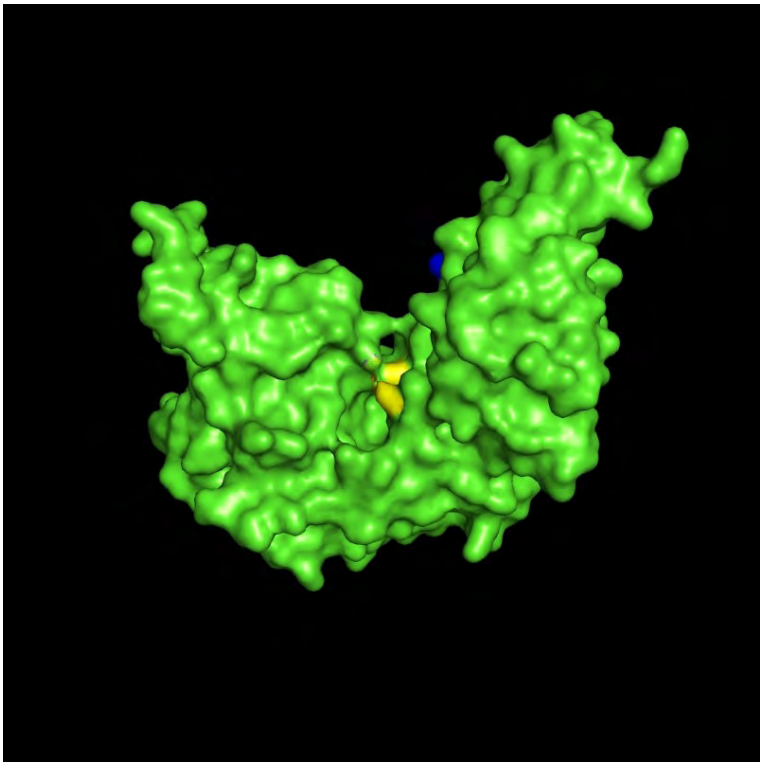
UPPSALA
UNIVERSITET

Efficient ligation of DNA on RNA templates using a mutated T4 DNA ligase



Saiful Islam

Degree project in biology, Master of science (1 year), 2008
Examensarbete i biologi 15 hp till magisterexamen, 2008
Biology Education Centre and Division of Molecular Neurobiology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Uppsala University
Supervisor: Sten Linnarsson, PhD, Asistant Professor

Contents

ABBREVIATIONS

AMP	Adenosine monophosphate
ATP	Adenosine-5'-triphosphate
cDNA	Complementary DNA
CWT	Commercial wild type
ddH_2O	Double distilled water
dNTPs	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
E.coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
GMP	Guanosine monophosphate
His	His tagged
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kb	Kilobase
KCl	Potassium cholaride
KDa	Kilodalton
LA	Luria agar
LB	Lysogeny broth
mRNA	Messenger ribonucleic acid
mM	Milimolar
μM	Micromolar
NaCl	Sodium chloride
NAD^+	Nicotinamide adenine dinucleotide
ndsDNA	Nick double strand DNA
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PMSF	Phenylmethanesulphonylfluoride
SDS PAGE	Sodium dodecyl sulfate poly acrylamide gel electrophoresis
TAMRA	Tetramethyl-6-Carboxyrhodamine
Tris	Tris (hydroxymethyl) aminomethane
WT	Wild type
U	Unit

SUMMARY

ATP-dependent DNA ligase has very essential roles in different important biological processes like DNA replication, DNA repair etc. In addition, DNA ligase is a workhorse of modern molecular biology and now widely used in genetic recombination. The most commonly used ligase is that from bacteriophage T4, which is commercially available, efficient and stable. T4 DNA ligase is an Mg^{2+} and ATP dependent enzyme that seals a DNA nick in three steps: activation of the enzyme by covalent binding of AMP, transfer of the nucleotide to the 5′ end of the nick, and finally formation of a phosphodiester bond that results in sealing of the nick and releasing free AMP. The strong preference of DNA ligases for perfectly base-paired substrates has been used to enable accurate detection of genetic polymorphisms; for DNA sequencing and for selective amplification of desired targets. T4 DNA ligase is very efficient for ligation of DNA on DNA template. It also ligates DNA on RNA template, but the reaction is inefficient and a large fraction of partially completed (adenylated) products are formed.

I tested whether a combination of T4 WT DNA ligase (which carried out first two steps of ligation efficiently) and its mutant K159A (which carried out third step of ligation efficiently) could ligate DNA on an RNA template efficiently. The mutant T4 DNA ligase cannot perform the first step of the ligation reaction since the mutated lysine residue is in the active site where AMP is attached to form ligase adenylate. In the presence of a low concentration of ATP (10 μM) and Mg^{2+} (10 mM), a combination of a high concentration of T4 WT DNA ligase (0.5 U/ μl) and the K159A mutant ligase (0.05 U/μl) could ligate DNA very efficiently on an RNA template within the first hour of the reaction. Beyond the first hour, no additional ligation of DNA on an RNA template was observed using the enzyme mix while 80% ligation was achieved in the RNA-template reaction in 4 hrs by T4 wild-type DNA ligase enzyme alone.

The ligation reaction stopped after two hours, which might be because of other proteins were present. With more highly purified enzymes, the novel DNA ligase could be used for efficient ligation of DNA-RNA hybrids to detect and distinguish RNA sequence variants and other broad applications.

Front-page picture: Mutant K159A (mutation shown in yellow color at 159 position) carrying additional mutation (mutation at 286 position shown in blue color). The picture is the structure of T7 DNA ligase but residues number refers to T4 DNA ligase. Picture prepared and kindly provided Sten Linnarsson.

1. INTRODUCTION

1.1 T4 DNA Ligase

The enzyme DNA ligase catalyzes the formation of phosphodiester bonds at single stranded or double stranded breaks between adjacent 5′-phosphate and 3′-hydroxyl site. The strong preference of DNA ligase for perfectly base-paired substrates has been used to enable accurate detection of genetic polymorphisms, for DNA sequencing, and for selective amplification of desired targets. One major group of DNA ligases enzymes comprises that require ATP as cofactor includes enzymes from eukaryotic cells as well as from bacteriophages of the T series, whereas enzymes in another major group, including eubacteria DNA ligase, require NAD^+ as cofactor (reviewed in Kornberg *et al.* 1991, Lindahl *et al.* 1992). The only homology between NAD^+ dependent DNA ligases and ATP-dependent DNA ligase is the AMP binding site. In contrast, several homologies were found between ATP-dependent DNA ligases and mRNA capping enzymes that carry out very similar transadenylation reactions using GMP (Shuman *et al.* 1995). The most commonly used ligase, which is commercially available, efficient and stable, is from bacteriophage T4. T4 DNA ligase can join either blunt ends or sticky ends of two doublestranded DNA fragments, or it can seal a break between two singlestranded DNA fragments annealed on the complementary DNA strand (called nick-ligation).

