

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/23656828>

Recent Patents of Gene Sequences Relative to DNA Polymerases

Article · February 2008

DOI: 10.2174/187221508786241729 · Source: PubMed

CITATIONS

4

READS

45

1 author:



[Linda J Reha-Krantz](#)

University of Alberta

68 PUBLICATIONS 1,827 CITATIONS

SEE PROFILE

Recent Patents of Gene Sequences Relative to DNA Polymerases

Linda J. Reha-Krantz*

Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9 Canada

Received: April 16, 2008; Accepted: May 5, 2008; Revised: May 5, 2008

Abstract: Organisms with DNA genomes encode one or more DNA polymerases that are essential enzymes for chromosome replication, DNA repair and recombination. The ability of DNA polymerases to copy DNA templates has been exploited in a variety of *in vitro* reactions to sequence, amplify, mutate, label and recombine DNA, and in several other applications that are fundamental to molecular biology. Because natural DNA polymerases may have activities that interfere with *in vitro* applications or their substrate specificity is too narrow, DNA polymerases have been modified for specific applications. Patents are reviewed here on natural and variant DNA polymerases and their uses.

Keywords: DNA polymerases, Variant DNA polymerases, DNA replication, DNA polymerase applications, DNA sequencing, DNA amplification (PCR), DNA recombination, DNA labeling, Mutating DNA, Evolving DNA polymerases.

INTRODUCTION

DNA POLYMERASES

During each cycle of cell growth and division, DNA polymerases efficiently and faithfully make copies of the DNA strands that are the repository of genetic information. As stated by Arthur Kornberg, "An enzyme, named DNA polymerase [1], was discovered to have the unprecedented property of taking instructions from this template [DNA] and copying it by assembling activated nucleotides into long stretches of DNA, virtually error-free" [2]. DNA from any organism can be replicated in *in vitro* reactions by DNA polymerases as long as there is a primer strand to provide the growing end (3'-hydroxyl group) that is complementary to the DNA to be copied (the template strand). DNA replication also requires the DNA building blocks - deoxynucleoside triphosphates (dNTPs) and Mg^{2+} , which is proposed to catalyze the phosphoryl transfer reaction by a two-metal ion mechanism [3]. The nucleotide incorporation reaction is diagrammed in Fig. (1).

DNA polymerases have been isolated from a multitude of organisms and many others have been inferred from genome sequences. Based on phylogenetic relationships, there are six families of DNA polymerases: A, B, C, D, X, and Y [4,5]. Many organisms have multiple DNA polymerases. For example, while Arthur Kornberg isolated DNA polymerase (pol) I from *Escherichia coli*, which is a family A DNA polymerase that functions in processing Okazaki fragments and in DNA repair, the bacterial chromosome is replicated by the DNA pol III holoenzyme in which the α subunit is a family C DNA polymerase. *E. coli* also has a family B DNA polymerase, DNA pol II, and two family Y DNA polymerases, dinB (DNA pol IV) and UmuC (DNA pol V). DNA pols IV and V replicate DNA with poor fidelity, but they can replicate past DNA lesions that normally block replication by other DNA polymerases [5]. In some cases the translesion replication is accurate as when DNA pol V

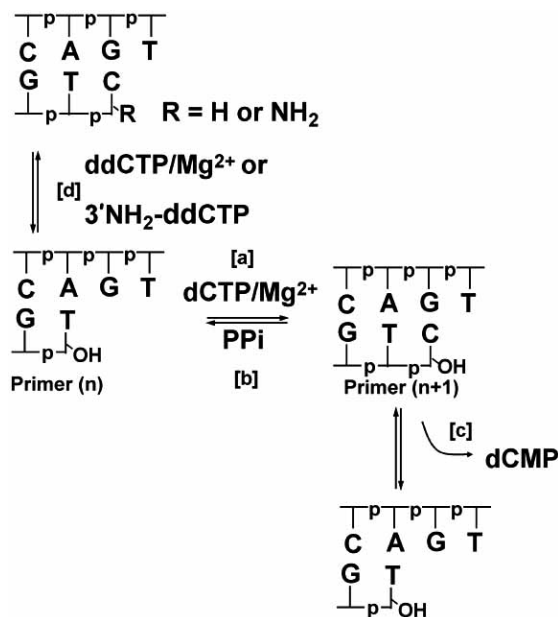


Fig. (1). DNA polymerase nucleotide incorporation, pyrophosphorolysis, and exonucleolytic proofreading activities. DNA polymerases catalyze the template-directed incorporation of nucleotides to synthesize a complementary copy of the template strand [a]. The complementary dNTP, in this case dCTP, is bound in the polymerase active center in a Mg^{2+} -dependent reaction [172,180]. The primer strand is extended by the addition of a dNMP (in this case dCMP) and the release of PPi. PPi is degraded in the cell, but in test tube reactions, PPi can drive the reverse reaction, pyrophosphorolysis [b]. In the absence of the next correct dNTP or if an incorrect nucleotide was incorporated, the end of the primer strand is transferred to the exonuclease active center and the terminal dNMP (in this case dCMP) is removed [c]; this reaction is 3'→5' exonucleolytic proofreading. When a chain-terminating nucleotide is incorporated [d], further primer elongation is not possible. Chain-termination is illustrated for dideoxy (dd)CTP and for the 3'NH₂-ddCTP [d]. The chain-terminating nucleotide may be removed by exonucleolytic proofreading or by pyrophosphorolysis.

*Address correspondence to this Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9 Canada; Tel: 780-492-5383; Fax: 780-492-9234; E-mail: Linda.Reha-Krantz@ualberta.ca

replicates past a cyclobutyl thymine dimer by incorporation of two dAMPs [6,7]. DNA pol II also appears to have a role in the replication of damaged DNA [7].

Human cells have even more DNA polymerases - at least 15, which include the following: three family B DNA polymerases (DNA pols α , δ and ϵ) that are required for chromosome replication; four Family Y DNA polymerases (DNA pols η , ι , κ and REV1) and one family B DNA polymerase (DNA pol ζ) that act in translesion synthesis; and a family X DNA polymerase, DNA pol β , which functions in the base excision repair pathway. Additional DNA polymerases function in the replication of mitochondrial DNA, in the development of antibody producing cells, and in roles that are still being determined. While *E. coli* and human DNA polymerases function optimally at 37 °C, organisms that live at 70 °C and higher temperatures encode thermostable DNA polymerases. DNA polymerases from all six families have been isolated from thermophiles [8,9].

Thus, nature has provided a variety of DNA polymerases, which researchers have used in several ingenious methods to manipulate DNA. Nobel prizes have been awarded for the discovery of a DNA polymerase, *E. coli* DNA pol I (A. Kornberg, [1]), and for the uses of DNA polymerases to sequence DNA (F. Sanger, [10]), to amplify DNA (K. Mullis, [11]), and with "mutant" oligonucleotides to site-specifically mutate DNA (M. Smith, [12]). DNA polymerases are also used in a variety of other molecular genetic applications including DNA labeling, blunt-ending DNA fragments, mutating DNA, isothermal DNA amplification, and in recombining DNA.

DNA POLYMERASE ACTIVITIES

Many DNA polymerases have activities in addition to 5'→3' template-directed nucleotide incorporation (Fig 1). *E. coli* DNA polymerase I, for example, has 5'→3' and 3'→5' exonuclease activities that reside in the same polypeptide as the polymerizing activity, but in separate domains [2,13]. The 3'→5' exonuclease activity is a common feature of many DNA polymerases that are responsible for genome replication. The presence of this exonuclease activity was puzzling at first. Why would a DNA polymerase that replicates DNA in the 5'→3' direction have an exonuclease activity that degrades DNA in the opposite direction? The puzzle was solved when researchers discovered that the 3'→5' exonuclease activity is a proofreading activity that acts preferentially on mismatched bases at the primer terminus [14,15]. Although the α subunit of the *E. coli* DNA pol III holoenzyme does not possess a 3'→5' exonuclease activity in the same polypeptide, the α subunit interacts with the ϵ subunit, which supplies exonucleolytic proofreading activity [2]. The α subunit interacts with the ϵ subunit via its N-terminal php domain [16]. The php domain was identified by sequence similarity to histidinol phosphatase [17]. Interestingly, the α php domain contains a Zn^{2+} -dependent 3'→5' exonuclease that can also function as a proofreading exonuclease [18,19].

DNA polymerases can also degrade DNA by pyrophosphorolysis, which is the reverse of the nucleotide incorporation reaction (Fig. 1). PPi can bind to the DNA polymerase-DNA complex and in the presence of Mg^{2+} a

dNTP is generated while the primer strand is shortened by one dNMP. While pyrophosphorolysis is not a problem in the cell since pyrophosphatases degrade PPi, pyrophosphorolysis is a problem in DNA sequencing reactions.

Archaeal family B DNA polymerases have an activity that so far has not been detected in any other DNA polymerase - the ability to scan ahead for damaged bases. Uracil and hypoxanthine are detected, which are the oxidative deamination products of cytosine and adenine, respectively [20-23]. These damaged bases are produced at all temperatures, but increased production is expected in the high temperature environment in which thermophiles live. If these pro-mutagenic lesions were replicated, GC to AT base substitutions would be produced in the case of uracil and AT to GC substitutions in the case of hypoxanthine. Archaeal DNA polymerases form tight complexes with DNA containing these damaged bases, which prevent mutagenic replication; however, DNA polymerase inhibition by small amounts of damaged DNA is a problem in PCR reactions. Damage recognition takes place via a "pocket" in the N-terminal domain; amino acid substitutions in the pocket of the Pfu DNA polymerase allow read through of uracils in the template strand [21].

DNA POLYMERASE ACCESSORY PROTEINS

Several DNA polymerases function alone in the cell, for example *E. coli* DNA pol I, but others interact with a variety of proteins to increase processivity, to coordinate activities at replication forks, and in repair and recombination complexes. At one extreme is the *E. coli* DNA pol III holoenzyme, which is composed of 15 subunits [2]. The core is composed of α , ϵ and θ subunits, which co-purify as a tight complex. The α and ϵ subunits provide polymerase and proofreading activities, respectively, but little is known about the function of the θ subunit. The remaining subunits form the β clamp and clamp loading complexes [24]. The β clamp is a ring shaped structure that encircles the DNA and confers high processivity to the DNA polymerase. Many eukaryotic replicative DNA polymerases also interact with a sliding clamp. The best known clamp is PCNA (proliferating cell nuclear antigen), which interacts with DNA polymerases and several other proteins at replication forks. At the other extreme of polymerase complexity is the bacteriophage T7 DNA polymerase, which gains high processivity by its interactions with a single protein, the bacterial encoded thioredoxin [25,26], while other replicative DNA polymerases are intrinsically processive, such as the bacteriophage T5 [27,28] and ϕ 29 DNA polymerases [29].

DNA POLYMERASE APPLICATIONS: DNA SEQUENCING

DNA carries the genetic information for each organism, but development of methods to sequence DNA and, thus, to reveal organism-to-organism differences as well as individual-to-individual variation began only about 30 years ago [30]. Since DNA polymerases copy the DNA strands of each chromosome during each cycle of cell growth and division, a logical approach to sequence DNA is to employ DNA polymerases. But simply copying a template DNA does not provide sequence information unless the individual nucleotides can be distinguished from one another. Frederick

Sanger developed an elegant method to observe each successive step of nucleotide incorporation by using 2',3'-dideoxynucleoside triphosphates (ddNTPs) [30]. Because ddNTPs lack the 3'-hydroxyl group, incorporation of a ddNMP produces a primer chain that cannot be elongated further Fig (1). When a nucleotide incorporation reaction is set up with the four standard dNTPs plus one ddNTP, for example ddTTP, a series of chain-terminated DNA strands is produced that can be separated by denaturing polyacrylamide gel electrophoresis as illustrated in Fig (2), which is adapted from the Nobel prize lecture given by F. Sanger [10]. With four separate reactions, each with a different chain-terminating ddNTP, the full complementary DNA sequence can be "read" after the products are separated by denaturing polyacrylamide gel electrophoresis Fig (2).

