ligase I in being able to join single-strand interruptions in an oligo(dT)·poly(rA) hybrid substrate. A third chromatographically distinct mammalian DNA ligase, designated DNA ligase III, can also join this substrate (10, 11). DNA ligase II has been compared with DNA ligase I by peptide fingerprinting of their AMP-binding regions (12). The authors concluded that the catalytic sites of the two enzymes were closely related or identical over an extended region, although differences between the two proteins were apparent in their DNA binding domains (12, 13). In the present work, we have compared active site peptides of bovine DNA ligases I and II after partial or complete digestion with several proteolytic enzymes. The active site regions of the two enzymes are clearly different, with no indication of extensive sequence homology.

EXPERIMENTAL PROCEDURES

Formation of DNA Ligase-AMP—Bovine DNA ligases I, II, and III were isolated essentially as described previously (11, 14) but with emphasis on separating the DNA ligases from each other rather than achieving high degrees of purity. DNA ligases I and III were obtained from calf thymus, whereas DNA ligase II was isolated from calf liver. In agreement with previous results (see Ref. 11), [^{32}P]AMP was transferred from DNA ligases II and III, but not DNA ligase I, on incubation of the adenylylated enzymes with an oligo(dT)·poly(rA) substrate, whereas all three enzymes were able to transfer the AMP moiety to oligo(dT)·poly(dA).

DNA ligases were adenylylated in reaction mixtures (30 μl) containing 60 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM dithiothreitol, 50 μg·ml^{-1} bovine serum albumin, and 0.5 μCi [α-^{32}P]ATP (3000 Ci/mmol, Amersham) at 20 °C for 10 min. After addition of EDTA to a final concentration of 15 mM, aliquots were removed for proteolytic digestion.

Partial Digestion of DNA Ligase-AMP with Trypsin or V8 Protease—Aliquots of adenylylation reaction mixtures (1–4 μg of total protein) were supplemented with 0–10 ng of trypsin (sequencing grade, Boehringer Mannheim) or 0–600 ng of V8 protease (endoproteinase Glu-C, sequencing grade, Boehringer Mannheim) and incubated at 37 °C for 15 min. Digestions were stopped by the addition of SDS sample buffer and heating at 90 °C for 10 min. The proteolytic products were separated by SDS-polyacrylamide gel electrophoresis (15) in 10–20% gradient gels, at a constant current of 12 mA. The dimensions of the gels (45 × 180 × 5 mm) allowed maximum separation of multiple products. Gels were fixed in 5% methanol, 7% acetic acid, 5% glycerol for 1 h prior to drying and autoradiography.

Complete Digestion of DNA Ligase-AMP by Trypsin or Lys-C Protease—Aliquots of adenylylation reaction mixtures (0.5–2 μg of total protein) were supplemented with 800 ng of either trypsin or Lys-C protease (sequencing grade, Boehringer Mannheim) and incubated at 37 °C for 4 h. Samples were dried and dissolved in 20 μl of loading buffer (8 M urea, 0.125 M Tris-HCl (pH 6.8) and 0.1% bromphenol blue), to which dithiothreitol was added to 12.5 mM. After incubation at 37 °C for 30 min, each sample was divided into two, one of which was treated with iodoacetamide (final concentration 23 mM) at pH 7.0 to block cysteine residues (16). Both samples were then incubated at 37 °C for 30 min prior to electrophoresis.

The products of the complete proteolytic digests were separated in gels prepared according to Pantazis and Bonner (17). These gels (pH 8.8) contained high polyacrylamide concentrations but no SDS, providing a system that allows efficient separation of small peptides according to both molecular mass and charge. The products of Lys-C digestion were separated in gels containing 25% acrylamide (of which 0.15% is bisacrylamide). For the separation of trypsin digestion products, the acrylamide concentration was increased to 30%. Electrophoresis was carried out at constant voltage: 100 V for 1 h at 0 °C, then 200 V at room temperature. The gels were dried without fixing and radioactive peptides visualized by autoradiography.

RESULTS

Partial Proteolysis of DNA Ligases I and II by Trypsin or V8 Proteases—Digestion of ^{32}P-adenylylated forms of DNA ligases I and II by trypsin or V8 was carried out using a range of protease concentrations, and the radioactive peptides charac-

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Fax: 44-707-649404.