DNA ligase generally has very important roles in DNA repair, DNA replication and DNA recombination (Lehman 1974). In genetic engineering, DNA ligase is now indispensable to make genetic recombination. For example, DNA ligase is widely used to insert DNA into plasmids. DNA ligase is a workhorse of modern molecular biology. During DNA replication, a short DNA strand, called Okazaki fragment, is created on the lagging strand. This lagging strand grows discontinuously in the opposite direction of the main leading strand. Discontinuous but adjoining small fragments are then linked to each other by DNA ligase to create a continuous strand of DNA (Alberts *et al.*, 2002).

1.2 The ligation reaction

The nick-sealing activity of T4 DNA ligase utilizes a nicked doublestranded DNA (ndsDNA), ATP and an inorganic cofactor Mg^{2+}. Ligation by T4 DNA ligase proceeds in three steps and involves two covalently joined reaction intermediates **(Figure 1)**. Step I involves adenylation of the ligase where AMP is linked covalently to the ε-amino group of a lysine (159 position) in the active site of the enzyme accompanied by the release of pyrophosphate (PP_i) from ATP (Rossi *et al.* 1997). In step II, the ligase-adenylate forms a transient complex with the nicked doublestranded DNA. In the presence of ATP, the adenylated enzyme searches for a 5′-phosphorylated end through the formation of successive transient complexes. Once it finds the 5′-phosphorylated nucleotide, it transfers the adenylate group to the 5′-phosphorylated site and a stable complex is formed which sits on the DNA until the 3′ end becomes available to complete the sealing reaction (Rossi *et al.* 1997). In step III, the ligase catalyzes an attack on this pyrophosphate bond by the OH group at the 3′ end of the nick resulting in sealing of the nick and release of free AMP and enzyme (Lehman 1974). The enzyme is then reloaded with fresh AMP and starts a new cycle of ligation reaction (Rossi *et al.* 1997).

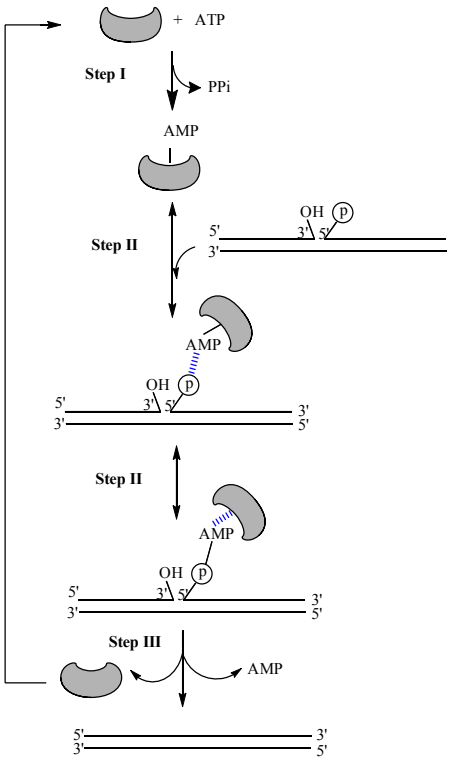


Figure 1: Mechanism of ligation by T4 DNA ligase. The ligation reaction of nicked nucleic acid performed in a three steps reaction of which step II is reversible shown in figure.

1.3 Ligation of RNA-templated DNA

The DNA-DNA ligation reaction and the mechanism to seal nicked DNA substrates by T4 DNA ligase are well studied in details (Higgins *et al.* 1979, Lehman 1974, Rossi *et al.* 1997). However, the mechanism of RNA-templated DNA ligation has not been studied in details. The kinetics of the DNA-joining reaction on an RNA template is very slow. It is found from the structural analysis that the nicked substrate must adopt a B-helical conformation in order to be sealed by DNA ligase (Sekiguchi *et al.* 1997). Usually, doublestranded RNA adopts an A-helical conformation and the RNA strands of RNA-DNA hybrid adopt a similar helical conformation that is not favorable for the ligase enzyme to sealed the nick (Sekiguchi *et al.* 1997). Though the reaction kinetics of RNA ligation are much slower than those of DNA ligation, according to the Rossi *et al.* model (Rossi *et al.* 1997), the ligation reaction is inhibited by an ATP concentration exceeding the K_m (Michaelis-Menten constant, corresponds to the substrate concentration at which half of the reaction rate is achieved) for ATP binding. That is, ligation of DNA on RNA template is inefficient due to the enzyme prematurely leaving the adenylated nick and becoming adenlyated again, resulting in an accumulation of adenylated nicks.

When the concentration of ATP is kept low (~ 10 μM), the nick sealing reactions on an RNA template molecule is more efficient (Nilsson *et al.* 2001). This is because at low concentration of ATP (< 40 μM), the ligase binds to ATP more slowly than it binds to the template (Cherepanov *et al*. 2003), and therefore unadenylated enzymes are available to seal already adenylated templates. In contrast, at high ATP concentration (1-5 mM), the ligase binds faster to ATP than to the template, and as a consequence, adenylated ndsDNA accumulates. This is similar to other difficult ligations, such as blunt end ligation, which are inhibited by premature binding of AMP to the ligase that leads to dissociation of enzymes from the substrates after 5′ adenylation step (Rossi *et al.* 1997).

1.4 Mutated T4 DNA ligase

The idea to use mutated T4 DNA ligase to seal the nick more efficiently in RNA-templated DNA came from the studies of vaccinia virus ligase. Vaccinia virus encodes a 552 amino acid DNA ligase that has a nick-sensing function. This enzyme has the ability to distinguish between nicked DNA substrate containing a 5′ phosphate and nicked DNA containing a 5′-hydroxyl group (Sekiguchi *et al.* 1997). It has been shown that vaccinia virus DNA ligase where the active site lysine-231 is replaced with alanine cannot form the covalent ligase-adenylate intermediate. As a consequence, it cannot bind to nicked DNA and cannot seal the nick. In contrast, it has been found that this mutated vaccinia virus DNA ligase can catalyze the formation of a phosphodiester bond if the nicked DNA substrate is preadenylated (Sekiguchi *et al.* 1997). In that mutant vaccinia virus ligase, the affinity to bind to nicked DNA adenylate is ten fold higher than for a plain nicked DNA (Sekiguchi *et al.* 1997).