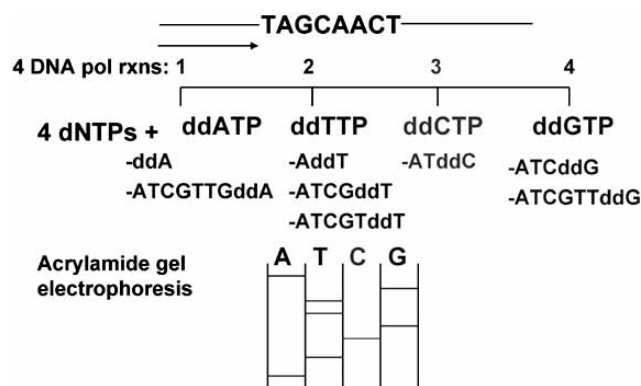


Fig. (2). The Sanger dideoxy/chain terminating DNA sequencing method, adapted from Sanger [10]. Four sequencing reactions were required in the original form of the Sanger method, with each individual reaction containing one of the chain-terminating ddNTPs. For the ddA reaction, reaction #1, the four dNTPs are present plus ddATP. At each position where there is a template T, either a ddAMP or a dAMP may be incorporated. With the optimal ratio of ddATP/dATP, a series of chain-terminated products will be synthesized that represent termination sites at all available template Ts. Chain-terminated products ending either in ddT, ddC, or ddG are synthesized in reactions #2, #3, and #4, respectively. The chain-terminated products from each reaction are then separated by denaturing polyacrylamide gel electrophoresis. In this example, the shortest product is terminated with a ddA, then the next product is terminated with a ddT, and the next product is terminated with a ddC and so on. The complementary DNA sequence is thus A, T, C, G, T, T, A. In the original Sanger sequencing reactions, the DNA products were labeled with radioactivity, but in current methods each chain-terminated nucleotide is labeled with a fluorescent tag, which means that all four chain-terminating reactions can be done in a single tube.

DNA sequencing by chain-termination today is still a powerful DNA sequencing method, but there have been several important improvements. Sanger used the Klenow fragment of *E. coli* DNA pol I in which the N-terminal domain containing the 5'→3' exonuclease activity was removed by protease treatment [31,32]. Thus, DNA polymerases were modified to suit the needs of *in vitro* applications even in the first stages of development of new DNA techniques. Modification of *E. coli* DNA pol I did not

stop with the Klenow fragment and many patents have been deposited on a variety of additional modifications and on several other native and modified DNA polymerases; some are described in this review.

Other improvements to Sanger's dideoxy method of DNA sequencing include the use of fluorophores to tag each chain-terminating nucleotide. The gel electrophoresis step now has increased resolution and sequences from 700 to >800 nucleotides are obtained routinely. The process is also largely automated, which has led to the development of genome centers that are essentially DNA sequencing factories in which numerous "sequencing machines" produce the raw DNA sequence data for entire genomes, including the human genome. New gel electrophoresis-independent sequencing methods, such as pyrosequencing, are also making an impact [33,34].

While current methods to sequence DNA have revealed a wealth of DNA sequence information for many organisms, these methods are expensive and slow. The sequence of the first human genome, J. Craig Venter's [35], is estimated to have cost in excess of \$10 million. However, > 20 gene variants were discovered in Venter's diploid genome that are associated with increased risks for diseases such as alcoholism, cardiovascular problems and Alzheimer's [35]. While this genetic profile could prove invaluable for prevention of disease or for prescribing treatments that could significantly prolong or increase the quality of life, the current price tag is much too high to be considered for routine, individual use. DNA sequencing costs are coming down, but are still expensive; James Watson's genome was sequenced for about \$1 million by the pyrosequencing method [33,34]. The possibility of personal genomics, however, is providing strong incentive to researchers to further improve current DNA sequencing methods and to invent novel methods with the goal to reduce the cost of sequencing an individual human genome to the more reasonable range of \$1000 to \$10,000. Since DNA polymerases are still regarded as excellent sequencing machines, numerous patents have been deposited on DNA polymerases that are being developed for the new DNA sequencing methods.

DNA POLYMERASE APPLICATIONS: DNA AMPLIFICATION (PCR)

The journal *Science* selected the polymerase chain reaction (PCR) as the major scientific development of 1989 and chose for its first "Molecule of the Year" the DNA polymerase molecule that drives the reaction [36]. PCR is a method to amplify selected DNA sequences by using repeated cycles of DNA synthesis from two opposing oligonucleotide primers that flank the target sequence [37]. Each cycle of DNA synthesis is followed by heat denaturation of the DNA strands (> 95 °C) and then cooling to allow the primers to anneal and prime the DNAs synthesized in previous cycles for the next round of primer extension. Thus, each successive cycle doubles the amount of target DNA. Initially, the Klenow fragment of *E. coli* DNA pol I was used to catalyze primer extension since the 5'→3' exonuclease activity of the full length *E. coli* DNA pol I can degrade the primers. But the use of a heat-sensitive DNA polymerase meant that fresh DNA polymerase had to

be added for each cycle, since the enzyme was inactivated during the heat denaturation step. The solution to this problem was to use a thermostable DNA polymerase - the Taq DNA polymerase, a family A DNA polymerase from the thermophile *Thermus aquaticus* [38-42]. The increased reaction temperature used with the Taq DNA polymerase also improved specificity, yield, length of the amplified products and, most importantly, opened the door to automation [40].

The applications of PCR in 1989 when *Science* awarded PCR the major scientific development of that year included disease diagnosis, detection of pathogens such as the agents for Lyme disease and AIDS, cancer diagnosis and research, detection of RNA expression via cDNAs (RT-PCR), detection of trace amounts of DNA at crime scenes, resolution of paternity cases, detection of ancient DNA, matching transplant donors and recipients, gene cloning, and DNA sequencing [36,40]. Today, it is not possible to imagine molecular biological research without PCR. As the list of PCR applications grows, so does the list of patents on improvements to the method and the DNA polymerases that drive the reaction.

One major improvement is to do PCR reactions with a combination of thermostable DNA polymerases, one with and one without exonucleolytic proofreading activity [43-46], in order to increase the length of PCR products and their accuracy. For example, one composition [43,44,46] comprises the Taq DNA polymerase (isolated from the thermophilic eubacterium *Thermus aquaticus*), which naturally lacks 3'→5' exonuclease activity [41,42] and the thermostable archaeal Pfu DNA polymerase (isolated from *Pyrococcus furiosus*), which is a proofreading DNA polymerase. Another composition consists primarily of a thermostable DNA polymerase from *Thermus thermophilus*, again lacking proofreading activity, and a lesser amount of a secondary thermostable DNA polymerase possessing a 3'→5' exonuclease activity from *Thermococcus litoralis*, *Pyrococcus* species GB-D or *Thermotoga maritime* [45]. Accurate amplifications of DNA sequences > 35 kilobases are reported.

DNA POLYMERASE APPLICATIONS: OLIGONUCLEOTIDE-DIRECTED MUTAGENESIS

Michael Smith received the Nobel Prize in Chemistry in 1993 for fundamental contributions to oligonucleotide-based site-directed mutagenesis [12]; the prize was shared with Kary Mullis for development of PCR, which was described above. Thus, two DNA polymerase based techniques were acknowledged in the same year. M. Smith cited in his Nobel Lecture the following quotation from Joshua Lederberg's Nobel prize lecture given in 1959 [47]: "The *ignis fatuus* of Genetics has been the specific mutagen, the reagent that would penetrate to a given gene, recognize it, and modify it in a specific way." *Ignis fatuus* can mean deceptive goal or delusion on one hand or, if one is more optimistic, this term can also mean hope. Lederberg certainly wanted to be hopeful because if geneticists were able to site-specifically mutate DNA, then there would be the potential to uncover the role of each individual DNA base pair in a genome. In 1959 the discovery of such a specific mutagen seemed to be an impossible task; however, with the development of DNA

sequencing and cloning methods, site-directed mutagenesis was within reach.

Oligonucleotide-directed mutagenesis is the mutagenesis method that Lederberg hoped for [48]. Mutation specificity is produced by using chemistry to synthesize an oligonucleotide that is complementary to the target DNA sequence except for the site of the mutation. When the mutagenic oligonucleotide is annealed to the target DNA sequence, which is usually cloned into a plasmid, a mismatch is formed (Fig. (3)). A DNA polymerase is then used to extend the oligonucleotide, which incorporates the mutagenic oligonucleotide into double-stranded DNA. The Klenow fragment was used for the first oligonucleotide-directed mutagenesis procedures for the same reasons that the Klenow fragment was used for DNA sequencing and PCR reactions - the 5'→3' exonuclease activity of the intact enzyme is known to excise most of the primer molecules [49], and would remove the mutagenic oligonucleotide in this application. The DNA polymerase must also replicate DNA with high fidelity to prevent the introduction of non-targeted mutations. The heteroduplex DNA with the incorporated mutagenic oligonucleotide is then transformed into *E. coli* where biological replication produces mutated and non-mutated progeny DNAs. A number of procedures have been developed to preferentially select for the mutated DNA [12,50]. The ease and robustness of oligonucleotide-directed mutagenesis has revolutionized genetic research and has provided the means to carry out precise studies of

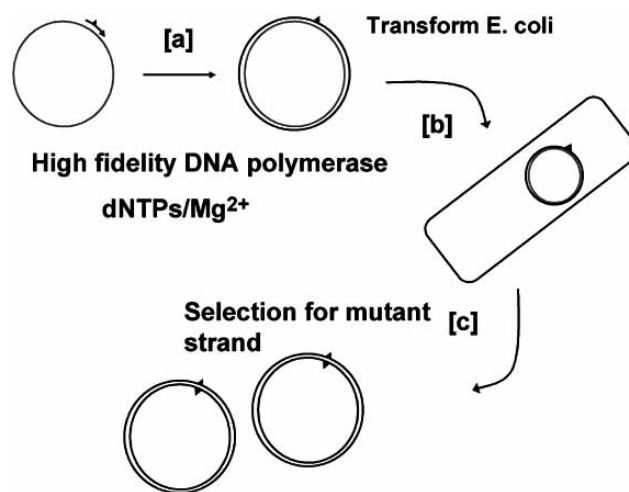


Fig. (3). Oligonucleotide-directed mutagenesis. A mutagenic oligonucleotide is synthesized that is complementary to the mutation target region except for the specific mutation(s). The mutagenic oligonucleotide is then annealed to the template DNA, which creates a mismatch (▲) at the mutant site. The primer is then extended by a DNA polymerase, especially a high fidelity DNA polymerase [a]. Ideally, the primer has a phosphate on the 5'-end so that once synthesis is complete, a DNA ligase will seal the nick. This process creates a heteroduplex DNA with one strand containing the wild type sequence and one strand containing the mutant sequence. When the heteroduplex DNA is replicated in bacteria [b], wild type and mutant copies will be produced, but selection strategies can be used to enrich for production of mutant DNA [c].

protein-structure relationships [51]. As concluded by M. Smith in his Nobel Lecture, "the dilemma posed by Joshua Lederberg in 1959 has been resolved" [12].

OTHER DNA POLYMERASE APPLICATIONS

Molecular biologists have used DNA polymerases for several other applications. A few will be described here to illustrate the versatility of the DNA polymerase as a fundamental tool for DNA manipulations.

Labeling DNA

DNA can be labeled to a high specific activity in *in vitro* DNA replication reactions with radioactively labeled dNTPs [52,53]. *E. coli* DNA pol I catalyzes a reaction at "nicks" in double-stranded DNA that is known as *nick translation*. While the 5'→3' exonuclease activity of the enzyme removes nucleotides from the 5' side of the nick, coupled incorporation of nucleotides at the primer-end results in movement of the nick in the 5' direction. If this reaction is carried out in the presence of [α -³²P]dNTPs, the newly synthesized DNA will be radioactively labeled. Instead of using nicks in the DNA to prime DNA replication, random primers as short as hexamers can be used to prime DNA replication by the Klenow fragment of *E. coli* DNA pol I [54]; the advantage of this method is that the DNA sequence does not need to be known. PCR is yet another method that can be used to label DNA.

Instead of radioactively labeled dNTPs, dNTPs with fluorescent tags can be used to produce a complementary fluorescent DNA. dNTPs with bulky fluorophores or other modifications, however, are not readily incorporated by most DNA polymerases and this has urged the use of other polymerases such as reverse transcriptases, which are more adept in incorporating modified nucleotides, and the development of novel DNA polymerases and patent applications on their use.

