AN ABSTRACT OF THE THESIS OF

Xiaolin Zhang for the degree of Doctor of Philosophy in Genetics presented on April 14, 1995. Title: Mutagenic Mechanisms Associated with DNA Cytosine Methylation, DNA Base Sequence Context and DNA Precursor Pool Asymmetry.

Abstract approved: Redacted for Privacy

Christopher K. Mathews

ABSTRACT

A major goal of genetics is to elucidate causes and mechanisms of mutation. Mutagenic mechanisms associated with DNA cytosine methylation, DNA base sequence context, and DNA precursor pool asymmetry were studied in this research. For the first time, the deamination rate for 5-methylcytosine residues in duplex DNA was measured. Under conditions approximating the biological situation, the rate of spontaneous deamination at G·mC exceeds that at G·C by more than 21-fold, indicating chemical instability toward deamination as a major causal factor relating DNA-cytosine methylation to spontaneous mutagenesis.

A major source of spontaneous mutations is errors made during DNA replication. Using a recombinant phagemid containing a sensitive mutagenic target, I investigated errors arising during replication of that phagemid in HeLa cell extracts. I found that natural DNA precursor pool biases observed in

the nuclei of human cells are mildly but significantly mutagenic, and the degree of this natural mutagenic effect is partly dependent on DNA base sequence context. DNA sequence context significantly affects the effectiveness of proofreading activity of DNA polymerase. A GC-rich sequence flanking a site of mutation confers great sensitivity to DNA precursor pool imbalance and the presence of a proofreading inhibitor. Apparently, DNA base sequence context and pool asymmetry both contribute to the natural fidelity of DNA replication.

Finally, the role of nucleoside diphosphate kinase in dNTP synthesis was investigated. *Escherichia coli* cells with a *ndk* deletional mutation showed an abnormally higher spontaneous mutation rate. I found that the dCTP pools in these cells were some 20-fold higher over normal, a sufficient bias to account for the observed mutator phenotype. Mutant cells can still support T4 infection both aerobically and anaerobically, although T4 phage must rely on the host to provide NDP kinase activity. Cells with disrupted pyruvate kinase genes failed to grow anaerobically, but further deletion of *ndk* in these cells makes them viable again in anaerobiosis, suggesting that nucleoside diphosphate kinase may play important regulatory roles in cells in addition to the synthesis of nucleoside triphosphates.

Mutagenic Mechanisms Associated with DNA Cytosine Methylation, DNA Base Sequence Context and DNA Precursor Pool Asymmetry

by

Xiaolin Zhang

A THESIS

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Doctor of Philosophy

Completed March 30, 1995

Commencement June 1995

Doctor of Philosophy thesis of Xiaolin Zhang presented on April 14, 1995

APPROVED:

Redacted for Privacy

Major Professor, representing Biochemistry and Biophysics

Redacted for Privacy

Director of Genetics program

Redacted for Privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Xiaolin Zhang, Author

To Mr. Zhongping Zhang:

my friend, my teacher, and my father.

ACKNOWLEDGMENTS

I am grateful to Christopher K. Mathews for providing me an intellectually stimulating environment and always encouraging me to pursue my scientific curiosities. I am fortunate to have an opportunity to observe a person who is dedicated to education and science. Dr. Mathews designed the DNA constructs and carried out the forward mutation assay described in Chapter 3. He also performed the anaerobic growth experiments and prepared anaerobic dNTP extracts for the work described in Chapter 4 of this thesis.

I wish to thank members of the Mathews research group for many helps and ideas. I also wish to thank my graduate committee members for their advice: Drs. Tony Chen, George D. Pearson, David Barnes, and Dale Mosbaugh.

My deepest appreciation is felt for the love, tolerance, patience, and support of my wife, Lian Yu. My appreciation is extended to her family and my uncle's family. Without their love and support, I would not be able to finish my graduate education here at Oregon State University.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

Mutagenic Mechanisms Associated with DNA Cytosine Methylation, DNA Base Sequence Context and DNA Precursor Pool Asymmetry

Chapter 1

Introduction

Maintaining the integrity of genetic materials has to be a top priority for organisms to survive. For a species to exist, it must pass on its genetic information with great accuracy. During evolution, cells have acquired various mechanisms to safeguard their genetic materials and replicate them with extremely high fidelity. On the other hand, without change of genetic information, there would be no source of variation, therefore no basis for selection and evolution. Change of genetic materials, or mutations, may arise spontaneously, that is, in the absence of any external agent. Most spontaneous mutations are the consequence either of errors made during DNA replication, or DNA damage generated from natural decay and reactive molecules produced during cellular metabolism (Kunz et al., 1994). A major goal in mutation research is to understand the causes of mutagenesis, and their effects on cellular function and molecular evolution.

MUTAGENIC EFFECT OF DNA CYTOSINE METHYLATION.

1. Biological significance of DNA cytosine methylation.

DNA base modification is an important source of mutations. One major type of DNA modification is DNA methylation. DNA methylation

plays an important role in many biological functions. In most eukaryotic cells, only cytosine is methylated. DNA cytosine methylation occurs mainly at CpG dinucleotides. In fact, most of the CpGs in mammalian genomes are methylated (Antequera and Bird, 1993). The significance of DNA cytosine methylation was directly demonstrated by Li et al (1992), when they showed that functional DNA methyltransferase, which catalyzes DNA cytosine methylation reaction, was essential for development in mice. Even partial disruption of this gene resulted in embryonic lethality.

DNA methylation is usually associated with suppression of gene expression. All known housekeeping genes and many tissue-specific genes have CpG islands at their 5' promoter regions (Larsen et al., 1992). CpG islands are stretches of DNA sequence of about 1 kb long, usually GC-rich; they showed the expected frequencies of CpG dinuleotide (Antequera and Bird, 1993). In a recent survey by Yeivin and Razin (1993), tissue-specific genes showed a correlation between gene activity and hypomethylation of the promoter region, while CpG islands in housekeeping genes have never been found to be methylated. In many permanent cell lines, ten of thousands of genes have been stably inactivated, owing to *de novo* methylation of CpG-island promoters that are normally methylation-free (Antequera et al., 1990). Only genes that are essential for life in culture conditions are spared from methylation (Bird, 1992).

Accumulating evidence indicates that repression of gene expression by DNA methylation is mediated through methyl-CpG binding proteins (MeCPs) (Lewis et al., 1992). The binding of MeCPs to the promoter regions could interfere with the binding of transcriptional machinery and so inactivate gene expression. The density of methyl-CpGs near promoters is a crucial determinant of repression, but full activity or complete repression also

depends on the strength of the promoter. The latter has been shown nicely by Boyes and Bird (1992). Weak promoters were fully repressed by sparse methylation, while strengthened promoter restored transcription without methylation loss. If the density of methylation was increased, even strengthened promoter could not prevail.

DNA cytosine methylation patterns change during normal development, and abnormal changes in methylation patterns have been seen as a characteristic trait in a great variety of tumor cells (Laird and Jaenisch, 1994). Cytosine methylation is "cleared" at the time of fertilization, and the somatic methylation pattern is established by a wave of *de novo* methylation at the time of implantation. Loss of methylation is usually thought to be the result of two rounds of replication without methylation. However, replication-independent enzymatic demethylation has recently been observed in chicken and mouse embryos (Jost, 1993; Kafri et al., 1993), and several types of cells in tissue culture (Vairapandi and Duker, 1993). A final wave of sex-specific *de novo* methylation occurs during gametogenesis (Razin and Cedar, 1993). An important biological phenomenon related to this methylation pattern change is genomic imprinting: the expression of a gene depends on whether it is inherited from the mother or the father. The difference is believed to be due to gamete-specific DNA methylation, which persists through the development after fertilization. In embryos homozygous for a DNA methyltransferase mutation, the expression pattern of three known imprinted genes, *H19*, insulin-like growth factor 2 (*Igf-2*) and *Igf-2* receptor (*Igf-2r*), was altered, and the normal imprinting pattern was lost (Li et al., 1993).

2. Mutagenic consequence of DNA methylation

Although DNA methyltransferase is essential for normal development in mice, DNA methylation adds a substantial mutagenic pressure to vertebrate genomes. High incidence of C → T or G → A transitions has long been known to be associated with cytosine methylation (Coulondre et al., 1978). The observed specific mutagenic pathway has been thought to be the result of deamination from 5-methylcytosine (Fig. 1.1). Cytosine methylation in vertebrates mainly takes place at CpG sites; as a result, CpG dinucleotide in vertebrate genomes is severely depleted, while the corresponding mutation products, TpG and CpA are over-represented (Bird, 1980). In the human genome, 5-methylcytosine represents less than 1% of DNA bases, yet 30-40% of human genetic disorders and cancers caused by point mutations are related to base substitution at 5-methylcytosines (Cooper and Youssoufian, 1988; Laird and Jaenisch, 1994).

3. Evolutionary significance of DNA methylation.

Most animals avoid this problem by keeping cytosine methylation away from their own DNA. For example, geneticists' favorite animal *Drosophila* does not methylate its DNA (Hergersberg, 1991). The vertebrates, however, have allowed DNA methylation to spread throughout their genomes. Could this be the evolutionary price that vertebrates have to pay, or is the mutational load outweighed by other evolutionary advantages? The evidence that DNA methylation is essential for mouse development suggests that vertebrates acquired this base modification by no accident.

Genetic studies with the fungus *Neurospora crassa* reveal that DNA methylation plays a pivotal role in keeping genome stability. *Neurospora*

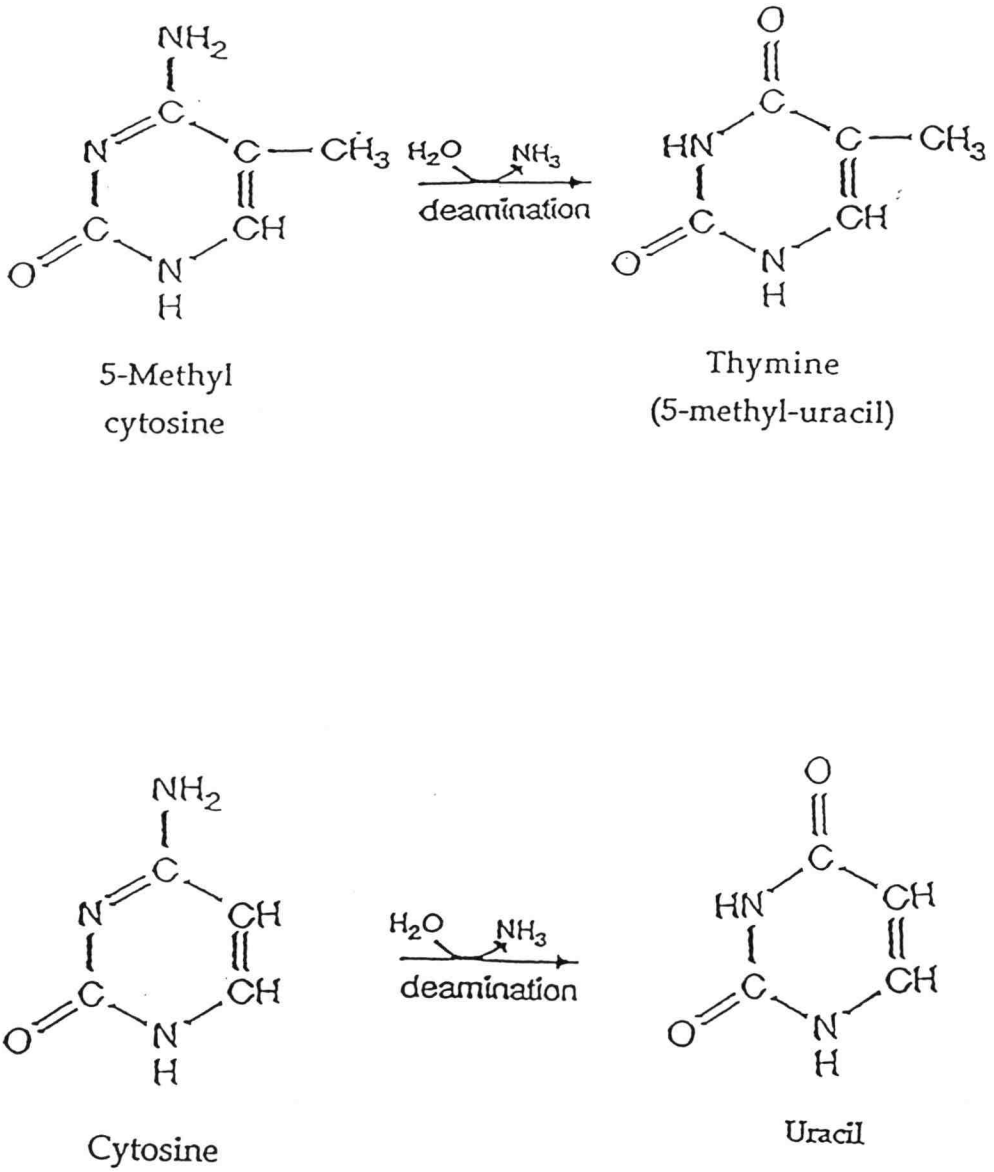


Figure 1.1. Mutagenic deamination pathways of DNA cytosine

crassa has very low abundance repetitive sequence in its genome (Selker, 1990). Artificial repetitive sequence can be generated by transforming *Neurospora crassa* with a plasmid carrying direct duplication of a 0.8-kb marker. The transforming DNA could integrate into host genome through recombination, and this integration was generally stable during vegetative growth. In striking contrast, it became very unstable during meiotic division. In some cases, over 80% of the progenies from a cross lost the duplication marker as a result of homologous recombination (Selker et al., 1987). The vast majority of the progenies (up to 100%) with the duplicated marker was found to have exclusive G·C to A·T transition mutations in the duplication region. Although the bulk of the Neurospora genome is unmethylated, most or all the sites examined in the duplicated sequence were methylated. When a single copy of the marker was integrated, no mutation and methylation occurred (Selker 1990). These results indicate that repetitive sequences were specifically recognized and disrupted through either homologous recombination or repeat-induced point mutation, RIP. Similar processes have been found in other lower eukaryotes (Selker, 1990). Vertebrate genomes commonly consist of about 50% repeated sequences such as mobile genetic elements and pseudogenes (Selker, 1990), and they are facing the threat of chromosomal scrambling by recombination between repetitive sequences. Studies have shown that reducing the sequence similarities by only a few percent can sharply decrease homologous recombination (Radman, 1988). Kricker et al., (1992) proposed a model that DNA methylation and subsequent mutation at CpG dinucleotides is a defensive strategy that vertebrates evolved to specifically mark and diversify duplicated sequences and thereby to protect against recombination-mediated chromosome rearrangement.

GENETIC CONSEQUENCE OF DNA PRECURSOR ASYMMETRY

The spontaneous mutation rate in microbial and mammalian cells has been estimated to be less than 1 mutation per billion base pairs per cell division (Drake, 1991; Loeb, 1991). This extraordinary accuracy is the combined result of at least three processes. First, the DNA replication apparatus selects correct nucleotides in accordance with Watson-Crick base-pairing rules before the coming nucleotide is added to the terminus of the growing chain at the replication fork. Second, if a mistake is made, DNA polymerase-associated exonuclease can come back and excise the incorrect nucleotide, a process called proofreading. Third, those errors that escape proofreading and are fixed in the duplex DNA may be corrected by post-replicational repair system. The effectiveness of each of the above three processes depends on a balanced supply of deoxyribonucleoside triphosphates (dNTPs). If cellular dNTP levels are perturbed, a wide range of genetic consequences may follow due to either aberrant DNA replication or impaired DNA repair. Those genetic consequences include mutation, recombination, enhanced sensitivity to mutagens/carcinogens, or chromosomal breakage (Kunz et al., 1994). Thus, cellular dNTP levels, both relative and absolute levels, are carefully regulated. A balanced synthesis of four dNTPs is the critical determinant in this regulation (Figure 1.2). This is achieved through delicate control of enzymes involved in dNTP synthesis (Mathews, 1989). Ribonucleoside diphosphate reductase, catalyzing the conversion of ribonucleoside diphosphate to doxyribonucleoside diphosphate, is one of the key enzymes in this regulation. The four activities of this enzyme must be balanced by the cell's need for the respective products. dCMP deaminase, as a branch point for pyrimidine synthesis, is allosterically regulated accordingly. CTP synthetase controls the

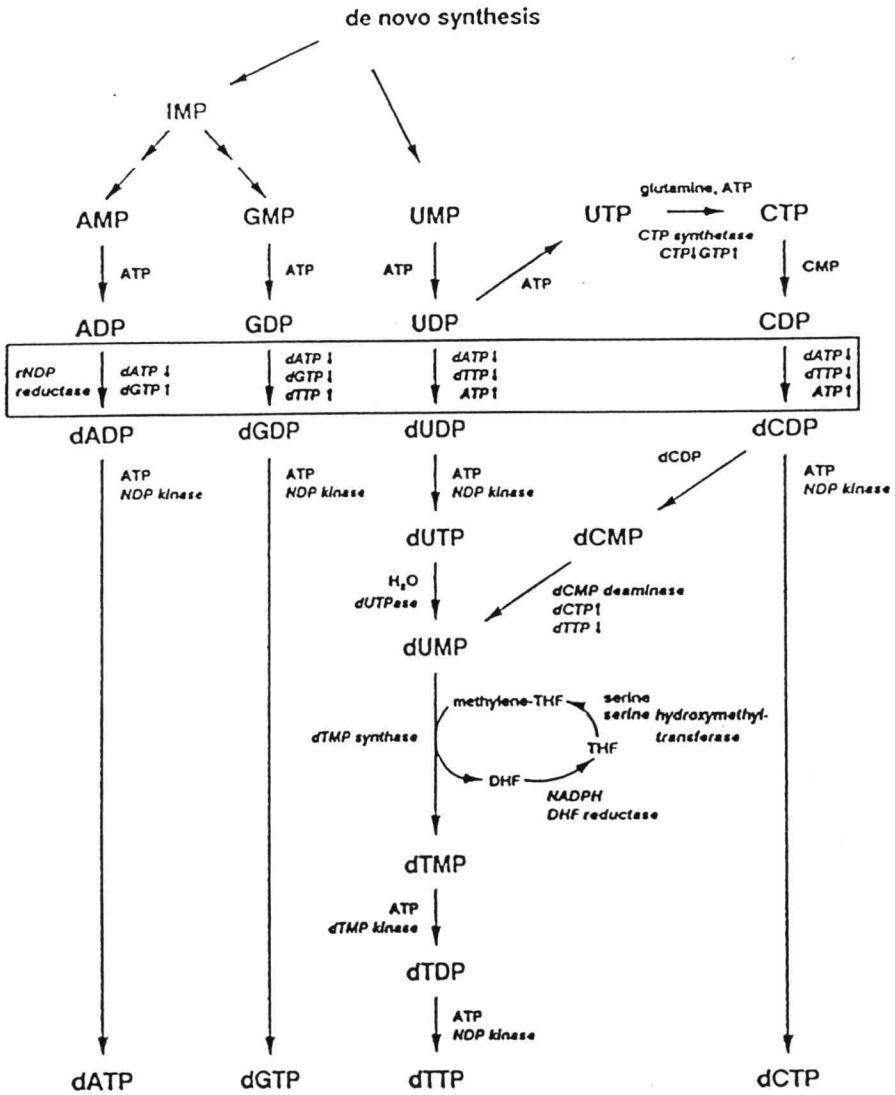


Figure 1.2. Pathways of dNTP synthesis in mammalian cells. Nucleotides that are allosteric activators or inhibitors are identified with adjacent upward and downward arrows, respectively. (From Kunz et al., 1994).

relative levels of cytidine/deoxycytidine and of uridine/deoxythymidine nucleotides, and represents another key regulatory site.

Effects of dNTP pool perturbations on DNA replication fidelity has been intensively studied *in vitro* using purified DNA polymerases and cell extracts (Kunkel, 1992). In the latter system, DNA is replicated in a bi-directional, discontinuous, and semiconservative manner (Roberts and Kunkel, 1988), similar to DNA replication *in vivo*. In both systems, the concentrations of dNTPs are critical determinants of rates for base substitution and frameshift errors. Imbalanced supply of dNTPs could affect replication fidelity in at least two ways. First, it could affect the outcoming of the competition for insertion of a correct versus an incorrect nucleotide. Misincorporation of the nucleotide in excess is usually the predominant effect observed when DNA is replicated with imbalanced dNTPs. Second, it could impair the effectiveness of proofreading activity. As the concentration of the next correct nucleotide after an error increases, so does the rate of polymerization from the mispaired (or misaligned) terminus. As a result, less time is available for the proofreading reaction and the probability of excision is decreased (Kunkel, 1988). The absolute concentration of dNTPs is also an important factor and affects this so-called "next-nucleotide effect" for the same reason: higher dNTP levels drive the polymerization reaction faster, and leave less time for proofreading reaction (Roberts et al., 1991).

Results from the *in vivo* studies are consistent with the conclusions based on *in vitro* replication assays. At least in one mammalian system, the genetic effects of imbalanced dNTP pools have been well studied *in vivo*. Some Chinese hamster ovary (CHO) mutant cell lines are deficient in feedback inhibition of CTP synthase, so that flux through CTP synthase is increased (Trudel et al., 1984). Cells bearing this mutation, like parental CHO

cells, usually have very low level of dCMP deaminase activity, so the utilization of dCMP for deoxyuridine nucleotide synthesis is greatly decreased. The combined effects elevate the pools of all cytidine and deoxycytidine nucleotides while depleting uridine and deoxythymidine nucleotide pools (Mun and Mathews, 1991). Indeed, pyrimidine pools in these cells are so low that the cells require exogenous thymidine for growth and survival (thus being termed Thy^-). Thy^- cells have a mutator phenotype. DNA sequencing analysis revealed that the majority of the mutations are attributable to the misincorporation induced by excess dCTP, and 69% of the transition mutations were found at sites where misincorporated dCTP is fixed by incorporation of two further dCTPs 3' to the error (Meuth, 1989).

The regulation of dNTP biosynthesis in eukaryotic cells is further complicated by the fact that a number of dNTP synthesis genes are regulated through the cell cycle (Merrill et al., 1992; Johnston and Lowndes, 1992). The expression of these genes is tightly linked with those of other genes whose products are involved in DNA replication. Previous data from our lab (Leeds et al., 1985) showed that dNTP levels within mammalian cell nuclei undergo significant changes as synchronized cells proceed through S phase of the cell cycle. But the effect of dNTP pool fluctuations on the fidelity of DNA replication through S phase remains to be determined. The biological significance of tight regulation of dNTP supply in S phase has recently been proposed to be a defense system that eukaryotes have evolved to limit the resources available for parasitic virus replication (McIntosh, 1993). This hypothesis is supported by the observations that reverse transcription of the HIV RNA genome is limited by dNTP pool sizes in infected lymphocytes and that viral replication can be limited further by reducing dNTP pool sizes through inhibition of ribonucleotide reductase by hydroxyurea (Gao et al.,

1993; Lori et al., 1994). Imbalanced dNTP pools from the host can decrease viral genome replication accuracy too. Recent data suggest that the hypermutation, mostly G to A transitions, displayed by HIV virus genomes may arise from a dramatic dCTP deficiency at the sites of reverse transcription (Vartanian et al., 1992 and 1994). These works underscore the importance of seeking further understanding of dNTP perturbation and their genetic consequences.