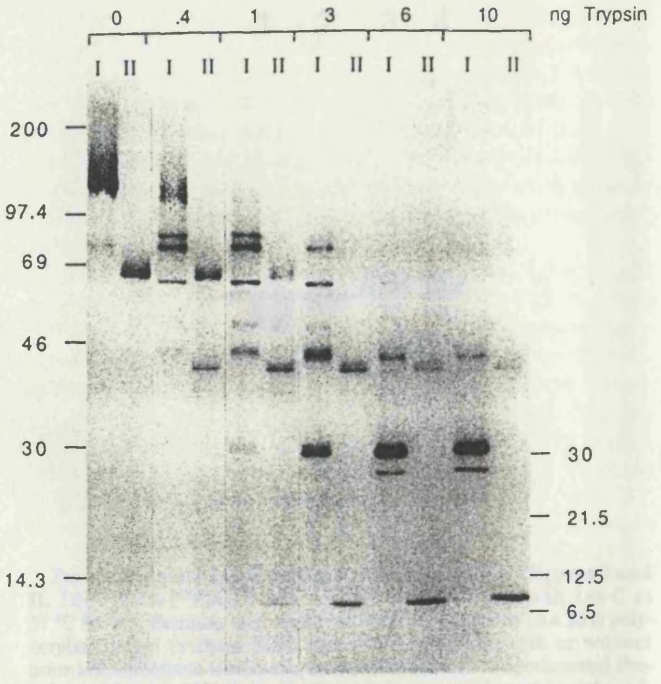


FIG. 1. **Partial trypsin digestion of DNA ligases I and II.** DNA ligase-[^{32}P]AMP was incubated at 37 °C for 15 min with trypsin as indicated. Products were separated by SDS-polyacrylamide gel electrophoresis in 10–20% gradient gels and radioactive peptides detected by autoradiography. ^{14}C-Methylated protein molecular mass markers (Amersham Corp.) were electrophoresed on the same gel and are shown in kDa.

teristic of each protein compared by SDS-polyacrylamide gel electrophoresis and autoradiography (18). DNA ligase I is a 102-kDa protein that migrates anomalously slowly as a 125-kDa polypeptide during SDS-polyacrylamide gel electrophoresis (14). A high proportion of charged amino acids and the presence of phosphoserine residues could account for this discrepancy (19). In the present experiments, which aimed at maximum electrophoretic resolution of proteolytic digestion products, the 125-kDa polypeptide migrated as a broad band (Figs. 1 and 2, *lane 1*). Trypsin digestion yielded major DNA ligase I degradation products of approximately 85, 80, 60, 50, 45, and 30 kDa (Fig. 1). The 85-kDa fragment apparently corresponds to the previously characterized 85-kDa catalytic domain of the enzyme that is readily generated by endogenous cellular proteases or after limited treatment of DNA ligase I with subtilisin, trypsin, chymotrypsin, or V8 protease (14). Digestion of DNA ligase II with trypsin yielded a completely different pattern of peptides (Fig. 1). DNA ligase II migrates with an apparent molecular mass of 68–72 kDa upon SDS-polyacrylamide gel electrophoresis, and here gave major tryptic peptides of approximately 42 and 8 kDa. The digestion products of the calf liver DNA ligase II were identical to those observed for the calf thymus enzyme (data not shown).

DNA ligases I and II also yielded different peptides after cleavage by V8 protease. In this case, DNA ligase I gave major proteolytic products of approximately 85, 38, 24, 14, 7, and 6 kDa, whereas the main products of DNA ligase II were 50, 37, 26, and 9 kDa (Fig. 2). A 22-kDa active site domain generated by partial V8 digestion and reported to be common to both DNA ligases I and II (12) was not observed in the present experiments.

