THE ATP-OPERATED CLAMP OF HUMAN DNA TOPOISOMERASE IIα

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Spencer Campbell BSc MPhil (Res)
Department of Biochemistry
University of Leicester

April 2001
To my parents
ACKNOWLEDGEMENTS

The work presented in this thesis was carried out in the laboratory of Professor A. Maxwell in the Department of Biochemistry, University of Leicester.

I would like to thank Tony Maxwell for his advice and guidance throughout my studies, and particularly during the last six months whilst I have been writing up. I am indebted to Laurence Gardiner for his assistance at the commencement of my PhD, and to Tim Hammonds for valuable advice and suggestions. Many thanks to Alison Howells for her scientific support, supplies of DNA, and birthday cakes. Thanks to all past and present members of the lab for advice and making the past three years an enjoyable experience: Jonathan Heddle for his scientific advice and valuable friendship; Lois Wentzell and Symon Erskine for advice and support; Nicola Williams and Christian Noble for being my bay-buddies. Thanks also to Melisa Wall, Susan Hockings, Susan Trigwell, Mike Smyth, Sotiris Kampranis, Clare Smith and Steve Blance for advice. I am grateful to Andrew Leech for his help with analytical ultracentrifugation experiments, and Kathryn Lilley for N-terminal sequencing of proteins.

I acknowledge the financial support of the BBSRC.
ABBREVIATIONS

A$_{260}$  Absorbance at 260 nm
ADP  Adenosine diphosphate
ADPNP  Adenosine-β,γ-imidodiphosphate
ATP  Adenosine triphosphate
ATPase  Adenosine triphosphatase
bp  Base pair
BSA  Bovine serum albumin
DNA  Deoxyribonucleic acid
DNTP  Deoxynucleotide triphosphate
ds  Double stranded
DTT  Dithiothreitol
E. coli  *Escherichia coli*
EDTA  Ethylenediaminetetraacetic acid
GyrA  Gyrase A subunit
GyrB  Gyrase B subunit
Hsp  Heat shock protein
IPTG  Isopropyl-β-D-thiogalactopyranoside
kb  Kilobase pairs
kDa  Kilodaltons
LB  Luria-Bertani broth
Lk  Linking number
PAGE  Polyacrylamide gel electrophoresis
PCR  Polymerase chain reaction
PEP  Phosphoenol pyruvate
PK/LDH  Pyruvate kinase/lactate dehydrogenase
RNA  Ribonucleic acid
SAR  Scaffold-attachment/associated region
*S. cerevisiae*  *Saccharomyces cerevisiae*
SDS  Sodium dodecyl sulphate
ss  Single stranded
TEMED  N, N, N', N'-tetramethylethylenediamine
Topo  Topoisomerase
Tris  Tris[hydroxymethyl]aminomethane
ABSTRACT

THE ATP-OPERATED CLAMP OF HUMAN DNA TOPOISOMERASE IIα

SPENCER CAMPBELL

The type II DNA topoisomerases operate via a complex multistep mechanism: one segment of DNA (the T segment) is captured by the N-terminal domains and directed through a transient break in another DNA segment held by the core of the enzyme (the G segment). In most reactions of type II enzymes, this process is dependent upon the binding and hydrolysis of ATP by the N-terminal domains. I have examined the catalytic properties of the N-terminal ATPase domain of human DNA topoisomerase IIα.

The intrinsic ATPase activity of this N-terminal domain is stimulated by DNA in a length-dependent manner. A DNA fragment of 10 base pairs is sufficient to stimulate the ATPase activity of this enzyme. With pBR322 as the DNA cofactor, the ATPase activity is hyperstimulated at a base pair to enzyme dimer ratio of ~50:1. With short DNA fragments of 28-44 base pairs, the hyperstimulation of the ATPase activity is maximal with a 38 base pair DNA fragment. This hyperstimulation phenomenon was previously attributed to interactions between adjacent enzyme dimers on a DNA strand, and we have interpreted our results in the light of this model. The physiological significance of this hyperstimulated ATPase activity is not yet known.

The ability of the N-terminal domain to bind DNA and to dimerise in the presence and absence of nucleotide has been investigated. The enzyme is a monomer in solution, but in the presence of ADPNP a large conformational shift occurs, such that the enzyme associates into a dimer and a higher order species, possibly a tetramer. The N-terminal domain binds DNA with a considerably lower affinity than the full-length enzyme, and the interaction with DNA is significantly increased by the presence of ADPNP i.e. in the dimeric state. The presence of DNA does not appear to cause the enzyme to dimerise, nor does it induce a large conformational change in the enzyme.

Taken together, these results help to clarify the role of the N-terminal domain in the catalytic cycle of human DNA topoisomerase IIα. The hyperstimulation phenomenon, previously seen with the full-length enzyme, is now known to be a property of the N-terminal domain, and suggests a significant role for the T segment in the ATPase cycle of this enzyme.
## TABLE OF CONTENTS

Acknowledgements ............................................................................................................. iii  
Abbreviations ...................................................................................................................... iv  
Abstract ................................................................................................................................. v  
Table of contents ................................................................................................................. vi  
List of figures ....................................................................................................................... x  

### CHAPTER 1 - INTRODUCTION

1.1 ENERGY TRANSDUCTION IN BIOLOGICAL SYSTEMS ............... 2  
1.1.1 DNA-targeted enzymes ................................................................. 3  
   1.1.1.1 The RecA protein of *E. coli* ......................................................... 4  
   1.1.1.2 The MutL DNA mismatch repair protein ................................. 6  
1.1.2 The GHKL ATPase superfamily ................................................. 7  
1.2 DNA STRUCTURE AND TOPOLOGY ................................................. 8  
   1.2.1 DNA supercoiling ........................................................................ 9  
   1.2.2 DNA knots and catenanes ........................................................... 10  
1.3 DNA TOPOISOMERASES ................................................................. 11  
   1.3.1 The roles of DNA topoisomerases ............................................. 12  
   1.3.2 Type I DNA topoisomerases ....................................................... 16  
      1.3.2.1 Type IA topoisomerases ......................................................... 17  
      1.3.2.2 Type IB topoisomerases ......................................................... 19  
   1.3.3 Type II DNA topoisomerases ..................................................... 21  
      1.3.3.1 *E. coli* DNA gyrase .............................................................. 22  
      1.3.3.2 Yeast topoisomerase II ......................................................... 24  
      1.3.3.3 Human type II topoisomerases ........................................... 26  
1.4 CATALYTIC CYCLE OF TYPE II DNA TOPOISOMERASES .......... 29  
   1.4.1 A model for the catalytic cycle of type II topoisomerases ............. 29  
   1.4.2 Introduction of negative supercoils by DNA gyrase ..................... 32  
   1.4.3 The role of ATP hydrolysis in the catalytic cycle ......................... 33  
1.5 THE ATPase REACTION OF TYPE II TOPOISOMERASES .......... 36  
   1.5.1 DNA gyrase ................................................................................. 36  
   1.5.2 Eukaryotic topoisomerase II ...................................................... 38  
   1.5.3 Coupling of ATP usage to T segment capture and strand passage ... 40  
1.6 ROLE OF THE N-TERMINAL ATPase DOMAIN ....................... 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6.1 ATP binding and hydrolysis</td>
<td>45</td>
</tr>
<tr>
<td>1.6.2 Capture of a DNA segment</td>
<td>47</td>
</tr>
<tr>
<td>1.7 AIMS</td>
<td>48</td>
</tr>
<tr>
<td><strong>CHAPTER 2 - MATERIALS AND METHODS</strong></td>
<td></td>
</tr>
<tr>
<td>2.1 BACTERIOLOGY</td>
<td>52</td>
</tr>
<tr>
<td>2.1.1 Bacterial strains</td>
<td>52</td>
</tr>
<tr>
<td>2.1.2 Bacterial growth media</td>
<td>52</td>
</tr>
<tr>
<td>2.1.3 Antibiotics</td>
<td>53</td>
</tr>
<tr>
<td>2.1.4 Transformation of competent <em>E. coli</em> cells</td>
<td>53</td>
</tr>
<tr>
<td>2.2 SYNTHESIS AND PURIFICATION OF OLIGONUCLEOTIDES</td>
<td>54</td>
</tr>
<tr>
<td>2.2.1 Oligonucleotides used for ligation-independent cloning</td>
<td>54</td>
</tr>
<tr>
<td>2.2.2 Oligonucleotides used for DNA sequencing</td>
<td>55</td>
</tr>
<tr>
<td>2.2.3 Oligonucleotides used for production of a 140 bp DNA fragment</td>
<td>55</td>
</tr>
<tr>
<td>2.3 PREPARATION OF DNA SUBSTRATES</td>
<td>56</td>
</tr>
<tr>
<td>2.3.1 Determination of DNA concentration</td>
<td>56</td>
</tr>
<tr>
<td>2.3.2 Preparation of plasmid pBR322</td>
<td>56</td>
</tr>
<tr>
<td>2.3.3 Preparation of a 140 bp DNA fragment</td>
<td>57</td>
</tr>
<tr>
<td>2.3.4 Preparation of short DNA fragments</td>
<td>57</td>
</tr>
<tr>
<td>2.4 GEL ELECTROPHORESIS OF DNA</td>
<td>58</td>
</tr>
<tr>
<td>2.4.1 Agarose gel electrophoresis</td>
<td>58</td>
</tr>
<tr>
<td>2.4.2 Polyacrylamide gel electrophoresis</td>
<td>58</td>
</tr>
<tr>
<td>2.5 LIGATION-INDEPENDENT CLONING</td>
<td>59</td>
</tr>
<tr>
<td>2.5.1 Ligation-independent cloning</td>
<td>59</td>
</tr>
<tr>
<td>2.5.2 Isolation of plasmid DNA</td>
<td>61</td>
</tr>
<tr>
<td>2.5.3 Restriction digests</td>
<td>61</td>
</tr>
<tr>
<td>2.5.4 DNA sequencing</td>
<td>61</td>
</tr>
<tr>
<td>2.6 EXPRESSION AND PURIFICATION OF 52 kDa PROTEIN</td>
<td>61</td>
</tr>
<tr>
<td>2.7 EXPRESSION AND PURIFICATION OF FUSION PROTEINS</td>
<td>63</td>
</tr>
<tr>
<td>2.7.1 Proteolytic cleavage of fusion proteins with Factor Xa</td>
<td>64</td>
</tr>
<tr>
<td>2.8 PROTEIN ANALYSIS</td>
<td>65</td>
</tr>
<tr>
<td>2.8.1 Determination of protein concentration</td>
<td>65</td>
</tr>
<tr>
<td>2.8.2 SDS-PAGE gel electrophoresis of proteins</td>
<td>65</td>
</tr>
<tr>
<td>2.8.3 Western blotting</td>
<td>66</td>
</tr>
<tr>
<td>2.8.4 N-terminal sequencing of proteins</td>
<td>67</td>
</tr>
<tr>
<td>2.9 ATPase ASSAYS</td>
<td>67</td>
</tr>
</tbody>
</table>
2.10 PROTEIN CROSSLINKING .............................................................. 69
2.11 DNA BINDING ASSAYS .............................................................. 69
   2.11.1 Filter binding ......................................................................... 69
   2.11.2 Gel retardation ........................................................................ 70
2.12 LIMITED PROTEOLYSIS ............................................................... 70
2.13 MOLECULAR WEIGHT STUDIES .................................................. 70
   2.13.1 Gel filtration ........................................................................... 70
   2.13.2 Analytical ultracentrifugation ................................................. 71

CHAPTER 3 - EXPRESSION AND PURIFICATION OF THE N-TERMINAL
DOMAIN OF DNA TOPOISOMERASE IIα AS A FUSION PROTEIN
3.1 INTRODUCTION ............................................................................... 73
3.2 RESULTS .......................................................................................... 76
   3.2.1 Fusion proteins with an Enterokinase cleavage site ................. 78
      3.2.1.1 Cloning and expression of Ek/LIC fusion proteins .......... 78
      3.2.1.2 Purification of fusion proteins ........................................... 81
         3.2.1.2.1 Immobilised metal affinity chromatography ............... 82
         3.2.1.2.2 Ion-exchange chromatography ................................. 84
      3.2.1.3 Cleavage of fusion proteins with Enterokinase ............... 85
   3.2.2 Fusion proteins with a Factor Xa cleavage site ....................... 87
      3.2.2.1 Cloning and expression of Xa/LIC fusion proteins ........... 87
      3.2.2.2 Optimisation of expression and purification ..................... 89
      3.2.2.3 Characterisation of fusion proteins of the N-terminal
         domain ...................................................................................... 93
      3.2.2.4 Cleavage of fusion proteins with Factor Xa ..................... 94
   3.2.3 Initial characterisation of the ATPase activity of the N-terminal
      fragments topoIIα-(1-420) and topoIIα-(1-453) ......................... 95
3.3 DISCUSSION .................................................................................. 96

CHAPTER 4 - DNA-DEPENDENT ATPase ACTIVITY OF THE N-TERMINAL
FRAGMENT TOPOIIα-(1-453)
4.1 INTRODUCTION .............................................................................. 101
4.2 RESULTS ........................................................................................ 101
   4.2.1 Initial characterisation of the ATPase activity of the N-terminal
      fragment topoIIα-(1-453) .............................................................. 102
      4.2.1.1 ATPase activity as a function of salt concentration ......... 102
# LIST OF FIGURES

## Chapter 1

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Mechanism of allosteric interaction between ATP hydrolysis and DNA binding in RecA</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>Schematic representation of the ATP-binding Bergerat fold in the GHKL ATPase superfamily</td>
<td>7</td>
</tr>
<tr>
<td>1.3</td>
<td>Topological forms of double-stranded DNA</td>
<td>9</td>
</tr>
<tr>
<td>1.4</td>
<td>The ‘enzyme-bridging’ model for the reaction mechanism of <em>E. coli</em> topoisomerase I</td>
<td>18</td>
</tr>
<tr>
<td>1.5</td>
<td>The ‘controlled rotation’ model for the reaction mechanism of human topoisomerase I</td>
<td>20</td>
</tr>
<tr>
<td>1.6</td>
<td>Domain structure of <em>E. coli</em> DNA gyrase, yeast DNA topoisomerase II and human DNA topoisomerase IIα</td>
<td>21</td>
</tr>
<tr>
<td>1.7</td>
<td>Crystal structure of the 59 kDa N-terminal fragment of the <em>E. coli</em> DNA gyrase GyrA protein</td>
<td>22</td>
</tr>
<tr>
<td>1.8</td>
<td>Crystal structure of the 43 kDa N-terminal fragment of the <em>E. coli</em> DNA gyrase GyrB protein</td>
<td>23</td>
</tr>
<tr>
<td>1.9</td>
<td>Crystal structure of the 92 kDa core fragment of yeast topoisomerase II</td>
<td>25</td>
</tr>
<tr>
<td>1.10</td>
<td>Schematic representation of the reaction mechanism of eukaryotic topoisomerase II</td>
<td>30</td>
</tr>
<tr>
<td>1.11</td>
<td>Model for the mechanism of strand passage by DNA gyrase</td>
<td>32</td>
</tr>
<tr>
<td>1.12</td>
<td>Hydrolysis of one ATP drives DNA transport by eukaryotic topoisomerase II</td>
<td>35</td>
</tr>
</tbody>
</table>

## Chapter 3

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Sequence alignment of the N-terminus of type II DNA topoisomerasers</td>
<td>73</td>
</tr>
<tr>
<td>3.2</td>
<td>Expression and purification of the 52 kDa protein</td>
<td>74</td>
</tr>
<tr>
<td>3.3A</td>
<td>Overview of the Ligation Independent Cloning (LIC) strategy for</td>
<td>76</td>
</tr>
</tbody>
</table>
3.3B Schematic representation of the pET-32 Ek/LIC and pET-32 Xa/LIC vector products

3.4 Plasmid maps of Ek/LIC vectors

3.5A Expression of Ek/LIC fusion proteins at 37°C

3.5B Soluble and insoluble protein from expression of fusion proteins

3.6 Purification of 1-420F(Ek) with TALON resin

3.7 Purification of 1-420F(Ek) on HiTrap SP column

3.8 ATPase activity of 1-420F(Ek)

3.9A Proteolytic cleavage of 1-420F(Ek) with recombinant enterokinase

3.9B Proteolytic cleavage of 1-420F(Ek) in the presence of ADPNP

3.10 Plasmid maps of Xa/LIC vectors

3.11A Expression of Xa/LIC fusion proteins at 37°C

3.11B Soluble and insoluble protein from expression of fusion proteins

3.12A Expression of Xa/LIC fusion proteins at 20°C

3.12B Soluble and insoluble protein from expression of fusion proteins

3.13A Expression of 1-420F in BL21 and BL21-CodonPlus cells

3.13B Soluble and insoluble protein from expression of fusion proteins

3.14 Purification of 1-420F from BL21 and BL21-CodonPlus cells

3.15 ATPase activity of fusion proteins

3.16 Proteolytic cleavage of 1-420F with Factor Xa

3.17 Summary of the expression and purification of proteins topoIIα-(1-420) and topoIIα-(1-453)

3.18 ATPase activity of fusion and cleaved proteins

Chapter 4

4.1A ATPase activity of topoIIα-(1-453) as a function of NaCl concentration

4.1B ATPase activity of topoIIα-(1-453) as a function of NaCl concentration

4.2A ATPase activity of topoIIα-(1-453) as a function of enzyme concentration
4.2B Schematic model for the ATPase cycle of the *E. coli* DNA gyrase 43 kDa protein

4.2C ATPase activity of topoIIα-(1-453) as a function of enzyme concentration in the presence of DNA

4.3A ATPase activity of topoIIα-(1-453) as a function of ATP concentration

4.3B ATPase activity of topoIIα-(1-453) as a function of ATP concentration in the presence of DNA

4.4 ATPase activity of topoIIα-(1-420) as a function of ATP concentration in the presence of DNA

4.5 Demonstration of the lag in the rate of ATP hydrolysis by topoIIα-(1-453) in the presence of DNA

4.6 Effect of preincubation on the lag in ATP hydrolysis by topoIIα-(1-453) in the presence of DNA

4.7 ATPase activity of topoIIα-(1-453) as a function of DNA concentration

4.8 ATPase activity of topoIIα-(1-420) as a function of DNA concentration

4.9 ATPase activity of topoIIα-(1-453) as a function of DNA fragment concentration

4.10 ATPase activity of topoIIα-(1-453) as a function of the concentration of a 140 bp DNA fragment

4.11 ATPase activity of topoIIα-(1-453) as a function of short DNA fragment concentration

4.12 ATPase activity of topoIIα-(1-453) at 25°C as a function of short DNA fragment concentration

4.13A ATPase activity of topoIIα-(1-453) as a function of long DNA fragment concentration

4.13B ATPase activity of topoIIα-(1-453) as a function of long DNA fragment concentration

4.14 Schematic model for the DNA-dependent ATPase activity of the N-terminal fragment topoIIα-(1-453)
4.15 Alternative model for the hyperstimulated ATPase activity of the N-terminal fragment topoIIα-(1-453)

Chapter 5

5.1 Cross-linking of the N-terminal fragment topoIIα-(1-453) using Dimethyl suberimidate (DMS)

5.2 Cross-linking of the N-terminal fragment topoIIα-(1-453) using Bis(sulfo)succinimidyl suberate (BS³)

5.3 Cross-linking of the N-terminal fragment topoIIα-(1-453) using Disuccinimidyl glutarate (DSG)

5.4 Gel filtration of the N-terminal fragment topoIIα-(1-453) in the presence and absence of ADPNP

5.5 Calibration curve for gel filtration

5.6 Filter binding assay for DNA binding by the N-terminal fragment topoIIα-(1-453)

5.7 Gel retardation assay using a 140 bp DNA fragment

5.8 Limited proteolysis of the fragment topoIIα-(1-453)

5.9 Summary gel for limited proteolysis of the fragment topoIIα-(1-453) without ligand, and in the presence of ATP or ADPNP.
CHAPTER 1

Introduction
1.1 ENERGY TRANSDUCTION IN BIOLOGICAL SYSTEMS

The adoption of the purine nucleotide adenosine triphosphate (ATP) as the central energy currency in living cells presumably occurred at an early stage of evolution, given its widespread use in biological systems (reviewed in Baltscheffsky, 1996). However, there is evidence that suggests that inorganic pyrophosphate (PPi) acted as an early energy carrier (Miller & Parris, 1964). Inorganic pyrophosphate is only one half to two thirds as energy rich as ATP; the hydrolysis of the γ-phosphate from ATP releases ~30 kJ/mol of free energy under standard conditions, or ~55 kJ/mol under normal cellular conditions (Harris, 1995). A living organism has a huge requirement for ATP, with a typical human synthesising an amount approaching their body mass on a daily basis (Harris, 1995). The principle means for the production of ATP is through the oxidation of organic nutrients via glycolysis, the citric acid cycle and hydrogen transport. The production of ATP by the ATP synthase enzyme occurs by a coupled process of proton translocation and ATP synthesis (Mitchell, 1979). The enzyme serves to link the flow of protons down an electrochemical gradient to the phosphorylation of ADP to ATP, the two processes being coupled so that they occur simultaneously. The production of ATP in this manner represents a form of cellular energy transduction, the essential characteristic of which is that an energetically favourable process is linked to, and in fact drives, an energetically unfavourable process (Rees & Howard, 1999).

Energy transduction processes are responsible for driving many diverse reactions within cells e.g. biosynthetic reactions, motility, vesicular transport and other metabolic processes. Cellular energy sources such as ATP (and other nucleoside triphosphates) or ion gradients are used to drive these energetically (or kinetically) unfavourable reactions to completion. The coupling process is thought to follow a set of rules (Jencks, 1980), which allows for the codependence of a step in one reaction on a step in the other reaction, and thus disfavours uncoupling. Generally speaking, enzymes which couple the hydrolytic bond energy of ATP hydrolysis to drive another reaction can be divided into two groups (Tanford, 1983). The first group operates via a phosphorylated chemical intermediate. For example, in the case of cytidine triphosphate (CTP) synthetase from Escherichia coli (E. coli), the reaction proceeds first by the phosphorylation of uridine triphosphate (UTP) on the 4-oxygen atom by ATP, followed
by the attack of ammonia on the 4-oxygen atom which leads to the formation of CTP (Von der Saal et al., 1985). The second group does not involve substrate phosphorylation, and instead uses the energy of ATP binding and hydrolysis to drive the protein through an ordered series of conformational changes, which impart directionality and irreversibility on the reaction. The occurrence of an ATPase-driven conformational cascade is commonplace amongst large multisubunit protein complexes or ‘molecular machines’, such as helicases in the deoxyribonucleic acid (DNA) replication apparatus and ribonucleic acid (RNA)-dependent ATPases within the spliceosome (reviewed in Alberts, 1984, 1998; Alberts & Miake-Lye, 1992). The energy of ATP hydrolysis may also be utilised to drive a reaction away from its equilibrium position. For example, in the presence of ATP, the enzyme nicotinic acid phosphoribosyltransferase generates a new steady state in which a thermodynamically unfavourable ratio of products:substrates is maintained (Vinitsky & Grubmeyer, 1993).

Proteins which couple ATPase activity to conformational changes can be further subdivided according to the biological roles they fulfill (Alberts, 1992). One such classification for ATPases is as motor proteins, which use the energy of ATP hydrolysis to translocate along a filament. Motor proteins such as myosins, kinesins and dyneins move along actin filaments or microtubules, and helicases translocate along single-stranded DNA or RNA strands. A second possible role for ATPases is as molecular clocks or timing devices that increase the fidelity of biological reactions. A third role for ATPases is to use ATP binding and hydrolysis to drive the enzyme through a cycle of affinity changes for binding and release of ligands and subsequent recycling of the enzyme. Examples of ATP-induced affinity changes in enzymes include the chaperones GroEL (Martin et al., 1991) and hsp70 (Pelham, 1986), nicotinic acid phosphoribosyltransferase (Rajavel et al., 1996), RecA (Kowalczykowski & Krupp, 1995), DNA helicases (Wong & Lohman, 1992) and DNA gyrase (Cullis et al., 1992).

1.1.1 DNA-targeted enzymes

There are many examples of enzymes that use the binding and hydrolysis of ATP to catalyse a reaction in which DNA represents a substrate or ligand. The
unwinding of double-stranded DNA (dsDNA) is catalysed by DNA helicases, which use the energy from the binding and hydrolysis of ATP to unwind the duplex DNA, and in many cases to translocate unidirectionally along the DNA strand (Lohman, 1993). During DNA replication, a ‘clamp loader’ complex uses the energy of ATP hydrolysis to assemble a circular sliding protein clamp around DNA, which in turn tethers the DNA polymerase enzyme to the DNA and allows for processive DNA replication (Baker & Bell, 1998). The clamp loader belongs to a class of ATPases known as ‘molecular matchmakers’ (Sancar & Hearst, 1993). These proteins use the energy of ATP hydrolysis to cause conformational changes in proteins that promote the formation of a complex between two or more macromolecules. The type II DNA topoisomerases are a class of enzyme which use the energy from the binding and hydrolysis of ATP to catalyse the interconversion of topological isomers of DNA (Wang, 1998). The following sections look at two other examples of DNA-targeted ATPases: RecA, which uses ATP to cycle through different affinity states for DNA during recombination; and MutL, which acts as a switch protein to coordinate DNA mismatch repair.

1.1.1.1 The RecA protein of E. coli

The E. coli RecA protein is involved in homologous DNA recombination and repair. The DNA strand exchange reaction comprises three stages: binding of RecA to single-stranded DNA (ssDNA) in an ATP-dependent manner (presynapsis); binding of dsDNA to the RecA-ssDNA complex (synapsis) and homologous pairing leading to joint molecule formation; and extension of the heteroduplex region by ‘branch migration’ to yield the gapped dsDNA products (Kowalczykowski, 1987). The RecA protein is a DNA-dependent ATPase, but only ATP binding and not hydrolysis is required for DNA strand exchange leading to the formation of joint molecules. This has been shown using the nearly non-hydrolysable ATP analogue adenosine 5'-[γ-thio]triphosphate (ATP[γ]S) (Menetski et al., 1990), and also using a noncovalent complex of adenosine diphosphate (ADP) and AlF4– (Kowalczykowski & Krupp, 1995). These results are interpreted in terms of ATP binding inducing an allosteric transformation in the enzyme, which forms a high-affinity DNA binding state that can stabilise a transition state essential for DNA strand exchange. The free energy derived
Figure 1.1 Mechanism of allosteric interaction between ATP hydrolysis and DNA binding in RecA.
The Walker A motif (\^{66}GPEGGKT\^{73}) contains a lysine (Lys72) that interacts with the phosphate groups of the nucleotide. Through its interaction with Gln194, the \(\gamma\)-phosphate of ATP is thought to stabilise a conformation of loop L2 (and helix G) that can bind ssDNA. Hydrolysis of ATP, by nucleophilic attack from a water molecule activated by Glu96, destroys these interactions and leads to generation of products. Therefore, the presence or absence of the \(\gamma\)-phosphate of ATP alters the conformation and binding properties of the region involved in DNA binding.
(Redrawn from Story & Steitz, 1992).
from ATP binding presumably overcomes the activation energy barrier required to form this intermediate. The role of ATP hydrolysis appears to be to generate the 'release factor' ADP, which induces dissociation of the RecA-DNA product complex by generation of a low-affinity DNA binding state. The hydrolysis of ATP may also play other roles e.g. impart directionality on the reaction. Therefore, by cycling through a high affinity vs. low affinity DNA-binding state mediated by ATP binding and hydrolysis, the RecA protein is able to promote first DNA pairing and exchange, and then product release.

Structural information on the interaction of RecA with nucleotides and DNA has provided further insight into the mechanism for allosterically coupling ATP binding and hydrolysis to DNA strand exchange. The crystal structure of the RecA protein in the presence of ADP (Story & Steitz, 1992) reveals that a loop (66-PESSGKT73), whose amino acid sequence matches the consensus sequence for a nucleotide-binding 'Walker A' motif (Walker et al., 1982) or phosphate-binding or 'P-loop' (Saraste et al., 1990), interacts with the α- and β-phosphates of ADP. Although not visualised in the crystal, molecular modelling suggests a role for Glu96 as a general base that activates a water molecule for attack on the γ-phosphate in the ATP hydrolysis reaction. The crystal structure of the RecA protein (Story et al., 1992) reveals two disordered regions forming loops, i.e. L1 (157-164) and L2 (195-209), that are implicated as DNA binding sites in the RecA protein. The close proximity of these regions in the major domain of the RecA protein is suggestive of an allosteric mechanism for ATP-mediated DNA binding (Story & Steitz, 1992; see Figure 1.1). In this mechanism, Gln194 interacts with the γ-phosphate of ATP to stabilise a conformation of loop L2 (and the following helix, helix G) that can bind ssDNA. Hydrolysis of ATP destroys the interactions between Gln194 and the nucleotide, causing a conformational shift back to the low affinity state that allows product release. More recently, photochemical crosslinking has been used to identify DNA binding sites within the RecA protein (Rehrauer & Kowalczykowski, 1996). Although loop L2 itself was not identified as a DNA binding site by photochemical crosslinking, a peptide corresponding to amino acids 61-72 of RecA was found to interact with DNA. This region contains the 'P-loop' for interaction with ATP,
which may therefore be directly involved in DNA binding mediated by ATP binding and hydrolysis.

1.1.1.2 The MutL DNA mismatch repair protein

The MutL protein of *E. coli* is involved in repairing DNA base pair mismatches, which occur primarily during DNA replication and recombination. Mismatch repair is initiated by MutL, MutS and MutH, which identify a mismatched base pair within a DNA strand and introduce an incision, and is then completed by DNA exonucleases, helicases and polymerases, which remove nucleotides from the nick and fill the gap (reviewed in Modrich, 1991). The MutL protein is an ATPase (Ban & Yang, 1998), and the rate of ATPase activity is stimulated ~3-fold by dsDNA and ~7-fold by ssDNA (Ban *et al*., 1999). The crystal structure of a 40 kDa N-terminal fragment of *E. coli* MutL (LN40) has been determined in the presence of ADP and a non-hydrolysable analogue of ATP, 5'-adenylyl-β,γ-imidophosphate (ADPNP) (Ban *et al*., 1999), and in the apoprotein form (Ban & Yang, 1998). This fragment shows extensive structural homology with an N-terminal 43 kDa fragment of the *E. coli* DNA gyrase GyrB protein (Wigley *et al*., 1991; see section 1.1.2). This 40 kDa fragment can hydrolyse ATP, albeit at a 10-fold lower rate than full-length MutL (Ban *et al*., 1999). Binding of ADPNP to LN40 induces a shift from a monomer to dimer form, as seen in the crystal structure. The intact MutL protein, which exists as a stable dimer due to the C-terminal domains, undergoes a shift from a large to a smaller Stoke's radius when bound to ADPNP, indicative of an ADPNP-induced interaction of the N-terminal (LN40) region. The crystal structure of the LN40-ADPNP complex reveals a groove that has been suggested to be a DNA binding surface because of its positive surface potential. The presence of the C-terminal domains of the MutL protein would convert this groove into a hole, capable of binding ssDNA, but not dsDNA, based on the dimensions of the hole. Thus, the MutL protein may operate as an ATP-operated clamp for capturing a ssDNA strand (Ban *et al*., 1999). The conformational change in MutL brought about by ADPNP is also sufficient to cause an interaction with, and activation of, the endonuclease MutH (Ban & Yang, 1998). In this regard, MutL can be thought of as
another example of a 'molecular matchmaker', coordinating the assembly of the mismatch repair complex on the DNA strand (Sancar & Hearst, 1993).

1.1.2 The GHKL ATPase superfamily

The elucidation of the three-dimensional structure of three ATP-binding protein fragments from *E. coli* GyrB (Wigley et al., 1991), *E. coli* MutL (Ban & Yang, 1998) and yeast and human Hsp90 (Prodromou et al., 1997a, 1997b; Stebbins et al., 1997), have revealed the presence of a novel ATP-binding fold, first postulated by Bergerat et al. (1997). All three proteins are capable of hydrolysing ATP, which led to them being grouped under the GHL ATPase superclass (Ban et al., 1999). Subsequently, the crystal structure of two histidine kinases (Tanaka et al., 1998; Bilwes et al., 1999) also revealed the presence of the same 'Bergerat' ATP-binding fold, and thus the GHL ATPase class was expanded to the GHKL ATPase/kinase superfamily (Dutta & Inouye, 2000). Remarkably, the four enzyme types which share this unique nucleotide-binding module have diverse biological functions: DNA gyrase (and other type II topoisomerases) and MutL both manipulate dsDNA, for control of DNA topology and DNA repair respectively; Hsp90 is a molecular chaperone which accommodates unfolded polypeptides; histidine kinases act as signal sensors and transducers in prokaryotes and lower eukaryotes.

The ATP-binding Bergerat fold consists of an α/β sandwich structure incorporating three α-helices and four β-strands, arranged to form the structural framework of the ATP-binding pocket (Figure 1.2). Four highly conserved motifs are located within the fold: Motif I (ExxxNxxD), Motif II (DxGxG), Motif III (GxxGxG/A) and Motif IV (Tx₄GT). With the exception of Motif I, the conserved motifs are found in surface loops that connect the structural elements, and the conserved amino acid residues within these motifs are involved in making direct contact to the ATP molecule or in maintaining the structural integrity of the pocket. The conserved asparagine of Motif I (the N box) coordinates a bound Mg²⁺ ion, which links all three phosphate groups of ATP to the protein, and the conserved glutamate (not present in histidine kinases) acts as a general base for water activation in ATP hydrolysis. The conserved
Figure 1.2 Schematic representation of the ATP-binding Bergerat fold in the GHKL ATPase superfamily.

The core elements of the novel ATP-binding Bergerat fold (Bergerat et al., 1997) are shown, based on the crystal structures of the DNA gyrase B protein, MutL, EnvZ and Hsp90. The fold has an α/β sandwich structure consisting of three α-helices and a four-stranded mixed β-sheet. The four conserved motifs of the Bergerat ATP-binding fold: N (Motif I), G1 (Motif II), G2 (Motif III) and G3 (Motif IV), are represented by orange rectangles. The ATP-binding pocket is bordered by helices α1, 2 and 3, and loops 3 and 4 (the ATP-lid); and the bottom of the pocket is formed by β-sheets β2, 3 and 4. Structural variations in the Bergerat ATP-binding fold within the GHKL superfamily localise to the loop regions (blue triangles).

(Redrawn from Dutta & Inouye, 2000).
aspartic acid of Motif II (G1 box) forms a hydrogen bond with the adenine-moiety of the ATP molecule, and accounts for the specificity for ATP over GTP. Conserved glycine residues in Motifs II and III (G1 and G2 boxes) form the two flexible hinges of an intervening loop referred to as the ‘ATP-lid’. The ATP-lid represents the most variable region of the Bergerat ATP-binding fold; its composition and conformation varies within the GHKL family, and according to the state of nucleotide bound. Motif III represents the phosphate-binding P-loop (Saraste et al., 1990); the glycine residues interact with the α- and γ-phosphates of ATP. The second threonine of Motif IV (G3 box) interacts with the adenine of ATP, and the preceding threonine makes hydrogen bonds with the isoleucine following Motif II to stabilise the ATP pocket.

Another commonality amongst the members of the GHKL superfamily is the presence of a dimerisation domain linked to the ATP-binding domain. In the case of DNA gyrase and MutL, binding of ATP is known to induce dimerisation of the N-terminal domains of these proteins, forming a ‘molecular clamp’ for the capture of DNA (Wigley et al., 1991; Roca & Wang, 1992; Ban et al., 1999). The crystal structure of the N-terminal domain of yeast Hsp90 also reveals the presence of a molecular clamp (Prodromou et al., 1997b), and ATP binding induces large conformational changes in the structure of Hsp90, indicative of a shift from an open to closed conformation (Csermely et al., 1993). Histidine kinases represent the most divergent of the GHKL superfamily, and are not recognised as true ATPases per se. However, histidine kinases presumably use ATP-dependent intra- and inter-subunit conformational changes to effect phosphoryltransfer to a response regulator protein (Bilwes et al., 1999).

1.2 DNA STRUCTURE AND TOPOLOGY

The double-helical arrangement of DNA, as elucidated by Watson and Crick (1953), is sufficient to account for the known properties of DNA. Further study showed that dsDNA can adopt a variety of helical conformations, and indeed some structures which deviate from the classic double helix e.g. cruciforms and Holliday junctions. These structural variations represent changes in the primary structure of DNA. In
addition to these, variations in the higher order structure of DNA are found in the cell. These variations in DNA structure, or DNA topology, represent the most significant structural alterations of DNA in vivo (reviewed in Bates & Maxwell, 1993). The most important of these structural variations is the supercoiling of DNA.

1.2.1 DNA supercoiling

An important property of a closed-circular DNA molecule is that it can coil around itself to form a more compact structure (Figure 1.3). This ‘supercoiling’ of circular DNA introduces torsional stress into the molecule, which can only be relieved by the breakage of one or both strands.

The topological state of a DNA circle can be defined by a series of parameters. The ‘linking number’ ($L_k$) is described as the number of times one strand of DNA winds around the other in the right-handed direction. If a linear double-stranded DNA molecule is closed into a circle, the two strands will be linked together a number of times equal to the number of double-stranded turns in the linear DNA molecule. Thus, the double-helical arrangement of DNA contributes to the value of the linking number, which must be an integer. Any closed-circular DNA molecule in a relaxed state will have a linking number which corresponds to the length of the DNA molecule in base pairs divided by the number of base pairs per turn of helix (10.5 bp/turn under standard conditions). This number is not always an integer, as a slight twisting of the DNA molecule may be required to align the strand ends. The linking number is taken as the closest integer to this value. This is the ‘standard linking number’ ($L_{km}$) of the DNA molecule.

Introducing supercoils into a DNA molecule changes the number of helical turns within the molecule, and is associated with a change in the linking number. Changes in linking number can only be made by breaking and rejoining one or both strands of DNA. Double-stranded DNA is a right-handed helix, breaking and unwinding of the helix (i.e. in the left-handed direction) leads to a decrease in the linking number and is termed ‘negative supercoiling’, whereas overwinding (i.e. in the
Figure 1.3 Topological forms of double-stranded DNA.
All these interconversions can be carried out by the passage of one double-stranded DNA segment through another.
(Bates & Maxwell, 1993).
right-handed direction) leads to an increase in linking number and is called 'positive supercoiling'. The change in linking number from $L_{km}$ ($L_k - L_{km}$) is called the linking difference, $\Delta L_k$. The change in linking number can also be normalised to the length of the DNA molecule to give the 'specific linking difference' ($\sigma$) of the DNA molecule:

$$\sigma = \frac{L_k - L_{km}}{L_{km}}$$

The specific linking difference of most natural DNA molecules lies in the range -0.03 to -0.09, indicating that most DNA circles are negatively supercoiled. Molecules differing only in linking number are topological isomers (topoisomers) of one another.

The linking number is a topological property of the DNA molecule, and can not be changed by geometric deformations of the molecule. Two geometric functions that are related to linking number are 'twist' ($T_w$) and 'writhe' ($W_r$). Twist is the number of Watson-Crick turns in a DNA helix, and writhe is the coiling of the helical axis in space. The sum of twist and writhe is the linking number. Changes in the linking number usually lead to changes in the writhing number, and thus the degree of supercoiling, rather than to changes in the twisting number.

1.2.2 DNA knots and catenanes

Aside from supercoiling, DNA can also exhibit several other higher order structural features, namely knots and catenanes (Figure 1.3). Knotting represents the entanglement of a single DNA molecule, single or double-stranded, that is in the form of a closed circle. The simplest form of knot is the trefoil, which has three lobes and three nodes (DNA strand crossovers). Knots can also be classified into 'torus' knots or 'twist' knots.

Catenanes are interlinked DNA rings, and are more common in nature than knots. The simplest form of catenane is a pair of singly intertwined covalently-closed double-helical DNA rings. The mitochondrial DNA of trypanosomes forms a compact
network, called kinetoplast DNA (kDNA), consisting of hundreds of interlinked DNA rings. Catenated DNA molecules may be produced during DNA replication, and will need to be resolved during chromosome segregation (Holm, 1994).

1.3 DNA TOPOISOMERASES

The DNA topoisomerases have evolved to overcome the topological complexities of DNA. They pass one DNA segment through a transient break in another DNA segment, and are therefore able to alter the topological state of DNA. The importance of DNA topoisomerases is testified to by their ubiquity in nature; they have been found in all organisms studied to date, and participate in nearly all cellular transactions of DNA. These enzymes are able to cleave and subsequently reseal the phosphodiester backbone of a DNA strand, which involves a transesterification reaction between an active site tyrosine hydroxyl group and a phosphoryl group on the DNA strand, forming a phosphotyrosine linkage. In this manner, the bond energy is conserved and no energy cofactor is required. Conformational changes within the enzyme-DNA covalent complex allow the DNA ends to be separated, which creates a gate through which another DNA strand can be passed. DNA topoisomerases can be classified into type I and type II enzymes (Liu et al., 1980). Type I enzymes are monomeric and can cleave only one DNA strand at a time, either in a single or double-stranded DNA molecule, whereas type II enzymes are generally dimeric, and can cleave both strands of a double-helical DNA molecule. As a result, type II enzymes are generally able to relax positively and negatively supercoiled DNA, knot/unknot and catenate/decatenate double-stranded DNA in an ATP-dependent manner. A unique bacterial type II topoisomerase, DNA gyrase, is also able to introduce negative supercoils into DNA. Type I enzymes also generally relax positively and negatively supercoiled double-stranded DNA. However, they are unable to knot/unknot or catenate/decatenate closed-circular DNA, except in the presence of a single stranded nick, and they are generally ATP-independent.
1.3.1 The roles of DNA topoisomerases

The biological roles of topoisomerases are primarily associated with maintenance of the topological state of DNA during normal cellular processes e.g. replication, transcription, recombination, and chromosome condensation and segregation (reviewed in Wang, 1996; Watt & Hickson, 1994; Nitiss, 1998). Evidence for these in vivo roles of topoisomerases comes principally from genetic studies using knockout systems.

