An abundance of DNA polymerases protects our genes

Faithful Copyists

Wojciech Kuban, a PhD student in the laboratory team lead by Dr. Iwona Fijałkowska and Dr. Piotr Jonczyk, explores the mechanisms of DNA replication fidelity in *Eschericha coli* bacteria

WOJCIECH KUBAN
Institute of Biochemistry and Biophysics, Warsaw
Polish Academy of Sciences
kubek@ibb.waw.pl

DNA polymerase is a kind of protein that performs an important function in copying genetic material – a fundamental process in the life of the cell. The high accuracy achieved in replication is only made possible by a wide diversity of polymerase types

If I cut myself while shaving in the morning, I just dab on a bit of cologne and forget about it immediately, quickly leaving for work. But as I stand idly on the commuter train my body is in fact hard at work, immediately starting to rebuild the damaged cuticle and produce new cells. Such things go on every day: our cells have to divide, and young daughter cells have to stand in for old and damaged ones. Nowadays we know that the information about proper cell structure and activity is contained in the cell nucleus. The creation of new cells requires that a copy of such information be made quickly and faithfully.

This is made possible by the special structure of deoxyribonucleic acid (DNA), whose molecules consist of two strands "coupled" to each other through hydrogen bonds. DNA is constructed of four types of nucleotides, symbolically represented by the letters A, C, G, and T. One very important rule for copying genetic information is that an A on one strand should always form a pair with a T on the other strand, while G should always pair with C. Abiding by this rule, the cellular machinery is able to copy genes faithfully: it unravels the DNA strands and by treating one of them as a template, it "writes on" a copy of the other strand.

The molecular scribes that are responsible for this most fundamental function of organisms, the truly painstaking work of replicating DNA, are enzymes called DNA polymerases. The two strands of the DNA "ladder" are first separated from each other, polymerases bind to the DNA chains, and through precise spatial tailoring they "read off" the successive nucleotides and build a second strand based on this template. Both of the separate template strands are copied simultaneously, and the structure of the DNA plus protein complexes during replication is known as a "replication fork" (figure p. 26).

The entire human genome consists of about 6 billion nucleotides. That number is

The fundamental work of copying genes performed by DNA polymerases may be compared to the tedious handwriting of medieval scribes – the fate of the invaluable material for future generations hinges upon their on their precise and faultless work

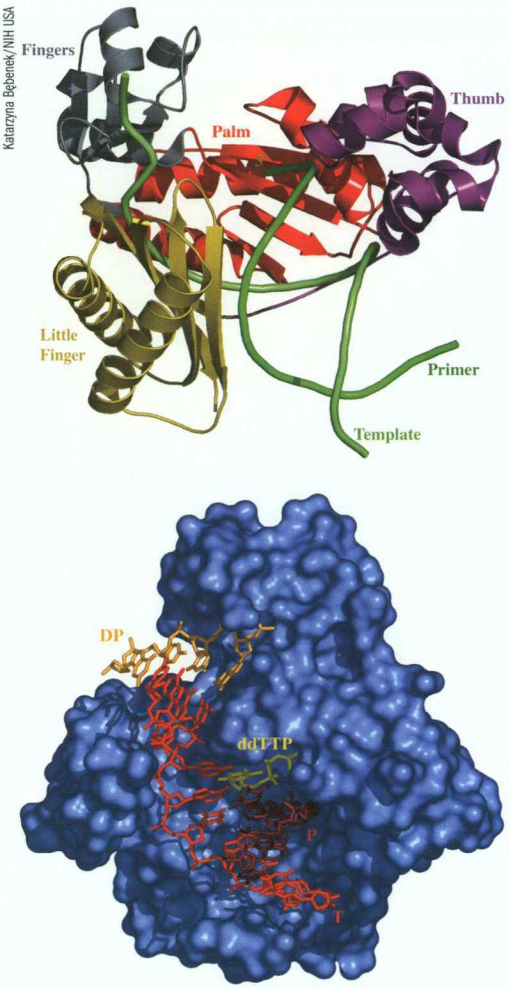
East News

quite abstract, so let's try to put it into more realistic terms: if we start counting at a pace of one number per second, we will reach 1 million after 11 days. The number of nucleotides to be copied in the human genome is nevertheless six thousand times greater than that, so we would have to spend 180 years just to count them all! Moreover, the work of DNA polymerases is further complicated by inter- and intra-cellular factors that harm the DNA template. They can block the synthesis of a newly-emerging DNA strand. How do polymerases cope with such difficulties?