Blunt-ending DNA Fragments

DNA fragments produced by shearing or cutting DNA with restriction endonucleases often have protruding 5'- or 3'-ends. The Klenow fragment of *E. coli* DNA pol I is used traditionally to fill in the recessed 3'-ends and labeled dNTPs can be used to label the ends [55]. For DNA with 3' extensions, the bacteriophage T4 DNA polymerase is used to remove the extension [55]. In the absence of dNTPs, the strong 3'→5' exonuclease of the T4 DNA polymerase degrades the 3' protrusion and continues into the duplex DNA. The T4 DNA polymerase can be shifted from being primarily an exonuclease to a polymerase by the addition of dNTPs. Thus, when labeled or a mixture of labeled and unlabeled dNTPs are added to the reaction, the degraded DNA will be "replaced" by new synthesis that can lead to the production of highly labeled DNA. If dNTPs are present at the beginning of the reaction, degradation is limited to removing the 3' protrusion and blunt-ended fragments will be produced.

Mutating DNA

Although DNA polymerases are known for their ability to replicate DNA with high fidelity, some DNA polymerases are error prone and highly accurate DNA polymerases can be

modified to reduce their accuracy. Certain reaction conditions can also increase DNA replication errors. Low fidelity DNA replication is useful if random mutations are wanted over a large portion of a gene. One way to generate one or a few mutations within a cloned gene is to use mutagenic PCR [56]. The Taq DNA polymerase is error prone already because it lacks exonucleolytic proofreading activity. If MnCl₂ is added to the reactions to reduce nucleotide incorporation accuracy [57] and if the dNTP pools are unbalanced by adding high concentrations of dCTP and dTTP, a library of mutant PCR fragments is produced. Repeated cycles of mutagenic amplification coupled with screening or selection for desired products can be used for the *in vitro* evolution of nucleic acids [58,59]. The potential to rapidly evolve nucleic acids and protein encoding genes is fueling further development of "mutator" DNA polymerases and patents for their use.

Isothermal DNA Amplification

PCR is a powerful method to amplify defined DNA sequences and the use of proofreading proficient thermostable DNA polymerases has greatly improved the fidelity of these reactions and the length of products. Isothermal DNA amplification methods, however, provide a useful alternative. One of these methods - rolling circle amplification (RCA) [60,61], exploits the high processivity of the ϕ 29 DNA polymerase [29] and the use of random primers to initiate multiple sites of DNA replication [54]. Circular DNA templates can be amplified 10,000-fold in just a few hours at 30 °C and the products are suitable for *in vitro* cloning and other molecular biology applications. Further improvements to RCA produce amplifications >10⁹-fold of individual DNA molecules of 5 – 7 kb in size [62], truly paving the way for cell-free cloning.

Recombining DNA

The vaccinia virus DNA polymerase is a family B DNA polymerase, but apparently unlike other family B DNA polymerase this enzyme has the ability to catalyze recombination-like reactions [63]. The recombinase activity has been used successfully to develop ligation-independent cloning methods for engineering fusion proteins, modular vectors, and in novel mutagenesis strategies [64-67].

PATENTS

PATENTS DEPOSITED FOR NATIVE DNA POLYMERASES

Because DNA polymerases are used in a variety of applications, patents have been deposited for a large number of DNA polymerases from several organisms from all three domains of life (*Archae*, *Eubacteria*, *Eukaryota*) and from viruses. In general, these patents provide information and/or claims on one or more of the following subjects: the DNA sequence of the DNA polymerase, the procedure for cloning the DNA polymerase into an expression vector, expression of the DNA polymerase, the purification protocol, biochemical characterization of the DNA polymerase, proposed applications, and proof-in-principle examples. A list of DNA polymerase patents is given below, which is provided to illustrate the variety of DNA polymerases that have been used in various applications or have the potential

to be used. The prototype DNA polymerase is the full length [68] and the large or Klenow fragment of *E. coli* DNA pol I [69].

Viral DNA Polymerases

Bacteriophage DNA polymerases.

Cloning and expression of the T4 DNA polymerase [70,71].

Cloning and expression of the T5 DNA polymerase [72].

T7 DNA polymerase [73,74].

φ29 DNA polymerase [29,75,76].

Archaeal virus DNA polymerase.

Thermostable DNA polymerase protein of the archaeal ampullavirus ABV (*Acidianus* bottle-shaped virus) [77,78].

DNA Polymerases from Mesophilic Bacteria

Plasmids containing the gene for DNA pol I from *Streptococcus pneumoniae* [79].

DNA polymerase type-III holoenzymes from a variety of bacteria including *Streptococcus*, *Staphylococcus*, *Yersinia*, and *Pseudomonas* [80-82].

DNA Polymerases from Thermophilic Bacteria

Purified thermostable DNA polymerase from *Thermus aquaticus* [38-41,83,84].

Thermostable DNA polymerase from *Thermus thermophilus* [85].

DNA encoding a thermostable DNA polymerase I from *Thermotoga maritima* [86,87].

Thermophilic DNA polymerases from *Thermotoga neapolitana* [88].

Thermostable DNA polymerase from *Thermoanaerobacter thermohydrosulfuricus* [89].

Thermostable DNA polymerase from *Carboxydotherrnus hydrogenoformans* [90].

Novel polymerases from *Ammonifex degensii*, *Pyrolobus furmarius*, *Archaeoglobus lithotrophicus*, *Metallo-sphaera prunae*, and *Aquifex* VF-5 [91].

Purified DNA polymerase from *Bacillus stearothermophilus* [92,93].

DNA pol III-type DNA polymerase and polymerase subunits from *Bacillus stearothermophilus* [94,95].

Thermus thermophilus polymerase III holoenzyme [96,97].

Thermophilic DNA polymerases from *Thermoactinomyces vulgaris* [98].

DNA pol III-type DNA polymerase and polymerase subunits from *Aquifex aeolicus* [99,100].

DNA polymerase III-type enzyme from a thermophilic bacterium [101].

Archaeal DNA Polymerases

Purified thermostable DNA polymerase from *Thermococcus litoralis* [102-107].

Thermostable DNA polymerase from *Thermococcus* 9°N-7 [108,109].

Purified DNA polymerase from *Thermococcus barossii* [110].

Thermococcus DNA polymerase JDF-3 [111]

Thermostable DNA polymerase from *Desulfurococcus* strain Tok12-S1 [112].

Thermostable DNA polymerases from *Pyrodictium* [113].

Thermostable DNA polymerase from *Pyrococcus furiosus* [114,115].

Thermostable DNA polymerase from a *Pyrococcus kodakaraensis* KOD 1 strain [116].

Eukaryotic DNA Polymerases

Human DNA polymerase alpha [117].

Mammalian DNA polymerase kappa [118].

Human Rev1 (dCMP transferase) [119].

Plant DNA polymerase eta [120].

PATENTS DEPOSITED FOR MODIFIED DNA POLYMERASES

Modified Family A DNA Polymerases and their Applications: *E. coli* DNA Pol I, Phage T7 DNA Polymerase, Phage T5 DNA Polymerase, and Taq and Taq-like DNA Polymerases

Natural DNA polymerases can be used for various DNA manipulations, but while DNA polymerases are exquisitely adapted for their cellular roles in DNA replication, repair and recombination, some DNA polymerase functions interfere with optimal performance in *in vitro* applications. The first modified DNA polymerase used was the large or Klenow fragment of *E. coli* DNA pol I. The Klenow fragment has 5'→3' polymerase and 3'→5' exonuclease activities, but lacks the N-terminal domain that is responsible for the 5'→3' exonuclease activity of the full-length enzyme [31,32]. As discussed above, the Klenow fragment rather than the full-length *E. coli* DNA pol I was used in the first DNA sequencing [30], PCR [37], and oligonucleotide-directed mutagenesis procedures [48] because the 5'→3' exonuclease activity can degrade the primers and product DNA.

The 3'→5' exonuclease activity of the Klenow fragment and other DNA polymerases can also degrade primers and product DNA; the latter is a significant problem for DNA sequencing reactions since removal of the chain-terminating nucleotides prevents the accumulation of chain-terminated products. The 3'→5' exonuclease activity can be inactivated by oligonucleotide-directed mutagenesis of the cloned gene. Mutations that encode one or more alanine (A) substitutions for the active site aspartate (D) residues in the exonuclease active center of the Klenow fragment largely abolish exonuclease activity since the Mg²⁺ ions required for

catalysis are bound by the aspartate residues [121]. This observation for the Klenow fragment was extended to a family B DNA polymerase, the T4 DNA polymerase, when mutational analysis revealed that the phage T4 DNA polymerase appeared to share the same functional organization as the Klenow fragment [122,123]. Exonucleolytic proofreading activity of the phage T4 DNA polymerase resides in a domain N-terminal to the polymerizing domain as observed for the Klenow fragment. Further analysis of the protein sequences of several proofreading DNA polymerases led to the discovery of three conserved exonuclease (Exo) motifs in the N-terminal region of family A and B DNA polymerases [4,124]. Thus, a general method to inactivate the 3'→5' exonuclease activity is to engineer mutations in the cloned gene to encode alanine substitutions for one or more essential residues in the exonuclease active center.

Inactivation of exonucleolytic proofreading activity, however, still leaves two problems for the use of the Klenow fragment for DNA sequencing: (1) *E. coli* DNA pol I and the Klenow fragment are not processive DNA polymerases as they incorporate only a few nucleotides (~ 15 on average) for each binding event [27,125], which results in incomplete strand extension unless reaction times are long and dNTP concentrations are high, and (2) these DNA polymerases have poor ability to incorporate chain-terminating nucleotides [126] for the Sanger method of DNA sequencing [30], which necessitates the use of high concentrations of ddNTPs in the reactions.

The solution to these problems was to develop another family A DNA polymerase, the bacteriophage T7 DNA polymerase, into a sequencing polymerase. The T7 DNA polymerase is naturally processive due to its association with the host-encoded thioredoxin [26]. In addition, this DNA polymerase discriminates only slightly in the incorporation of ddNMPs [127-129]. This left only one remaining problem in the way of converting the T7 DNA polymerase into a sequencing polymerase - the 3'→5' exonuclease activity. A largely exonuclease deficient T7 DNA polymerase was first produced by chemical treatment with ferrous ammonium sulfate in a saturated oxygen environment, which preferentially inactivated exonuclease, but not polymerase activity [130,131]. A completely exonuclease-deficient T7 DNA polymerase was engineered by *in vitro* mutagenesis, in this case 84 bps encoding N-terminal residues lysine (K) 118 to arginine (R)145 were deleted from the cloned T7 DNA polymerase gene [132,133]. Several patents have been deposited for modified or variant T7 DNA polymerases [73,133-136], mostly for use in the Sanger chain-terminating DNA sequencing method. The patents provide detailed information about DNA sequencing reactions with the exonuclease-deficient form of the T7 DNA polymerase, known as Sequenase, and documentation of the use of Sequenase for manual DNA sequencing.

The bacteriophage T5 DNA polymerase is also a processive enzyme, but processivity is independent of thioredoxin [27,28]. The 3'→5' exonuclease activity was inactivated by oligonucleotide-directed mutagenesis of the cloned T5 DNA polymerase gene following the general method - to engineer mutations that encode alanine

substitutions for essential residues in the exonuclease active center, for example D138A [137-139].

While excellent DNA sequence information can be obtained with Sequenase, inverted repeat DNA sequences in the template strand that form stable secondary structures (hairpins) under typical 30 to 37 °C sequencing reaction conditions cause premature termination bands across all four sequencing lanes. This problem was countered by developing a thermostable family A DNA polymerase, the Taq DNA polymerase, into a sequencing enzyme so that reactions could be performed above temperatures that favor formation of DNA hairpins [140,141].