Eubacterial cells, such as *E. coli*, contain two type I topoisomerases (topoisomerase I and topoisomerase III), and two type II topoisomerases (DNA gyrase and topoisomerase IV). During DNA transcription and replication, the translocation process of the polymerase complex along the DNA strand generates positive supercoils in front of the complex and negative supercoils behind the complex; such supercoils can only be released if the DNA or protein complex are free to rotate about the axis of the other molecule. This is known as the ‘twin supercoiled-domain’ model of supercoiling (Liu & Wang, 1987). DNA gyrase and topoisomerase I have diametrically opposed functions in dealing with this problem: DNA gyrase removes the positive supercoils ahead of the enzyme complex, whereas topoisomerase I removes negative supercoils behind the enzyme complex (prokaryotic topoisomerases I and III are unable to remove positive supercoils from DNA). This function of topoisomerases is known as a DNA swivel activity. Deletion of the topoisomerase I gene results in an increase in negative supercoiling and a slow growing phenotype (DiNardo et al., 1982), whereas deletion of the topoisomerase III gene does not affect cell viability (Schofield et al., 1992), this enzyme being more active in decatenation than in relaxation of negative supercoils (DiGate & Marians, 1988). Both type II topoisomerases are essential for cell viability. DNA gyrase is required during DNA replication for both initiation (Baker et al., 1986), and elongation (Drlica et al., 1980), unwinding the DNA template and removing positive supercoils ahead of the replication fork to allow forward progression of the fork. The expression of the gyrase proteins themselves is increased by a reduction in negative supercoiling (Menzel & Gellert, 1983). DNA gyrase also has a weak decatenase activity in vivo (Zechiedrich & Cozzarelli, 1995), and is thought to be involved in homologous and illegitimate DNA recombination (reviewed in Wang,
Topoisomerase IV unlinks DNA catenanes following DNA replication (Adams et al., 1992; reviewed in Baker, 1993), and deletion mutants are non-viable due to the accumulation of large nucleoids in the midcell and failure of the chromosomes to partition (Kato et al., 1990).

The contributions of topoisomerases I and IV and DNA gyrase to the steady-state DNA supercoiling level in *E. coli* was examined by inhibiting one or combinations of these enzymes *in vivo* (Zechiedrich et al., 2000). Unexpectedly, topoisomerase IV was found to play a significant role, alongside topoisomerase I and DNA gyrase, in maintaining the level of intracellular supercoiling. If DNA gyrase activity was blocked, then either topoisomerase I or IV removed negative supercoils to a $\sigma \approx -0.05$, but only topoisomerase IV was able to relax the DNA to a $\sigma \approx -0.015$ *i.e.* topoisomerase I only relaxes very negatively supercoiled DNA whereas topoisomerase IV relaxes DNA almost completely, implying that topoisomerase I plays a more important role in the relaxation of transcription-induced supercoiling than topoisomerase IV.

Yeast cells contain two type I topoisomerases (topoisomerase I and topoisomerase III) and one type II topoisomerase (topoisomerase II). Topoisomerase I is able to relax positively and negative supercoiled DNA, whereas topoisomerase III is like its prokaryotic counterpart in that it only relaxes negative supercoils (reviewed in Watt & Hickson, 1994; Nitiss, 1998). Neither of the two type I topoisomerases are essential for cell viability (Uemera & Yanagida, 1984), however, growth of double mutants is markedly retarded, which suggests that there is some degree of redundancy within these enzymes. Yeast topoisomerase I is required in the chain elongation step of DNA replication (Kim & Wang, 1989), presumably where it functions as a DNA swivel. Topoisomerase III relaxes negative supercoils only weakly, and thus is unlikely to play a major role in the maintenance of DNA supercoiling levels (Kim & Wang, 1992). Deletion mutants of topoisomerase III show a hyper-recombination phenotype (Bailis et al., 1992), implying a role for this enzyme in suppression of recombination and the maintenance of genome stability. The presence of only one type II topoisomerase in yeast makes it essential for cell viability due to its role in segregating replicated chromosomes during mitosis and meiosis (DiNardo et al., 1984). Although
topoisomerase II can function as the DNA swivel in topoisomerase I deletion mutants, the absence of any significant effect on DNA chain elongation in topoisomerase II mutants suggests that this enzyme does not normally perform this role in vivo (Kim & Wang, 1989).

Mammalian cells contain at least five topoisomerases (reviewed in Nitiss, 1998): three type I topoisomerases (topoisomerase I and topoisomerases IIIα and IIIβ) and two type II topoisomerases (topoisomerase IIα and IIβ). Topoisomerase I knockouts in mice are embryonic lethal (Morham et al., 1996), and this enzyme is also thought to play a role in the initiation (Kretzschmar et al., 1993) and elongation (Zhang et al., 1988) phases of DNA replication and transcription, acting as a swivel to relieve torsional strain in the template. More recently, human topoisomerase I has been reported to have a protein kinase activity in the presence of ATP (Rossi et al., 1996), and may play a structural role in chromosome organisation (Mo et al., 2000). Topoisomerase IIIα knockouts in mice are also embryonic lethal (Li & Wang, 1998), and the purified enzyme relaxes negatively supercoiled DNA in a distributive manner and with a relatively low catalytic efficiency (Goulaouic et al., 1999), which suggests that this enzyme does not play a significant role in the removal of negative supercoils in vivo. Topoisomerase IIIβ is expressed as three different gene products in humans, and like yeast topoisomerase II and topoisomerase III, human topoisomerase IIIβ interacts with a yeast DNA helicase (SGS1), and thus may be involved in regulation of DNA recombination (Ng et al., 1999). The two type II topoisomerases, IIα and IIβ, are both able to complement yeast strains that lack a functional type II topoisomerase (Jensen et al., 1996), and therefore both isoforms can support mitotic chromosome condensation and segregation in yeast, which has been proposed as their major role in vivo. However, human cell lines expressing a mutant topoisomerase IIα which lacks its C-terminal nuclear localisation signal, and is therefore located mostly outside the nucleus, showed that topoisomerase IIβ did not adopt the mitotic functions played by the α-isoform in wild-type cells (Grue et al., 1996). Instead, mitosis was sustained by mutant topoisomerase IIα that entered the chromatin after breakdown of the nuclear envelope. These results suggest that topoisomerase IIα is responsible for chromosome segregation and disjunction following mitosis. Topoisomerase IIβ knockouts in mice revealed a
surprise role for this enzyme in neural development (Yang et al., 2000). Although neurogenesis was normal in these mice, motor axons failed to contact skeletal muscle and sensory axons failed to enter the spinal chord, suggesting a defect in neuromuscular function. The authors propose these effects may be due to a plausible role of topoisomerase IIβ in DNA repair in neurons. The precise roles of the topoisomerase IIα and IIβ isoforms, which share significant homology at the amino acid level, remain to be delineated. It is likely that differences in function will be mediated by their cellular location and expression levels, and by the nature of their interactions with other proteins (Jensen et al., 1996; Austin & Marsh, 1998).

DNA topoisomerases are also thought to play a structural role in chromosomes (reviewed in Poljak & Käs, 1995; Warburton & Earnshaw, 1997). Topoisomerase II is the major protein component of the ‘chromosome scaffold’, a subfraction of nonhistone chromosomal proteins (Earnshaw et al., 1985). Staining of human chromosomes with anti-topoisomerase II antibodies reveals an axial distribution for topoisomerase II along the long axis of the chromatid arms (Earnshaw & Heck, 1985). Given the ‘radial-loop’ model of genomic DNA, where the chromosome scaffold is thought to tether topologically-independent DNA loops of ~30-100 kilobases (kb) emanating from a central scaffold (Earnshaw & Laemmli, 1983), it has been suggested that topoisomerase II is located at the base of these DNA loops, at AT-rich DNA sequences called ‘scaffold-attachment/associated regions’ (SARs), where it anchors the DNA loops to the nuclear scaffold (Earnshaw & Heck, 1985; Gasser et al., 1986). These SAR sequences are several hundred base pairs long and contain many (dA•dT) tracts which mediate the interaction with the nuclear scaffold (Käs et al., 1989). Topoisomerase II has been shown to interact with DNA containing SAR sequences in a highly specific and cooperative manner (Adachi et al., 1989), and in fact causes extensive protein-DNA aggregation and precipitation, suggestive of DNA-dependent protein-protein interactions. Phosphorylation of the topoisomerase II C-terminal domain has been shown to increase the stability of enzyme-SAR-containing DNA fragment complexes (Dang et al., 1994), and also to promote the formation of higher order complexes of topoisomerase dimers (Vassetzky et al., 1994). Sites of preferential topoisomerase II-mediated DNA cleavage have been mapped to intergenic SAR sequences in eukaryotic
DNA fragments (Udvardy et al., 1985). Analysis of long-range cleavage sites in an amplified c-myc gene locus in the presence of several topoisomerase II-directed antitumour drugs revealed that cleavage occurred within regions of ~5 kb separated by regions of ~70-100 kb, consistent with the radial-loop model for chromatin organisation (Gromova et al., 1995). Topoisomerase II has also been shown to play a structural role in chromosome condensation during meiosis and mitosis, and subsequent decondensation (Uemera et al., 1987; Adachi et al., 1991).

A recent study by Meyer et al. (1997) sought to determine the locations of topoisomerase I and topoisomerases IIα and IIβ during mitosis and interphase in human A431 epidermoid cells. During interphase, the majority of all three topoisomerases was found in the nucleoplasm. During mitosis, topoisomerase I and topoisomerase IIα were found to be associated with the chromosomes, moreover, topoisomerase I showed a diffuse distribution whereas topoisomerase IIα was concentrated in threadlike structures along the chromosome arms. Topoisomerase IIα was also found associated with the centrioles during mitosis. However, topoisomerase IIβ diffused completely into the cytosol, and was not detectable bound to chromatin. Although other studies have revealed different staining patterns for the type II topoisomerases in mammalian cells, probably due to the use of different antibodies, there is mounting immunohistochemical (and biochemical) evidence that it is topoisomerase IIα and not topoisomerase IIβ that is essential in chromosome condensation and disjunction.

Taken together, these results suggest that the role of topoisomerase II located in SAR sequences at the base of chromatin loops is to tether these topologically-independent DNA loops and presumably to use its enzymic activity to control the topology of these loops, which will undergo transcription-dependent supercoiling according to the twin-supercoiled domain model (Liu & Wang, 1987).

1.3.2 Type I DNA topoisomerases

As previously stated, type I topoisomerases can cleave only one DNA strand in a ss- or dsDNA molecule, and pass a second DNA strand (either ss- or dsDNA) through
this break before resealing the gap. Such a reaction results in a change in linking number in steps of one, whereas type II topoisomerases cause linking number changes in steps of two. The type I topoisomerases can be further subdivided into types IA and IB on the basis of amino acid sequence homology and reaction characteristics. For members of the IA subfamily the active site tyrosyl residue becomes linked to a DNA 5'-phosphoryl group during the DNA cleavage reaction; whereas for the type IB enzymes the tyrosyl residue becomes linked to a DNA 3'-phosphoryl group. The type IA subfamily includes bacterial DNA topoisomerases I and III and eukaryotic DNA topoisomerase III. The type IB subfamily includes eukaryotic DNA topoisomerase I and pox virus DNA topoisomerases (reviewed in Wang, 1996). Reverse gyrase, a unique type IA topoisomerase isolated first from the hyperthermophilic archaeabacterium Sulfolobus acidocaldarius, is able to introduce positive supercoils into DNA in the presence of ATP (Kikuchi & Asai, 1984). The amino acid sequence of this enzyme (Confalonieri et al., 1993) reveals the presence of two distinct domains: an N-terminal ATP-binding domain that contains sequence motifs commonly found in DNA helicases, and a C-terminal domain with an amino acid sequence homologous to other type IA topoisomerases. A model for the introduction of positive supercoils by reverse gyrase was proposed based on the twin supercoiled-domain model (Liu & Wang, 1987): translocation of a helicase along the DNA generates positive supercoils ahead of the enzyme and negative supercoils behind it; relaxation of the negative supercoils by the topoisomerase I domain would result in a net introduction of positive supercoils (Confalonieri et al., 1993). Presumably, the role of this enzyme is to stabilise the genomic DNA from high-temperature-induced denaturation by tightening the DNA pitch; however, its exact roles are still to be determined, and it may play a role in maintaining genetic stability (Duguet, 1997).

1.3.2.1 Type IA topoisomerases

Of the type IA topoisomerases, *E. coli* DNA topoisomerase I is the best characterised, and shall be used here as a representative example of this group. This enzyme binds preferentially to short regions of ssDNA (Depew et al., 1978), and therefore is able to relax highly negatively supercoiled DNA (in which ssDNA regions
can be formed), knot/unknot ssDNA rings, link two complementary ssDNA rings into a dsDNA ring, and catenate and knot dsDNA in the presence of a ssDNA nick (reviewed in Champoux, 1990). *E. coli* topoisomerase I is a monomeric 97 kDa metalloprotein (865 residues) consisting of three domains: a 67 kDa N-terminal DNA-cleavage domain, a central domain which binds three Zn$^{2+}$ atoms, each coordinated by a tetracysteine motif (Tse-Dinh & Beran-Steed, 1988), and a 14 kDa C-terminal DNA-binding domain (Yu *et al.*, 1995). The crystal structure of the 67 kDa N-terminal fragment of this protein (Lima *et al.*, 1994) reveals a four-domain architecture (see Figure 1.4), which fold to form an elongated toroidal shape with a hole large enough to accommodate dsDNA. The active site tyrosine (Tyr319) is located in domain III, and is buried at the interface of this domain with domain I.

The available structural and biochemical evidence support an 'enzyme-bridging' model for DNA strand passage, shown in Figure 1.4 (Brown & Cozzarelli, 1981; Lima *et al.*, 1994; reviewed in Champoux, 1990; Roca, 1995). In this model, the catenation or decatenation activity of the enzyme is shown, where one of the dsDNA rings has a gap that provides the ssDNA region. In the first stage, the ssDNA strand binds to the interface between domains I and III, presumably requiring a slight opening of the interface. The ssDNA strand is then cleaved by nucleophilic attack of the active site tyrosine on the 5'-phosphoryl group, forming a covalent link. The 3'-OH group is stabilised by noncovalent interactions with domain I, and thus the enzyme forms a protein 'bridge' between the DNA 3' and 5' ends. The opening of the gate via a conformational change in the protein allows the second DNA strand to pass through the break in the first strand and enter the cavity of the enzyme. The cleaved DNA strand is then resealed by the closure of the gate, transiently trapping the second DNA strand within the enzyme cavity. The gate must then reopen (without DNA cleavage) to release the second DNA strand, while the first ssDNA strand remains noncovalently associated with one side of the protein bridge only. This model must be modified slightly for the relaxation of negative supercoils, since the cleaved and passed DNA strands are complementary strands of the same DNA duplex. One prediction for DNA relaxation according to this model is that supercoils can only be removed in steps of one, corresponding to one cycle of nicking, strand passage and closing. This has been
The model shows the catenation/decatenation of double-stranded rings where one ring contains a ssDNA gap region. The enzyme structure shown is that of the 67 kDa N-terminal domain of E. coli topoisomerase I (Lima et al., 1994). In this reaction, the ssDNA strand (green) is cleaved by the enzyme (B); the 5'-phosphoryl end is linked to the active site tyrosine, whereas the 3'-OH end interacts noncovalently with domain I. A conformational change in the protein allows the DNA ends to move apart (C), and the duplex DNA molecule (seen in cross-section as a coloured circle) passes through the break and enters the protein cavity (D). The broken DNA strand is then resealed (E) before the enzyme opens again to allow the duplex DNA molecule to exit from the protein cavity (F,G). (Lima et al., 1994).
shown to be the case under conditions where the enzyme acts in a distributive manner (Brown & Cozzarelli, 1981). This mechanism is conceptually similar to that of the type II topoisomerase (see section 1.5).

1.3.2.2 Type IB topoisomerases

Eukaryotic type I topoisomerases differ significantly from their prokaryotic counterparts in that they bind preferentially to dsDNA, form a covalent link with the 3' end of the broken strand, and can relax both positive and negative supercoils. Human topoisomerase I is a 91 kDa protein (765 residues) consisting of four domains (Stewart et al., 1996b): a 24 kDa N-terminal domain that is dispensable for catalytic activity and is largely unfolded (Stewart et al., 1996a), a 54 kDa core domain, a 7 kDa linker region, and a 6 kDa C-terminal domain that contains the active site tyrosine. The crystal structure of a 70 kDa N-terminally truncated form of the enzyme in a noncovalent complex with a 22 bp duplex DNA molecule has been solved at a resolution of 2.8 Å (Stewart et al., 1998). The protein has a bi-lobed structure that forms a clamp that completely encircles the DNA molecule (see Figure 1.5), which has been likened to a 'nut around a bolt' (Nash, 1998). One lobe (core subdomains I and II) forms a cap that sits on top of the DNA duplex, while the second lobe (core subdomain III, C-terminal domain and linker) sits below the DNA and contains the catalytic residues implicated in DNA cleavage and religation.

Viral type I topoisomerases are significantly smaller in size than their cellular counterparts, ranging from 314 to 333 amino acids (reviewed in Shuman, 1998), with vaccinia virus topoisomerase I the smallest at 36 kDa (Shuman & Moss, 1987). The vaccinia enzyme is composed of two domains (Sekiguchi & Shuman, 1995): a tightly folded N-terminal domain (residues 1-80), and the catalytic C-terminal domain (residues 81-314), which contains a loosely folded ‘hinge’ region at residues 135-147. The presence of DNA reduces the extent of proteolytic cleavage at the interdomain and hinge regions, indicative of either direct DNA binding or DNA binding-induced conformational changes in these regions. Vaccinia topoisomerase binds and cleaves dsDNA with a preference for the conserved sequence element 5'-CCCTT↓ in the scissile
strand (Shuman & Prescott, 1990). Molecular modelling using the crystal structures of the N-terminal and C-terminal domains and a B-form DNA molecule containing the preferred cleavage site for vaccinia topoisomerase suggests that the enzyme forms a C-shaped protein clamp around the DNA molecule (Cheng et al., 1998; Shuman, 1998).

Despite obvious differences in size and domain organisation, and only weak sequence similarity, the vaccinia and human topoisomerase I enzymes display similar biochemical and catalytic properties and share several mechanistic features (Shuman, 1998). These enzymes both have common catalytic core structures and a common active site, and in both cases the DNA duplex is bound circumferentially by a C-shaped protein clamp. Several features of the reaction cycle of type IB topoisomerases argue against an enzyme-bridging mechanism of supercoil relaxation for these enzymes: these enzymes show no base preference at the 5' end of the broken DNA strand (Shuman & Prescott, 1990), and short segments of the cleaved strand 3' to the cleavage site dissociate from the enzyme after cleavage (Shuman, 1991). Together, these suggest that the enzyme does not interact with the broken DNA strand on the noncovalent (5') end of the cleaved DNA. Also, unlike type IA topoisomerases, the type IB enzymes are unable to knot ssDNA rings or nicked dsDNA circles (Champoux, 1990).

This evidence led to an alternate model for supercoil release by the type IB topoisomerases being proposed: the 'free rotation' model (Champoux, 1990). In this model, the end of the broken strand that is not covalently attached to the enzyme (the 5' end) is unconstrained by the protein and is therefore free to rotate about the phosphodiester bond in the unbroken strand, which releases the DNA superhelical tension. More recently, with the determination of the structure of a 70 kDa N-terminally truncated form of the enzyme in complex with a 22 bp DNA duplex (Stewart et al., 1998), this model has been modified slightly and renamed the 'controlled rotation' model (Figure 1.5). In this modified model, the DNA duplex downstream of the cleavage site rotates about the phosphodiester bond in the intact strand within a positively charged cavity formed by the cap of the enzyme and the linker domain. This cavity is shaped in such a way so as to guide the rotating DNA duplex through interaction with two 'nose-cone' helices, within the cap, and the linker domain. A large
Figure 1.5 The ‘controlled rotation’ model for the reaction mechanism of human topoisomerase I.
The model shows the relaxation of highly negatively supercoiled DNA (red) into a less supercoiled state (green). The crystal structure shown is that of a 70 kDa fragment of human topoisomerase I in complex with dsDNA (Stewart et al., 1998). The cap of the enzyme (subdomains I and II) is coloured blue, and the catalytic core (subdomain III, C-terminal domain and linker) is coloured purple. The supercoiled DNA substrate binds to an ‘open’ conformation of the enzyme (A), such that the protein completely encircles the DNA duplex (B). Following the cleavage event, the release of superhelical tension can occur through one or more cycles of controlled rotation within the positively charged cavity provided by the cap and linker domain (D). The rotating DNA duplex downstream of the cleavage site is shown in different coloured rotation states at 30° increments. Subsequently, the covalent intermediate is religated (F), and the relaxed DNA molecule is released from the enzyme (G). (Stewart et al., 1998).
degree of conformational flexibility has been shown to occur in these two regions (Redinbo et al., 1999), consistent with the structural flexibility invoked in the controlled rotation mechanism. According to this model, multiple unwinding events should be possible for each cleavage-religation cycle. This has been confirmed for the vaccinia enzyme (Stivers et al., 1997), which was shown to remove an average of five supercoils per cleavage event.

1.3.3 Type II DNA topoisomerases

Type II DNA topoisomerases are structurally and mechanistically distinct from the type I topoisomerases, although some similarities are observed with the type IA enzymes. The type II enzymes are DNA-dependent ATPases that couple ATP binding and hydrolysis to the topological conversion of double-stranded segments of DNA. These enzymes differ in their quaternary structure: eukaryotic type II enzymes function as homodimers, prokaryotic enzymes have an A₂B₂ structure, and bacteriophage T4 topoisomerase has three subunits in an A₂B₂C₂ formation (reviewed in Wang, 1998). The type II enzymes are evolutionarily conserved across the three kingdoms of life. However, archaeabacteria contain one additional type II topoisomerase quite distinct from the type II topoisomerases of prokaryotes and eukaryotes. Topoisomerase VI of the archaeabacterium Sulfolobus shibatae exhibits little sequence homology with other type II topoisomerases, except for the presence of an ATP-binding domain in the N-terminus of one of the two subunits (Bergerat et al., 1997; see section 1.1.2). This has led to the classification of type II topoisomerases into types IIA (classical type II topoisomerases from archaeabacteria, prokaryotes and eukaryotes) and IIB (topoisomerase VI from archaeabacteria). The reconstituted enzyme is able to relax both negatively and positively supercoiled DNA and decatenate DNA circles in an ATP-dependent manner (Buhler et al., 1998), and may play a role in meiotic recombination.

Like the type IA topoisomerases, the type II topoisomerases form a phosphotyrosine bond with a 5'-phosphoryl group on a DNA strand. However, because of their dyadic structure, the type II topoisomerases are able to cleave both strands of a duplex DNA molecule. The two active site tyrosine residues of the enzyme become
*E. coli* DNA gyrase

Human DNA topoisomerase IIα

Yeast DNA topoisomerase II

Figure 1.6 Domain structure of *E. coli* DNA gyrase, yeast DNA topoisomerase II and human DNA topoisomerase IIα.

DNA gyrase is a heterotetramer of A₂B₂ configuration, whereas the eukaryotic enzymes are homodimers. The proteins are aligned to indicate regions of homology. The N-terminal regions have the most homology, the C-terminal regions have the least homology. The eukaryotic enzymes can be considered as fusions of the prokaryotic GyrB and GyrA proteins. The length of each protein in amino acid residues is indicated on the right. The position of the active site tyrosines are indicated (Y). ATP indicates the ATP binding site. Black bars indicate regions of known structure (Xtal). Sites A, B and C on the yeast enzyme indicate SV8 proteolysis sites in the absence (A & C) and presence (B & C) of the non-hydrolysable ATP analogue ADPNP. A' and B' indicate known domains of DNA gyrase. The loop in the DNA gyrase GyrB protein indicates a region of unique sequence.
covalently linked to 5'-phosphoryl groups four base pairs apart (Morrison & Cozzarelli, 1979). The cleaved dsDNA strand, the gate or G segment, is then transiently opened to allow a second dsDNA segment, the transported or T segment, through this break, which is subsequently resealed (Liu et al., 1980). If the two DNA segments are from the same molecule, this reaction would result in supercoil relaxation (or introduction in the case of DNA gyrase) and change the linking number of the DNA molecule in steps of two, whereas if the two segments are from different DNA molecules it would result in either the catenation or decatenation of DNA rings.

1.3.3.1 *E. coli* DNA gyrase

The prokaryotic enzyme DNA gyrase (prokaryotic topoisomerase II) was discovered in 1976 by Gellert and coworkers as an essential host factor in the integrative recombination of phage lambda (Gellert et al., 1976). The purified enzyme from *E. coli* was able to introduce negative supercoils into a relaxed dsDNA substrate in the presence of ATP and Mg$^{2+}$, presumably via a coupled pathway linking ATP hydrolysis to the introduction of supercoils (Gellert et al., 1976). The gyrase enzyme was subsequently found to consist of two separate proteins (Mizuuchi et al., 1978), now referred to as GyrA and GyrB and encoded by the *gyrA* and *gyrB* genes respectively; the functional enzyme is an A$_2$B$_2$ heterotetramer.

In *E. coli*, DNA gyrase consists of a GyrA and GyrB protein of length 875 amino acids and 804 amino acids respectively, with molecular weights of 97 kDa and 90 kDa (Adachi et al., 1987; Swanberg & Wang, 1987). Proteolytic cleavage of these two proteins separates each into two functional domains (Figure 1.6). The GyrA protein is cleaved by trypsin into two fragments with molecular weights of 64 kDa and 33 kDa (Reece & Maxwell, 1989). The 64 kDa N-terminal fragment comprises the DNA breakage-reunion domain of GyrA. The active site tyrosine in the transesterification reaction of DNA breakage/religation has been mapped to Tyr122 (Horowitz & Wang, 1987); when reconstituted with the GyrB protein, the 64 kDa fragment was found to support both DNA cleavage and limited supercoiling (Reece & Maxwell, 1989). A 59 kDa portion of the 64 kDa fragment, representing residues 2-523, has been crystallised
Figure 1.7 Crystal structure of the 59 kDa N-terminal fragment of the *E. coli* DNA gyrase GyrA protein.

The structure of the 59 kDa N-terminal fragment of the *E. coli* DNA gyrase GyrA protein was determined at 2.8 Å resolution (Morais Cabral *et al.*, 1997). This fragment of the protein comprises residues 2-523, although residues 2-29, 251-256, 307-312 and 426-429 were disordered and are not shown. The protein is a heart-shaped dimer with one monomer shown in yellow and the other in blue. The N- and C-terminus of the blue monomer are indicated. Each monomer folds into two sub-fragments, an N-proximal head and a C-proximal tail. The head region contains a motif which has the same fold as the DNA-binding domain of the catabolite activator protein (CAP), a helix-turn-helix motif (green). The active site tyrosines (Tyr122) are shown in space filling mode (red). The diameter of the central hole is ~30 Å.
and its structure solved at 2.8 Å (Figure 1.7; Morais Cabral et al., 1997). This dimeric fragment has a heart-shaped structure with a large central cavity. Each monomer has an N-proximal ‘head’ region and a C-proximal ‘tail’ region. The head region contains the same DNA-binding helix-turn-helix motif as the catabolite activator protein (CAP) (Schultz et al., 1991). DNA has been modelled into this structure based on the positions of the active site tyrosines and the recognition helices of the CAP domain. This structure is thought to represent the protein in a conformation where the G segment is bound but not cleaved or separated. The purified 33 kDa C-terminal domain can bind to DNA in such a manner so as to introduce positive supercoils, and when reconstituted with the 64 kDa fragment and GyrB is able to increase the efficiency of the supercoiling reaction compared to in its absence (Reece & Maxwell, 1991). Proteolytic cleavage of the GyrB protein yields an N-terminal fragment of 43 kDa and a C-terminal fragment of 47 kDa (Adachi et al., 1987). The 43 kDa fragment binds and hydrolysates ATP (Ali et al., 1993), while the 47 kDa fragment interacts with GyrA and DNA. The 47 kDa fragment was originally isolated from *E. coli* cells and shown to have a DNA relaxation activity in the absence of ATP when reconstituted with the GyrA protein; this enzyme was termed topoisomerase II' (Brown et al., 1979).

The crystal structure of the 43 kDa N-terminal fragment of GyrB (residues 2-393) in the presence of the non-hydrolysable ATP analogue ADPNP was solved at 2.5 Å resolution by Wigley et al. (1991) (Figure 1.8). The protein monomer consists of two subdomains: an N-terminal domain that contains the bound ADPNP (residues 2-220), and a C-terminal domain. The catalytic residue in ATP hydrolysis has been mapped by site-directed mutagenesis to Glu42 (Jackson & Maxwell, 1993). The N-terminal subdomain has been shown to be the binding site for the coumarin antibiotics (Gilbert & Maxwell, 1994). The 43 kDa fragment dimerises in the presence of ADPNP (Ali et al., 1995), with most of the dimer interactions occurring in the N-terminal region of the protein, as visualised in the crystal structure. The C-terminal helices also appear to make a dimer contact, but this may represent a crystallographic artifact rather than a native interaction because the 47 kDa fragment is missing. The N-terminal arm of one monomer makes contact with the ADPNP molecule of the other monomer through an interaction between Tyr5 on the protein and the 2' hydroxyl group on the ribose ring of
Figure 1.8 Crystal structure of the 43 kDa N-terminal fragment of the *E. coli* DNA gyrase Gyrb protein.
The 43 kDa N-terminal fragment of the Gyrb protein of *E. coli* was crystallised in the presence of ADPNP and the structure solved at 2.5 Å resolution (Wigley *et al.*, 1991). The protein is a dimer with one monomer shown in red and the other in yellow; the ADPNP molecule (one per monomer) is shown in blue in space filling mode. The N- and C-terminus of the yellow monomer are indicated. The N-terminal arm of one monomer makes contact with the ADPNP molecule of the other monomer. The central cavity is approximately 20 Å in diameter and is proposed to form part of a protein clamp which captures a segment of DNA.
ATP. The crystal structure also reveals the presence of a large hole approximately 20 Å in diameter running through the protein dimer. The walls of this cavity are the arginine-rich C-terminal helices of the second subdomain, which may constitute a DNA-binding surface. It has been proposed that the N-terminal domains form a protein clamp that captures a segment of DNA during the strand-passage event (Wigley et al., 1991).

The interaction of the gyrase enzyme with DNA is markedly different to that of the other more conventional type II topoisomerases that are unable to introduce supercoils into DNA. DNA gyrase wraps a segment of DNA ~130 bp in length around the core of the enzyme in a positive superhelical sense, of which a region of ~13 bp adjacent to the cleavage site is protected to a higher degree, which presumably represents a region buried in the active site of the enzyme (Orphanides & Maxwell, 1994). In contrast, eukaryotic topoisomerase II enzymes bind DNA without wrapping it and protect only 25-35 bp from digestion in DNase I footprinting experiments (Lee et al., 1989). It is the 33 kDa C-terminal domain of the GyrA protein that is responsible for the right-handed wrapping of DNA around the enzyme; deletion of this domain results in an enzyme that cannot wrap DNA. This complex is unable to supercoil DNA but relaxes DNA in an ATP-dependent manner, akin to other type II topoisomerases (Kampranis & Maxwell, 1996). The role of the wrapped DNA segment in the negative supercoiling reaction of DNA gyrase will be discussed in a later section.

1.3.3.2 Yeast topoisomerase II

The type II DNA topoisomerase from the yeast *Saccharomyces cerevisiae* is a single chain polypeptide of length 1429 amino acids, with a molecular weight of 164 kDa (Giaever et al., 1986). This enzyme is a stable dimer in solution, as judged by sedimentation equilibrium studies and subunit exchange assays (Tennyson & Lindsley, 1997). Limited proteolytic digestion with SV8 protease reveals the presence of three distinct protease sites (Lindsley & Wang, 1991; Figure 1.6). In the absence of the ATP analogue ADPNP, the enzyme is cleaved at sites A and C, on the carboxyl side of Glu410 and Asp1202 respectively. In the presence of ADPNP, the enzyme is cleaved at site B instead of site A, which is on the carboxyl side of Glu680. These results suggest
there is an allosteric interdomainal movement in the enzyme following the binding of ADPNP, and presumably ATP. Alignment of the sequences of *E. coli* GyrA and GyrB with that of *S. cerevisiae* topoisomerase II reveals that protease sites A and C correspond to the two principle proteolytic sites in the GyrB and GyrA proteins respectively (Caron & Wang, 1994). Site B corresponds to the GyrB/GyrA junction, and lies in the N-terminal region of GyrA. The active site tyrosine has been mapped to Tyr782 (Worland & Wang, 1989).

A 92 kDa core fragment of *S. cerevisiae* DNA topoisomerase II has been crystallised, and its structure determined at 2.7 Å resolution (Berger et al., 1996; Figure 1.9). This fragment of the protein comprises residues Glu-410 to Asp-1202 i.e. from protease site A to site C. Functionally, this 92 kDa fragment is able to cleave DNA, but cannot carry out the duplex transport step. The overall structure of this fragment is that of a heart-shaped dimeric protein of dimensions 120 x 120 x 55 Å, with a large central cavity. The protein folds into two sub-fragments, termed B' and A' since they align approximately with regions of the GyrB and GyrA proteins of *E. coli* DNA gyrase respectively. The A' region is divided into the N-proximal head region which contains the same CAP DNA-binding domain as the GyrA protein, and the C-proximal tail region, which is largely α-helical and forms the primary dimer interface at the ‘bottom’ of the molecule. A segment of DNA has been modelled into this structure taking account of the positions of the active site tyrosines and the CAP-like domain, and suggests that an enzyme with one bound DNA segment can admit a second DNA duplex which can be transported through a break in the first (Berger et al., 1996). The gross structure of this fragment is similar to that of the 59 kDa fragment of GyrA (Morais Cabral et al., 1997). Differences in the two structures are proposed to be representative of them being different conformations that occur in the enzymatic pathway. More recently, an alternate structure of the same 92 kDa yeast fragment has been determined in which the enzyme occupies yet another conformational state differing from that seen previously with GyrA59 or the same fragment (Fass et al., 1999). The principle difference between these three structures is the location of the A' heads: the GyrA59 structure (T2C), new yeast structure (T2M) and original yeast structure (T2O) appear to
Figure 1.9 Crystal structure of the 92 kDa core fragment of yeast topoisomerase II. Shown is the crystal structure of the 92 kDa core fragment of *S. cerevisiae* DNA topoisomerase II (Berger *et al.*, 1996). This fragment of the protein comprises residues Glu410 to Asp1202 *i.e.* from protease site A to site C. The protein is a heart-shaped dimer with one monomer shown in yellow and the other in blue. The N- and C-terminus of the yellow monomer are indicated. Each monomer folds into two sub-fragments, termed B' and A' which align approximately with regions of the GyrB and GyrA proteins of *E. coli* DNA gyrase respectively; the A' region consists of a head region and a tail region. Like the GyrA59 structure, the head region also contains a CAP DNA-binding domain (green). The active site tyrosines (Tyr783) are shown in space filling mode (red). The large central hole is ~55 Å wide at its base, ~25 Å wide at the top, and ~60 Å long.
represent snapshots of the protein in a conformation with the DNA gate closed, midway, and open, respectively.

Further structural information comes from electron microscopy data of yeast DNA topoisomerase II and truncated forms of the protein (Benedetti et al., 1997). The whole enzyme appears as a tripartite structure, with an ~170 Å diameter globular core flanked by two ~80 Å diameter spherical masses. Comparisons with truncated proteins revealed that the two smaller spheres are the N-terminal domains of the dimeric enzyme. The addition of ADPNP was found to bring the two smaller spheres into contact with one another, presumably via a conformational change in a flexible linker region brought about by the binding of ADPNP. This is consistent with the dimerisation of the 43 kDa fragment of DNA gyrase GyrB in the presence of ADPNP (Ali et al., 1995), and also limited proteolysis of the full-length yeast enzyme (Lindsley & Wang, 1991).

1.3.3.3 Human type II topoisomerases

Drake et al. (1987) first reported the presence of two distinct forms of type II DNA topoisomerases in human cells. Initial biochemical analysis (Drake et al., 1989) of the two isoforms (p170 and p180) showed that they could both unknot DNA in an ATP-dependent manner. In the ATP-dependent relaxation of negatively supercoiled pBR322, the progress of the reaction differed for the two forms; for p170 (now the α-isoform) the reaction was distributive, while p180 (the β-isoform) was processive.

The α- and β-isoforms were subsequently cloned and sequenced by Tsai-Pflugfelder et al. (1988) and Jenkins et al. (1992), respectively. The α-isoform is 1531 amino acids in length while the β-isoform is 1621 amino acids; the two isoforms are 68% identical at the amino acid level. However, this homology is not evenly distributed: the N-terminal three quarters of the proteins are 78% identical, whereas the C-terminal quarter is only 34% identical. Based on the exon-intron structure for the two genes, they appear to have diverged following a recent gene duplication event (Sng et al., 1999). Both type II topoisomerases function as homodimers, however, recent evidence suggests that α/β heterodimers exist naturally in mammalian cells or when
coexpressed in yeast (Biersack et al., 1996; Gromova et al., 1998). These α/β heterodimers were found to be active in the decatenation of kinetoplast DNA (Biersack et al., 1996), and as much as 25% of the total topoisomerase IIβ population may be involved in heterodimer formation (Gromova et al., 1998). Although the significance of these results is not yet clear, these α/β heterodimers may represent a subclass of topoisomerase II. This subclass may have a role distinct from the α and β homodimers within the nucleus e.g. they may be involved in DNA metabolism or in ‘fine-tuning’ the activities of topoisomerase II (Biersack et al., 1996), or they may be involved in DNA recombination as a result of subunit exchange (Bae et al., 1988).

Both of the human type II topoisomerases have been cloned and overexpressed in yeast (Wasserman et al., 1993; Austin et al., 1995). The first 28 amino acids of the α-isoform and 45 amino acids of the β-isoform were replaced with the first 5 of the yeast type II topoisomerase. The human enzymes were found to be catalytically active and biologically functional in yeast cells. Limited proteolytic digestion of the two isoforms indicated that these proteins share a domain structure similar to that seen for yeast topoisomerase II (Austin et al., 1995; see Figure 1.6). Each enzyme monomer consists of three functional domains: an N-terminal ATPase domain, a central DNA-cleavage/religation domain, and a C-terminal domain. The active site tyrosine residues are Tyr805 and Tyr821 for topoisomerase IIα and IIβ, respectively. The C-terminal regions of type II DNA topoisomerases are highly divergent; in contrast, the N-terminal ATPase domains and breakage-reunion domains show high sequence conservation throughout all type II DNA topoisomerases (Caron & Wang, 1994). The C-terminal region of eukaryotic type II topoisomerases has been implicated in three potential roles: nuclear localisation, dimerisation and regulation of enzyme activity (reviewed in Watt & Hickson, 1994). It is the divergent C-terminal domain that may give rise to the isoenzyme-specific differences.

There is no high resolution structural information available on human type II DNA topoisomerases. Scanning transmission electron microscopy has been used to examine the structure and conformational changes in the DNA topoisomerase IIα enzyme (Schultz et al., 1996). The structure reveals a globular core of ~90 Å diameter.
flanked by two smaller domains of ~50 Å diameter. The smaller domains represent the N-terminal region of the protein. In the presence of ADPNP, the structure changes to consist of one ~90 Å domain and another of ~60 Å in diameter, due to interaction of the N-terminal domains. This is in agreement with similar work carried out on the yeast enzyme (Benedetti et al., 1997). The larger domain can be seen to have an ~25 Å internal tunnel, which may represent the channel through which the T segment is passed. The structure of the full-length topoisomerase IIα enzyme has also been examined by atomic force microscopy (AFM) under a tapping mode (Nettikadan et al., 1998). The enzyme had a heart or doughnut-like appearance with a large off-centred axial hole, which in ~10% of the molecules was ‘filled’ by a protein fragment. The different images may represent two different surface views of the enzyme.

The α- and β-isoenzymes are differentially expressed during the cell cycle (Woessner et al., 1991; Negri et al., 1992). Topoisomerase IIα is present at higher levels during the logarithmic phase of cell growth, whereas topoisomerase IIβ is present at relatively constant levels throughout the cell cycle, although a small increase is seen in exponentially growing cells. The type II topoisomerases of eukaryotes are phosphoproteins, the proteins become hyperphosphorylated as the cell enters the G2/M phase of the cell cycle (reviewed in Watt & Hickson, 1994; Austin & Marsh, 1998). Most of the sites of phosphorylation have been mapped to serine residues in the C-terminal domain of the protein (Wells et al., 1994; Wells & Hickson, 1995). However, a serine residue in the N-terminal domain of the protein, Ser29, is phosphorylated by protein kinase C (Wells et al., 1995). Furthermore, it has been shown that phosphorylation of topoisomerase II by protein kinase C and casein kinase II enhances the DNA relaxation activity of the enzyme by increasing its rate of ATP hydrolysis (Corbett et al., 1992a, 1993). It is likely that the cell-cycle dependent phosphorylation and activation of topoisomerase II is required to fulfill its role in chromosome segregation. For a detailed review of human topoisomerases IIα and IIβ, see Austin & Marsh (1998).
1.4 CATALYTIC CYCLE OF TYPE II DNA TOPOISOMERASES

The catalytic cycle of the type II DNA topoisomerases must be a complex, multistep one in which one DNA duplex is transiently broken and another DNA duplex is passed through this break. Furthermore, the ATP-dependent DNA relaxation cycle of DNA by eukaryotic type II topoisomerases must differ from the ATP-independent DNA relaxation cycle of DNA gyrase, and from the ATP-dependent introduction of negative supercoils by the same enzyme. Although the exact details of these cycles are still unclear, information from structural studies coupled with biochemical and enzymological findings have led to a generally accepted model for the strand passage event in each of these three catalytic cycles of type II DNA topoisomerases (reviewed in Watt & Hickson, 1994; Orphanides & Maxwell, 1994; Berger et al., 1996; Wang, 1998; Kampranis et al., 1999). The following sections describe the proposed mechanism for ATP-dependent DNA relaxation by eukaryotic type II topoisomerases, and the proposed mechanism for the ATP-dependent introduction of negative supercoils by DNA gyrase, and for the relaxation of supercoils by DNA gyrase. The final section deals with the role of ATP hydrolysis in the catalytic cycle.