Complete Proteolysis of DNA Ligases I and II by Trypsin or Lys-C Proteases—Additional experiments were performed to characterize the active site peptides of DNA ligases I and II following complete digestion with trypsin or Lys-C proteases. High protease concentrations were preferred to prolonged incubation periods in order to minimize slow non-enzymatic hydrolysis of the Lys-AMP phosphoamide bond. The neutral gel system of Pantazis and Bonner (17), with high polyacrylamide concentrations and no SDS allowed efficient resolution of small proteolytic products according to both their molecular mass and charge. The tryptic peptide containing the Lys-AMP has

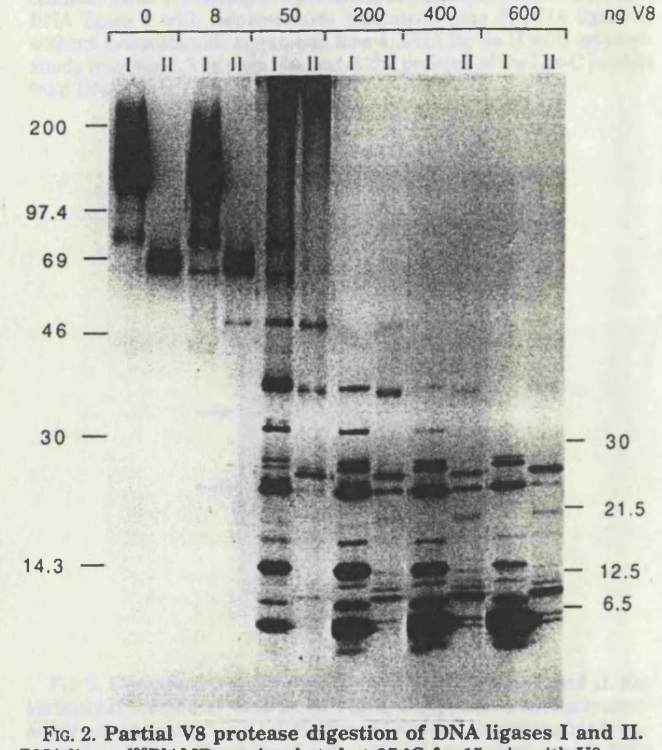


FIG. 2. **Partial V8 protease digestion of DNA ligases I and II.** DNA ligase-[^{32}P]AMP was incubated at 37 °C for 15 min with V8 protease as indicated. Experimental procedures and size markers were as for Fig. 1.

been sequenced for bovine DNA ligase I and is identical to a 16-amino acid (1.9-kDa) peptide defined by arginine residues in the human cDNA (4). The probable molecular masses of other active site peptides of bovine DNA ligase I can be estimated from the predicted amino acid sequence of the human cDNA (Fig. 3).

Lys-C cuts at the C-terminal side of lysine residues but does not recognize a Lys-AMP moiety; digestion of human DNA ligase I-[^{32}P]AMP would yield a 29-amino acid radioactive peptide of 3.3 kDa (Fig. 3). The single radiolabeled Lys-C peptides derived from DNA ligase I *versus* DNA ligase II were different, with DNA ligase II yielding a more slowly migrating peptide

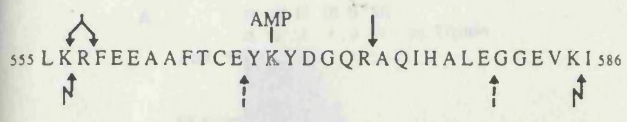


FIG. 3. **Amino acid residues 555–586 in mammalian DNA ligase I, derived from the human DNA ligase I cDNA.** Residues 558–573 have also been determined by amino acid sequencing of a bovine DNA ligase I tryptic peptide and shown to be identical with the human sequence (4). The active site Lys-568, as well as cleavage sites around this residue for trypsin (↓), Lys-C protease (⇃), and V8 protease (↑), are shown.

(Fig. 4, *lanes 2* and *4*). A 3-fold increase or decrease in the amount of Lys-C protease used for digestion produced no change in migration of the DNA ligase-derived peptides. The DNA ligase I digest initially produced two detectable bands; however, blocking of cysteine residues using iodoacetamide converted these to a single band (Fig. 4, *lanes 1* and *2*). This is in agreement with the DNA ligase I sequence, which shows that this active site peptide contains a Cys residue (Fig. 3).