The Taq DNA polymerase is naturally devoid of 3'→5' exonuclease activity [41,42], but there is a 5'→3' exonuclease activity. Deletion of 289 amino acid residues from the N-terminus, which contains the 5'→3' exonuclease activity, produces what is called the Stoffel fragment [83,84]. The Stoffel fragment is a useful enzyme for both DNA sequencing and PCR applications; however, this polymerase is like *E. coli* DNA pol I and its Klenow fragment and discriminates against the use of ddNTPs. Tabor and Richardson solved this problem [129]. They discovered that a single aromatic amino acid residue in the O helix of *E. coli* DNA pol I, phenylalanine (F) 762, and analogous residues in the phage T7 (tyrosine (Y) 526) and Taq (F667) DNA polymerases are critical for enabling these DNA polymerases to distinguish between dNTPs and ddNTPs. While the wild type T7 DNA polymerase with Y526 readily incorporates ddNMPs, the mutant Y526F-T7 DNA polymerase does not. The wild type Taq and *E. coli* DNA polymerases have a F residue instead of a Y residue at the critical position and these DNA polymerases have poor ability to incorporate ddNMPs; however, substituting Y in place of F converts both DNA polymerases into enzymes that can now efficiently use ddNTPs [129,142]. One embodiment of these modifications is Thermo Sequenase [143], which is a recombinant Taq DNA polymerase containing the F667Y substitution and the N-terminal deletion, which removes 5'→3' exonuclease activity. Another embodiment is AmpliTaq DNA pol. FS (Applied BioSystems); this enzyme also has the F667Y substitution and another amino acid substitution, glycine (G) 46 to aspartate (G46D), which removes almost all 5'→3' exonuclease activity. Another refinement of the Taq DNA polymerase is to replace arginine (R) 660 with another amino acid residue, for example aspartate (D) in order to reduce biased utilization of ddGTP [144]. Other advantageous modifications include H, Y, or K substitutions for R772 in the O-helix region of the *Thermotoga neapolitana* DNA pol I, which nearly abolishes template-independent addition of dAMP to the 3'-ends of PCR products [145]. These substitutions also increase replication fidelity from 5- to 50-fold [145].

Thermostable DNA polymerases also provide the means to couple asymmetric PCR and direct sequencing in a single tube, which facilitates automation of the Sanger DNA sequencing method [140,141,146,147]. Cycle sequencing was further advanced with the use of fluorescently labeled dideoxy chain-terminating nucleotides. Additional amino acid substitutions to G46D and F667Y are reported to enhance the ability of the Taq DNA polymerase to incor-

porate labeled nucleotides, particularly fluorescently labeled nucleotides [148]. These amino acid substitutions include, for example, arginine (R) 660 to aspartate (R660D), which also improves a more uniform utilization of ddNTPs [144], R595D and glutamate 681 to glycine (E681G). Several additional patents have been deposited for modified Taq and Taq-like DNA polymerases [44,85,149-156]. These DNA polymerases are proposed to be useful for DNA sequencing and/or PCR.

Another important application of PCR is in RNA detection and in cloning DNA copies (cDNA) from messenger RNA (mRNA). Small amounts of RNA can be amplified by first converting RNA to cDNA by using a reverse transcriptase (RT), a polymerase which synthesizes a complementary DNA from an RNA template. The cDNA can then be amplified by standard PCR; this method is called RT-PCR. Unfortunately, secondary structures (hairpins) in RNA can impede synthesis of the cDNA. The hairpins would be destabilized if cDNA synthesis could be done at the high temperatures of standard PCR reactions; however, RTs isolated from retroviruses cannot withstand these conditions. High temperature RT-PCR was achieved, however, by using the recombinant DNA polymerase derived from *Thermus thermophilus* (*Tth* pol), which possesses efficient reverse transcriptase activity in the presence of $MnCl_2$ [157,158]. Interestingly, amino acid substitutions for E683 (but not A, P, Q or D) in the *Tth* pol and for the corresponding residue E681 in the Taq DNA polymerase enable efficient RT-PCR in a Mg^{2+} containing buffer [158,159]. Alternatively, a DNA polymerase from the thermophilic eubacterium *Carboxydothermus hydrogenoformans* is claimed to have Mg^{2+} -dependent RT activity and, in addition, to have 3'→5' exonuclease to enable high fidelity RT-PCR [90]. The patents claim that these thermostable DNA polymerases and the described conditions allow combined reverse-transcriptase/amplification (RT-PCR) in a single tube (one pot) reaction.

Modified Family B DNA Polymerases and their Applications: Phage ϕ 29 DNA Polymerase, Phage T4 DNA Polymerase, and Archaeal DNA Polymerases

Focus on DNA Sequencing

Family B DNA polymerases can also be used for DNA sequencing and DNA amplification as well as several other applications; however, until recently thermostable family A polymerases have been the preferred enzymes for DNA sequencing. In order to convert a family B DNA polymerase into a sequencing polymerase, the intrinsic 3'→5' exonuclease of family B DNA polymerases must be inactivated. Proofreading deficient DNA polymerases are typically produced by the general method of engineering mutations in the cloned polymerase gene that encode alanine substitutions for essential residues in the exonuclease active center. For example, for the ϕ 29 DNA polymerase alanine (A) substitutions for D or E residues in the Exo I (D12A, E14A) or Exo II (D66A) motifs or in both motifs (D12A/D66A, E14A/D66A) reduce 3'→5' exonuclease activity to < 10% of the activity present in the wild type enzyme [160,161]. The proofreading-deficient ϕ 29 DNA polymerases are proposed to be useful for DNA sequencing by the Sanger method.

The first family B DNA polymerase developed as a commercial sequencing enzyme was the bacteriophage T4 DNA polymerase. A manual sequencing kit called the FIDELITY™ DNA Sequencing System was marketed by Oncor (Gaithersburg, MD, USA; Appligene, France). This kit had several components that enabled the sequencing of difficult DNA sequences that are rich in AT or GC base pairs and/or have repeated sequences. These components will be described because they teach the fundamentals of a DNA sequencing system.

The starting point was a proofreading-deficient phage T4 DNA polymerase. Exonuclease deficiency can be produced by the following amino acid substitutions alone or in combination: D112A/E114A in the Exo I motif, D219A in the Exo II motif, and D324A in the Exo III motif [162-165]. While the phage T4 DNA polymerase is not intrinsically processive [27], association of the T4 DNA polymerase with the gp45 sliding clamp produces a highly processive replicase [166]. The T4 DNA polymerase holoenzyme (DNA pol plus clamp) was reconstituted in the kit by providing the purified gp45 sliding clamp and the clamp loader proteins, gps 44/62. The kit also included the phage T4 encoded single-strand DNA binding protein, gp32, which hinders formation of DNA hairpin structures on single-stranded DNA [167].

Like many other DNA polymerases, the phage T4 DNA polymerase cannot efficiently incorporate chain-terminating 2',3'-dideoxynucleotides. This problem was solved by using another chain-terminating nucleotide that is incorporated more readily by the T4 DNA polymerase: 3'-amino-2',3'-dideoxynucleotides [162]. The amino group at the 3' position on the sugar Fig (1) prevents further nucleotide incorporation, but the amino group enables the T4 DNA polymerase to utilize this chain terminator in preference over 2',3'-dideoxynucleotides. Efficient incorporation of 3'-amino-2',3'-dideoxynucleotides was demonstrated for exonuclease deficient forms of the T4 DNA polymerase and another family B DNA polymerase, *E. coli* DNA pol II. Efficient use of other chain terminating nucleotides with arabinose in place of the 2'-deoxyribose, araCTP and araUTP, was also claimed. The claims regarding the use of 3'-amino-2',3'-dideoxy- and arabinose nucleotides were extended to all proofreading deficient family B DNA polymerases.

The T4 DNA polymerase in the FIDELITY™ DNA Sequencing System was additionally modified with a methionine substitution for leucine 412 (L412M) in the polymerase active center. L412 resides within the highly conserved Motif A in the polymerase active center Fig (4), which has the following amino acid sequence in family B DNA polymerases: DxxSLYPSI, where x is one of several amino acid residues [4,70,168-170]. The phage T4 L412M-DNA polymerase has increased intrinsic processivity compared to the wild type T4 DNA polymerase [171-175], which enables the modified DNA polymerase to efficiently copy difficult DNA sequences including A+T- and G+C-rich DNA sequences and DNA sequences with repeats that cannot be sequenced by other enzymes, for example inverted repeat sequences that produce particularly strong secondary structures (hairpins) [173].

Another useful property is that the L412M-DNA polymerase can incorporate nucleotides with modifications, including bulky fluorophores [174]. This means that the multiply mutant T4 DNA polymerase with amino acid substitutions to remove 3'→5' exonuclease activity, for example D112A/E114A, coupled with the L412M substitution produce a variant T4 DNA polymerase that is useful for sequencing DNAs with difficult sequences, for labeling DNA with modified nucleotides, and for sequencing DNA with fluorophore-labeled nucleotides, which is discussed below. Because Motif A is highly conserved, the claims for the L412M substitution were extended to similar modifications in other family B DNA polymerases [162,163,174] and to polymerases in general since the aspartate (D) residue in Motif A is conserved in the polymerase active center of all polymerases Fig (4) [169,170]. While protein sequence comparisons and DNA polymerase structures led researchers to Motif A as a likely target to produce variant DNA polymerases that can incorporate non-conventional nucleotides [169,170], *in vivo* genetic selection strategies led to the discovery of the T4 L412M-DNA polymerase [171,175]. The use of genetic selection methods to identify variant DNA polymerases is discussed in the section on EVOLVING DNA POLYMERASES. Additional evidence for the role of Motif A in nucleotide interactions is provided by experiments with eukaryotic DNA pols α and δ [176-178] and from observations for a variety of DNA polymerases that are discussed below.

The drive to develop novel DNA sequencing techniques that are rapid and cost effective is also urging development of novel DNA polymerases, particularly modified family B DNA polymerases, which are suited for the new methods. Richard Keller proposed one of the first new sequencing methods [179]. In this method, a DNA polymerase synthesizes the complementary strand using fluorophore-labeled nucleotides, with the base of each of the four dNTPs tagged with a unique fluorescent molecule. A single fluorophore-labeled DNA is then suspended in a flow stream and degraded sequentially, nucleotide-by-nucleotide by an exonuclease; the order that the excised fluorescent nucleotides reach the detector gives the DNA sequence.

In theory, this method could produce very long sequence reads at a rapid rate since polyacrylamide gel electrophoresis is not needed. However, there are technical difficulties, one of which is the synthesis of complementary fluorophore-labeled DNA because DNA polymerases have difficulty in synthesizing DNA with unnatural, modified nucleotides. First, the bulky fluorophore can interfere with the ability of the DNA polymerase to bind the nucleotide in the polymerase active center and then can hinder subsequent steps in the nucleotide incorporation pathway [180]. Second, further primer extension will be hindered unless the DNA polymerase can form stable complexes with DNA having a modified primer terminus. If binary complexes can be formed, then the next fluorescent nucleotide must be bound and incorporated. Thus, primer extension requires that the

Polymerase	Motif A	Phenotype
Family B		
T4	408 D I t S L Y P S I I	
	V	strong antimutator, low dGTP^S
	M V	antimutator, low dGTP^R
	M	weak mutator, increased processivity, incorporate non-conventional nucleotides, PAA^S
	I	weak antimutator, PAA^R
Ec pol II	419 D y k S L Y P S I I	
ϕ29	249 D v n S L Y P a q	
Sc pol δ	608 D f n S L Y P S I	
Sc pol α	863 D f n S L Y P S I I	
h pol α	860 D f n S L Y P S I I	
Pfu	405 D f r a L Y P S I I	
9°N-7	404 D f r S L Y P S I I	
VENT	407 D f r S L Y P S I I	
Sso	514 D f a S L Y P S I I	
Family A		
Ec pol I	705 D y s Q I E L r	
Taq pol I	610 D y s Q I E L r	
RT		
HIV	265 D v g d A Y	

Fig. (4). The polymerase conserved Motif A sequence, adapted from Delarue *et al.* [169] Joyce [170].

DNA polymerase accommodate bulky modifications at two sites, modifications in the primer-terminal region and in the incoming nucleotide. The multiply mutant phage T4 D112A/E114A/L412M-DNA polymerase can incorporate adjacent fluorophore-labeled nucleotides [174,181,182].

Another new DNA sequencing method is sequence-by-synthesis (SBS). In one variation of this method, 1000s of "spots" of DNA templates (either single molecule or amplified templates from a single molecule) are arrayed on the surface of a slide. The DNA spots are then interrogated with a single, fluorescently-labeled nucleotide, for example *dGTP. In one embodiment, the * indicates a bulky fluorophore that is attached to the base in a non-base pairing position. The *G nucleotide is incorporated by the DNA polymerase only in spots where the template is the complimentary C. The slide is then rinsed to remove unincorporated fluorescent nucleotides and the array is imaged to record each spot where fluorescent *G nucleotides are incorporated. Then the fluorophore is attacked chemically to remove fluorescence and the process is repeated with another fluorescent nucleotide [183,184]. This method is a high-throughput system because 1000s of spots can be sequenced in parallel. One drawback with this method is that the base may remain partially modified after the chemical treatment, which then could impede further primer extension; however, the phage T4 D112A/E114A/L412M-DNA polymerase has been used with some success [183,184].