1.4.1 A model for the catalytic cycle of type II topoisomerases

It has been known for some time that type II DNA topoisomerases can catalyse the formation of a transient double-stranded break in a DNA duplex, and allow the passage of another strand through this break (Liu et al., 1980). A gate or G segment becomes cleaved in both DNA strands via a transesterification reaction between the active site tyrosine residues and 5' phosphoryl groups four base pairs apart (Morrison & Cozzarelli, 1979). The transport or T segment is then moved through the enzyme-operated gate. Two kinds of model were originally proposed for the accomplishment of this feat: the 'one gate' model and the 'two gate' model (Mizuuchi et al., 1980; Morrison et al., 1980; reviewed in Roca & Wang, 1994). In the one gate model, the T segment exits from the enzyme, after passage through the G segment, via the same opening through which it entered; whereas, in the two gate model, the T segment exits through a second exit point on the other side of the G segment.
A series of experiments by Roca & Wang (1992) shed further light on the mechanism of strand passage. In the absence of ADPNP, yeast topoisomerase II was found to bind to both linear and circular DNA, with a preference for negatively or positively supercoiled DNA over linear, nicked circular or relaxed DNA. In contrast, in the presence of ADPNP yeast DNA topoisomerase II was found to bind only to linear DNA, and not to circular DNA of any form. This linear DNA could be ligated to form a stable ring around the enzyme, which can only be disrupted by denaturation of the protein with SDS. The preferential binding of supercoiled DNA to the protein in the absence of ADPNP suggests that it can interact with two DNA molecules simultaneously. These results were interpreted in terms of the enzyme functioning as an ATP-operated molecular clamp, which is open in the absence of ATP and closed in its presence, and is therefore able to capture a circular DNA ring in a topological link. With the G segment bound to the enzyme in the open state (seen as the binding of supercoiled DNA to the enzyme in the absence of ADPNP), the binding of ADPNP (and presumably ATP) may or may not lead to the capture of another DNA segment, the T segment. If the protein clamp captures a T segment, it is transported through a break in the G segment. This model agrees with the known crystal structure of the ATP-hydrolysing domain of *E. coli* GyrB (Wigley *et al.*, 1991; Figure 1.8). Although the authors of this paper proposed that the cavity in the 43 kDa fragment holds the T segment after transport through the G segment and not before, this new evidence suggests that it is the N-terminal ATPase domains that clamp shut in the presence of ADPNP and capture the T segment.

Evidence for the two gate mechanism came from further experiments from Roca & Wang (1994). The yeast enzyme was able to decatenate two singly-linked DNA rings even with the N-terminal domains clamped shut in the presence of ADPNP. This suggests that there is a second gate through which the T segment can exit, and also showed that ATP hydrolysis is not necessary for the release of the T segment following the strand passage reaction. The protein clamp formed by the N-terminal domains of eukaryotic topoisomerase II enzymes, and of the GyrB protein, was termed the N gate (Roca & Wang, 1994). The crystal structure of the 92 kDa fragment of the yeast enzyme revealed a probable location for the exit gate in the C-terminal portion of the
Figure 1.10 Schematic representation of the reaction mechanism of eukaryotic topoisomerase II.

Shown is a model for the catalytic mechanism of topoisomerase II, based on the crystal structures of the N-terminal domain of the DNA gyrase GyrB protein (Wigley et al., 1991) and the 92 kDa fragment of yeast topoisomerase II (Berger et al., 1996), and on available biochemical evidence. The ATPase domains are shown in yellow, the B' and A' subfragments are coloured red and blue, respectively; the G segment DNA is grey and the T segment DNA is green. In (1), a G segment binds to the open enzyme conformation, causing a conformational change which brings the A' subfragments together (2). Binding of ATP (asterisks) causes the ATPase domains (the N gate) to clamp shut (3), and the captured T segment is transported through a transient break in the G segment and into the interior of the enzyme (4) (shown to proceed via a hypothetical intermediate in square brackets). The G segment is then resealed and the T segment is released from the enzyme through the bottom dimer interface (the C gate) (5). The ATP is hydrolysed and the enzyme is reset (2). (Berger et al., 1996).
enzyme (Berger et al., 1996). To confirm this, Roca et al. (1996) introduced cysteine residues in this region of the full-length enzyme such that disulphide links can form across the C-terminal dimer interface and effectively lock this gate, the C gate, shut. The disulphide-linked enzyme was able to bind and cleave a G segment, suggesting that the G segment enters through the N gate. In one-step decatenation experiments, the ADPNP-driven closure of the N gate caused the unlinked DNA ring containing the T segment to be trapped within the cross-linked enzyme. This ring could be released from the enzyme by reduction of the disulphide links. These results confirmed that the type II topoisomerases function via a two gate mechanism, with two gates or jaws at opposite ends of the enzyme.

With the crystal structures of a large fragment of the yeast DNA topoisomerase II molecule (Figure 1.9) and the ATPase domain of E. coli DNA gyrase (Figure 1.8), the first three-dimensional picture of a functional type II topoisomerase became available (Berger et al., 1996; reviewed in Wigley, 1996). The model proposed, consistent with structural and biochemical observations, is shown in Figure 1.10 (Berger et al., 1996). In the fully open (unliganded) state the enzyme binds to the DNA segment to be cleaved (the G segment), which induces a conformational change in the enzyme that brings the A' sub-fragments together and the active site tyrosines into their attack positions (Figure 1.10; 1 & 2). Binding of ATP to the N-terminal domains causes the jaws of the clamp (N gate) to close, which may lead to the capture of a weakly held T segment (Figure 1.10; 3). The dimerisation of the N-terminal domains pushes the T segment through the transiently cleaved G segment and into the interior of the enzyme via an as yet unknown conformational cascade (Figure 1.10; 4). The G segment is then resealed and the second dimer interface (C gate) opens to allow the T segment to exit (Figure 1.10; 5). Hydrolysis of the bound ATP and release of the hydrolytic products allows the N gate to reopen which resets the enzyme (Figure 1.10; 2). Although this model is undoubtedly a vast oversimplification of the precise events that occur in the catalytic cycle, it is likely that the essential features of this model hold true for the ATP-dependent DNA relaxation mechanism of type II topoisomerases.
1.4.2 Introduction of negative supercoils by DNA gyrase

The ability of DNA gyrase to introduce negative supercoils into its DNA substrate is due to its ability to wrap the DNA in a right-handed, positive superhelical manner; deletion of the C-terminal domain responsible for this wrap abolishes supercoiling in favour of the relaxation reaction (Kampranis & Maxwell, 1996). In the strand passage reaction, the captured T segment forms a crossover with the G segment; the supercoiling ability of gyrase is due to its ability to dictate the orientation of this crossover. By wrapping the DNA in such a manner so as that only positive crossovers are formed, gyrase ensures that the strand passage reaction results in a reduction of linking number. By contrast, in the case of eukaryotic type II topoisomerases, the orientation of the crossovers depends upon the topology of the DNA substrate, and results in relaxation of positive or negative supercoils.

The model shown in Figure 1.11 is a representation of the gyrase tetramer, based on available crystal structure and electron microscopy information (Kampranis et al., 1999). In this representation, the 47 kDa C-terminal domain of the GyrB protein is replaced with the equivalent domain from the 92 kDa fragment of yeast topoisomerase II. The structure and location of the 33 kDa C-terminal domain is also unknown, and is depicted only as a solid object. In the strand passage reaction of DNA gyrase, the segment of DNA that is contiguous with the G segment is wrapped around the C-terminal domains of the enzyme and presented to the ATP-operated clamp, forming the T segment (Figure 1.11a). Although both DNA arms are wrapped around the gyrase enzyme, it is unknown whether one or both T segments are presented to the ATP-operated clamp. Binding of ATP closes the clamp and captures the T segment (Figure 1.11b). Hydroxyl radical footprinting of the DNA-gyrase complex, and the DNA-gyrase-ADPNP complex, has revealed that upon binding of ADPNP an additional 14 bp is protected on one of the two flanking DNA arms, presumably reflecting the capture of one arm by the ATP-operated clamp (Orphanides & Maxwell, 1994). It was originally proposed that the efficiency of capture depended upon the topological state of the DNA, and that once a T segment is captured it is transported through the G segment with a high probability (Bates et al., 1996). However, more recent evidence suggests that this is not the case i.e. the T segment is captured with high probability irrespective of the
Figure 1.11 Model for the mechanism of strand passage by DNA gyrase.
Shown is a representation of the gyrase tetramer; the model consists of the 59 kDa domain of GyrA (light blue), the 43 kDa domain of GyrB (dark blue), the B' subfragment of yeast topoisomerase II (red) is shown to represent the 47 kDa domain of GyrB, the 33 kDa C-terminal domain of GyrA is represented by a solid yellow crescent. In (a), the bound DNA strand containing the G segment is wrapped around the enzyme by the GyrA C-terminal domains such that one or both of the two DNA arms are presented to the ATP-operated clamp. Binding of nucleotide causes the ATP-operated clamp to shut and capture a T segment with high probability (b). This sets up an equilibrium of the transported DNA segment across the DNA gate, which depends upon the topology of the DNA substrate: with negatively supercoiled DNA, the T segment would equilibrate within the ATP-operated clamp (b), whereas with positively supercoiled DNA, the T segment would equilibrate beyond the gate (c). (Kampranis et al., 1999).
topology of the DNA substrate (Kampranis & Maxwell, 1999). This sets up an equilibrium of the transported DNA segment across the DNA gate (Figure 1.11c). The position of this equilibrium will depend upon the topology of the DNA substrate: with positively supercoiled DNA, the T segment would equilibrate beyond the gate (post-strand passage), whereas with negatively supercoiled DNA, the T segment would equilibrate within the ATP-operated clamp (pre-strand passage). The presence of the T segment within the ATP-operated clamp shifts the cleavage-religation equilibrium in favour of cleavage, which in turn stimulates the intrinsic ATPase activity and completes the catalytic turnover.

In experiments similar to those previously described with the yeast enzyme (Roca et al., 1996), the bottom dimer interface of DNA gyrase has been cross-linked at a pair of novel cysteine residues (Williams & Maxwell, 1999). The cross-linked enzyme was able to carry out a single round of strand passage in the presence of ATP; however, enzyme turnover was blocked, presumably because the T segment is trapped within the bottom cavity and cannot be released through the cross-linked C gate. It would thus appear that the strand passage event for DNA relaxation by eukaryotic type II topoisomerases and for the introduction of negative supercoils by DNA gyrase both operate via a similar two gate mechanism. In contrast, the ATP-independent removal of supercoils by DNA gyrase was completely abolished by cross-linking of the C gate. This suggests that the T segment enters through the bottom cavity, is passed through the G segment and expelled through the ATP-operated clamp i.e. via a reverse strand passage event.

1.4.3 The role of ATP hydrolysis in the catalytic cycle

In the case of the introduction of negative supercoils by DNA gyrase, an energetically unfavourable reaction, the requirement for an energy input from the hydrolysis of ATP is clear. Experimental evidence suggests that the extent of DNA supercoiling is limited by the free energy of hydrolysis of ATP i.e. a thermodynamic limit is imposed (Cullis et al., 1992). Each successive superhelical turn introduced requires an increasing amount of energy until the energy requirement exceeds that
available from the hydrolysis of ATP, where DNA gyrase is thought to hydrolyse two ATP molecules per round of supercoiling (Sugino & Cozzarelli, 1980). Using an ATP analogue, ATPαS (S<sub>p</sub>), with an estimated free energy of hydrolysis ~1 kJ/mol greater than for ATP, gyrase is able to introduce one additional superhelical turn into its DNA substrate. The additional ~2 kJ/mol provided by the hydrolysis of two molecules of ATPαS is consistent with the calculated energy change between consecutive steps in the supercoiling reaction (see Cullis et al., 1992, and references therein). However, the situation appears more complex than this, since the non-hydrolysable ATP analogue ADPNP is able to support a single round of supercoil introduction (Tamura et al., 1992). The precise mechanism of energy coupling is therefore unclear.

The requirement for ATP hydrolysis in eukaryotic type II topoisomerases is less obvious, since the relaxation of supercoils is an energetically favourable process. Several possible roles have been suggested for the involvement of ATP in the catalytic cycle of these enzymes. Firstly, ATP binding and hydrolysis are known to cause the closing and opening of the ATP-operated clamp, respectively. The energy input from ATP hydrolysis may serve to drive the enzyme through an ordered series of conformational changes that drive the strand passage reaction (Liu et al., 1980). The hydrolysis of ATP may serve to make the reaction more efficient, or to impart directionality on the reaction. Alternatively, ATP binding and hydrolysis may be used by the enzyme to transiently disrupt dimer interfaces, whilst maintaining a high fidelity in the reaction (Tennyson & Lindsley, 1997).

Alternatively, it has been suggested that the energy from ATP hydrolysis is required to drive the reactions of these enzymes away from their equilibrium positions (Rybenkov et al., 1997). The equilibrium distribution of DNA topoisomers was determined in the presence of ATP-dependent type II topoisomerases and an ATP-independent type I topoisomerase. The type II enzymes were able to reduce the level of supercoiling below the equilibrium level established by the type I enzymes, and also to generate a distribution of topoisomers with a narrower range of linking numbers. Also, the type II enzymes were able to decatenate singly-linked dimers to a steady-state level ~16-fold lower than the value at equilibrium in the absence of enzyme. Therefore, the
type II topoisomerases use the energy of ATP hydrolysis to perturb the equilibrium distribution of DNA topoisomers: DNA gyrase shifts the equilibrium towards supercoiled DNA, whereas eukaryotic enzymes shift the equilibrium towards relaxed DNA. To account for this behaviour, Rybenkov et al. (1997) proposed that type II enzymes bind to three DNA segments in order to specifically recognise G and T segments on different DNA rings of a dimeric catenane. Initially, the enzyme binds to the G segment and another, distant DNA segment at a DNA crossover. The crossover complex then slides along the DNA, presumably using the energy from the hydrolysis of ATP; in the case of decatenation, movement of the crossover complex may lead to a part of the second DNA ring destined to become the T segment being constrained within a shorter loop of the crossover-bound circle and thus make T segment capture more probable.

Recently, compelling evidence has been provided for a direct role for ATP hydrolysis in DNA transport by type II topoisomerases. Following on from pre-steady-state analysis of ATP hydrolysis by yeast topoisomerase II (Harkins et al., 1998; Harkins & Lindsley, 1998; reviewed in section 1.5.2), in which the enzyme was shown to rapidly bind two ATP molecules and hydrolyse them sequentially, Baird et al. (1999) examined the rate of single turnover DNA decatenation under conditions where the enzyme can hydrolyse zero, one, or two molecules of ATP. These experiments showed that topoisomerase II hydrolyses one of its two bound ATP molecules ~5-fold faster than it unlinks DNA catenanes. In a heterodimeric enzyme with one wild type subunit and one mutant subunit unable to hydrolyse ATP, the decatenation rate constant is essentially the same; however, with a homodimeric mutant enzyme, or using ADPNP, the rate of decatenation drops ~20-fold. Taken together, these results suggest that topoisomerase II hydrolyses one ATP before transporting the T segment; although ATP binding is sufficient for DNA transport, hydrolysis of one ATP accelerates the rate of DNA transport. Furthermore, in the presence of vanadate, which inhibits the release of ADP, the decatenation rate constant was unchanged. This result suggests that DNA transport occurs after hydrolysis of the first ATP and before release of the first ADP. The model shown in Figure 1.10 for the catalytic mechanism of type II topoisomerases has therefore been adapted to take account of these new results, and is shown in Figure 1.12.
Figure 1.12 Hydrolysis of one ATP drives DNA transport by eukaryotic topoisomerase II.

The topoisomerase dimer is represented the same way as in Figure 1.10; the domains of each monomer are numbered starting from the N-terminus (the catalytically dispensable C-terminal domains are not shown). The G and T segments are green and purple, respectively. Binding of one ATP is sufficient to cause the N-terminal domains to dimerise and capture a T segment (B); binding of a second ATP to the other monomer does not cause a change in conformation (C). Hydrolysis of one of the ATPs and/or release of the P_1 formed causes a conformational change in both subunits of the enzyme that drives the T segment through the broken G segment (D). The ADP is released before the second ATP is hydrolysed. Exit of the T segment is not shown as it is currently not clear at which stage this occurs. (Baird et al., 1999).
It has been shown that both monomers of topoisomerase II adopt the same overall conformation whether one or two ATPs are bound (Lindsley & Wang, 1993a). Therefore, binding of ATP to one ATPase domain causes a conformational change in both domains that allows them to dimerise and capture the T segment (Figure 1.12; B). Hydrolysis of one ATP triggers a conformational change that also affects both monomers, and drives the T segment through the transiently cleaved G segment (Figure 1.12; D). Hydrolysis of the second ATP may be blocked until another conformational change returns the ATPase domains to a hydrolysis-competent state. This model is also sufficient to explain the limited strand passage activity seen with the nonhydrolysable analogue ADPNP; the conformational state shown in Figure 1.12 (B) persists in the presence of ADPNP, and the enzyme-DNA complex may slowly sample altered conformations, including that shown in Figure 1.12 (D). Alternatively, ATP binding may stabilise the gate open conformation, and ATP hydrolysis may be responsible for driving the T segment through the gate.

1.5 THE ATPase REACTION OF TYPE II TOPOISOMERASES

The type II topoisomerases represent a class of DNA-dependent ATPases that function as relatively simple molecular machines. Although the precise series of events that occur in the ATP hydrolysis cycle are unknown, and indeed, in the case of eukaryotic type II topoisomerases, the role for ATP hydrolysis remains unclear, the ATPase activity of these enzymes has been extensively studied in the hope of understanding more about their mode of operation.

1.5.1 DNA gyrase

The ATPase activity of the E. coli DNA gyrase protein has been investigated (Mizuuchi et al., 1978; Sugino & Cozzarelli, 1980; Staudenbauer & Orr, 1981; Maxwell & Gellert, 1984). In the absence of DNA, the GyrB protein has an intrinsic ATPase activity that is very low; this activity is greatly increased by pre-treatment of the enzyme with urea or heat, which may be due to a conformational change resulting in a more
active ATPase (Maxwell & Gellert, 1984). The intrinsic ATPase activity of the GyrB protein has a greater than first-order dependence on enzyme concentration, suggesting an oligomerisation process occurs during the hydrolysis cycle (Staudenbauer & Orr, 1981). Calculated values for $K_m$ and $k_{cat}$ are 1.7 mM and 1 s$^{-1}$, respectively (Staudenbauer & Orr, 1981), although the value for the $k_{cat}$ is likely to be an overestimate because urea was used in the purification of the enzyme. Indeed, Maxwell & Gellert (1984) report an intrinsic ATPase rate of $\sim$0.05 s$^{-1}$ at 25°C. Addition of the GyrA protein or DNA individually to the GyrB protein does not affect the low level of ATPase activity. However, in the presence of GyrA and dsDNA, the ATPase activity of GyrB is increased more than 20-fold (Maxwell & Gellert, 1984). Relaxed or linear DNA stimulate the ATPase activity to a greater extent than supercoiled DNA, and no sequence specificity was observed for the stimulation (Sugino & Cozzarelli, 1980). Single-stranded DNA was found to be a poor stimulator of the ATPase activity (Sugino & Cozzarelli, 1980). In the presence of either two linear DNA fragments of 117 bp or 171 bp, or lambda DNA, calculated values of $K_m$ and $k_{cat}$ ranged from 0.21 mM to 0.45 mM, and 0.65 s$^{-1}$ to 1.4 s$^{-1}$, respectively.

The DNA length dependence of the ATPase activity has also been investigated (Maxwell & Gellert, 1984). At low DNA concentrations, the minimum length of DNA required to stimulate the ATPase activity is approximately 70-85 bp, with greater than $\sim$100 bp required to give full stimulation. However, stimulation of the ATPase rate is seen using much higher concentrations of a 55 bp DNA fragment. Furthermore, a sigmoidal dependence of ATPase activity upon DNA concentration was observed with this 55 bp fragment, whereas a hyperbolic relationship exists with longer DNA fragments. Analysis of protein-DNA binding by gel retardation assays showed that DNA gyrase bound DNA fragments of $>100$ bp with high efficiency, but bound fragments $<70$ bp with very low efficiency, and that a similar sigmoidal relationship also exists between DNA binding and DNA concentration for the 55 bp fragment. These results are consistent with a model in which binding of DNA to a minimum of two sites on the enzyme is required to maximally stimulate the ATPase reaction. A single DNA fragment of greater than 100 bp is able to contact both DNA binding sites, whereas with a 55 bp fragment, two DNA molecules must bind to the enzyme in order
to occupy both binding sites. The sigmoidal relationships for ATPase activity and DNA binding suggest that positive cooperativity exists between the two binding sites.

1.5.2 Eukaryotic topoisomerase II

Yeast DNA topoisomerase II, like all type II DNA topoisomerases with the notable exception of DNA gyrase, requires the energy from the hydrolysis of ATP in order to relax supercoiled DNA. The ATPase reaction of the whole yeast enzyme has been studied, and the coupling to DNA transport investigated (Lindsley & Wang, 1993b). In the absence of DNA the enzyme has weak ATPase activity which exhibits Michaelis-Menten kinetics. Values for $K_m$ and $k_{cat}$ are calculated to be 0.3 mM and 0.4 $s^{-1}$/dimer at 30°C, respectively, although the actual $k_{cat}$ is expected to be two to three times higher (i.e. $\sim 1$ $s^{-1}$/dimer) because of the presence of inactive enzyme molecules. Binding of DNA to the enzyme increases the $k_{cat}$ by approximately 20-fold, and decreases the $K_m$ to 0.13 mM. This stimulation of the ATPase activity seen in the presence of DNA has been largely attributed to binding of the G segment to the enzyme (Wang, 1998). In the DNA-bound enzyme, the reaction kinetics deviate from classic Michaelis-Menten kinetics; the binding of ATP to the two ATPase sites appears to be cooperative (Lindsley & Wang, 1993b). Binding of nucleotide to DNA gyrase also occurs via cooperative interactions between the two nucleotide binding sites in the DNA-bound enzyme (Tamura et al., 1992).

The ATPase activity of yeast topoisomerase II has also been subjected to pre-steady state kinetic analysis using chemical quench and pulse-chase techniques (Harkins & Lindsley, 1998; Harkins et al., 1998). In the presence of DNA, yeast topoisomerase II displays burst kinetics in hydrolysing ATP. At saturating ATP concentrations, the enzyme binds to two ATP molecules and hydrolysers them both before either can dissociate. However, chemical quench experiments revealed a rapid burst in ADP production before a steady-state rate was reached, and showed that ATP hydrolysis is sequential, the enzyme rapidly hydrolysers at least one ATP before the rate determining step is reached (Harkins & Lindsley, 1998). No such burst phase was seen in the absence of DNA. Therefore, in the absence of DNA either ATP binding or hydrolysis is
rate determining; however, the presence of DNA increases the rate of either ATP binding or hydrolysis such that a later step becomes rate determining. Further analysis revealed that following hydrolysis of the first ATP, the products (ADP and P_i) are released before the second is hydrolysed; in fact, release of P_i occurs before ADP (Harkins et al., 1998). Release of ADP and P_i represents the rate determining step for ATP hydrolysis, and release of the first ADP and P_i may be coupled to a conformational change that results in DNA transport. The presence of DNA was shown to stimulate the rate of ATP binding, which is reasonable given that in the absence of DNA slow binding of ATP would increase the probability of binding to DNA before the N-terminal clamp closes.

The ATPase reaction of human topoisomerase IIα has been studied in this laboratory (Hammonds & Maxwell, 1997). The enzyme was found to have an intrinsic ATPase rate of approximately 0.05 ATP molecules hydrolysed/s/dimer at 37°C. This rate was found to be stimulated approximately 10-fold by DNA at >1000 bp per enzyme dimer; no difference was seen when using supercoiled, relaxed or linear DNA. Both the DNA-dependent and independent ATPase activities obey the Michaelis-Menten equation; no evidence of cooperativity in ATP binding was found. This is in contrast to the yeast enzyme, which does show cooperativity in ATP binding (Lindsley & Wang, 1993b; Harkins et al., 1998); given that there are also two ATP binding sites in the human enzyme, some degree of cooperativity might be expected, if not observed. The DNA-stimulated ATPase rate was found to be dependent on DNA concentration i.e. the bp:enzyme dimer ratio. With plasmid pBR322, the DNA-stimulated ATPase rate increases as the bp:dimer ratio increases, up to ~150 bp/dimer, after which the ATPase rate decreases and reaches a plateau level (~10-fold higher than the intrinsic level) at ~400-500 bp/dimer. A similar DNA-dependent effect was seen in experiments with the yeast enzyme (Hammonds & Maxwell, 1997). Using two different bp:dimer ratios, calculated values for K_m and k_cat are: at 130 bp/dimer, 0.78 mM and 2.17 s⁻¹; at 2000 bp/dimer, 0.56 mM and 0.59 s⁻¹. A similar pattern of DNA length-dependent ATPase is seen when using short DNA fragments of 40-890 bp. No stimulation of the ATPase rate is seen with a 40 bp fragment, except at greater than 1000 DNA molecules per dimer. Approximately 100 bp are required to stimulate ATPase activity; maximal
stimulation was seen with DNA of approximately 300 bp, while DNA of greater than 400 bp led to a decrease in the ATPase rate.

These results are interpreted in terms of the enzyme requiring DNA fragments to be bound to both the DNA gate and the ATP-operated clamp (the ‘G&T’ complex) to stimulate the ATPase activity, as was seen with DNA gyrase (Maxwell & Gellert, 1984). As the DNA length is increased from 108 bp to 263 bp, the ATPase rate increases gradually, this may be due to the increased flexibility of the DNA allowing the DNA to bind to both the gate and clamp, forming the G segment and T segments, respectively. The hyperstimulation with ~300 bp linear fragments (or ~150 bp/dimer with relaxed pBR322) is thought to be a consequence of protein-protein interactions due to the saturation of DNA with enzyme (Hammonds & Maxwell, 1997). Under conditions of saturation, two topoisomerase dimers may be forced close enough together so as they interact with one another, leading to the hyperstimulated ATPase rate, an effect which diminishes as the bp:dimer ratio increases. Alternatively, the effect may not be directly due to protein-protein interactions, but rather to stabilisation of the G&T complex, leading to a permanently captured T segment and an increased ATPase rate.

1.5.3 Coupling of ATP usage to T segment capture and strand passage

It has been shown for both E. coli DNA gyrase and yeast topoisomerase II that the efficiency of strand passage is dependent, at least in part, upon the topological state of the DNA (Bates et al., 1996; Roca & Wang, 1996; Kampranis et al., 1999). For DNA gyrase, with positively supercoiled DNA as substrate, the coupling efficiency of ADPNP binding to strand passage (DNA supercoiling) approaches 100% (Bates et al., 1996). The extent of coupling is reduced with relaxed DNA to ~30%, and becomes undetectable with negatively supercoiled DNA with a specific linking difference of −0.046. For yeast topoisomerase II, coupling of ADPNP binding to strand passage (DNA relaxation) approaches 100% with a moderately negatively supercoiled DNA substrate (Roca & Wang, 1996); this falls to ~25% with a relaxed DNA substrate. Furthermore, for DNA gyrase, the efficiency of strand passage depends upon an on-enzyme equilibrium that is dependent upon DNA topology, since T segment capture occurs with
high efficiency irrespective of the topological state of the DNA (Kampranis & Maxwell, 1999).

The number of ATP molecules hydrolysed per strand passage event has been measured for yeast topoisomerase II (Lindsley & Wang, 1993b). At a saturating ATP concentration (1 mM) and low temperatures (7-17°C) to slow down the rate of reaction, it was calculated that 7.4 ± 1 ATP molecules were hydrolysed per DNA transport event. At low ATP concentrations (5-25 μM) and at 30°C, the enzyme hydrolysed an average of 1.95 ± 0.5 ATP molecules per DNA transport event, with a lowest value of 1.2 ATPs per transport event. Thus, it is likely that at high concentrations of ATP, the rate limiting step in the catalytic cycle is in the capture and transport of the DNA segment. At low concentrations of ATP, the rate limiting step is in the hydrolysis of ATP and hence the degree of coupling is high. Several lines of evidence suggest that binding of ATP to only one of the two N-terminal domains is sufficient to cause these domains to dimerise and for strand passage to occur. Firstly, using a heterodimer of yeast topoisomerase II consisting of one wild-type and one mutant polypeptide unable to bind ATP, it was shown that binding of ADPNP to one subunit is sufficient to trigger the closure of the N-terminal clamp (Lindsley & Wang, 1993a). The results described above at low ATP concentrations suggest that hydrolysis of 1-2 ATP molecules is sufficient for DNA transport. Furthermore, a mutant E. coli DNA gyrase containing one wild-type GyrB and one mutant GyrB able to bind but not hydrolyse ATP, has shown that hydrolysis of ATP can proceed at one site only, and that this is sufficient to drive catalytic DNA supercoiling, albeit at a lower limit (Kampranis & Maxwell, 1998b). This scenario is unlikely to occur in vivo however, due to the demonstrated positive cooperativity in ATP binding (Tamura et al., 1992; Lindsley & Wang, 1993b), and the relatively high physiological ATP concentration.

No information is available on the coupling efficiency of the human enzymes. However, in studies of the ATPase and decatenation activities of human topoisomerase IIα (Hammonds & Maxwell, 1997) and yeast topoisomerase II (Lindsley & Wang, 1993b), it has been observed that human topoisomerase IIα hydrolyses ATP at a 15-fold slower rate than yeast topoisomerase II, yet the human enzyme is approximately 5-fold
more active in decatenation. These results would appear to suggest that the human enzyme is more efficient in coupling ATP hydrolysis to strand passage, at least under the experimental conditions employed.

The role of ATP binding and hydrolysis in the cleavage and religation of the G segment, and *vice versa*, is presently not clear due to conflicting reports (reviewed in Wang, 1998). Covalent adduct formation between DNA and topoisomerase II is often enhanced by the presence of ATP or ADPNP (Sugino *et al.*, 1978; Osheroff, 1986), which suggests that G segment cleavage is stimulated by nucleotide. However, G segment cleavage is not dependent upon the presence of ATP (see for example, Sander & Hsieh, 1983), and mutants lacking ATPase activity (Jackson & Maxwell, 1993; Lindsley & Wang, 1993b), or indeed lacking the ATPase domain itself (Berger *et al.*, 1996), have DNA cleavage activities similar to the wild-type enzyme. Important, but contrasting, information has come from an active site mutant of yeast topoisomerase II where the tyrosine involved in DNA cleavage has been replaced by phenylalanine (Y782F). This protein was originally reported to have no DNA-stimulated ATPase activity (Hammonds & Maxwell, 1997), suggesting that DNA-stimulated hydrolysis of ATP is dependent upon G segment cleavage. However, more recent kinetic analysis of the DNA-dependent ATP hydrolysis by this mutant provides evidence to the contrary (Morris *et al.*, 1999). The $k_{cat}$ for ATP hydrolysis for the yeast Y782F mutant enzyme was increased 2-3-fold in the presence of DNA, compared to 10-11-fold for the wild-type enzyme. The mutant enzyme displayed a similar burst in ADP production as the wild-type enzyme in the presence of DNA, but was shown to hydrolyse the first ATP 3-4-fold slower than the wild-type enzyme, despite no significant decrease in ATP binding. Wild type and mutant proteins bound to a single DNA segment (the G segment) with equal affinities, but the mutant was unable to trap a second DNA segment (the T segment), suggesting that G segment cleavage and gate opening are required for T segment capture. To explain these results, two possibilities were proposed. Firstly, the mutant protein cannot undergo the normal series of conformational changes that usually accompanies strand passage, and ATP hydrolysis (but not ATP binding) may be coupled to a DNA cleavage-dependent conformational change. Alternatively, the inability of the mutant protein to productively bind the T segment may account for the reduction in
DNA-stimulated ATPase activity: binding of the G segment stimulates ATP binding, whereas binding of the T segment stimulates ATP hydrolysis. It can be seen that these two hypotheses are not mutually exclusive.

The role of the T segment in stimulating cleavage of the G segment, and in separating the two ends, is equally controversial. In the clamp-closed conformation, yeast topoisomerase II can bind to and cleave a single linear DNA molecule (Roca & Wang, 1992), suggesting that T segment binding is not required for G segment cleavage. However, experiments with Drosophila topoisomerase II and a 40 bp DNA fragment show that the extent of DNA cleavage has a sigmoidal dependence on the DNA concentration (Corbett et al., 1992b), suggesting that the enzyme interacts with more than one DNA fragment (as was seen with DNA gyrase and a 55 bp DNA fragment), and that cleavage of the G segment is stimulated by the presence of the T segment. This difference may represent the two different steps in DNA cleavage; namely, covalent attachment of the enzyme to the DNA ends (chemical cleavage), and the pulling apart of the DNA ends to allow the passage of the T segment. Furthermore, the results presented above with the DNA cleavage mutant Y782F (Morris et al., 1999) suggest that G segment cleavage is required for, and may be coupled to, T segment capture. Analysis of the crystal structure of the N-terminal fragment of E. coli DNA gyrase (Figure 1.8) suggests that the central cavity is not large enough to accommodate a duplex DNA (Tingey & Maxwell, 1996); steric repulsion may be responsible for driving the opening of the DNA gate when the T segment is bound (Berger et al., 1996).

1.6 ROLE OF THE N-TERMINAL ATPase DOMAIN

The role of the N-terminal domains in the catalytic cycle of type II topoisomerases, as originally suggested by Roca & Wang (1992, 1994), is to act as a molecular clamp, and in so doing capture a DNA segment destined to be transported through the G segment; this process is coupled to binding of ATP to the N-terminal domains. The exact role for ATP hydrolysis in the strand passage reaction is unknown, and may serve to accelerate the strand passage reaction (Baird et al., 1999), but
hydrolysis of ATP is required for subsequent opening of the molecular clamp so that the enzyme may be reset for another cycle. Furthermore, the intrinsic ATPase activity of these enzymes is stimulated by DNA.

Evidence for the involvement of the T segment in the DNA-dependent ATPase activity of type II topoisomerases can be summarised as follows. Firstly, in the presence of short (≤ 55 bp) DNA fragments, stimulation of the ATPase activity of both human topoisomerase IIα and *E. coli* DNA gyrase displays a sigmoidal relationship with respect to DNA concentration, suggesting that the G and T segments must be bound for maximum stimulation (Maxwell & Gellert, 1984; Hammonds & Maxwell, 1997). Further evidence comes from an *E. coli* DNA gyrase enzyme bearing a mutation within the 43 kDa ATPase domain (R286Q); this region of the protein lines the cavity seen in the crystal structure of the same fragment (Figure 1.8), and contains several arginine residues that were proposed to form a DNA-binding surface (Wigley et al., 1991). This residue was chosen because its side chain projects into the cavity seen in the crystal structure of this domain, constricting the diameter of the hole from >20 Å to ~13 Å, and may interact with the T segment during strand passage (Tingey & Maxwell, 1996). The R286Q mutant has almost no DNA supercoiling activity and an intrinsic ATPase activity that cannot be stimulated by DNA, whereas relaxation and DNA cleavage activities are near-normal (Tingey & Maxwell, 1996). These results suggest that binding of DNA within this clamp may be required for stimulation of the ATPase activity. Also, as described above, it has been suggested that the reduced rate of DNA-stimulated ATPase activity for the DNA cleavage mutant of yeast topoisomerase II (Y782F) is due to its inability to capture the T segment (Morris et al., 1999). Although it would seem beneficial to ensure that the maximal rate of ATP hydrolysis could not occur until both the G and T DNA segments were present (Tingey & Maxwell, 1996), it is also worthy of note that if this were the case, a G segment-bound enzyme which binds ATP and closes the clamp without capturing a T segment would be reliant upon the intrinsic ATPase activity to free itself from this nonproductive state (Wang, 1998).

The N-terminal domains of three type II topoisomerases have been cloned, overexpressed and characterised (Ali et al., 1993, 1995; Gardiner et al., 1998; Olland &
Table 1.1 Comparison of N-terminal fragments of *E. coli* DNA gyrase, *Saccharomyces cerevisiae* topoisomerase II and human topoisomerase IIα.

Shown in the table is a summary of the biochemical properties of the N-terminal fragments of *E. coli* DNA gyrase GyrB (Ali et al., 1993), *Saccharomyces cerevisiae* topoisomerase II (Olland & Wang, 1999) and human topoisomerase IIα (Gardiner et al., 1998).

<table>
<thead>
<tr>
<th>PROPERTIES</th>
<th>DNA GYRASE</th>
<th>YEAST</th>
<th>HUMAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residues</td>
<td>1-393 (GyrB)</td>
<td>1-409 (+ His-tag)</td>
<td>1-439 (+ His-tag)</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>43 kDa</td>
<td>47 kDa</td>
<td>52 kDa</td>
</tr>
<tr>
<td>Expression:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host</td>
<td>E. coli Soluble</td>
<td>S. pombe Soluble</td>
<td>E. coli Insoluble</td>
</tr>
<tr>
<td>Form</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrinsic ATPase</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DNA-stimulated ATPase</td>
<td>No</td>
<td>Yes ([NaCl] dependent)</td>
<td>Yes</td>
</tr>
<tr>
<td>Oligomeric state:</td>
<td>Monomer Dimer</td>
<td>?</td>
<td>Dimer Dimer</td>
</tr>
<tr>
<td>– ADPnP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ ADPnP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binds DNA</td>
<td>No</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>
Wang, 1999); the properties of these fragments are summarised in Table 1.1. All three protein fragments were found to have an intrinsic ATPase activity. In the following two sections, the enzymic properties of these three protein fragments are discussed in terms of the proposed role of the N-terminal ATPase domain in the catalytic cycle of type II topoisomerases, namely ATP binding and hydrolysis, and capture of the T segment.

1.6.1 ATP binding and hydrolysis

The first N-terminal fragment to be cloned and characterised was the 43 kDa N-terminal domain of the GyrB protein of *E. coli* DNA gyrase (Ali et al., 1993), of which the crystal structure has been previously determined in the presence of ADPNP (Wigley et al., 1991). This fragment was found to have an intrinsic ATPase activity with a greater than first-order dependence on enzyme concentration, indicative of non-Michaelian kinetics, and suggesting that ATP hydrolysis might be dependent upon an oligomerisation process. Molecular weight studies showed that the protein is a monomer in solution in the absence of nucleotide, but dimerises in the presence of ADPNP. In the presence of ATP, the protein eluted at a position intermediate between monomer and dimer, possibly suggesting a monomer-dimer equilibrium; ADP and P\textsubscript{i} led to the protein being eluted as a monomer. These results are consistent with a scheme whereby ATP binding to the 43 kDa monomer causes dimerisation into the active ATP-hydrolysing form, dimerisation represents the rate-limiting step in the ATPase reaction of this 43 kDa fragment. Addition of DNA had no effect on the ATPase activity of this fragment, alone or in the presence of the GyrA protein.

Subsequently, the N-terminal domains of human topoisomerase II\textalpha{} and yeast topoisomerase II were cloned and overexpressed, and these fragments were also shown to have an intrinsic ATPase activity (Gardiner et al., 1998; Olland & Wang, 1999). In the case of the human enzyme, the position of the domain boundary was determined based on sequence similarity with the equivalent region of *E. coli* DNA gyrase (Gardiner et al., 1998). The yeast N-terminal fragment has an ATPase activity which, like the equivalent region in *E. coli* DNA gyrase, shows a parabolic dependence on protein concentration, again indicating a monomer-dimer shift, with the active form of
the enzyme being a dimer (Olland & Wang, 1999). The ATPase activity was significantly higher at 10 mM NaCl than at 100 mM NaCl, although the same parabolic dependence on protein concentration was seen. However, with the human N-terminal fragment, and an NaCl concentration of 100 mM, the ATPase activity was linearly dependent on protein concentration (Gardiner et al., 1998). Furthermore, chemical cross-linking revealed that the protein is a dimer in solution in both the presence and absence of nucleotide (ADPNP, ATP or ADP). These results suggest that the human fragment is predominantly a dimer in solution and that dimerisation is not the rate-limiting step in the ATPase catalytic cycle. The catalytic residue in the ATP hydrolysis reaction was identified as Glu86 by site-directed mutagenesis (Gardiner et al., 1998). Calculated values of $K_M$ and $k_{cat}$ for the human fragment are 0.47 mM and 0.018 s$^{-1}$ respectively. A second N-terminal fragment of yeast topoisomerase II, representing residues 1-660, has an ATPase activity linearly dependent on protein concentration, suggesting that this longer fragment is dimeric (Olland & Wang, 1999).