EFFECT OF DNA BASE SEQUENCE CONTEXT ON DNA REPLICATION FIDELITY

The fidelity of DNA synthesis depends not only on DNA polymerase, but also on template DNA sequence context. Significant site-to-site variability in misinsertion rates, by an order of magnitude or more, depending on surrounding nucleotide sequence, has been reported (Mendelman et al., 1989). Kinetic studies using DNA polymerases without proofreading activity showed that base pairs immediately 5' to the nucleotide to be inserted have a significant effect on both insertion kinetics and misinsertion efficiencies; generally a 10- to 50- fold difference was observed (Mendelman et al., 1989; Bloom et al., 1993).

For misinsertion to result in a base substitution, DNA polymerase must be able to extend from the mismatched terminus before it is attacked by 3' to 5' proofreading exonuclease. Indeed, the kinetic barriers to chain extension from a mismatch is a key factor in determining the efficiency of proofreading (Donlin et al., 1991). There are twelve possible mispairs when one considers only the template nucleotide and incoming dNTP. Partly because of the structure constraint from neighboring sequence, all twelve

mispairs are not formed with equal efficiency. At least in one case, results from systematic studies showed that the geometric stability of G·A mismatch depends entirely on the neighboring sequence (Cheng et al., 1992). Goodman and his collaborators (Goodman et al., 1993) have developed a very sensitive system to study sequence context effects on replication fidelity *in vitro*. They use a mutagenic nucleotide analog, 2-aminopurine deoxynucleoside triphosphate (dAPTP) as a substrate to follow chain extension from an annealed template-primer. 2-Aminopurine competes with adenine to form a Watson-Crick type base pair with T with two hydrogen bonds; it can also form "incorrect" base pairs with C. The misinsertion frequency for 2-aminopurine (AP) is much higher than for natural nucleotides. In addition, AP is much more fluorescent and its fluorescence properties are sensitive to its environment. These advantages make the system very sensitive for measuring kinetics of base insertion and chain extension (Bloom et al., 1993). Using this system, they demonstrated that base mispair extension by *Drosophila* polymerase α and AMV reverse transcriptase, both of which lack detectable 3'-exonuclease activity, was dependent on surrounding sequence and that from a given mismatched terminus, some polymerases can extend while others cannot (Mendelman et al., 1989).

Five different DNA polymerases have been found in eukaryotes. Data from different assay systems all indicate that they replicate DNA with quite different accuracies (Kunkal, 1992; Wang, 1991). These experimental data have raised an interesting question: are two strands of DNA synthesized with different accuracies since it is generally accepted that in eukaryotes two DNA strands are synthesized by different DNA polymerases (Kornberg and Baker, 1992)? Roberts et al. (1991) observed no significant base substitution rates between the two strands when they replicated DNA *in vitro* using HeLa cell

extracts. But more detailed examination of this issue are obviously needed since there is also evidence indicating inequality in mutation rates of the two strands of DNA (Wu and Maeda, 1987; Filipski, 1990; Trinh and Sinden, 1991; Veaute and Fuchs, 1993). The common explanation about the observed difference between the polymerases is based on the observation that polymerase may be in direct contact with up to eight bases in primer/template. Enzymes with different amino acid sequences may have differences in protein-DNA interactions and those differences may result in different replication accuracies (Kunkel, 1992; Goodman et al., 1993).

Sequence context also affects the effectiveness of proofreading activity. Using a mutant T4 DNA polymerase with highly active proofreading exonuclease (antimutator phenotype), Pless and Bessman (1983) found that local base sequence had a significant effect on substitution of 2-aminopurine for adenine. The wild-type T4 polymerase followed the same pattern but in a less pronounced form. Petruska and Goodman (1985) further analyzed the data and found that there was a marked correlation between proofreading and the average DNA stability in the ten-base-pair region around 2-aminopurine incorporation sites. Proofreading appeared most effective in AT-rich and least effective in GC-rich regions. Since structural evidence showed that duplex DNA must be melted to send a mismatched terminus to a separate proofreading domain or subunit, a simple explanation of the above results is that DNA duplexes at AT-rich regions are easier to be separated and encourage proofreading while G·C base pairs stabilize DNA and inhibit proofreading.

Sequence context also influences polymerase dissociation rate after an error and nucleotide selection opposite abasic lesions on the template

(Randall et al., 1987; Goodman et al., 1993), and these influences can also change DNA replication fidelity.

It should be mentioned that all these studies regarding sequence effects on replication fidelity were performed on synthetic primer/templates and DNA was synthesized from a fixed primer by purified DNA polymerases, T4 DNA polymerse in particular. Obviously, this mode of DNA synthesis does not represent typical DNA replication *in vivo*. Although results generated are very informative, to generalize these conclusions, more studies in a system close to the *in vivo* situation are required.

NUCLEOSIDE DIPHOSPHATE KINASE: A FUNCTIONAL LINK BETWEEN dNTP SYNTHESIS AND DNA REPLICATION

1. Metabolic role of NDP kinase

NDP kinase is a ubiquitous enzyme that transfers γ-phosphate from virtually any ribo- and deoxyribonucleoside triphosphate (NTP and dNTP) to almost any ribo- and deoxyribonucleoside diphosphate (NDP and dNDP). It plays a pivotal role in the supply of precursors for the synthesis of nucleic acids since it catalyzes the synthesis of all but one of the eight common ribo- and deoxyribonucleoside triphosphates. In aerobic cells, ATP can be generated through oxidative phosphorylation. Because ATP is usually the most abundant nucleoside triphosphate and NDP kinase has high activity with quite low specificity, γ-phosphate from ATP is rapidly redistributed to other NDPs and dNDPs with the synthesis of NTPs and dNTPs.

NDP kinase proteins have been purified and characterized from various sources, both prokaryotic and eukaryotic. The protein as studied in most organisms is an oligomer composed of up to six 15- to 18- kilodalton

(kDa) subunits. It requires Mg^{+2} for its catalytic reaction and no allosteric effector has been described (Ray and Mathews, 1992). All NDP kinases studied so far react through a phosphoenzyme intermediate in which a histidine is phosphorylated, and follow a ping-pong reaction mechanism (Parks and Agarwal, 1973):

$$N_1TP + E \rightleftharpoons N_1DP + E \sim P$$

$$N_2DP + E \sim P \rightleftharpoons N_2TP + E$$

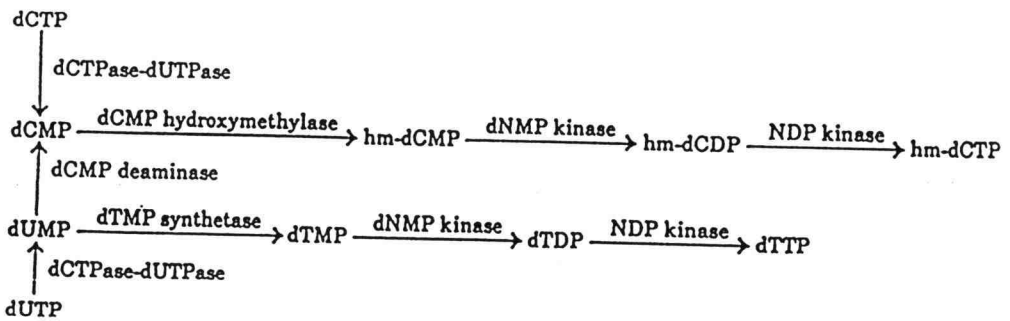
2. Roles of NDP kinase in T4 bacteriophage dNTP synthesis and DNA replication.

a. *Escherichia coli* host NDP kinase is a part of T4 dNTP synthesis machinery.

After bacteriophage T4 infection the flow of precursors into nucleic acids is dramatically altered. This large DNA virus causes the rate of DNA synthesis to increase about 10-fold relative to the preinfection rate. To accomplish this rapid DNA synthesis, the T4 genome encodes almost all the enzymes required for both DNA replication and dNTP synthesis. NDP kinase is the only *E. coli* protein known to play an indispensable role in the synthesis of dNTPs for T4 DNA replication (Mathews and Allen, 1983). Bello and Bessman (1963) first observed an increase of deoxyribonucleoside monophosphate (dNMP) kinase activity in T-even phage infected cells, but they did not find any increase in NDP kinase activity. Subsequent searches found a phage-encoded dNMP kinase gene, but failed to find a phage-encoded NDP kinase gene. Results from activity assays revealed that NDP kinase activity in uninfected *E. coli* cells was already some 20-fold higher than T4 dNMP kinase activity in T4-infected cells. Apparently, there is no need for T4 to carry its own NDP kinase gene.

b. T4 dNTP synthesis machinery is a multienzyme complex.

The earliest indication that dNTP synthesis in T4 infection might occur via a multienzyme complex came from the report by Chiu and Greenberg (1968) that dCMP hydroxymethylase was found to be in a high-molecular-weight aggregate in the extracts of phage-infected bacteria. Multienzyme aggregates containing at least ten enzymes of dNTP synthesis were later partially purified. The aggregate still retained through fractionation the ability to kinetically facilitate multi-step reactions, with appreciable shortening of transient times and lowering of intermediate concentrations, when compared with the behavior expected for equivalent mixture of noncomplexed enzymes. In one example, Reddy and Mathews (1978) isolated the enzyme aggregate and demonstrated the kinetic coupling in the three-step sequential reactions converting dUMP to dTTP:



More interestingly, this kinetic coupling is disrupted by mutational inactivation of gene 42, even though the mutant gene product, dCMP hydroxymethylase, plays no direct role in converting dUMP to dTTP. They further showed that mutants that make nearly full length dCMP hydroxymethylase show normal kinetic behavior (Thylén and Mathews,

1989). The existence of this multienzyme complex has now been confirmed from the results collected through a dozen experimental approaches (Mathews, 1993 for recent review).

c. T4 dNTP synthesis complex is linked to DNA replication machinery

Maintaining a high DNA synthesis rate after T4 infection requires rapid dNTP synthesis. It has been estimated that, in a T4 infected *E. coli* cell, the dNTP pools in the immediate vicinity of replication sites must turn over completely about 10 times per second (Ji and Mathews, 1991), suggesting that simple diffusion of dNTPs from a remote site of synthesis cannot provide sufficient supplies of dNTPs for DNA synthesis. Accumulating evidence indicates that dNTPs used for DNA synthesis are synthesized near the replication sites and directly channeled to the replication forks (Figure 1.3). The initiation of dNTP synthesis and DNA replication occurs simultaneously and the rate increases exponentially up to 18 minutes post infection, indicating the tight coupling of dNTP synthesis and DNA replication (Tomich et al., 1974). When *E. coli* cells are permeablized after T4 infection to allow passage of exogenous nucleotides, dNMPs and rNDPs are two- to threefold more efficiently incorporated into T4 DNA than dNTPs, suggesting that dNTPs formed through the dNTP synthase complex are more efficiently utilized (reviewed by Greenberg, 1994). The importance of the integrity of this coupling between dNTP synthesis and DNA replication is further demonstrated by the observations that cells infected by T4 phage defective in hm-dCTP synthesis showed little or no DNA synthesis even when they are supplied with all four dNTPs, including hm-dCTP (Wovcha et al., 1973; North et al., 1976; Stafford et al., 1977).

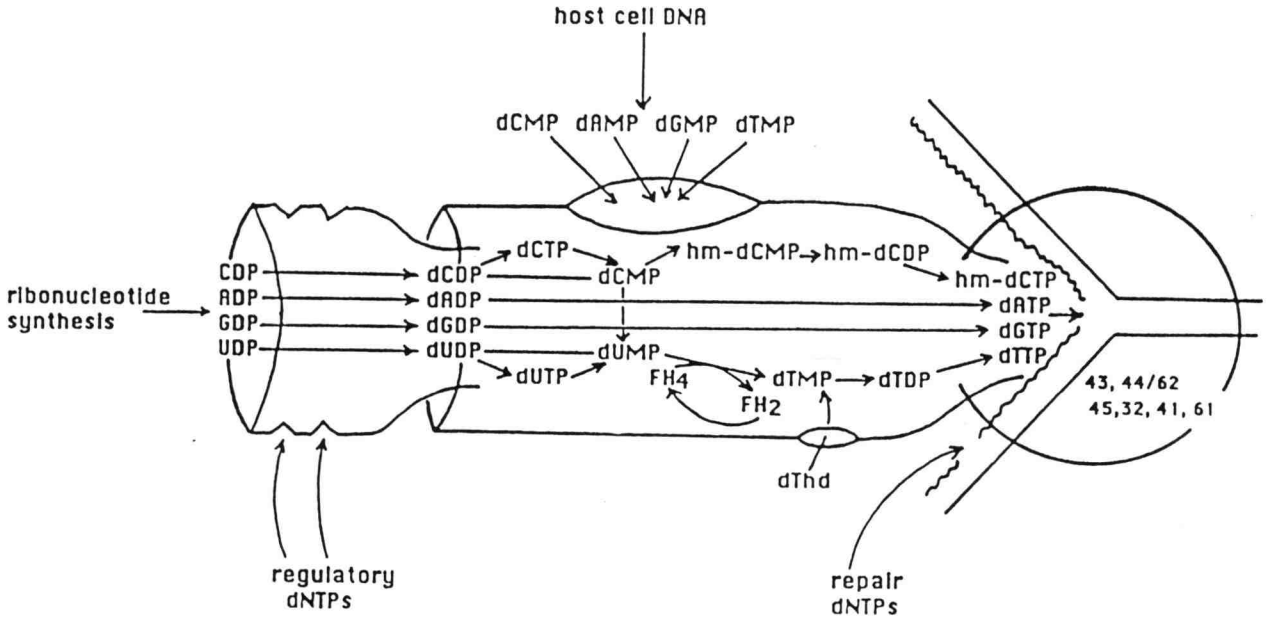


Figure 1.3. A model of dNTP biosynthesis in T4 phage infection in which a multienzyme complex for dNTP synthesis is juxtaposed with the replication machinery. The numbers in the circle refer T4 replication proteins. (From Mathews, 1993).

Genetic evidence also suggests specific enzyme-enzyme interactions that link dNTP synthesis and DNA replication. A dCMP hydroxymethylase gene mutation can be suppressed by a mutation of gene 43, which encodes T4 DNA polymerase (Chao et al., 1977). Similar interactions have been found between phage-coded dihydrofolate reductase and primosome-associated helicase, and between small subunit of ribonucleotide reductase and a subunit of topoisomerase (Chiu et al., 1980; Macdonald and Hall, 1984). Although the biochemical nature of these interactions is not clear, the observed genetic interactions do suggest an intracellular protein organization that coordinates dNTP synthesis with DNA replication.

Another informative approach to analyze protein-protein interactions is protein affinity chromatography. Immediately after T4 infection the expression of host genes is turned off and newly synthesized proteins are almost exclusively T4 proteins, so they can be easily labeled. Purified proteins of interest can be immobilized on column matrix, and radiolabeled T4 proteins which interact with the fixed protein and are retained on the column can be identified through two-dimensional gel electrophoresis and radioautography. Columns with immobilized protein of dNTP synthesis and DNA replication retained a number of proteins which are associated with both dNTP synthesis and DNA replication. Some of these protein-protein interactions are rather strong. The proteins remain bound at 0.2 M NaCl and are eluted from the column at 0.6 M NaCl, indicating that these protein-protein interactions can exist at salt concentrations comparable to, or somewhat exceeding, those encountered *in vivo* (Formosa et al., 1983; Wheeler et al., 1992).

d. NDP kinase could be a functional and physical link between dNTP synthesis complex and DNA replication apparatus.

Several observations suggest that NDP kinase could be a link between dNTP synthesis complex and DNA replication machinery. First, bacterial NDP kinase activity remained associated with several dNTP synthesis enzymes even after three chromatographic purification. When cells were infected by T4 with mutations that disrupted the formation of dNTP synthetase complex, this tight association disappeared (Thylén and Mathews, 1989), indicating that this enzyme complex was a specific biological assembly. Second, NDP kinase is found to associate with purified DNA polymerase (Miller and Wells, 1971). Third, immobilized NDP kinase specifically bound several T4 enzymes of dNTP synthesis as well as some T4 DNA replication proteins. Among those T4 proteins fairly tightly bound to a NDP kinase column were those identified as DNA single-strand-binding protein, DNA polymerase accessory protein, DNA primase and DNA repair protein (Ray, Ph.D thesis, 1992). Since NDP kinase interacts with so many dNTP synthesis enzymes and DNA replication proteins, NDP kinase may play the role as a link from the dNTP synthetase complex to the replication machinery (Ray and Mathews, 1992; Greenberg et al., 1994).

3. Regulatory functions of NDP kinase.

In addition to its role in nucleotide synthesis, NDP kinase has also been indicated to play roles in several other biological processes. *Drosophila* developmental gene, *abnormal wing discs* (*awd*) encodes NDP kinase protein. A null mutation of *awd* causes abnormalities in development, leading to death in the prepupal stage (Rosengard et al., 1989). Crude extracts of mutant

larvae had less than 2% NDP kinase activity detected in nonmutant larvae of the same stage (Biggs et al., 1990). That the lethal effect was only observed at later stage seems to suggest that insufficient supply of dNTPs may not be the reason since mutant larvae were able to support rapid DNA synthesis during embryonic and larval development.

A special mutation of *awd* gene, called *killer of prune* (*awd*$^{k\text{-}pn}$), on its own has no phenotype. Loss-of-function mutations at *prune* locus result in altered eye color to prune instead of purple. But when in combination with *awd*$^{k\text{-}pn}$, prune mutations are lethal (Biggs et al., 1988). *Killer of prune* mutation is a serine to proline substitution and does not affect catalytic activity of the enzyme (Lascu et al., 1992), so the lethal effect is not due to loss of NDP kinase. Teng et al. (1991) found that the *prune* locus has significant homology with ras GTPase-activating proteins (GAPs), suggesting that *ras* proteins may be regulated by NDP kinase. *Killer of prune* mutation in prune animals may disrupt this regulation and cause a lethal effect. NDP kinase has also been found to associate with GTP-binding proteins, further suggesting that NDP kinase could play a role in G-protein-mediated signal transduction pathway (Randazzo et al., 1991, Bominar et al., 1993).

Reduced expression of *nm23*, mammalian gene coding for NDP kinase, in certain rodent model systems and human breast tumors has been correlated with high tumor metastatic potential. Constitutive expression of wild type nm23 in highly metastatic murine melanoma cells significantly reduces the incidence of primary tumor formation and tumor metastatic potential, and alters cell responses to transforming growth factor b, suggesting a suppressive effect of *nm23* on several aspects of the cancer process (Leone et al., 1991; Liotta et al., 1991). Further study from the same group shows that NDP kinase activity is not required for this tumor suppressive function

(MacDonald et al., 1993). The biochemical mechanism by which nm23 regulates tumor development remains to be established. One step in that direction is the recent report that human c-myc transcription factor PuF is identified as nm23, and again, the NDP kinase activity is not required for its DNA binding activity and *in vitro* transcriptional functions (Postel et al., 1993, Postel and Ferrone, 1994). These results suggest that NDP kinase/nm23 is a bifunctional protein with NDP kinase on the one hand, and DNA binding and transcriptional activity on the other.

NDP kinase copurified with succinyl-CoA synthetase in prokaryotes and in mitochondria of eukaryotes (Kadrmas et al., 1991; Kavanaugh-Black et al., 1994). It has been proposed that *ndk* may act as a protein kinase and directly phosphorylate succinyl-CoA synthetase to modulate sugar metabolism during developmental changes. The idea that NDP kinase can act as a protein kinase is interesting since it could explain many regulatory roles attributed to the enzyme.

PHOSPHOTRANSFERASE SYSTEM AND PYRUVATE FORMATION DURING THE CATABOLISM OF GLUCOSE IN *E. COLI* WITH PYRUVATE KINASE MUTATIONS

Results from *in vitro* assays demonstrate that NDP kinase is not the only enzyme that can catalyze the formation of nucleoside triphosphates. Pyruvate kinase has a broad substrate specificity and can transfer energy-rich phosphate from phosphoenolpyruvate (PEP) to any other nucleoside diphosphates (Saeki et al., 1975). Pyruvate kinase normally plays a key role in cell metabolism. It catalyzes the formation of pyruvate from PEP with production of one molecule of ATP. Pyruvate can be converted to acetyl-CoA by pyruvate dehydrogenase complex. Both pyruvate and acetyl-CoA can be directly used for many biosynthesis. Under aerobic conditions, pyruvate can

be oxidized via the tricarboxylic acid (TCA) cycle to yield more ATP. During anaerobic growth, a number of end products are formed, almost all of them are derived from pyruvate (Figure 1.4).

There are two pyruvate kinase genes in *E. coli* cells, *pyk*A and *pyk*F. Although pyruvate kinases play a major role in the formation of pyruvate from PEP, cells with mutations of both pyruvate kinase genes still grow well on certain sugars because pyruvate can be formed through other pathways (Figure 1.5). Among them, the PEP:sugar phosphotransferase system (PTS) is probably the most significant one since cells totally devoid of pyruvate kinase activity grow poorly on non-PTS sugars (Pertierra and Cooper, 1977). Figure 1.6 shows schematically the components and reactions of PTS (Postma, 1987). The net reaction for the uptake of sugars via the PTS is as follows.

$$\text{Sugar}_{\text{out}} + \text{PEP}_{\text{in}} \rightleftharpoons \text{sugar-P}_{\text{in}} + \text{pyruvate}_{\text{in}}$$

Enzyme I and HPr are the two general proteins of the PTS. They both can be phosphorylated at histidine residues. Phosphorylated HPr can react in two different ways, either directly with an enzyme II or with an enzyme III and its membrane-bound enzyme II. As shown in Figure 1.6, mannitol is transferred directly through the enzyme II specific for mannitol, while proteins II and III, specific for glucose, are involved in glucose transport.