1.5 Aim of this study

The aim of this project was to determine if a mutated T4 DNA ligase in combination with wild type enzyme might efficiently ligate DNA-RNA hybrid. By analogy with vaccinia DNA ligase, the wild type T4 DNA ligase enzyme was expected to efficiently adenylate the nick (step I and step II) but not to seal the DNA nick efficiently (step III). On the other hand, mutated T4 DNA ligase might be able to seal the adenylated nick (step III) though it could not adenylate the nick at all. With a mixture of these two enzymes, I predicted that all three steps of nick sealing reaction would be performed more efficiently.

2. RESULT

2.1 Construction of plasmids expressing wild type and mutant T4 DNA ligase

The mutant (K159A) as well as wild type T4 DNA ligase was obtained via site directed mutagenesis on a gene encoding mutant T4 DNA ligase (K159L). The plasmid pTrcHis carrying the coding region of T4 ligase K159L was PCR-amplified using forward and reverse primer designed to introduce the mutation at the desired site. The PCR products were analyzed on an agarose gel **(Figure 2a)**. A clear band appeared at the expected position for the entire construct at about 6.3 Kb (pTrcHis plasmid 4.4 Kb and T4 DNA ligase gene 1.9 Kb). After transformation of plamids into Top10 *E.coli* cells, successfully transformed plasmids were digested by restriction enzymes (EcoRV) and analyzed on agarose gels **(Figure 2b)**. After cutting with EcoRV, an expected size (6.3 Kb) of band appeared (Lane 4, 6, 8 and 10). When plasmids were doubly digested, two bands appeared at the expected size (3.7 Kb and 2.6 Kb).

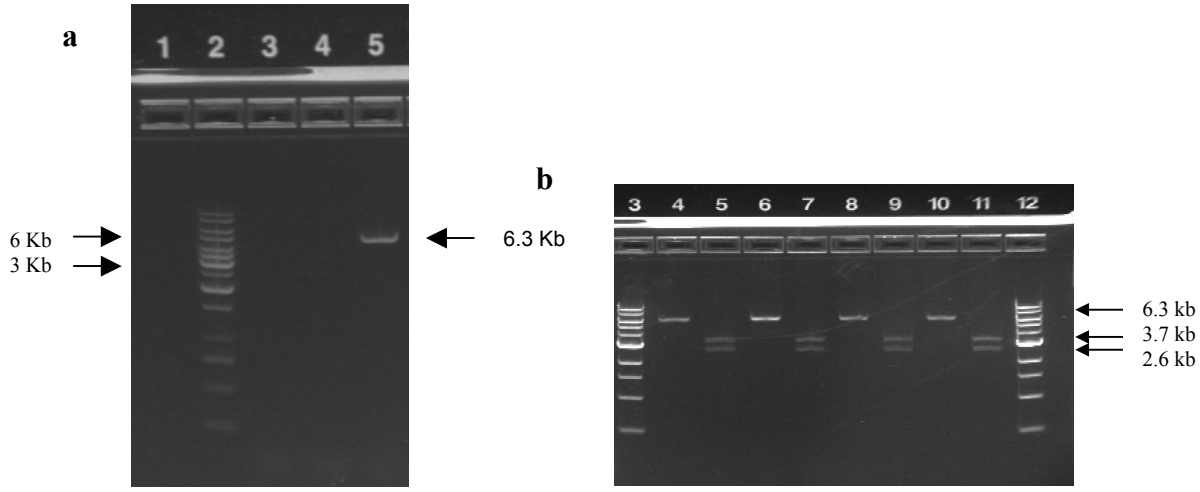


Figure 2: Verification of plasmids. a) PCR products of gene encoding T4 ligase K159A after site directed mutagenesis. Lane 2, 1 kb DNA ladder; lane 3, negative control; lane 5, PCR amplicon. b) Single and double restriction digestion of plasmids. Lanes 4 to 7, plasmids expressing T4 WT ligase; lanes 8 to 11, plasmids expressing mutant K159A ligase; lane 4, 6, 8 and 10, plasmids were cut with EcoRV; lanes 5, 7, 9 and 11, plasmids cut with EcoRV andPvuI. Samples were run in 1.2% agarose gel

The constructs were sent for sequencing to confirm the expected point mutation (**Figure 3a**). After alignment the sequence in Vector NTI advance 10 program, I found that the beginning of the coding region (ATG…) showed perfect sequence for both K159A (2nd line) and T4 WT (3rd line) compared with the reference T4 DNA ligase sequence (1st line). Sequence around the mutation showed that alanine (AAA) and lysine (GCT) were introduced correctly in the K159A and T4 WT clones **(Figure 3b)**. The entire coding regions before and after the mutations were perfect matches with each other and with the original T4 DNA ligase sequence, except for a substitution V286A that turned out to have been present in the original K159L plasmid. This substitution was outside any known catalytic sites and I would expect it to be neutral.

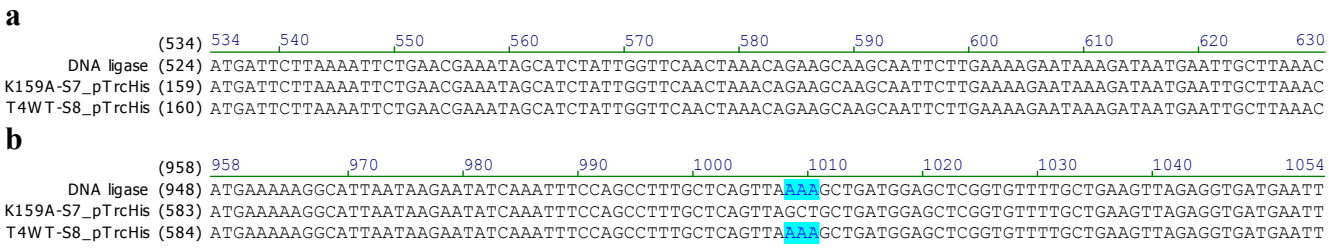


Figure 3: Sequencing of the constructs after site directed mutagenesis. Sequence of the original coding region of T4 DNA ligase starting with ATG codon (GenBank identification NC_000866) and the constructs T4 WT ligase and mutant T4 ligase K159A genes. S7 and S8 indicate sample numbers.

2.2 Protein purification

2.2.1 T4 wild type and mutant (K159A, K159L) proteins

The bacteria were grown for three hours before induction and then induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG) for five hours. The induction was not quite successful and the desired proteins were not overexpreesed. His-tagged T4 DNA ligases were purified on a cobalt affinity column. HisPur cobalt resin columns were used to purify the His-tagged proteins because this cobalt column has no metal contamination during elution. Most of the His-tagged proteins were bound in the presence of 10 mM imidazole and subsequently eluted with high concentration of imidazole (150 mM). In most of the cases almost all His-tagged proteins were obtained in the first 2 ml elution. A trace amount of His-tagged proteins were found in second 2 ml elution. A limited number of few other proteins were also observed to elute along with His-tagged proteins. The eluted proteins were analyzed by polyacrylamide gel electrophoresis. Bands of the expected size (57 KDa for T4 DNA ligase) were obtained from protein preparations **(Figure 4)**.