In contrast to the N-terminal fragment of *E. coli* DNA gyrase, the intrinsic ATPase activity of the yeast and human N-terminal fragments are stimulated by DNA (Gardiner et al., 1998; Olland & Wang, 1999). For the human fragment, the ATPase activity was found to be stimulated by DNA approximately 5-10-fold, with the rate linearly dependent on enzyme concentration. Values of $K_M$ and $k_{cat}$ are 0.40 mM and 0.11 s$^{-1}$ respectively, indicating an ~6-fold increase in $k_{cat}$ (Gardiner et al., 1998). For the yeast N-terminal fragment, the ATPase activity is stimulated by DNA in the 10 mM NaCl buffer but not in the 100 mM NaCl buffer, and in the low salt buffer the ATPase rate becomes linearly dependent upon protein concentration in the presence of DNA (Olland & Wang, 1999). These results can be interpreted in terms of the role of the ATPase domains in capturing the DNA T segment. For the DNA-stimulated ATPase activity of the human fragment, the calculated $k_{cat}$ value is less than that for the full-length enzyme (Hammonds & Maxwell, 1997), therefore, binding of DNA at the DNA gate may also be required for full stimulation, as has been previously postulated (Maxwell & Gellert, 1984; Hammonds & Maxwell, 1997). Alternatively, stimulation of the ATPase rate by DNA may not be due to direct interaction between the protein and
the DNA, but rather due to DNA-induced dimerisation of the ATPase domains (Olland & Wang, 1999).

Given the role of the N-terminal domains in the catalytic cycle, clearly it is essential that dimerisation precedes ATP hydrolysis. To this end, the ATPase catalytic centre of each N-terminal domain contains residues from the N-terminal ‘arm’ of the other protomer in E. coli DNA gyrase (Wigley et al., 1991; see Figure 1.8). This N-terminal arm of 14 residues interacts with the second GyrB fragment and stabilises the dimeric form, and Tyr5 interacts with the ribose sugar of the bound ADPNP. The role of this N-terminal arm in the dimerisation and ATPase activity of the 43 kDa N-terminal fragment of E. coli DNA gyrase has been investigated (Brino et al., 2000). Mutation of Tyr5 to phenylalanine or serine reduces, but does not abolish, the ability of the N-terminal fragments to dimerise. The ATPase activity of these mutants were also reduced but not abolished. Mutation of a second highly conserved residue, Ile10, significantly reduced the ability of the protein to dimerise and to hydrolyse ATP. Deletion of the N-terminal 14 residues completely abolished the ability of the N-terminal fragment to dimerise and hydrolyse ATP. This mechanism of ATP hydrolysis helps to ensure the fidelity of the strand passage reaction i.e. bound ATP cannot be hydrolysed until the N-terminal domains dimerise, and in so doing capture the T segment which is subsequently transported through the G segment.

1.6.2 Capture of a DNA segment

We have seen above how ATP binding to one or both N-terminal domains of a type II topoisomerase causes the closing of the protein clamp in these enzymes. Despite strong evidence that suggests that this ATP-operated clamp closure is required for capture of the DNA segment that is destined to become the T segment (Roca & Wang, 1992, 1994), very little is understood about the way in which the N-terminal domains interact with the T segment. For E. coli DNA gyrase, capture of the T segment following addition of ADPNP occurs with high efficiency irrespective of the topological state of the DNA (Kampranis & Maxwell, 1999); whether the same situation obtains with eukaryotic topoisomerase II is unknown. Addition of ADPNP to yeast
topoisomerase II bound to a relaxed DNA substrate leads to strand passage in ~25% of the enzyme molecules (Roca & Wang, 1996). This suggests that there is only a weak interaction between the N-terminal domains and the T segment; in the open conformation, the T segment may move in and out of the protein clamp unhindered (Wang, 1998).

The 43 kDa domain of *E. coli* DNA gyrase does not bind DNA, as determined by gel retardation assays, consistent with the inability of DNA to stimulate the ATPase activity of this fragment (Ali *et al.*, 1993). However, the reduced DNA supercoiling activity and DNA-independent ATPase of the gyrase mutant R286Q (see above; Tingey & Maxwell, 1996) can be rationalised in terms of this residue being involved in binding to a DNA segment (the T segment) trapped within the clamp. The simplest interpretation of the DNA-stimulated ATPase activity of the human and yeast N-terminal fragments is that binding of the T segment to these domains, in isolation and presumably in the full-length enzyme, leads to an increase in the rate of ATP hydrolysis which is somehow coupled to strand passage.

### 1.7 AIMS

A great deal is known about the biochemistry and enzymology of the DNA topoisomerases, and information gleaned from these studies coupled with information from structural studies provides clues to the mechanism of these enzymes. One of the most interesting aspects of these enzymes is how they couple the binding and hydrolysis of ATP to their manipulation of dsDNA. Functionally, the type II topoisomerases consist of an N-terminal ATPase domain, a DNA cleavage/religation domain and a catalytically dispensable C-terminal domain. The N-terminal domains act as ATP-operated clamps that capture a DNA segment and direct it through a second cleaved DNA strand.

In studying such complex, multimeric enzymes as the DNA topoisomerases, it is often easier and more informative to study each functional domain in isolation, and
relate this back to the whole enzyme. The N-terminal domain of the GyrB protein of *E. coli* DNA gyrase has been identified as the ATP-hydrolysing domain of this enzyme. This 43 kDa domain has been characterised extensively in this laboratory (Ali *et al.*, 1993, 1995), and the crystal structure solved (Wigley *et al.*, 1991), and this has provided vital information on the functioning of this enzyme. The N-terminal domain of human DNA topoisomerase IIα is homologous to that of GyrB of *E. coli*, and has also been shown to be the ATP-hydrolysing domain of this enzyme (Gardiner *et al.*, 1998). Despite problems encountered with recovering suitable levels of protein for biochemical analysis, these studies have shown marked differences between this domain and the equivalent domain in GyrB. It is likely that these differences are due to the differing reaction mechanisms for these two proteins: DNA gyrase introduces supercoils into a DNA substrate in the presence of ATP and relaxes supercoils in the absence of ATP, whereas eukaryotic topoisomerase II relaxes supercoiled DNA only in the presence of ATP.

The type II topoisomerases are DNA-stimulated ATPases, but the mechanics of this reaction are poorly understood. The catalytic cycle requires that a minimum of two DNA segments be manipulated by the enzyme to effect the strand passage reaction, which must be coupled to the ATPase activity. The initial aim of this project was to determine the role that one of these DNA segments, the T segment, plays in the ATPase reaction of the N-terminal ATPase domain of human DNA topoisomerase IIα. Given the problems encountered with the recovery of significant amounts of consistently active 52 kDa protein, it was decided that the N-terminal domain of human DNA topoisomerase IIα would be recloned as a fusion protein in an attempt to increase the yield and solubility of this protein fragment. The ATPase activity of a variety of N-terminal fragments have been examined in the presence and absence of DNA of varying lengths. This study has again highlighted an interesting relationship between DNA length and the extent of stimulation of the ATPase activity, a phenomenon previously referred to as hyperstimulation (Hammonds & Maxwell, 1997).

Since N-terminal clamp closure and DNA capture are paramount to the strand passage reaction of eukaryotic type II topoisomerases, these properties of the N-terminal
domain have also been investigated. The oligomerisation properties of the N-terminal domain have been determined in the light of the monomer-dimer equilibrium seen with the equivalent domain of the *E. coli* DNA gyrase protein (Ali *et al.*, 1993). The ability of the human N-terminal domain to bind DNA, predicted by the stimulation of ATPase activity seen in the presence of DNA, is also tested.
CHAPTER 2

Materials and Methods
2.1 BACTERIOLOGY

2.1.1 Bacterial strains

The following *E. coli* strains were used during this project.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>B834(DE3)</td>
<td><em>E. coli</em> B F⁻ <em>ompT hsdS₉ (rB⁻ mB⁻) gal dcm met</em> (DE3)</td>
<td>General expression host (methionine auxotroph)</td>
</tr>
<tr>
<td>B834(DE3)pLysS</td>
<td>As B834(DE3) plus pLysS (Cm')</td>
<td>Stringent expression host (methionine auxotroph)</td>
</tr>
<tr>
<td>BL21(DE3)</td>
<td><em>E. coli</em> B F⁻ <em>ompT hsdS₉ (rB⁻ mB⁻) gal dcm</em> (DE3)</td>
<td>General expression host</td>
</tr>
<tr>
<td>BL21(DE3)pLysS</td>
<td>As BL21(DE3) plus pLysS (Cm')</td>
<td>Stringent expression host</td>
</tr>
<tr>
<td>BL21-CodonPlus</td>
<td>As BL21(DE3) plus <em>endA Hte [argU ileY leuW Cm']</em></td>
<td>Codon-optimised expression host</td>
</tr>
<tr>
<td>XL1-Blue</td>
<td><em>E. coli</em> K-12 recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacF7ΔM15 Tn 10 (Tet')]*</td>
<td>General cloning host</td>
</tr>
</tbody>
</table>

Strains were stored either as glycerol stocks at -80°C or according to the manufacturer’s instructions for commercially available competent cells.

2.1.2 Bacterial growth media

Media were prepared as described below and sterilised by autoclaving at 15 lb/sq. in. for 15 minutes. For agar plates, 15 g/l of bacteriological agar (OXOID) was added to the media before autoclaving.

Luria-Bertani (LB) medium: 
- Tryptone (OXOID) 10 g
- Yeast extract (OXOID) 5 g
- NaCl 10 g
- Water to 1 litre

52
Nutrient Broth (NB): Nutrient broth (OXOID) 13 g
Water to 1 litre

2YT medium: Tryptone (OXOID) 16
Yeast extract (OXOID) 10 g
NaCl 5 g
Water to 1 litre

Terrific Broth (TB): Tryptone (OXOID) 12 g
Yeast extract (OXOID) 24 g
Glycerol 4 ml
Water to 900 ml

This media was sterilised by autoclaving, and then 100 ml of a sterile solution of 0.17 M KH₂PO₄, 0.72 M K₂HPO₄ was added.

2.1.3 Antibiotics
Ampicillin (Sigma) was made up as a stock solution of 50 mg/ml in water and used at a working concentration of 50 μg/ml in growth media and agar plates. Chloramphenicol (Sigma) was made up as a stock solution of 34 mg/ml in ethanol and used at 34 μg/ml. All antibiotics were stored at −20°C.

2.1.4 Transformation of competent E. coli cells
Commercially available XL1-Blue cells (Stratagene) and BL21-CodonPlus cells (Stratagene) were transformed with plasmid DNA according to the manufacturer’s instructions. For other cell types (purchased from Novagen as a glycerol stock), competent cells were prepared and transformed as follows. A 10 ml LB broth culture was inoculated with a single colony from an LB-agar plate and grown with shaking at 37°C for ~6 hours (for strains containing pLysS chloramphenicol was present in the plates and growth media). A 1 ml aliquot of this culture was used to inoculate a 9 ml LB broth culture that was grown at 37°C for 1 hour. Cells were pelleted in a bench top
centrifuge at 2,500 revolutions per minute (rpm) for 10 minutes at 4°C. The cell pellet was drained and resuspended in 1 ml of ice cold 100 mM CaCl₂. The suspension was then transferred to 1.5 ml tubes and incubated on ice for 15 minutes. The cells were pelleted and resuspended in 400 μl of ice cold 100 mM CaCl₂. Competent cells were stored on ice until use. To transform competent cells prepared in this manner, 1 μl of plasmid DNA (~0.2 μg) from a standard plasmid miniprep (see section 2.5.2) was added to 200 μl of competent cells and incubated on ice for 30 minutes. Cells were then heat shocked at 42°C for 90 seconds in a water bath and then placed on ice for 2 minutes. An 800 μl aliquot of LB broth was added to the cells, which were then incubated at 37°C for 1 hour with shaking for phenotypic expression. An appropriate volume of transformed cells (typically 75 μl) were then spread onto LB-agar plates containing the appropriate antibiotics and grown overnight (~16 hours) at 37°C.

2.2 SYNTHESIS AND PURIFICATION OF OLIGONUCLEOTIDES

Oligonucleotides were synthesised either by PNACL (University of Leicester) on an ABI 394 model DNA synthesiser at 0.2 μmole or 1 μmole scale, or were synthesised at 40 nmole scale and supplied ready purified by PE-Applied Biosystems (Cheshire, UK). Oligonucleotides from PNACL were purified from synthetic by-products by butanol extraction as follows (Sawadogo & Van Dyke, 1991). 100 μl of synthesised oligonucleotide was mixed with 1 ml of butanol and vortexed before being centrifuged at 13,000 rpm for 1 minute. After the butanol phase was removed the DNA pellet was dried under vacuum and resuspended in TE buffer (10 mM Tris-HCl (pH 8), 1 mM EDTA).

2.2.1 Oligonucleotides used for ligation-independent cloning

Oligonucleotides used for cloning Eκ/LIC vectors:

SC01  GAC GAC GAC AAG ATG GAA GTG TCA CCA TTG GAC C
SC02  GAG GAG AAG CCC GGT TCA CTG GAC TTG GGC CTT AAA C
SC03  GAG GAG AAG CCC GGT TCA CTT CTT GTT TAA CTG GAC
SC04  GAG GAG AAG CCC GGT TCA TTG GGC CTT AAA CTT CAC C
SC05  GAG GAG AAG CCC GGT TCA CTT GTT TAA CTG GAC TTG

Oligonucleotides used for cloning Xa/LIC vectors:

SC10  GGT ATT GAG GGT CGC ATG GAA GTG TCA CCA TTG CAG C
SC11  AGA GGA GAG TTA GAG CCT CAT TGG GCC TTA AAC TTC ACC
SC12  AGA GGA GAG TTA GAG CCT CAC TTG TTT AAC TGG ACT TG
SC16  AGA GGA GAG TTA GAG CCT CAA GTG GAG TTT CGG C
SC17  AGA GGA GAG TTA GAG CCT CAC TCA GTG GAG TTT CGG
SC18  AGA GGA GAG TTA GAG CCT CAA CAC TCA GTG GAG TTT CG

2.2.2 Oligonucleotides used for DNA sequencing

Oligonucleotides used for DNA sequencing throughout the topoisomerase II-encoding region of expression plasmids:

SC06  ACA ATA CCA CAG CCA ATG GC
SC07  AGG AGT TAC CAG TTT CAT CC
SC08  CAC AAT TTG GCT CCA TAG CC
SC09  CTT CCC TAT AGT TAA TGC C
S-Tag  CGA ACG CCA GCA CAT GGA CA
T7 term. GCT AGT TAT TGC TCA GCG G

2.2.3 Oligonucleotides used for production of a 140 bp DNA fragment

Oligonucleotides used for PCR production of a 140 bp DNA fragment based around the preferred cleavage site for DNA gyrase within pBR322 (the 990 site):

140L   TCG GGG AAT TCG CAT GGC G
140R   TGG ACA GCA TGG CCT GCA A
2.3 PREPARATION OF DNA SUBSTRATES

2.3.1 Determination of DNA concentration
The concentration of DNA in solution was determined spectrophotometrically by measuring the $A_{260}$ of a suitable dilution in TE buffer, where an absorbance of 1.0 corresponds to a DNA concentration of 50 µg/ml for dsDNA (Sambrook et al., 1989). For ssDNA oligonucleotides used in section 2.3.4, DNA concentrations were determined by measuring the $A_{260}$ of a suitable dilution in TE buffer and by calculating the Molar Extinction Coefficient of the oligonucleotide using:

$$
epsilon (M^{-1} cm^{-1}) = nG \times 11800 M^{-1} cm^{-1} + nT \times 8700 M^{-1} cm^{-1} + nC \times 6100 M^{-1} cm^{-1} + nA \times 14700 M^{-1} cm^{-1}$$

(where $n =$ number of bases in the oligonucleotide sequence)

2.3.2 Preparation of plasmid pBR322
Supercoiled, relaxed and linear plasmid pBR322 were generous gifts of Mrs. A.J. Howells (University of Leicester). In brief, supercoiled plasmid pBR322 was produced by the large-scale alkaline-lysis method (Sambrook et al., 1989) from E. coli strain JM109. To separate the closed-circular plasmid DNA from contaminating nicked or linear plasmid, the sample was subjected to two rounds of CsCl density gradient centrifugation in the presence of ethidium bromide. The DNA was extracted with water-saturated butanol to remove the ethidium bromide before being precipitated with isopropanol and resuspended in TE. Relaxed plasmid pBR322 was prepared by relaxing negatively supercoiled pBR322 with chicken erythrocyte DNA topoisomerase I, followed by CsCl density gradient centrifugation to remove any supercoiled or nicked DNA. Linear plasmid pBR322 was prepared by linearisation of relaxed pBR322 with the restriction enzyme EcoRI, followed by precipitation of the DNA, which was subsequently resuspended in TE.
2.3.3 Preparation of a 140 bp DNA fragment

A 140 bp DNA fragment was prepared by PCR amplification using the following reaction conditions:

Primer 140L 1 µM
Primer 140R 1 µM
dNTPs 250 µM each
Template 140mer 3 nM
10 x Pfu buffer 10 µl
Pfu DNA polymerase 2.5 units
Water to 100 µl

The following PCR cycles were used in a Perkin-Elmer Gene Amp 9600 PCR machine:

<table>
<thead>
<tr>
<th>No. of cycles</th>
<th>Temperature (°C)</th>
<th>Time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95</td>
<td>30</td>
</tr>
<tr>
<td>25</td>
<td>95</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>68</td>
<td>30</td>
</tr>
</tbody>
</table>

The DNA was purified using a PCR purification kit (Qiagen), and the purity was assessed by electrophoresis on a 5% polyacrylamide gel (see section 2.4.2).

2.3.4 Preparation of short DNA fragments

Short dsDNA fragments of 8-70 bp were made by annealing complementary oligonucleotides based symmetrically around the preferred DNA gyrase binding and cleavage site within pBR322 (the 990 site). Shown below is the sequence of the 70 bp region around this site, with the DNA gyrase cleavage positions shown by arrows:
The complementary oligonucleotides (PNACL, University of Leicester) were mixed at 25 μM final concentration in a 200 μl volume in TE buffer and placed in a waterbath at 85°C and allowed to cool gradually to <10°C. The DNA fragments were then analysed by electrophoresis on a 5% polyacrylamide gel (see section 2.4.2).

2.4 GEL ELECTROPHORESIS OF DNA

2.4.1 Agarose gel electrophoresis

Agarose gel electrophoresis was used to determine the size, quantity and purity of plasmid DNA and DNA fragments produced by restriction enzyme analysis and of PCR fragments produced in LIC cloning. Agarose gels of 1 or 1.5% were made using TAE buffer (40 mM Tris-acetate, 2 mM EDTA). Samples were mixed with 0.5 volumes of STEB (40% sucrose, 100 mM Tris-HCl (pH 8.0), 100 mM EDTA, 0.5 mg/ml bromophenol blue), and gels were typically run in TAE buffer at 100 V for 1 hour. The gels were stained for approximately 15 minutes in an ethidium bromide solution (0.5 μg/ml in TAE), and destained in TAE for approximately 5 minutes. DNA bands were visualised on a UV light box and photographed using a Syngene gel documentation system.

2.4.2 Polyacrylamide gel electrophoresis

Short DNA fragments of 140 bp or less were analysed on 5% polyacrylamide gels (19:1 acrylamide:bis, UltraPure Accugel, National Diagnostics) using the Mini-Protean II gel system (Bio-Rad). DNA samples were mixed with 0.5 volumes of STEB
before loading. Gels were made up with TBE buffer (90 mM Tris-borate, 2 mM EDTA) and were run in TBE buffer for approximately 1 hour before staining with ethidium bromide (in TBE buffer) as for agarose gels. For gel retardation assays (see section 2.11.2), 5% polyacrylamide gels (37.5:1 acrylamide:bis, UltraPure Protogel, National Diagnostics) were made up with TBM buffer (90 mM Tris-borate, 5 mM MgCl₂) and run in TBM buffer.

2.5 LIGATION-INDEPENDENT CLONING

2.5.1 Ligation-independent cloning

Expression vectors were cloned using a Ligation-Independent Cloning (LIC) technique (Aslandis & de Jong, 1990), using a pET LIC vector kit from Novagen. LIC-cloning negates the need for a ligation reaction during cloning by using complementary single-stranded overhangs in the insert and vector that allow annealing to occur in vitro with ligation occurring in vivo following transformation into competent E. coli cells. PCR amplification of the desired region of the human topoisomerase IIα gene was performed using the primers given in section 2.2.1 and the following reaction conditions:

<table>
<thead>
<tr>
<th>Primers</th>
<th>25 pmoles each</th>
</tr>
</thead>
<tbody>
<tr>
<td>dNTPs</td>
<td>200 µM each</td>
</tr>
<tr>
<td>Template</td>
<td>20 ng</td>
</tr>
<tr>
<td>10 x <em>Pfu</em> buffer</td>
<td>5 µl</td>
</tr>
<tr>
<td><em>Pfu</em> DNA polymerase</td>
<td>1.5 units</td>
</tr>
<tr>
<td>Water</td>
<td>to 50 µl</td>
</tr>
</tbody>
</table>

The following PCR cycles were used in a Perkin-Elmer Gene Amp 9600 PCR machine:
<table>
<thead>
<tr>
<th>No. of cycles</th>
<th>Temperature (°C)</th>
<th>Time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95</td>
<td>30</td>
</tr>
<tr>
<td>30</td>
<td>95</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>68</td>
<td>240</td>
</tr>
</tbody>
</table>

The DNA was purified using a PCR purification kit (Qiagen), and the size, purity and yield were assessed by electrophoresis on a 1.5% agarose gel alongside a ladder of DNA markers of known size and quantity of DNA. The PCR product was then treated with T4 DNA polymerase to expose the single-stranded overhangs using the following reaction conditions:

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR product</td>
<td>0.2 pmoles (determined by comparison to DNA marker)</td>
</tr>
<tr>
<td>25 mM dGTP</td>
<td>2 µl</td>
</tr>
<tr>
<td>100 mM DTT</td>
<td>1 µl</td>
</tr>
<tr>
<td>10 x T4 poly. buffer</td>
<td>2 µl</td>
</tr>
<tr>
<td>T4 DNA polymerase</td>
<td>1 unit</td>
</tr>
<tr>
<td>Water</td>
<td>to 20 µl</td>
</tr>
</tbody>
</table>

The reaction mix was incubated at room temperature for 30 minutes, followed by 20 minutes at 75°C to heat-inactivate the enzyme. In the annealing reaction, 2 µl of the T4 DNA polymerase-treated insert (approximately 0.02 pmoles) was added to 1 µl (0.01 pmoles) of the LIC vector and the reaction incubated at room temperature for 5 minutes. 1 µl of 25 mM EDTA was then added to the reaction and the incubation continued for a further 5 minutes. 3 µl of this annealing mix was used to transform competent *E. coli* XL1-Blue cells according to the manufacturer’s instructions. Transformants were selected on LB-agar plates containing ampicillin, and overnight cultures were set up from single colonies to obtain plasmid DNA.
2.5.2 Isolation of plasmid DNA

Plasmid DNA was purified from an overnight culture of *E. coli* strain XL1-Blue harbouring the plasmid using QIAprep Spin Miniprep kits (Qiagen) according to the manufacturer’s instructions. This procedure is based on the alkaline lysis method (Birnboim & Doly, 1979). Purified plasmid DNA was resuspended in 10 mM Tris-HCl (pH 8.0).

2.5.3 Restriction digests

Restriction enzyme analysis was used to screen for correct incorporation of the PCR insert into the LIC vector. The two restriction endonucleases used were *Hinc*II and *Bam*HI (Pharmacia). A typical restriction digest contained 600-800 ng of plasmid DNA, 2 units of restriction enzyme and 2 μl of 10x reaction buffer (supplied with enzyme) in a total volume of 20 μl. The reaction was incubated for approximately 2 hours at 37°C and terminated by the addition of 0.5 volumes of STEB before being electrophoresed on a 1.5% agarose gel.

2.5.4 DNA sequencing

DNA sequencing reactions were performed by PNACL (University of Leicester). A cell pellet of *E. coli* strain XL1-Blue harbouring the plasmid to be sequenced, and the appropriate primers (section 2.2.2), were provided for automated DNA purification and sequencing using an ABI 377 DNA sequencer. Sequence data was analysed using SeqEd software (Applied Biosystems).

2.6 EXPRESSION AND PURIFICATION OF 52 kDa PROTEIN

The expression and purification of a 52 kDa N-terminal fragment of human DNA topoisomerase IIα was as previously described (Gardiner *et al.*, 1998). Plasmid pTOPSTOP/His (D. Roper, University of York) was transformed into competent *E. coli* B834(DE3)pLysS cells and plated onto LB-agar plates containing ampicillin for
overnight growth at 37°C. A test expression was performed before scaling-up to large scale expression. A single colony was used to inoculate 10 ml of LB broth containing ampicillin, which was shaken at 37°C until the OD$_{600}$ reached approximately 0.4. The culture was then split in two, and IPTG was added to one to a final concentration of 0.05 mM and growth continued for a further 3 hours at 37°C. The cells were then pelleted by centrifugation and resuspended in 50 mM Tris-HCl (pH 7.5), 10% (w/v) sucrose. A sample of the cells was mixed with an equal volume of Sample Application Buffer, SAB (125 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (w/v) glycerol, 10% (v/v) β-mercaptoethanol, 0.002% (w/v) bromophenol blue) and placed in a boiling waterbath for approximately 15 minutes to lyse the cells. The sample was centrifuged to pellet insoluble material and an aliquot was subjected to SDS-PAGE on a 12% gel to check for protein overexpression. If successful, 1 ml of the original uninduced culture was used to inoculate 200 ml of LB broth containing ampicillin which was grown overnight at 37°C with shaking. For large scale preparation of the protein, eight 2 litre baffled flasks containing 500 ml each of NB broth were inoculated with 10 ml of the overnight culture and grown with shaking at 37°C until the OD$_{600}$ reached approximately 0.4. Protein expression was induced by the addition of IPTG to a final concentration of 0.05 mM and growth continued for a further 3 hours at 37°C. The cells were harvested by centrifugation at 6,000 rpm for 10 minutes in a Beckman JA-10 rotor using a Beckman J2-21 centrifuge. The cell pellet was resuspended in a minimum volume (typically 5 ml per 500 ml culture) of 50 mM Tris-HCl (pH 7.5), 10% (w/v) sucrose, frozen in liquid nitrogen, and stored at -70°C.

After thawing, PMSF was added to a final concentration of 1 mM, and the cells were disrupted by three cycles of sonication (30 seconds on/30 seconds off) at 10 microns amplitude on ice in a Soniprep 150 sonicator (MSE). Most of the 52 kDa protein was found to be in inclusion bodies, and was extracted with the following steps, performed at 4°C. The sample was centrifuged at 40,000 rpm in a TFT 50.38 rotor in a OTD65B Sorvall Ultracentrifuge to pellet the inclusion bodies. The supernatant was discarded and the pellet resuspended in 1% Triton X-100, sonicated and recentrifuged as above. The pellet was then washed with water, resuspended in Binding Buffer (8 M urea, 50 mM Tris-HCl (pH 8.0), 250 mM NaCl), and again sonicated and recentrifuged
as above. The supernatant was then applied to a pre-equilibrated 'TALON' Co²⁺-based immobilised metal affinity column (Clontech) and allowed to bind with gentle agitation for 1 hour. The column was washed several times with one bed volume of Binding Buffer, and subsequently with Binding Buffer plus 5 mM imidazole. The protein of interest was then eluted with several washes of Binding Buffer plus 50 mM imidazole. The eluted 52 kDa protein was then refolded by overnight dialysis at 4°C in refold buffer (50 mM Tris-HCl (pH 8.8), 100 mM KCl, 1 mM EDTA, 1 mM DTT, 10% (w/v) glycerol). Any remaining insoluble material was subsequently removed by centrifugation.

2.7 EXPRESSION AND PURIFICATION OF FUSION PROTEINS

Expression plasmids were transformed into competent E. coli BL21(DE3) cells and test expressions were performed in the same manner as that described in section 2.6. 1 ml of a successful test expression was used to inoculate 200 ml of 2YT broth containing ampicillin which was grown overnight at 37°C with shaking. For large scale preparation of the protein, eight 2 litre baffled flasks containing 500 ml each of 2YT broth were inoculated with 10 ml of the overnight culture and grown with shaking at 25°C until the OD₆₀₀ reached approximately 0.8. Protein expression was induced by the addition of IPTG to a final concentration of 0.1 mM and growth continued for a further 16 hours at 20°C. The cells were harvested by centrifugation and the cell pellet was resuspended in 50 mM Tris-HCl (pH 7.5), 10% (w/v) sucrose, frozen in liquid nitrogen, and stored at -70°C.

After thawing, the following protease inhibitors were added: AEBSF to 0.5 mM, benzamidine to 1 mM, soybean trypsin inhibitor to 5 µg/ml. The cells were disrupted by passage three times through a pre-cooled French pressure cell at 8,000-12,000 psi, and by sonication using two cycles of sonication (30 seconds on/30 seconds off) at 10 microns amplitude. A significant amount of the fusion protein was found in the soluble fraction under these conditions, and was purified by the following steps performed at 4°C. The cell debris was pelleted by centrifugation at 40,000 rpm in an
ultracentrifuge as before. The supernatant was then applied to a pre-equilibrated 'TALON' Co\textsuperscript{2+}-based immobilised metal affinity column (Clontech) and allowed to bind with gentle agitation for 1 hour. The column was washed several times with five bed volumes of high salt Wash Buffer (50 mM Tris-HCl (pH 8.0), 500 mM NaCl), and subsequently with low salt Wash Buffer (50 mM Tris-HCl (pH 8.0), 250 mM NaCl) plus 5 mM imidazole. The fusion protein was then eluted with several washes of low salt Wash Buffer plus 50 mM imidazole. Fractions were analysed by SDS-PAGE electrophoresis, and peak fractions were pooled. The pooled fractions were diluted 5-fold with 50 mM sodium phosphate buffer (pH 7.0), 0.5 mM DTT, 0.5 mM EDTA and then loaded onto a HiTrap SP ion-exchange chromatography column (Pharmacia) using a pump at a flow rate of \(-5\) ml/min at 4°C. The column was then washed with 50 mM phosphate buffer (pH 7.0), 50 mM NaCl, 0.5 mM DTT, 0.5 mM EDTA, and then eluted with a stepwise salt gradient of 75 mM to 1 M NaCl in the same phosphate buffer. Peak fractions, typically greater than 95% pure as determined by SDS-PAGE analysis, were pooled and dialysed overnight at 4°C in storage buffer (50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 50% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA), and stored at -20°C.

\subsection{2.7.1 Proteolytic cleavage of fusion proteins with Factor Xa}

Proteolytic cleavage of fusion proteins to release the human topoisomerase II\(\alpha\) were performed using biotin-labelled Factor Xa (Boehringer Mannheim). Test cleavage experiments were conducted with 10 \(\mu\)g of fusion protein in a 50 \(\mu\)l volume of 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 10% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA using 0 to 2 \(\mu\)g of Factor Xa. Samples were left at 4°C for 16 hours and then analysed by SDS-PAGE electrophoresis. The optimum amount of Factor Xa to use was determined for each batch of fusion protein and Factor Xa, and large scale proteolytic cleavage of fusion proteins was performed under similar conditions. Following proteolytic cleavage, the sample was incubated with an immobilised streptavidin-agarose resin (200 \(\mu\)l of a 50% suspension of the resin per 100 \(\mu\)g Factor Xa) for 1 hour at 4°C to bind the Factor Xa. The sample was then centrifuged to pellet the resin, and the supernatant was removed and diluted 5-fold with 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 10% (v/v) glycerol. This sample is incubated with pre-equilibrated TALON
resin for 1 hour at 4°C, and the resin is pelleted and the supernatant removed (the 'flow-
off'). The resin is then washed with a high salt buffer (50 mM Tris-HCl (pH 8.8), 1 M 
NaCl, 10% (v/v) glycerol) to maximise the recovery of the human topoisomerase IIα 
fragment. The flow-off and high salt washes were then pooled and dialysed overnight 
into storage buffer and stored at -20°C.

2.8 PROTEIN ANALYSIS

2.8.1 Determination of protein concentration

Protein concentration was determined by the method of Bradford (1976). 
Bradford reagent was prepared by dissolving 100 mg of Coomassie Brilliant Blue G 
(Sigma) in 50 ml of 95% ethanol, 100 ml of 85% (w/v) phosphoric acid, made up to 1 
litre with water. The solution was filtered through Whatman filter paper (No. 1) prior to 
use. To determine protein concentration, 100 µl of a suitable dilution of protein sample 
was mixed with 1 ml of Bradford reagent and left for 5 minutes at room temperature. 
The absorbance was measured at 595 nm against a blank solution containing no protein, 
and compared to a calibration curve prepared using bovine serum albumin (BSA).

2.8.2 SDS-PAGE gel electrophoresis of proteins

Proteins were analysed on 10, 12 or 15% discontinuous polyacrylamide gels 
(37.5:1 acrylamide:bis, UltraPure Protogel, National Diagnostics) using the Mini-
Protean II gel system (Bio-Rad). Gels of 0.75 mm thickness were made up as follows:
<table>
<thead>
<tr>
<th></th>
<th>4% stacking gel</th>
<th>10% resolving gel</th>
<th>12% resolving gel</th>
<th>15% resolving gel</th>
</tr>
</thead>
<tbody>
<tr>
<td>30% acrylamide</td>
<td>0.65 ml</td>
<td>3.33 ml</td>
<td>4 ml</td>
<td>5</td>
</tr>
<tr>
<td>Tris-HCl (pH 8.8)</td>
<td>-</td>
<td>3.7 ml</td>
<td>3.7 ml</td>
<td>3.7 ml</td>
</tr>
<tr>
<td>Tris-HCl (pH 6.8)</td>
<td>0.63 ml</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10% SDS</td>
<td>50 µl</td>
<td>100 µl</td>
<td>100 µl</td>
<td>100 µl</td>
</tr>
<tr>
<td>10% APS</td>
<td>50 µl</td>
<td>100 µl</td>
<td>100 µl</td>
<td>100 µl</td>
</tr>
<tr>
<td>TEMED</td>
<td>5 µl</td>
<td>10 µl</td>
<td>10 µl</td>
<td>10 µl</td>
</tr>
<tr>
<td>Water</td>
<td>to 5 ml</td>
<td>to 10 ml</td>
<td>to 10 ml</td>
<td>to 10 ml</td>
</tr>
</tbody>
</table>

Protein samples to be electrophoresed were mixed with an equal volume of Sample Application Buffer, SAB (125 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (w/v) glycerol, 10% (v/v) β-mercaptoethanol, 0.002% (w/v) bromophenol blue) and placed in a boiling waterbath for approximately 5 minutes prior to loading, or 15 minutes to lyse cells. The sample was centrifuged to pellet insoluble material. The gels were run at 200 V until the dye front reached the bottom of the gel (approximately 1 hour). Gels were stained for approximately 15 minutes in Coomassie Brilliant Blue stain (15% (v/v) glacial acetic acid, 10% (v/v) isopropanol, 0.25% (w/v) Coomassie Brilliant Blue) and destained in the same buffer without Coomassie Brilliant Blue until the background staining had disappeared.

### 2.8.3 Western blotting

The protein sample was electrophoresed on an SDS-PAGE gel as described in section 2.8.2. The proteins were then transferred onto a Hybond-C Extra nitrocellulose membrane (Amersham) using a Mini Trans-Blot Electrophoretic Transfer cell (Bio-Rad) and Transfer Buffer (12.5 mM Tris, 200 mM glycine and 10% (v/v) methanol). Electrotransfer was performed at 200 mA for 45 minutes. The nitrocellulose membrane was stained with a non-permanent stain (0.1% Ponceau S (w/v) in 5% acetic acid) to ensure protein transfer had occurred and to allow the position of the protein markers to be marked on the membrane. The membrane was then destained by washing with TBS buffer (25 mM Tris-HCl (pH 7.5), 154 mM NaCl) plus 0.1% Tween-20. The membrane
was then incubated on a shaking platform overnight at 4°C in Block Buffer (TBS buffer containing 0.1% Tween-20 and 5% dried milk) to block excess protein binding sites. The membrane was washed in Wash Buffer (TBS buffer containing 0.1% Tween-20 and 1% dried milk) prior to incubation for 1 hour at room temperature with the primary antibody, a rabbit polyclonal antibody raised against the 52 kDa fragment of human DNA topoisomerase IIα (generous gift of J.R. Jenkins, University of Liverpool), diluted 1:5000 in Wash Buffer. The membrane was washed in Wash Buffer with several changes of buffer prior to incubation with a 1:5000 dilution of the secondary antibody, a mouse anti-rabbit antibody conjugated to horseradish peroxidase (DAKO) in Wash Buffer for 1 hour at room temperature. The membrane was then washed as before. Proteins were detected using an ECL chemiluminescence kit (Amersham) according to the manufacturer's instructions. The membrane was dried between two sheets of 3MM paper (Whatman) and soaked for 1 minute in 2 ml of each of the two luminescence reagents, mixed immediately prior to use. The membrane was dried and wrapped in Saran wrap, and then exposed to Fuji X-ray film for 10 seconds to 1 minute depending on the strength of the chemiluminescent signal.

2.8.4 N-terminal sequencing of proteins

N-terminal sequencing of proteins was carried out by K. Lilley (University of Leicester). Protein samples to be sequenced were electrophoresed on an SDS-PAGE gel as described in section 2.8.2, which was subsequently blotted to a PVDF membrane to transfer proteins. The membrane was stained to identify protein bands, and the band of interest was excised and the protein isolated and subjected to N-terminal sequence analysis by Edman degradation using an ABI 476 protein sequencer.

2.9 ATPase ASSAYS

ATPase assays were performed using an ADP-sensitive linked enzyme assay (Tamura & Gellert, 1990). During this assay (summarised below), ADP produced by the ATPase action of the enzyme is phosphorylated back to ATP by pyruvate kinase.
(PK) using phosphoenolpyruvate (PEP) as the phosphate donor, which produces pyruvate as a by-product. The pyruvate is then reduced to lactate by the action of lactate dehydrogenase (LDH), using NADH as the reducing agent, which is oxidised to NAD⁺. Since NADH absorbs strongly at 340 nm whereas NAD⁺ does not, ATP hydrolysis is coupled to a decrease in absorbance at 340 nm, which can be followed continuously in a spectrophotometer.

This assay therefore has two advantages: firstly, it prevents the build-up of ADP, which inhibits the ATPase activity, and secondly, it maintains a constant concentration of ATP within the reaction. The change in absorbance is related to ADP production using \( A_{\text{340}} = 6.22 \text{ cm}^{-1} \).

ATPase reactions were typically performed in a volume of 150 μl at 37°C in the following buffer (unless otherwise stated): 50 mM Tris-HCl (pH 7.5), 10 mM NaCl, 5% glycerol (v/v), 0.4 mM PEP, 0.2 mM NADH, 5 μl PK/LDH (in 50% (w/v) glycerol, 100 mM KCl, 10 mM HEPES (pH 7.0), 0.1 mM EDTA), 2 mM Mg-ATP (in sodium acetate buffer at pH 7.5), 0.5 mM DTT, 0.05 mM EDTA. DNA, where present, was added at the indicated concentration or bp:dimer ratio. Reactions were initiated by the addition of enzyme and monitored continually for 1 hour in an absorbance plate reader (either a Bio-Tek EL340 Bio-Kinetics reader or a Molecular Devices SpectraMax Plus Microplate Spectrophotometer). The initial rate of ATP hydrolysis was calculated following a 5 minute equilibration period using a minimum of 10 minutes of data collection. Error bars, where shown, indicate ± 1 standard deviation from the mean, where \( n = 3 \).
2.10 PROTEIN CROSSLINKING

Crosslinking of protein was performed using three crosslinking reagents (Pierce): dimethyl suberimidate (DMS), disuccinimidyl glutarate (DSG) and bis(sulfosuccinimidyl) suberate (BS³). Stock solutions of crosslinkers were made up as follows: DMS, 20 mg/ml in water; DSG, 5 mg/ml in dimethyl sulfoxide (DMSO); BS³, 10 mg/ml in water. Protein (7.5 μM final concentration) was incubated in the presence or absence or nucleotide (2 mM) or relaxed pBR322 at 50 bp/dimer for 1 hour at 25°C in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 10% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA. Crosslinker was then added at a 10-fold dilution of the stock solution and the samples incubated at 25°C for a further 16 hours. An aliquot of the samples was then mixed with an equal volume of SAB, placed in a boiling waterbath for 5 minutes, and then analysed by SDS-PAGE electrophoresis on a 10% gel.

2.11 DNA BINDING ASSAYS

2.11.1 Filter binding

Protein-DNA interactions were assessed using a filter binding technique. Samples (50 μl) containing up to 1 μM protein in 50 mM Tris-HCl (pH 8.8), 20 mM NaCl, 5 mM MgCl₂, 10% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA, 0.1 mg/ml BSA were incubated at 25°C in the presence or absence of 3.62 nM relaxed ³H-pBR322 (generous gift of A.J. Howells) or 2 mM ADPNP. After 30 minutes, either DNA or ADPNP was added as above, or no addition was made, and the samples were incubated for a further 30 minutes at 25°C. A 25 mm filter disc (0.45 μm NC 45 membrane filter, Schleicher & Schuell) was washed with 2 x 500 μl of Binding Buffer (50 mM Tris-HCl (pH 8.8), 20 mM NaCl, 5 mM MgCl₂, 10% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA). The sample was added to 200 μl of Binding Buffer, and this was filtered through the filter disc at a flow rate of approximately 1 ml/minute. The filter disc was washed with 2 x 500 μl of Binding Buffer, and was then left to air dry for approximately 30 minutes. 4 ml of scintillation liquid was added and the amount of tritiated DNA retained on the filter disc was calculated by scintillation counting.
2.11.2 Gel retardation

Gel retardation experiments were carried out with a 140 bp DNA fragment based around the preferred DNA gyrase cleavage site in pBR322. The DNA fragment (50 nM) was incubated with protein in 50 mM Tris-HCl (pH 8.0), 20 mM NaCl, 5 mM MgCl₂, 10% (v/v) glycerol, 5 mM DTT, 0.1 mM EDTA at 25°C for 30 minutes. ADPNP (or water) was then added to the samples to a final concentration of 2 mM and the samples were further incubated for 30 minutes at 25°C. Samples were loaded onto a 5% TBM gel (pre-run for >30 minutes) and run at 100 V for 1 hour. DNA was visualised by staining with ethidium bromide.