The peptides of DNA ligase I, obtained consistently after a 4-h incubation with excess trypsin, are shown in Fig. 5. Four different peptides were detected, which were reduced to two by iodoacetamide treatment (*lanes 1* and *2*), consistent with the presence of a Cys residue in the active site peptide. These migrated more rapidly than the Lys-C peptides of DNA ligase I under identical conditions, which would be expected for smaller peptides of similar charge (Fig. 3). DNA ligase II (iodoacetamide-treated) also yielded two tryptic peptides (Fig. 5, *lane 4*), but these migrated significantly more slowly than the peptides from DNA ligase I (*lane 2*). The occurrence of two different tryptic peptides of DNA ligase I might be explained by the presence of two adjacent basic amino acids at the N-terminal cleavage site (Fig. 3). Trypsin treatment could generate two peptides; one with an N-terminal Phe residue being the DNA ligase I active site tryptic peptide previously sequenced (4), the other a slightly longer peptide with an N-terminal Arg residue. The latter would be less negatively charged and would migrate more slowly in the gel electrophoresis system employed here. Complete digestion of DNA ligase II with trypsin yielded two peptides, one of which apparently contains a Cys residue. Lys-C digestion of DNA ligase II yielded only a single peptide. These data could be explained by potential Lys-C and trypsin sites within a sequence context such as *X*RCK*X* at the N termini of the active site peptides.

Complete digestion of DNA ligase I-[^{32}P]AMP with V8 protease yielded a single radioactive peptide which migrated slightly more rapidly than the Lys-C-derived peptide shown in Fig. 4 (*lane 2*) under identical conditions. No difference in migration of this peptide was observed after iodoacetamide treatment. These data (not shown) are consistent with the amino acid sequence around the active site, which indicates that a small peptide of 14 residues without a Cys residue would be generated from DNA ligase I by V8 protease treatment (Fig. 3). In contrast, no active site peptide of V8-digested DNA ligase II was observed in repeated attempts. It seems likely that extensive V8 protease treatment of DNA ligase II generated a very small or highly charged peptide that was not detected in the electrophoresis system employed. In this regard, the Gln residue at position 572 in the mammalian DNA ligase I is unusual; other DNA ligases with evolutionarily conserved active site regions have a Glu residue at this position (4). Thus, the sequence at the active site of the *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* DNA ligases is -EYKYDGE-, whereas that of mammalian DNA ligase I is -EYKYDGQ-. A DNA ligase with the former sequence would yield a very short V8 peptide containing the active site Lys residue.

Proteolytic Digestion Products of DNA Ligase III—Partial

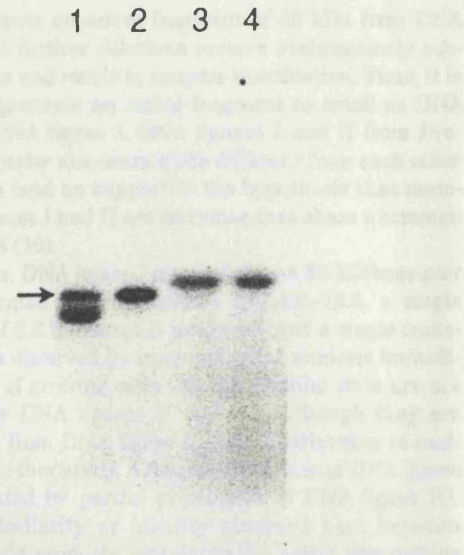


FIG. 4. **Complete Lys-C protease digestion of DNA ligases I and II.** DNA ligase-[^{32}P]AMP was digested to completion with Lys-C at 37 °C for 4 h. Samples were analyzed by electrophoresis in a 25% polyacrylamide gel (without SDS) and autoradiography, with or without prior iodoacetamide treatment, as described under "Experimental Procedures." *Lane 1*, DNA ligase I without iodoacetamide treatment; *lane 2*, DNA ligase I with iodoacetamide treatment; *lane 3*, DNA ligase II without iodoacetamide treatment; *lane 4*, DNA ligase II with iodoacetamide treatment. The *arrow* indicates the position of the Lys-C peptide from DNA ligase I (see Fig. 3).