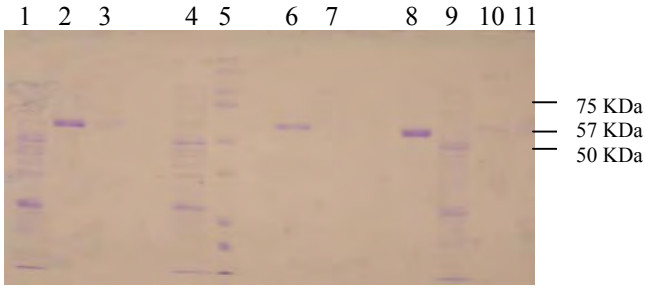


Figure 4: Polyacrylamide gel electrophoresis of T4 WT and mutant (K159A, K159L) proteins. A 10% Sodium dodecyl sulfate poly acrylamide gel electrophoresis with 0.1% Coomassie brilliant blue staining was used to check T4 DNA ligase purified proteins. Lane 1, bacterial (carry gene of WT ligase enzyme) lysate flow through; lanes 2 and 3, His-T4 WT elution 1 and 2; lane 4, bacterial (carry gene of K159A mutant) lysate flow through; lane 5, protein ladder where two bands 50 KDa and 75 KDa are indicated; lanes 6 and 7, His-mutant K159A elution 1 and 2; lane 8, commercial T4 WT DNA ligase protein with a 57 KDa band indicated; lanes 10 and 11, His-mutant K159L elution 1 and 2. The gel was run at 140 volt for 90 min.

Nilsson *et al.* 2001 had used a very high concentration of WT T4 DNA ligase (0.5 U/µl) to ligate DNA on an RNA template. A commercial protein concentrator was used to concentrate proteins **(table 1)**. Since most of the proteins were obtained in elution 1, only proteins in elution 1 were concentrated. During concentration of proteins, storage buffer was used to exchange imidazole with storage buffer. The concentration of the proteins was improved slightly but not so much as I expected (at least10 times).

Table 1: Concentration of different purified enzymes before and after concentration

Proteins	Before concentration (mg/ml)	After concentration (mg/ml)
T4 WT DNA ligase	0.1068	0.2382
Mutant K159A	0.0799	0.1374
Mutant K159L	0.1050	0.0922

2.2.2 Enzyme activity

Wild type T4 DNA ligase can join DNA nicks, but the other two mutants cannot ligate DNA nicks alone. To verify that my purification protocol resulted in active proteins, I compared the activity of purified T4 wild type DNA ligase with commercial T4 DNA ligase. Different dilutions of purified WT DNA ligase and commercial T4 DNA ligase were used. All dilutions tested both purified and commercial wild type T4 DNA ligase (200X dilution was not used in commercial T4 DNA ligase) successfully ligated HindIII treated λ DNA within 30 minutes **(Figure 5)**. In case of 400x dilution of purified T4 DNA ligase, the ligated λ DNA became smeared (ligation started but was not complete) (lane 5) as was found also with the same dilution of commercial T4 DNA ligase (lane 10). This shows that the DNA ligation activity of my purified T4 DNA ligase was similar to the ligation activity of the commercial enzyme. It also demonstrates that the accidental V286A mutation had no effect on ligation efficiency.

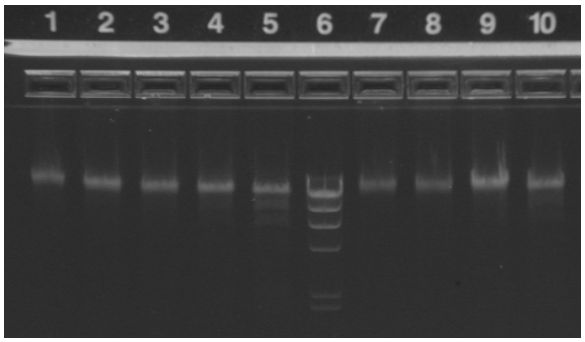


Figure 5: DNA ligation activity of purified T4 DNA ligase. Lanes 1-5, HindIII treated λ DNA reacted with 40X, 80X, 100X, 200X and 400X times dilution of purified T4 WT DNA ligase enzyme respectively; lane 6, HindIII treated λ DNA standard; lanes 7-10 HindIII treated λ DNA reacted with 40X, 80X, 100X and 400X times dilution of commercial WT T4 DNA ligase enzyme respectively. The samples were analyzed in 1.2% agarose gel for 26 min.

2.3 Experimental setup for ligation reaction

I used in vitro ligation assays to monitor the ligation of DNA on a DNA template and DNA on an RNA template at different time intervals. In addition, a single enzyme or a combination of enzymes was used in the reaction mix. This experimental setup is originally described by

Nilsson *et al.* 2001, where a 3′ labeled oligonucleotide and a 5′ unlabeled oligonucleotide are annealed to an RNA template that has a complementary sequence, forming a ligatable nick. One oligonucleotide presents a 5′ phosphate at the nick, and is fluorescently labeled at its 3′ end; the other oligo presents a 3′-OH at the nick and is unlabeled **(Figure 6).** A 3′-labeled [tetramethyl-6-carboxyrhodamine dye (TAMRA)] 5′-phosphorylated oligonucleotide (PHO-Dye), RNA template and 5′-unlabeled oligonucleotide (dARK) were mixed in 1:2:4 molar ratio as ligation substrate. The excess RNA template and 5′ unlabeled oligonucleotide was used to ensure that all labeled oligonucleotides would be able to react. Capillary electrophoresis can distinguish a single nucleotide addition to the labeled oligonucleotide.

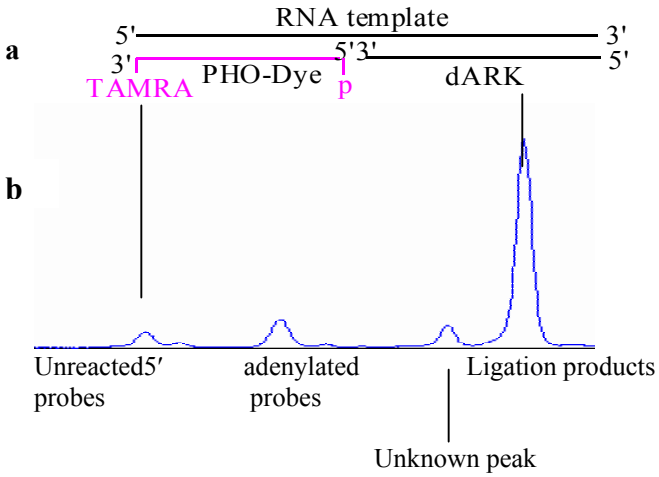


Figure 6: Schematic illustration of the experimental set-up. a) Set up of oligonucleotides and RNA or DNA template. For further details, see text. **b)** The three different ligation reaction products were identified in capillary electrophoresis. An unknown peak was identified in each experiment just before ligation products.