In another embodiment, the * indicates the presence of a bulky fluorophore on the base and a chain-terminating modification and both modifications are *reversible* [185-187]. After each cycle of interrogation and imaging with one fluorescent, chain-terminating nucleotide, the chain-terminator and fluorophore are removed, leaving only a small remnant of the modification on the base. Removal of most of the fluorophore facilitates primer extension compared to synthesis of a complimentary DNA strand in which all of the bases are modified as required for the R. Keller method [179]. Two thermostable DNA polymerases, Thermo Sequenase (a family A DNA pol) and a variant form of the archaeal 9°N-7 DNA polymerase (a family B DNA pol) have been reported for this SBS method [186,187]. Thermo Sequenase, a variant Taq DNA polymerase has already been described [143]; the variant 9°N-7 DNA polymerase is described below.

As discussed above for the general development of DNA sequencing enzymes, 3'→5' exonuclease activity must be inactivated and the polymerase must be able to efficiently incorporate chain-terminating nucleotides. A proofreading-deficient form of the 9°N-7 DNA polymerase was produced following the general protocol, which is to engineer mutations into the cloned gene that encode alanine substitutions for essential residues in the exonuclease active center [108,109]. Mutational analyses of several similar family B archaeal DNA polymerases (Vent, Deep Vent, 9°N-7 and Pfu) were carried out to discover the determinants of nucleotide sugar discrimination [188,189]. The A485L substitution in the 9°N-7 DNA polymerase and the analogous A488L substitution in the Vent DNA polymerase significantly increase utilization of dideoxy terminating

nucleotides. These residues are near the conserved Motif B sequence that is found in all polymerases [169]. The proofreading deficient (exo⁻) A485L-9°N-7 DNA polymerase is marketed by New England BioLabs as TerminatorTM DNA polymerase. An additional amino acid substitution, Y409A, which resides in the conserved Motif A sequence Fig (4), converts the TerminatorTM DNA polymerase to the TerminatorTM II DNA polymerase, and this was the mutant form of the 9°N-7 DNA polymerase that was used for the SBS method described by Ju *et al.* [187]. The Y409A substitution decreases discrimination against ribose nucleotides and other sugars as observed for substitutions for the invariant glutamate residue (E) which occupies a similar position in Motif A of family A DNA polymerases (Fig. 4) [170]. The TerminatorTM DNA polymerase can also efficiently synthesize long oligonucleotides on a DNA template using threonucleoside triphosphates, in which (3'-2') α-L-threose replaces the standard deoxyribose sugar [190].

Mutational analysis of another archaeal family B DNA polymerase, the *Thermococcus* JDF-3 DNA polymerase, was also carried out to discover variants with reduced 3'→5' exonuclease and/or reduced discrimination against non-conventional nucleotides including chain-terminating nucleotides and fluorescently labeled nucleotides [191]. The variant *Thermococcus* JDF-3 DNA polymerases constructed include mutant DNA polymerases with L408H (histidine) and L408F (phenylalanine) substitutions, where L408 is analogous to residue L412 in Motif A of the phage T4 DNA polymerase Fig (4). Substitutions for the nearby proline residue, P410 were also proposed. Mutant DNA polymerases were also constructed with A485T (threonine), C (cysteine), S (serine), L, I (isoleucine), F or V (valine) substitutions. Note that A485 in the *Thermococcus* JDF-3 DNA polymerase is analogous to A488 in the 9°N-7 DNA polymerase. Several other mutant DNA polymerases were constructed including variants with the following amino acid substitutions: Y497C, I630V, E645K, E578K, Y409V etc. Variant forms of the *Thermococcus* JDF-3 DNA polymerase containing combinations of these amino acid substitutions and other substitutions listed in the patent [191] are claimed to be useful for DNA sequencing and/or PCR and/or DNA labeling.

Other modified archaeal DNA polymerases are also claimed to be useful for DNA sequencing and/or PCR and/or DNA labeling. For example, several patents have been filed for exonuclease-deficient variants of *Pyrococcus* DNA polymerases [192,193] and a variety of exonuclease-deficient archaeal DNA polymerases were screened for their ability to incorporate modified nucleotides in reactions in which one or more of the standard dNTPs was replaced with a non-conventional dNTP [194]. Mutational analysis of the Pfu DNA polymerase was carried out to discover variant enzymes with improved incorporation of dideoxynucleotides. In one study, the target region was the Motif B region, amino acid residues 488 to 496 [195]; and as found in similar studies of Motif B that are reviewed above, variant DNA polymerases with amino acid substitutions in this region, the A491Y substitution in this study, have improved ability to incorporate dideoxynucleotides. In another study of the Pfu DNA polymerase, utilization of dideoxynucleotides

was improved for variant DNA polymerases with one of the four following amino acid substitutions: Q472H, A486Y, L490Y and Y497A [196].

Another SBS method employs a different form of fluorescent nucleotide in which the fluorophore is attached to the γ -phosphate rather than to the base, and a new strategy to perform single molecule sequencing reactions using zero-mode waveguide technology to reduce the observation volume to the zeptoliter range (10^{-21}) [197,198]. Single DNA polymerase-DNA complexes are interrogated in the small reaction volume with a γ -phosphate-labeled nucleotide; stable ternary complexes are imaged in real-time before the labeled-PPi diffuses away. Thus, while the DNA polymerase must be able to form a stable ternary complex specifically with the correct γ -phosphate-labeled nucleotide, none of the bulky fluorophore remains after nucleotide incorporation and no chemical treatment is needed to remove the fluorescent tag.

Several patents have been deposited for DNA sequencing with the novel nucleotides with the fluorophore attached to the γ -phosphate of the nucleotide and the zero-mode waveguide technology, for example [199], and for polymerases that can be used with this method [200,201]. In one embodiment, the DNA polymerase is a modified form of the highly processive, family B phage ϕ 29 DNA polymerase [29]. As for sequencing enzymes in general, the 3'→5' exonuclease is largely abolished by introducing mutations into the cloned gene that produce amino acid substitutions for essential residues in the exonuclease active site, for example D12A in the Exo I motif or N62D in the Exo II motif. Mutations encoding these substitutions are then combined with mutations that encode amino acid substitutions to enhance binding of non-conventional nucleotides; residues were targeted that may interfere with binding the γ -phosphate labeled nucleotides. These amino acid substitutions include E375H, which is near Motif B, several other substitutions for E375, the L384R substitution which is in Motif B, the E486A substitution, deletion of residues 505-525, and combinations of these substitutions [200,201]. The variant ϕ 29 DNA polymerases are claimed to be useful for DNA sequencing as well as for PCR and RT-PCR applications.

Focus on PCR

While thermostable family A DNA polymerases, such as the modified Taq DNA polymerase in which the 5'→3' exonuclease is removed, are useful PCR enzymes the absence of 3'→5' exonucleolytic proofreading activity prevents mismatch correction and as a result limits the length and accuracy of PCR products. When an incorrect nucleotide is incorporated, there are two possible outcomes: mutation or termination. If the mismatch can be extended, subsequent PCR cycles will amplify the mutation. If mismatch extension occurs early in the cycling process, a large clone of mutant PCR product will be produced [56]. If the mismatch cannot be extended, then primer extension is terminated and since the "terminated" DNA cannot prime subsequent cycles, there is a reduction in the yield of PCR product. Both problems are exacerbated as the length of the PCR fragment increases, but both problems can be solved by adding a small amount of a thermostable proofreading proficient DNA polymerase to

PCR reactions with a larger amount of a proofreading-deficient thermostable DNA polymerase [43-46]. When an incorrect nucleotide is incorporated, the proofreading-deficient DNA polymerase will likely dissociate from the mismatched DNA and this provides an opportunity for the proofreading-proficient DNA polymerase to bind and excise the incorrect nucleotide at the primer-terminus. Thus, the combined action of proofreading-deficient and -proficient thermostable DNA polymerases enables the synthesis of long and accurate PCR products.

A further possible improvement would be to use an even more proofreading proficient archaeal family B DNA polymerase. Higher fidelity archaeal DNA polymerases (*Pyrococcus kodakaraensis* KOD1 DNA polymerase) have been engineered in which a variety of amino acid substitutions were made near the Exo I motif [202]. The H147R and H147K substitutions increased exonucleolytic proofreading 3- and 4-fold, respectively, compared to the wild type enzyme. Replication fidelity in PCR reactions was also increased 3- to 4-fold. These variant DNA polymerases and similar archaeal DNA polymerases are claimed to be useful for high fidelity PCR reactions [202].

Note, however, that accurate amplification of long DNA templates is possible with a single family B DNA polymerase, the phage ϕ 29 DNA polymerase [60-62]. The ϕ 29 DNA polymerase is a proofreading DNA polymerase [203]. What prevents archaeal family B DNA polymerases from being stand-alone PCR enzymes?

One problem was discussed in the INTRODUCTION under DNA polymerase activities – the ability of archaeal DNA polymerases to read-ahead and detect pro-mutagenic uracils in the template strand [20-23,204]. During PCR reactions at high temperatures that may range from >50 to >95 °C, oxidative deamination of dCTP to dUTP occurs at a significant level. When dUMP is incorporated opposite template adenine, uracils are introduced into the PCR products. Uracil may also be produced from deamination of cytosine in DNA, particularly in denatured, single-stranded DNA. Since archaeal family B DNA polymerases form tight complexes with uracil-containing DNAs, uracil in DNA deters DNA amplification. Thus, uracil detection must be prevented to enable archaeal DNA polymerases to be PCR enzymes.

The detection of uracil and other damaged bases has been localized to a "pocket" in the amino terminal domain of archaeal DNA polymerases [21-23]. Structure-guided mutagenesis of pocket residues identified several amino acid substitutions that hinder recognition of uracil in DNA. For the Pfu DNA polymerase, amino acid substitutions in the pocket region including residues 1-40 and 78-130, and specifically one or more of the following amino acid substitutions Y7A, Y37A, V93Q/R, and I114R/Q enable the variant DNA polymerases to replicate DNA with uracil [205]. These variant Pfu DNA polymerases are claimed to be useful for PCR reactions. These claims are extended to all archaeal DNA polymerases in general because of the sequence and functional similarities of these enzymes.

RT-PCR is a useful technique, as described in the section about modified family A DNA polymerases, which can be

used to detect and quantitate an RNA species and to clone cDNAs synthesized from mRNAs. However, family A DNA polymerases with RT activity lack exonucleolytic proofreading activity and cannot replicate DNA with sufficient accuracy to produce error-free PCR products. Thus, there is a need to develop thermostable DNA polymerases with proofreading and RT activities and, additionally, it would be useful for such enzymes to be able to incorporate non-conventional nucleotides. These goals have been achieved with variant forms of the archaeal Pfu and related DNA polymerases [206]. Specifically, amino acid substitutions in Motif A in the LYP region Fig (4), such as L to an aromatic amino acid F, Y, H or W create variant DNA polymerases with RT activity and the ability to incorporate non-conventional nucleotides.

Although archaeal family B DNA polymerases are known in general as high fidelity DNA polymerases, variant forms of the Pfu DNA polymerase have been developed as proposed agents for mutagenic PCR [207]. Inactivation of exonucleolytic proofreading reduces replication fidelity, but nucleotide incorporation is still accurate. However, mutations that encode glycine or alanine amino acid substitutions for T471, Q472, or D473 in the Pfu DNA polymerase produce error prone DNA polymerases that misincorporate a greater number of incorrect nucleotides compared to the wild type enzyme. Combination of one of the amino acid substitutions that increases nucleotide misincorporation with the D215A substitution in the Exo II motif, which inactivates exonucleolytic proofreading, generates a mutator DNA polymerase that is claimed to be useful for mutagenic PCR. Claims in this patent [207] are extended to archaeal family B DNA polymerases in general.