2.12 LIMITED PROTEOLYSIS

Limited proteolysis was carried out using trypsin (Sigma). Samples contained protein (4 μM, 0.2 mg/ml) in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 5 mM MgCl₂, 10% (v/v) glycerol, 5 mM DTT, 0.5 mM EDTA. Samples were incubated for 1 hour at 25°C in the presence or absence of nucleotide (2 mM) or relaxed pBR322 at 50 or 250 bp/dimer. Trypsin was then added to the samples at 15 μg/ml final concentration and the reactions incubated at 25°C for a further 2 hours. An aliquot was removed after 0, 1, 5, 15, 30, 60 and 120 minutes and mixed with an equal volume of SAB, boiled for 5 minutes, and analysed by SDS-PAGE electrophoresis on a 15% gel.

2.13 MOLECULAR WEIGHT STUDIES

2.13.1 Gel filtration

A Sephacryl S-200 HR (16/60) FPLC column (Pharmacia) was equilibrated overnight at a flow rate of 0.2 ml/minute in gel filtration buffer (50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 5 mM MgCl₂, 10% glycerol (w/v), 0.5 mM DTT, 0.5 mM EDTA) in the presence or absence of 0.1 mM ADPNP. The column was calibrated using a set of molecular weight markers and also Blue Dextran to determine the void volume of the
column. Protein (0.4 mg/ml in gel filtration buffer) was incubated for 1 hour at 25°C in the presence or absence of 2 mM ADPNP. A 500 µl aliquot of the sample (~200 µg) was loaded onto the column and eluted with gel filtration buffer at a flow rate of 0.8 ml/minute in the presence or absence of 0.1 mM ADPNP. The absorbance was monitored at 280 nm, and 1 ml fractions were collected and analysed by SDS-PAGE.

2.13.2 Analytical ultracentrifugation

Analytical ultracentrifugation was carried out by A. Leech (University of East Anglia). Samples were prepared for analytical ultracentrifugation by gel filtration in an identical manner to that described in section 2.13.1 except that the elution buffer used was 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 5% glycerol (w/v), 5 mM MgCl₂, 0.5 mM EDTA ± 0.1 mM ADPNP. Fractions corresponding to the most prominent peak in each sample were pooled and concentrated to ~0.1 mg/ml and stored on ice until being used in analytical ultracentrifugation experiments. Analytical ultracentrifugation was carried out in a Beckman Optima XL-1 Analytical Ultracentrifuge using an 8 hole rotor and a double-sector cell with an optical length of 12 mm. Equilibrium sedimentation runs were performed initially at 15,000 rpm for 28 hours, after which time the speed was increased to 20,000 rpm. The concentration distribution for the protein within the cell was determined using an absorbance optical system.
CHAPTER 3

Expression and purification of the N-terminal domain of DNA topoisomerase IIα as a fusion protein
3.1 INTRODUCTION

The sequence homology between both prokaryotic and eukaryotic type II topoisomerases is highest in the N-terminal and DNA cleavage regions of these enzymes (Caron & Wang, 1994). An N-terminal fragment of the GyrB protein of *E. coli*, corresponding to residues 2-393 (the N-terminal methionine is presumably lost by posttranslational modification), has been cloned and overexpressed and shown to be an ATPase (Ali et al., 1993). The crystal structure of this fragment has been solved (Wigley et al., 1991), and based on this and site-directed mutagenesis studies (Tingey & Maxwell, 1996), this fragment has been proposed to operate as an ATP-operated clamp which captures a DNA segment during the catalytic cycle.

By analogy, the N-terminal region of human DNA topoisomerase IIα was thought to be an ATPase. A variety of constructs encoding this N-terminal region, ranging from residues 1-400 to 1-440, were constructed and used in expression studies in *E. coli* BL21(DE3) cells (Gardiner et al., 1998). The C-terminus of these clones was based on the putative domain boundary defined by homology with the *E. coli* GyrB protein. The C-terminal residue of the N-terminal domain of GyrB is Arg393, which corresponds to Lys425 in human DNA topoisomerase IIα (Caron & Wang, 1994; see Figure 3.11). The best expression was found with a clone (pTOPSTOP/His) expressing residues 1-440; a 52 kDa protein that also included an 11 amino acid polylinker and a C-terminal hexa-histidine tag.

Large scale overexpression of this 52 kDa protein proved to be problematic (L.P. Gardiner, personal communication). A large inconsistency in expression levels was found between protein preparations, which initially could not be attributed to any one single factor. A variety of measures were taken to maximise the level of protein

---

1 The sequence shown for human DNA topoisomerase IIα is that of Tsai-Pflugfelder et al. (1988). However, according to the results of Hinds et al. (1991), and M. Tsai-Pflugfelder & J.C. Wang (unpublished results, cited in Wasserman et al., 1993), the original sequence at residues 109-113 (MIRKQ) has been replaced with the sequence IDPENN. This discrepancy is due to a sequencing error in this region that resulted in the omission of three nucleotides from the original sequence. The clone pTOPSTOP/His actually encodes the residues MIRKQ rather than IDPENN, and therefore expresses only a 439 residue protein representing the N-terminal domain. However, to avoid confusion I have used the numbering system used in Figure 3.1 throughout this thesis (although quoted sizes of plasmids and molecular weights for proteins are correct for the individual plasmid/protein).
| Human TopoIIa | MEVSPHQFYNR-ENMQVNIKKETAKLRLSVEQH | 34 |
| Human TopoIIb | MAKSGCCGAGAGCAGNNCTWNADKKESETNKKSKKSLVEQQ | 50 |
| Yeast Top2 | KTTPEWVIAHDQ | 12 |
| E. coli GyrB | MSNYDSSSI | 10 |
| Human TopoIIa | OKKTCLERHILRPITYETGGVEQKTVFHE | 81 |
| Human TopoIIb | KKTCLERHILRPITYETGGVEQKTVFHE | 97 |
| Yeast Top2 | KTSQCLERHILRPITYETGGVEQKTVFHE | 60 |
| E. coli GyrB | KVLKQDVARFPEGMCGTDD | 36 |
| Human TopoIIa | YKVFYTVPLTTGGLSSTSYDDEKETKTVGEE | 130 |
| Human TopoIIb | YKVFYTVPLTTGGLSSTSYDDEKETKTVGEE | 146 |
| Yeast Top2 | NKKNTFETRMHGLLSSNYDDEKETKTVGEE | 110 |
| E. coli GyrB | PMVSEAVEVRVTVHGKEQDATEGH | 132 |
| Human TopoIIa | ASREYKLFKQTFDKGRGAKRKN-SEDTCTFLPDLSKFM | 34 |
| Human TopoIIb | ASREYKLFKQTFDKGRGAKRKN-SEDTCTFLPDLSKFM | 244 |
| Yeast Top2 | TADDLVGKKTHKMNNSICHPLTDYKKPS | 209 |
| E. coli GyrB | VQREBGKEAMRENQHGPO | 177 |
| Human TopoIIa | QGLQDIDALHRAY | 272 |
| Human TopoIIb | QGLQDIDALHRAY | 268 |
| Yeast Top2 | QGLQDIDALHRAY | 253 |
| E. coli GyrB | NVTEKPIAKRPSLESNSGVSGLYVRKGDHFRYEGK | 227 |
| Human TopoIIa | MLKDKKDLT | 311 |
| Human TopoIIb | MLKDKKDLT | 327 |
| Yeast Top2 | MLKDKKDLT | 303 |
| E. coli GyrB | YLNNKTKFPI-HP | 266 |
| Human TopoIIa | FVSNISATSKGCGGGVDDYQDIIYKLVLKKV-KGCGVAVK-AHVK | 357 |
| Human TopoIIb | FVSNISATSKGCGGGVDDYQDIIYKLVLKKV-KGCGVAVK-AHVK | 237 |
| Yeast Top2 | FVSNISATSKGCGGGVDDYQDIIYKLVLKKV-KGCGVAVK-AHVK | 347 |
| E. coli GyrB | YGFRKTPQDQDGLAGFRAAMRSLNAYELDGKSMKVSATGDDAR | 316 |
| Human TopoIIa | NHMWIFVKNALNPNDPSSTKEMTMLOPKPSGDFSLEKFI | 359 |
| Human TopoIIb | NHMWIFVKNALNPNDPSSTKEMTMLOPKPSGDFSLEKFI | 415 |
| Yeast Top2 | NHMWIFVKNALNPNDPSSTKEMTMLOPKPSGDFSLEKFI | 389 |
| E. coli GyrB | EGLIAVSYVKDLPKSSTOKLVSSESVKAVEQOMVNAELAYLENPT | 366 |
| Human TopoIIa | NAAACGGIVESLNM-VKFKACVGLNK-SAVKMRIRKGGPKLDDANDA | 447 |
| Human TopoIIb | NAAACGGIVESLNM-VKFKACVGLNK-SAVKMRIRKGGPKLDDANDA | 463 |
| Yeast Top2 | NAAACGGIVESLNM-VKFKACVGLNK-SAVKMRIRKGGPKLDDANDA | 436 |
| E. coli GyrB | DAKTVYGGKLDDAREAAREAREMT-RKGG-ALDLAGLPLTAC | 411 |
| Human TopoIIa | GQGRPTECTLITTEGEDSALKLAULSLGVVGLKUYVPPGRKILNVEES | 497 |
| Human TopoIIb | GQGRPTECTLITTEGEDSALKLAULSLGVVGLKUYVPPGRKILNVEES | 513 |
| Yeast Top2 | GQGRPTECTLITTEGEDSALKLAULSLGVVGLKUYVPPGRKILNVEES | 456 |
| E. coli GyrB | EPDTLYSGMVGEDSAG-GSASQGHRKNOATPGKILNVEES | 486 |

**Figure 3.1 Sequence alignment of the N-terminus of type II DNA topoisomerases.**

Alignment of the N-terminal region of human DNA topoisomerase IIα (IIα) with that of human DNA topoisomerase IIβ (IIβ), yeast topoisomerase II and *E. coli* GyrB, according to the sequence of Caron & Wang (1994). Identical residues are indicated by black shading; similar residues are indicated by grey shading. Residue numbers are shown on the right.
expression, which resulted in switching the expression strain to *E. coli* B834(DE3) cells grown in Nutrient Broth. This strain of *E. coli* (the parent strain of BL21) has been reported to be a superior host strain for the expression of certain recombinant proteins under the control of T7 promoters (Doherty *et al.*, 1995). Strain B834, a methionine auxotroph, was shown to give increased expression levels of some toxic gene products as a result of an improved growth rate following induction and a higher level of production of soluble protein.

The 52 kDa protein was subsequently found to be expressed almost entirely in the insoluble fraction of the cell lysate in the form of inclusion bodies. Inclusion body formation is believed to be caused by the self-association of partially folded or misfolded proteins, and is a consequence of high-level expression of a heterologous recombinant eukaryotic protein in a prokaryotic cell (Schein, 1989). Attempts to increase the recovery of soluble protein were deemed unsuccessful, and it was decided to develop a strategy for refolding the protein following its isolation from inclusion bodies. Expression of the protein of interest as inclusion bodies does offer several advantages over the normal soluble expression of proteins (Rudolph & Lilie, 1996). The protein is present in a highly enriched form, and the relatively high density of inclusion proteins allows them to be efficiently separated from cell debris in a relatively homogenous state. Inclusion body formation also protects the protein from proteolytic degradation by intracellular proteases. However, the formation of inclusion bodies results in inactive aggregates of protein which require solubilising and subsequent refolding *in vitro*, a procedure which must be determined empirically.

In the case of the 52 kDa protein, the inclusion bodies are first isolated by repeated rounds of sonication of the cell suspension in the presence of the detergent Triton X-100 to further disperse the cell membranes. The resultant pellet following centrifugation is resuspended and sonicated in Tris-buffered 8 M urea to solubilise the protein. The His-tagged protein can then be purified on a cobalt-based immobilised metal affinity column (TALON\textsuperscript{TM}, Clontech). Removal of the denaturant to allow the protein to refold is accomplished by overnight dialysis into a refold buffer (50 mM Tris-HCl (pH 8.8), 100 mM KCl, 10 % (w/v) glycerol, 1 mM EDTA, 1 mM DTT).
Figure 3.2 Expression and purification of the 52 kDa protein.
Analysis of the expression and purification of the 52 kDa protein according to the method of Gardiner et al. (1998). Expression of the protein in B834(DE3)pLysS cells transformed with pTOPSTOP/His at 37°C was induced by the addition of IPTG to a final concentration of 0.05 mM. The 52 kDa protein was purified from the fractionated cell lysate on a metal affinity column (TALON), and dialysed overnight in refold buffer (50 mM Tris-HCl (pH 8.8), 100 mM KCl, 10% glycerol (w/v), 1 mM DTT, 1 mM EDTA). Fractions were analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lanes 2 & 3, uninduced (Un) and induced (In) total cell protein extract respectively; lane 4, protein eluted from the TALON column; lane 5, eluted protein following overnight dialysis.
3.2 shows a summary of the expression and purification of the 52 kDa protein, which has been prepared according to the method of Gardiner et al. (1998). Following induction with IPTG, the 52 kDa protein accumulates in the cell as inclusion bodies (Figure 3.2, Lane 3). When the inclusion bodies are isolated and the 52 kDa protein solubilised and purified, the resultant protein is >80% pure (Figure 3.2, Lane 4). Following overnight dialysis into refold buffer, a large amount of the protein precipitates and can be removed by centrifugation, then resuspended in Tris-buffered 8 M urea and redialysed to maximise recovery. The concentration of soluble 52 kDa protein recovered in the supernatant (Figure 3.2, Lane 5) is typically less than 0.15 mg/ml, with a total yield of less than 0.5 mg per litre of culture; any attempt to further concentrate this protein results in precipitation. This protein was found to be active in ATPase assays. Previously, it had been reported that there was considerable variation in activity between different batches of purified protein (Gardiner et al., 1998). The activity was found to differ by as much as 5-fold for different preparations of the 52 kDa protein for both the intrinsic and DNA-stimulated ATPase rate. I have also found there to be significant differences in activity between different batches of 52 kDa protein that I have prepared. It is possible that this is due to variations in the refolding step; solubility does not necessarily imply that a protein is correctly folded, and soluble proteins may still form inactive aggregates.

It is apparent from the above that the expression and purification of the 52 kDa protein represents a sub-optimal system for characterising the N-terminal ATPase domain of human DNA topoisomerase IIα. The position of the C-terminal residue derived from DNA topoisomerase IIα is only loosely based on sequence alignment with the E. coli GyrB protein, not limited proteolysis of the whole protein as would be more informative. It contains 17 extra amino acids at its C-terminus which are foreign to this protein, including the hexa-histidine tag. The protein is expressed as insoluble inclusion bodies in a foreign host, indicative of an incorrectly folded protein. This may be due to the protein sequence ending in a secondary structure element that is important for solubility or structural integrity. The recovery of the protein requires it to be denatured with strong chaotropic agents and subsequently refolded in vitro, where misfolding and aggregation compete with the correct folding pathway (Rudolph & Lilie, 1996). The
final yield of protein is only a small fraction of that which is expressed in the cell, and its low solubility and variable activity suggest that the soluble fraction is contaminated with inactive aggregates.

There are many strategies that can be employed to optimise the level of protein production in a host such as *E. coli*, most of which must be determined on a case-by-case basis (Schein, 1989; Hannig & Makrides, 1998). These include: growth of cultures at lower temperatures to increase solubility, optimisation of codon usage (relevant here as a human protein is being expressed in an *E. coli* host), and the use of fusion proteins. There are many advantages that can be gained by fusing the gene of interest to a second gene that is known to be well expressed in the host cell. This chapter describes the cloning, expression, purification and initial characterisation of the N-terminal domain of human DNA topoisomerase IIα as a cleavable thioredoxin fusion protein.

3.2 RESULTS

A novel Ligation-Independent Cloning (LIC-cloning) strategy has been used to produce fusion constructs of the gene encoding the putative N-terminal domain of human DNA topoisomerase IIα. LIC-cloning allows the directional cloning of PCR products with high efficiency and reduced non-recombinant background levels, but without the need for restriction digestion or ligation reactions (Aslanidis & de Jong, 1990). This procedure utilises single-stranded complementary overhangs engineered into PCR product and vector to allow them to non-covalently anneal to one another *in vitro*. In order to generate the single-stranded overhangs in the PCR-amplified gene, the primers are designed with 5' extensions of 12 to 17 bases. The PCR product is treated with the (3' → 5') exonuclease activity of T4 DNA polymerase in the presence of one dNTP, which degrades back the DNA from the 3' ends until it encounters the chosen dNTP (Figure 3.3A). The linearised vector is also treated with T4 DNA polymerase to reveal single-stranded overhangs complementary to those on the insert, but that are not self-complementary. The vector and insert are then annealed to each other in a rapid 10 minute reaction, not the typical overnight ligation reaction. These recombinants are
Figure 3.3
3.3A Overview of the Ligation Independent Cloning (LIC) strategy for Xa/LIC vectors.
The gene of interest is PCR-amplified using primers with 5' extensions to introduce complementary overhangs to the LIC vector after treatment with T4 DNA polymerase. The insert and vector are then annealed and transformed into competent \textit{E. coli} cells where the backbone is ligated \textit{in vivo}.

3.3B Schematic representation of the pET-32 Ek/LIC and pET-32 Xa/LIC vector products.
The LIC-vectors encode the thioredoxin protein of \textit{E. coli} (Trx-tag), a hexa-histidine tag (His-tag), and the 15 amino acid S-tag. The gene of interest is inserted at the LIC-site, which is followed by a multiple cloning site (MCS) and a second His-tag. The thrombin and Ek/Factor Xa cleavage sites are shown.
then transformed into \textit{E. coli} cells where the insert-vector junction is ligated by the host cell DNA ligase \textit{in vivo}. Due to the poor transformation efficiency of linear plasmid DNA, the background level of non-recombinant vector-only transformants is significantly reduced.

The LIC-vectors used carry the T7 \textit{lac} promoter, T7 transcription terminator and \textit{lacI} gene for T7 RNA polymerase-driven protein expression, and the \textit{bla} gene for ampicillin resistance. The correct incorporation of the insert into the LIC-vector causes the expression of a fusion protein consisting of, from N-terminus to C-terminus, the 109 amino acid thioredoxin protein (Trx) from \textit{E. coli}, a hexa-histidine tag and a 15 amino acid S-tag (which interacts with the 104 amino acid S-protein portion of ribonuclease A), and finally the protein of interest (Figure 3.3B). The fusion protein can also contain a second hexa-histidine tag C-terminal to the protein of interest (depending upon whether the antisense primer contains a stop codon or allows readthrough to the vector-encoded stop codon). The fusion of the protein of interest to thioredoxin has been found to increase the solubility of some proteins when expressed in the \textit{E. coli} cytoplasm (LaVallie \textit{et al.}, 1993). The 11,675 Da thioredoxin protein of \textit{E. coli} can accumulate to \(~40\%\) of the total cellular protein when overexpressed, and still remain in the soluble fraction. The crystal structure of \textit{E. coli} thioredoxin (Katti \textit{et al.}, 1990) reveals that the C-terminus is accessible on the surface of the molecule, which is ideal for fusion to another protein. Also, many of the proteins retain some or all of their biological activity whilst fused to thioredoxin (LaVallie \textit{et al.}, 1993). A further advantage of using an N-terminal fusion protein is that it helps to avoid the problem of translational stalling, which is particularly important if the gene of interest contains multiple rare codons near the beginning of the coding sequence (Chen & Inouye, 1990). The hexa-histidine and S-tags can each be used for both detection and purification of the fusion protein using commercially available antibodies and affinity resins.

The fusion protein also contains within it the recognition site for two proteases, which can be used to proteolytically cleave it into the topoisomerase II\(\alpha\) subfragment and the thioredoxin tag subfragment. Thrombin can be used to cleave the fusion protein between the His-tag and the S-tag, and thus leaves the S-tag attached to the N-terminus.
of the protein of interest. The second cleavage site occurs at the LIC site itself, and is
encoded by a region contained in the sense primer. Two proteases that are commonly
used to cleave fusion proteins are Enterokinase (Ek) and Factor Xa, because they cleave
on the C-terminal side of their recognition site. The sense primer contains the first
codon of the protein of interest immediately after the codon encoding the final amino
acid of the protease recognition site, and therefore following cleavage of the fusion
protein the released protein of interest contains no extra amino acids at its N-terminus.
LIC-vectors are available that use either enterokinase or Factor Xa to cleave the fusion
protein and so release the protein of interest in its native state (pET-32 Ek/LIC and pET-
32 Xa/LIC respectively).

3.2.1 Fusion proteins with an Enterokinase cleavage site

3.2.1.1 Cloning and expression of Ek/LIC fusion proteins

The pET-32 Ek/LIC vector (Novagen) was used to produce fusion constructs of
the N-terminal domain of human DNA topoisomerase IIα. It is a 5917 bp linearised
vector with 12 to 14 nucleotide 5' overhangs that are complementary to those
incorporated into the insert. A sense PCR primer was designed that incorporated a 14
bp extension to generate a vector-compatible overhang that would anneal to the
complementary sequence on the vector, and a 20 bp region of the target gene (primer
SC01, 5' - GAC GAC GAC AAG ATG GAA GTG TCA CCA TTG GAC C - 3'; the 14
bp extension is underlined). The first nucleotide of the insert-specific sequence is a G,
which completes the ATG initiation codon and encodes the first methionine of human
DNA topoisomerase IIα. Antisense primers were then designed, with 15 bp 5'
extensions to generate a vector-compatible overhang, and translational stop codons that
would lead to the production of a variety of fusion constructs encoding the expression of
residues 1-420 (primer SC04, 5' - GAG GAG AAG CCC GGT TCA TTG GGC CTT
AAA CTT CAC C - 3'), 1-422 (primer SC02, 5' - GAG GAG AAG CCC GGT TCA
CTG GAC TTG GGC CTT AAA C - 3'), 1-425 (primer SC05, 5' - GAG GAG AAG
CCC GGT TCA CTT GTT TAA CTG GAC TTG C - 3') and 1-426 (primer SC03, 5' -
GAG GAG AAG CCC GGT TCA CTT CTT GTT TAA CTG GAC - 3') of human
DNA topoisomerase IIα. The exact location of the domain boundary between the N-terminal domain and the central DNA-cleavage domain is unknown, and so the final residue was based upon sequence homology between the N-terminal domain of human DNA topoisomerase IIα, *E. coli* GyrB and yeast DNA topoisomerase II (Caron & Wang, 1994; see Figure 3.1). The final residue of the N-terminal domain of *E. coli* GyrB is Arg393 (Adachi *et al*., 1987), which corresponds to Lys425 in human DNA topoisomerase IIα; and the final residue of the N-terminal domain of yeast DNA topoisomerase II is Glu410 (Lindsley & Wang, 1991), which corresponds to Gln420 in human DNA topoisomerase IIα.

These oligonucleotides were used in PCR reactions (Saiki *et al*., 1988) using pTOPSTOP/His as template. The sizes of the PCR fragments were verified by agarose gel electrophoresis (data not shown), and once purified were cloned into the Ek/LIC vector according to the manufacturer’s instructions (Novagen), and transformed into commercially available competent *E. coli* XL1-Blue cells (Stratagene). Plasmid DNA from overnight cultures of recombinants was isolated by alkaline-lysis miniprep (Qiagen Miniprep kit) and the correct incorporation of the insert was confirmed by restriction enzyme digest (data not shown). Restriction enzymes used were *HincII* (which cuts twice in the insert and twice in the pET-32 Ek/LIC vector sequence; see Figure 3.4) and *BamHI* (which cuts once in the insert and once in the pET-32 Ek/LIC vector sequence). These four constructs were designated pTOP1-420Ek, pTOP1-422Ek, pTOP1-425Ek and pTOP1-426Ek (Figure 3.4):

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Total bp</th>
<th>TopoIIα residues Expressed</th>
<th>MW of TopoIIα fragment (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTOP1-420Ek</td>
<td>7174</td>
<td>1-420</td>
<td>47,747</td>
</tr>
<tr>
<td>pTOP1-422Ek</td>
<td>7180</td>
<td>1-422</td>
<td>47,974</td>
</tr>
<tr>
<td>pTOP1-425Ek</td>
<td>7189</td>
<td>1-425</td>
<td>48,329</td>
</tr>
<tr>
<td>pTOP1-426Ek</td>
<td>7192</td>
<td>1-426</td>
<td>48,457</td>
</tr>
</tbody>
</table>

Initial expression studies for these plasmids were carried out in *E. coli* strains BL21(DE3) and B834(DE3), both of which are lysogens of bacteriophage DE3 (a
Figure 3.4 Plasmid maps of Ek/LIC vectors.
Plasmid maps for pTOP1-420Ek, pTOP1-422Ek, pTOP1-425Ek and pTOP1-426Ek are shown. The positions of the thioredoxin (Trx) and human DNA topoisomerase IIα gene (TopoII) are shown, as well as the bla gene for ampicillin resistance. Also indicated are the restriction sites for HincII.
Figure 3.4 (cont) Plasmid maps of Ek/LIC vectors.
lambda derivative), and contain a chromosomal copy of the gene for T7 RNA polymerase under the control of the lacUV5 promoter, which is inducible by the addition of IPTG (Studier & Moffatt, 1986). Expression levels were consistently higher in strain BL21(DE3) than in B834(DE3), and the growth rate of strain BL21(DE3) was significantly higher than that of B834(DE3) both pre- and post-induction (data not shown). Therefore, strain BL21(DE3) was used as the general expression host. Expression of the four fusion proteins in BL21(DE3) cells grown at 37°C is shown in Figure 3.5A. All four fusion proteins are overexpressed to approximately the same level in both the induced and uninduced samples (predicted molecular weights are: 1-420 fusion (F) protein (1-420F(Ek)), 64,950 Da; 1-422F(Ek), 65,177 Da; 1-425F(Ek), 65,532 Da; 1-426F(Ek), 65,660 Da). Expression of the fusion protein without induction is due to 'leaky' expression of T7 RNA polymerase from the lacUV5 promoter. There appears to be no deleterious effect on cell growth from this leaky expression, which suggests that the fusion proteins are in no way toxic to the host cells. The high level of protein expression with and without induction would also suggest that the plasmids are stable in the host cells, since unstable plasmids would not be able to establish themselves in the host cells, and thus expression levels would be low (Studier & Moffatt, 1986). However, in some cases it was observed that there was no visible overexpression of the fusion protein in the uninduced sample, although there was a high level of overexpression in the corresponding induced sample. This phenomenon did not appear to correlate with the presence of any one particular expression plasmid, and may reflect slight differences in the physiological state of the expression host cells. Trial expressions of the fusion proteins were also carried out in E. coli BL21(DE3)pLysS cells, which contain a plasmid (pLysS) expressing a small amount of T7 lysozyme, which is a natural inhibitor of T7 RNA polymerase (data not shown). As expected, this reduced the level of uninduced protein expression to an undetectable level, but also significantly reduced the level of induced protein expression.

To determine to what extent the fusion proteins are expressed as soluble protein, the induced fractions shown in Figure 3.5A were separated into soluble and insoluble protein by first sonicating the cell suspension, which is then centrifuged and the supernatant (containing the soluble protein) is removed. The pellet is resuspended
Figure 3.5

3.5A Expression of Ek/LIC fusion proteins at 37°C.
Cultures (10 ml LB) of pTOP1-420Ek, pTOP1-422Ek, pTOP1-425Ek and pTOP1-426Ek in BL21(DE3) cells were grown at 37°C until an OD$_{600}$ of ~0.4 was reached. Cultures were then split into two, one was induced with 0.1 mM IPTG and growth was continued for 3 hours. Total cell protein was then analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lanes 2-9, uninduced (Un) and induced (In) protein from cells expressing the indicated fusion protein.

3.5B Soluble and insoluble protein from expression of fusion proteins.
The induced fractions from Figure 3.5A were separated into soluble and insoluble protein by sonication and analysed by 12% SDS-PAGE. Lane 1, molecular weight markers; lanes 2-9, soluble (S) and insoluble (P) protein, resuspended in urea, from cells expressing the indicated fusion protein.
in Tris-buffered 8 M urea, and is centrifuged to pellet cell debris. This supernatant contains the insoluble protein, which has now been solubilised in 8 M urea. As can be seen from Figure 3.5B, the vast majority of the expressed fusion protein is present in the insoluble protein fraction, presumably in the form of inclusion bodies. This appears to be the case for all four of the expressed fusion proteins, and would suggest that the fusion protein is not folding correctly and is aggregating \textit{in vivo}.

Based on the above expression studies, it was decided that initial work would focus on the fusion protein l-420F(Ek), which would be prepared from the soluble fraction following cell disruption. Despite the relatively low yield of soluble protein compared to insoluble protein, it was hoped that the soluble protein would represent correctly folded and more consistently active protein than was obtained with the 52 kDa protein. There appeared to be little difference in expression levels and soluble protein production between clones, but the pTOP1-420Ek plasmid appeared to transform with the greatest efficiency and consistency.

3.2.1.2 Purification of fusion proteins

The fusion proteins contain three protein or peptide regions that can be used for affinity purification. A hexa-histidine tag within the linker region between the N-terminal thioredoxin protein and the C-terminal topoisomerase IIα fragment can be used to purify the fusion protein by immobilised metal affinity chromatography, as is the case with the 52 kDa protein. Also within the linker region is the S-tag, a 15 amino acid peptide which interacts with the 104 amino acid S-protein portion of ribonuclease A, which is available covalently immobilised on agarose beads. The thioredoxin protein itself can also be used to purify the fusion protein. The active site of thioredoxin contains two vicinal thiol groups (cysteine residues), which can reversibly bind metal ions such as cadmium or arsenic. A covalently modified agarose-based support (Thiobond™ resin, Invitrogen) containing phenylarsine oxide can bind the vicinal dithiols in thioredoxin; the resin can then be washed to remove contaminants, and the fusion protein can then be eluted by the addition of a reducing agent such as β-mercaptoethanol. Given the high degree of purity obtained for the 52 kDa protein using
the cobalt-based TALON metal affinity resin (Clontech), this strategy was adopted for the initial purification of the thioredoxin fusion proteins.

3.2.1.2.1 Immobilised metal affinity chromatography

The affinity purification of His$_6$-tagged proteins is based on the interaction between electropositive metals, e.g., Ni$^{2+}$ and Co$^{2+}$, which have six available coordination sites arranged in an octahedral configuration, and an electron-rich ligand such as histidine. Histidine binds by sharing the electron density of the imidazole nitrogen with the electron-deficient orbitals of the transition metal, and can be eluted using a pH gradient, EDTA, histidine or imidazole. The advantage of the TALON Co$^{2+}$-based system over some commercially available Ni$^{2+}$ systems is that the metal ion is held firmly in place by a tetradeinate chelator. This chelator coordinates the octahedral Co$^{2+}$ ion at four positions, and orients the remaining two sites away from the solid support such that they are accessible to the histidine residues. This overcomes the problem of metal ion leakage during purification that has been encountered with some affinity systems; this was encountered with the 52 kDa protein, where the nickel contamination was found to inhibit the ATPase activity of the 52 kDa protein (L.P. Gardiner, personal communication). However, the fact that only two coordination sites are available to bind histidine residues means that His$_6$ proteins will bind less tightly to TALON resin than to some nickel-based resins, but then so will non-His$_6$ contaminant proteins.

Four litres of LB broth (8 x 500 ml in 2 litre baffled flasks) were inoculated with an overnight culture of BL21(DE3)[pTOP1-420Ek] cells from a fresh transformation plate and grown at 37°C until an OD$_{600}$ of ~0.4 was reached. Protein expression was then induced by the addition of IPTG to a final concentration of 0.1 mM and growth was continued for a further 3 hours. The cells were then harvested and resuspended in 50 mM Tris-HCl (pH 7.5), 10% sucrose. After disruption of the cell wall by passage through a French pressure cell, the clarified cell lysate was applied to 4 ml of prewashed TALON resin, and protein was bound by gentle shaking at 4°C for 1 hour. The resin was then washed with several column volumes of high salt wash buffer.
(50 mM Tris-HCl (pH 8.0), 500 mM NaCl) and poured into a 10 ml disposable column and the flow-through was collected. The column was then washed with five column volumes of low salt wash buffer (50 mM Tris-HCl (pH 8.0), 250 mM NaCl) plus 5 mM imidazole, and then eluted with five column volumes of low salt wash buffer plus 50 mM imidazole.

Figure 3.6 shows the purification of the I-420F(Ek) fusion protein in this manner. The protein eluted from the column is contaminated with many other proteins of higher and lower molecular weight than the fusion protein. It was expected that this system would not give as pure protein as was obtained for the purification of the 52 kDa protein, as the isolation of a protein from inclusion bodies represents a significant purification step in itself, and the use of native conditions increases the frequency of non-specific biomolecular interactions. The soluble fraction contains a great many more proteins which may bind to the resin, and several histidine-rich proteins from E. coli have been reported to persistently bind to a variety of metal affinity columns (Hengen, 1995). It is also possible that some of the lower molecular weight contaminants are proteolysis products of the fusion protein, even though protease inhibitors have been added this may still be occurring. Some of the contaminants may be caused by ribosomal stalling and premature termination of translation, but with the E. coli thioredoxin protein itself at the N-terminus this should be kept to a minimum until the translational machinery reaches the part of the transcript encoding the N-terminal fragment of human topoisomerase IIα. It is also possible that some of the contaminants are artificial translation products starting from an internal methionine. Some of the larger contaminants may be caused by translational read-through past the stop codon introduced to terminate translation after residue Gln420, perhaps to the vector-encoded stop codon after the second C-terminal hexa-histidine tag. This would lead to the production of a fusion protein with a molecular weight of 68,480 Da. It can also be seen from Figure 3.6 that some of the fusion protein is eluted from the resin when 5 mM imidazole is present in the wash buffer (this was also seen with the purification of the 52 kDa protein), which suggests that the interaction between the His-tag and the resin is only very weak.
The cell pellet following protein expression was resuspended in 50 mM Tris-HCl (pH 7.5), 10% sucrose and lysed by French press. The clarified supernatant was applied to 4 ml of prewashed resin and protein was bound for 1 hour at 4°C. The resin was then washed with 4 x 5 column volumes of high salt wash buffer and decanted into a 10 ml disposable column. The resin was then washed with 5 x 1 column volume of low salt wash buffer plus 5 mM imidazole, and then eluted with 5 x 1 column volume of low salt wash buffer plus 50 mM imidazole. Fractions were analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lane 2, sample of whole cell protein extract from induced cells (C); lane 3, supernatant (S) following lysis and clarification of the cell suspension; lane 4, flow-off (FO) from the resin; lane 5, first wash fraction (W₁) with five column volumes of high salt wash buffer (50 mM Tris-HCl (pH 8.0), 500 mM NaCl); lanes 6-10, wash fractions (W₂-W₉) with one column volume of low salt wash buffer (50 mM Tris-HCl (pH 8.0), 250 mM NaCl) plus 5 mM imidazole; lanes 11-15, eluted fractions (E₁-E₅) with one column volume of low salt wash buffer plus 50 mM imidazole.
Given that the TALON resin did not yield sufficiently pure protein, a second purification scheme was required to purify the fusion protein to homogeneity. Thiobond resin was used to affinity purify the fusion protein following TALON purification using the reversible interaction between the thioredoxin protein and a metal ion covalently attached to an agarose-based support. Although this resin proved competent in binding the fusion protein, and many contaminants were removed in the washing stage, elution of the column with increasing concentrations of β-mercaptoethanol led to the co-elution of several persistent contaminating bands: a doublet at ~50 kDa, and a single band at ~24 kDa (data not shown). These bands also represent the major contaminating bands that are present in the eluant from the TALON column along with the fusion protein (see Figure 3.6, lanes 11-13). This would suggest that these contaminating bands are in fact proteolytic fragments of the fusion protein. This was partially confirmed by western blotting using a rabbit polyclonal antibody raised against the 52 kDa protein (J.R. Jenkins, University of Liverpool), which recognised the fusion protein and also the doublet at ~50 kDa, but not the band at ~24 kDa (data not shown).

3.2.1.2.2 Ion-exchange chromatography

As an alternative to a second affinity-tag purification scheme for the fusion proteins, the technique of ion-exchange chromatography was investigated. Using computer prediction software (Isoelectric Point service, EMBL Computational Services), the fusion protein 1-420F(Ek) has a predicted isoelectric point (pI) of 8.37. Therefore, purification would be best achieved with a strong cation exchanger. This is based on the fact that at a pH below its pI, the fusion protein will carry a positive charge and thus bind to a negatively charged matrix. The column selected, HiTrap SP (Pharmacia), uses a sulphopropyl (SP) functional group as the cation exchanger, and has a volume of 5 ml with a capacity of ~250 mg of protein.

The pooled fractions containing the fusion protein from the TALON column were diluted 5-fold with 50 mM sodium phosphate buffer (pH 7.0), 0.5 mM DTT, 0.5 mM EDTA to reduce the pH to ~7.0 and to reduce the salt concentration to 50 mM to allow the protein to bind to the column. This sample was then loaded onto the column
Figure 3.7 Purification of 1-420F(Ek) on HiTrap SP column.
The pooled fractions containing 1-420F(Ek) from the immobilised metal affinity column were purified on a HiTrap SP cation exchange column. The protein was diluted 5-fold with 50 mM phosphate buffer (pH 7.0), 0.5 mM DTT, 0.5 mM EDTA, and the sample was loaded onto the column at 4°C using a pump. The column was washed using phosphate buffer plus 50 mM NaCl, and then eluted with a stepwise NaCl gradient from 75 mM to 1 M. Fractions were analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lane 2, pooled fraction from TALON column (C); lane 3, flow-off (FO) from the column; lanes 4-13, eluted fractions with 50 mM phosphate buffer (pH 7.0) plus [NaCl] of 50 mM, 75 mM, 100 mM, 125 mM, 150 mM, 200 mM, 250 mM, 300 mM, 500 mM and 1M.
Figure 3.8 ATPase activity of 1-420F(Ek).
The ATPase activity of the fusion protein 1-420F(Ek) was assayed at 50 nM and 200 nM with ATP concentrations of 0 to 4 mM. The reaction sample (200 μl) consisted of 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 3% glycerol (w/v), 0.4 mM phosphoenolpyruvate, 0.2 mM NADH, 5 μl PK/LDH (in 50% (w/v) glycerol, 100 mM KCl, 10 mM HEPES (pH 7.0), 0.1 mM EDTA), 2 mM MgCl\(_2\), 0.5 mM DTT, 0.05 mM EDTA, 3.5 nM supercoiled pBR322 and ATP concentrations as indicated. The reaction was started by the addition of enzyme and incubated at 37°C for 1 hour, and the change in absorbance at 340 nm was related to the ATPase rate (nM/sec) using \(A^{1\text{mM}}_{340} = 6.22\).
using a pump at a flow rate of ~5 ml/min at 4°C. The column was then washed with 50 mM phosphate buffer (pH 7.0), 50 mM NaCl, 0.5 mM DTT, 0.5 mM EDTA, and then eluted with a stepwise salt gradient of 75 mM to 1 M NaCl in the same phosphate buffer. Figure 3.7 represents a typical purification of the fusion protein 1-420F(Ek) in this way. The fusion protein is characteristically eluted at 150-200 mM NaCl and is judged to be greater than 95% pure by SDS-PAGE analysis. Fractions containing fusion protein were dialysed overnight at 4°C in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 10% (w/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA (this buffer is similar to that used to store the 52 kDa protein), and stored at -70°C.

The purified fusion protein was assayed for ATPase activity, as it had been previously reported that many proteins expressed as a thioredoxin fusion retain their activity (LaVallie et al., 1993). This would indicate whether the protein expression and purification procedure used was leading to a correctly folded protein. The fusion protein was indeed found to have an intrinsic ATPase activity that is stimulated approximately 5-fold by DNA, as was the case with the 52 kDa protein. Figure 3.8 shows the ATPase activity of the fusion protein as a function of substrate (ATP) concentration. The initial rate was found to be hyperbolically dependent on substrate concentration, indicative of Michaelis-Menten kinetics, with calculated values of $K_M$ and $k_{cat}$ at 50 nM enzyme of 0.52 mM and 0.006 s$^{-1}$ in the absence of DNA; and 0.23 mM and 0.033 s$^{-1}$ in the presence of DNA.

### 3.2.1.3 Cleavage of fusion proteins with Enterokinase

The 1-420F(Ek) fusion protein is cleaved by enterokinase at the sequence DDDDK↓ to release the ~48 kDa N-terminal topoisomerase IIα subfragment and the ~17 kDa thioredoxin tag. Trial cleavage experiments with recombinant enterokinase (Novagen) were performed on 10 µg of fusion protein in Ek Cleavage Buffer (20 mM Tris-HCl (pH 7.4), 50 mM NaCl, 2 mM CaCl$_2$) at room temperature overnight for 16 hours. Enterokinase was titrated in from 0 to 0.5 units (where 1 unit is the amount required to cleave 50 µg of fusion protein in 16 hours at 23°C). The extent of cleavage was then observed by SDS-PAGE electrophoresis (Figure 3.9A). Proteolytic cleavage
Figure 3.9
3.9A Proteolytic cleavage of 1-420F(Ek) with recombinant enterokinase.
10 μg of 1-420F(Ek) was cleaved overnight at room temperature for 16 hours in Ek Cleavage Buffer (20 mM Tris-HCl (pH 7.4), 50 mM NaCl, 2 mM CaCl$_2$) with 0 to 0.5 units Ek and cleavage was monitored by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lanes 2-9, cleavage with 0U (control), 0.1U, 0.15U, 0.2U, 0.25U, 0.3U, 0.4U and 0.5U Ek as indicated.