2.3.1 Ligation of DNA nick on DNA and RNA templates by WT T4 DNA ligase

The rates of nick sealing reactions in DNA by T4 DNA ligase on a DNA template and DNA on an RNA template are faster at lower concentration of ATP (Sekiguchi *et al.*1997, Nilsson *et al.* 2001). I observed the rate of ligation of a single nick in DNA on a DNA template and DNA on an RNA template by T4 DNA ligase in the presence of 10 µM ATP and 10 mM Mg^{2+}. Products were analyzed by capillary electrophoresis. 36% ligated labeled probe was detected within 15 seconds of the start of the ligation reaction on the DNA template and almost 70% ligation was completed within 15 minutes. On the other hand, no ligated labeled probe was detected until 15 min on the RNA template. A small amount of ligation product template was detected one hour after start of the reaction **(Figure 7)**.

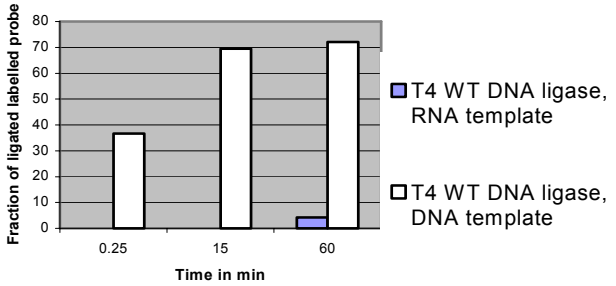


Figure 7: The ligation rate on DNA and RNA templates. The reaction mix contained 10 µl oligonucleotides (40 nanomolar), 10 µl T4 WT DNA ligase enzyme (0.5 U/µl on RNA and 0.005 U/µl on DNA template) 10 µM ATP and 10 mM Mg^{2+}.

2.3.2 Ligation of DNA nick on an RNA template by a combination of T4 WT and K159A DNA ligase

According to Rossi *et al*. (1997), the ligation of DNA on an RNA template by T4 DNA ligase is slow due to the ATP concentration exceeding the K_m for ATP. One of the mutants (K159A) was combined with T4 WT DNA ligase to increase the ligation. The K159A mutant enzyme alone was completely inactive in the ligation assay (data not shown), likely due to its inability to bind AMP in step I of the reaction. The combination of K159A with T4 DNA ligase resulted in a small fraction of ligated labeled probe after 15 min, whereas no labeled probe was detected in the reaction with T4 DNA ligase alone at this time (Fig 8, of Fig 7). Around 35% ligated labeled probe was found within one hour of reaction with the combination of WT and K159A DNA ligase, which was almost 30% more ligated labeled probe with the combination of enzymes than with WT DNA ligase alone.

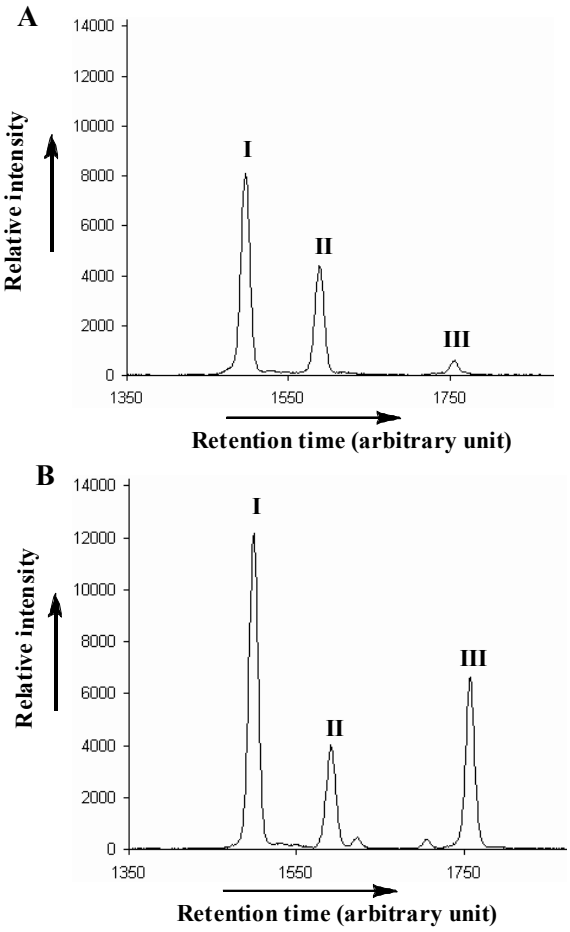


Figure 8: Kinetics of ligation of DNA on an RNA template. a) Ligation of DNA on an RNA template by T4 WT DNA ligase alone at 1 hr. b) Ligation of DNA on an RNA template by a combination of T4 WT and K159A DNA ligase at 1 hr. The final concentration of WT T4 DNA alone or in combination with the K159A mutant was 0.5 U/µ The final concentration of K159A ligase was 0.1 U/µl (approximately). The peak I show unreacted probes, the peak II (closest to first peak) adenylated probes and the peak III fully ligated products. Some unknown additional peaks are also seen.

On the other hand, a significant portion of nicked DNA was ligated after 4 hr in the presence of only WT T4 DNA ligase and 80% of the labeled probe was ligated at this time. The fraction of the labeled probe ligated by the combined enzymes changed very little from 1 h to 18 h (**Figure 9**).

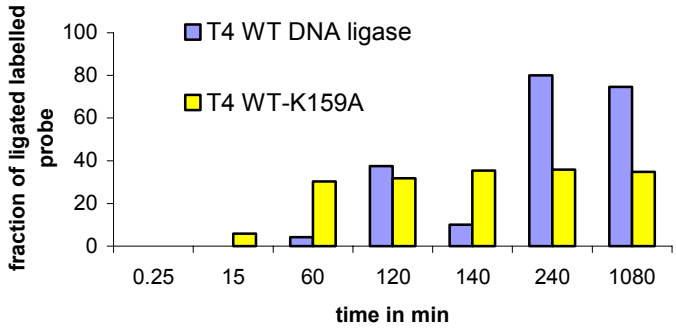


Figure 9: Effect of K159A mutant ligase on ligation reaction of DNA on an RNA template. In this graph, gray and white bar represent ligation products of DNA on RNA template where a combination of enzymes (T4 WT and K159A) or T4 WT DNA ligase were used respectively. The reaction conditions were same as described before.