During anaerobic glycolysis, one molecule of glucose generates two molecules each of pyruvate, ATP and NADH. NADH must be oxidized to regenerate NAD^+ for further glycolysis. As shown in Figure 1.4, NADH can be consumed at several steps during anaerobic dissimulation of pyruvate. If sufficient pyruvate is not produced to permit reoxidation of the NADH

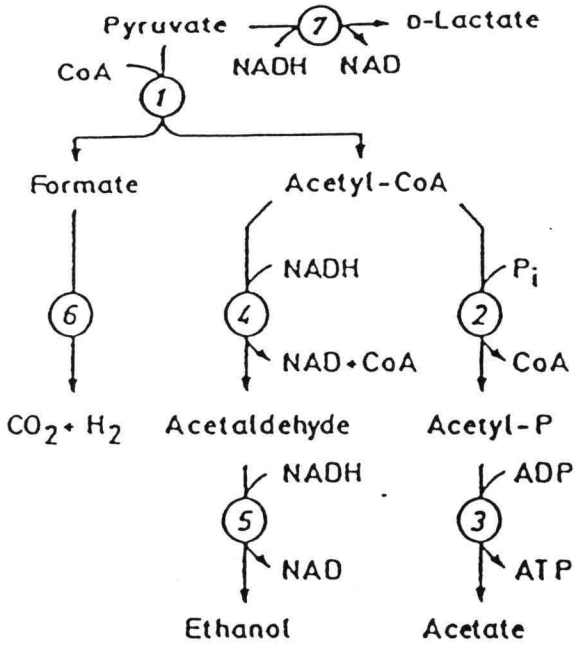


Figure 1.4. Terminal reactions of anaerobic glycolysis in *E. coli* cells. 1, pyruvate formate-lyase; 2, phosphotransacetylase; 3, acetate kinase; 4, acetaldehyde dehydrogenase; 5, alcohol dehydrogenase; 6, formate hydrogen-lyase complex; 7, lactate dehydrogenase. (From Knappe, 1987).

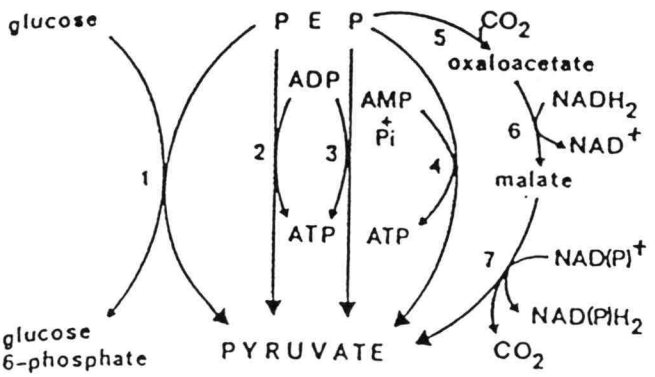


Figure 1.5. Possible reactions for the formation of pyruvate from PEP in *E. coli*. Reaction 1 is the PEP:sugar PTS; reaction 2 is pyruvate kinase A; reaction 3 is pyruvate kinase F; reaction 4 is PEP synthase; reaction 5 is PEP carboxylase; reaction 6 is malic dehydrogenase; reaction 7 is malic enzyme. (From Pertierra and Cooper, 1977 with modifications).

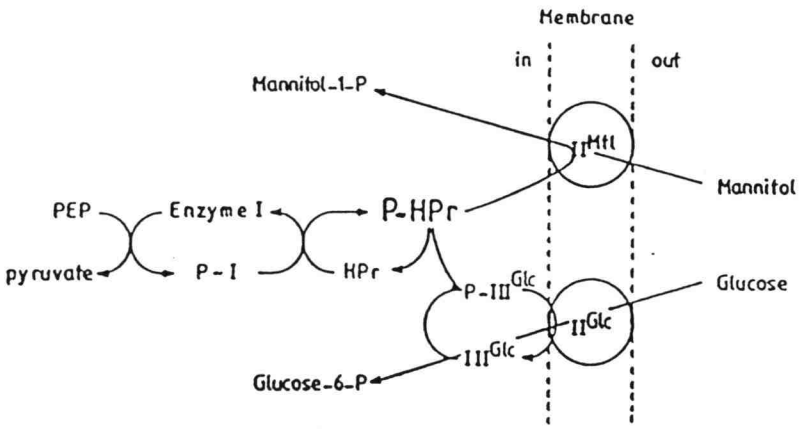


Figure 1.6. The PTS. Enzyme I and HPr are the general PTS proteins. Of the many different enzymes II only two are shown. II^{Mtl} is specific for mannitol, and III^{Glc}, together with II^{Glc}, is specific is specific for glucose. P~I, P~HPr, and P~III^{Glc} are the phosphorylated forms of enzyme I, HPr, and III^{Glc}.

formed during glycolysis, an additional electron acceptor must be provided to support growth.

PRESENT WORK

In this thesis, the work of three projects is presented. The objective of this work is to further our understanding of mutagenic mechanisms associated with DNA cytosine methylation, DNA base sequence context, and DNA precursor pool asymmetry. The first project is to investigate why 5-methylcytosine residues in DNA are mutation hotspots. Is it the result of the intrinsic chemical instability of 5-methylcytosine toward deamination or some other biological processes, such as the enzymatic methylation reaction itself, that is to be blame? To answer this question, I designed a sensitive genetic assay method and for the first time measured the deamination rates for cytosine and 5-methylcytosine in duplex DNA. Results from this study implicate difference in chemical stability toward deamination as a major causal factor relating DNA-cytosine methylation to spontaneous mutagenesis.

The second project is designed to seek better understanding of the relationship between nucleotide composition of a gene and the rate of the gene's evolutionary variation. Previous studies showed a complex relationship between G·C content of a gene and its base substitution rate, and the change of dNTP pools during cell cycle has been proposed to be partly responsible. To investigate the underlying mechanism, I constructed a reversion assay system, in which an opal codon in *lacZα* region is flanked either by nine adenine·thymine base pairs on each side, or by nine guanine·cytosine base pairs, or by the wild-type sequence context. Reversions due to replication errors can be easily detected based on the color change of

the plaques. I found that dNTP pools approximating those in HeLa cell nuclei are mildly mutagenic and sequence context can significantly affect the effectiveness of proofreading activity. Apparently, both base sequence context and base sequence context contribute to the natural fidelity of DNA replication.

The third project is focused on investigating the role of nucleoside diphosphate kinase in dNTP synthesis. In *E. coli* cells with a deletional mutation of the NDP kinase gene (*ndk*), some 20-fold expansion of the dCTP pool over normal was observed, a sufficient bias to account for the mutator phenotype associated with the *ndk* mutation. Mutant cells can still support T4 infection, although T4 phage does not have its own *ndk* gene and must rely on host NDP kinase to synthesize nucleoside triphosphates. A novel source of NDP kinase was suggested. Change of dNTP pools after anaerobic T4 infection was measured for the first time. A surprising finding was that deletion of *ndk* gene in cells with disrupted pyruvate kinase genes enables cells to grow anaerobically again, while cells with the same mutations on pyruvate kinase genes alone fail to grow. These results suggest that NDP kinase plays a regulatory role in cellular functions in addition to dNTP synthesis, and those regulatory effects exerted by NDP kinase could be related to its role in tumor suppressing and metastatic regulation observed in mammalian cells.

Chapter 2

Effect of DNA-Cytosine Methylation upon Deamination-induced Mutagenesis in a Natural Target Sequence in Duplex DNA

Xiaolin Zhang and Christopher K. Mathews§

Published in *Journal of Biological Chemistry*
Vol. 269, No. 10, Issue of March 11, pp. 7066-7069,1994
by The American Society for Biochemistry and Molecular Biology, Inc

Effect of DNA-Cytosine Methylation upon Deamination-induced Mutagenesis in a Natural Target Sequence in Duplex DNA

Xiaolin Zhang and Christopher K. Mathews§

Running title: DNA Methylcytosine is a mutagenic target

Acknowledgments and abbreviations

* This work was supported by National Science Foundation Research Grant DMB 9119854. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed. Tel.: 503-737-1865.

[1]The abbreviations used are: mC, 5-methylcytosine; ss, single-stranded; ds, double-stranded; IPTG, isopropyl b-D-thiogalactoside; X-gal, 5-bromo-4-chloro-3-indolyl b-D-galactoside; RF, replicative form; AdoMet, S-adenosyl-L-methionine

Acknowledgments—We thank Dr. Thomas A. Kunkel and Dr. David C. Thomas of the National Institute of Environmental Health Sciences for providing *E. coli* strains and M13mp2SV DNA. Dr. Dale Mosbaugh kindly provided us with uracil-DNA glycosylase inhibitor. We thank members of this lab for helpful discussion.

SUMMARY

Are 5-methylcytosine residues in DNA hotspots for transition mutagenesis? Numerous studies identify (1) structural changes induced by DNA methylation, (2) high percentages of human mutations that result from GC to AT transition pathways, and (3) differences between G·C and G·mC base pairs in susceptibility to nonenzymatic deamination. However, investigations of chemical stability necessarily involve non-physiological conditions for chemical analysis of deamination. Here we describe an experiment that compares rates of deamination-induced mutagenesis between a G·C and G·mC base pair, when both are present in duplex DNA, incubated at 37 °C and pH 7.4, within identical sequence contexts, in a natural mutational target (the *E. coli lacZα* gene) which selects for mutagenesis at the specific site under investigation. Under these conditions the rate of spontaneous deamination at G·mC exceeds that at G·C by more than 21-fold. Our data implicate differences in chemical stability toward deamination as a major causal factor relating DNA-cytosine methylation to spontaneous mutagenesis.

INTRODUCTION

Vertebrate genomes are heavily methylated at C-5 of cytosine residues within CpG dinucleotides. In most eukaryotic DNAs only cytosine is methylated. While the specific roles of this DNA methylation are still not well understood, many important functions are involved (1). The recent finding that interference with DNA methylation *in vivo* blocks embryonic development in the mouse (2) dramatically underscores the importance of DNA methylation.

However, methylation of DNA-cytosine residues is highly mutagenic. Methylated cytosines were first found to be base substitution hot spots in *Escherichia coli,* and when the respective cytosines were not methylated the hot spots disappeared (3). Subsequently the analysis of transition mutations in vertebrates similarly showed a correlation with methylation at CpG sites (4). The mutation products are specifically associated with C→T or G→A transitions. Various investigators have proposed that these transition mutations are the combined results of the elevated spontaneous deamination of 5-methylcytosine to thymine and slower repair of the consequent G·T mispairs relative to G·U mispairs arising from deamination of cytosine. This is in agreement with the results from the analysis of the nearest neighbor dinucleotide frequencies and the level of DNA methylation in animals: CpG dinucleotides are under-represented, while the resultant deamination products, 5'-TpG-3' and 5'-CpA-3' are over-represented (5). Although 5-methylcytosine comprises fewer than 1% of human DNA bases, (6), 30-40% of point mutations leading to human genetic disorders or cancer are related to base substitution at 5-methylcytosines (7, 8). These transitions have been

found in mutant forms of Factor IX, human LDL receptor, cyclic AMP-dependent protein kinase, and the tumor suppressor gene product p53 (9).

5-Methylcytosine residues in ss[1] DNA have been shown to be more susceptible to heat-induced deamination than are cytosine residues. Based upon analysis at non-physiological temperatures, the deamination rate for 5-methylcytosine was calculated to be about 4-fold higher than that for cytosine at 37 °C and pH 7.4 (10, 11). Using a sensitive genetic assay system, Frederico *et al.* (12) measured the deamination rate for cytosine residues in ss DNA at 37 °C as 1×10^{-10} s^{-1}. On the other hand, the double helix structure provides good protection against cytosine deamination. The measured half life for cytosine in ss DNA at 37 °C is about 200 years, while in ds DNA it is of the order of 30,000 years (10, 12, 13). Although all evidence indicates that 5-methylcytosine in double-stranded DNA should deaminate faster also, to our knowledge, its rate has not been measured, partly because the reaction is too slow to measure accurately (4).

Patterns of DNA methylation are actively maintained by postreplicative methylation of newly synthesized DNA by DNA methyltransferase (14). A proposal for the mechanism of DNA methylation by methyltransferase suggests the formation of dihydrocytosine as an intermediate. Dihydrocytosine is very labile to hydrolytic deamination, with a half-life of approximately 105 minutes at 25 °C (15). This suggests that the observed cytosine deamination at CpG sites could also be the result of the enzymatic methylation reaction instead of deamination from 5-methylcytosine. If the deamination occurs before the methyl group is bonded to the 5 position, uracil is produced; if after, thymine is the product. This pathway was initially proposed as a possible mechanism for DNA enzymatic deamination with methyltransferase thereby acting as a "DNA deaminase"

(16). Shen *et al.* (17) provided experimental evidence supporting this pathway when the AdoMet concentration is very low. They found, in the absence of AdoMet, that *Hpa*II methylase (recognition sequence CCGG with methylation at 5 position of the second cytosine residue) induced an approximately 10^4-fold increase of C to U transitions.

Does mutagenesis associated with methylation result from inherent instability of a G·mC base pair in duplex DNA, compared to G·C, or is deamination an enzymatic process, such as a consequence of a side reaction of DNA methylase activity at limiting AdoMet levels? We describe two approaches to this question: (1) construction of duplex DNAs identical except for methylation of the cytosine residue at one site where mutagenesis can be followed at 37 °C and pH 7.4; (2) analysis of mutagenesis induced *in vitro* by the action of CpG methylase, in the absence of AdoMet.

RESULTS AND DISCUSSION

Measurement of Cytosine and 5-Methylcytosine Deamination Rate.

We used a sensitive genetic assay to detect the products of deamination from cytosine and 5-methylcytosine. Site-directed mutagenesis by an oligonucleotide 5' AGGGTTTTCCCGGTCACGACGTTG 3' (mutation target base underlined) was used to create a *Hpa*II methylase site (CCGG) in the *E. coli lacZ*α gene in M13mp2SV (18). Consequently, tyrosine 17 (methionine as 1) was changed to arginine. This mutation inactivated the *lacZ*α function. When DNA was transfected into an *E. coli* host and plated with cells providing the rest of the *lacZ* gene, a functional α-complementation cannot occur. This caused a phenotypic change from dark blue plaques to faint blue

plaques on minimal medium containing IPTG and X-gal. The transfection and plating were performed as described by Kunkel (19). The mutation was confirmed by DNA sequencing. Mutant supercoiled replicative DNA was prepared by alkali lysis and CsCl gradient centrifugation (20).

The mutant RF1 DNA was methylated by *Hpa*II methylase *in vitro* with AdoMet as methyl donor. That methylation was complete was confirmed by resistance of the DNA to *Hpa*II restriction digestion. Both methylated and unmethylated DNAs were dissolved in 0.1 M KCl/0.05 M HEPES-KOH/0.01 M $MgCl_2$/0.001 M EDTA, pH 7.4 (13) at 40 ng/μl each. 200 μl of each DNA solution was sealed in a capillary tube and incubated at 37 °C in the dark for the indicated times. Conditions were chosen so that temperature and ionic strength closely correspond to conditions *in vivo*.

The experimental scheme is presented Figure 2.1. To detect the deamination product of cytosine, that is, uracil, DNA was transfected into *E. coli* NR8052, which is DNA-uracil glycosylase-deficient strain to prevent the repair of G·U mismatches. Since the deamination of 5-methylcytosine gives rise to a G·T mispair, methylated DNA was transfected into *E. coli* NR9162, which is deficient in DNA mismatch repair (*mutS*$^-$). *E. coli* CSH50 (F'*lac*ZΔM15) was used as plating host. The genotypes of *E. coli* NR8052, NR9162 and CSH50 have been described (18, 21). Deamination of the target cytosine or 5-methylcytosine in mutant DNA restores the wild type *LacZα* function, giving rise to dark blue plaques which are readily distinguished from faint blue background. Revertants were selected after 16 hours incubation of plates at 37 °C. To eliminate false positives, blue plaques were picked and replated along with wild-type phage. This replate gave direct color comparison on the same plate. Fifteen revertants were then randomly chosen

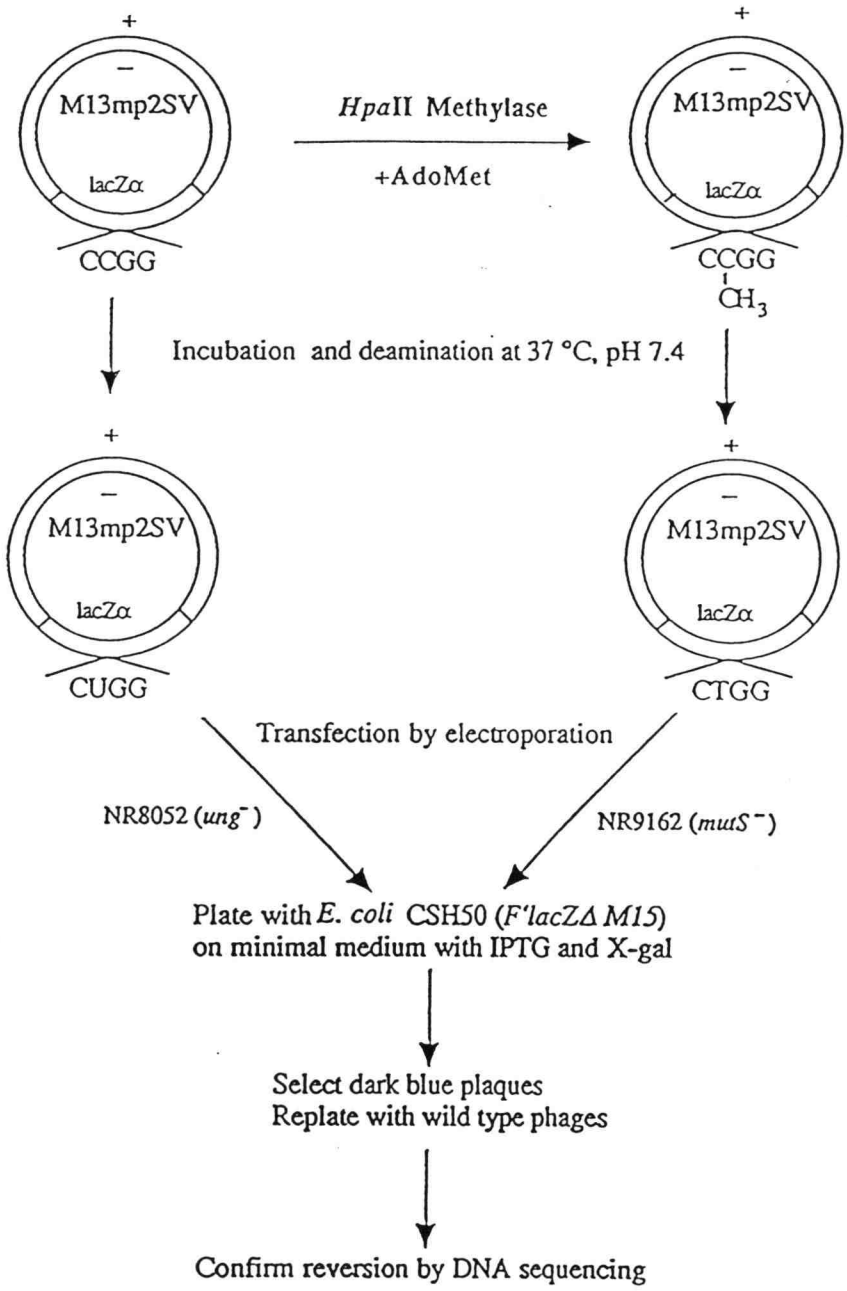


Figure 2.1. Experimental design of the genetic assay for cytosine and 5-methylcytosine deamination.

and their DNAs were found to have wild-type sequence, confirming the genetic reversion.

Theoretically, without repair of the G/T mispair, the heteroduplex should produce both mutant and wild-type phenotypes—one from each strand. Due to the asymmetry of rolling circle replication, the expression of the two strands is not equal. In our experiment, only deamination of cytosine or 5-methylcytosine on the (+)-strand (viral strand) can restore *lacZ*α function and be detected. The frequency of expression of T-containing strand from G(–)/T(+) heteroduplex was measured after the heteroduplex was transfected into *E. coli* NR9162. The preparation of the heteroduplex was essentially the same as described by Thomas *et al.* (22). Ten independent transfections were performed. The frequency of expression of the (+)-strand was calculated by dividing the sum of blue and mixed plaques by the total plaque-forming units. The frequency of expression was found to be 55.4% ± 5.4% (data not shown), which is consistent with published data (22, 23). This analysis is necessary to accurately calculate the reversion rate (24).

Figure 2.2 shows the results of this genetic assay. After 120 days' incubation of methylated DNA at 37 °C, a 21-fold increase of revertant fraction (the percentage of blue plaques among total plaques screened) was observed. In contrast, no detectable change of the revertant fraction was found for unmethylated DNA during the same incubation period. Since the increase of revertant fraction was linear with the incubation time ($r = 0.99$), the deamination rate constant for methylated DNA in this sequence content was calculated as 1.5×10^{-11} s^{-1} from the equation

$$K = \frac{-\ln (1 - F')}{t}$$

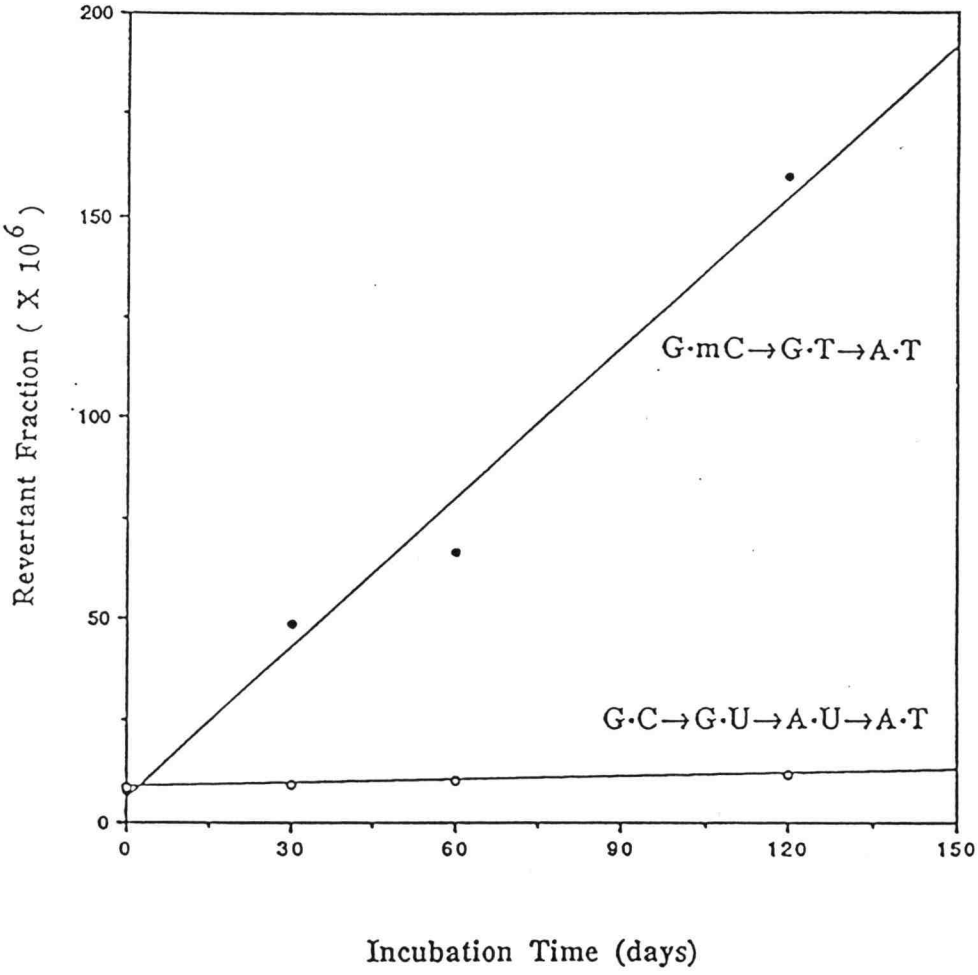


Figure 2.2. Time course for cytosine and 5-methylcytosine deamination in ds DNA at 37 °C. 200 µl of 40 ng/µl mutant RF1 DNA was sealed in capillary tubes. The tubes were incubated at 37 °C in the dark for the indicated times. DNA was then desalted and transfected into *E. coli* by electroporation. Data represent the average results of 2 to 4 determinations for each time point, with more than 1,000,000 plaques screened or at least 60 revertants selected. Values from each determination agreed within 50%. for each time point. The methylated DNA was transfected into NR9162 (*mutS*$^-$). Its background (time 0) revertant fraction was 7.7 x 10^{-6}. The observed revertant fraction after incubation was divided by 0.554 to account for strand expression as described in the text. The unmethylated DNA was transfected into NR8052 (*ung*$^-$); its background revertant fraction was 8.7 x 10^{-6}. Revertant fraction for methylated DNA in NR8052 was 10.9 x 10^{-6}, the revertant fraction for unmethylated DNA in NR9162 was 6.1 x 10^{-6}.

where K is the rate constant, t is the incubation time, and F' is the net increase of revertant fraction obtained by subtracting revertant fraction at time 0 from revertant fraction at time t (12, 24). Because DNAs were transfected into the methyl-directed DNA-repair deficient strain (NR9162) and DNA-uracil glycosylase-deficient strain (NR8052) in this experiment, the interference from different repair efficiencies toward the G·U and G·T mispairs should be minimal. In addition, the fact that revertant fraction increased with time of incubation further substantiates that the difference of reversion rate observed is not due to repair processes after the transfection.