Right-hand protein

The entire process of replication can most easily be compared to typing out a text on a computer. We do our best not to make typos, and this is similar to the first characteristic of DNA polymerases: the precise selection of the proper nucleotide. Polymerases rarely make errors - the most precise enzymes make one mistake for every 100,000 correctly entered nucleotides! Nevertheless, errors do occur. The proofreading function of polymerase allows the erroneous nucleotides to be excised - just like when we press "backspace" on the keyboard. Such action improves the accuracy of replication 100-fold. Amazingly, DNA polymerases are not only very precise, they also perform their work with impressive speed, copying some 1,000 nucleotides per second! When proofreading our text for the final time we make sure to select the "spell-check" function. The cell, too, has such a third-level mechanism for boosting the accuracy of replication, called the "mismatch repair system" (MMR), which augments the precision of the copying work a further 1000 times.

The great precision of DNA replication is chiefly facilitated by the structure and properties of polymerases. They have a shape similar to the right hand, within which several parts, called "domains," can be distinguished: the palm, the fingers, the thumb, and in certain cases an exonuclease domain. The palm contains the active center of the polymerase, which binds nucleotides and bonds them into a DNA chain. The "thumb" and "fingers" help the enzyme stay on the DNA chain, position the template correctly in the active center, and have a significant influence on the pace of replication. Mistakes made by polymerase are corrected by the proofreading exonuclease