3.9B Proteolytic cleavage of 1-420F(Ek) in the presence of ADPNP.
Prior to Ek-cleavage as indicated in Figure 3.9A, samples were incubated in the presence or absence of 2 mM ADPNP for 30 minutes at room temperature, and then cleaved as before and analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lanes 2 & 3, cleavage with 0U (control) in the absence and presence of ADPNP respectively; lanes 4 & 5, cleavage with 0.1U Ek in the absence and presence of ADPNP respectively; lanes 6 & 7, cleavage with 0.2U Ek in the absence and presence of ADPNP respectively; lanes 8 & 9, cleavage with 0.4U Ek in the absence and presence of ADPNP respectively.
of 1-420F(Ek) gave rise to three unexpected bands, as well as the cleaved (C) N-terminal fragment of topoisomerase IIα (1-420C(Ek)), caused by non-specific degradation, at approximately 43 kDa, 30 kDa and 18 kDa. These bands appear even in the presence of the lowest amount of enterokinase used (0.1 units). By inspection of the protein sequence of the N-terminal domain of topoisomerase IIα and using the estimated molecular weights of these bands, it was deduced that the non-specific cleavage was occurring at a site approximately one third of the way in from the N-terminus of the protein. This corresponds to the sequence ‘...DDDEK...’ at residues 152-156 of the protein (see Figure 3.1). Cleavage at this site would produce two proteins of ~30 kDa and ~18 kDa, and the ability of glutamate to substitute for aspartic acid (both acidic residues) at the cleavage site for enterokinase has been previously reported (Light & Janska, 1989). The origin of the ~43 kDa protein is unknown, since it is not obvious from the protein sequence. Attempts to optimise the cleavage reaction to the specific site using an alternative source of Ek, length of incubation and volume of reaction were unsuccessful.

It was observed that the non-specific cleavage site at residues 152-156 is close to the residues that are implicated in forming the ATP-binding site in DNA topoisomerase IIα by homology with the *E. coli* 43 kDa domain crystal structure (Wigley *et al.*, 1991) and sequence alignment (Caron & Wang, 1994). In all type II topoisomerases, a Walker motif (Walker *et al.*, 1982) for ATP coordination defined by a ‘G-loop’, also known as a phosphate-binding or ‘P-loop’ (Saraste *et al.*, 1990), occurs in the ATPase domain, and in human DNA topoisomerase IIα this corresponds to residues Gly161, Gly164 and Gly166 (see Figure 3.1). Therefore, it is possible that the binding of ATP or its analogue ADPNP may induce a conformational change in the region of the non-specific site that reduces the extent of cleavage at this site. The fusion protein 1-420F(Ek) was preincubated with ADPNP at a final concentration of 2 mM for 30 minutes at room temperature, and then cleaved overnight with enterokinase as before. Figure 3.9B indicates that preincubation with ADPNP does reduce the amount of non-specific cleavage that is occurring, judging by the reduced level of the band at ~30 kDa, and also the band at ~43 kDa. However, preincubation with ADPNP did not reduce the non-specific cleavage to an acceptable level for quantitative recovery of the 48 kDa N-
terminal domain of human topoisomerase IIα, and there was also the added problem of removing the ADPNP following cleavage.

As an alternative to using an enterokinase-cleaved fusion protein, there was also the option of using the Factor Xa-cleaved protein using the same system. To test the validity of this system, the fusion protein 1-420F(Ek) was used in trial cleavage experiments with Factor Xa according to the manufacturer's instructions (Boehringer Mannheim). These test cleavage experiments revealed that the cleavage of a Factor Xa-based clone would be more successful, as the extent of non-specific cleavage by Factor Xa was very low under standard cleavage conditions, although there was some proteolytic cleavage at high ratios of Factor Xa to fusion protein (data not shown).

3.2.2 Fusion proteins with a Factor Xa cleavage site

3.2.2.1 Cloning and expression of Xa/LIC fusion proteins

The pET-32 Xa/LIC vector (Novagen) encodes the same fusion protein as the Ek/LIC vector, except that the Xa/LIC site encodes a Factor Xa cleavage site (IEGR↓) to replace the Ek cleavage site. The pET-32 Xa/LIC vector is 5926 bp in length with 12 to 15 nucleotide 5’ overhangs that are different to those used in Ek/LIC cloning, and so new primers were designed. A sense PCR primer was designed with a 15 bp 5’ extension to generate the overhang required for LIC cloning, and a 22 bp region of the target gene (primer SC 10, 5’ - GGT ATT GAG GGT CGC ATG GAA GTG TCA CCA TTG CAG C - 3’; the 15 bp extension is underlined). The first nucleotide of the insert-specific sequence is an A, which begins the ATG initiation codon. Unlike with the Ek/LIC vector, this means that the first amino acid encoded following the Factor Xa cleavage site does not have to be a methionine (or isoleucine), and can in fact be any amino acid. Antisense primers were then designed, with 17 bp 5’ extensions, that would lead to the production of fusion constructs encoding the expression of residues 1-420 (primer SC11, 5’ - AGA GGA GAG TTA GAG CCT CAT TGG GCC TTA AAC TTC ACC - 3’), 1-425 (primer SC12, 5’ - AGA GGA GAG TTA GAG CCT CAC TTG TTT AAC TGG ACT TG - 3’), 1-453 (primer SC16, 5’ - AGA GGA GAG TTA GAG CCT
Figure 3.10 Plasmid maps of Xa/LIC vectors.
Plasmid maps for pTOP1-420Xa, pTOP1-425Xa, pTOP1-453Xa, pTOP1-454Xa and pTOP1-455Xa are shown. The positions of the thioredoxin (Trx) and human DNA topoisomerase IIα gene (Topoll) are shown, as well as the bla gene for ampicillin resistance. Also indicated are the restriction sites for HincII.
Figure 3.10 (cont) Plasmid maps of Xa/LIC vectors.
CAA GTG GAG TTT CGG C - 3'), 1-454 (primer SC17, 5' - AGA GGA GAG TTA
GAG CCT CAC TCA GTG GAG TTT CGG - 3') and 1-455 (primer SC18, 5' - AGA
GGA GAG TTA GAG CCT CAA CAA CAC TCA GTG GAG TTT CG - 3') of human
DNA topoisomerase IIα. The clones expressing residues 1-453/454/455 were
constructed in light of evidence from limited proteolysis of human DNA topoisomerase
IIβ (Austin et al., 1995). This enzyme, which shows high sequence homology to the α
isoform, is cleaved by SV8 protease following residue Glu470, which corresponds to
residue Glu454 in the α isoform (Figure 3.1). These oligonucleotides were used in PCR
reactions as for the Ek/LIC vectors, using pYES2αR as template (a full-length yeast
expression vector for human DNA topoisomerase IIα; T. R. Hammonds, University of
Leicester), and the correct incorporation of the insert was confirmed by restriction digest
and DNA sequencing throughout the coding region for human DNA topoisomerase IIα.
These five constructs were designated pTOP1-420Xa, pTOP1-425Xa, pTOP1-453Xa,
pTOP1-454Xa and pTOP1-455Xa (Figure 3.10):

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Total bp</th>
<th>TopoIIα residues expressed</th>
<th>MW of TopoIIα fragment (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTOP1-420Xa</td>
<td>7192</td>
<td>1-420</td>
<td>47,929</td>
</tr>
<tr>
<td>pTOP1-425Xa</td>
<td>7207</td>
<td>1-425</td>
<td>48,111</td>
</tr>
<tr>
<td>pTOP1-453Xa</td>
<td>7291</td>
<td>1-453</td>
<td>51,458</td>
</tr>
<tr>
<td>pTOP1-454Xa</td>
<td>7294</td>
<td>1-454</td>
<td>51,587</td>
</tr>
<tr>
<td>pTOP1-455Xa</td>
<td>7297</td>
<td>1-455</td>
<td>51,690</td>
</tr>
</tbody>
</table>

Expression of these fusion proteins in BL21(DE3) cells grown at 37°C is
shown in Figure 3.11A. As was seen with the Ek-vectors, all five fusion proteins are
overexpressed to approximately the same level in both the induced and uninduced
samples (predicted molecular weights are 1-420F fusion protein, 65,132 Da; 1-425F,
65,714 Da; 1-453F, 68,661 Da; 1-454F, 68,790 Da; 1-455F, 68,893 Da). To determine
the extent of expression as soluble protein, the induced fractions shown in Figure 3.11A
were separated into soluble and insoluble protein by sonicating as before. As can be
seen from Figure 3.11B, the vast majority of the expressed fusion protein is present in
the insoluble protein fraction. However, there appears to be more soluble protein
3.11A Expression of Xa/LIC fusion proteins at 37°C.
Cultures (10 ml LB) of pTOP1-420Xa, pTOP1-425Xa, pTOP1-453Xa, pTOP1-454Xa and pTOP1-455Xa in BL21(DE3) cells were grown at 37°C until an OD₆₀₀ of ~0.4 was reached. Cultures were then split into two, one was induced with 0.1 mM IPTG and growth was continued for 3 hours. Total cell protein was then analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lane 2, untransformed BL21(DE3) cells (BL21); lanes 3-12, uninduced (Un) and induced (In) protein from cells expressing the indicated fusion protein.

3.11B Soluble and insoluble protein from expression of fusion proteins.
The induced fractions from Figure 3.11A were separated into soluble and insoluble protein by sonication and analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lanes 2-11, soluble (S) and insoluble (P) protein, resuspended in urea, from cells expressing the indicated fusion protein.
recovered for the longer fusion proteins i.e. 1-453F, 1-454F and 1-455F, than for the shorter 1-420F and 1-425F (compare lanes 6, 8 & 10 with lanes 2 & 4). This would suggest that a greater proportion of the longer fusion proteins are folding correctly, perhaps because the extra ~30 amino acids contains the domain boundary or allows the protein to adopt a more stable structure.

Test expressions were also carried out by growing cells at 20°C, which is known to increase the soluble expression of some proteins otherwise found as inclusion bodies at 37°C (Hannig & Makrides, 1998). Expression of the protein at 20°C resulted in reduced levels of overexpression of the fusion proteins compared to 37°C (Figure 3.12A). When separated into soluble and insoluble protein (Figure 3.12B), the proportion of soluble protein recovered was found to be greater than from cells grown at 37°C. To confirm this, the protein 1-420F was expressed and purified from a 500 ml LB culture grown at 20°C and 37°C using the method as given in section 3.2.1.2.1, except that at 20°C the cells were grown for 16 hours after inducing protein expression. The yield of purified soluble protein was found to be ~3-fold greater at 20°C compared to 37°C. Thus, a growth temperature of 20°C was used in all subsequent protein expressions.

3.2.2.2 Optimisation of expression and purification

Based on the fact that the fusion protein 1-420F(Ek) was found to be active in ATPase assays, it was decided that the equivalent Factor Xa-cleaved protein would be used to try to further optimise the expression of the fusion proteins. One litre cultures (2 x 500 ml in 2 l baffled flasks) of BL21(DE3)[pTOP1-420Xa] were grown at 20°C in each of four different growth media: Luria-Bertani Broth (LB), Nutrient Broth (NB), Terrific Broth (TB) and 2YT broth (Sambrook et al., 1989); and the fusion protein was purified separately from each culture. The final recoveries of fusion protein were: LB, ~0.53 mg; NB, <0.1 mg; TB, ~0.85 mg; 2YT, ~0.93 mg. The yield of fusion protein correlates with the cell density of the culture at harvesting (16 hours post-induction) i.e more protein is recovered from TB and 2YT cultures than from LB and NB cultures because there are a greater number of cells present at harvesting, not because the protein
Figure 3.12

3.12A Expression of Xa/LIC fusion proteins at 20°C.
Expression of the Xa/LIC fusion proteins was the same as in Figure 3.11A, except the cultures were grown at 20°C and induced for 16 hours. Protein was analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lane 2, untransformed BL21(DE3) cells (BL21); lanes 3-12, uninduced (Un) and induced (In) protein from cells expressing the indicated fusion protein.

3.12B Soluble and insoluble protein from expression of fusion proteins.
The induced fractions from Figure 3.12A were separated into soluble and insoluble protein by sonication and analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lanes 2-11, soluble (S) and insoluble (P) protein, resuspended in urea, from cells expressing the indicated fusion protein.
is expressed in a more soluble form. Based on this, 2YT broth was used as growth media in subsequent protein expressions.

One of the problems with expressing a recombinant eukaryotic protein in a prokaryotic host such as *E. coli* is that of codon usage, whereby the codon bias of the foreign gene differs from that of the host organism. The problem arises because of the rarity of certain tRNA species in the host organism compared to the native organism, which can lead to depleted levels of these tRNA species during protein expression and stalled translation. One way of overcoming this problem is to artificially introduce these tRNA species into the host organism. Such a system is used in commercially available BL21-CodonPlus(DE3)-RIL cells (Stratagene), which contain a plasmid bearing the *argU, ileY* and *leuW* genes, which encode tRNA species that recognise the arginine codons AGA and AGG, the isoleucine codon AUA, and the leucine codon CUA (Carstens & Waesche, 1999). The arginine AGA and AGG codons are the rarest codons used in *E. coli*, but are relatively common in human genes. Analysis of the gene encoding residues 1-420 of human DNA topoisomerase IIα reveals that there are a total of 11 AGA/AGG codons, 4 CTA codons and 6 ATA codons present which may be having an effect on protein translation. To test for any effect, BL21-CodonPlus(DE3)-RIL cells were compared to BL21(DE3) cells in test expressions of 1-420F as before. At 37°C, there was no visible difference between the two cell types, both giving a high level of overexpression in both the uninduced and induced sample (data not shown). However, at 20°C there is more expression of 1-420F in the CodonPlus cells compared to the standard BL21(DE3) cells (Figure 3.13A). This suggests that at a slower growth rate, the level of these three tRNA species in normal BL21(DE3) cells is depleted to a point that reduces the level of translation of the fusion protein. The induced fractions from Figure 3.13A were separated into soluble and insoluble protein as before. Figure 3.13B shows that the additional protein produced in the CodonPlus cells is present in the insoluble fraction, and that there is no visible difference in the recovery of soluble protein between the two cell types.

Figures 3.13A and 3.13B show that the extra copies of the three tRNA species give rise to increased expression of the fusion protein, albeit in an insoluble form.
Figure 3.13

3.13A Expression of 1-420F in BL21 and BL21-CodonPlus cells.
10 ml LB cultures of pTOPl-420Xa in BL21(DE3) and BL21-CodonPlus (DE3) cells were grown at 20°C and protein expression was induced as before. Total cell protein was analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lane 2, untransformed BL21(DE3) cells (BL21); lanes 3-6, uninduced (Un) and induced (In) protein from BL21(DE3) cells, in duplicate; lanes 7-10, uninduced (Un) and induced (In) protein from BL21-CodonPlus (DE3) cells, in duplicate; lane 11, untransformed BL21-CodonPlus (DE3) cells (C+).

3.13B Soluble and insoluble protein from expression of fusion proteins.
The induced fractions from Figure 13.3A were separated into soluble and insoluble protein by sonication and analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lanes 2-5, soluble (S) and insoluble (P) protein from BL21(DE3) cells, in duplicate; lane 6-9, soluble (S) and insoluble (P) protein from BL21-CodonPlus (DE3) cells, in duplicate.
However, they do not show whether these extra tRNA species also cause decreased translational stalling which may be occurring, and would thus reduce the levels of any abortive translation products that have the appearance of proteolytic products (as have been seen during the purification stages). To test for this, a 500 ml culture (2YT) of pTOP1-420Xa in BL21(DE3) and BL21-CodonPlus(DE3)-RIL cells was grown and protein expression was induced as before. From each culture, three 150 ml samples were removed and centrifuged to pellet the cells. Three different methods were then used to lyse the cells, to determine the best method for the release of soluble protein. One fraction from both cell types was resuspended in 50 mM Tris-HCl (pH 7.5), 10% (w/v) sucrose and passed through a pre-cooled French press three times at 8,000-12,000 psi. One fraction was resuspended in 50 mM Tris-HCl (pH 7.5), 10% (w/v) sucrose and lysed by sonication on ice for 3 cycles of 30 seconds on/30 seconds off. The final sample was resuspended in a non-ionic detergent designed for protein extraction from *E. coli* (BugBuster, Novagen). The sample is gently mixed for 10 minutes at room temperature, which perforates the cell wall without denaturing soluble protein. All lysates were subsequently clarified by centrifugation and partially purified using TALON resin. Figure 3.14 shows a comparison of the pooled fractions from the TALON column containing the eluted protein, and a sample of the whole cell extract from the induced cells. From the induced cell extracts, the CodonPlus cells and the normal BL21 cells have overexpressed the fusion protein to an approximately similar level. For both cell types, cell lysis using the French press led to the highest recovery of fusion protein, sonication led to a slightly reduced level of recovery, and detergent-lysis led to a significantly reduced recovery of fusion protein. Detergent-lysis also led to the accumulation of a protein product of ~40 kDa not seen with French pressing and sonication. It is possible that this band represents a proteolytic fragment of the fusion protein, which is being rapidly degraded, perhaps because the room temperature incubation causes the proteases to be significantly more active than at 4°C. Comparing the purified protein from cell lysates prepared using the French press and sonication does not reveal any significant difference in the proportion of contaminant bands present in these fractions. This would suggest that using the CodonPlus(DE3)-RIL cells does not lead to reduced levels of any abortive translation products which bind to the TALON resin and are carried through the purification steps with the appearance of
Figure 3.14 Purification of 1-420F from BL21 and BL21-CodonPlus cells.
A 500 ml culture (2YT) of pTOP1-420Xa in BL21(DE3) and BL21-CodonPlus(DE3) cells was grown at 20°C and protein expression was induced as before. Each 500 ml culture was split into three 150 ml volumes, and the cells were disrupted either by French press, sonication, or using a non-ionic detergent (BugBuster, Novagen). The clarified lysates were partially purified using TALON resin, peak fractions were pooled and analyzed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lane 2, sample of whole cell protein extract from induced BL21(DE3) cells (BL21); lanes 3-5, purified protein following French press (FP), sonication (Son) and detergent-lysis (BB) of the cell suspension respectively for BL21(DE3) cells; lane 6, sample of whole cell protein extract from induced BL21-CodonPlus(DE3) cells (C+); lanes 7-9, purified protein following French press (FP), sonication (Son) and detergent-lysis (BB) of the cell suspension respectively for BL21-CodonPlus(DE3) cells.
proteolytic products. Also, the total yield of fusion protein from the CodonPlus cells is lower than that for the normal BL21 cells, despite their approximately equal level of expression of the fusion protein. Taken as a whole, these data suggest that there is no advantage to be gained in using the BL21-CodonPlus(DE3)-RIL cells as an alternative expression host to BL21(DE3) cells.

To investigate the native state of the purified fusion protein 1-420F, gel filtration was used. The technique of gel filtration is useful for separating a relatively homogeneous protein sample into monomeric, dimeric (or other multimeric structures) and higher order aggregate fractions. The latter may be particularly relevant in this case as it may separate the correctly folded soluble protein from any aggregated or incorrectly folded protein. During gel filtration, smaller molecules are retained in the pores of the gel matrix and are eluted at a later time than the larger molecules, which are not retained to the same extent. Aggregated proteins are generally not retained by the column at all and are the first to be eluted. Gel filtration is also useful as a final ‘polishing’ step in the purification of proteins. A sample of 1-420F was concentrated to ~1.3 mg/ml and 500 μl was applied to a HiPrep 16/60 Sephacryl S-200 HR gel filtration column (Pharmacia Biotech) in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 10% glycerol (w/v), 0.5 mM DTT, 0.5 mM EDTA. The elution profile revealed the presence of one peak corresponding to the fusion protein being eluted as a monomer (data not shown). A very small ‘shoulder’ peak was also observed, which corresponds to the minor contaminant of ~24 kDa seen before as a carry over from the TALON column. This would suggest that the purified fusion protein does not contain any aggregated material, and exists primarily as a monomer in solution.

In summary, the best expression conditions for the fusion proteins have been found to be in BL21(DE3) cells grown at 20°C in 2YT media, and with cell disruption by passage through a French pressure cell. However, the French pressed cell lysate is also sonicated for 30 seconds on/30 seconds off for 2 cycles to shear the genomic DNA and reduce the viscosity of the suspension. Due to the higher cell density reached in 2YT media, cultures are grown until an OD$_{600}$ of ~0.8 is reached, and then protein expression is induced by the addition of 0.1 mM IPTG and growth continued for a
further 16 hours. Also, the purified fusion proteins were dialysed into a storage buffer containing 50% glycerol (i.e. 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 50% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA) and stored at -20°C. The activity of protein was found to be unchanged after greater than 3 months when stored in this manner.

3.2.2.3 Characterisation of fusion proteins of the N-terminal domain

To characterise the five fusion proteins 1-420F, 1-425F, 1-453F, 1-454F and 1-455F, a one litre culture (2 x 500 ml) of all 5 clones in BL21(DE3) cells was grown at 20°C and protein expression was induced as before. The fusion proteins were subsequently purified from each culture. The final recovery of each fusion protein was: 1-420F, ~0.79 mg; 1-425F, ~0.36 mg; 1-453F, ~1.13 mg; 1-454F, ~1.14 mg; 1-455F, ~1.08 mg. This confirms the results seen in section 3.2.2.1 i.e. expression of the longer fusion proteins (1-453F, 1-454F and 1-455F) results in a significantly increased yield of purified soluble protein.

The ATPase activity of the five fusion proteins was determined at 500 nM fusion protein and 2 mM ATP (Figure 3.15). All five fusion proteins displayed an intrinsic ATPase rate which was found to be stimulated by DNA (relaxed pBR322 at 250 bp/dimer). Interestingly, the shorter fusion proteins (1-420F and 1-425F) displayed a smaller intrinsic ATPase rate, which was stimulated ~5-fold by DNA, whereas the longer fusion proteins (1-453F, 1-454F and 1-455F) had a higher intrinsic ATPase rate, but the level of DNA stimulation was only ~3-fold. This may suggest that the extra ~30 amino acids stabilise the protein in a conformation which is more competent to bind and hydrolyse ATP; perhaps by promoting dimer formation, leading to a higher intrinsic ATPase rate. The extra amino acids are reducing the DNA stimulation from ~5-fold to ~3-fold, and therefore may be altering the interaction with DNA or reducing the apparent stimulation of catalytic turnover that is usually seen in the presence of DNA because these proteins have a higher intrinsic rate.
Figure 3.15 ATPase activity of fusion proteins.
Each of the five fusion proteins was concentrated to ~0.5 mg/ml and the ATPase activity assayed at 500 nM. The reaction sample (200 µl) consisted of 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 5% glycerol (v/v), 0.4 mM phosphoenolpyruvate, 0.2 mM NADH, 5 µl PK/LDH (in 50% (w/v) glycerol, 10 mM HEPES (pH 7.0), 0.1 mM EDTA), 2 mM Mg-ATP (in sodium acetate buffer at pH 7.5), 0.5 mM DTT, 0.05 mM EDTA, 14 nM relaxed pBR322 (~250 bp/dimer). The reaction was started by the addition of enzyme and incubated at 37°C for 1 hour, and the change in absorbance at 340 nm was related to the ATPase rate (nM/sec) using $A^{1 \text{mM}}_{340} \text{NADH} = 6.22$. 
3.2.2.4 Cleavage of fusion proteins with Factor Xa

The optimal conditions for proteolytic cleavage with Factor Xa must be determined empirically, since the cleavage reaction can be performed at a temperature between 4°C and room temperature and with no strict requirements for buffer conditions. Test cleavage experiments were initially set up with 10 μg of 1-420F in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 10% (w/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA in a 50 μl volume, and Factor Xa was added from 0 to 2 μg. Samples were left at 4°C overnight for 16 hours and subsequently analysed by SDS-PAGE (Figure 3.16). With 0.25 μg of Factor Xa, a ratio of 1:40, the fusion protein is fully cleaved into the N-terminal fragment of topoisomerase IIα (1-420C) and the thioredoxin tag, with no non-specific degradation products visible (Figure 3.16, lane 5). At higher ratios of Factor Xa to fusion protein, a non-specific degradation product can be seen at ~40 kDa (Figure 3.16, lane 9), but considerably more protease is required before this appears. The band visible at ~33 kDa corresponds to Factor Xa, which is composed of two polypeptides that run at ~33 kDa and ~28 kDa. The ~48 kDa fragment seen following proteolytic cleavage was N-terminally sequenced (K. Lilley, PNACL, University of Leicester), and produced the sequence Met-Glu-Val-Ser-Pro-Leu, which corresponds to the first six amino acids of human DNA topoisomerase IIα. This confirmed that proteolytic cleavage is occurring at the correct site, and that the ~48 kDa fragment represents the released N-terminal domain of human DNA topoisomerase IIα (predicted molecular weight 47,929 Da). Following proteolytic cleavage, the biotin-labelled Factor Xa can be removed using an immobilised streptavidin-agarose resin which is incubated with the sample for 1 hour at 4°C and then removed by centrifugation; and PMSF is added to 1 mM to inhibit any remaining Factor Xa. Proteolytic cleavage of all five fusion proteins was found to require a ratio of Factor Xa to fusion protein of ~1:40, and an incubation time of 16 hours at 4°C to give full or near full cleavage.

The released fragment of human DNA topoisomerase IIα can be separated from the His-tagged fusion peptide (~17 kDa) using TALON resin, which will bind the fusion peptide and any uncleaved fusion protein. Before the cleaved sample is applied to the TALON resin, it must first be diluted with buffer to reduce the concentration of EDTA and DTT to <0.1 mM, since these reagents are incompatible with the TALON resin.
Figure 3.16 Proteolytic cleavage of 1-420F with Factor Xa.

10 µg of 1-420F was cleaved overnight at 4°C for 16 hours in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 10% glycerol (w/v), 0.5 mM DTT, 0.5 mM EDTA with 0 to 2 µg of Factor Xa, and the extent of cleavage was monitored by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); lanes 2-9, cleavage with 0 µg (control), 0.05 µg, 0.1 µg, 0.25 µg, 0.5 µg, 0.75 µg, 1.0 µg and 2.0 µg of Factor Xa respectively. The ratio of Factor Xa to fusion protein is also indicated.
After the protein is allowed to bind to the resin, the sample is centrifuged and the supernatant is removed (the 'flow-off'). The resin is then washed with a high salt buffer to maximise the recovery of the cleaved human DNA topoisomerase IIα fragment, some of which binds to the resin despite the His-tag being removed. The flow-off and high salt wash are then pooled and dialysed into storage buffer (50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 50% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA) and stored at -20°C. This leads to the recovery of approximately 60% of the released N-terminal fragment of human DNA topoisomerase IIα.

3.2.3 Initial characterisation of the ATPase activity of the N-terminal fragments topoIIα-(1-420) and topoIIα-(1-453)

From the expression studies and ATPase assays, there appears to be a significant difference between the two groups of proteins i.e. the short (1-420F/425F) and long (1-453F/454F/455F) proteins; but there is little difference between proteins within each group. On the whole, the two fusion proteins 1-420F and 1-453F were the best expressed and most active in terms of ATP hydrolysis for each group. To determine which of these proteins (in the cleaved state) would be used to characterise the N-terminal domain of human DNA topoisomerase IIα, each of the fusion proteins was prepared under identical conditions according to the procedure given in Materials and Methods. The purified fusion protein was cleaved with Factor Xa, and the released topoisomerase IIα fragment purified away from the thioredoxin tag using TALON resin. A summary gel representing the purification, Factor Xa cleavage, and subsequent repurification of each of these proteins is shown in Figure 3.17. In both cases, the N-terminal fragments representing residues 1-420 (47,929 Da) and 1-453 (51,458 Da) of human DNA topoisomerase IIα, topoIIα-(1-420) and topoIIα-(1-453) respectively, are recovered in a highly purified form (Figure 3.17, lane 7). The N-terminal sequence of both the cleaved proteins was verified by sequencing (K. Lilley, PNACL, University of Leicester).

The rate of ATP hydrolysis of each of the two N-terminal fragments was determined in the fusion and cleaved (native) state at 250 nM in the absence of DNA.
Figure 3.17 Summary of the expression and purification of proteins topoIIα-(1-420) and topoIIα-(1-453).

A representative gel for the expression and purification of the N-terminal fragments topoIIα-(1-420) (*top*) and topoIIα-(1-453) (*bottom*), according to the procedure given in Materials and Methods. Fractions were analysed by 12% SDS-PAGE. Lane 1, molecular weight markers (kDa); Lanes 2 & 3, total cell protein from uninduced (Un) and induced (In) BL21 cells; lane 4, clarified protein extract following French pressing (Supe); lane 5, purified fusion protein (420F and 453F); lane 6, Factor Xa-based cleavage of fusion proteins (Cleave); lane 7, purified N-terminal fragment of human topoisomerase IIα (420C and 453C).
Figure 3.18 ATPase activity of fusion and cleaved proteins.
The ATPase activity of the fusion (F) and cleaved (C) proteins representing residues 1-420 and 1-453 of the N-terminal domain were determined at 250 nM under standard conditions given in Materials and Methods. The reaction sample (150 μl) consisted of 50 mM Tris-HCl (pH 7.5), 10 mM NaCl, 5% glycerol (v/v), 0.4 mM phosphoenolpyruvate, 0.2 mM NADH, 5 μl PK/LDH (in 50% (w/v) glycerol, 100 mM KCl, 10 mM HEPES (pH 7.0), 0.1 mM EDTA), 2 mM Mg-ATP (in sodium acetate buffer at pH 7.5), 0.5 mM DTT, 0.05 mM EDTA and relaxed pBR322 at 50 bp/dimer and 250 bp/dimer as indicated. The reaction was started by the addition of enzyme and incubated at 37°C for 1 hour, and the change in absorbance at 340 nm was related to the ATPase rate (nM/sec) using A1 mM⁻³₄₀ = 6.22.
and in the presence of relaxed pBR322 at a ratio of 50 bp/dimer and 250 bp/dimer (Figure 3.18). Both the fusion and cleaved protein fragments display similar ATPase rates with and without DNA, although for the shorter fragment (1-420), the ATPase rates for the cleaved protein are approximately half of that for the fusion protein. For the shorter fragment topoIIα-(1-420), the protein has a low intrinsic ATPase rate that is increasingly stimulated by increasing the concentration of DNA. However, the longer fragment topoIIα-(1-453) has a significantly higher intrinsic ATPase rate (as seen in Figure 3.15) which is stimulated by DNA in a different manner from the shorter fragment. The ATPase rate of the enzyme increases sharply in the presence of DNA at 50 bp/dimer, but from there the rate decreases in the presence of DNA at 250 bp/dimer. This 'hyperstimulation' of the ATPase rate at high enzyme:DNA ratios has been seen before with the full-length human enzyme, and is attributed to protein-protein interactions (Hammonds & Maxwell, 1997).

3.3 DISCUSSION

The expression of a 52 kDa protein, the putative N-terminal ATPase domain of human DNA topoisomerase IIα, an 11 amino acid polylinker and His6 tag, has again highlighted some of the problems associated with this protein. Although the isolation of proteins from inclusion bodies represents a key purification step, the denaturing conditions required necessitates a refolding process that may or may not be efficient at regenerating soluble, active protein. The inherent insolubility of this protein makes it a difficult protein to work with, and, along with the inconsistent activity found between preparations, raises the question as to how the activity of this protein compares to the native protein in vivo. Expression of a variety of fragments of human DNA topoisomerase IIα in E. coli have also led to the production of insoluble aggregates, with no detectable activity in the crude extract, despite their relatively high degree of overexpression (Kikuchi & Miyake, 1993). The exact location of the domain boundary between the N-terminal ATPase domain and central DNA-cleavage domain is unknown for human DNA topoisomerase IIα, and this may account for the insolubility of the 52 kDa protein. To overcome this problem, I have constructed fusion proteins of a variety
of fragments of the putative N-terminal domain of human DNA topoisomerase IIα in a covalent link to the thioredoxin protein of *E. coli*.

Based on the N-terminal domains of yeast topoisomerase II, *E. coli* GyrB and human DNA topoisomerase IIβ, a variety of constructs were made that fused an N-terminal fragment of human DNA topoisomerase IIα homologous to the N-terminal domain of these three enzymes to the thioredoxin protein. All of these fusion proteins were found to be expressed as insoluble inclusion bodies when cell growth occurred at 37°C. Lowering the growth temperature to 20°C increased the recovery of soluble protein to a level that was acceptable for biochemical and structural studies. The best levels of soluble expression were found with the fragments of human DNA topoisomerase IIα that correspond to the putative ATPase domain of human DNA topoisomerase IIβ (Austin *et al*., 1995), a homologous protein with a high degree of sequence identity. Based on this evidence it is impossible to say whether these longer fragments represent the complete ATPase domain of the enzyme, but the extra ~30 amino acids does appear to stabilise the fragment in a more soluble form.

The two step purification procedure employed yielded highly purified protein, but also suggests that a high degree of proteolytic degradation and/or translational stalling may be occurring *in vivo* and *in vitro*. Use of codon-bias adjusted BL21(DE3) cells did not overcome the possible problem of translational stalling, even though they enhanced the rate of protein expression over that of normal BL21(DE3) cells. The accumulation of significant amounts of truncated forms of protein has also been observed with the expression of human DNA topoisomerase I in *E. coli* (Kikuchi & Miyake, 1993). When expressed in *E. coli* this protein is very unstable, and proteolytic breakdown products accumulate to levels similar to the full-length protein. However, the activity of the human DNA topoisomerase I in the crude extract was found to be quite high, even in the uninduced state. The topoisomerase I gene is also believed to contain cryptic prokaryotic promoter elements that cause the expression of truncated forms of the protein from artificial translation at internal methionines. To overcome the problem of low yield in the expression of human DNA topoisomerase I, Stewart & Champoux (1999) have used the baculovirus-insect cell expression system to produce
milligram quantities of various forms of this protein. A typical yield of 20-30 mg of purified protein is obtained, with an activity comparable to that of protein purified from HeLa cells.

Proteolytic cleavage of the enterokinase-cleaved fusion proteins at the non-specific site was used as a tool to monitor the conformational change that occurs in the protein upon the binding of ADPNP. This conformational change is likely to be due to dimerisation, as seen in the case of the N-terminal fragment of E. coli GyrB (Ali et al., 1993). Cleavage of alternative fusion proteins with Factor Xa was more successful. The cleaved N-terminal topoisomerase IIα could be easily separated from the thioredoxin tag and recovered in a highly purified form. The purified proteins were stored in a 50% glycerol buffer at -20°C, in which they retained their activity for a minimum of 3 months and were routinely stored at a concentration of ~1 mg/ml with no ill effects. Furthermore, the protein topoIIα-(1-453) from four different preparations revealed a less than 30% variation in its ATPase activity in the presence and absence of DNA i.e. this protein has a much more consistent activity than the 52 kDa protein.

Both the fusion and cleaved proteins were found to display similar rates of ATP hydrolysis. The stimulation of the ATPase activity of the fragment topoIIα-(1-420) by DNA was similar to that seen with the 52 kDa protein (Gardiner et al., 1998), in that increasing the concentration of DNA leads to an increase in the ATPase rate. In the case of the fragment topoIIα-(1-453), the presence of the extra 33 residues significantly increased the intrinsic ATPase activity of this protein, and had a profound effect on the DNA-stimulated ATPase activity. The hyperstimulation of the ATPase activity seen with this fragment at 50 bp/dimer is similar to that seen with the full-length protein at ~150 bp/dimer (Hammonds & Maxwell, 1997). This hyperstimulation was attributed to protein-protein interactions occurring between adjacent enzyme dimers on a saturated DNA molecule. At the higher DNA concentration of 250 bp/dimer, this molecular crowding effect diminishes as the distance between enzyme dimers increases, and therefore the ATPase rate decreases. The hyperstimulation effect that is seen with the fragment topoIIα-(1-453) at 50 bp/dimer, but not with topoIIα-(1-420), suggests that the
extra 33 residues are interacting with DNA and/or with the same enzyme dimer or an adjacent enzyme dimer to cause an increase in the ATPase rate.

In view of the increased yield and ATPase activity of the N-terminal fragment topoIIα-(1-453) over that of topoIIα-(1-420), and because of the interesting phenomenon of hyperstimulation of the ATPase rate that is seen with the longer protein in the presence of DNA, the fragment topoIIα-(1-453) was used in the study of the DNA-dependent stimulation and hyperstimulation of the ATPase activity of the N-terminal domain of human DNA topoisomerase IIα (Chapter 4), and its interaction with DNA and nucleotides (Chapter 5).
CHAPTER 4

DNA-dependent ATPase activity of the N-terminal fragment topollα-(1-453)
4.1 INTRODUCTION

The type II topoisomerases are DNA-dependent ATPases; the magnitude of DNA stimulation varies from approximately 3-20-fold (Maxwell & Gellert, 1984; Lindsley & Wang, 1993b; Hammonds & Maxwell, 1997; reviewed in Wang, 1998). The relative contributions of the G and T segments to the DNA-stimulated ATPase rate are controversial and, although there is evidence that suggests that the T segment does play a role in stimulating the ATPase activity of the enzyme, it is presently unclear whether the DNA-stimulated ATPase rate is a result of interaction of the enzyme with the G segment alone or both the G and T segments together (reviewed in Wang, 1998).

In Chapter 1, the hyperstimulation of the ATPase rate of human topoisomerase IIα in the presence of low concentrations of DNA was described (Hammonds & Maxwell, 1997); this hyperstimulation effect was also seen with the N-terminal protein topoIIα-(1-453) in Chapter 3, but not with the shorter fragment topoIIα-(1-420), nor with the 52 kDa protein (Gardiner et al., 1998). As has been previously discussed, this hyperstimulated ATPase rate was attributed to either protein-protein interactions occurring on enzyme-saturated DNA molecules, or to the stabilisation of the enzyme complex in a state with the G segment bound and the T segment permanently trapped within the ATPase domains (Hammonds & Maxwell, 1997). The fact that this hyperstimulation effect also occurs with the N-terminal fragment topoIIα-(1-453) at low DNA concentrations raises interesting questions as to what causes this increased ATPase rate and what role it may play in the normal catalytic cycle of the full-length enzyme. In this chapter, the fragment topoIIα-(1-453) has been used to study the effect that DNA, presumably corresponding to the T segment, has on the ATPase activity of the N-terminal domain of human DNA topoisomerase IIα, and more specifically, the dependence of the hyperstimulated ATPase rate on DNA length.

4.2 RESULTS

The ATPase activities of the N-terminal fragments of human topoisomerase IIα have been determined using the PK/LDH linked enzyme assay, which is described in the
Materials and Methods section. The standard conditions used in ATPase assays throughout this chapter (unless otherwise stated) are: 50 mM Tris-HCl (pH 7.5), 10 mM NaCl, 5% glycerol (v/v), 0.4 mM PEP, 0.2 mM NADH, 5 µl PK/LDH (in 50% (w/v) glycerol, 100 mM KCl, 10 mM HEPES (pH 7.0), 0.1 mM EDTA), 2 mM Mg-ATP (in sodium acetate buffer at pH 7.5), 0.5 mM DTT, 0.05 mM EDTA. DNA, where present, was added at the indicated concentration or bp:dimer ratio. Reactions were initiated by the addition of enzyme and monitored continually for 1 hour at 37°C in an absorbance plate reader. The initial rate of ATP hydrolysis was calculated following a 5 minute equilibration period using a minimum of 10 minutes of data collection.

4.2.1 Initial characterisation of the ATPase activity of the N-terminal fragment topollα-(1-453)

4.2.1.1 ATPase activity as a function of salt concentration

The ATPase activity of topollα-(1-453) in the presence and absence of DNA was found to be significantly affected by the concentration of NaCl in the reaction buffer; increasing the concentration of NaCl from 5 mM to 300 mM produced a rapid decline in the ATPase activity of the protein (Figures 4.1A and 4.1B). Most affected by increasing the NaCl concentration was the sample containing DNA at 50 bp/dimer, where the ATPase rate decreased rapidly even at low NaCl concentrations. The ATPase activity at 50 bp/dimer dropped below that at 250 bp/dimer at an NaCl concentration of between 25 and 50 mM i.e. the hyperstimulation effect is abolished at approximately 50 mM NaCl or higher. The hyperstimulation in the ATPase rate of the full-length enzyme was also previously reported to be highly sensitive to the monovalent cation concentration in the assay mixture (Hammonds & Maxwell, 1997). In the absence of DNA it is likely that the effect of increasing NaCl concentration is to reduce the monomer-dimer association constant; whereas, in the presence of DNA, and particularly at 50 bp/dimer, the increasing NaCl concentration reduces the interaction between the DNA and enzyme as well as the enzyme-enzyme interaction, leading to a lower ATPase rate.
Figure 4.1

4.1A ATPase activity of topoIIα-(1-453) as a function of NaCl concentration.