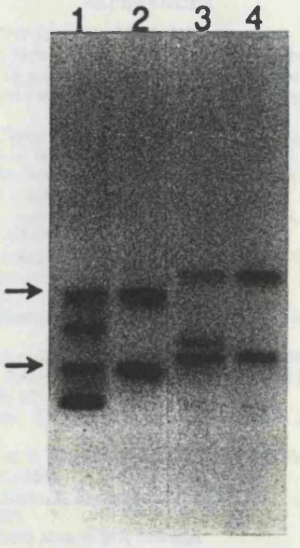


FIG. 5. **Complete trypsin digestion of DNA ligases I and II.** Experimental procedures were as for Fig. 4, except that a 30% polyacrylamide gel was used. *Lane 1*, DNA ligase I without iodoacetamide treatment; *lane 2*, DNA ligase I with iodoacetamide treatment; *lane 3*, DNA ligase II without iodoacetamide treatment; *lane 4*, DNA ligase II with iodoacetamide treatment. The *arrows* indicate the positions of the two DNA ligase I tryptic peptides (see Fig. 3).

digestion of the 97-kDa DNA ligase III-[^{32}P]AMP with trypsin yielded major fragments of approximately 87, 57, and 43 kDa (Fig. 6*A*). These did not resemble the initial digestion products of either DNA ligase I or II. An 87-kDa fragment of DNA ligase III is also generated by endogenous proteases during enzyme purification and by limited subtilisin treatment (11). More extensive digestion of DNA ligase III with trypsin yielded an 8-kDa fragment of the same size as a fragment from DNA ligase II (Fig. 6*A*). Partial digestion of the ^{32}P-adenylylated forms of DNA ligases II and III with V8 protease showed more clearly that smaller peptides (<35 kDa) were similar or identical,

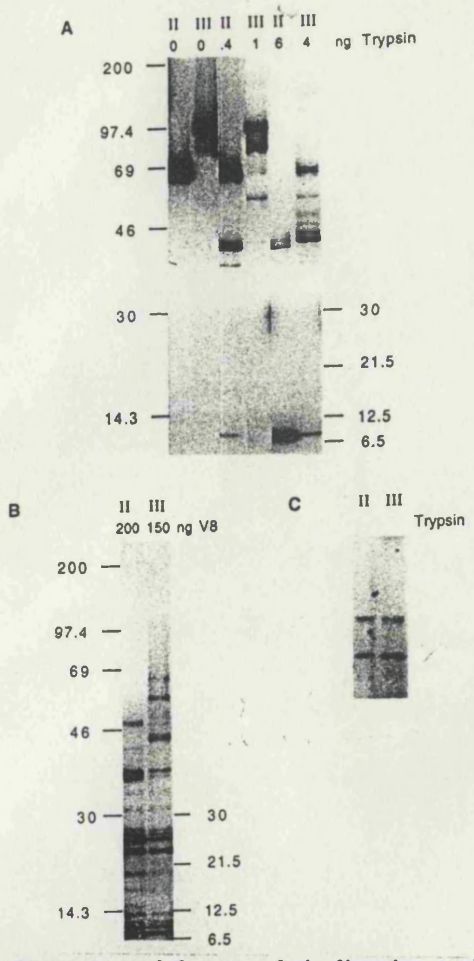


FIG. 6. **Comparison of the proteolytic digestion products of DNA ligases II and III.** DNA ligase-[^{32}P]AMP was subjected to limited digestion at 37 °C for 15 min with trypsin (*A*) or V8 protease (*B*). Partial digestion products were separated and visualized as for Fig. 1. *C*, complete digestions with trypsin as for Fig. 5; iodoacetamide-treated ^{32}P-labeled tryptic peptides are shown.

whereas larger fragments differed (Fig. 6*B*). Moreover, complete digestion of DNA ligase III-[^{32}P]AMP with trypsin yielded active site peptides indistinguishable from those of DNA ligase II (Fig. 6*C*), but different from DNA ligase I (Fig. 5). These data indicate close similarity or identity of the region around the active site Lys residue in DNA ligases II and III.

DISCUSSION

This comparison of proteolytic digestion products of the ^{32}P-adenylylated forms of bovine DNA ligases I and II indicates that the two proteins are not directly related, since common proteolytic fragments were not observed. This is supported by previous immunological data. An antibody that effectively neutralized DNA ligase I from bovine, human, murine, and rabbit cells did not inhibit DNA ligase II (20); conversely, a neutralizing antibody against DNA ligase II did not affect the activity of DNA ligase I (9). Moreover, deletion mutagenesis at both termini of the DNA ligase I cDNA (21) has shown that it is possible to generate an active fragment of 80 kDa from DNA ligase I but that further deletions remove evolutionarily conserved sequences and result in enzyme inactivation. Thus, it is not possible to generate an active fragment as small as DNA ligase II from DNA ligase I. DNA ligases I and II from *Drosophila melanogaster* also seem quite different from each other (22). These data lend no support to the hypothesis that mammalian DNA ligases I and II are isozymes that share a common catalytic domain (12).