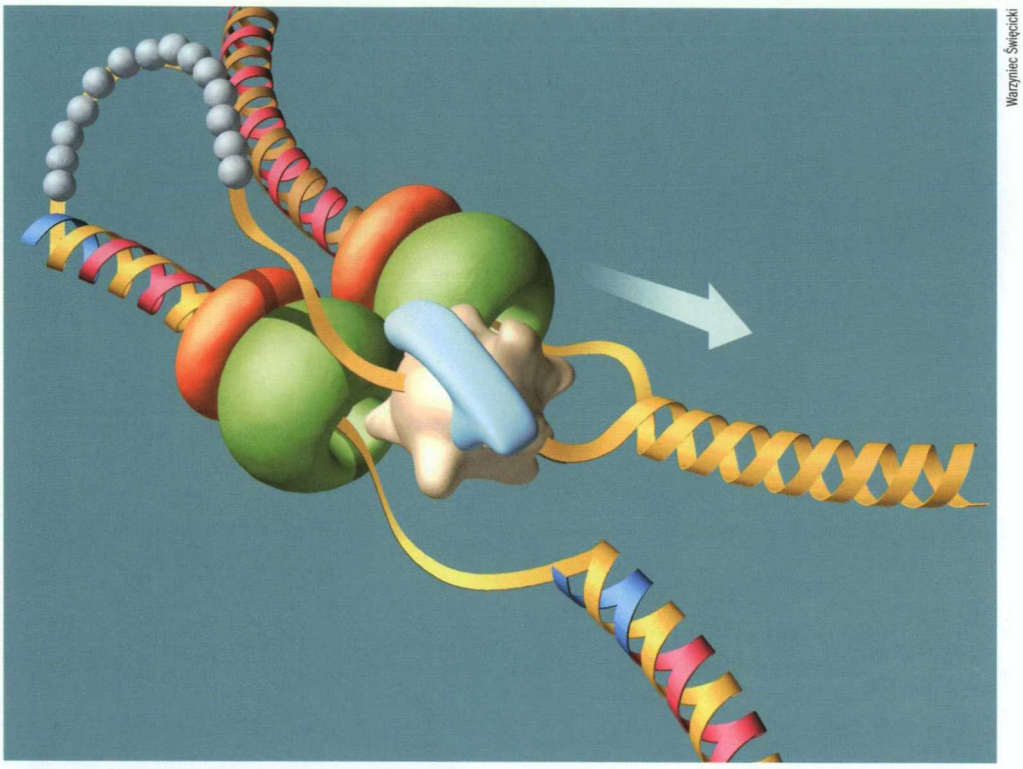


These images of Dpo4 and human Pol λ were created based on their X-ray crystal structures in complex with duplex DNA. (Figure kindly provided by Katarzyna Bebenek, NIH):
A. The structure of Dpo4, a Family Y polymerase from a bacterial virus T7. DNA strands are colored green.
B. DNA Pol λ from human cells, a member of Family X, in complex with a one nucleotide DNA gap. The polymerase, colored blue, is in surface representation. The temple strand (T), the primer strand (P) and the downstream primer (DP) are colored bright red, dark red and gold, respectively. The incoming ddTTP is yellow

domain, which causes the DNA chain to "backtrack" one step and to excise the mistakenly added nucleotide.

Until recently, replication seemed to be a process carried out by only a few enzymes. Nowadays we are discovering a rich array of such proteins, a fact that might seem a considerable extravagance of nature. But is this truly so? Why should cells need such a broad diversity of DNA polymerases?

One big obstacle to replication and the correct expression of genes lies in the modification and damage to DNA caused by physical and chemical factors. DNA damage poses a dramatic threat to the coherence, copying, and conferral of genetic information to daughter cells. In order to counteract various DNA modifications, cells have not just one, but a whole array of types of DNA polymerases, each of them with special properties needed under very specific and exceptional circumstances. Over the past decade, a whole range of new



Warzyniec Święcicki

DNA synthesis is performed in a "replication fork," shown here. One of the strands (the "lagging" strand) is copied discontinuously, the other continuously (the "leading" strand). DNA polymerase (green) is the main replicase, DNA helicase (white) unwinds the parental double helix, DNA primase (blue) produces starters for synthesis, a sliding clamp (red ring) helps hold the polymerase on the DNA, and single-strand DNA binding proteins (white pearls) protect the single strands from breaking. The parental DNA, newly synthesized strands and RNA primers are shown in yellow, pink and blue, respectively

proteins showing DNA polymerase activity have been discovered, and each polymerase discovered is given a name, usually a letter of the Greek alphabet. In total, there are five DNA polymerases in *Escherichia coli* bacteria, more than a dozen in the cells of higher organisms, and 16 in human cells!

The family tree

In basic functional terms, we can divide DNA polymerases into those that replicate (taking responsibility for most of the genome synthesis before the division of the cell) and those that repair (taking part in proofreading processes or able to copy DNA on the basis of a damaged matrix). A more precise characterization, however, will classify them into classes.

DNA polymerases from the A class, similar (homological) to DNA polymerase I in *E. coli*, are involved in excision repair and processing of Okazaki fragments generated during lagging strand synthesis. Class B includes enzymes similar to DNA polymerase II from *E. coli*, and are the basic cellular replicases. Although polymerase II is not essential for bacteria, members of this family in higher organisms' cells are absolutely crucial for the life, as their loss results in cell death. Although they work very rapidly and precisely, they are "not happy" when they encounter DNA damage along their way, they often stop and disconnect from the DNA strand. Class X consists of polymerases homological to the human Pol ß, which specialize in repairing DNA damage. Recent research has led to the identification of a new class Y, which includes polymerases highly specialized in repairing DNA damage. Their structure does not allow them to fit well to the DNA, and so they are less "sensitive" to deformations (damage) in the template. However, this also means that they frequently make mistakes in synthesis and very easily "fall off" the DNA. As a result, they insert a single nucleotide opposite to the damage (even at the expense of an error), enabling the continuity of the DNA strand to be preserved and replication to be continued by other polymerases.

Bacteria as models

Processes as fundamental for cell survival as DNA replication and cell division have been preserved through evolution with relatively little change. That is why they are studied in simple organisms, such as bacteria, but the conclusions reached may likewise hold for human cells. The bacteria *E. coli* was the

cell model used in the present author's own research, concentrating on polymerase IV DNA (DinB) from family Y of DNA polymerases. It does not have proofreading function, it frequently makes errors and easily "falls off" the DNA template. Until now, it was unclear what role this polymerase plays in the replication of chromosomal DNA or what impact it had on the fidelity of replication - even though as many as 250 molecules of DinB are to be found in a single *E. coli* cell.