The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at 250 nM at varying NaCl conditions, using the standard conditions given in Materials and Methods and described in the legend to Figure 3.18. Where present, DNA (relaxed pBR322) was added at 50 bp/dimer or 250 bp/dimer as indicated.
Figure 4.1 (cont)

4.1B ATPase activity of topoIIα-(1-453) as a function of NaCl concentration.
The same data from the Figure 4.1A are presented in a different format, with the results for each NaCl concentration grouped together for samples in the presence and absence of DNA.
4.2.1.2 ATPase activity as a function of enzyme concentration

In the absence of DNA, the ATPase activity of topoIIα-(1-453) exhibits a non-linear, greater than first order dependence upon enzyme concentration (Figure 4.2A), indicative of an oligomerisation process occurring during the course of the ATPase cycle. This result differs from that seen with the 52 kDa fragment of human topoisomerase IIα, where the rate of ATP hydrolysis was shown to be linearly dependent upon enzyme concentration (Gardiner et al., 1998). This non-linear dependence of ATPase activity upon enzyme concentration was also seen with the 43 kDa fragment of E. coli DNA gyrase (Ali et al., 1993) and the N-terminal domain of yeast topoisomerase II (Olland & Wang, 1999), and is assumed to be a result of the dimerisation step in the ATPase cycle, which represents the rate-limiting step in the reaction. The curve drawn through the data in Figure 4.2A is a theoretical curve based on the rate equation given in Ali et al. (1993) (see Figure 4.2B), and was fitted to the data using MacCurveFit (Kevin Raner Software, Mt. Waverly, Australia).

In contrast, in the presence of DNA at 50 bp/dimer and 250 bp/dimer, the rate of ATP hydrolysis shows a linear dependence upon enzyme concentration (Figure 4.2C). This result suggests that in the presence of DNA the dimeric form of the enzyme is stabilised.

4.2.1.3 ATPase activity as a function of ATP concentration

At a constant enzyme concentration and varying substrate (ATP) concentration, the fragment topoIIα-(1-453) displays a hyperbolic dependence of ATPase rate on ATP concentration both in the absence (Figure 4.3A) and presence (Figure 4.3B) of DNA. In the absence of DNA, and using the data obtained at 250 nM enzyme, the turnover number of topoIIα-(1-453) is 0.072 s⁻¹/dimer, with a $K_M$ of 0.69 mM. With DNA at 250 bp/dimer the $k_{cat}$ increases to 0.71 s⁻¹/dimer, with a $K_M$ 0.89 mM, which represents an increase in turnover number of approximately 10-fold. At 50 bp/dimer i.e. under conditions of hyperstimulation of the ATPase rate, the $k_{cat}$ is increased to 1.1 s⁻¹/dimer, with a $K_M$ of 0.76 mM, which represents a stimulation of the intrinsic ATPase rate by approximately 15-fold. The values of $k_{cat}$ in the absence of DNA and at 250 bp/dimer.
Figure 4.2
4.2A ATPase activity of topoIIα-(1-453) as a function of enzyme concentration.
The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at varying enzyme concentrations under standard conditions in the absence of DNA. The line shown is a theoretical curve based on the scheme and equation given in Ali et al. (1993).
Figure 4.2B Schematic model for the ATPase cycle of the *E. coli* DNA gyrase 43 kDa protein.

In this scheme (top), ATP is shown binding to the monomeric enzyme (orange), which induces a conformational change that allows the enzyme to dimerise. The active dimer complex hydrolyses ATP, which leads to product release and dissociation into the monomeric form. A steady-state rate equation based on this scheme was derived and is shown (bottom).

(Redrawn from Ali *et al.*, 1993).
Figure 4.2 (cont)

4.2C ATPase activity of topoIIα-(1-453) as a function of enzyme concentration in the presence of DNA.
The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at varying enzyme concentrations under standard conditions in the presence of DNA (relaxed pBR322) at 50 bp/dimer and 250 bp/dimer.
The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at 250 nM with varying substrate (ATP) concentrations under standard conditions in the absence of DNA. The total Mg$^{2+}$ concentration was maintained at 1 mM greater than the ATP concentration in the assay. The data were fit to the Michaelis-Menten equation by non-linear regression. Shown in the table are the derived values of $K_M$ (mM) and $k_{cat}$ (nM/second).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vmax (nM/s)</td>
<td>9.0469</td>
<td>0.3327</td>
</tr>
<tr>
<td>Km (mM)</td>
<td>0.6867</td>
<td>0.0637</td>
</tr>
</tbody>
</table>

**Figure 4.3**
4.3A ATPase activity of topoIIα-(1-453) as a function of ATP concentration. The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at 250 nM with varying substrate (ATP) concentrations under standard conditions in the absence of DNA. The total Mg$^{2+}$ concentration was maintained at 1 mM greater than the ATP concentration in the assay. The data were fit to the Michaelis-Menten equation by non-linear regression. Shown in the table are the derived values of $K_M$ (mM) and $k_{cat}$ (nM/second).
Figure 4.3 (cont)

4.3B ATPase activity of topoIIα-(1-453) as a function of ATP concentration in the presence of DNA.

The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at 250 nM with varying substrate (ATP) concentrations under standard conditions in the presence of DNA (relaxed pBR322) at 50 bp/dimer and 250 bp/dimer. The total Mg$^{2+}$ concentration was maintained at 1 mM greater than the ATP concentration in the assay. The data were fit to the Michaelis-Menten equation by non-linear regression. Shown in the table are the derived values of $K_M$ (mM) and $k_{cat}$ (nM/second).
Figure 4.4 ATPase activity of topoIIα-(1-420) as a function of ATP concentration in the presence and absence of DNA.

The ATPase activity of the N-terminal fragment topoIIα-(1-420) was determined at 250 nM with varying substrate (ATP) concentrations under standard conditions in the absence of DNA and in the presence of DNA (relaxed pBR322) at 250 bp/dimer. The total Mg$^{2+}$ concentration was maintained at 1 mM greater than the ATP concentration in the assay. The data were fit to the Michaelis-Menten equation by non-linear regression. Shown in the table are the derived values of $K_M$ (mM) and $k_{cat}$ (nM/second).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Std. Error</th>
<th>Value</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vmax (nM/s)</td>
<td>1.5396</td>
<td>0.0629</td>
<td>8.6722</td>
<td>0.3349</td>
</tr>
<tr>
<td>Km (mM)</td>
<td>0.2785</td>
<td>0.0380</td>
<td>0.3455</td>
<td>0.0431</td>
</tr>
</tbody>
</table>
are similar to those determined for the full-length enzyme: 0.055 s\(^{-1}\)/dimer in the absence of DNA and 0.59 s\(^{-1}\)/dimer in the presence of excess DNA (Hammonds & Maxwell, 1997). The \(K_m\) of the full-length enzyme in the presence of excess DNA is 0.56 mM (Hammonds & Maxwell, 1997), which compares to 0.89 mM seen here with the fragment topoII\(\alpha\)-(1-453) and 250 bp/dimer DNA.

The ATPase rate of the shorter fragment topoII\(\alpha\)-(1-420) was also determined at varying ATP concentrations in the absence of DNA and with DNA at 250 bp/dimer (Figure 4.4). The derived values of \(k_{cat}\) and \(K_M\) are: no DNA, 0.012 s\(^{-1}\)/dimer and 0.28 mM, respectively; 250 bp/dimer, 0.069 s\(^{-1}\)/dimer and 0.35 mM, respectively. Therefore, the presence of 250 bp/dimer DNA causes the \(k_{cat}\) to increase approximately 6-fold whilst the \(K_M\) is largely unaffected. This compares with derived values for the 52 kDa protein of 0.018 s\(^{-1}\)/dimer and 0.47 mM in the absence of DNA and 0.11 s\(^{-1}\)/dimer and 0.40 mM in the presence of excess DNA (Gardiner et al., 1998). Overall, it can be seen that the fragment topoII\(\alpha\)-(1-453) displays similar catalytic properties, in terms of ATP turnover, to the full-length enzyme, whereas the shorter fragment topoII\(\alpha\)-(1-420) displays similar catalytic properties to the 52 kDa protein.

4.2.2 Time-dependent lag of the hyperstimulated ATPase activity

An interesting phenomenon that occurs during the course of the ATPase reaction of topoII\(\alpha\)-(1-453) with DNA at 50 bp/dimer is shown in Figure 4.5. In the ATPase assays described here, the initial rate of ATP hydrolysis is coupled stoichiometrically to the oxidation of NADH, which is measured as a decrease in absorbance at 340 nm. In Figure 4.5, the decrease in absorbance at 340 nm for a standard ATPase assay is shown for the intrinsic and DNA-stimulated activity. For the intrinsic and 250 bp/dimer samples, the absorbance at 340 nm, corresponding to the concentration of NADH, decreases linearly with time, indicating that ATP hydrolysis is occurring at a constant rate. However, for the sample at 50 bp/dimer, the reduction in the absorbance at 340 nm is not linear during the first 10-15 minutes of the assay, indicating that there is a gradual increase in the rate of ATP hydrolysis before a constant (linear) rate is reached. That this lag occurs only for the hyperstimulated sample and
Figure 4.5 Demonstration of the lag in the rate of ATP hydrolysis by topoIIα-(1-453) in the presence of DNA.
Shown is the ATPase trace (plot of change in absorbance versus time) for a typical ATPase assay with protein at 250 nM and DNA (relaxed pBR322) either absent or present at 50 bp/dimer or 250 bp/dimer. The assay was performed under standard conditions except that before the reaction was started by adding the enzyme to the sample, both the sample and a small aliquot of enzyme were placed in a waterbath at 37°C for 15 minutes to allow their temperatures to reach the assay temperature (i.e. to minimise the effect of temperature equilibration in the early stages of the reaction). Using 10 minutes of continuous data a straight line fit was obtained for each sample to give the maximal rate in the time period shown. The ATPase rates for the samples are: no DNA, 5.327 nM/second; 50 bp/dimer, 60.300 nM/second; 250 bp/dimer, 31.517 nM/second.
Figure 4.6 Effect of preincubation conditions on the lag in ATP hydrolysis by topoIIα-(1-453) in the presence of DNA.
Shown is the ATPase trace (plot of change in absorbance versus time) for a typical ATPase assay with protein at 250 nM and DNA (relaxed pBR322) present at 50 bp/dimer. The assay was performed under standard conditions except that the samples and missing reagents were incubated at 37°C for 15 minutes before starting the reaction (as in Figure 4.5). Using 10 minutes of continuous data a straight line fit was obtained for each sample to give the maximal rate in the time period shown. The ATPase rates for the samples are: started with enzyme, 61.082 nM/second; started with DNA, 64.468 nM/second; started with ATP, 64.154 nM/second.
only during the initial phase of the reaction suggests that it is a result of an equilibration step in the interaction of the enzyme and the DNA under the conditions of hyperstimulation.

To further explore the lag in ATPase rate that occurs under conditions of hyperstimulation, the effect of preincubating different combinations of the three main reactants \textit{i.e.} enzyme, DNA and ATP, before adding the missing reactant to start the reaction, was investigated. For example, in one of the samples a reaction mixture was prepared containing ATP and DNA (at 50 bp/dimer), but no enzyme. A separate aliquot of enzyme, and the reaction mixture, were incubated at 37°C for 15 minutes, and then the enzyme was added to the reaction mixture to start the reaction and the absorbance was monitored at 340 nm as normal. This was repeated with ATP and DNA as the missing reagent used to start the reaction, and the results are shown in Figure 4.6. For all three samples there is still a considerable lag in the rate of ATP hydrolysis that occurs during the initial phase of the reaction before a constant ATPase rate is reached. However, the extent and duration of the lag differs for the three samples: the lag is greatest for the sample started with enzyme and least for the sample started with ATP. Therefore, preincubation of the enzyme with the DNA before the ATP is added to start the reaction reduces the extent of the lag, which suggests that this lag in ATPase rate is due to an equilibration step between the enzyme and the DNA. The actual ATPase rates of the three samples are given in the legend to Figure 4.6; it can be seen that there is little difference in the final ATPase rate of the three samples after a constant rate of ATP hydrolysis is achieved.

\textbf{4.2.3 The DNA-length dependence of the ATPase activity}

In Chapter 3, the ATPase activity of the N-terminal fragment topol-α-(1-453) was seen to be dependent upon the DNA concentration in the assay \textit{i.e.} the bp/dimer ratio. In the following sections, the dependence of the ATPase activity on DNA length is investigated, either using different bp/dimer ratios with relaxed pBR322, or using short DNA fragments of various lengths from 8 bp to 140 bp.
4.2.3.1 ATPase activity as a function of pBR322 concentration

In Figure 4.7, the ATPase activity of topoIIα-(1-453) is shown as a function of pBR322 DNA concentration, expressed as the bp/dimer ratio. The ATPase activity is stimulated by DNA up to a maximum, which occurs at ~50 bp/dimer, whereupon the ATPase rate decreases as the DNA concentration increases, reaching a plateau at ~150 bp/dimer. Above a DNA concentration of 150 bp/dimer the ATPase rate remains constant. This pattern of ATPase rate versus bp/dimer ratio is remarkably similar to that seen with the full-length enzyme (Hammonds & Maxwell, 1997), albeit with the peak and plateau levels being at different bp/dimer ratios than those seen with the full-length enzyme.

The ATPase activity of the shorter fragment topoIIα-(1-420) was also determined as a function of pBR322 concentration (Figure 4.8). This fragment, like the 52 kDa protein, does not show any hyperstimulation of the ATPase rate at low DNA:enzyme ratios, and requires a large excess of DNA to maximally stimulate the ATPase activity. It is clear from the above data that the presence of residues 421-453 is not only having a significant effect on the intrinsic ATPase activity of the N-terminal fragment of human topoisomerase IIα, but also causes the hyperstimulated ATPase rate which peaks at a bp/dimer ratio of 50 bp/dimer using pBR322.

4.2.3.2 ATPase activity as a function of DNA oligonucleotide length

To further investigate the DNA length-dependence of the DNA-stimulated ATPase activity of topoIIα-(1-453), DNA fragments of various lengths from 20 to 140 bp were used in ATPase assays. DNA fragments of 20-70 bp were prepared by annealing complementary oligonucleotides (PNACL); a 140 bp fragment was produced by PCR. All DNA oligonucleotides were based symmetrically around the preferred DNA gyrase binding and cleavage site within pBR322, the ‘990’ site (see Materials and Methods). Experiments were performed with a range of DNA fragment concentrations, from 2 enzyme dimers/DNA fragment (i.e. excess enzyme), up to 4 DNA fragments/dimer (i.e. excess DNA); the results are shown in Figure 4.9. In the case of the 20 and 30 bp fragments, increasing the amount of DNA present caused an increase
Figure 4.7 ATPase activity of topoIIα-(1-453) as a function of DNA concentration. The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at 250 nM under standard conditions with varying concentrations of DNA (relaxed pBR322), expressed as bp/dimer ratio.
Figure 4.8 ATPase activity of topoIIα-(1-420) as a function of DNA concentration. The ATPase activity of the N-terminal fragment topoIIα-(1-420) was determined at 250 nM under standard conditions with varying concentrations of DNA (relaxed pBR322), expressed as bp/dimer ratio.
Figure 4.9 ATPase activity of topoIIα-(1-453) as a function of DNA fragment concentration.
The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at 250 nM under standard conditions with varying concentrations of DNA fragments as indicated above, expressed as number of DNA molecules/dimer.
Figure 4.10 ATPase activity of topoIIα-(1-453) as a function of the concentration of a 140 bp DNA fragment.

The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at 250 nM under standard conditions with varying concentrations of a 140 bp DNA fragment, expressed as bp/dimer ratio. Also shown is the data from Figure 4.7.
in the rate of ATP hydrolysis. However, for fragments of 40 bp and longer, increasing the concentration of DNA caused the ATPase rate to decrease. Therefore, the hyperstimulation of the ATPase activity would appear to require a DNA fragment of \( \sim 40 \) bp or greater. These results are discussed in more detail in the Discussion section of this chapter.

In Figure 4.10, the ATPase activity of topoII\(\alpha\)-(1-453) in the presence of different concentrations of the 140 bp fragment, expressed as bp/dimer ratio, is shown, alongside the data for pBR322 (Figure 4.7). The ATPase rate rises sharply with increasing concentrations of the 140 bp fragment, up to a maximum at a bp/dimer ratio of \( \sim 40 \) bp/dimer, at which point the ATPase rate decreases rapidly with increasing DNA concentration, and reaches a plateau level at \( \sim 200 \) bp/dimer. The initial sharp rise and then fall in the ATPase rate is similar to that seen in the presence of pBR322 at low bp/dimer ratios. However, high concentrations of the 140 bp DNA fragment cause the ATPase rate to fall to a level which is significantly lower than that seen with high concentrations of pBR322, and represents only an \( \sim 75\% \) stimulation over the intrinsic ATPase rate.

4.2.3.2.1 DNA length required to stimulate the ATPase activity

To determine the minimum length of DNA required to stimulate the ATPase activity of the fragment topoII\(\alpha\)-(1-453), ATPase assays were performed with DNA fragments of 8 bp to 24 bp. These experiments were conducted at 37°C and also 25°C, because the shorter fragments are likely to 'melt' at 37°C. At 37°C (Figure 4.11), addition of DNA fragments of 12-24 bp stimulate the ATPase rate, although the 12 bp and 14 bp DNA fragments are required in a large excess over the enzyme in order to maximally stimulate the ATPase rate. The 10 bp DNA fragment, however, causes only a slight increase in the ATPase rate in the presence of excess DNA, since this fragment is likely to be almost completely dissociated into ssDNA at this assay temperature. At 25°C, the longer DNA fragments give the same pattern of stimulation of the ATPase rate as at 37°C, although the actual rates are reduced because of the lower temperature. The 10 bp fragment stimulates the ATPase rate to a greater extent than was seen at
Figure 4.11 ATPase activity of topoIIα-(1-453) as a function of short DNA fragment concentration.
The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at 250 nM under standard conditions (37°C) with varying concentrations of DNA fragments (10-24 bp) as indicated above, expressed as number of DNA molecules/dimer.
Figure 4.12 ATPase activity of topoIIα-(1-453) at 25°C as a function of short DNA fragment concentration.

The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at 250 nM under standard conditions, except that the temperature was reduced to 25°C, with varying concentrations of DNA fragments (8-24 bp) as indicated above, expressed as number of DNA molecules/dimer.
37°C, presumably because more of the DNA is in the double-stranded form at the lower temperature. An 8 bp fragment causes a slight stimulation of the ATPase activity at 25°C, similar to that seen with the 10 bp fragment at 37°C. Although it is not possible to say for certain whether the lack of significant stimulation of the ATPase rate with the 8 bp fragment is due to this DNA fragment being largely in the single-stranded form in the assay, given the results seen with the 10 bp fragment it seems likely that a DNA fragment of 8-10 bp is sufficient to stimulate the ATPase activity of topoIIα-(1-453). However, it is possible that a longer DNA fragment i.e. 12-14 bp, is required to maximally stimulate the ATPase activity of this N-terminal fragment.

4.2.3.2.2 DNA length required to hyperstimulate the ATPase activity

In Figure 4.9 we saw how increasing the concentration of a 30 bp DNA fragment caused the DNA-dependent ATPase rate to increase, whereas increasing the concentration of a 40 bp DNA fragment caused the DNA-dependent ATPase rate to decrease. This would suggest that the hyperstimulation phenomenon requires a DNA fragment with a minimum length of somewhere between 30 and 40 bp. To determine the exact length of DNA required to cause hyperstimulation of the ATPase activity, DNA fragments of 28 bp to 44 bp were used in ATPase assays (Figure 4.13A, and the same data are shown on two separate graphs in Figure 4.13B). Overall, the hyperstimulation effect is absent with a DNA fragment of 28 bp, peaks with a DNA fragment of 38 bp, and is diminished, but not abolished altogether, with a DNA fragment of 44 bp. Furthermore, at high concentrations of the longer DNA fragments (≥32 bp), the ATPase rate drops significantly to a level similar to that seen with high concentrations of a 140 bp DNA fragment (Figure 4.10). This effect was not seen with shorter DNA fragments of less than 30 bp (Figures 4.9 and 4.11), although with these DNA fragments the ATPase rate did not reach the same level as that seen with high concentrations of pBR322 (Figure 4.7). Clearly, the length of the DNA fragment determines whether or not the hyperstimulation phenomenon occurs, but also the plateau level of DNA-stimulated ATPase activity.
ATPase activity of topoIIα-(1-453) as a function of long DNA fragment concentration.

The ATPase activity of the N-terminal fragment topoIIα-(1-453) was determined at 250 nM under standard conditions with varying concentrations of DNA fragments (28-44 bp) as indicated above, expressed as number of DNA molecules/dimer.
Figure 4.13 (cont)
4.13B ATPase activity of topolα-(1-453) as a function of long DNA fragment concentration. The same data from Figure 4.13A are presented on two separate graphs.
4.3 DISCUSSION

In this chapter we have investigated the kinetics and DNA stimulation of the ATPase activity of the N-terminal fragment topoIIα-(1-453). In the previous chapter we saw that the intrinsic ATPase activity is stimulated approximately 10-fold by DNA (relaxed pBR322) at 50 bp/dimer, but only approximately 5-fold by DNA at 250 bp/dimer (Figure 3.18), and in this chapter we have seen that this 'hyperstimulated' ATPase rate does indeed reach its maximum at 50 bp/dimer, and the ATPase rate at 250 bp/dimer represents the 'basal' DNA-stimulated ATPase rate (Figure 4.7). Therefore, these two concentrations of DNA were used in ATPase assays to investigate the kinetics of ATP hydrolysis by the N-terminal fragment topoIIα-(1-453), along with the intrinsic ATPase activity.

The ATPase rate at 50 bp/dimer is more affected by higher concentrations of salt than the intrinsic ATPase rate or the rate in the presence of 250 bp/dimer DNA. This suggests that the interaction between the enzyme and DNA that results in the hyperstimulation phenomenon is weaker and therefore particularly sensitive to salt concentration; alternatively, it may be the dimerisation step into the ATP-hydrolysing active form that is most affected in the presence of DNA at 50 bp/dimer. The intrinsic ATPase activity displays a parabolic dependence upon enzyme concentration (Figure 4.2A). The curve in Figure 4.2A was obtained by fitting the data to a steady-state rate equation derived from the mechanism proposed for the ATPase activity of the 43 kDa N-terminal fragment of *E. coli* DNA gyrase (Ali *et al.*, 1993). In this reaction mechanism, ATP binding to the monomeric enzyme induces a conformational change that allows the enzyme to dimerise; it is the dimer complex which hydrolyses ATP, leading to product release and dissociation into the monomeric form. According to this scheme (Figure 4.2B), it is the dimerisation of the enzyme-ATP monomer complex that represents the rate-limiting step of the reaction, and this accounts for the greater than first order dependence of the ATPase rate upon enzyme concentration. The fact that the DNA-dependent ATPase activity is linearly dependent upon enzyme concentration is indicative of a scheme whereby the monomeric form of the enzyme, which may predominate in the absence of DNA, is induced to dimerise in the presence of DNA, forming the ATPase active sites. However, it is also possible that, in the presence of
DNA, dimerisation is not a prerequisite for ATP hydrolysis, or dimerisation is not the rate-limiting step in ATP hydrolysis. Previous results for the full-length enzyme (Hammonds & Maxwell, 1997), and for the 52 kDa N-terminal fragment (Gardiner et al., 1998), indicated that the ATPase rate is linearly dependent upon protein concentration both in the presence and absence of DNA. The difference seen in the relationship between intrinsic ATPase rate and protein concentration for the two N-terminal fragments is likely to be due to the different C-terminus of these protein fragments i.e. the 52 kDa protein represents only residues 1-440, but more importantly, also contains 11 foreign amino acids and a hexa-histidine tag which may affect the multimeric state of the protein. In the full-length enzyme, the N-terminal domains are tethered to the rest of the topoisomerase molecule, and although they are in the 'open clamp' conformation in the absence of ATP, the existence of the topoisomerase molecule as a dimer leads to a linear dependence of the intrinsic ATPase rate upon enzyme concentration. The oligomeric properties of the N-terminal fragment topoIIα-(1-453) are further examined in Chapter 5.

At a constant enzyme concentration the rate of ATP hydrolysis was hyperbolically dependent upon ATP concentration. Under conditions where the Mg$^{2+}$ concentration was maintained at 1 mM greater than the concentration of ATP, i.e. non-standard conditions, the intrinsic ATPase rate was found to be stimulated ~15-fold by DNA at 50 bp/dimer and ~10-fold by DNA at 250 bp/dimer. The shorter fragment topoIIα-(1-420), which does not show any hyperstimulation of the ATPase activity, has a lower intrinsic ATPase activity which is stimulated ~6-fold by excess DNA. Although direct comparison of the results for the full-length enzyme (Hammonds & Maxwell, 1997) and the 52 kDa fragment (Gardiner et al., 1998) with the results shown here are complicated by the non-linear dependence of the intrinsic ATPase rate upon enzyme concentration, what is clear is that the longer N-terminal fragment topoIIα-(1-453) is distinct from the shorter N-terminal fragments (topoIIα-(1-420) and the 52 kDa protein) in terms of the level of intrinsic and DNA-dependent ATP hydrolysis and its affinity for substrate, and does in fact display similar catalytic properties to the full-length enzyme.
A significant finding of this study is that the ATPase activity of the N-terminal fragment topoIIα-(1-453) displays a similar hyperstimulation effect at high enzyme:DNA ratios to that seen with the full-length human protein (Hammonds & Maxwell, 1997). This hyperstimulation effect is not seen with the shorter N-terminal fragment topoIIα-(1-420). With the full-length enzyme, the hyperstimulated ATPase rate was seen at ~150 bp/enzyme dimer with relaxed pBR322, whereas with the fragment topoIIα-(1-453) the hyperstimulated ATPase rate peaks at ~50 bp/dimer (Figure 4.7), which is consistent with the smaller size of the N-terminal fragment. In both cases, when the DNA is present in excess, the ATPase rate plateaus at a 'basal' level that is approximately half that of the hyperstimulated ATPase rate. The hyperstimulation effect was originally attributed to one of two possibilities: firstly, the saturation of DNA with enzyme molecules could lead to a molecular crowding effect where neighbour enzyme molecules could interact with one another to cause a hyperstimulated ATPase rate, possibly by the formation of tetramers; or secondly, enzyme molecules adjacent to one another on a DNA strand may become trapped in a configuration where they have both the G and T segments captured (the 'G&T' complex) e.g. if they are both trying to pass the same T segment in opposite directions, leading to a permanently stimulated ATPase rate (Hammonds & Maxwell, 1997). Under conditions where the DNA is in excess, the molecular crowding effect giving rise to either of the above situations is absent, resulting in the basal level of DNA stimulation. The fact that this hyperstimulation effect is also seen with the N-terminal fragment topoIIα-(1-453) suggests that this phenomenon is not due to stabilisation of the G&T complex, since the fragment topoIIα-(1-453) represents only the ATP-operated clamp and presumably is not able to bind a G segment. Furthermore, the residues 421-453 appear to be essential for the occurrence of the hyperstimulation effect and are presumably involved in either protein-protein or protein-DNA interactions. Interaction of the enzyme with DNA at 50 bp/dimer also produced a lag in the ATPase rate; no such lag was seen with DNA at 250 bp/dimer (Figure 4.5). The extent of this lag could be reduced by pre-incubating the enzyme with DNA before the reaction was started. Therefore, the lag in ATPase rate occurs specifically under conditions of hyperstimulation, and is presumably a result of an equilibration step in the binding of enzyme to DNA.
Based on the results discussed above and the previous explanation for the hyperstimulation phenomenon (Hammonds & Maxwell, 1997), the scheme shown in Figure 4.14 may be used as a preliminary hypothetical model to account for the existence of the hyperstimulated ATPase rate. In this model, the ATP-hydrolysing topoIIα-(1-453) dimers are shown interacting with one another under conditions where the enzyme is saturated with DNA i.e. 50 bp/dimer. Exactly how the enzyme dimers may interact with one another or with DNA to cause the hyperstimulation effect is unclear, but it is likely that the region encompassing residues 421-453 is required for this interaction. Furthermore, given the role of the N-terminal domains as a DNA clamp, the most likely position of the DNA is within the protein clamp, but it is also possible that DNA binds to the outside of the clamp or wraps around the clamp (this may or may not be artefactual due to the use of a truncated enzyme fragment). At 250 bp/dimer the enzyme dimers are sufficiently spaced out along the DNA strand for the dimer-dimer interactions not to occur, and as the DNA concentration increases further the ATPase rate remains at a basal DNA-stimulated level. Olland & Wang (1999) proposed that the apparent hyperstimulation effect seen with a 76 kDa fragment of yeast topoisomerase II, Top2-(1-660), which is equivalent to the *E. coli* DNA gyrase GyrB protein, could be caused by the occupation of slightly weaker DNA binding sites at higher DNA concentrations, leading to a reduction in the ATPase activity. However, the results shown here suggest that it is the interaction of the enzyme with DNA at 50 bp/dimer that is the weakest, as evidenced by the response to increasing salt concentration. An alternative proposal from Olland & Wang (1999) was that at higher DNA concentrations the monomer form of the enzyme may bind to DNA, leading to a large separation of monomers bound to DNA which are therefore unlikely to dimerise. However, this does not account for the hyperstimulation effect seen with the full-length enzyme, which is a permanent dimer.

At first glance the hyperstimulation effect seen with the N-terminal fragment topoIIα-(1-453) appears similar to that seen with full-length enzyme, but in actual fact it is more complicated. Using short linear dsDNA fragments it was shown that a linear DNA fragment of 8-10 bp is sufficient to stimulate the ATPase activity of the fragment topoIIα-(1-453) (Figures 4.11 and 4.12). This is consistent with the crystal structure of
Figure 4.14 Schematic model for the DNA-dependent ATPase activity of the N-terminal fragment topoIIα-(1-453).

Enzyme molecules (orange) are shown in the dimeric form. In the absence of DNA (top) the fragment topoIIα-(1-453) has an intrinsic ATPase activity. With relaxed pBR322 at 50 bp/dimer (middle) the DNA (black) is saturated with enzyme dimers which may interact with one another to give a hyperstimulated ATPase rate. The dimers are shown in the same orientation but their exact orientation is unknown, although residues 421-453 appear to be important. With excess DNA (bottom) interactions between enzyme dimers occur with much lower probability and a basal DNA-stimulated ATPase rate is seen.
the 43 kDa fragment of *E. coli* DNA gyrase (Wigley *et al.*, 1991). If we assume that a 10 bp DNA fragment is sufficient to stimulate the ATPase activity, it is significant that a DNA fragment:enzyme dimer ratio of five times this value is required to maximally hyperstimulate the ATPase activity. If the enzyme dimer occupies ~10 bp, this implies that the distance between enzyme dimers is ~40 bp, which would make protein-protein interactions such as those depicted in Figure 4.14 (*middle*) unlikely. Using a linear DNA fragment of 140 bp, the initial rise and fall in ATPase rate as the base pair:enzyme dimer ratio increases is similar to that seen with relaxed pBR322, albeit peaking at a slightly lower 40 bp/dimer (Figure 4.10). However, at high base pair:enzyme dimer ratios (*i.e.* >200 bp/dimer) the ATPase rate is significantly lower than at high concentrations of pBR322. With linear pBR322, the results are similar to that seen with relaxed pBR322 (data not shown). If molecular interactions such as those shown in Figure 4.14 (*middle*) were responsible for the hyperstimulated ATPase rate seen with the 140 bp DNA fragment, then the hyperstimulated ATPase rate should be abolished when the ratio of DNA molecules:enzyme dimers exceeds 1:1 *i.e.* >140 bp/dimer (unless the enzyme dimers bind cooperatively to DNA). From Figure 4.10 this can be seen to be the case, but this does not explain why the basal DNA-stimulated ATPase rate is only ~50% higher than the intrinsic ATPase rate and significantly lower than that seen with excess pBR322.

In Figure 4.9 we saw that increasing the concentration of short DNA fragments (<30 bp) caused the ATPase rate to increase, whereas increasing the concentration of longer DNA fragments (>40 bp) had the opposite effect, causing the ATPase rate to decrease. This result suggests that a minimum of ~40 bp is required to hyperstimulate the ATPase activity, and raising the bp/dimer ratio (*i.e.* increasing the DNA concentration) abolishes the hyperstimulation effect, as seen in Figure 4.10 with both relaxed pBR322 and a 140 bp DNA fragment. DNA fragments of 28-44 bp showed that 28 bp does not cause any hyperstimulation of the ATPase rate, the hyperstimulation effect peaks with a DNA fragment of 38 bp, and subsides as the DNA length increases to 44 bp. These results are consistent with a scheme, based on Figure 4.14, whereby the molecular interactions resulting in the hyperstimulated ATPase rate are optimal with a DNA fragment of 38 bp. With a shorter DNA fragment, where the enzyme dimers are
closer together, or a longer DNA fragment, where the enzyme dimers are further apart, the molecular interactions giving rise to the hyperstimulated ATPase rate are reduced. With high concentrations of DNA fragments >30 bp the ATPase rate is again reduced to a level which represents only an ~50% increase over the intrinsic ATPase rate, as was seen with the 140 bp DNA fragment. Furthermore, the hyperstimulated ATPase rate with the 38 bp DNA fragment peaks at ~25 nM/s, whereas with the 140 bp DNA fragment and relaxed pBR322 the hyperstimulated ATPase rate peaks at ~90 nM/s at ~40-50 bp/dimer.

Overall, it is unclear as to the exact nature and cause of the hyperstimulation effect seen with the variety of DNA fragments. With relaxed pBR322, the hyperstimulation effect peaks at ~50 bp/dimer, and there is a pronounced lag in the ATPase rate during the first 10-15 minutes of the reaction before the maximum rate of ATP hydrolysis is reached. With a 140 bp DNA fragment, the hyperstimulation effect peaks at ~40 bp/dimer, and plateaus at a very low level. With shorter DNA fragments, the hyperstimulation effect is most pronounced with a DNA fragment of 38 bp. If we assume that the hyperstimulated effect is caused by the molecular interactions indicated in Figure 4.14, this would suggest that a 38 bp DNA fragment could accommodate two enzyme dimers that can interact with one another. This corresponds to ~20 bp/dimer, which is less than half that required with relaxed pBR322 to give maximal hyperstimulation. With the full-length enzyme, maximum hyperstimulation was seen with relaxed pBR322 at ~150 bp/dimer, or a linear DNA fragment of ~300 bp, or assuming the above to be true, ~150 bp/dimer (Hammonds & Maxwell, 1997) i.e. the DNA length:enzyme dimer ratios giving maximum hyperstimulation are identical. In Figure 4.9, we saw that as the length of the DNA fragment increases the maximum hyperstimulated ATPase rate increases, which is consistent with more enzyme dimers binding to the DNA and interacting with one another.

Any model that is to explain these results must be able to account for two apparent anomalies. Firstly, with relaxed pBR322, the hyperstimulation peaks at ~50 bp/dimer, whereas with short DNA molecules and assuming the model in Figure 4.14 to be true, the hyperstimulation peaks at ~20 bp/dimer. Secondly, the plateau ATPase
Figure 4.15 Alternative model for the hyperstimulated ATPase activity of the N-terminal fragment topoIIα-(1-453).

Shown is an alternative hypothetical model, in which interaction of two enzyme dimers, bound to a short DNA fragment, in the N-terminal region of the ATPase domains could give rise to a hyperstimulated ATPase rate.
level with excess DNA fragments of 140 bp or less is significantly lower than with excess pBR322. In the model shown in Figure 4.14, the enzyme dimers are depicted to interact with an adjacent enzyme dimer on the same DNA molecule, in a ‘side-to-side’ manner, to give the hyperstimulated ATPase rate. It is possible that enzyme dimers on different DNA molecules can interact with one another, perhaps in a different orientation than that suggested in Figure 4.14 e.g. in an end-on manner (Figure 4.15). Although residues 421-453 are known to be important for the hyperstimulated ATPase rate, the interaction has been depicted as occurring at the N-terminal arm region of the ATPase domains and not the C-terminal region, because such an interaction would be unlikely to occur in the full-length enzyme. However, if the hyperstimulation effect was due to DNA-bound enzyme dimers interacting in the manner depicted in Figure 4.15, one might expect the hyperstimulation phenomenon to occur with any DNA fragment greater than ~10 bp, which is not the case. Another possibility is that the monomer form of the N-terminal fragment may interact with DNA, as postulated by Olland and Wang (1999). According to this hypothesis, the ATPase rate decreases as the DNA concentration increases because of the increasing separation of enzyme monomers bound to DNA i.e. there is an inhibition of ATPase rate at high DNA concentrations, rather than a stimulation at low DNA concentrations.

In summary, the ATPase activity of the N-terminal fragment topoIIα-(1-453) has been shown to be stimulated by a DNA fragment of ~10 bp in length. Using longer DNA molecules, this fragment gave a similar pattern of hyperstimulation at low DNA:enzyme concentrations as that observed with the full-length enzyme. The absence of this effect with the shorter fragment topoIIα-(1-420) suggests an involvement of the C-terminal part of this protein i.e. residues 421-453. The degree of hyperstimulation depends both on the length of the DNA fragment (or the bp/dimer ratio), and the nature of the DNA molecule i.e. linear or circular. We have proposed, based on the original work of Hammonds and Maxwell (1997), that this hyperstimulation effect is due to protein-protein interactions between enzyme dimers. However, the nature of the interaction and the physiological significance of the increased ATPase rate is not clear at the present time.
CHAPTER 5

Interaction of the N-terminal domain with DNA and nucleotides: an ATP-operated DNA-binding clamp
5.1 INTRODUCTION

The catalytic cycle of the type II topoisomerases requires that the N-terminal domains of the enzyme function as a 'molecular clamp'. The clamp must capture the T segment, which is accomplished by ATP binding and a conformational shift leading to dimerisation. The clamp must then aid in the passage of the T segment through the G segment, and the DNA transport event may be accelerated by the hydrolysis of bound ATP (Baird et al., 1999). The clamp must then be reopened following strand passage, and presumably this occurs after the release of ADP and P_i and a resultant conformational change.

In light of the above, the conformation of the N-terminal domain of type II topoisomerases can be predicted to be sensitive to the presence of ATP (and therefore ADPNP) and also ADP±P_i, and also to bind or interact with DNA. In the case of the 43 kDa protein of *E. coli* DNA gyrase, this N-terminal fragment was shown to dimerise in the presence of ADPNP, but was found not to bind to DNA, nor did DNA stimulate the ATPase activity of the 43 kDa protein (Ali et al., 1993). In contrast, the N-terminal fragment of human DNA topoisomerase IIα was previously shown to be predominantly a dimer in the presence and absence of ADPNP, and to have an intrinsic ATPase activity stimulated by DNA (Gardiner et al., 1998). In this chapter, the fragment topoIIα-(1-453) has been investigated for its DNA binding properties and the effect of nucleotides on the oligomeric state of the protein, and limited proteolysis has been used to identify conformational changes resulting from interaction with nucleotides and DNA.

5.2 RESULTS

5.2.1 Molecular weight studies

The oligomeric state of the protein fragment topoIIα-(1-453) has been investigated using three techniques: protein cross-linking, gel filtration and analytical ultracentrifugation. These techniques are the most widely used to determine the molecular weight and oligomeric state of a protein, and in the order given above, they
offer increasing accuracy as to the molecular weight of a protein in solution in the presence and absence of ligands.

5.2.1.1 Cross-linking

Three protein cross-linking agents were used to try to resolve the oligomeric state of the protein topoIIα-(1-453) in the presence and absence of nucleotides and DNA. These cross-linking agents: dimethyl suberimidate (DMS, an imidoester cross-linker), disuccinimidyl glutarate (DSG, an NHS-ester cross-linker) and bis(sulfosuccinimidyl) suberate (BS³, an NHS-ester cross-linker), react with primary amine groups (most notably the ε-amine of lysine) to form stable covalent bonds. The protein was first incubated with nucleotide or DNA for 1 hour at 25°C; cross-linker was then added and the samples incubated at 25°C for a further 16 hours. An aliquot of the samples was then analysed by SDS-PAGE electrophoresis on a 10% gel (the results are shown in Figures 5.1, 5.2 and 5.3; the predicted molecular weight of the protein topoIIα-(1-453) is 51,458 Da). With all three cross-linking agents, no high molecular weight cross-linked species were seen in the absence of nucleotide, whereas cross-linked species were seen in the presence of ADPNP. With BS³ and DSG there was a single cross-linked species at ~150 kDa, but with DMS several cross-linked bands were seen ranging from ~100 kDa to ~160 kDa. The existence of several cross-linked products may represent different multimeric species e.g. dimers and trimers/tetramers, or may be due to cross-linking at different sites within the same multimeric species. Several cross-linked species were also seen with the 43 kDa protein of *E. coli* DNA gyrase in the presence of ADPNP (Ali *et al.*, 1995); this was attributed to cross-linking at multiple sites within the dimeric species. However, the range of bands seen with the 43 kDa protein was much narrower than that seen here with the fragment topoIIα-(1-453). With ATP and ADP+P; cross-linked bands with similar electrophoretic mobility to that seen with ADPNP were seen with BS³, and to a lesser extent with DMS. Based on the electrophoretic mobility of the cross-linked bands it is not possible to say for certain whether they represent dimers (the most likely explanation) or a mixture of oligomeric species.
Protein (7.5 µM) was incubated in the presence or absence of nucleotide (2 mM) or relaxed pBR322 at 50 bp/dimer (as indicated) for 1 hour at 25°C in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 5 mM MgCl₂, 10% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA. Crosslinker (DMS) was then added to 2 mg/ml and the samples incubated at 25°C for a further 16 hours. Samples were then boiled for 5 minutes, and analysed by SDS-PAGE electrophoresis on a 10% gel. The sizes of the markers (LMW and HMW) are indicated (kDa).
Figure 5.2 Cross-linking of the N-terminal fragment topoIIα-(1-453) using bis(sulfosuccinimidyl) suberate (BS³).