Modified Family C DNA Polymerases: Eubacterial Replicases

Family C DNA polymerases are part of the DNA pol III holoenzymes of mesophilic gram positive and negative bacteria and the pol III replicases from thermophiles. As discussed in the INTRODUCTION, family C DNA polymerases function as multisubunit holoenzymes [2]. One patent is discussed here to illustrate the types of modifications that have been done or are proposed [208]. One proposal is to reduce the complexity of the pol III holoenzyme to just two subunits. The simplest form of the DNA pol III replicase envisioned consists of just two components - the α and β subunits. The α subunit has polymerase activity and the β subunits form the sliding clamp, which provides processivity. Clamp loading components are not included because the α subunit is modified to have increased affinity for the β subunit sliding clamp, which is proposed to facilitate formation of the processive α - β complex. These replicases are proposed to be useful for PCR applications. Binding motifs within α subunits for the β subunit of thermolabile and thermostable DNA pol III replicases are provided as well as the sequences of mutated α subunits with greater affinity for β subunits [208]. The patent also provides methods to modulate 3'→5' exonuclease activity [208].

EVOLVING DNA POLYMERASES

Variant DNA polymerases with reduced exonuclease activity or new abilities to incorporate chain-terminating and other non-conventional nucleotides have been engineered starting with the cloned wild type gene and knowledge about active sites, structure, and protein sequence alignments. However, amino acid substitutions away from active sites and in unpredicted locations can have large impacts on function. Furthermore, multiple amino acid substitutions could individually have little effect on function, but cumulatively could produce a variant DNA polymerase with the desired properties. In order to discover these less obvious amino acid substitutions, genetic selection and screening strategies must be used.

One of the first genetic strategies developed was an *in vivo* scheme to select for mutant T4 phage, which express *mutator* DNA polymerases and replicate DNA with low fidelity [123,209]. These experiments were done to locate amino acid residues needed for exonucleolytic proofreading, which were not known at that time. The selection scheme was based on the rationale that spontaneous mutation rates would be higher in phage expressing proofreading-deficient DNA polymerases. The goal of all genetic selection strategies is to design conditions where only organisms with the desired phenotype survive. Thus, if the goal is to select for T4 phage that express mutator DNA polymerases, then only phage with a mutator phenotype should survive. This criterion was achieved by starting with T4 phage that carried an *rII* allele that reverts to *rII*⁺ only rarely. Since experimental conditions exist that prevent replication of *rII* deficient phage, millions of phage can be plated on a single Petri dish and only phage in which the *rII* allele has reverted to a functional state (*rII*⁺) will grow. But the wild type T4 DNA polymerase makes very few replication errors at this site, the *rII* revertant frequency is just 1×10^{-9} ; however, a mutator, less accurate DNA polymerase will increase the mutant frequency. The conditions of the selection scheme allowed just a single round of phage replication under permissive conditions, which highly enriches for mutant T4 phage expressing mutator DNA polymerases [123,209]. Most of the mutator DNA polymerases were defective in exonucleolytic proofreading due to amino acid substitutions in the exonuclease active site, for example D112N in the Exo I motif, but also at unpredicted locations outside of the exonuclease active site, for example D131G [210] and G255S [211,212]. A similar selection strategy was used for the yeast (*S. cerevisiae*) DNA pol δ and resulted in the identification of mutator DNA polymerases with amino acid substitutions in the exonuclease active site, for example D321G, E323Q and D407N, but also amino acid substitutions away from the active center, for example L479S [213].

Another *in vivo* selection strategy was used in phage T4 to identify *mutator* (low fidelity) and *antimutator* (high fidelity) DNA polymerases [171,175,209,214,215]. Antimutator DNA polymerases are useful when high accuracy is needed, for example in PCR reactions, since an error made in an early cycle will be clonally expanded. Highly accurate DNA polymerases are also needed for single molecule

DNA sequencing. Mutator DNA polymerases, on the other hand, can be used as mutagenizing agents, but some “weakly” mutator DNA polymerases have additional properties that are useful for *in vitro* applications as explained below.

Amino acid substitutions that generate mutator T4 DNA polymerases were identified in genetic selection schemes that began with mutant T4 DNA polymerases that catalyzed excessive proofreading and/or could not catalyze sufficient DNA replication on a bacterial host that reduced the concentration of dGTP [171,209,214]. The conditional lethality produced by a low dGTP pool provided experimental conditions to select for second-site mutations that encode amino acid substitutions that allow phage to replicate under the non-permissive conditions of a low dGTP pool. Several amino acid substitutions were identified and in general these amino acid substitutions decreased the ability of the mutant DNA polymerases to proofread, hence the mutator phenotype, but many of the variant DNA polymerases have additional phenotypes. For example, one of the amino acid substitutions, L412M Fig (4), enables the T4 DNA polymerase to have increased intrinsic processivity [175], to form DNA polymerase complexes with higher stability than complexes formed with the wild type enzyme [172], and to incorporate non-conventional nucleotides [174]. The L412M-DNA polymerase also proofreads less, but since there is only a modest 10-fold increase in the spontaneous mutant frequency [175], this DNA polymerase still replicates DNA with sufficient accuracy for most applications.

Note that residue L412 resides in the conserved Motif A sequence Fig (4), which has proven to be an important target for DNA polymerase modifications that enable variant DNA polymerase from several polymerase families to incorporate non-conventional nucleotides [170,189,191]. Note, however, that not all amino acid substitutions in the Motif A sequence confer relaxed discrimination in nucleotide incorporation; some substitutions have the opposite effect. The I417V and L412I substitutions (Fig. 4) generate variant T4 DNA polymerases with antimutator phenotypes and the inability to replicate under conditions of a low dGTP pool [171, 175,214]. The ability of certain amino acid substitutions in Motif A to increase or decrease replication fidelity is exemplified by the finding that the L412M substitution rescues the inability of phage expressing the I417V-DNA polymerase to replicate DNA when the dGTP pool is low [175,214]. Thus, an amino acid substitution (I417V) that produces an antimutator DNA polymerase can be countered by a second amino acid substitution (L412M) that by itself produces a weakly mutator DNA polymerase.

The L412M-DNA polymerase displays yet another phenotype - sensitivity to the antiviral drug, phosphonoacetic acid (PAA) and to ddNTPs [171,175]. The PAA-sensitivity of the T4 L412M- and other mutant DNA polymerases [175] can be used to select for second site mutations that encode amino acid substitutions that suppress PAA-sensitivity; these suppressor mutations in turn encode amino acid substitutions that confer the antimutator phenotype [215]. Thus, the conditional lethality provided by either a low dGTP pool or the presence of PAA to sensitive T4 strains provides the

means to discover amino acid residues that are important for normal DNA polymerase function and to generate variant DNA polymerases with new properties that may be useful for *in vitro* applications [209,216].

In vivo genetic screens for altered DNA polymerase function can also be useful. For genetic screens in general, non-permissive conditions are established and then the ability of individual organisms to proliferate or not is tested. In practice, the first step is mutagenesis and the survivors are plated under various experimental conditions. For genetic screens with bacteria, thousands to tens of thousands of bacteria may be screened for a specific phenotype. We performed a genetic screen to discover variant *E. coli* DNA pol III replicases that have increased ability to incorporate chain-terminating, dideoxynucleotides [217]. Although, Tabor and Richardson discovered a single amino acid residue in family A DNA polymerases that is critical for distinguishing between deoxy- and dideoxynucleotides [129], the critical residue(s) is not known for family C DNA polymerases and until recently there was no structural information to guide mutational analysis of potential dNTP binding residues [218]. However, dideoxynucleotides are predicted to kill bacteria with variant forms of the Pol III replicase that increase the ability of the replicase to utilize chain-terminating nucleotides; these are conditions that are well-suited for a genetic screen.

This rationale was proven to be correct. The *dnaE* gene, which encodes the polymerase (α) subunit of the DNA pol III holoenzyme, was selectively mutagenized and bacteria were screened for sensitivity to dideoxyadenosine (ddA), which is converted to ddATP in cells [217]. The parental cells had reduced DNA pol I activity, which increases sensitivity to killing by chain-terminating nucleotides [219]. Two amino acid substitutions were identified in the α subunit that confer increased sensitivity to ddA: L329F and H417Y [217]. Since these residues are located near the polymerase active center and are conserved in bacterial Pol III α subunits from mesophiles and thermophiles, these amino acid substitutions provide a starting point for converting family C DNA polymerases into DNA sequencing enzymes. The genetic approach is particularly useful for studies of DNA pol III replicases because the structure of one of these replicases, *T. aquaticus* DNA pol III [218], is distinct from family A and B DNA polymerases and, thus, information learned from family A and B polymerases cannot necessarily be transferred directly to family C DNA polymerases.

Another powerful genetic screen in bacteria exploits the observation that exogenous DNA polymerases can functionally substitute for a temperature sensitive *E. coli* DNA pol I, *polA12*, at elevated temperatures; temperature sensitivity is enhanced by addition of the *recA718* allele [220,221]. Expression of the wild type Taq DNA polymerase from a plasmid complements the growth defect of the *polA12 recA718* strain [222-226], which thus provides a method to screen a mutagenized plasmid library for mutant DNA polymerases that can or cannot complement. The mutant DNA polymerase genes are then sequenced to identify mutations and then the expressed mutant DNA polymerases are further screened for other phenotypes such

as replication fidelity. Methods for all of these steps are described in the relevant patents and referenced papers [225-230].

In one example, variant Taq DNA polymerases were identified from a mutagenized pool of plasmids in which the region encoding Motif A was mutagenized randomly (Fig. 4). Non-mutable, nearly non-mutable and highly mutable amino acid residues within Motif A were identified [222-226]. Conservative amino acid substitutions were detected for the invariant E residue in family A DNA polymerases (Fig. 4), which was previously shown to play a pivotal role in nucleotide discrimination as discussed above [170]. Variant DNA polymerases with high-fidelity (antimutator) phenotypes were also discovered, as predicted from the phage T4 DNA polymerase studies [175,209]; multiple amino acid substitutions were encoded in some variant DNA polymerases. The high fidelity variants are proposed to be useful for PCR and also for PCR-type methods that are used to detect genomic mutations. For example, if a primer is provided that is perfectly complementary to a mutant DNA sequence but there are two mismatches at the 3'-end when annealed to the wild type sequence, primer extension will ideally be detected only if the genomic DNA contains the mutant sequence. In practice, however, the wild type Taq DNA polymerase may occasionally extend the mismatch and, thus, produce false positives for the presence of the mutant sequence. A high fidelity DNA polymerase is predicted to decrease mismatch extension and, thus, to decrease false positives [225,226].

A similar screen was carried out for another target region in the Taq DNA polymerase, residues in the O helix (R659-Y671) [227]. Note that the F667Y substitution enables the Taq DNA polymerase to utilize ddNTPs [129]. While several amino acid residues in the O helix were found to be essential for DNA polymerase function and were immutable or nearly immutable based on the genetic complementation assay, amino acids adjacent to these residues altered replication fidelity [227-230]. High and low fidelity variants were discovered and a method to use a high fidelity DNA polymerase to detect a genetic disease is provided, which is based on the premise that high fidelity DNA polymerases will only extend perfectly complementary primers (see discussion in the above paragraph). The low fidelity DNA polymerases are proposed to be useful for mutagenic PCR.

Phage display methods have also been used to select for catalytically active DNA polymerases with novel activities [231] and, specifically, for obtaining variants of the Taq DNA Pol I which are claimed to be useful for RT-PCR [232]. A number of amino acid substitutions were identified that are reported to give enhanced activity including H203R, F205L, T232S, E253G, Q257R, D274G, L275H or P, I276F, V309I, I322N, A331V, S335N, I361F, R374Q, A384T, T387A, Y419C or N, E465K, M479K or R, F472Y or S, A487T, K490E, P493T, M498T, G499E, M502K, L503P, V506I, R518G, A523V, A526V, P639S, E543K, and W550R. In another phage display experiment [233], variants of the Stoffel fragment of the Taq DNA pol I were selected that can function as thermostable RNA polymerases. The multiply mutant polymerases have a glycine substitution for the invariant E residue in Motif A of family A DNA

polymerases Fig (4) as well as the A596T substitution and some isolates also have amino acid substitutions for W604 and I614.

In vitro genetic selection schemes have several advantages over selections with organisms since a wide range of experimental conditions can be applied that may not be compatible with cell viability. A series of elegant polymerase directed-evolution schemes have been devised for the Taq DNA polymerase, which can be extended to DNA polymerases in general [234-236]. The essence of the method is called compartmentalized self-replication (CSR) where individual polymerase variants are expressed in separate compartments. The variant DNA polymerases are produced from a library of mutagenized cloned genes that were generated by mutagenic PCR [237]. The polymerase and the gene are released in the presence of amplification reagents (dNTPs, primers), which allows self-amplification to proceed, but only if the polymerase is active under the experimental conditions. Amplified genes, produced only by the active polymerases, are then recloned and subject to another cycle of CSR.