Incubation of DNA with *Hpa*II Methylase and CpG Methylase in the Absence of AdoMet.

Shen *et al.* (17) concluded that deamination at a G·C base pair is a specific consequence of DNA methylase action in the absence of AdoMet. Their analysis involved methylation of DNA *in vitro* by *Hpa*II methylase. Since this enzyme acts upon only CpG dinucleotides within a CCGG sequence, we carried out a similar experiment, using *Spiroplasma* CpG methylase instead. This enzyme methylates at all CpG sites in duplex DNA, thus showing sequence specificity similar to that of mammalian DNA methyltransferase. 400 ng of mutant RF1 DNA was incubated at 37 °C for 16 hours with 0.4 U/μl *Hpa*II methylase or 0.8 U/μl CpG methylase in the 1X reaction buffer provided by the enzyme supplier (New England Biolabs). The incubation time and enzyme concentrations were based on those used by Shen *et al.* (17) and the maximum concentrations of enzymes possible while keeping the glycerol concentration at no more than 5% in the reaction to avoid unspecific activity. Incubation of the same amount of DNA without either methylase in *Hpa*II methylase reaction buffer was used as background

control. After incubation, DNA was purified by phenol-chloroform extraction and transfected into *E. coli* NR8052 (*ung* $^-$). Plating and selecting revertants were the same as described above. The results are summarized in Table 2.1.

To our surprise, our results showed no increase of revertant fraction after 16 hours of incubation. The revertant fraction for the control is consistent with that for unmethylated DNA at time 0 (see Figure 2.2). To test the possibility that uracil-DNA glycosylase contamination in commercial methylase preparations contributed to our failure to observe methylase-induced cytosine deamination in Table 2.1, we added an excessive amount of uracil-DNA glycosylase inhibitor (10,000 units, Bennet et al., 1993) to the incubation reaction. We also incubated the DNA with 5-fold higher concentration of *HpaII* DNA methylase (2 units per μl of the reaction volume). The results are summarized in Table 2.2. Addition of uracil-DNA glycosylase inhibitor did not increase the reversion frequency, but a 2.7- to 4.6-fold increase was observed when DNA was incubated with five times more *HpaII* methylase. We are not sure if the nature of this effect is the same as reported by Shen et al.(17) because of the significant quantitative difference. They observed a 10^4-fold increase of C → U transitions when DNA was incubated with 0.4 unit/μl of DNA*HpaII* methylase. Another potential source of difference is the presence of AdoMet in enzyme preparations since DNA *HpaII* methylase was prepared by New England Biolabs using affinity chromatograph with AdoMet as binding ligand. According to Shen et al (17), methylase-catalyzed deamination reaction is very sensitive to SAM inhibition.

Evidence from *in vivo* experiments strongly suggests that cytosine deminations catalyzed by DNA cytosine methyltransferases could not be the

Table 2.1

Reversion assay of deamination induced by HpaII and CpG methylase. 400 ng of mutant RF1 DNA was incubated with (0.4U/μl) or without *Hpa*II methylase in *Hpa*II methylase reaction buffer (50 mM Tris-HCl, pH 7.5/10 mM EDTA/5 mM 2-mercaptoethanol). The same amount of DNA was incubated with 0.8 U/μl CpG methylase in its reaction buffer (50 mM NaCl/10 mM Tris-HCl, pH 7.9/10 mM $MgCl_2$/1 mM dithiothreitol). Incubations were carried out at 37 °C for 16 hours. DNA was then purified by phenol-chloroform extraction and ethanol precipitation. Purified DNA was transfected into *E. coli* NR8052 (*ung*$^-$) by electroporation. Deamination of cytosine residue at the target site in *LacZ*α gene restores the functional α-complementation and produces plaques with dark blue color. Plating and screening of revertants was the same as described above. All enzymes and buffers were from New England Biolabs. According to the supplier, one unit of either methylase was defined as the amount of enzyme required to protect 1 μg of l phage DNA against cleavage by *Hpa*II endonuclease at 37 °C in an hour.

DNA treatment	Plaques scored		Revertant fraction (x 10^6)
	Total	Revertant	
DNA incubation without enzyme	807,180	8	9.9
DNA incubated with *Hpa*II methylase	3,151,800	32	10.2
DNA incubated with CpG methylase	2,951,340	29	9.8

Table 2.2.

DNA cytosine deamination induced by HpaII methylase at higher concentration and in the presence of uracil-DNA glycosylase inhibitor (Ugi). Incubation conditions are the same as described in Table 2.1. One unit of uracil-DNA glycosylase inhibitor inactivates 1 unit of uracil-DNA glycosylase as defined by Bennett et al (1993).

Treatment	Plaques screened		Revertant
	Total	Revertant	fraction (x 10^5)
No incubation	441,530	9	2.0
Buffer and BSA only	108,540	3	2.8
2 units/μl *HpaII* DNA methylase	911,790	117	12.8
2 units/μl *HpaII* DNA methylase and 10, 000 units of Ugi.	1,250,890	68	5.4

major cause of mutational hot spots at the site of cytosine methylation. Wyszynski et al (1994) reported that C to T transition rate at the targeted methylation site was 100-fold higher in *ung*$^+$ cells where uracil-DNA glycosylase is very active and should suppress the occurrence of mutations due to C to U deaminations. Furthermore, the deamination rate at the same methylation site in *ung*$^-$ cells was found to be comparable to that in *ung*$^+$ cells.

The sulfhydryl group in a Pro-Cys dipeptide is absolutely conserved in all enzymes that methylate C-5 of cytosine residues and in thymidylate synthase (25). Considerable evidence supports the reaction mechanism of DNA methyltransferase proposed by Wu and Santi (26). The conserved Cys residue acts as the catalytic nucleophile, which adds at C-6 of cytosine to generate the dihydrocytosine intermediate with anionic character at the 5-position. This carbanion equivalent is sufficiently nucleophilic to condense with one-carbon units. Subsequent β-elimination of 5-H and the enzyme nucleophile generates the methylated cytosine and free enzyme. One of the smallest DNA methyltranferases, *Hha*I, was recently crystallized (27). Its crystal structure indicates that a cleft between two conserved domains is likely the DNA-binding site. The putative nucleophilic Cys residue is within this cleft. This cleft might provide a hydrophobic environment which prevents exposure of the unstable dihydrocytosine intermediate to water. If so, attack at C-6 of cytosine by methylase might not necessarily cause an increase of hydrolytic deamination of cytosine, despite the formation of an inherently labile intermediate. This hypothesis is in agreement with a previous report that no thymine was produced by human methyltranferase during the methylation reaction (28). Although *Hha*I methylase was found to catalyze the release of 5-^{3}H from cytosine residues in DNA into water in the absence

of AdoMet (26), this does not necessarily contradict the idea. The hydrogen exchange may be indirect. 5-^{3}H can be transferred to an amino acid residue of methylase first, and exchange with water after the dissociation of DNA substrate from the enzyme. Further support for this idea must await the solution of a crystal structure for a DNA-bound enzyme.

Based on our results, the deamination rate for 5-methylcytosine in ds DNA is 1.5×10^{-11} s^{-1} while the DNA-cytosine deamination rate was less than 4×10^{-13} s^{-1}. Although this latter rate was too slow to be measured accurately in our experiment, our maximal estimate is close to 7×10^{-13} s^{-1} reported by Frederico *et al.* (12), who used a similar assay and much longer incubation times. Comparing their value for G·C deamination with ours for G·mC, we estimate that methylation at C-5 increases the cytosine deamination rate more than 21-fold. Deamination of cytosine at neutral pH proceeds via an intermediate in which cytosine is protonated at N-3 position (29). 5-Methylation increases cytosine pK value slightly, from 4.45 to 4.6 (30). The higher pK value increases the fraction of protonated molecules at physiological pH. This effect could be further amplified if methylation distorts the structure of a DNA duplex (31), so that normal H-bonding interaction is changed, which could increase the protonation of cytosine. Cytosine has been found to deaminate 1-2 orders of magnitude faster in mismatched base pairs, and it is believed that mismatch increases the accessibility for protonation and so increases the deamination rate (24). Methylation at N-3 position of cytosine increases pK value to 7.4; consequently, the hydrolytic deamination rate is up to 5000 times faster (32, 33).

Cooper and Youssoufian (7) estimated that the frequency of C → T or G → A transitions in coding regions of the human genome is 42-fold higher than that predicted from random mutation. Under the assumption that the

present CpG deficiency represents an equilibrium state between the rate of CpG loss due to deamination and the rate of CpG creation due to random mutation, Sved and Bird (34) estimated that the observed CpG deficiency in vertebrate genomes can be explained by mutation from methylated CpG to TpG/CpA at about 12 times the normal transition rate. The observed high mutation frequency of 5-methylcytosine relative to cytosine is considered to be the sum of the higher deamination rate of 5-methylcytosine and relatively inefficient repair of the resulting G/T mismatch (35). On the other hand, deamination of cytosine generates a G/U mispair, and the uracil is rapidly excised by the abundant and ubiquitous uracil-DNA glycosylase, followed by repair synthesis (36). Recently, a mismatch-specific thymine-DNA glycosylase has been purified from HeLa cells. This enzyme is capable of hydrolyzing the thymine from G/T, C/T, and T/T mismatches (37). It has been proposed that the biological role of this enzyme is the correction of G/T mispairs arising from hydrolytic deamination of 5-methylcytosine (38). However, the lack of sequence specificity (39) suggests that it could also be associated with the G/T mismatches from replication errors. It is not clear whether this enzyme exists in other organisms whose DNA is not methylated at C-5 of cytosine residues. That information would provide a good indication about its biological function. Since we do not know whether this thymine-DNA glycosylase works as efficiently as uracil-DNA glycosylase *in vivo*, the answer to the question whether the difference in deamination rate is amplified by repair processes is still unclear. DNA replication and transcription generate ss DNA; since cytosine residues in ss DNA deaminate much more rapidly than in ds DNA, deamination might occur preferentially. However, this possibility has not been studied.

Although we did not find enzymatic deamination by *Hpa*II and CpG methylase in our experiments, we do not rule out the existence of a "DNA deaminase". dCMP deaminase (dCTP deaminase in *E. coli* and some other bacteria) is an important enzyme in the biosynthesis of thymine nucleotides. Recently, a gene whose product is involved in the posttranscriptional RNA editing of apolipoprotein B was cloned. This gene encodes an "RNA deaminase" catalyzing the deamination of a cytosine residue in apolipoprotein B mRNA to produce a UAA termination codon (40). DNA enzymatic deamination by "DNA deaminase" is still an attractive idea to explain phenomena like repeat-induced point mutation (RIP) (16, 41) or some extreme mutation hot spots at CpG sites found in factor IX (34).

Addition

Shortly after our paper was published, Shen et al (1994) reported their results of a similar experiment. They found that the rate constant for spontaneous deamination of 5-methylcytosine in double-stranded DNA at 37 °C was about 2- to 3-fold higher than that for cytosine. Although our experimental design was similar in principle, there is one major difference. They did not electroporate methylated DNA into mismatch repair-deficient strain (*mutS*$^-$, see Fig. 2.1) after incubation at 37 °C. Their rationale was based on the assumption that methyl-directed mismatch repair is the only general mismatch repair system, and since plamids were prepared from a *dam*$^+$ strain and both stands should be methylated at *dam* sites, therefore methyl-directed mismatch repair will not repair G·T mismatch generated from deamination of 5-methylcytosine. For several reasons, I believe that their rationale is flawed and could resulted in artifactually lower deamination rates for 5-

methylcytosine. First, Claverys and Méjean (1988) reported that in the absence of GATC sites 50% to 70% mismatches can be repaired in a *mutS*- and *mutL*-dependent manner in *E. coli*, indicating the existence of other repair pathway(s). Second, eukaryotic cells do not methylate their DNA at *dam* sites, but they have homologues for both *mutS* and *mutL*. Defects in either gene caused deficiency in DNA mismatch repair and hypermutability (Parsons et al., 1993; Umar et al., 1994), suggesting that MutS and MutL can participate in efficient DNA mismatch repair without a methyl-directed signal. Third, at least mammalian topoisomerase I can nick DNA at the first phosphodiester bond 5' to all eight possible mismatched bases and provide a signal for DNA mismatch repair (Yeh et al., 1994). Fourth, they did not actually test the survival of G·T mismatch in their strain as we did.

REFERENCES

1. Bird, A. (1992) *Cell* **70,** 5–8

2. Li, E., Bestor, T. H., and Jaenisch, R. (1992) *Cell* **69,** 915–926

3. Coulondre, C., Miller, J. H., Farabaugh, P. J., and Gilbert, W. (1978) *Nature* **274,** 775–780

4. Cooper, D. N., and Krawczak, M. (1989) *Hum. Genet.* **83,** 181–188

5. Bird, A. P. (1980) *Nucleic Acids Res.* **8,** 1499–1504

6. Rideout III, W. M., Coetzee, G. A., Olumi, A. F., and Jones, P. A. (1990) *Science* **249,** 1288–1290

7. Cooper, D. N., and Youssoufian, H. (1988) *Hum. Genet.* **78,** 151- 155

8. Nigro, J. M., Baker, S. J., Preisinger, A. C., Jessup, J. M., Hostetter, R., Cleary, K., Bigner, S. H., Davidson, N., Baylin, S., Devilee, P., Glover, T., Collins, F. S., Weston, A., Modali, R., Harris, C. C., and Vogelstein, B. (1989) *Nature* **342,** 705–708

9. Jones, P. A., Rideout (III), W. M., Shen, J., Spruck, C. H., and Tsai, Y. C. (1992) *BioEssays* **14,** 33–36

10. Wang, R. Y-H., Kuo, K. C., Gehrke, C. W., Huang, L-H., and Ehrlich, M. (1982) *Biochim. et Biophys. Acta* **697,** 371–377

11. Ehrlich, M., Norris, K. F., Wang, R. Y-H., Kuo, K. C., and Gehrke, C. W. (1986) *Bioscience Reports* **6,** 387–393

12. Frederico, L. A., Kunkel, T. A., and Shaw, B. R. (1990) *Biochemistry* **29,** 2532–2537

13. Lindahl, T., and Nyberg, B. (1974) *Biochemistry* **13,** 3405–3410

14. Leonhardt, H., Page, A. W., Weier, H-U., and Bestor, T. H. (1992) *Cell* **71,** 865–873

15. Evans, B. E., Mitchell, G. N., and Wolfenden, R. (1975) *Biochemistry* **14,** 621–624

16. Selker, E. U. (1990) *Annu. Rev. Genet.* **24,** 579–613

17. Shen, J-C., Rideout III, W. M., and Jones, P. A. (1992) *Cell* **71,** 1073–1080

18. Roberts, J. D., and Kunkel, T. A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85,** 7064–7068

19. Kunkel, T. A. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81,** 1494–1498

20. Sambrook, J., Fritsch., E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

21. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol.* **154,** 362–387

22. Thomas D. C., Roberts, J. D., and Kunkel, T. A. (1991) *J. Biol. Chem.* **266,** 3744–3751

23. Shenoy, S., Ehrlich, K. C., and Ehrlich, M. (1987) *J. Mol. Biol.* **197,** 617–626

24. Frederico, L. A., Kunkel, T. A., and Shaw, B. R. (1993) *Biochemistry* **32,** 6523–6530

25. Chen, L., MacMillan, A. M., Chang, W., Ezaz-Nikpay, K., Lane, W. S., and Verdine, G. L. (1991) *Biochemistry* **30,** 11018–11025

26. Wu, J. C., and Santi, D. V. (1987) *J. Biol. Chem.* **262,** 4778–4786

27. Cheng, X., Kumar, S., Posfai, J., Pflugrath, J. W., and Roberts, R. J. (1993) *Cell* **74,** 299–307

28. Smith, S. S., Kaplan, B. E., Sowers, L. C., and Newman, E. M. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89,** 4744–4748

29. Shapiro, R. (1981) in *Chromosome Damage and Repair* (Seeberg, E. & Kleppe, K., eds.) 3–18, Plenum, New York

30. Sober, H. A. (ed.) (1968) *Handbook of Biochemistry*: Selected Data for Molecular Biology G12–13, CRC, Cleveland

31. Hodges-Garcia, Y., and Hagerman., P. J. (1992) *Biochemistry* **31,** 7595–7599

32. Sowers, L. C., Shaw, B. R., Veigl, M. L., and Sedwick, W. D. *Mutation Res.* **177,** 201–218

33. Sowers, L. C., Sedwick, W. D., and Shaw, B. R. (1989) *Mutation Res.* **215,** 131–138

34. Sved, J., and Bird, A. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87,** 4692–4696

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

36. Duncan, B. K., Miller, J. H. (1980) *Nature* **287,** 560–561

37. Neddermann, P., and Jiricny, J. (1993) *J. Biol. Chem.* **268,** 21218– 21224

38. Wiebauer, K., and Jiricny, J. (1989) *Nature* **339,** 234–236

39. Brown, T., and Jiricny, J. (1987) *Cell* **50,** 945–950

40. Teng, B., Burant, C. F., and Davidson, N. O. (1993) *Science* **260,** 1816–1819

41. Kricker, M. C., Drake, J. W., and Radman, M. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89,** 1075–1079

Chapter 3.

Natural DNA Precursor Pool Asymmetry and Base Sequence Context as Determinants of Replication Fidelity

Xiaolin Zhang and Christopher K. Mathews

Submitted to *Journal of Biological Chemistry*.

American Society for Biochemistry and Molecular Biology, Bethesda, MD,

January 1995, 23 pages, in press, April, 1995.

Natural DNA Precursor Pool Asymmetry and Base Sequence Context as Determinants of Replication Fidelity*

Xiaolin Zhang and Christopher K. Mathews#

Department of Biochemistry and Biophysics

Oregon State University

2011 Agricultural and Life Sciences Building

Corvallis, Oregon 97331-7305, USA

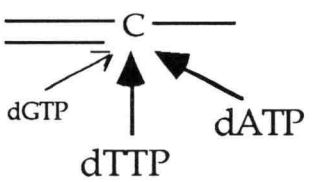
Running title: dNTP pool asymmetries, sequence context, and mutagenesis

SUMMARY

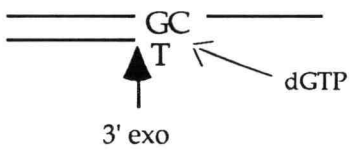
Previous studies showed a complex relationship between nucleotide composition of a gene and the rate of the gene's evolutionary variation. We have investigated mechanisms by constructing M13 phagemids containing part of the *E. coli lacZ* gene, in which an opal codon is flanked either by nine adenine·thymine base pairs on each side, or by nine guanine·cytosine pairs, or by its wild-type sequence context. Reversions or pseudoreversions within the opal codon yield a *lacZ* α-peptide that can undergo α-complementation and yield a blue plaque when plated with a chromogenic substrate. When these constructs were replicated in HeLa cell extracts, in the presence of equimolar deoxyribonucleoside triphosphate (dNTP) mixtures, reversion was near background levels in both the AT-rich and GC-rich contexts. By contrast, when the DNAs were replicated at dNTP concentrations approximating those in HeLa cell nuclei, increases over background were seen in all three contexts. Replication of the phagemids *in vivo* led to even higher mutation frequencies. Replication in the presence of dGMP, added to inhibit proofreading, caused extraordinarily high reversion frequencies in the G·C-flanked opal codon. Apparently, dNTP concentrations approximating intracellular concentrations are mildly but significantly mutagenic, and pool asymmetries and base sequence context both contribute to the natural fidelity of DNA replication.

INTRODUCTION

The studies described in this paper were inspired by a report (Wolfe, Sharp, and Li, 1989) that the rate at which mammalian DNA sequences undergo evolutionary variation is a complex function of the guanine·cytosine content of the sequence, with the highest rates observed in sequences containing about 50 percent G+C. It seemed likely that natural asymmetries in intracellular concentrations of deoxyribonucleoside triphosphates (dNTPs) could be at least partly responsible for the variations observed (Mathews and Ji, 1992). Specifically, dGTP accounts for only five to ten percent of the total pool of the four common dNTPs in most mammalian cell lines that have been studied. Thus, misinsertion opposite template dCMP residues might be relatively frequent, causing increased mutation rates with increasing G+C content. At the same time, replication immediately 5' to a template dCMP residue might be more accurate if dGMP insertion at that site were slow enough to increase the probability of excision of misinserted nucleotides at the upstream site; this effect would decrease the replication error rate as a function of increased G+C content, as schematized below.