Protein (7.5 µM) was incubated in the presence or absence or nucleotide (2 mM) or relaxed pBR322 at 50 bp/dimer (as indicated) for 1 hour at 25°C in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 5 mM MgCl₂, 10% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA. Crosslinker (BS³) was then added to 1 mg/ml and the samples incubated at 25°C for a further 16 hours. Samples were then boiled for 5 minutes, and analysed by SDS-PAGE electrophoresis on a 10% gel. The sizes of the markers (LMW and HMW) are indicated (kDa).
Figure 5.3 Cross-linking of the N-terminal fragment topoIIα-(1-453) using disuccinimidyl gluterate (DSG).

Protein (7.5 μM) was incubated in the presence or absence of nucleotide (2 mM) or relaxed pBR322 at 50 bp/dimer (as indicated) for 1 hour at 25°C in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 5 mM MgCl₂, 10% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA. Crosslinker (DSG) was then added to 0.5 mg/ml and the samples incubated at 25°C for a further 16 hours. Samples were then boiled for 5 minutes, and analysed by SDS-PAGE electrophoresis on a 10% gel. The sizes of the markers (LMW and HMW) are indicated (kDa).
No high molecular weight species were seen with any of the cross-linking agents in the presence of DNA (relaxed pBR322) at 50 bp/dimer, or at 250 bp/dimer (data not shown), which suggests that the interaction of the N-terminal fragment topoIIα-(1-453) with DNA does not cause the protein to dimerise. This result is interesting given that in the presence of DNA the ATPase activity is linearly dependent upon protein concentration, which could indicate the protein exists as a dimer (see Chapter 6 for further discussion).

Although these experiments have provided some information on the dimerisation properties of the N-terminal fragment topoIIα-(1-453), they are qualitative rather than quantitative, and although high molecular weight species are seen, the results are ambiguous as to the exact multimeric nature of the molecule.

5.2.1.2 Gel filtration

Gel filtration was used to determine the native molecular weight of the protein topoIIα-(1-453) alone or in complex with ADPNP. The protein was incubated with or without 2 mM ADPNP for one hour and then applied to a pre-equilibrated Sephacryl S-200 HR (16/60) FPLC column (Pharmacia) and eluted at a flow rate of 0.8 ml/min in the presence or absence or 0.1 mM ADPNP. The absorbance was monitored at 280 nm, and 1 ml fractions were collected and analysed by SDS-PAGE. The column was calibrated using a set of molecular weight markers and also Blue Dextran to determine the void volume of the column. The protein alone eluted as a single main peak (Figure 5.4); whereas, in the presence of ADPNP, two main peaks were observed that eluted earlier than the peak seen for protein alone, indicating an ADPNP-induced increase in molecular weight. The small peak which elutes at ~70 ml corresponds to minor contamination by the 17 kDa thioredoxin fusion tag carried over from the Factor Xa cleavage and purification reaction. Using the calibration curve (Figure 5.5) the molecular weight of the single peak for protein alone was estimated to be ~83 kDa, and ~136 kDa and ~231 kDa for the two peaks in the presence of ADPNP (the elution point for the 17 kDa thioredoxin tag corresponds to a molecular weight of ~26 kDa). Given that no high molecular weight species were seen in cross-linking studies without
Figure 5.4 Gel filtration of the N-terminal fragment topoIIα-(1-453) in the presence and absence of ADPNP.

Protein (0.4 mg/ml) in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 5 mM MgCl₂, 10% glycerol (w/v), 0.5 mM DTT, 0.5 mM EDTA was incubated in the presence or absence of 2 mM ADPNP for 1 hour at 25°C. A 500μl aliquot (~200 μg) was loaded onto a pre-equilibrated Sephacryl S-200 column (Pharmacia) and eluted with the above buffer in the presence or absence of 0.1 mM ADPNP. Fractions of 1 ml were collected, and the absorbance at 280 nm was continually monitored. The void volume of the column was found to be 37.3 ml.
Figure 5.5 Calibration curve for gel filtration.
The gel filtration column was calibrated using Sigma molecular weight gel filtration marker kit (Blue Dextran, 2,000,000 kDa; Amylase, 200,000 kDa; Alcohol Dehydrogenase, 150,000 kDa; Carbonic Anhydrase, 29,000 kDa; Cytochrome C, 12,400 kDa). The ratio of the elution volume (Ve) to the void volume (Vo) for each of the protein standards (Ve/Vo) was plotted against the log of molecular weight (white circles), and a straight line was fitted to the points. From the elution volume of the peaks for the protein samples the apparent molecular weights were calculated.
nucleotide (and also that the 17 kDa thioredoxin tag eluted at the higher molecular weight of ~26 kDa), it is likely that the single peak eluting at ~83 kDa corresponds to the monomeric species. The peak at ~136 kDa would therefore be consistent with the protein being a dimer, and the peak at ~231 kDa could correspond to a tetramer (or a higher order oligomeric species). However, the single peak eluting at ~83 kDa could also correspond to a monomer-dimer equilibrium.

5.2.1.3 Analytical ultracentrifugation

Samples were prepared for analytical ultracentrifugation by gel filtration in a similar manner to that described in the previous section. Fractions corresponding to the most prominent peak for each sample (i.e. the peak at ~83 kDa for protein alone and the peak at ~136 kDa for protein-ADPNP from Figure 5.4) were pooled, concentrated to ~0.1 mg/ml and used in analytical ultracentrifugation experiments (performed by A. Leech, University of East Anglia). The molecular weights that were obtained for the protein ± ADPNP at the two different speeds used in equilibrium sedimentation studies are shown in the table below. These studies show conclusively that, under the experimental conditions employed, the protein alone is a monomer whereas in the presence of ADPNP the protein forms a dimer. After 24 hours at 15,000 rpm the molecular weights of 52,204 ± 1460 Da for protein alone and 103,338 ± 2150 Da for protein-ADPNP are consistent with a scheme whereby ADPNP binding is required for the N-terminal fragment to dimerise. After a further 16 hours at a speed of 20,000 rpm the apparent molecular weight for protein alone is 40,137 ± 1380 Da, and for the protein-ADPNP complex is 100,622 ± 3230 Da. The values at 20,000 rpm (i.e. after greater time) indicate that there is some decomposition of the sample over time, particularly for the protein only sample.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Apparent molecular weight (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15,000 rpm</td>
</tr>
<tr>
<td>TopoIIα-(1-453) alone</td>
<td>52,204 ± 1460</td>
</tr>
<tr>
<td>TopoIIα-(1-453) + ADPNP</td>
<td>103,338 ± 2150</td>
</tr>
</tbody>
</table>
5.2.2 DNA binding studies

The nature of the interaction of the N-terminal domains of type II DNA topoisomerases with DNA is unknown, but some degree of interaction is predicted by the proposed model for strand passage by these enzymes (see Figure 1.10). Furthermore, the DNA binding characteristics of the N-terminal domains would be expected to be altered by the presence of nucleotide. The DNA binding properties of the fragment topoIIα-(1-453) were investigated using filter-binding and gel shift assays.

5.2.2.1 Filter-binding

The filter-binding assay is a simple and reproducible technique for detecting protein-DNA interactions (Stockley, 1994). Samples with up to 1 μM protein were pre-incubated in the presence and absence of relaxed ³H-pBR322 or 2 mM ADPNP. After 30 minutes, either DNA or ADPNP was added as above, or no addition was made, and the samples were incubated for a further 30 minutes. The sample was added to 200 μl of buffer and filtered through a pre-washed filter disc. The filter disc was then washed with buffer and after drying the amount of radioactivity retained on the filter was determined by scintillation counting. The results (Figure 5.6) are expressed as the percent of DNA retained on the filter. The interpretation of the results is complicated by the fact that multiple proteins can bind to the same DNA molecule. Protein alone binds DNA with a relatively low affinity, and DNA binding does not reach saturation at 1 μM protein (Figure 5.6, white squares). Where the protein is first incubated with DNA, and then ADPNP to dimerise the protein, the extent of DNA binding is increased and approaches saturation at 1 μM protein (Figure 5.6, grey circles). When the protein clamp is first locked with ADPNP, and then DNA is added (Figure 5.6, black triangles), the extent of DNA binding is similar to that seen when DNA is added before the ADPNP. This result suggests that the nature of the protein-DNA interaction is the same whether the DNA is added before or after the protein dimerises and the clamp is locked.
Figure 5.6 Filter binding assay for DNA binding by the N-terminal fragment topoIIα-(1-453).

Samples (50 μl) containing up to 1 μM protein in 50 mM Tris-HCl (pH 8.8), 20 mM NaCl, 5 mM MgCl₂, 10% (v/v) glycerol, 0.5 mM DTT, 0.5 mM EDTA, 0.1 mg/ml BSA were incubated at 25°C in the presence or absence of 3.62 nM relaxed ³H-pBR322 or 2 mM ADPNP. After 30 minutes, either DNA or ADPNP was added as above, or no addition was made, and the samples were incubated for a further 30 minutes at 25°C. The sample was added to 200 μl of Binding Buffer, and this was filtered through a pre-washed filter disc at a flow rate of approximately 1 ml/minute. The filter disc was washed with 2 x 500 μl of Binding Buffer, and was then left to air dry for approximately 30 minutes. 4 ml of scintillation liquid was added and the amount of tritiated DNA retained on the filter disc was calculated by scintillation counting.
5.2.2.2 Gel shift

To further examine the DNA-binding characteristics of the fragment topoIIα-(1-453), gel shift assays (Taylor et al., 1994) were performed using a 140 bp DNA fragment prepared by PCR. Protein up to 3 μM was incubated first with the 140 bp DNA fragment (50 nM), and then either with or without ADPNP at 2 mM. The sample was loaded directly onto a pre-run 5% TBM gel (DNA was visualised by staining with ethidium bromide). A typical result from these experiments is shown in Figure 5.7. The results were largely unsatisfactory due to problems associated with the methodology, but probably also due to the nature of the DNA-protein interaction. In the absence of ADPNP, the protein caused the DNA band to smear in the direction of a higher molecular weight species. This could be caused by the protein-DNA complex dissociating within the gel, which is likely given the apparent weak interaction between the protein and DNA as suggested by the filter-binding experiments. Furthermore, it is possible that more than one protein molecule or dimer is binding to each DNA molecule, which, along with the dissociation of complexes described above, could give rise to the smeared DNA band. In the presence of ADPNP, the increasing concentration of protein resulted in the DNA not being observed on the gel, although in some cases a smearing effect like that seen in the absence of ADPNP was observed. This suggests that the ADPNP-induced dimerisation of the N-terminal fragment results in a complex that interacts with the DNA and blocks its entry into the gel. It is possible that the protein-ADPNP complex is aggregating and preventing the DNA from entering the gel. In an attempt to optimise the assay, several different reaction conditions were employed: the pH was varied from ~8.0 to ~9.5, the reaction and gel were run at room temperature or 4°C, several buffers were tried, the DNA concentration was varied, BSA was added to try to stabilise the complex, the gel composition and strength were varied. All of these conditions failed to produce any increase in DNA binding as evidenced by a more resolute shifted DNA band. Overall, the results were unsatisfactory, but do suggest that the protein interacts with a 140 bp DNA fragment, and this interaction is altered by the presence of ADPNP.
Figure 5.7 Gel retardation assay using a 140 bp DNA fragment.

A 140 bp DNA fragment (50 nM) was incubated with protein at the indicated concentration (from 0.1 μM to 3 μM) in 50 mM Tris-HCl (pH 8.0), 20 mM NaCl, 5 mM MgCl₂, 10% (v/v) glycerol, 5 mM DTT, 0.1 mM EDTA at 25°C for 30 minutes. ADPNNP (or water) was then added to the samples to a final concentration of 2 mM and the samples were further incubated for 30 minutes at 25°C. Samples were loaded onto a pre-run 5% TBM gel and run at 100 V for 1 hour. DNA was visualised by staining with ethidium bromide.
5.2.3 Conformational mapping using limited proteolysis

Limited proteolysis of a protein in its native state can be used to probe the higher order structure of the protein, and to monitor ligand-induced conformational changes in protein structure (Hubbard, 1998). This technique was used to show that yeast topoisomerase undergoes allosteric interdomainal movements following ATP binding (Lindsley & Wang, 1991). Furthermore, limited proteolysis of the 43 kDa N-terminal fragment of *E. coli* DNA gyrase revealed that the presence of ADPNP protects a 33 kDa N-terminal fragment from the further digestion seen in the absence of nucleotide (Ali *et al.*, 1995). The difference in proteolysis pattern seen in these two cases is interpreted as being caused by the ADPNP-induced dimerisation of the N-terminal domains (Roca & Wang, 1992). Limited proteolysis has also been used to examine conformational changes in *E. coli* DNA gyrase, as evidenced by different proteolytic signatures (Kampranis & Maxwell, 1998a). The gyrase A_2B_2 tetramer has a proteolytic fingerprint consisting of an N-terminal 62 kDa fragment of GyrA and a C-terminal 25 kDa fragment of GyrB; this pattern was not altered by the presence of DNA. In the presence of ADPNP, two additional bands were seen, corresponding to the 43 kDa N-terminal domain and the 33 kDa subfragment of this domain mentioned above; again, the presence of DNA did not alter the proteolytic pattern observed.

To probe the conformation of the fragment topoIIα-(1-453) in the presence of nucleotides and DNA, the protein was first incubated with ligand for 1 hour at 25°C, following which trypsin was added to 15 μg/ml and samples were removed over a 2 hour period and analysed by SDS-PAGE electrophoresis. In the absence of ligand (Figure 5.8A) most of the protein is degraded to small peptides within the 2 hour period. With the protein-ADPNP complex, however, two protein fragments of ~40 kDa and ~35 kDa are protected from further digestion (Figure 5.8B). In the presence of ATP (Figure 5.8C), the proteolysis pattern observed appears to be ‘intermediate’ between that obtained in the presence and absence of ADPNP. This partial protection is to be expected as ATP hydrolysis would lead to dissociation of enzyme dimers and loss of the protein conformation leading to proteolytic protection. Incubation with ADP alone (Figure 5.8D) or ADP+P_i (data not shown) produced a similar pattern of proteolytic digestion as that seen with ATP. The presence of DNA (relaxed pBR322) at 50
Figure 5.8 Limited proteolysis of the fragment topoIIα-(1-453).
Samples containing protein (4 μM, 0.2 mg/ml) in 50 mM Tris-HCl (pH 8.8), 100 mM NaCl, 5 mM MgCl₂, 10% (v/v) glycerol, 5 mM DTT, 0.5 mM EDTA were incubated for 1 hour at 25°C in the presence or absence of nucleotide (2 mM) or relaxed pBR322 at 50 or 250 bp/dimer. Trypsin was then added to the samples at 15 μg/ml final concentration and the reactions incubated at 25°C for a further 2 hours. An aliquot was removed after 0, 1, 5, 15, 30, 60 and 120 minutes and analysed by SDS-PAGE electrophoresis on a 15% gel. Shown in this figure are tryptic digestion in the absence of any ligand (A) and the presence of ADPNP at 2 mM (B).
Figure 5.8 (cont) Limited proteolysis of the fragment topoIIα-(1-453).
Shown in this figure are tryptic digestion in the presence of 2 mM ATP (C) or the presence of 2 mM ADP (D).
Figure 5.8 (cont) Limited proteolysis of the fragment topoIIα-(1-453).
Shown in this figure are tryptic digestion in the presence of relaxed pBR322 at 50 bp/dimer (E) or 250 bp/dimer (F).
Figure 5.9 Summary gel for limited proteolysis of the fragment topoIIα-(1-453) without ligand, and in the presence of ATP or ADPNP.

Tryptic digest experiments were performed as described previously, in the presence of no ligand, 2 mM ATP or 2 mM ADPNP. An aliquot was removed after 30 and 60 minutes and analysed by SDS-PAGE electrophoresis on a 15% gel, along with a sample of purified protein (453C). A duplicate gel was blotted to a PVDF membrane and the indicated bands were subjected to N-terminal sequence analysis (carried out by K. Lilley, University of Leicester). The N-terminal amino acid sequence of the excised bands are indicated, along with the corresponding residue number in the human DNA topoisomerase IIα sequence. Also shown is the amino acid sequence of the N-terminal region of human DNA topoisomerase IIα, the sequences found in N-terminal sequencing of peptides are underlined.
bp/dimer (Figure 5.8E) and 250 bp/dimer (Figure 5.8F) did not alter the proteolysis pattern from that seen with protein alone, while the presence of ADPNP and DNA did not alter the proteolysis pattern from that seen with ADPNP alone (data not shown).

To enable the protein bands to be identified, digestion experiments containing either no ligand, or ADPNP or ATP at 2 mM, were repeated and the 30 and 60 minute samples were run on an SDS-PAGE gel along with the undigested protein (Figure 5.9). The five bands numbered 1 to 5 were subjected to N-terminal sequence analysis (K. Lilley, University of Leicester). In the absence of any ligand or the presence of ATP, the highest molecular weight band seen (~38 kDa, #2) had an amino acid sequence corresponding to residues 37-42. In the presence of ADPNP, the two protected bands at ~40 kDa (#4) and ~35 kDa (#5) had an amino acid sequence corresponding to residues 26-31. This suggests that the stretch of residues from 26-36 is protected from proteolytic removal in the protein-ADPNP complex. Such protection is consistent with the crystal structure of the 43 kDa fragment of *E. coli* DNA gyrase in the presence of ADPNP (Wigley *et al.*, 1991), where the N-terminal residues form an extending ‘arm’ that interacts with the adjacent enzyme molecule in the dimer (Figure 1.8).

5.3 DISCUSSION

Central to the proposed model for the catalytic cycle of type II DNA topoisomerases is the role played by the N-terminal domains in capturing a DNA segment and ‘directing’ it through an enzyme-mediated break in a second DNA segment (Wigley *et al.*, 1991; Roca & Wang, 1992, 1994). In this chapter we have tested this hypothesis by investigating the ability of the N-terminal fragment topoIIα-(1-453) to dimerise in the presence of nucleotides and bind DNA. Earlier work on the N-terminal domain of human DNA topoisomerase IIα using the 52 kDa protein suggested that it exists as a dimer in solution and interacts with DNA, due to the ability of DNA to stimulate its intrinsic ATPase activity (Gardiner *et al.*, 1998). This was intriguing given that the 43 kDa fragment of *E. coli* DNA gyrase dimerises only in the presence of nucleotide and does not bind to DNA (Ali *et al.*, 1993), and it was thought that these
differences in the properties of the N-terminal domains of human and *E. coli* type II topoisomerases may, at least in part, account for the different reactions catalysed by these enzymes.

The data obtained from analytical ultracentrifugation experiments suggest that, under the experimental conditions used, the N-terminal fragment topoIIα-(1-453) is in fact a monomer in solution, and dimerises only in the presence of ADPNP. This is consistent with the results obtained in cross-linking studies, where high molecular weight species were seen in the presence of ADPNP but not in its absence, and also in the greater than first order dependence of ATPase rate on enzyme concentration (see Chapter 4). Gel filtration in the absence of ADPNP revealed a single species (Figure 5.4), now presumed to be the monomer, although it may still represent a monomer-dimer equilibrium. The presence of ADPNP caused two peaks to elute at a higher molecular weight, the lower of which is likely to represent the dimer species. The higher molecular weight species, therefore, may represent a tetramer. The existence of such tetramers has been previously observed with yeast topoisomerase II (Vassetzky *et al.*, 1994) and with full-length human DNA topoisomerase IIα (Hammonds & Maxwell, 1997), and the data seen here suggest that the N-terminal domains of these enzymes may play a role in the formation of these tetramers. Limited proteolysis confirmed that a large conformational shift accompanies ADPNP, and presumably ATP, binding to the enzyme, leading to dimerisation and resulting in proteolytic protection of two protein fragments of ~40 kDa and ~35 kDa (Figure 5.8B).

In the presence of ATP, therefore, one would expect that ATP binding to the enzyme would induce it to dimerise; ATP hydrolysis and/or product release would cause the dimer to dissociate. The proteolysis patterns obtained with ATP and ADP±P_i are similar (Figure 5.8C and 5.8D). If the bound ATP is rapidly hydrolysed to ADP+P_i, and the dissociation rate of this from the enzyme is relatively slow, then this would account for the similar proteolysis pattern observed. In cross-linking experiments, high molecular weight bands were seen with ATP and ADP+P_i, albeit to a lesser extent than were seen with ADPNP. This suggests that the enzyme can exist in a dimer state with ADP+P_i bound, and that full product release may be necessary for dimer dissociation.
Alternatively, a conformational change involving another part of the enzyme may also be involved in the dissociation step. It is not clear why the addition of ADP+P_i should cause the N-terminal fragment topoIIα-(1-453) to dimerise; the 43 kDa N-terminal fragment of *E. coli* DNA gyrase is a monomer in the presence of ADP+P_i (Ali et al., 1993). This may reflect an increased affinity of the human enzyme for ADP+P_i.

The N-terminal fragment topoIIα-(1-453) was shown to bind to DNA in filter-binding assays (Figure 5.6). The affinity of the fragment topoIIα-(1-453) for DNA is much weaker than that of the full-length enzyme in similar assays (A. Howells, personal communication). This is to be expected, as the interaction of DNA with the N-terminal domains must only be a transient one; if the interaction were too strong then strand passage might not occur. Cross-linking studies did not reveal any high molecular weight bands in the presence of DNA. Furthermore, limited proteolysis in the presence of DNA produced a pattern of digestion similar to that seen for protein alone. These results suggest that the interaction of the N-terminal fragment with DNA does not induce any significant conformational change in the protein, nor does it induce it to dimerise.

How does ADPNP alter the DNA-binding characteristics of the N-terminal fragment topoIIα-(1-453)? In the absence of ADPNP *i.e.* in the monomeric form, the fragment topoIIα-(1-453) was able to bind DNA. Addition of ADPNP, either before or after the DNA, caused an increase in DNA binding. Dimerisation presumably localises the DNA binding surface in a conformation which is more competent to bind DNA. Where the clamp is pre-locked with ADPNP, the DNA may enter up through the bottom of the clamp in this truncated fragment. This assumes that the DNA interacts with the inside of the clamp, which may not be the case, although it is the most likely scenario given the crystal structure of the 43 kDa fragment of *E. coli* DNA gyrase (Wigley et al., 1991).

In the full-length homodimeric enzyme, the N-terminal domains are joined to the core of the enzyme. Electron microscopy revealed that each domain extends from the globular core via a flexible linker (Schultz et al., 1996). Addition of ADPNP caused
the two N-terminal domains to interact to form a single large domain, consistent with the ADPNP-induced dimerisation of topoIIα-(1-453) seen in analytical ultracentrifugation experiments. The images from electron microscopy suggest that dimerisation is brought about by reorientation of polypeptide chains in the junction between the N-terminal domains and the core of the enzyme (Schultz et al., 1996). These images, and the data shown here, suggest that the N-terminal domains do not interact with one another in the absence of ADPNP; the jaws of the clamp are held wide open to facilitate the capture of the T segment. When ADPNP (or ATP) is bound, the N-terminal domains dimerise by means of interactions around the nucleotide binding site, including the N-terminal arm, which was shown to be protected by ADPNP binding. If the N-terminal fragment topoIIα-(1-453) is able to bind DNA in the absence of ADPNP, it may suggest that each of the N-terminal domains in the homodimeric enzyme is able to interact with DNA in the open clamp formation, although presumably only one DNA duplex is captured.
CHAPTER 6

Discussion
6.1 GENERAL DISCUSSION

This project was instigated following the initial characterisation of an N-terminal 52 kDa fragment of human DNA topoisomerase IIα (Gardiner et al., 1998). This original study revealed some interesting insights into the ATPase activity and catalytic functioning of the N-terminal domain of this enzyme, and eukaryotic type II topoisomerases in general, and highlighted some differences between the N-terminal domain of prokaryotic and eukaryotic type II topoisomerases. However, this initial study was dogged by difficulties in expressing and recovering a soluble form of the 52 kDa protein. At the outset of my studies the goal was to reclone the N-terminal fragment of human topoisomerase IIα to obtain sufficient protein of a more consistent biochemical activity to use in further characterisation of the N-terminal domain of this enzyme.

Based on the N-terminal domains of three type II DNA topoisomerases, *E. coli* DNA gyrase GyrB (Adachi et al., 1987), yeast DNA topoisomerase II (Lindsley & Wang, 1991), and human DNA topoisomerase IIβ (Austin et al., 1995), a variety of plasmid constructs representing the N-terminal domain of human DNA topoisomerase IIα were produced. Although the sequence similarity of *E. coli* DNA gyrase GyrB and yeast DNA topoisomerase II with human DNA topoisomerase IIα in the boundary region between the ATPase domain and the core-domain of the enzyme is quite low (Caron & Wang, 1994; see Figure 3.1), more information is available on the N-terminal domain of the *E. coli* and yeast proteins, and hence it was decided to include constructs that ended in a homologous region to the N-terminal domains of these two proteins. The gene for the N-terminal region of human DNA topoisomerase IIα was fused to the gene encoding the thioredoxin protein of *E. coli* and a short polylinker sequence encoding a hexa-histidine tag. Expression of the fusion proteins in BL21(DE3) cells at 20°C gave a significant amount of soluble protein that could be isolated to high purity using a two-step purification procedure. The fusion proteins could be successfully cleaved with Factor Xa and repurified to leave a native N-terminal fragment with no foreign amino acids. All of the fusion proteins were found to possess a DNA-stimulated ATPase activity (Figure 3.15), as were the cleaved and purified N-terminal fragments. The intrinsic and DNA-dependent ATPase activity of two of these proteins in both the
fusion and cleaved states, topoIIα-(1-420) and topoIIα-(1-453) (Figure 3.18), revealed that the longer protein topoIIα-(1-453) has a higher intrinsic ATPase activity that is stimulated by DNA in a complex manner, the so-called 'hyperstimulation' effect seen before with the full-length enzyme (Hammonds & Maxwell, 1997), but not the 52 kDa protein (Gardiner et al., 1998), nor the shorter N-terminal fragment topoIIα-(1-420). Based on the results presented in this thesis as a whole, it is not possible to determine whether the N-terminal fragment topoIIα-(1-453) is a true representation of the N-terminal domain of human DNA topoisomerase IIα, although it does contain all the elements required for ATP hydrolysis and most likely encompasses the complete ATPase domain. Although soluble protein has been obtained using a fusion protein and lower growth temperature, it is likely that the N-terminal domain is relatively unstable and prone to misfolding when separated from the remainder of the enzyme.

Kinetic studies of the ATPase activity of topoIIα-(1-453) showed that the presence of DNA induces a change from a parabolic to a linear dependence of ATPase rate upon enzyme concentration (Figures 4.2A and 4.2C). This suggests that the protein exists predominantly in the monomeric form in the absence of DNA, which is the same as was observed for the 43 kDa domain of *E. coli* DNA gyrase (Ali et al., 1993) and the N-terminal fragment of yeast DNA topoisomerase II (Olland & Wang, 1999). In contrast, in the presence of DNA, the linear dependence of ATPase activity upon protein concentration for topoIIα-(1-453), but not for the 43 kDa domain of *E. coli* DNA gyrase, suggests that DNA may induce the human N-terminal fragment to dimerise. However, in cross-linking studies, no high molecular weight species were seen when cross-linking was performed in the presence of DNA, whereas cross-linked species were seen with ADPNP. The interaction of enzyme monomers with DNA, leading to dimerisation, may occur only transiently, particularly in the absence of ATP, and thus the dimers species may not have been detected by cross-linking. Alternatively, DNA may not induce the enzyme to dimerise; it is possible that, in the presence of DNA, dimerisation is not a prerequisite for ATP hydrolysis, or dimerisation is not the rate-limiting step in ATP hydrolysis.
At a constant enzyme concentration, the rate of ATP hydrolysis was hyperbolically dependent upon ATP concentration (Figures 4.3 and 4.4). The intrinsic ATPase rate of topoIIα-(1-453) was found to be stimulated ~15-fold by DNA at 50 bp/dimer and ~10-fold by DNA at 250 bp/dimer. The shorter fragment topoIIα-(1-420) has a lower intrinsic ATPase activity which is stimulated ~6-fold by excess DNA. Direct comparison of the results for the full-length enzyme (Hammonds & Maxwell, 1997) and the 52 kDa fragment (Gardiner et al., 1998) with the results shown here are complicated by the non-linear dependence of the intrinsic ATPase rate upon enzyme concentration; however, the longer N-terminal fragment topoIIα-(1-453) is distinct from the shorter N-terminal fragments (topoIIα-(1-420) and the 52 kDa protein) in terms of the level of intrinsic and DNA-dependent ATP hydrolysis and its affinity for substrate ($K_m$), and does in fact display similar catalytic properties to the full-length enzyme. The different properties of these two N-terminal fragments is presumably due to the longer C-terminus of topoIIα-(1-453), which increases the intrinsic ATPase rate and alters the DNA-stimulation and substrate affinity of the enzyme. These residues may stabilise the enzyme in a configuration that is more competent to hydrolyse ATP, and alter the interaction of the enzyme with DNA either directly or indirectly.

The intrinsic ATPase activity of the N-terminal fragment topoIIα-(1-453) was found to be stimulated by a DNA fragment of ~10 bp (Figure 4.12), although a DNA fragment of ~14-16 bp may be required for full stimulation. This is consistent with the size and shape of the cavity in the structure of the 43 kDa-ADPNP complex (Wigley et al., 1991), and suggests that the DNA fragment stimulates the ATPase activity by binding within the cavity rather than interacting with, or wrapping around, the outside of the protein. Given the above, it is interesting to note that a DNA fragment of 40 bp is not sufficient to stimulate the ATPase activity of the full-length enzyme (except at >1000-fold excess over enzyme); a DNA fragment of >100 bp is required (Hammonds & Maxwell, 1997). In the full-length enzyme, DNA stimulation of the ATPase rate may be dependent upon both the T and G segments binding, which may be favoured by a segment of DNA where the T and G segments are contiguous, i.e. requires a DNA fragment of >100 bp (Hammonds & Maxwell, 1997).
The DNA-length dependence of the hyperstimulation effect was investigated using relaxed pBR322 and linear DNA fragments. With relaxed pBR322 the ATPase rate peaked at 50 bp/dimer and dropped to a basal level of DNA-stimulation by ~150 bp/dimer (Figure 4.7). It was originally thought that this inhibition of DNA-stimulated ATPase activity at high DNA concentrations was due to binding of a third DNA segment, possibly within the C-terminal domains (Hammonds & Maxwell, 1997). However, a C-terminal deletion mutant of the yeast enzyme still exhibited a similar hyperstimulation pattern to the full-length enzyme (Hammonds & Maxwell, 1997). The similarity in the hyperstimulation pattern seen with the fragment topoIIα-(1-453) and the full-length human enzyme (albeit at different DNA:enzyme ratios), suggests that this hyperstimulation effect is principally a function of the N-terminal domains of the type II eukaryotic topoisomerases. The hyperstimulated ATPase rate was previously attributed to protein-protein interactions between adjacent dimers bound to DNA, and this is depicted for the N-terminal domains in Figures 3.14 and 3.15. The exact nature of these interactions is unclear, but presumably is dependent upon residues at the C-terminal end of topoIIα-(1-453). Incongruous with this model is the finding that a DNA base pair:enzyme dimer ratio of five times that required to stimulate the ATPase activity \( i.e. \sim 10 \text{ bp} \), is required to maximally hyperstimulate the ATPase activity. If the enzyme dimer occupies \( \sim 10 \text{ bp} \), this implies that the distance between enzyme dimers on a molecule of relaxed pBR322 is \( \sim 40 \text{ bp} \), which would make protein-protein interactions such as those depicted in Figure 4.14 (middle) unlikely. This is particularly interesting given that a base pair:enzyme dimer ratio of \( \sim 20 \) gives rise to a hyperstimulated ATPase rate when short DNA molecules are used (assuming the model in Figure 4.14 to be correct). It therefore appears that the bp:dimer ratio required to hyperstimulate the ATPase activity (and the maximum ATPase rate achieved) is proportional to the length of DNA used (up to a maximum) \( i.e. \sim 20 \text{ bp/dimer with short (<40 bp) DNA fragments, rising to } \sim 40 \text{ bp/dimer with a 140 bp DNA fragment, and } \sim 50 \text{ bp/dimer with pBR322 (relaxed and linear). This may imply that as the length of the DNA increases, the maximum hyperstimulated ATPase rate occurs as enzyme dimers become increasingly separated along the DNA molecule, up to a maximum at } \sim 50 \text{ bp/dimer.} \)
Using a linear DNA fragment of 140 bp at high base pair:enzyme dimer ratios (i.e. >200 bp/dimer) the ATPase rate is significantly lower than at high concentrations of pBR322 (Figure 4.10). With DNA fragments of 32-44 bp, where the hyperstimulation effect was apparent, the ATPase rate also dropped to a very low level at high concentrations of DNA (Figure 4.13). In both these cases, the ATPase rate drops below the plateau level that is attained with high concentrations of DNA fragments <30 bp, where no hyperstimulation effect is apparent. Therefore, the hyperstimulated ATPase rate is dependent upon bp/dimer ratio (and DNA length), and the basal DNA-stimulated ATPase rate is also dependent upon DNA length. The presence of the hyperstimulated ATPase effect at low DNA concentrations leads to a relatively low basal DNA-stimulated ATPase rate at high DNA concentrations. The reason for this drop in ATPase rate at high concentrations of DNA fragments >32 bp is not known. It is possible that, as has been previously postulated (Hammonds & Maxwell, 1997; Olland & Wang, 1999), another segment of the DNA fragment binds to the enzyme at high DNA concentrations and inhibits the ATPase activity; this interaction requiring a DNA fragment of >32 bp. This mode of inhibition itself would be dependent on DNA length, and may also involve protein-protein interactions. Furthermore, if the monomer form is capable of binding to DNA, as suggested by the filter binding experiments, the separation of enzyme monomers would increase as the DNA concentration increases, and therefore dimerisation would be less favourable.

The cause of the hyperstimulated ATPase activity seen at low DNA:enzyme ratios is unknown; indeed, we do not know whether it represents a stimulation at low DNA:enzyme ratios, or an inhibition at high DNA:enzyme ratios. This effect was not seen with the N-terminal domain of yeast topoisomerase II, Top2-(1-409), but was seen with the longer fragment Top2-(1-660), which includes the B'-subfragment of the enzyme (Olland & Wang, 1999). It is possible that the human fragment topoIIα-(1-453) includes part of the B'-subfragment of the human enzyme, and this B'-subfragment may be responsible for the hyperstimulation effect. If, as we have suggested, the hyperstimulated ATPase rate is a result of protein-protein interactions occurring between enzyme dimers on the DNA molecule, there are several possibilities that may account for this. The interactions between adjacent enzyme molecules bound to DNA
may lead to an increased rate of dimerisation, resulting in multiple cycles of ATP binding, hydrolysis and product release/dissociation without the enzyme releasing the DNA and/or its association with its neighbours. Given that the rate-limiting step in the ATPase cycle of the 43 kDa protein of DNA gyrase is predicted to be dimerisation (Ali et al., 1993), such an interaction could lead to a significant increase in the rate of ATP hydrolysis. Alternatively, the interactions may enhance the rate of product release, and, in a manner opposite to that above, may actually cause the enzyme to be displaced from the DNA molecule so that it can rebind ATP.

In the context of the catalytic cycle of the full-length enzyme, the N-terminal domains operate as ATP-operated DNA-binding clamps (Roca & Wang, 1992, 1994). We have shown that the N-terminal domains do not associate into dimers when isolated from the remainder of the enzyme, and it is possible that this is also the case in the full-length enzyme. This would enable the jaws of the clamp to be held open to facilitate capture of a T segment. The boundary between the ATPase domains and the B'-subfragment of human DNA topoisomerase IIα is unknown, and this region is known to undergo large conformational changes during the course of the catalytic cycle (Schultz et al., 1996). We have shown that the N-terminal domains bind DNA in the presence and absence of ADP·NP, which is interpreted as the dimeric and monomeric form, respectively. Therefore, both N-terminal domains of one full-length enzyme may bind to a DNA T segment, although presumably only one is captured. Binding of ATP causes the N-terminal domains to dimerise and clamp around the DNA, leading to its capture within the cavity seen in the crystal structure of the 43 kDa fragment of E. coli DNA gyrase (Wigley et al., 1991). Given the size of this cavity, and the finding that ~10 bp is sufficient to stimulate the ATPase activity of the N-terminal domains, it seems reasonable to assume that a length of DNA of ≈10 bp is encompassed by the N-terminal clamp. The presence of DNA within the clamp, and also the G segment at the DNA gate, increases the rate of ATP hydrolysis; binding of both the T and G segments are required for full DNA stimulation of the ATPase rate. It is not clear how the T segment stimulates the ATPase activity, it may merely induce dimerisation of the N-terminal domains, or it may directly accelerate the rate of ATP hydrolysis by promoting a conformational change at the active site. It is likely that regions outside of the N-
terminal domains are also involved in coupling ATP hydrolysis to DNA-cleavage and strand passage, most notably the B'-subfragment of the enzyme.

The role of the hyperstimulated ATPase activity in the normal catalytic cycle of the type II topoisomerases remains to be determined. It is not known whether the increased ATPase rate translates to an increased rate of strand passage, or is a result of uncoupling of ATP hydrolysis from DNA transport. In section 1.3.1, the role of type II topoisomerases in the maintenance of chromosome structure was described, whereby topoisomerase II is located at scaffold-attachment regions (SARs) at the base of chromatin loops. Topoisomerase II interacts with DNA containing SAR sequences and causes extensive protein-DNA aggregation, suggestive of DNA-dependent protein-protein interactions (Adachi et al., 1989), and also results in the formation of higher order complexes of topoisomerase dimers (Vassetzky et al., 1994). The occurrence of the hyperstimulated ATPase rate in these higher order complexes of topoisomerase dimers may lead to increased catalytic activity at these sites where DNA is highly concentrated.

6.2 FUTURE PERSPECTIVES

One important goal in any enzyme characterisation study is the elucidation of the detailed three-dimensional structure of the protein being investigated. Although we already have the structure of several fragments of type II DNA topoisomerases, we do not have the complete structure of any single topoisomerase. Given the structural homology within the type II topoisomerases, it is possible to piece together a structural model of a functional type II topoisomerase, but one must be cautious in this approach. This work and that of others support the model of the N-terminal domains of type II topoisomerases as ATP-operated DNA-binding clamps, but these N-terminal domains also show important differences in biochemistry which presumably have a structural basis, and therefore we cannot always extrapolate the known structure of one fragment to all other fragments. To this end, the determination of the detailed structure of the N-terminal domain of human DNA topoisomerase IIα would go a long way to resolving
the basis of the differences in the ATPase activity and catalytic properties of the prokaryotic and eukaryotic type II topoisomerases studied to date. Crystallisation trials were set up during the course of this project with the N-terminal fragment topoIIα-(1-453) alone or in the presence of ADPnP or a DNA fragment of 10 bp; however, no crystals were obtained (data not shown). Any structural determination project requires a relatively large amount of homogeneous soluble protein, which has not yet been achieved for fragments of human type II topoisomerases. The baculovirus-insect cell expression system has been used to overcome the problem of low yield in the expression of human DNA topoisomerase I (Stewart & Champoux, 1999). Such an approach in the expression of the N-terminal domain of human DNA topoisomerase IIα may prove fruitful in obtaining large quantities of soluble, active protein suitable for structure determination and biochemical characterisation.

The nature of the interaction between the N-terminal domains and DNA requires further investigation. Resolution of the structure of an N-terminal domain-DNA complex at the atomic level would go a long way to answering many of the questions about the interaction of the N-terminal domains of type II topoisomerases with DNA. The interaction between enzyme and DNA appears to be a complex one, which is dependent upon the length and/or concentration of DNA. The DNA sequence specificity of the interaction could also be examined; SAR sequences at the base of DNA loops are AT-rich. The site of interaction of the DNA on the enzyme is more difficult to determine. Mutagenesis studies would give some insight into residues involved in this interaction, but without the structure of this domain any mutagenesis would have to be done ‘blind’. Computer modelling of the N-terminal domain of the human enzyme based on the N-terminal 43 kDa fragment of E. coli DNA gyrase was attempted, but the only region within the two enzymes with sufficient homology for modelling was close to the ATP binding site, not in the C-terminal part of the domains where DNA is thought to interact. The question of how DNA increases the ATPase rate of the N-terminal domain is also unanswered. The effect of the presence and length of DNA on the oligomeric state of the N-terminal domain could be determined by analytical ultracentrifugation. It would also be useful to study the N-terminal domains within the full-length enzyme itself, which can carry out the strand passage reaction and
therefore provide us with more information on the role of the ATPase domains in the catalytic cycle. One question worthy of further investigation in the full-length enzyme is whether the hyperstimulated ATPase rate produce an increased rate of strand passage.

In summary, we have examined the role of the N-terminal domain of human DNA topoisomerase IIα as an ATP-operated DNA-binding clamp. We have found that this domain fulfills all the requirements placed on it by the proposed catalytic cycle of the full-length enzyme, *i.e.* it can capture DNA by an ATP-promoted dimerisation step, and its interaction with each of these is affected by the presence of the other. The N-terminal domain of this enzyme is like the N-terminal domain of its prokaryotic counterpart from *E. coli*, DNA gyrase, in that it exists as a monomer in solution and dimerises into an active form when bound to ATP. However, the human fragment’s interaction with DNA has highlighted the usefulness of examining individual domains of a complex protein to dissect its catalytic function. The complex dependence of ATPase rate on DNA length is markedly similar to that seen with the full-length enzyme (Hammonds & Maxwell, 1997), and this hyperstimulation effect is therefore attributable to the N-terminal domains. Further analysis of the N-terminal domain, either in isolation or within the full-length enzyme, is required to determine the role of the hyperstimulation phenomenon in the catalytic function of this enzyme.
CHAPTER 7

References
7.1 REFERENCES