3. DISCUSSION

It was recognized early on that the T4 DNA ligase can ligate DNA oligonucleotides hybridizing to RNA strands (Kleppe et al. 1970). However, no analysis has been presented of optimal reaction conditions for RNA-templated DNA ligation. My results show a promising tool to detect an RNA target molecule efficiently via ligation of oligonucleotides by T4 DNA ligase. During expression of WT and mutant DNA ligase enzyme, a few of other proteins also induced along with that desired enzymes. A second purification using Sephadex G 75 could have been done to obtain more purified enzyme. An additional mutation was found in the coding region in all versions of the cloned T4 DNA ligase gene. The normal ligation activity of this enzyme proved that this additional mutation has no effect on ligation mechanism of DNA ligase enzyme.

3.1 Ligation reaction

After peak analysis resulting from capillary electrophoresis the second peak was from the adenylated probe. But the peak was shifted more than expected for a single nucleotide added to the 5′ position of the labeled oligonucleotide due to unknown reasons. To identify adenylated probe, the protein, aprataxin, have been used. Aprataxin can remove AMP from the adenylated-nick DNA intermediate and back to non-ligated DNA form (Ahel *et al.* 2006). Finally the third peak originated from the complete ligation product. An unknown peak in capillary electrophoresis was always detected just before the ligation products. I assumed from the position of the peak (see **Figure 6**) that it contained labeled oligonucleotide that joined to each other in ligation substrate mix and gave signal as ligation products. Some scattered peaks were also detected in capillary electrophoresis and that might be due to loading problem in capillary electrophoresis. According to Nilsson *et al.* (Nilsson *et al.* 2001), the ligation reaction of DNA on RNA template by T4 DNA ligase was efficient when DNA ligase was used in molar excess over oligonucleotide substrate, the ATP concentration was kept low, 10 mM magnesium or manganese was used and sufficient time provided for ligation. After 2 h of reaction, the ligation efficiency of DNA on RNA template by WT DNA ligase was increased (**Figures 8 & 9**). However, the extent of ligation by combined ligases did not increase further, while the T4 WT yielded 80% completely ligated probe after 4 h. The mutant enzyme may not have been completely purified and phosphatase might be present in the purified enzyme (western *et al.* 1991) that removed the phosphate group from the 5′-labeled DNA. This would have caused dephosphorylation of a fraction of the labeled templates, which therefore would not be able to participate in a ligation. The experiment will be repeated with more highly purified enzyme to test this hypothesis

3.2 Application and hypothesis

Genes are always expressed to mRNA first and then translated to proteins. In applications involving RNA, it would be desirable to avoid conversion of RNA into cDNA before detection/amplification reaction, i.e. to ligate a detection probe (DNA) directly to RNA. Direct analysis of RNA sequences without a preceding cDNA synthesis steps may more faithfully report the relative abundance of specific mRNA in a single cell or in cellular extracts. The use of padlock probe to ligate with a targeted mRNA directly to quantify gene expression is not possible due to the slow ligation rate of DNA probe with RNA. In my method, a combination of T4 WT DNA ligase and mutant T4 ligase can ligate a DNA probe with RNA very efficiently. By using ligase-mediated probe circularization (padlock probe), reacted probes can be replicated through rolling circle amplification, doing quantitative and

sensitive detection of RNA sequence variants. By quantifying the mRNA, gene expression can also be quantified in a single cell or insitu using padlock probe and rolling circle amplification Moreover, ligase-assisted probe ligation could be used to differentiate members of gene families more accurately compared to traditional hybridization-based analyses. If the hypothesis, ligase mediated RNA detection using DNA probe by a combination of T4 WT DNA and mutant T4 ligase, is true, then a powerful novel hybrid ligase enzyme will be produced. Such a ligase could be used to detect single nucleotide polymorphism directly on RNA, to quantify mRNA abundance by ligase dependent amplification and to detect single mRNA molecules in tissue sections using padlock probe ligation and rolling-circle amplification.

4. MATERIALS AND METHODS

4.1 Plasmid and oligonucleotides

The sequence of K159L T4 DNA ligase had been cloned in expression vector pTrcHis (**Figure 10**). The construct (pTrcHis vector + gene of K159L T4 DNA ligase) was the generous gift of Dr. Alessandra Montecucco (Rossie et al. 19997). *E. coli* TOP10 (Invitrogen) cells were used to expressed the constructs.

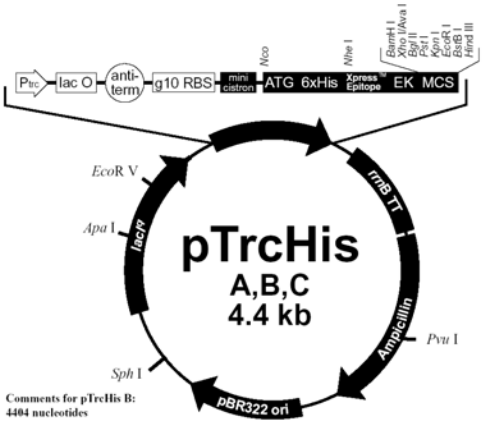


Figure 10: pTrcHis plasmid: The map of the plasmid reproduced from the Invitrogen site with kind permission.

A list of mutants and corresponding forward and reverse primers for site directed mutagenesis and oligonucleotides used for ligation reaction are listed in the **Table 2**. All primers were phosphorylated at the 5′ end.

Table 2: Mutant, oligonucleotides and forward and reverse primers used for site directed mutagenesis

Use	Oligonucloetide sequence[1]	Modification	Source
K159L	F 5′ GCTCAGTTAACTGCTGATGGA 3′ R 5′ AAAGGCTGGAAATTTGATATTCTT 3′		Rossie *et. al.* 1997
K159A	F 5′ GCTCAGTTA**GCT**GCTGATGGA 3′ R 5′AAAGGCTGGAAATTTGATATTCTT 3′	5′ phosphate	Invitrogen, USA
Wild Type	F 5′ GCTCAGTTA**AAA**GCTGATGGA 3′ R 5′ AAAGGCTGGAAATTTGATATTCTT 3′	5′ phosphate	Invitrogen, USA
PHO-Dye	5′ GCCTTATGCAGTT 3	3′ TAMRA 5′ phosphate	Eurofins MWG GmbH
dARK	5′-GCGTATCTCTTCATA- 3′		Eurofins MWG Gmb
RNA template	5′AACUGCAUAAGGCUAUGAAGAGAUACGC 3′	RNA	Eurofins MWG GmbH
DNA template	5′ AACTGCATAAGGCTATGAAGAGATACGC 3′		Invitrogen, USA

[1]Nucleotides shown in bold in the forward primer was the sequence to be changed by site directed mutagenesis. Altered nucleotides are underlined in the forward primer where the desired mutation was introduced.