Under cover

The objective of this work, therefore, was to identify what role DinB plays in the fidelity of replication. We carried out a series of experiments employing the chromosome system previously developed by Dr. Iwona Fijałkowska, making it possible to measure the frequency of errors appearing during bacteria DNA replication. The absence of DNA polymerase IV in the cell turned out not to affect the frequency of the mutations studied. Perhaps DNA polymerase IV performs a special function and its activity is closely controlled by as-yet unrecognized factors? Or perhaps a surplus of the protein, rather than an absence, would affect replication fidelity? Indeed, we confirmed that overproduction of DinB caused a large increase in the number of mutations. Moreover, the unique system for measuring the frequency of changes enabled us to ascertain that DNA polymerase IV more often takes part in the replication of one of the two strands (the so-called lagging strand) of unraveled DNA.

It had previously been postulated that polymerase IV could help extend a strand in cases when the main replicase "gets into trouble," inserts an erroneous nucleotide, and "falls off" the template. To verify this hypothesis, we tested DNA polymerase IV's involvement in chromosome replication when the main portion of replication is carried out by a handicapped version of the main replicase (polymerase III). We employed various mutants: mutators, which make the most errors in DNA synthesis, and antimutators, which more frequently dissociate from the DNA matrix. The idea was that DNA polymerase IV in such mutants has greater access to the end of the strand being synthesized. Experimental results confirm that DinB is more involved in replication in cells with faulty DNA polymerase III, most likely facilitating the extension of the synthesized strand.

When factors harmful to DNA, causing the so-called SOS response, come into action, the number of DinB polymerase molecules in cells abruptly increases to about 2500. We were the first to show that many mutations arising during SOS induction in *E. coli* cells are a consequence of DinB activity, and that it then most likely acts together with another polymerase (V).

The static, textbook scheme of DNA replication is therefore now known to be a simplification of a dynamic, multifaceted process involving a group of protein complexes precisely performing individual tasks - like a well-coordinated team of doctors carrying out a complex operation lasting many hours. The success of this procedure hinges upon their ability to work together, to coordinate their actions in time and space, and to quickly react to the changing condition of the patient and progress of work. Many details of how such molecular teamwork operates remain unknown, but the evidence being discovered of the incredible wealth and abundance of the polymerases employed by nature continue to astound and amaze us. ■

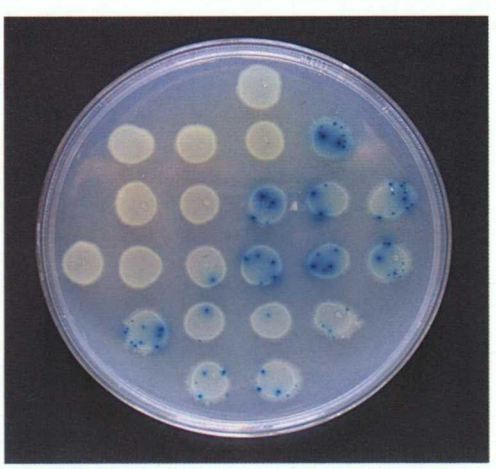
Further reading:

Bebenek K., Kunkel T.A. (2004). Functions of DNA polymerases. *Advances in Protein Chemistry*, 69, 137-65.

Shcherbakova P.V., Fijalkowska I.J. (2006). Translesion synthesis DNA polymerases and control of genome stability. *Frontiers in Biosciences*, 11, 2496-517.

Kuban W., Banach-Orlowska M., Schaaper R.M., Jonczyk P., Fijalkowska I.J. (2006). Role of DNA Polymerase IV in the *E. coli* SOS mutator activity. *Journal of Bacteriology*.

Wojciech Kuban



Processes fundamental for cell survival are studied in simple organisms but the conclusions reached may likewise hold for human cells. In the modified strain of *Escherichia coli* shown here, the number of blue dots reflects the number of errors arising during DNA replication