Misinsertion



Next-nucleotide effect

It is well known that dNTP pool biases are mutagenic during DNA replication both *in vitro* and *in vivo* (*cf.* Kunkel, 1992; Kunz et al, 1994). The relationship between dNTP pool biases and replication fidelity has become of

special interest with the report that pool imbalances during reverse transcription may be responsible for hypervariability of the HIV genome (Vartanian *et al*, 1994).

Most *in vitro* studies have been carried out either at rather extreme dNTP pool biases, or with unnatural DNA templates, or both. The present studies were designed to ask (1) whether replication of a natural gene sequence, at dNTP concentrations approximating natural pool asymmetries, is mutagenic, and (2) whether the immediate base sequence context influences replicative error rates in ways that would help explain the observed relationships between base composition of a gene and its evolutionary variation.

EXPERIMENTAL PROCEDURES

These studies used the modified phage M13mp2SV, decribed by Roberts and Kunkel (1988). This phage contains in its genome an SV40 DNA replication origin and a mutational target consisting of the first 45 codons for *Escherichia coli* β-galactosidase plus 115 nucleotides of upstream sequence. Expression of the 45 codons of wild-type sequence generates a peptide that can undergo α-complementation when introduced into a host strain, *E. coli* CSH50, which expresses the remainder of the *lacZ* gene. Complementation is scored by plating in the presence of X-gal, a chromogenic substrate for β-galactosidase. A deep blue plaque is scored as wild-type, while mutants yield white or light blue plaques. Constructs for the reversion assays described here were prepared from M13mp2SV by site-directed mutagenesis, using the methods of Kunkel, Roberts, and Zakour (1987).

Other methods were also as described by Roberts and Kunkel (1988) and by Roberts *et al* (1991), including preparation of HeLa cell extracts, preparation and purification of double-strand replicative form DNAs, conditions for SV40 origin-dependent DNA replication catalyzed by HeLa cell extracts, *DpnI* digestion to eliminate unreplicated DNA from analyses, electroporation of replicated DNA, plating on *E. coli* CSH50, and scoring mutations on the basis of plaque color. The host strain for electroporation, *E. coli* NR9162, was *mutS*$^{-}$, to minimize loss of replicational heterozygotes due to mismatch repair. In the reversion assays, plates were incubated for 15-18 hours at 37^{o}, followed by an additional 48-hour incubation at room temperature, to allow detection of the maximum number of mutational events.

dNTP pool measurements were carried out essentially as described by North et al (1980). These analyses, when carried out on the concentrated HeLa

cell extracts used for the replication reactions, confirmed that the dNTPs in these extracts contributed negligibly toward the dNTP concentrations in each reaction mixture (Table 3.1). Also, similar analyses confirmed that dNTP degradation in the replication reactions was negligible (less than five percent of the starting values) over the course of the replication reactions (Table 3.2). Data presented in Table 3.3 shows that no inhibitory effect on dNTP assay reactions in the nucleotide extract was observed, confirming the validity of assay results.

For analysis of the M13mp2SV derivatives that had replicated *in vivo*, COS7 cells were grown in Dulbecco's modified Eagle medium (DMEM) plus 10% fetal bovine serum to about 40% confluency, then washed twice with Opti-MEM reduced serum medium (BRL). For each 100-mm culture dish, 3 μg of RFI DNA and 10 μl of Transfectase reagent (BRL) were diluted to 300 μl with Opti-MEM I reduced serum medium, then mixed together. After standing 20 minutes at room temperature for formation of lipid-DNA complexes, each mixture was diluted to 3.0 ml with the same medium and added to cells treated as described above. Cells were incubated for 10 hours at 37°, then 3.0 ml of DMEM plus 20% fetal bovine serum was added, and incubation was continued for 24 hours. At that point medium was replaced with DMEM containing 10% fetal bovine serum and incubation continued for 12 hours more. After trypsin treatment and centrifugation, each cell pellet was washed with phosphate-buffered saline and resuspended in 200 μl of 50 mM Tris-HCl, pH 7.5, 10 mM EDTA, and 100 μg/ml RNase A. Cells were lysed by adding an equal volume of 0.2 M NaOH and 1% SDS. Chromosomal DNA and cell debris were precipitated by adding potassium acetate, pH 4.8, to a final concentration of 0.44 M, followed by centrifugation. Episomal DNA was purified through Wizard mini-columns (Promega), and unreplicated DNA was eliminated from each mixture by digestion with *DpnI*

prior to electroporation into *E. coli* NR9162 and subsequent analysis for revertants.

Table 3.1

dNTP levels in HeLa cell cytoplasmic extract used for in vitro DNA replication. Proteins and nucleic acids in the cytoplasmic extract were precipitated by 5% ice-cold trichloroacetic acid twice. Trichloroacetic acid was then neutralized by vigorous extraction with 0.5M tri-N-octylamine/Freon 113 until pH was over 6. The aqueous phase was then dried in a Speed-Vac, and the dried powder was dissolved in water for dNTP assays. For a typical 25-µl in vitro DNA replication reaction, 10 µl HeLa cell extract, about 75 µg total proteins, was used.

dNTP	pmole/µl HeLa extract
dATP	0.65
dTTP	2.07
dGTP	0.27
dCTP	0.14

Table 3.2.

dNTP recovery and dNTP degradation during the extraction process. 1.1×10^7 dpm ^{3}H-dATP was added to the sample before the start of the extraction. Total radioactivity in the final solution of the extract was measured and recovery was determined. Labeled dAMP, dADP and dATP in the final extracts were separated by cellulose PEI-F TLC with excess cold dAMP, dADP and dATP (Randerath and Randerath, 1967). Corresponding spots were cut out and radioactivities counted.

^{3}H-dATP added to HeLa cell cytoplasmic extract	1.10×10^7 dpm
^{3}H-dATP recovered in nucleotide extracts	1.05×10^7dpm
Recovery	95.5%

Ratios of mono-, di-, and triphosphate in the final nucleotide extract.

^{3}H-dATP	95.3%
^{3}H-dADP	4.2%
^{3}H-dAMP	0.5%

Table 3.3

Results of internal standard assay. The assay values (in cpm) for standard and nucleotide extract alone were presented. Since the values for the mixture of standard and extract were equal to the addition of them measured separately, these data demonstrate that no inhibitory effects are present in the nucleotide extract sample, and the standard curves are valid.

dNTP	Standard	Extract	Standard + extract
dATP	3,200	9,600	12,000
dTTP	2,500	47,700	50,000
dGTP	1,700	7,000	8,100
dCTP	3,300	17,700	20,300

RESULTS AND DISCUSSION

Originally we contemplated a "global" approach to analyzing the relationship between DNA base composition, dNTP pool asymmetry, and mutagenesis. We planned to modify the G+C content of the 135-b.p. protein-coding part of the *lacZ* gene in M13mp2SV, to values as high and low as possible without changing the amino acid sequence of the encoded *lacZ* α-peptide. We would then replicate these modified constructs *in vitro*, in the presence of dNTP concentrations chosen to represent the approximate levels within HeLa cell nuclei (Leeds *et al*, 1985) and determine the extent to which the natural asymmetry in dNTP levels influenced replication error frequencies. This latter analysis would involve a forward mutation assay (*lacZ*$^+$→*lacZ*$^-$), where mutations anywhere in the 135-b.p. target could be scored as a change in plaque color from dark blue to white or light blue.

However, a preliminary analysis (Table 3.4) indicated that this approach would not be feasible. M13mp2SV DNA was replicated either in the presence of an equimolar dNTP mixture (100 μM each dNTP) or an asymmetric mixture representing the estimated dNTP concentrations in S-phase HeLa cell nuclei (60 μM dATP, 60 μM dTTP, 30 μM dCTP, 10 μM dGTP; Leeds *et al*, 1985). In both cases the replicated DNA samples showed significantly more mutants than unreplicated controls, which were treated identically to the experimental reaction mixtures, except for the omission of SV40 T-antigen during incubations with HeLa cell extract. However, we did not see a significant difference in mutation frequencies between the equimolar and asymmetric dNTP mixtures, and it was apparent that an extremely large number of plates would have to be counted, if we were to learn whether the small difference we did observe was significant.

Table 3.4

Forward mutation assay: *lac*$^{-}$ → *lac*$^{+}$· 0.5 μg of M13mp2SV replicative form DNA was replicated in each assay for six hours, as described by Roberts and Kunkel (1988), at the specified dNTP concentrations. Incorporation of radioactivity from [α-^{32}P]dCTP confirmed that replication was undetectable in the controls incubated in the absence of SV40 T antigen.

dNTP concentrations	T antigen present	Mutant/total plaques counted	Mutant fraction (x 10^4)
Equimolar (100 μM each)	no	10/11,608	8.6
Equimolar	yes	23/13,220	17.4
Biased (60 μM dATP,60 μM dTTP, 30 μM dCTP, 10 μM dGTP)	no	7/9,039	7.7
Biased	yes	20/9,166	21.8

Accordingly, we turned from the global to a more local approach, involving reversion and pseudoreversion events within one codon. Reversion analysis in this system involves scoring blue plaques against a white-plaque background, and this allows inspection of a far larger number of plaques per plate than does the forward mutation assay. For our analysis we chose a serine codon (residue no. 7) in a flexible part of the *lacZα* mutational target. The TCA encoding this serine was changed to an opal codon (TGA), and revertants or pseudorevertants were scored as dark blue or light blue plaque formers.

Extensive sequence analysis of mutants generated during *in vitro* replication of the *lacZα* target in M13mp2 and its derivatives has revealed few null (white-plaque) mutations within this region (*cf.* Kunkel and Alexander, 1986), suggesting that most mutations occurring here allow some retention of wild-type protein function. This means that (1) we can alter the sequences flanking this codon and expect relatively little effect on protein function, and (2) we can expect most single-base substitution errors involving the engineered opal codon to generate a wild or pseudo-wild phenotype, and hence, to be scored as mutational events in a reversion analysis.

Because we are interested in sequence context as a determinant of replication fidelity in the presence of biologically biased dNTP concentrations, we wished to alter the base sequences flanking the local mutational target, namely, the opal codon introduced in place of the serine-7 codon. Accordingly, we designed two sets of flanking sequences, as shown in Table 3.5. In one construct, $(AT)_9TCA(AT)_9$, the opal codon was flanked on each side by nine adenine·thymine base pairs, which generated two conservative changes in the six codons from the wild-type flanking sequence. Design of the other construct, $(GC)_9TCA(GC)_9$, required more changes in order to flank the opal codon with nine guanine·cytosine pairs on each side. However, the changes apparently had

Table 3.5

DNA constructs used in the reversion assay. Aside from the sequence alterations shown, each construct is identical to M13mp2SV. Each altered sequence extends from codon 4 through 10 of the coding sequence for the *lacZ* α-peptide.

DNA Construct	Codon number in *lacZα* peptide						
	4	5	6	7	8	9	10
$(AT)_9TGA(AT)_9$	ATT	AAT	AAT	TGA	TTA	AAT	ATT
	ile	asn	asn	opal	leu	asn	ile
(wt)TGA(wt)	ATT	ACG	AAT	TGA	CTG	GCC	GTC
	ile	thr	asn	opal	leu	ala	val
$(GC)_9TGA(GC)_9$	GCG	CGC	GCC	TGA	CCG	GCC	GGC
	ala	arg	ala	opal	pro	ala	gly

little effect on function of the gene product, because many of the revertant plaques seen in analysis of all three constructs had a deep blue color indistinguishable from that given by the wild-type sequence.

Results of two experiments are summarized in Table 3.6. Several noteworthy results can be seen. First, as noted elsewhere (*cf.* Roberts *et al*, 1991), replication in this *in vitro* system is quite accurate. The DNAs replicated in equimolar dNTP mixtures showed error rates comparable to those seen in the unreplicated controls ("background" in the table). Second, the biological asymmetry in DNA precursor pools apparently does contribute toward the natural mutation rate. In all three constructs the mutant fraction was significantly higher when the DNA was replicated at "biological" dNTP concentrations, biased as described in Table I. This effect was particularly striking in the $(AT)_9TCA(AT)_9$ construct, where the mutant fractions in biological and equimolar dNTP mixtures differed by a factor of 3.5.

The third noteworthy result is the higher mutant fractions observed when the three DNA constructs were replicated *in vivo*, as compared with *in vitro* replication under biologically biased conditions. In the experiment labeled *"in vivo,"* each DNA construct was replicated after transfection into living mammalian cells, rather than *in vitro*. The mutant fractions observed were higher than the corresponding values for the same DNAs replicated *in vitro* at "biological" dNTP levels by about twofold. Of course, factors other than dNTP asymmetries may well contribute toward the error rates seen during replication in living cells. Mismatch repair, for example, could occur *in vivo*, but this would tend to decrease the mutant fractions to values lower than those seen after incubation *in vitro*. In any event, the results are consistent with the hypothesis that biological dNTP pool asymmetries contribute toward the natural replication error rate.

Table 3.6

Reversion and pseudoreversion mutations generated during DNA replication. The DNA constructs described in Table 3.5 were replicated *in vitro* as described for Table I, except that the incubation period was three hours. All dark-blue and light-blue plaques were scored as mutations, and the mutant fraction is the ratio of mutant to total plaques counted. These actual numbers of mutant and total plaques seen after each incubation are shown in parentheses. "Background" denotes DNAs incubated in the absence of SV40 T antigen, where no detectable replication occurred. "Replicated *in vivo*" means replicated in, and isolated from, COS7 cells. ND, not determined. Assay errors are within 50%.

Replication conditions	Mutant Fraction x 10^5		
	(wt)TGA(wt)	$(AT)_9TGA(AT)_9$	$(GC)_9TGA(GC)_9$
		Experiment I	
Equimolar (100 μM each dNTP)	ND	2.2 (9/416,990)	1.1 (9/832,830)
dGTP excess (1mM dGTP, 100 μM dATP, dCTP, dTTP)	ND	5.2 (60/1,158,700)	7.5 (14/186,390)
		Experiment II	
Background	1.6 (15/949,740)	1.1 (8/700,800)	0.58 (17/2,906,400)
Equimolar (100 μM eachdNTP)	3.0 (15/492,360)	1.1 (8/718,200)	0.44 (6/1,365,600)
dGTP excess (1 mM dGTP, 10 μM dATP, dCTP, dTTP)	11.8 (30/255,000)	5.2 (48/928,800)	24.4 (185/759,000)
Inhibited proofreading (100 μM each dNTP plus 0.5 mM dGMP)	17.2 (31/180,000)	9.6 (14/145,800)	61.4 (430/700,800)
Inhibited proofreading (100 μM each dNTP plus 2.0 mM dGMP)	37.8 (106/280,250)	9.7 (17/174,600)	137.5 (706/513,500)
"Biological" dNTP levels (60 μM dATP and dTTP, 30 μM dCTP, 10 μM dGTP)	4.0 (29/724,000)	3.8 (56/1,468,990)	0.8 (7/894,300)
Replicated *in vivo*	7.0 (115/1,641,900)	8.0 (66/826,980)	1.8(26/1,471,800)

Fourth, whether replicated *in vivo* in bacterial cells s flanked by G·C base pairs than with either A·T base pairs or with the natural nucleotides. This may reflect the stability of guanine·cytosine base pairs, which could ("background") or mammalian cells, or *in vitro* in equimolar or biologically biased dNTP pools, replication was significantly more accurate when the opal codon walower the tolerance for insertion of incorrectly base-paired nucleotides. This interpretation is consistent with the relatively large difference in mutant fraction between equimolar and biologically biased pools for the AT-flanked opal codon, described above. Formation of a C·dGTP pair occurs during normal replication of the 3'–ACT–5' trinucleotide in the antisense strand at the opal codon. If mispairing in this site at low dGTP concentrations occurs more readily in an AT-rich sequence context, then the results described in the previous two paragraphs are readily understood. It seems unlikely that variations in mismatch repair are involved, because mismatch repair activities are thought to be low during replication in HeLa cell extracts (Roberts *et al*, 1991).

Fifth, although a GC-rich sequence context seems to promote correct base pairing at the insertion step, maintenance of high fidelity is highly dependent upon proofreading of insertion errors that do occur. Note from the fourth and fifth lines from Experiment 2 in Table III the extraordinary sensitivity of the $(GC)_9TGA(GC)_9$ target to inhibition of proofreading, brought about by addition of a deoxyribonucleoside monophosphate at high concentrations. Error rates increased in all three constructs, but the severalfold *increment* in replication accuracy caused by GC-rich flanking sequences when proofreading was not inhibited was replaced by a *decrement* in replication accuracy, by about an order of magnitude, when proofreading was inhibited.

Essentially the same conclusion can be drawn from the "dGTP excess" experiments. Presumably, the mutations here were caused largely by the next-

nucleotide effect, which involves pool-driven incorporation of nucleotides past the site of a substitution error before that error can be repaired exonucleolytically (Roberts *et al*, 1991). Again, if helix unwinding is slower when the site of an error is flanked by guanine·cytosine pairs, the sensitivity of the $(GC)_9TGA(GC)_9$ construct to mutagenesis under these conditions is easily understood.

To propose an effect of flanking helix stability upon proofreading efficiency in this system is to propose that proofreading in eukaryotic DNA replication involves significant helix unwinding to place the primer terminus in the 3' exonuclease site, as evidently occurs in prokaryotic DNA replication (Beese *et al*, 1993). Whereas structural studies on eukaryotic replication proteins make this a reasonable expectation (*cf.* Wang, 1991; Beckman and Loeb, 1994), it has not been explicitly demonstrated. However, our results are consistent with this model.

The influence of base sequence context upon replication fidelity has long been apparent, simply from the existence of hotspots for spontaneous mutagenesis. However, systematic analysis of this phenomenon has begun only recently. Of particular interest is a recent study of Bloom *et al* (1994), who used pre-steady-state kinetic analysis to analyze 3' exonucleolytic proofreading, and who showed also the influence of helix stability at the primer terminus upon replication accuracy. The system used by these authors involves proofreading of a nucleotide analog, in replication of synthetic DNA templates by a purified DNA polymerase. By contrast, our system involves replication of natural or near-natural DNA sequences by a multiprotein replication apparatus using natural DNA precursors at concentrations that can be adjusted to near-natural levels. Both kinds of analyses should be mutually supportive as investigations of spontaneous mutagenesis continue.

The preliminary results reported here demonstrate, we believe, the utility of this approach to understanding the effects of natural dNTP asymmetries upon replication accuracy and spontaneous mutagenesis. The results suggest a variety of informative approaches to be taken in subsequent investigations, including sequence analysis of the revertants, analysis of different sequence contexts (*e.g.*, AT upstream, GC downstream), and more definitive analyses of the effective dNTP concentrations at eukaryotic DNA replication sites.

Acknowkedgment.Much of the work described in this paper was carried out in the laboratory of Dr. Thomas A. Kunkel, National Institute of Environmental Health Sciences, during separate visits by each of the two authors to that laboratory. We are grateful for hospitality, instruction, and guidance to Dr. Kunkel and to Drs. John D. Roberts and David C. Thomas of that laboratory.

*This work was supported by National Science Foundation Research Grant DMB 9119854. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

#Corresponding author. Tel.: 503-737-1865; Fax: 503-737-0481.

REFERENCES

Beckman, R. A., and Loeb, L. R. (1994) *Quart. Rev. Biophys.* **26,** 225–331

Beese, L. S., Derbyshire, V., and Steitz, T. A. (1993) *Science* **260,** 352–355

Bloom, L. B., Otto, M. R., Eritja, L. J., Reha-Krantz, L. J., Goodman, M. F., and

Beechem, J. M. (1994) *Biochemistry* **33,** 7576–7586

Kunkel, T. A. (1992) *BioEssays* **14,** 303–308

Kunkel, T. A., and Alexander, P. S. (1986) *J. Biol. Chem.* **261,** 160–166

Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods in Enzymol.* **154,** 367–382

Kunz, B. A., Kohalmi, S. E., Kunkel, T. A., Mathews, C. K., McIntosh, E. M., and

Reidy, J. A. (1994) *Mutation Res.* **318,** 175–239

Leeds, J. M., Slabaugh, M. B., and Mathews, C. K. (1985) *Mol. Cell. Biol.* **5,** 3443–3450

Mathews, C. K., and Ji, J. (1992) *BioEssays* **14,** 295–301

North, T. W., Bestwick, R. K., and Mathews, C. K. (1980) *J. Biol. Chem.* **255,** 6640–6645

Roberts, J. D., and Kunkel, T. A. (1988) *Proc. Natl. Acad Sci. USA* **85,** 7064–7068

Roberts, J. D., Thomas, D. C., and Kunkel, T. A. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 3465–3469

Vartanian, J-P., Meyerhans, A., Sala, M., and Wain-Hobson, S. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 3092–3096

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

Wolfe, K. H., Sharpe, P. M., and Li, W-H. (1989) *Nature* **337,** 283–285

Chapter 4

Effects of nucleoside diphosphate kinase deletional mutation on nucleoside triphosphate synthesis and spontaneous mutation rate in *Escherichia coli.*

SUMMARY

Nucleoside diphosphate kinase, encoded by gene *ndk*, is the major enzyme responsible for the synthesis of all nucleoside triphosphates except ATP. There is not a bacteriophage T4 counterpart to the host cell enzyme, and T4 phage must rely on the host to provide the activity for its own synthesis of nucleoside triphosphates. Surprisingly, a deletional mutation of nucleoside diphosphate kinase gene in *Escherichia coli* was neither lethal nor did affect the apparent growth rate, and the mutant cells could still support T4 infection. However, *ndk* mutant cells showed an abnormal expansion of the dCTP pool and a mutator phenotype. In T4-infected mutant cells, the dCTP pool was close to normal, and the host mutator phenotype did not extend to T4 phage. Cells deficient in both NDP kinase and pyruvate kinase still possess 18% of wild-type activity, indicating the existence of a novel source of NDP kinase activity. dNTP pools of anaerobically grown cells with or without T4 infection are reported for the first time. In contrast to previous reports, results from this study demonstrated that pyruvate kinase is unlikely to play a major role in nucleoside triphosphate synthesis either aerobically or anaerobically. The possible relationship between the mutator phenotype and biased dCTP pool in *ndk* mutant cells is discussed.

INTRODUCTION

Nucleoside diphosphate (NDP) kinase catalyzes the last step of all nucleoside triphosphate synthesis except ATP.

$$N_1TP + N_2DP \rightleftharpoons[\text{NDP Kinase}] N_1DP + N_2TP$$

It has long been known as a housekeeping enzyme for the maintenance of cellular levels of ribo- and deoxyribonucleoside triphosphates. In recent years, new aspects for the cellular functions of NDP kinase have attracted much attention. Mammalian NDP kinase, nm23, has been identified as a candidate for tumor suppressor (Steeg et al., 1988). *Drosophila awd* locus, initially isolated as a recessive lethal mutation causing *a*bnormal *w*ing *d*isks and other severe developmental defects (Dearolf et al., 1988a and 1988b), has been found to encode *Drosophila* NDP kinase (Biggs et al., 1990). NDP kinase has also been cloned as a transcriptional factor for *c-myc* in human cells (Postel et al., 1993). Interaction of NDP kinase with G-protein-mediated signal transduction has been reported (Bominar et al., 1993). These observations suggest that NDP kinase plays regulatory roles in cellular functions in addition to being a housekeeping gene.