4.2 Site directed mutagenesis

Site directed mutagenesis was performed on the plasmid pTrcHis (**Figure 3**) encoding the K159L mutant T4 DNA ligase. The gene encoding the K159L ligase was amplified using two different pairs of oligonucelotides (**Table 2**) to produce mutant K159A and wild type T4 DNA ligase. The 50 µl master mix contained 35.5 µl of ddH_2O, 10 µl of 5x Phusion buffer, 1 µl of 10 mM dNTPs, 1 µl of 25 µM forward primer, 1 µl of 25 µM reverse primer, 1 µl of 5 ng/µl template DNA and 0.5 µl of 2 U/µl Phusion hot start DNA polymerase. All reagents except primers were from New England Biolabs Inc. The sample was mixed and amplified by PCR as follows: 98°C 30 sec, 25 cycles (98°C 10 sec, 64°C 30 sec, 72°C 3 min), 72°C 7 min. The PCR products were circularized with commercial T4 DNA ligase (Fermentas) using a reaction mix containing 2 µl of PCR product from mutagenesis reaction (~25 ng of PCR product), 5 µl of 2x Quick ligation buffer (Fermentas) and 0.5 µl of T4 DNA ligase (Fermentas). Ligation products were transformed into one shot chemically competent *E. coli* TOP10 (Invitrogen) and plated on a Luria agar (LA) plates (1% tryptone, 0.5% yeast extract, 1% NaCl, 1.5% agar in water, pH 7.0), prepared by lab technician supplemented with 100 µg/ml ampicillin (Sigma). The plates were incubated overnight at 37°C. Plasmid DNA was purified from bacteria using the QIAprep spin miniprep kit (Qiagen) as described by the manufacturer.

4.3 Restriction digestion and gel electrophoresis

Restriction enzymes EcoRV (Fermentas) and PvuI (New England Biolabs Inc.) were used for single and double digestion. A 20 µl reaction mixture containing 2 µl of 10x NEB3 buffer (New England Biolabs Inc.), 2 µl of 10x BSA (New England Biolabs Inc.), 5 µl of plasmid DNA (~50 ng DNA), 1 µl of restriction enzymes (1 µl of each restriction enzymes in case of double digestion) and 10 µl ddH_2O (make total volume 20 µl) was prepared and mixed.

PCR, ligation and restriction digestion products were analyzed along with a 1 kb DNA ladder (Fermentas) in a 1.2% agarose E-Gel (Invitrogen). The samples were run for 26 min using the E-Gel iBase power system (Invitrogen). Gel pictures were taken with a High performance ultraviolet transilluminator (Scion Corporation).

4.4 Sequencing

Plasmids DNA was sent for sequencing to Eurofins MWG, Germany. Sequences were aligned and analyzed using Vector NTI® Advance 10 software (Invitrogen).

4.5 Purification of recombinant protein

4.5.1 Wild type and mutant T4 DNA ligase protein

Recombinant protein was produced in *E. coli* Top10. 5 ml of overnight bacterial culture in LB (10 gm bacto-trypton, 5 gm bacto-yeast extract, 10 gm NaCl in 1 liter water and then autoclave) with 100 µg/ml ampicillin were added to 400 ml of LB medium containing 100 µg/ml ampicillin and grown at 37°C with vigorous shaking to an OD_{600} of 0.6-0.7. Then IPTG (Sigma) was added to 2 mM final concentration to induce the expression of recombinant protein for 5 hours. Then cells were harvested by centrifugation at 10000g for 10 min. After this step, all remaining steps of protein purification were performed at 4°C. The

pellets (bacterial cells) were resuspended in 20 ml of lysis buffer [20 mM potassium phosphate buffer pH 7.0, 1 mM phenylmethanesulfonyl fluoride solution (PMSF) (Sigma), 1 μg/ml pepstatin (Sigma)]. Resuspended pellets were frozen in liquid nitrogen and the frozen lysate thawed at 42°C. The freeze-thawing was repeated 3 more times and then the samples were centrifuged at 16000g for 30 min.

The clear supernatant was removed and purified using a HisPur Cobalt spin column (Pierce) according to the instructions from the manufacturer. The lysate was loaded into a HisPur cobalt column and washed with 2 ml of washing solution [1x Phosphate Buffered Saline (PBS), 10 mM imidazole] 6 times until the absorbance of the flow-through fraction at 280 nm reached baseline. PBS was made by dissolving 8 g of NaCl, 0.2 g of KCl, 1.44 g of Na_2HPO_4, 0.24 g of KH_2PO_4 in 800 ml distilled H_2O. pH was adjusted to 7.4 and volume made upto to 1L with additional distilled H_2O. The solution was then sterilized by autoclaving. Finally the protein was eluted with 2x2 ml of elution solution (1x PBS, 150 mM imidazole). The eluate was dialysed against storage buffer [20 mM Tris-HCl pH 7.4, 100 mM KCl, 0.2 mM EDTA and 2 mM dithiothreitol (DTT)] and concentrated to a volume of 200 μl using an iCON concentrator 7ml/9K (Pierce) as instructed by manufacturer.

4.6 Sodium dodecyl sulphate poly acrylamide gel electrophoresis (SDS PAGE)

Purified proteins were analyzed by SDS PAGE using 10% acrylamide-bisacrylamide (29:1) gel (National Diagnostics). Samples were prepared by mixing equal volume of protein and 2x Laemmli buffer (4 % SDS, 20 % (v/v) glycerol, 10 % (v/v) 2-mercaptoethanol, 0.004% (w/v) bromophenol blue and 0.125 M Tris HCl pH 7.5) and boiled at 95°C for 5 min. 18 μl of each sample was loaded into the gel and 3 μl molecular weight marker (Bio-Rad) was run on the same gel. The electrophoresis was performed for 80 min at 150 volt. The gel was stained with 0.2% Coomassie Brilliant Blue (Sigma) and then destained with destaining solution (methanol: water: acetic acid in a 25:65:10 ratio).

4.7 Determination of protein concentration

Protein concentration was determined using Bradford reagent (Sigma) with bovine serum albumin (New England Biolabs Inc.) as the standard. 5 μl of protein solution were mixed with 250 μl of Bradford reagent in a microtiter plate (Applied Biosystems) and read in the analyzer after 20 min.