Different types of mutant DNA polymerases can be selected depending on the reaction conditions. If increased thermostability is desired, then the selective condition is the duration of the denaturation step at 99 °C; only polymerases tolerant to prolonged incubation at high temperature will survive and be able to self-amplify their genes. The following amino acid substitutions were identified in the Taq Pol I after three cycles of CSR: F73S, R205K, K219E, M236T, E434D, and A608V [237]. Since this selection procedure requires retention of catalytic activity, improved thermostability was not allowed to evolve at the cost of reduced catalytic activity. Similar strategies were used to isolate variants of the Taq DNA polymerase with increased ability to utilize non-conventional nucleotides and/or to have reverse transcriptase activity [238,239].

In yet another application of CSR, novel Taq DNA polymerases were selected that could replicate ancient DNA, which suffers from a variety of different types of DNA damage [240,241]. For many samples of ancient DNA the large extent of DNA damage prevents amplification; however, use of the CSR or molecular breeding approach identified Taq DNA pol variants with the ability to replicate DNA with abasic sites. These DNA polymerase variants, however, have a downside - is decreased replication fidelity. One possibility, suggested by Thomas *et al.* [241] is to carry out reactions with a mixture of a high fidelity and a novel Taq DNA polymerase with ability to replicate damaged DNA, which is analogous to the general approach of combining a lesion repair/tolerant polymerase with a second polymerase [242]. Another proposed method for amplification of damaged DNA is to compose an Enzyme Blend with proofreading and proofreading-deficient DNA polymerases and DNA repair enzymes, for example enzymes for the repair of abasic sites, for uracil in DNA, a photolyase to remove pyrimidine cyclobutane dimers, and a DNA ligase to repair nicks in the DNA [243,244]. Additional patents have been filed that use enzymes to repair DNA damage in order to improve the extent and fidelity of DNA replication [245,246].

FAMILY X AND Y DNA POLYMERASES

Family X and Y DNA polymerases are proposed to be good agents for mutagenic replication [247] and family Y DNA polymerases are also proposed to be useful for mutagenic PCR and for the replication of damaged DNA [248,249].

CHIMERIC DNA POLYMERASES

Increased Processivity and Applications to PCR and Oligonucleotide-Directed Mutagenesis

Another method to modify DNA polymerases for various DNA manipulation applications is to fuse a desirable feature of one protein to a DNA polymerase to form a chimeric DNA polymerase. For example, the highly basic, non-specific DNA-binding protein Sso7d from the hyperthermophilic Archaea *S. solfataricus* [250] was fused with the Taq or Pfu DNA polymerases to increase processivity of DNA replication [251]. In one Taq-Sso7d hybrid recombinant protein, the Sso7d protein replaced the first 289 amino acid residues of Taq. In a second hybrid, the Sso7d gene was joined to the 5'-end of the Taq DNA polymerase gene, which produced a chimeric Taq DNA polymerase with the Sso7d protein fused to the N-terminus of the full-length Taq DNA polymerase via a 4 amino acid peptide linker. Both chimeric Taq-Sso7d DNA polymerases are more processive than the Taq DNA polymerase. While the processivity of the Taq DNA polymerase is about 15-20 nucleotides, the processivity of the full-length Taq-Sso7d chimera is reported to be about 10-fold higher, 130-160 nucleotides [251].

Using standard recombinant DNA technology, a hybrid Pfu DNA polymerase-Sso7d chimera was constructed with the N-terminus of the Sso7d protein fused to the C-terminus of the Pfu DNA polymerase with an intervening 6 amino acid peptide linker [251]. The chimeric Pfu-Sso7d polymerase was ~10-fold more processive than the Pfu DNA polymerase. The enhanced processivity of the chimeric Pfu DNA polymerase increased the yield of long PCR products and in shorter extension times. The chimeric Pfu DNA polymerase is also useful for various applications of oligonucleotide-directed mutagenesis methods Fig (3), where a high fidelity, processive DNA polymerase is useful for incorporating mutagenic oligonucleotides into the complementary strand.

DNA polymerase processivity can also be increased by joining high-affinity, DNA-binding, helix-hairpin-helix motifs (HhH) to the Stoffel fragment of the Taq DNA polymerase and to the Pfu DNA polymerase [9,252]. The HhH motifs, identified in topoisomerase V, also increase polymerization rates and thermostability of the hybrid proteins at high salt concentrations. Because the salt tolerant chimeric DNA polymerases are more resistant to various inhibitors, these enzymes may be of use for PCR and for sequencing DNA in crude samples. The fusion of HhH motifs to proteins is proposed to be a general method to increase DNA binding at extreme experimental conditions [252].

Another example of a DNA polymerase chimera is to construct a recombinant hybrid composed of the exonuc-

leolytic proofreading domain of a family B, archaeal DNA polymerase to a proofreading deficient DNA polymerase such as the Taq DNA polymerase [253]. This patent provides methods in general to produce chimeric DNA polymerases composed of functional protein domains from different DNA polymerases. The chimeric DNA polymerases are proposed to be useful for PCR.

OTHER DNA POLYMERASE APPLICATIONS

Patents for the DNA sequence, amino acid sequence and purification of an alternative DNA pol kappa were deposited for use in the fields of autoimmunity, cancer, and apoptosis [254,255]. The genes encoding DNA pol kappa and GPBP (Goodpasture antigen binding protein) are present in a head-to-head arrangement in the human genome (5q12-13). The genes share a common promoter for divergent expression that is regulated by the tumor necrosis factor (TNF (α/β)). The patents provide a method to detect an autoimmune condition that is based on the increased expression of alternative DNA pol kappa mRNA or protein. Treatment methods are suggested that inhibit TNF induction of DNA pol kappa gene expression.

Another useful application of DNA polymerases is to improve their replication fidelity in order to reduce spontaneous mutation rates in the context of gene expression systems and for production of fermentation products. A 50-fold reduction in spontaneous mutation frequencies was achieved in *E. coli* by an antimutator allele of *dnaE* (encodes the α subunit of the pol III holoenzyme) and an antimutator allele of *dinB* (encodes DNA pol IV), and by up-regulating *mutL*, which is a component of mismatch repair [256].

CURRENT & FUTURE DEVELOPMENTS

DNA polymerases used today for DNA sequencing, DNA amplification, site-directed mutagenesis, and many other DNA manipulations have been altered in many ways to enhance their performance for these applications. Since current techniques are being improved and several novel methods are being developed, additional DNA polymerase modifications will be needed. Thus, continuing DNA polymerase research and further "improved" DNA polymerases will continue to increase patents deposited on DNA polymerases and their uses for several years to come.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. L.R-K is a Scientist of the Alberta Heritage Foundation for Medical Research.

REFERENCES

- [1] Kornberg A. Biologic synthesis of deoxyribonucleic acid. Science 1960; 131(3412): 1503-1508.
- [2] Kornberg A, Baker TA. DNA replication, second ed. WH Freeman and Co 1992.
- [3] Steitz TA. DNA- and RNA-dependent DNA polymerases. Curr Opin Struct Biol 1993; 3(1): 31-38.
- [4] Braithwaite DK, Ito J. Compilation, alignment, and phylogenetic relationships of DNA polymerases. Nucleic Acids Res 1993; 21(4): 787-802.
- [5] Ohmori H, Friedberg EC, Fuchs RPP, et al. The Y-family of DNA polymerases. Mol Cell 2001; 8(1): 7-8.

- [6] Hogg M, Wallace SS, Doublé. Bumps in the road: how replicative DNA polymerases see DNA damage. *Curr Opin Struct Biol* 2005; 15(1): 86-93.
- [7] Goodman MF. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem* 2002; 71: 17-50.
- [8] Böhlke K, Pisani FM, Rossi M, Antranikian X. Archaeal DNA replication: spotlight on a rapidly moving field. *Extremophiles* 2002; 6:1-14.
- [9] Pavlov AR, Pavlova NV, Kozyavkin SA, Slesarev AI. Recent developments in the optimization of thermostable DNA polymerases for efficient applications. *Trends Biochem Sci* 2004; 22(5): 253-260.
- [10] Sanger F. Determination of nucleotide sequences in DNA. *Science* 1981; 214(4526): 1205-1210.
- [11] Mullis K. The polymerase chain reaction. Nobel Lectures, Chemistry 1991-1995. http://nobelprize.org/nobel_prizes/chemistry/laureates/1993/mullis-lecture.html.
- [12] Smith M. Synthetic DNA and biology. Nobel Lectures, Chemistry 1991-1995. http://nobelprize.org/nobel_prizes/chemistry/laureates/1993/smith-lecture.html.
- [13] Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA. Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* 1985; 313(6005): 762-766.
- [14] Brutlag D, Kornberg A. Enzymatic synthesis of deoxyribonucleic acid. XXXVI. A proofreading function for the 3'→5' exonuclease activity in deoxyribonucleic acid polymerases. *J Biol Chem* 1972; 247(1): 241-248.
- [15] Muzyczka N., Poland RL, Bessman MJ. Studies on the biochemical basis of spontaneous mutation. I. A comparison of the deoxyribonucleic acid polymerases of mutator, antimutator, and wild type strains of bacteriophage T4. *J Biol Chem* 1972; 247(22): 7116-7122.
- [16] Wieczorek A, McHenry CS. The NH₂-terminal php domain of the α subunit of the *Escherichia coli* replicase binds the ϵ proofreading subunit. *J Biol Chem* 2006; 281(18): 12561-12567.
- [17] Aravind L, Koonin EV. Phosphoesterase domains associated with DNA polymerases of diverse origins. *Nucleic Acids Res* 1998; 26(16): 3746-3752.
- [18] Stano N, Chen J, McHenry CS. A coproofreading Zn²⁺-dependent exonuclease within a bacterial replicase. *Nat Struct Mol Biol* 2006; 13(5): 458-459.
- [19] Bailey S, Wing RA, Steitz TA. The structure of *T. aquaticus* DNA polymerase III is distinct from eukaryotic replicative DNA polymerases. *Cell* 2006; 126(5): 893-904.
- [20] Lasken RS, Schuster DM, Rashtchian A. Archaeobacterial DNA polymerases tightly bind uracil-containing DNA. *J Biol Chem* 1996; 271(30): 17692-17696.
- [21] Fogg MJ, Pearl LH, Connolly BA. Structural basis for uracil recognition by archaeal family B DNA polymerases. *Nat Struct Biol* 2002; 9(12): 922-927.
- [22] Shuttleworth G, Fogg MJ, Kurpiewski MR, Jen-Jacobson L, Connolly BA. Recognition of the pro-mutagenic base uracil by family B DNA polymerases from Archaea. *J Mol Biol* 2004; 337(3): 621-634.
- [23] Gruz P, Shimizu M, Pisani FM, De Felice M, Kanke Y, Nohmi T. Processing of DNA lesions by archaeal DNA polymerases from *Sulfolobus solfataricus*. *Nucleic Acids Res* 2003; 31(14): 4024-4030.
- [24] Johnson J, O'Donnell M. Cellular DNA replicases; components and dynamics at the replication fork. *Annu Rev Biochem* 2005; 74: 283-315.
- [25] Modrich P, Richardson CC. Bacteriophage T7 deoxyribonucleic acid replication *in vitro*. Bacteriophage T7 DNA polymerase: an enzyme composed of phage- and host-specific subunits. *J Biol Chem* 1975; 250(14): 5515-5522.
- [26] Tabor S, Huber HE, Richardson CC. *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7 DNA polymerase and primed templates. *J Biol Chem* 1987; 262(33): 16224-16232.
- [27] Das SK, Fujimara RK. Processiveness of DNA polymerases. A comparative study using a simple procedure. *J Biol Chem* 1979; 254(4): 1227-1232.
- [28] Andraos N, Tabor S, Richardson, CC. The highly processive DNA polymerase of bacteriophage T5. *J Biol Chem* 2004; 279(48): 50609-50618.
- [29] Blanco L, Bernad A, Lázaro JM, Martin G, Garmendia, C, Salas, M. Highly efficient DNA synthesis by the phage ϕ 29 DNA polymerase. *J Biol Chem* 1989; 264(15): 8935-8940.
- [30] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 1977; 75(12): 5463-5467.
- [31] Brutlag D, Atkinson MR, Setlow P, Kornberg A. An active fragment of DNA polymerase produced by proteolytic cleavage. *Biochem Biophys Res Commun* 1969; 37(6): 982-989.
- [32] Klenow H, Henningsen I. Selected elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from *Escherichia coli* B by limited proteolysis. *Proc Natl Acad Sci USA* 1970; 65(1): 168-175.
- [33] Ronaghi M, Uhlén M, Nyrén P. A sequencing method based on real-time pyrophosphate. *Science* 1998; 281(5375): 363-365.
- [34] Leamon JH, Braverman MS, Rothberg JM. High-throughput, massively parallel DNA sequencing technology for the era of personalized medicine. *Gene Ther Reg* 2007; 3(1): 15-31.
- [35] Levy S, Sutton G, Ng PC, *et al.* The diploid genome sequence of an individual human. *PLoS Biol* 2007; 5(10): e254 doi:10.1371/journal.pbio.0050254.
- [36] Guyer RL, Koshland Jr. DE. The molecule of the year. *Science* 1989; 246(4937): 1543-1546.
- [37] Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor Symp Quant Biol* 1986; 51:263-273.
- [38] Chien A, Edgar DB, Trela JM. Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J Bacteriol* 1976; 127(3): 1550-1557.
- [39] Kaledin AS, Sliusarenko AG, Gorodetskii SI. Isolation and properties of DNA polymerase from extreme thermophilic bacteria *Thermus aquatiacus* YT-1. *Biokhimiia* 1980; 45(4): 644-651.
- [40] Saiki RK, Gelfand DH, Stoffel S, *et al.* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 1988; 239(4839): 487-491.
- [41] Lawyer FC, Stoffel S, Saiki RK, Myambo K, Drummond R, Gelfand DH. Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J Biol Chem* 1989; 264(11): 6427-6437.
- [42] Eom SH, Wang J, Steitz TA. Structure of *Taq* DNA polymerase with DNA at the polymerase site. *Nature* 1996; 382(6588): 278-281.
- [43] Barnes WM. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc Natl Acad Sci USA* 1994; 91(6): 2216-2220.
- [44] Barnes, W. M.: US5436149 (1995).
- [45] Cheng, S.: US5512462 (1996).
- [46] Sorge, J.A., Mullinax, R.L.: US20077312052 (2007).
- [47] Lederberg, J. A view of genetics. In: *Les Prix Nobel en 1958, 1959*. Norstedt and Sons, Stockholm pp. 170-189.
- [48] Zoller MJ, Smith M. Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. *Nucleic Acids Res* 1982; 10(20): 6487-6500.
- [49] Goulian M. Initiation of the replication of single-stranded DNA by *Escherichia coli* DNA polymerase. *Cold Spring Harbor Symp Quant Biol* 33: 11-20.
- [50] Smith M. *In vitro* mutagenesis. *Annu Rev Genet* 1985; 19: 423-462.
- [51] Hutchison III CA, Phillips S, Edgell MH, *et al.* Mutagenesis at a specific position in a DNA sequence. *J Biol Chem* 1978; 253(18):6551-6560.
- [52] Rigby PWJ, Dieckmann M, Rhodes C, Berg P. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J Mol Biol* 1977; 113(1): 237-251.
- [53] Maniatis T, Fritsch EF, Sambrook J. In: *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982; 109-112.
- [54] Feinberg AP, Vogelstein B. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* 1983; 132(6): 6-13.
- [55] Maniatis T, Fritsch EF, Sambrook J. In: *Molecular Cloning, A Laboratory Manual*, 1982. Cold Spring Harbor Laboratory, 113-121.