E. coli and yeast cells have only one known *ndk* gene, while at least two *ndk* genes have been reported for higher eukaryotes (Hama et al., 1991; Dearolf et al., 1988; Kimura et al., 1990; Shimada et al., 1993; and Stahl 1991). Bacteriophage T4 does not have its own NDP kinase gene although it encodes the rest of the genes whose products are required for deoxyribonucleoside triphosphate (dNTP) synthesis (Ray and Mathews, 1992 for review). *E. coli* NDP kinase and T4-encoded enzymes of dNTP biosynthesis function within organized multienzyme complexes that are physically and functionally linked

to the DNA replication apparatus. This multienzyme complex, also called dNTP synthetase complex, helps direct channeling of intermediates and maintaining high local supply of dNTPs for DNA replication forks. Genetic evidence with T4 indicated that dNTP synthetase also plays a role in maintaining DNA replication fidelity (Ji et al., 1989). When cells were infected under conditions where adequate total dNTP levels were maintained but where the complex did not form, there was an increase in spontaneous mutation rate.

In considering the essential roles that NDP kinase plays in *E. coli* and T4 infection, I was surprised to learn that Dr. M. Inouye's laboratory at Rutgers University had isolated a *ndk* deletion mutant which was viable, and more interestingly, possessed of a mutator phenotype. It has been reported that pyruvate kinase has a broad substrate specificity and can transfer phosphate from phosphoenolpyruvate to nucleoside diphosphates in *in vitro* assays. To test whether pyruvate kinase substitutes for NDP kinase in NDP kinase-negative cells, his group deleted the NDP kinase gene in cells where both of the known pyruvate kinase genes, *pykA* and *pykF* were disrupted. I was told that pyruvate kinase- and NDP kinase-negative cells were still viable and also possessed of a mutator phenotype during aerobic growth.

Dr. Inouye kindly sent us the *ndk* mutants, so that I could assess their dNTP pools in hopes of obtaining some clues about the synthesis of nucleoside triphosphates in these cells and the basis of the mutator phenotype. Another interesting question raised is whether the observed mutator phenotype in *E. coli* host will extend to T4 phage infecting these cells. Santos and Drake (1994) reported their results from an intensive screening of host mutator mutations for the effects on spontaneous mutation rates in T4, and they found that rates of spontaneous mutation in T4 were independent

of host fidelity determinants. However, they did not include host mutations that affect dNTP pools in their study. Mutations in T4 causing biased dNTP pools have been known to be mutagenic (Ji and Mathews, 1992). Should abnormal dNTP synthesis be responsible for the host mutator phenotype, the mutagenic effects could extend to T4 phage infecting these cells. I therefore measured T4 spontaneous mutation rates in *ndk* mutant cells and dNTP pool changes after T4 infection.

These studies were later extended to anaerobic growth and infection primarily for two reasons. First, I wanted to further investigate the proposed role for pyruvate kinase in nucleoside triphosphate synthesis to anaerobic growth and infection. Saeki et al (1972) found that *E. coli* cells growing under anaerobic conditions were resistant to the antibiotic desdanine, a specific inhibitor of NDP kinase, but cells growing under aerobic conditions were sensitive to desdanine inhibition. They also found that less amount of NDP kinase was generated while pyruvate kinase production was increased during anaerobic growth. They thus proposed that pyruvate kinase takes the place of NDP kinase in supplying nucleoside triphosphates in *E. coli* at anaerobiosis. But previous results from our laboratory showed that desdanine can completely inhibit T4 DNA synthesis and anaerobic infection was actually more sensitive than aerobic infection (Reddy and Mathews, 1978). Second, I have been curious about the regulatory mechanisms of the newly discovered anaerobic nucleotide reductase (Sun et al., 1993; Young et al., 1994). Anaerobic reductases from either *E. coli* or T4 phage differ from their aerobic counterparts in reducing rNTPs instead of rNDPs. In previous studies of the T4 aerobic ribonucleoside diphosphate reductase, novel insights into allosteric control mechanisms were gained through analysis of dNTP pools in infected cells (Mathews, 1972; Ji et al., 1991). To my knowledge, dNTP pools of

anaerobically grown *E. coli* cells with or without T4 infection have not been reported.

MATERIALS AND METHODS

Media:

LB broth contained 10 g of Bacto-tryptone, 5 g Bacto-yeast extract, and 10 g NaCl per liter. One liter of nutrient broth contained 8 g of Difco nutrient broth and 5 g NaCl, and one liter of nutrient agar contained 4 g Difco nutrient broth, 5 g NaCl and 23 g of Difco nutrient agar. SM9 medium contained 1 g NH_4Cl, 3 g KH_2PO_4, 6 g Na_2HPO_4, 1 g NaCl, 0.01 g $CaCl_2$, 0.2 g $MgSO4{\cdot}7H_2O$, 3 g glucose and 2 g Casamino acids per liter.

SM9 medium was used for all the anaerobic growth, while both LB and nutrient broth were used for the aerobic cultures.

Cells and Phages:

E. coli NDP kinase mutant and their parental strains were kindly provided by Dr. M. Inouye, Rutgers University Medical School. Deletional mutations of NDP kinase gene were confirmed by Southern blotting. E. *coli* B, *E. coli* BB, *E. coli* K38λ, T4D (wild-type T4 phage) and all the T4 phage *rII* mutants were from the collections of this laboratory. *E. coli* strains used in this study and their relevant phenotypes are summarized in Table 4.1

Table 4.1

Cell strains

E. coli cells	Comments	Source
B	Wild type	Our collection
BB	Suppresses *rII* mutations, resulting in *rII*$^+$ growth and normal-sized plaques	Our collection
K38λ	Lysogenic for phage λ, restrictive host for *rII* phage	Our collection
JC7623	parental strain for JC7623Cm^r. *ndk*$^+$*leu*6, *ara*14 *his*4, *thr*1, *thi*1, *lac*Y1, *mte*1, *xyl*5, *gal*K2, *pro*A2, *arg*E3, *str*31, *tsx*3, *sup*37-amber, *rec*B, *rec*C, *sbc*B.	Dr. Inouye
JC7623Cm^r	deletion of NDP kinase gene and insertion of chloramphenicol resistance gene.	Dr. Inouye
HW1387	Parental strain of QL1387. Insertional inactivation of both pyruvate kinase genes (*pykA*$^-$ and *pykF*$^-$) by kanamycin resistance genes. *ndk*$^+$	Dr. Inouye
QL1387	Derived from HW1387, but with deletion of NDP kinase gene and insertion of chloramphenicol resistance gene.	Dr. Inouye

340 nm by a coupled assay as described by Agarwal et al. (1978), in which the oxidation of NADH was followed.

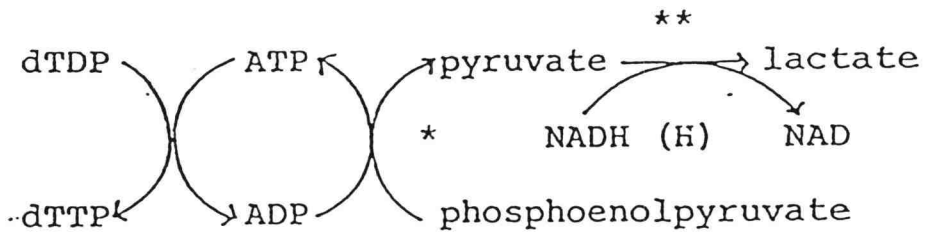


Figure 4.1. Coupled enzymatic reactions for nucleoside diphosphate kinase assay. *, pyruvate kinase; **, lactate dehydrogenase.

The reaction mixture was incubated in cuvettes at room temperature in the presence of ATP as phosphate donor and dTDP as phosphate acceptor nucleotide. 0.5 ml assay buffer contained 50 mM Tris-HCl, pH 7.4, 50 mM KCl, 6 mM $MgCl_2$, 0.1 mM phospoenolpyruvate, 0.1 mg/ml NADH, 0.5 mM ATP, 0.1 mM dTDP, 1 units of pyruvate kinase, 1.5 units of lactate dehydrogenase, and 1 mg/ml bovine serum albumin. All reagents were purchased from Sigma. Water was used to calibrate the Beckman DU-64 spectrophotometer. Self-oxidation of NADH in control reactions without enzyme was negligible, but these values were subtracted from the sample values. Protein concentrations were estimated by the method of Bradford (1976), using bovine serum albumin as a standard. The extinction coefficient for NADH used for the calculation of specific activities was $6.3 \times 10^3\ M^{-1}$.

Measurement of *E. coli* spontaneous mutation frequencies.

E. coli NDP kinase mutants and their parental strains were maintained on LB plates with appropriate antibiotics. Well-isolated single colonies were picked into LB medium and grown to near saturation at 37 °C under aerobic conditions or in an anaerobic growth chamber for anaerobic growth. Aliquots of proper dilutions of the above cultures were plated on LB plates for the determination of total cell counts, and on LB-rifampicin (150 μg/ml) and LB-nalidixic acid (20 μg/ml) plates for selecting rifampicin- and nalidixic acid-resistant mutants.

dCTPase treatment of nucleotide extract.

Crude extract of dCTPase was prepared from an *E. coli* strain overexpressing T4 dCTPase (Ungermann, 1993). After 30 min heat shock induction at 42 °C, cells were harvested by centrifugation, and disrupted by sonication on ice. The supernatant was collected and used as enzyme source.

The treatment of nucleotide extracts with T4 dCTPase was carried out as described by Sargent (1987). 10 μl of the crude extract of T4 dCTPase (about 10 μg total protein) was incubated with an equal volume of nucleotide extract at 37 °C for 30 min. 80 μl of dNTP assay reaction mix was then added to start dNTP assay reactions.

Nucleotide extraction and measurements of dNTPs and rNTPs.

The preparation of nucleotide extracts from cells with or without T4 phage infection was essentially as described by Sargent and Mathews (1987).

Briefly, overnight cultures were diluted into nutrient broth or SM9 medium, and grown to mid-log phase. Cells were then either harvested or infected by T4D at a multiplicity of 6 phages per bacterium, then harvested after indicated infection times. Cells were harvested by rapid filtration through 0.45 μM Millipore type HA filters. After the filtration, filters were immediately immersed in ice-cold 60% methanol/1% toluene for four hours at -20 °C. Methanol extracts were then collected and methanol was evaporated in a Speed-Vac. Acid-soluble nucleotides were extracted with 5% trichloroacetic acid on ice. Trichloroacetic acid was neutralized by vigorous extraction with 0.5 M tri-N-octylamine/Freon-113 until the pH was above 6.0. Extracts in aqueous phase were dried in a Speed-Vac and redissolved in water for dNTP assays.

dNTP measurements were performed using the enzymatic assay essentially as described by North et al. (1980). rNTP levels were determined using the HPLC method (Chen et al., 1977). The four common rNTPs were separated on Whatman Partisil-10 SAX columns and the peaks were detected by a UV detector at 254 nm and integrated electronically.

rII Reversion test.

Measurement of T4 phage *rII* reversion rates were conducted as described by Drake (1970). Fresh *E. coli* BB cells were grown to a cell density of 2 to 5 x 10^8 cells per ml and diluted to 2 x 10^7 cells/ml. 10 to 100 *rII* mutant phages were then added to 5-ml diluted cultures. Infections were allowed to proceed for 4 hours at 37 °C and stopped by adding a few drops of chloroform. Lysates were diluted in nutrient broth and plated with permissive host *E. coli*

BB for lysate titers and on non-permissive host *E. coli* K38λ for *rII* mutant phage for determining the titers of wild-type revertants.

RESULTS

Growth of *E. coli* cells with ndk deletional mutation.

Cells were grown under either aerobic or anaerobic conditions, and their growth was followed as increase of absorbance at 595 nm with time. Although NDP kinase is considered to be a house-keeping gene, deletion of the gene does not seem to affect cell growth rate. As shown in Table 4.2, the doubling times for *ndk* mutants and their wild-type parental strains were comparable during both aerobic and anaerobic exponential growth.

Strain HW1387, defective in both known pyruvate kinase genes, grew normally under aerobic conditions, but failed to grow under anaerobic conditions, indicating that functional pyruvate kinase is essential for anaerobic metabolism. Surprisingly, strain QL1387, HW1387 with deletional mutation of NDP kinase gene, showed normal growth rates under both aerobic and anaerobic conditions.

NDP kinase activity in *ndk* mutant cells.

The above growth results were surprising in light of presumed essential functions of NDP kinase. They also suggest that enzyme(s) other than pyruvate kinase must provide the nucleotide kinase activity needed to support cells' growth. I therefore measured the NDP kinase activity of QL1387 cells, which are deficient in NDP kinase and both of known pyruvate kinases. As Table 4.3 shows, QL1387 cells still had 18% of the NDP kinase activity of its *ndk* wild-type parental strain, as determined by the coupled spectrophotometric assay. The remaining activity, presumably due to a new

Table 4.2

Growth rate of *ndk* mutants and their parental strains. Overnight aerobic and anaerobic cultures of each strain were diluted in fresh LB (for aerobic growth) or SM9 (for anaerobic growth) medium. Cell growth was followed by measuring absorbance at 595 nm. Doubling times were calculated for cultures in their exponential growth.

E. coli Strains	*ndk*	Doubling time (minutes)	
		Aerobic growth	Anaerobic growth
JC7623	+	27	143
JC7623Cm^r	-	29	134
HW1387	+	30	No growth
QL1387	-	34	120

Table 4.3

Nucleoside diphosphate kinase activity in extracts of *E. coli* B, HW1387 and QL1387. The specific activities for dTDP to dTTP conversion were measured by the pyruvate kinase-lactate dehydrogenase coupling assay, as described in Materials and Methods. ATP was used as phosphate donor. Results are the average of two independent assays.

E. coli Strains	Specific activity (mmole/mg/min)
E. coli B	5.16
HW1387	3.30
QL1387	0.60

enzyme, was apparently sufficient to maintain normal growth rate. Interestingly, Fukuchi et al (1993) recently reported that yeast with a *ndk* knockout also possessed about 10% of wild-type activity and grew without apparent phenotypic defects. The source of the remaining activity was not clear.

Mutator phenotypes of *ndk* mutants during aerobic and anaerobic growth.

The spontaneous mutation rates for *ndk* mutants and their parental strains during aerobic and anaerobic growth were measured, based on the frequencies of spontaneous production of rifampicin- and nalidixic acid-resistant mutants. Table 4.4 summarizes the results. Both *ndk* mutant strains showed 10- to 16.8-fold higher spontaneous mutation rates over their parental strains when they were grown under aerobic condition, confirming their mutator phenotypes (Inouye and Lu, personal communications). The mutator phenotype of JC7623Cm^r was also observed when cells were grown anaerobically. Since HW1387 failed to grow, direct comparison between QL1387 and its *ndk* positive parental strain was not possible.

Nucelotide pools of *ndk* mutants and their parental strains under aerobic growth.

The dNTP and rNTP pools in these cells were determined, and the results are summarized in Table 4.5. The dNTP pools of *E. coli* B were close to values previously measured in this laboratory (Sargent, 1987). The rNTP pools of the mutants were comparable to their parental strains. The most striking difference was the 23-fold elevation of the dCTP pool in JC7623Cm^r

Table 4.4

Spontaneous mutation frequencies of *ndk* mutants and their parental strains. Cultures from single, well isolated colonies were grown in LB (at aerobiosis) or SM9 (at anaerobiosis) to near saturation. 100 ml to 200 ml of each cultures were then spread on LB-rifampicin or LB-nalidixic acid plates to select resistant mutants. Aliquots of appropriate dilutions of the same cultures were plated on LB plates without antibiotics to determine cell counts. Plates were incubated at 37 ^{0}C for at least 20 hours. ND: not determined. NG; no growth.

E. coli	Spontaneous mutant fraction (x 10^8)		
	Aerobic growth		Anaerobic growth
Strains	Rifampicin^r	Nalidixic acid^r	Nalidixic acid^r
E. coli B	1.20	0.60	0.63
JC7623	1.30	0.40	0.18
JC7623Cm^r	16.70	6.70	1.22
HW1387	ND	0.60	NG
QL1387	ND	5.90	0.43

Table 4.5

Nucleotide pools of *ndk* mutants and their parental strains. Cells were grown in LB medium to mid-log phase and harvested by rapid filtration as described in Materials and Methods. dNTPs were measured by the enzymatic assay method. dNTP extractions and dNTP pool assays are repeated 3-5 times. rNTPs were analyzed by HPLC method. ND: not determined. Unit: pmole per 10^8 cells.

Nucleotide	*E. coli* strain				
	B	JC7623	JC7623Cm^r	HW1387	QL1387
dATP	21.1	25.7	12.4	12.0	9.3
dTTP	48.6	34.4	55.7	20.8	47.2
dGTP	11.6	10.1	64.0	14.5	15.0
dCTP	58.1	57.5	1322.0	39.0	107.0
ATP	1540	1600	1540	ND	ND
UTP	1940	1940	1860	ND	ND
GTP	2010	2000	2000	ND	ND
CTP	1840	1830	2050	ND	ND

over its parental strain JC7623, a sufficient bias to account for the observed mutator phenotype. On the other hand, only about a 3-fold increase of dCTP level over normal was observed in QL1387. I am not certain whether this relatively mild dCTP bias could suffice to be responsible for the somewhat weaker mutator phenotype of QL1387 observed (Table 4.4).

Support of T4-infection by *ndk* mutants.

Both *ndk* mutants and their parental strains supported aerobic and anaerobic T4 infection abundantly (Table 4.6). Phage production from a single infection cycle seems not to be affected at all in *ndk* mutant cells. It should be pointed out that aerobic infection (in LB medium, 4 hours infection) and anaerobic infection (in SM9, over night infection) were carried out under different conditions, and the results presented in Table 4.6 do not necessarily indicate that phage production under anaerobic infection were higher.

rII Reversion test.

I next asked whether the mutator phenotypes of *ndk* mutants extended to T4 phage infecting these cells. The reversion frequencies of six T4 *rII* mutants, representing different reversion pathways at different sequence contexts in *rII* A and B genes, were measured. T4 *rII* mutants cannot form plaques on the restrictive host *E. coli* K38λ, but they can infect the permissive host *E. coli* BB and form normal plaques. Reversions of *rII* mutations to wild-type restore their plaque-forming ability on *E. coli* K38λ cells. Therefore, the ratios of relative titers between permissive (BB) and non-permissive (K38λ) measure the *rII* reversion fractions. Comparison of revertant fractions between *rII* phages produced in wild-type and *ndk* mutant cells can tell

Table 4.6

T4 phage production in wild-type and *ndk* mutant cells. Cultures were grown in LB aerobically or SM9 anaerobically to their mid-log phase. T4D (wild-type T4 phage) were then added at a multiplicity of 6 phages per bacterium. Aliquots of the infected cultures were taken 10 minutes after infection, and uninfected bacterial colony counts were determined by plating on LB plates. Since infected cells will be lysed in about 25 minutes after infection (Abedon, 1994), by counting the number of colonies formed from an infected culture, infection efficiencies were determined, and the results confirmed that virtually all the cells were infected, and the data therefore represent phage production from single infection cycles.

E. coli	T4 phage yield per infected cells (pfu/cell)	
Strains	Aerobic infection	Anaerobic infection
E. coli B	367	520
JC7623	333	215
JC7623Cm^r	533	299
HW1387	267	NG
QL1387	367	78

relative mutation rates of T4 phages in these cells. According to Drake (1970), three-fold or greater changes in reversion rates are considered to be significant. By this criterion, there was no significant increase of *rII* reversion rates after their infection of *ndk* mutant cells, as shown in Table 4.7.

dNTP pools of *ndk* mutants after T4 phage infection.

Several reasons prompted me to measure the dNTP pools in these *ndk* mutants and their parental strains after T4 phage infection. First, NDP kinase has been thought to play an essential role in the synthesis of dNTPs. Since T4 phages do not encode their own NDP kinase and have to rely on the host NDP kinase(s) to synthesize nucleoside triphosphates, what levels of dNTPs can the residual NDP kinase activity maintain in these *ndk* mutant cells after T4 infection? Second, since dNTP synthesis enzymes in T4 infected cells form a multi-enzyme complex, and since host NDP kinase is normally a part of this protein complex, will the absence of this host enzyme disrupt the normal channeling of the dNTP synthesis in the complexes and so severely affect dNTP synthesis?

T4 phage synthesizes 5-hydroxymethyl-dCTP (hm-dCTP) through degradation of host dCTP to dCMP by its own dCTPase. Phage-coded dCMP hydroxymethylase then catalyzes the formation of 5-hydroxymethyl-dCMP from dCMP, and this is further phosphorylated to hm-dCTP. So the third and most compelling reason was whether the phage-coded dCTPase activity would be sufficient enough to wipe out the observed dCTP accumulation in these *ndk* mutants, and whether similar hm-dCTP expansion would be observed in infected cells. That the host mutator phenotype did not extend to T4 phage (Table 4.7) seems to suggest that dNTP pools in infected mutant cells

Table 4.7

T4 phage *rII* reversion rates in *ndk* mutants and their parental cells

rII mutations	*E. coli* strains	*rII* reversion pathways	rII^+ revertants per 10^8 phage
HB74	JC7623	AT to GC and transversions	1.4
	JC7623Cm^r		2.3
	HW1387		2.0
	QL1387		<2.0
HB84	JC7623	AT to GC and transversions	<0.5
	JC7623Cm^r		<0.3
	HW1387		1.7
	QL1387		<0.5
UV183	JC7623	AT to GC	8.9
	JC7623Cm^r		10.2
	HW1387		11.4
	QL1387		<10
UV215	JC7623	AT to GC	4.8
	JC7623Cm^r		4.2
	HW1387		21.9
	QL1387		11.0
UV363	JC7623	GC to AT	35.1
	JC7623Cm^r		53.3
	HW1387		13.8
	QL1387		19.4
SN103	JC7623	GC to AT	14.9
	JC7623Cm^r		18.3
	HW1387		8.5
	QL1387		8.0

may not be as biased. Figure 4.2 shows the change of dNTP pools after T4 infection. Nine minutes after the addition of T4 phage, the hm-dCTP level in JC7623Cm^r was less than 4-fold in excess over its parental strain JC7623, and hm-dCTP in QL1387 was comparable to the level in its parental strain HW1387. Nucleotide extracts from cells 9 minutes post-infection were resistant to T4 dCTPase treatment (see Materials and Methods), suggesting that the data represent the true levels of hm-dCTP. These data demonstrated that dCTP accumulated in the *ndk* mutants to high levels was degraded after T4 infection. Although the influx of dCMP through dCMP hydroxymethylase may be increased due to the T4 dCTPase reaction, the hm-dCTP pool was only mildly expanded, which was consistent with the above results that the mutator phenotype of *ndk* mutants did not extend to T4 phage (Table 4.6).

dNTP pools of wild-type and *ndk* mutant cells under anaerobic growth and infection.