4.8 Ligase activity determination

4.8.1 Wild type DNA ligase

Ligation mix was prepared by mixing 1 μl of HindIII treated λ DNA (Finnzymes), 4 μl of polyethylene glycol 8000 containing 5x ligation buffer (Invitrogen), 1 μl of different dilutions of either commercial (Fermentas) or purified wild type DNA ligase and made up to 20 μl with water. Both purified enzyme and commercial enzymes were diluted in 1x NEB3 buffer (New England Biolabs Inc.). The reaction mix was mixed thoroughly and incubated for 30 min at room temperature followed by heat inactivation at 65°C for 10 min. 20 μl sample were analyzed in 1.2% agarose E-Gel (Invitrogen) and run for 40 min.

4.8.2 Nick sealing activity of ligase enzyme between DNA and RNA template

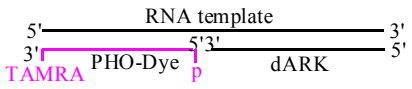


Figure 12: **Schematic illustration of the experimental setup.**

Ligation substrates were made in 1x ligation buffer (10 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 µM ATP) by combining the 3´ oligonucleotide labeled with TAMRA, RNA template and the 5´ oligonucleotide dARK (**Figure 12**) at 40, 80 and 120 nanomolar respectively (molar ratio 1:2:4).

The ligation substrates were mixed gently and incubated at 65°C for 3 min and cooled to room temperature and then transferred to ice. Enzymes and enzyme mixes were prepared separately for each ligation reaction. Commercial T4 DNA ligase (CWT) and a ligase mixture of CWT with K159A were prepared at final concentrations of 0.5 U/µl of CWT ligase and 0.1 U/µl of K159A ligase.

For the time course experiments, the reaction between ligation substrates and enzymes were started in the optical 96 well reaction plate (Applied Biosystems). In the reaction, 10 µl of substrate and 10 µl of enzyme or enzyme mix were mixed and 2 µl of sample was withdrawn from the reaction at different time intervals. Ligation reactions were terminated by adding 2 µl of ligation reaction samples into the 18 µl of 1 mM formamide and 2 mM EDTA (pH 8.0). After sampling, the samples were diluted again by taking 2 µl of samples to 23 µl of 1 mM formamide in a barcoded optical 96 well reaction plate (Applied Biosystems). The samples were mixed and kept at 4°C until analyzed by capillary electrophoresis. The samples were then sent to the structural biology lab of Medical Biochemistry and Biophysics Department (MBB) of Karolinksa Institutet. The capillary electrophoresis machine, ABI3130xl (Applied Biosystems) has 16 capillary arrays and can run the sequence of 16 samples at a time in 2 columns i.e. columns 1+2, 3+4 and 5+6 etc. of 96 well plates. After getting the result from capillary electrophoresis, the sequencing results converted to different peaks by using an internal visual basic program. As a control of the ligation reaction, ligation substrates were also prepared using DNA template in place of the RNA template, and in the same ratio with the other two nucleotides.

5. ACKNOWLEDGEMENT

I would like to express my special gratitude to Sten Linnarsson, Ph.D. for giving me chance to work with his group. He trained me the required techniques of molecular biology in my project starting from the beginning to ins and outs in every possible way. Sten's friendly attitude during my project work always inspired me to do this research work without any tiring. I am also grateful to my supervisor, Sten for managing time to read and correct my thesis that took considerable amount of his summer vacation time.

Very special thanks to Dr. Alessandra Montecucco (Rossi *et al.* 1997) for a generous gift of pTrcHis vector carrying a muatant sequence that was our starting of this project.

I also specially thanks to Arno Pihlak who taught me some basic knowledge of research and guiding me through difficult concepts and statistical analysis. I am also grateful to Ats Metsis in my lab who guided me through out my whole project by sharing his research experience. Also thanks to Una Kjällquist to give me valuable advice during my lab work.

I also thanks to Ann-Sofie Nilsson of Matrix biology department in MBB to run the Capillary electrophoresis.

I would like to give thanks all of the lab members of Molecular neurobiology division to create a friendly and working environment in lab.

It would be completely incomplete to finish my acknowledgement without thanking and loving my wife, Tayeba and my one-half year daughter, Tuba. Sometimes their presence in Sweden forced me to finish my lab work early but I refueled for next working day after spending time with them.

Thank you all.

6. REFERENCES

Ahel I., Rass U., El-Khamisy S. F., Katyal S., Clements P. M., McKinnon P. J., Caldecott K. W. &. West S. C. 2006. The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates. Nature. **443:** 713-716

Alberts B., Johnson A., Lewis J., Raff M., Roberts K., and Walter P. Molecular Biology of the Cell 2002. 4th edition, Garland science, New York.

Cherepanov A.V., and Vries S. 2003. Kinetics and thermodynamis of nick sealing by T4 DNA ligase. Eur. J. Biochem. **270:** 4315-4325

Higgins N.P., and Cozzarelli N.R. 1979. DNA-joining enzymes: a review. Method Enzymol. **68:** 50-71

Kleppe K., van de Sande J.H. and Khorana H.G. 1970. Polynucleotide ligase-catalyzed joining of deoxyribo-oligonucleotides on ribopolynucleotide templates and of ribo-oligonucleotides on deoxyribopolynucleotide templates. Proc. Natl. Acad. Sci. USA **67**: 68–73

Kornberg A., and Baker T.A. 1991. DNA replication. W.H. Freeman and Co., New York, NY.

Lehman I.R. 1974. DNA ligase: structure, mechanism, function. Science **186:** 790-797

Lindhal T., and Barnes D.E. 1992. Mammalian DNA ligases. Annu Rev Biochem. **61:** 251-281

Nilsson M., Antson D., Barbany G., and Landegren U. 2001. RNA-templated DNA ligation for transcript analysis. Nucleic Acid Research. **29:** 578-581.

Pascal J.M. 2008. DNA and RNA ligases: structural variations and shared mechanisms. Curr. Opinion in struc. Biol. **18:** 96-105

Rossi R., Montecucco1 A., Ciarrocchi G., and Biamonti G. 1997. Functional characterization of the T4 DNA ligase: a new insight into the mechanism of action. Nucleic Acids Research. **25:** 2106–2113

Sekiguchi J., and Shuman S. 1997. Ligation of RNA-containd duplexex by vaccinia DNA ligase. Biochemistry **36:** 9073-9079.

Sekiguchi J., and Shuman S. 1997. Nick sensing by vaccinia virus DNA ligase requires a 5′ phosphate at the nick and occupancy of the adenylated binding site on the enzyme. Journal of Virology. **71:** 9679-9684.

Shuman S., and Schwer B. 1995. RNA capping enzyme and DNA ligase: A superfamily of covalent nucleotidyltransferases. Mol. Microbiol. **17:** 405-410

Western L.M., and Rose S.J. 1991. A novel DNA joining activity catalyzed by T4 DNA ligase. Nucleic Acids Research. **19:** 809-813