In human cells, DNA ligase I is encoded by a 53-kilobase pair unique gene located on chromosome 19q13.2–13.3, a single mRNA species of 3.2 kilobases is produced, and a single translation product is observed by immunological analysis immediately after lysis of growing cells (19, 23). Similar data are not yet available for DNA ligases II and III. Although they are clearly different from DNA ligase I, their relationship to each other requires further study. A fragment the size of DNA ligase II is not generated by partial proteolysis of DNA ligase III. However, the similarity or identity observed here between smaller proteolytic products containing the active site regions indicate that DNA ligases II and III could be generated from the same gene by alternative splicing or by differential processing of an initial translation product. A similar suggestion has been made by Elder *et al.* (24). Molecular cloning of appropriate cDNA sequences is required to clarify this issue.

Acknowledgments—We thank Drs. D. E. Barnes, G. Evan, and G. G. Poirier for valuable discussions.

REFERENCES

1. Lehman, I. R. (1974) *Science* **186,** 790–797
2. Engler, M. J., and Richardson, C. C. (1982) in *The Enzymes* (Boyer, P. D., ed) 3rd Ed., Vol. 15B, pp. 1–29, Academic Press, New York
3. Gumport, R. I., and Lehman, I. R. (1971) *Proc. Natl. Acad. Sci. U. S. A.* **68,** 2559–2563
4. Tomkinson, A. E., Totty, N. F., Ginsburg, M., and Lindahl, T. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 400–404
5. Barnes, D. E., Johnston, L. H., Kodama, K., Tomkinson, A. E., Lasko, D. D., and Lindahl, T. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 6679–6683
6. Thorgersen, H. C., Morris, H. R., Rand, K. N., and Gait, M. J. (1985) *Eur. J. Biochem.* **147,** 325–329
7. Cong, P., and Shuman, S. (1993) *J. Biol. Chem.* **268,** 7256–7260
8. Arrand, J. E., Willis, A. E., Goldsmith, I., and Lindahl, T. (1986) *J. Biol. Chem.* **261,** 9079–9082
9. Teraoka, H., Sumikawa, T., and Tsukada, K. (1986) *J. Biol. Chem.* **261,** 6888–6892
10. Elder, R. H., and Rossignol, J.-M. (1990) *Biochem.* **29,** 6009–6017
11. Tomkinson, A. E., Roberts, E., Daly, G., Totty, N. F., and Lindahl, T. (1991) *J. Biol. Chem.* **266,** 21728–21735
12. Yang, S.-W., Becker, F. F., and Chan, J. Y.-H. (1990) *J. Biol. Chem.* **265,** 18130–18134
13. Yang, S.-W., Becker, F. F., and Chan, J. Y.-H. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 2227–2231
14. Tomkinson, A. E., Lasko, D. D., Daly, G., and Lindahl, T. (1990) *J. Biol. Chem.* **265,** 12611–12617
15. Laemmli, U. K. (1970) *Nature* **227,** 680–685
16. Lane, L. C. (1978) *Anal. Biochem.* **86,** 655–664
17. Pantazis, P., and Bonner, W. M. (1981) *J. Biol. Chem.* **256,** 4669–4675
18. Cleveland, D. W., Fischer, S. G., Kirschner, M. W., and Laemmli, U. K. (1977) *J. Biol. Chem.* **252,** 1102–1106
19. Lindahl, T., and Barnes, D. E. (1992) *Annu. Rev. Biochem.* **61,** 251–281
20. Soderhall, S., and Lindahl, T. (1975) *J. Biol. Chem.* **250,** 8438–8444
21. Kodama, K., Barnes, D. E., and Lindahl, T. (1991) *Nucleic Acids Res.* **19,** 6093–6099
22. Takahashi, M., and Tomizawa, K. (1990) *Eur. J. Biochem.* **192,** 735–740
23. Noguiez, P., Barnes, D. E., Mohrenweiser, H. W., and Lindahl, T. (1992) *Nucleic Acids Res.* **20,** 3845–3850
24. Elder, R. H., Montecucco, A., Ciarrocchi, G., and Rossignol, J.-M. (1992) *Eur. J. Biochem.* **203,** 53–58