Since NDP kinase-deficient JC7623Cm^r also showed a mutator phenotype under anaerobic growth conditions (Table 4.4), I measured dNTP pools of anaerobic cells to investigate whether abnormal dNTP pool could be a contributing factor. I also analyzed dNTP pools in anaerobically infected cells, hoping to get some insights into the regulatory mechanisms of the newly found anaerobic ribonucleotide reductases in *E. coli* and T4 phage (Sun et al., 1993; Young et al., 1994). The results are presented in Figure 4.3. Overall, cells or T4-infected cells showed lower dNTP pools under anaerobic conditions. The dCTP level in uninfected JC7623Cm^r cells was about 4-fold higher than in its parental strain, while other nucleotide levels in *ndk*

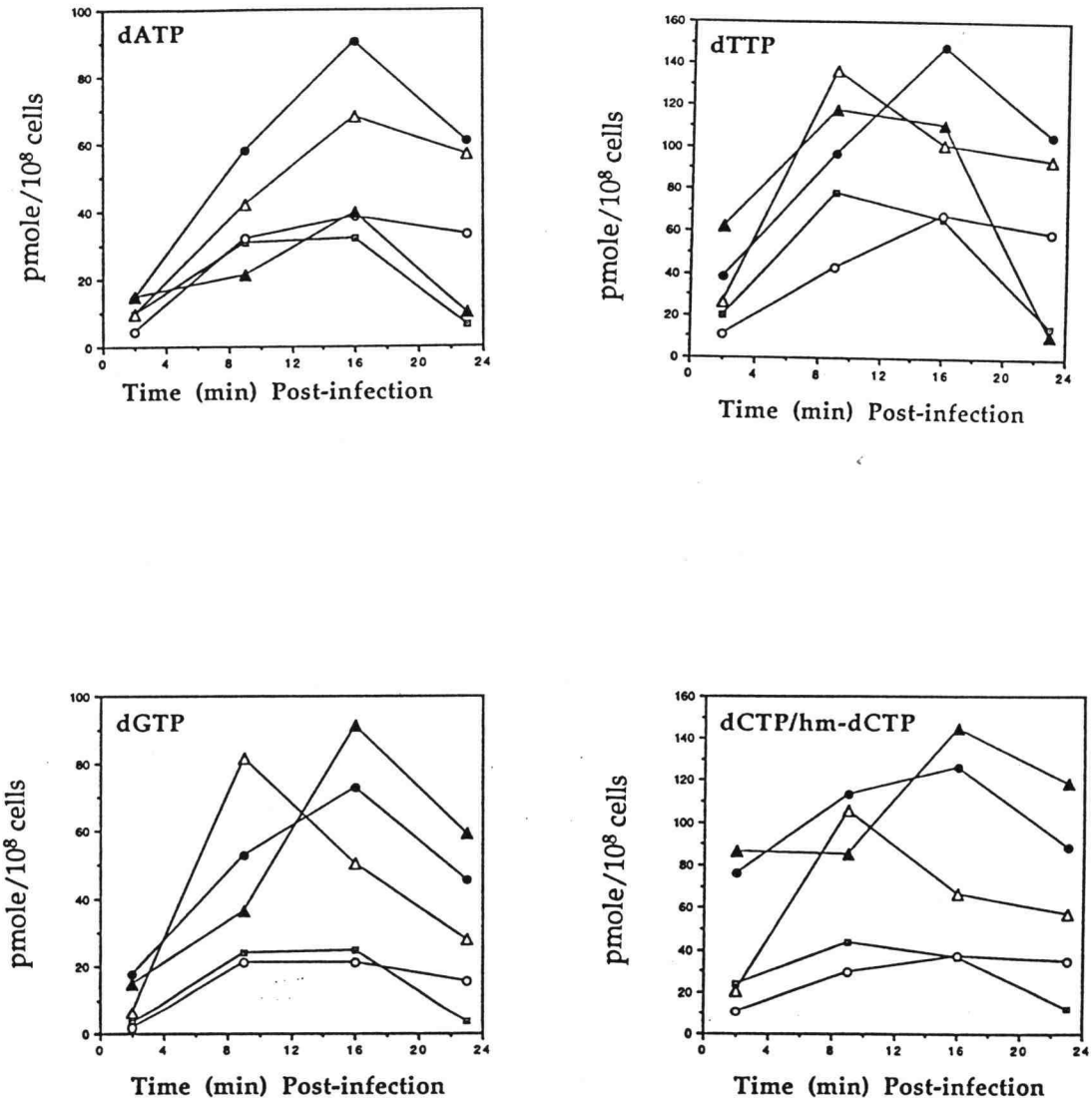


Figure 4.2. dNTP pools after T4D infection under aerobic conditions. Cultures of E. coli B (-▪-), JC7623 (-○-), JC7623Cm^r (-●-), HW1387 (-△-), and QL1387 (-▲-) were grown to a density of 3 x 10^8 cells/ml in nutrient broth. T4D infections were carried out at a multiplicity of 6 phages per bacterium. Infected cells were harvested by rapid filtration after the indicated times. Nucleotide extraction and dNTP pool assays were performed as described by Sargent and Mathews (1987). Unit: pmole per 10^8 cells.

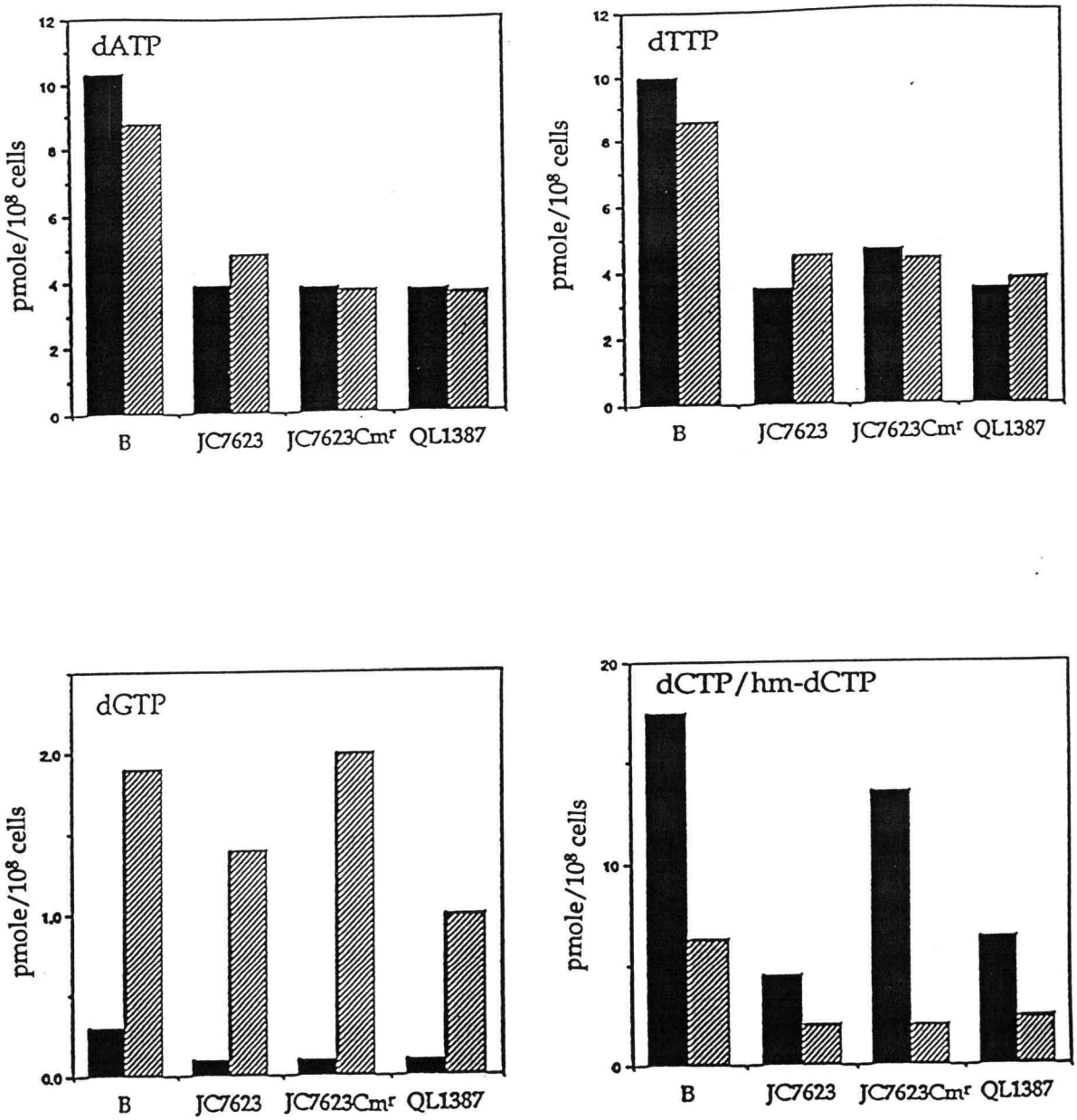


Figure 4.3. dNTP pools before and after T4D infection under anaerobic conditions. Cells were grown and infected anaerobically in SM9 medium. T4D infection and dNTP pool assays were the same as described in legend for Figure 4.2. Data represent the average of two assays. Unit: pmole per 10^8 cells.

mutant and wild-type cells were comparable. A notable change after anaerobic T4 infection was the 6- to 20- fold expansion of dGTP pools in all the strains.

DISCUSSION

The results from this study raised several interesting questions. *E. coli* QL1387 with mutations in both NDP kinase and pyruvate kinase genes is still viable and has about 18% of wild-type NDP kinase activity, suggesting the existence of a novel source of NDP kinase activity in the *ndk* deletion cells. Two NDP kinase genes have been reported in *Dictyostelium discoideum* and in higher animal cells (Troll et al., 1993; Ishikawa et al., 1992; Shimada et al., 1993). In the slime mold *Dictyostelium discoideum*, the two *ndk* genes are separately located in cytosol and mitochondria (Troll et al., 1993). In contrast, only one *ndk* gene was found in yeast when using conserved gene sequences as probes (Fukuchi et al., 1993). It has been suggested that the two *ndk* genes found in slime mold and animal cells were the results of two separate events of gene duplication that occurred recently since the precise conservation of gene structures, both primary sequence and positions of introns, were only found between the duplicated genes themselves (Troll et al., 1993). The finding that the two *ndk* genes in rat cells are arrayed in tandem and only about 3 kilobases apart from each other further enforces the duplication theory (Troll et al., 1993). All known *ndk* genes are well-conserved throughout evolution. For example, 43% identity at amino acid sequence level was found between *E. coli* and human enzymes (Hama, et al., 1991). Since low stringency hybridization using conserved *ndk* sequence as probes failed to pick up other *ndk* genes in *E. coli and* yeast (Hama et al., 1991; Fukuchi et al., 1993; Inouye, personal communication), this novel protein with NDP kinase activity should have little homology with the known *ndk* genes. It has been suggested that pyruvate kinase could take place of NDP kinase in supplying nucleoside triphosphates (Saeki et al., 1975; Fukuchi et al.,

1994). The fact that QL1387, defective in both *ndk* and *pyk* genes shows no growth rate abnormality either aerobically or anaerobically, indicates that pyruvate kinase is unlikely to play any major role in the synthesis of nucleoside triphosphate.

Pyruvate kinase catalyzes the conversion of phosphoenolpyruvate to pyruvate with concomitant generation of one molecule of ATP during glycolysis. Pyruvate plays important roles in cell metabolism, since it can be used directly for biosynthesis of many cell metabolites which have origins in pyruvate or acetyl-CoA. Phosphoenolpyruvate can also be converted to pyruvate through several other pathways, as discussed in Chapter 1. Perhaps the most significant of these is the phosphoenolpyruvate:sugar phosphotransferase system (PTS). Under aerobic conditions, HW1387 and QL1387, which are devoid of pyruvate kinase activity can grow well on sugars catabolized by the PTS, such as glucose, but they grow poorly when a non-PTS sugar is used as energy source (Pertierra and Cooper, 1977; Dr. Inouye, personal communications). The failure of HW1387 to grow anaerobically is consistent with the earlier report by Pertierra and Cooper (1977). They believed that not enough pyruvate was generated in pyruvate kinase mutant cells to permit reoxidation of NADH formed during anaerobic glycolysis, possibly the result of lower PTS activity in anaerobic cells. When they included fumarate (20 mM) as an additional electron acceptor, growth of pyruvate kinase mutant on glucose was obtained.

Of cause, an obvious question then is why QL1387 cells grow anaerobically. Wild-type NDP kinase could somehow suppress a protein, let's call it protein X, whose catalytic activity could produce directly or indirectly more electron acceptors. In *ndk* mutant cells, the suppression was relieved, more protein X could be produced or the activity of protein X could be

enhanced, or both. As a result, sufficient NADH can be oxidized, so QL1387 can survive. At least in animal cells, accumulating evidence indicates that NDP kinase plays important regulatory roles in gene expression and signal transduction (see Chapter 1). In a somewhat related report, Ginther and Ingraham (1974) found surprisingly that NDP kinase activity and concentration fell significantly as the growth rate of *Salmonella typhimurium* was increased. After 3.6-fold increase of growth rate by growing cells on a richer medium, only about 20% of original NDP kinase activity remained. Since the demands for triphosphates must increase at faster growth rate, they believed that this change might have very important effects on the regulation of cellular processes.

NDP kinase activities in *E. coli* and *Salmonella typhimurium* are usually more than 20-fold higher than monophosphate kinase activities which phosporylate nucleoside monophosphates to diphosphates (Ginther and Ingraham, 1974). So even five-fold dimished NDP kinase activity left in *E. coli* and yeast *ndk* mutant cells or in rapidly growing *Salmonella typhimurium* cells, should not be rate-limiting. The fact that no significant difference in growth rates between *ndk* mutants and their wild-type parents further supports the above argument.

What is the basis for the mutator phenotypes of these *ndk* mutant cells? The dramatic expansion of the dCTP pool in JC7623Cm^r, some 20-fold over normal, is biased enough to be mutagenic. That leads to the question why QL1387, which also possesses a mutator phenotype, showed only about a 3-fold increase in the dCTP pool. It may well be that whatever residual or recruited enzyme responsible for nucleoside triphosphate synthesis in *ndk* cells simply has insufficient activity to sustain large pools. It could also be possible that cells defective in both of the known pyruvate kinases have

lower energy levels and are not be able to provide the energy required for the synthesis of excess triphosphates. The assessment of rNTP pools in QL1387 could provide some clues. If energy is the limiting reason, mutant cells may have lower rNTP levels. In conformity with the pool data, QL1387 showed a weaker mutator phenotype when compared with JC7623Cm^r. Similarly, the overall lower dNTP pools in cells grown under anaerobic conditions could also be the consequence of the limited energy supply. Only about a 4-fold increase in dCTP pool was observed in JC7623Cm^r cells in anaerobiosis in contrast to the 23-fold expansion in aerobiosis. Mutational spectra from sequencing enough mutants in a forward assay system could yield informative clues since the majority of the mutations should result from misinsertion of C, or misinsertion followed by C due to next-nucleotide effects should dCTP pool imbalance indeed be responsible.

Why do *ndk* mutants have higher dCTP pools? It could simply be that the novel protein with NDP kinase activity has a higher substrate specificity for dCDP than for other dNDPs. Deletion mutation of *ndk* could also disrupt the control mechanisms of dCTP synthesis. NDP kinase has been reported to be associated with ribonucleotide reductase (RNR) in *E. coli* cells (von DÖbeln, 1976) and with T4 RNR in T4 infected cells (Mathews, 1993). RNR plays a crucial role in balancing the synthesis of all four dNTPs. Coordinated reduction of all four common nucleotides depends on a complicated allosteric control of substrate specificity. The enzyme activity can be modulated by the binding of dATP, dGTP and dTTP. Surprisingly, dCTP is inert, showing neither positive nor negative effect on RNR enzyme activity (Reichard, 1993). Obviously, dCTP synthesis is regulated by other means, and it is tempting to think that the interaction between NDP kinase and RNR plays a role in that regulation. In *ndk* mutants that regulation is missing and that leads to

excessive reduction of CDP to dCDP, which is subsequently phosphorylated to dCTP under aerobic condition, or CTP to dCTP directly in the case of anaerobic growth, since anaerobic RNR reduces nucleotides at triphosphate level. The lethal effect of the *Drosophila awd*$^{k\text{-}pn}$ mutation in *prune* animals (see Chapter 1) provides an example that an *ndk* mutation could result in an unexpected phenotype due to the change of protein-protein interactions with NDP kinase. The measurement of the biochemical parameters of this novel protein and RNR activity in the presence or absence of NDP kinase will certainly shed light on the question. The isolation of the protein with NDP kinase activity in QL1387 is currently being pursued in Dr. Inouye's laboratory and our laboratory.

For several years, reduced or absence of NDP kinase, or nm23 in mammalian cells, has been found to be associated with several types of tumor development and metastasis regulation (Dearolf et al., 1988; Kantor et al., 1993), but the mechanisms remain unclear. It would be interesting to assess the dNTP and rNTP pools of these cells, and check whether they also possess a mutator phenotype since increased mutation rate is associated with carcinogenesis. To my knowledge, a mammalian cell line with disrupted *ndk* gene is not yet available, but *ndk* knockout yeast has been reported recently (Fukuchi et al., 1993). Like *E. coli ndk* mutant, *ndk* yeast showed normal growth rate, spore formation, mating ability and morphology, and also possessed about 10% of the wild-type NDP kinase activity. Informative data such as pool size and mutation rate in mutant yeast cells should be obtainable in a short time. In addition, the yeast system has clear advantages in genetics for further studies.

The ability to support T4 infection (Table 4.6) does not seem to be affected by an *ndk* deletion mutation in the host under either aerobic or

anaerobic conditions. This finding is somewhat surprising in considering that host NDP kinase is the only host enzyme required for *de novo* dNTP synthesis, and it is part of the complex of phage-coded enzymes that carries out dNTP synthesis. NDP kinase normally interacts with other enzymes in the dNTP synthetase complex, and channels dNTPs directly to the replication fork (Ray and Mathews, 1992). How this new protein interacts with T4 machinery for dNTP synthesis and DNA replication remains to be an interesting question.

About 10 minutes after T4 infection of *ndk* mutant cells, the higher dCTP pool disappeared, suggesting high activity of T4-encoded dCTPase. Consistent with the pool data, the mutator phenotype of *ndk* mutants did not extend to T4 phage (Table 4.7). In agreement with the previous report from our laboratory that pyruvate kinase could not substitute for NDP kinase under anaerobic infection conditions (Reddy and Mathews, 1978), *ndk* and *pyk* mutant cells can still support anaerobic T4 infection. Since desdanine, an NDP kinase specific inhibitor, inhibited anaerobic T4 infection, the novel NDP kinase activity must also be sensitive to desdanine.

Conclusions and Future Directions

In this thesis, results from three projects are described. Although the three projects are somewhat distinct, they have all been pursued for one purpose: to understand causes and mechanisms of spontaneous mutation. The first two projects are primarily concerned with *in vitro* investigations, while in the third project, mutagenic effects associated with dNTP pool imbalance are investigated *in vivo*.

Results from the first project demonstrate that chemical instability of 5-methylcytosine toward deamination is the major causal factor relating DNA cytosine methylation to spontaneous mutagenesis. In duplex DNA, the deamination rate for G·mC was found to be 21-fold higher than that for G·C. Two reasons prompted us to explore the structural basis for the apparent difference in chemical reactivity. First, only about a 4-fold difference in deamination rate between 5-methylcytosine and cytosine was observed in single-stranded DNA (Ehrlich et al., 1986). Our results suggest that the difference is amplified in duplex DNA, presumably due to distortion of the duplex structure by the methyl group. Second, double helix structure is known to protect cytosine from deamination, but cytosine deaminates 1-2 orders of magnitude faster in mismatched base pairs, close to the rate observed in single-stranded DNA (Frederico et al., 1993). It is well known that mismatches disrupt the normal double helix structure. During his sabbatical leave at Stockholm University, Dr. Mathews, my major professor, initiated a collaboration with Dr. Astrid Gräslund, Professor of Biophysics at Stockholm University, in Sweden, to investigate the structural change after the methylation, using the sequence about the target G·C and G·mC base pairs in our structure described in Chapter 2: 5' TTCCCCGGTCACG and 5'

TTCCCmCGGTCACG and the complementary strand. Preliminary circular dichroism analysis of the duplexes shows significant spectral difference between those products. The precise nature of the structural difference is currently being investigated by NMR analysis.

Two major conclusions can be drawn from the results of the second project. First, the natural dNTP pool asymmetry in cells is mildly but significantly mutagenic. The "biological" dNTP concentrations used in the study were based on previous measurements of nuclear pool sizes in exponentially growing HeLa cell culture (Leeds et al., 1985). More accurate estimate of intranuclear dNTP pools from synchronized cells and their change through S phase are needed. A new technique using nitrous oxide to block HeLa cells at mitosis (Hozák et al., 1993) could help achieve that goal. Second, DNA sequence context affects DNA replication fidelity. A GC-rich sequence is especially sensitive to the pool imbalance and the presence of a proofreading inhibitor. Our current model is that higher GC content increases helix stability, which could correlate with increased time needed to move a mispaired nucleotide into the 3' exonuclease site for proofreading. On the other hand, our data also suggest that increased flanking sequence stability could enhance insertional accuracy. The critical sequence context for proofreading is that on the 5' side of the incoming incorrect nucleotide—upstream from the site of mutation, but at least DNA polymerase γ can come backward and excise an internal mismatched nucleotide up to 4 residues from the 3' end (Longley and Mosbaugh, 1991), suggesting that sequence on 3' side of mutation could also affect proofreading. It is not clear whether other polymerases can do so too. A "mix and match" experiment could be informative for further dissecting the effect of sequence context on replication fidelity. If the sequence on both sides of the mutation are important, then a

structure with all-AT on one side of the opal codon and all-GC on the other side would be less sensitive to proofreading inhibitor than all-GC on both sides. If only sequence on the 5' side is critical, mutations with all-AT on 5' side and all-GC on the 3' side would be as resistant to inhibition of proofreading as all-AT sequence. Experiments of this type would need to be designed so as to force replication errors to occur primarily on one strand, which can be achieved by adjusting specific nucleotide pool. For example, consider a construct of the following type: (GC)9TGA(AT)9. Replication of the two strands should produce the following intermediates.

```
3'.....(GC)9ACT(AT)9.....        3'....(AT)9AGT(GC)9....
5'.....(GC)9T                    5'....(AT)9T
          ↑                                ↑
        dGTP                             dCTP
```

Errors can be forced to occur predominantly as misincorporation opposite template C, simply by replicating this phagemid in low dGTP. Errors can also be forced during replication of the other strand, by running reactions in low dCTP.

My results about dNTP pools in *E. coli* with *ndk* deletional mutation indicate the existence of a novel NDP kinase activity. Isolation and characterization of the responsible enzyme is currently under the way in both our laboratory and Dr. Inouye's laboratory. Insights can be obtained through accessing some of the basic biochemical parameters of the mystery protein with NDP kinase activity, such as allosteric regulation and substrate specificity. The spectrum of mutations generated from a plasmid after replication in the *ndk* mutant cells could help firmly establish the relationship of the observed mutator phenotype and expansion of dCTP pool. Extension of the work carried out in this research to *ndk* mutant eukaryotic

cells, especially the assessment of dNTP pools and spontaneous mutation frequencies, could potentially yield some important clues about the regulatory roles of NDP kinase plays in tumor development and metastasis regulation.

Bibliography

Agarwal, R. P., Robinson, B. and Parks, R. E. (1978) *Methods Enzymol.* 51, 376-386.

Antequera, F. and Bird, A. (1993) In Jost, J. P. and Saluz, H. P. (ed.), *DNA methylation: molecular biology and biological significance.* Birkhauser Verlag, Basel (Switzerland). p. 169-85.

Antequera, F. and Bird, A. (1993) *Proc. Natl. Acad. Sci. USA* 90, 11995-9.

Antequera, F., Boyes, J., and Bird, A. (1990) *Cell* 62, 503-14.

Antequera, F., and Bird, A. (1992) *EMBO J.* 11, 327-333.

Beese, L. S., Derbyshire, V., and Steitz, T. A. (1993) *Science* 260, 352–355.

Beckman, R. A., and Loeb, L. R. (1994) *Quart. Rev. Biophys.* 26, 225–331.

Bello, L. J. and Besman, M. J. (1963) *J. Biol. Chem.* 238, 1777-1787.

Bennet, S. E., Schimerlik, M. I., and Mosbaugh, D. W. (1993) *J. Biol. Chem.* 268, 26879-26885.

Bestor, T. H. (1992) *Embo J.* 11, 2611-7.

Biggs, J., Trpoulas, N., Hersperger, E., Dearolf, C., and Shearn, A. (1988) Genes Dev. 2, 1333-1343.

Biggs, J., Hersperger, E., Steeg, P. S., Liotta, L. A., and Shearn, A. (1990) *Cell* 63, 933-940.

Bird, A. (1992) *Cell* 70, 5-8.

Bird, A. P. (1980) Nucleic Acids Res 8, 1499-504.

Bloom, L. B., Otto, M. R., Eritja, L. J., Reha-Krantz, L. J., Goodman, M. F., and Beechem, J. M. (1994) *Biochemistry* 33, 7576–7586.

Bloom, L. B., Otto, M. R., Beechem, J. M., and Goodman, M. F. (1993) *Biochemistry*, 32, 11247-11258.

Bominar, A. A., Molijin, A. C., Postel, M., Veron, M. and von Haastert, P. J. M. (1993) *EMBO J.* 12, 2275-2794.

Boyes, J. and Bird, A, (1992) *EMBO J.* 11, 327-333

Bradford, M. M. (1976) *Anal. Bichem.* 72, 248-254.

Brown, T., and Jiricny, J. (1987) *Cell* 50, 945–950.

Chao, J., Leach, M. and Karam, J. (1977) *J. Viol.* 24, 557-563.

Chen, L., MacMillan, A. M., Chang, W., Ezaz-Nikpay, K., Lane, W. S., and Verdine, G. L. (1991) *Biochemistry* 30, 11018–11025.

Chen, S-C., Brown, P. R., and Rosie, D. M. (1977) *J. Chromat. Sci.* 15, 218-221.

Cheng, X., Kumar, S., Posfai, J., Pflugrath, J. W., and Roberts, R. J. (1993) *Cell* 74, 299–307.

Cheng, J-W., Chou, S-H., and Reid, B. R. (1992) *Biochemistry* 228, 1037-1041.

Claverys, J-P and Méjean, V. (1988). *Mol. Gen. Genet.* 214, 574-578.

Cooper, D. N. and Krawczak, M. (1989) *Hum Genet.* 83, 181-8.

Cooper, D. N., and Youssoufian, H. (1988) *Hum. Genet.* 78, 151-155.

Coulondre, c., Miller, J. H., Farabaugh, P. J., and Gilbert, W. (1978) *Nature* 274, 775-780.

Chiu, C. S. and Greenberg, G. R. (1968) *J. Virol.* 46, 629-631.

Dearolf, C. R., Tripolus, N., Biggs, J., and Shearn, A. (1988a) Dev. Biol. 129, 159-168

Dearolf, C. R., Tripolus, N., Biggs, J., and Shearn, A. (1988b) Dev. Biol. 129, 169-178

Donlin, M. J., Patel, S. S., and Johnson, K. A. (1991) *Biochemistry* 30, 538-546.

Drake, J. W. (1970) *The Molecular Basis of Mutation.* Holden-Day. USA.

Drake, J. W. (1991). *Proc. Natl. Acad. Sci. USA* 88, 7160-7164

Duncan, B. K. and Miller, J. H. (1980) *Nature* 287, 560-1.

Ehrlich, M., Norris, K. F., Wang, R. Y., Kuo, K. C. and Gehrke, C. W. (1986) *Biosci. Rep.* 6, 387-93.

Evans, B. E., Mitchell, G. N. and Wolfenden, R. (1975) *Biochemistry* 14, 621-24.

Filipski, J. (1990) In Obe, G. (ed). *Advances in Mutagenesis Research.* 2, 1-54.

Formosa, T., Burke, R. L.,and Alberts, B. M. (1983) *Proc. Natl. Acad. Sci. USA* 80, 2442-2446.

Frederico, L. A., Kunkel, T. A. and Shaw, B. R. (1990) *Biochemistry* 29, 2532-7.

Frederico, L. A., Kunkel, T. A. and Shaw, B. R. (1993) *Biochemistry* 32, 6523-30.

Fukuchi, T., Nikawa, J-I., Kimura, M., and Watanabe. (1993) *Gene* 129, 141-146

Gao, Q-Y., Cara, A., Gallo, R. C., and Lori, F. (1993) *Proc. Natl. Acad. Sci. USA* 90,8925-8928.

Ginther, C. L., and Ingraham, J. L. (1974) *J. Biol. Chem.* 249, 3406-3411

Goodman, M. F., Creighton, S., Bloom, L. B., and Petruska, J. (1993) CRC *Critical Rew. in Biochem. Mol. Bio.* 28, 83-126.

Greenberg, G. R., He, P., Hilfinger, J., and Tseng, M-J. (1994) In Karam, J. D. (ed). Molecular Biology of Bacteriophage T4. American Society for Miocrobiology, Washington, D.C. (USA) p. 14-27.

Hama, H., Almaula, N., Lerner, C. C., Inouye, S., and Inouye, M. (1991) *Gene* 105, 31-36.

Hergersberg, M. (1991) *Experientia* 47, 1171-1181

Hodges-Garcia, Y., and Hagerman., P. J. (1992) *Biochemistry* 31, 7595–7599

Ishikawa, N, Shimada, N., Monakata, Y., Watanabe, K., and Kimura, N. (1992) *J. .Biol. Chem.* 267, 14366-14372.

Johnston, L. H. and Lowndes, N. F. (1992) *Nucleic Acids Res.* 20, 2403-2410.

Ji, J. and Mathews, C. K. (1991) *Mol. Gen. Genet.* 226, 257-264

Ji, J., Sargent, R. G., and Mathews, C. K. (1991) *J. Biol. Chem.* 266, 16289-16292

Jones, P. A. and Buckley, J. D. (1990) *Adv. Cancer Res.* 54, 1-23.

Jones, P. A., Rideout, W. III., Shen, J. C., Spruck, C. H. and Tsai, Y. C. (1992) *Bioessays* 14, 33-6.

Jost, J. P. (1993) *Proc. Natl. Acad. Sci. USA* 90, 4684-8.

Kadrmas, E. F., Ray, P. D. and Lambeth, D. O. (1991) *Biochim. Biophys. Acta* 1047, 339-346.

Kafri, T., Gao, X. and Razin, A. (1993) *Proc. Natl. Acad. Sci. USA* 90, 10558-62.

Kantor, J. D., McCormic, B., Steeg, P. S., and Zetter, B. M. (1993) *Cancer Res.* 53, 1971-1973.

Kavanaugh-Black, A., Connolly, D. M., Chugani., S. A., and Chakrabarty. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 5883-5887.

Kimura, N., Shimada, N., Nomura, K. and Watanabe, K. (1990) *J. Biol. Chem.* 265, 15744-15749.

Konberg, A. A.and Baker, T, A. (1992) DNA Replication. W. H. Freeman and Company. NY, USA.

Knappe, J. (1986) In Neidhardt, F. C. (ed.) *Escherichia coli and Salmonella trphimurium, Cellular and Molecular Biology*. American Society for Miocrobiology, Washington, D.C. (USA) p. 151-155

Kricker, M. C., Drake, J. W., and Radman, M. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 1075–1079

Kunkel, T. A. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 1494–1498

Kunkel, T. A. (1988) *Cell* 53,837-840

Kunkel, T. A. (1992) *BioEssays* 14, 303–308

Kunkel, T. A., and Alexander, P. S. (1986) *J. Biol. Chem.* 261, 160–166

Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods in Enzymol.* 154, 367–382

Kunz, B. A., Kohalmi, S. E., Kunkel, T. A., Mathews, C. K., McIntosh, E. M., and Reidy, J. A. (1994) *Mutation Res.* 318, 175–239

Laird, P. W. and Jaenisch. (1994) *Human Mol. Gent.* 3, 1487-1495

Larsen, F., Gundersen, G., Lopez, R. and Prydz, H. (1992) *Genomics* 13, 1095-1107.

Lascu, I., and Chaffotte, A., and Limbourg-Bouchong, B., and Véron, M. (1992) *J. Biol. Chem.* 267, 12775-12781

Leeds, J. M., Slabaugh, M. B., and Mathews, C. K. (1985) *Mol. Cell. Biol.* 5, 3443–3450

Leeds, J. M. and Mathews, C. K. (1987) *Mol. Cell. Biol.* 7, 532-534.

Leone, A., Flatow, U., King, C. R., Sandeen, M. A., Margulies, I. M. K., Liotta, L. A., and Steeg, P. S. (1991) *Cell* 65, 25-35

Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992) *Cell* 69, 905-914.

Leob, L. A. (1991) *Cancer Res.* 51,1-5

Leonhardt, H., Page, A. W., Weiser, H-U., and Bestor, T. H. (1992) *Cell* 71, 865-873.

Liotta, L. A., Steeg, P. S., and Stetler-steveson, W. G. (1991) *Cell* 64, 327-336.

Longley, M. J. and Mosbaugh, D. W. (1991) *J. Biol. Chem.* 266, 24702-24722

Lor, F., Malykh, A., Cara, A., Sun, D., Weinstein, J. N., Lisziewicz, J., and Gallo, R. C. (1994) *Science* 266, 801-805.

Li, E., Beard, C. and Jaenisch, R. (1993) *Nature* 366, 362-5.

Li, E., Bestor, T. H. and Jaenisch, R. (1992) *Cell* 69, 915-26.

Lindahl, T. (1993) *Nature* 362, 709-15.

Lindahl, T. and Nyberg, B. (1974) *Biochemistry* 13, 3405-10.

MacDonald, P. M. and Hall, D. H. (1984) *Genetics* 107,343-353.

MacDonald, N J., Rosa, A. D. L., Benedict, M. A., Freije, J. M. P. , Krutsch, H and Steeg. P. S. (1993) *J. Biol. Chem.* 268, 25780-25789.

McIntosh, E. M. (1993) *Curr. Genet.* 24, 185-192.

Mathews, C. K. (1972) *J. Biol. Chem.* 247, 7430-7438.

Mathews, C. K. (1989) In Adolph, K. W. (ed) *Molecular Biology of Chromocome Function*. Springer, New York, P461-470.

Mathews, C. K. (1993) *J. Bacterol.* 175, 6377-6381.

Mathews, C. K. and Allen, J. R. (1983) In Mathews, C. K. Kutter, E. M., Mosig, G., and Berget, P. B. (eds) *Bacteriophage T4*. American Sciety of Microbiology, Washinton D. C. p59-70.

Mathews, C. K., and Ji, J. (1992) *BioEssays* 14, 295–301.

Mendelman, L. V., Boosalis, M. S., Petruska, J., and Goodman, M. F. (1989) *J. Biol. Chem.* 264, 14415-14423.

Merril, G. F., Morgan, B. A., Lowndes, N. F., and Johnston. (1992) *BioEssays* 14, 823-830

Meuth, M. (1989) *Exp. Cell Res.* 181, 305-316.

Miller, L. K. and Wells, R. D. (1971) *Proc. Nat. Acad. Sci. USA* 68, 2298-2302

Mun, B-J. and Mathews, C. K. (1991) *Mol. Cell. Biol.* 11,20-26.

Neddermann, P. and Jiricny, J. (1993) *J. Biol. Chem.* 268, 21218-24.

Nigro, J. M., Baker, S. J., Preisinger, A. C., Jessup, J. M., Hostetter, R., Cleary, K., Bigner, S. H., Davidson, N., Baylin, S., Devilee, P., Glover, T., Collins, F. S., Weston, A., Modali, R., Harris, C. C., and Vogelstein, B. (1989) *Nature* 342, 705–708.

North, T. W., Bestwick, R. K., and Mathews, C. K. (1980) *J. Biol. Chem.* 255, 6640–6645.

Parsons, R., Li, G-M, Longley, M. J., Fang, W-H., Papadopoulos, Jen, J, de la Chapelle, A., Kinzler, K. W., Vogelstein, B., and Modrich, P. (1993) *Cell* 75, 1227-1236

Petrska, J. and Goodman, M. F. (1985) *J Biol. Chem.* 260, 7533-7539.

Pertierra, A. G. and Cooper, R. A. (1977) *J. Bacteriol.* 129, 1208-1214

Pless, R. C. and Bessman. (1983) *Biochemistry* 22, 4905-4915.

Postel, E. H., Berberich, S. J., Flint, S. J. and Ferrone, C. A. (1993) *Science* 261, 478-480.

Postel, E. H. and Ferrone, C. A. (1994) *J. Biol. Chem.* 269, 8627-8630

Postma, P. W. (1987) In Neidhardt, F. C. (ed.) *Escherichia coli and Salmonella trphimurium, Cellular and Molecular Biology*. American Society for Miocrobiology, Washington, D.C. (USA) p. 127-141.

Radman, M. (1988) In Kucherlapati, R. and Smith, G. R. (eds.) *Genetic Recombination* American Society for Miocrobiology, Washington, D.C. (USA) p. 169-192.

Randall, S. K., Eritja, R., Kaplan, B. E., Petruska, J., Goodmand, M. F. (1987) *J Biol. Chem.* 262, 6864-6869.

Ray, N. and Mathews, C. K. (1992) *Current Topics in Cellular Regulation.* 33, 343-357

Radazzo, P. A., Northup, J. K., and Kahn, R. A. (1991)*Science* 254, 850-853

Randerath, K. and Randerath, E. (1967) in Grossman, L. (ed.) Methods in Enzymology. 12, 323-347

Razin, A. and Cedar, H. (1993) In Jost, J. P. and Saluz, H. P. (ed.), *DNA Methylation: molecular biology and biological significance* Birkhauser Verlag, Basel (Switzerland). p. 343-57.

Reddy, G. P. V. and Mathews, C. K. (1978) *J. Biol. Chem.* 253, 3461-3467

Reichard, P. (1993) *Science* 260, 1773-1777.

Rosengard, A. M., Krutzsch, H. C., Shearn, A., Biggs, J. R., Barker, E., Margulies, I. M. K., King, C. R., Liotta, L. A., and Steeg, P. S. (1989) *Nature* 342, 177-180

Rideout, W. 3., Coetzee, G. A., Olumi, A. F. and Jones, P. A. (1990) *Science* 249, 1288-90.

Roberts, J. D., and Kunkel, T. A. (1988) *Proc. Natl. Acad Sci. USA* 85, 7064–7068

Roberts, J. D., Thomas, D. C., and Kunkel, T. A. (1991) *Proc. Natl. Acad. Sci. USA* 88, 3465–3469.

Ruggieri, R. and McCormick. (1991) *Nature* 353, 390-391.

Saeki, T., Hori, M., and Umezawa, H. (1975) *J. Antibiotics* 18, 974-981.

Saeki, T., Hori, M., and Umezawa, H. (1972) *J. Antibiotics* 15, 343-349.

Saeki, T., Hori, M., and Umezawa, H. (1974) *J. biochem.* (Japan) 76, 631-637.

Saier, M. H. Reizer, J. (1994) *Mol. Microbiol.* 13, 755-764

Sargent, R. G. and Mathews, C. K. (1987) *J. Biol. Chem.* 262, 5546-5553.

Sargent, R. G. (1987) Ph.D thesis. Oregon State University

Sambrook, J., Fritsch., E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Santos, M. E. and Drake, J. (1994) *Genetics* 138, 553-564

Selker, E. U., Cambareri, E. B., Jensen, B. C., Haack, K. R. (1987) *Cell* 51, 741-752.

Selker, E. U. (1990) *Annu. Rev. Genet.* 24, 579-613.

Shen, J. C., Rideout, W. III. and Jones, P. A. (1992) *Cell* 71, 1073-80.

Shen, J-C., Rideout III, W. M., and Jones, P. A. (1994) *Nucleic Acids Res.* 22, 972-976

Shapiro, R. (1981) in *Chromosome Damage and Repair* (Seeberg, E. & Kleppe, K., eds.) 3–18, Plenum, New York

Shenoy, S., Ehrlich, K. C., and Ehrlich, M. (1987) *J. Mol. Biol.* 197, 617–626

Shimada, N., Ishikawa, N., Munakata, Y., Yuda, T., Watanebe, K. and Kimura, N. (!993) *J. Biol. Chem.* 268, 2583-2589.

Smith, S. S., Kaplan, B. E., Sowers, L. C. and Newman, E. M. (1992) *Proc. Natl. Acad. Sci. USA* 89, 4744-8.

Sober, H. A. (ed.) (1968) *Handbook of Biochemistry*: Selected Data for Molecular Biology G12–13, CRC, Cleveland

Sowers, L. C., Shaw, B. R., Veigl, M. L., and Sedwick, W. D. *Mutation Res.* 177, 201–218

Stafford, M. E., Reddy, P. G. and Mathews, C. K. (1977) *J. Viol.* 23, 53-60.

Stal, J. A., Leone, A., Rosengard. A. M., Porter, L., King, C. R., and Steeg, P. S. (1991) *Cancer Res.* 51, 445-449.

Steeg, P. S., Bevilacqua, G., Kopper, L., Thorgeirsson, U. P., Talmadge, J. E., Liotta, L. A., and Sobel, M. E. (1988) *J. Natl. Cancer Inst.* 200-204

Steeg, P. S., Bevilacqua, G., J. E., Liotta, L. A., and Sobel, M. E. (1988) Cancer Res. 48, 6550-6554.

Sowers, L. C., Sedwick, W. D., and Shaw, B. R. (1989) *Mutation Res.* 215, 131–138

Sun, X., Harder, J., Krook, M., Jörnvall, H., Sjöberg, B-M., and Reichard, P. (1993) *Proc. Natl. Acad. Sci. USA* 90, 577-581.

Sved, J. and Bird, A. (1990) *Proc. Natl. Acad. Sci. USA* 87, 4692-6.

Teng, B., Burant, C. F., and Davidson, N. O. (1992) *Science* 260, 1816-1819.

Teng, D. H. F., Engele, C. M. and Venkatash, T. R. (1991) *Nature* 353,437440.

Thylén, C. and Mathews, C. K. (1989) *J. Biol. Chem.* 264, 15169-15172.

Thomas, D. C., Roberts, J. D., and Kunkel, T. A. (1991) *J. Biol. Chem.* 266, 3744–3751.

Tomich, P. K., Chiu, C. S., Wovcha, M. G. and Greenberg, G. M. (1974) *J. Biol. Chem.* 249,7613-7622.

Trinh, T. Q. and Sinden, R. R. (1991) *Nature* 352, 544-547.

Troll, H., Winckler, T., Lascu, I., Müller, N., Saurin, W., Véron, M., and Mutzel. (1993) *J. Biol. Chem.* 268, 25469-25475.

Trudel, M., Van Genechten, T., and Meuth, M. (1984) J. Biol. Chem. 259, 2355-2359

Ungermann, C. (1993) M.S. thesis. Oregon State University.

Umar, A., Boyer, J. C., Thomas, D. C., Nguyen, D. C., Risinger, J. I., Boyd, J., Inov, Y., Perucho, M., and Knukel, T. A. (1994) *J. Biol. Chem.* 269, 14367-14370

Vairapandi, M. and Duker, N. J. (1993) *Nucleic Acids Res.* 21, 5323-7.

Vartanian, J-P., Meyerhans, A., Åsjö, B., and Wain-Hobson, S. (1991) *J. Virol.* 65, 1779-1788

von Döbeln, U. (1976) *Biochem. Biophys. Res. Commun.* 72, 1160-1168.

Wang, R. Y., Kuo, K. C., Gehrke, C. W., Huang, L. H. and Ehrlich, M. (1982) *Biochim. Biophys. Acta* 697, 371-7.

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

Wheeler, L., Wang, Y., and Mathews, C. K. (1992) *J. Biol. Chem.* 267, 7664-7670.

Wiebauer, K., and Jiricny, J. (1989) *Nature* 339, 234–236

Wu, J. C. and Santi, D. V. (1987) *J. Biol. Chem.* 262, 4778-86.

Wolfe, K. H., Sharpe, P. M., and Li, W-H. (1989) *Nature* 337, 283–285

Wu, C-I. and Maeda, N. (1987) *Nature* 327, 169-170.

Wyszynski, M., Gabbara, S., and Bhagwat, A. S. (1994) *Proc. Natl. Acad. Sci. USA* 91, 1574-1578

Veaute, X. and Fuchs, R. P. P. (1993) *Science* 261, 598-600.

Yeh, Y-C., Liu, H-F., Ellis, C. A., and Lu, A-L. (1994) *J. Biol. Chem.* 269, 15498-15504

Yeivin, A. and Razin, A. (1993) In Jost, J. P. and Saluz, H. P. (ed.), *DNA Methylation: molecular biology and biological significance.* Birkhauser Verlag, Basel (Switzerland). p. 523-68.

Young, P., Öhman, M., Sjöberg, B-M. (1994) *J. Biol. Chem.* 269, 27815-27818