- [56] Cadwell RC, Joyce GF. Mutagenic PCR. *PCR Methods Appl* 1994; 3: 136-140.
- [57] Beckman RA, Mildvan AS, Loeb LA. On the fidelity of DNA replication: manganese mutagenesis *in vitro*. *Biochemistry* 1985; 24(21): 5810-5817.
- [58] Beaudry AA, Joyce GF. Directed evolution of an RNA enzyme. *Science* 1992; 257(5070): 635-641.
- [59] Bartel DP, Szostak JW. Isolation of new ribozymes from a large pool of random sequences. *Science* 1993; 261(5127): 1411-1417.
- [60] Dean FB, Nelson JR, Giesler TL, Lasken RS. Rapid amplification of plasmid and phage DNA using phi29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res* 2001; 11(6): 1095-1099.
- [61] Dean, F.B., Lasken, R.S., Fang, L., Faruqi, A.F., Alsmadi, O.A., Driscoll, M.D., Hosono, S., Wisniewski, M., Song, W.: US20030143587A1 (2003).
- [62] Hutchison III CA, Smith HO, Pfannkock C, Venter JC. Cell-free cloning using phi29 DNA polymerase. *Proc Natl Acad Sci* 2005; 102(48): 17332-17336.
- [63] Hamilton MD, Nuara AA, Gammon DB, Buller RM, Evans DH. Duplex strand joining reactions catalyzed by vaccinia virus DNA polymerase. *Nucleic Acids Res* 2007; 35(1): 143-151.
- [64] Evans, D.H., Willer, D.O., Xiao-dan, Y.: US20030162265 (2003).
- [65] Evans, D.H., Willer, D.O., Xiao-dan, Y.: EP1263973 (2006).
- [66] Zhu B, Cai G, Hall EM, Freeman GJ. In-Fusion™ assembly: seamless engineering of multidomain fusion proteins, modular vectors, and mutations. *Biotechniques* 2007; 43(3): 354-359.
- [67] Berrow NS, Alderton D, Sainsbury S, *et al.* A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res* 2007; 35(6): e45.
- [68] Kelly WS, Chalmers K, Murray NE. Isolation and characterization of a λ polA transducing phage. *Proc Natl Acad Sci USA* 1977; 74(12): 5632-5636.
- [69] Joyce CM, Grindley NDF. Construction of a plasmid that overproduces the large proteolytic fragment (Klenow fragment) of DNA polymerase I of *Escherichia coli*. *Proc Natl Acad Sci USA* 1983; 80(4): 1830-1834.
- [70] Spicer EK, Rush J, Fung C, Reha-Krantz LJ, Karam JD, Konigsberg, WH. Primary structure of T4 DNA polymerase. *J Biol Chem* 1988; 263(16): 7478-7486.
- [71] Lin, T.-C., Konigsberg, W., Spicer, E., Rush, J.: US4935361 (1990).
- [72] Chatterjee, D.K.: US5047342 (1991).
- [73] Tabor, S., Richardson, C.C.: US4795699 (1989).
- [74] Tabor, S., Richardson, C.C.: US4946786 (1990).
- [75] Garmendia C, Bernad A, Esteban, JA, Blanco L, Salas M. The bacteriophage phi29 DNA polymerase, a proofreading enzyme. *J Biol Chem* 1992; 267(4): 2594-2599.
- [76] Blanco, L., Bernad, A., Salas, M.: US5001050 (1991).
- [77] Xu, P., Haering, M., Garrett, R. Prangishvili, D.: WO07132358A2 (2007).
- [78] Xu, P., Haering, M., Garrett, R. Prangishvili, D.: EP1854875 (2007).
- [79] Lacks, S.A., Martinez, S., Lopen, P., Espinosa, M.: US5002875 (1991).
- [80] Bullard, J.J., Janjic, N., McHenry, C.S.: EP1816213 (2007).
- [81] Bullard, J.J., Janjic, N., McHenry, C.S.: US20040235766 (2004).
- [82] O'Donnell, M.E., Bruck, I., Zhang, D., Whipple, R.: WO01009164 (2001).
- [83] Gelfand, D. H., Stoffel, S., Lawyer, F. C., Saiki, R. K.: US4889818 (1989).
- [84] Gelfand, D. H., Stoffel, S., Lawyer, F. C., Saiki, R. K.: US5079352 (1992).
- [85] Oshima, T., Sakashita, H., Matsumoto, H., Maekawa, Y.: US5192674 (1993).
- [86] Gelfand, D. H., Lawyer, F. C.: US5374553 (1994).
- [87] Callen, W., Mathur, E.J.: US5948666 (1999).
- [88] Slater, M.R., Huang, F., Hartnet, J.R., Bolchakova, E., Storts, D.R., Otto, P., Miller, K.M., Novikov, A., Velikodvorskaya, G.A.: US20006077664 (2000).
- [89] Mamone, J.A., Davis, M., Sha, D.: US5744312 (1998).
- [90] Ankenbauer, W., Markau, U., Svetlichny, V., Schmitz-Agheguian, G., Reiser, A., Angerer, B., Ebenbichler, C., Laue, F., Bonch-Osmolovskaya, E.: US20026468775 (2002).
- [91] Callen, W., Mathur, E.J.: US20026492511 (2002).
- [92] Kong, H., Pelletier, J.J., Aliotta, J.M.: US5814506 (1998).
- [93] Riggs, M.G., Sivaram, M., Tudor, S.D.: US20006100078 (2000).
- [94] O'Donnell, M.E., Yuzhakov, A., Yurieva, O., Jeruzalmi, D., Bruck, I., Kuriyan, J.: US20050112580 (2005).
- [95] O'Donnell, M.E., Yuzhakov, A., Yurieva, O., Jeruzalmi, D., Bruck, I., Kuriyan, J.: US20050153299 (2005).
- [96] Bullard JM, Williams JC, Acker WK. DNA polymerase III holoenzyme from *Thermus thermophilus*. Identification, expression, purification of components, and use to reconstitute a processive replicase. *J Biol Chem* 2002; 277(16): 13401-13408.
- [97] Janjic, N., Bullard, J.M., McHenry, C.S., Kery, V.: US20046677146 (2004).
- [98] Gu, T., Huang, F., Hartnett, J.R.: US20077214522 (2007).
- [99] O'Donnell, M.E., Yuzhakov, A., Yurieva, O., Jeruzalmi, D., Bruck, I., Kuriyan, J.: US20050048510 (2005).
- [100] Bruck I, Yuzhakov A, Yurieva O, *et al.* Analysis of a multicomponent thermostable DNA polymerase III replicase from an extreme thermophile. *J Biol Chem* 2002; 277(19): 17334-17348.
- [101] O'Donnell, M.E., Yurieva, O.: US20056897053 (2005).
- [102] Perler FB, Comb, DG, Jack, WE, *et al.* Intervening sequences in an Archaea DNA polymerase gene. *Proc Natl Acad Sci USA* 1992; 89(12): 5577-5581.
- [103] Kon H, Kucera RB, Jack WE. Characterization of a DNA polymerase from the hyperthermophile archaea *Thermococcus litoralis*. Vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities. *J Biol Chem* 1993; 268(3): 1965-1975.
- [104] Comb, D.G., Perler, F., Kucera, R., Jack, W.E.: US5322785 (1994).
- [105] Comb, D.G., Perler, F., Kucera, R., Jack, W.E.: US5352778 (1994).
- [106] Comb, D.G., Perler, F., Kucera, R., Jack, W.E.: US5500363 (1996).
- [107] Comb, D.G., Perler, F., Kucera, R., Jack, W.E.: US5834285 (1998).
- [108] Southworth MW, Kong H, Kucera RB, Ware J, Jannasch HW, Perler FB. Cloning of thermostable DNA polymerases from hyperthermophilic marine Archaea with emphasis on *Thermococcus* sp. 9^N-7 and mutations affecting 3'-5' exonuclease activity. *Proc Natl Acad Sci USA* 1996; 93(11): 5281-5285.
- [109] Perler, F.B., Southworth, M.W.: US5756334 (1998).
- [110] Luhm, R.A., D'Hennezel, O.B., Duffaud, G.D., Jolly, J.F., Kelly, R.M., Ting, E.Y.: US5602011 (1997).
- [111] Sorge, J.A., Hogrefe, H.H., Hansen, C.J.: US20056946273 (2005).
- [112] Blumentals, I.I., Lasken, R.S., Schmidt, B.J., Longo, M.C., Hughes Jr., J.A., Chatterjee, D.K.: USH001531 (1996).
- [113] Gelfand, D.H., Wang, A.M.: US5491086 (1996).
- [114] Mathur EJ, Adams MW, Callen WN, Cline JM. The DNA polymerase gene from the hyperthermophilic marine archaeobacterium, *Pyrococcus furiosus*, shows sequence homology with alpha-like DNA polymerases. *Nucleic Acids Res* 1991; 19(24): 6952.
- [115] Mathur, E.J.: US5866395 (1999).
- [116] Imanaka, T., Takagi, M., Morikawa, M.: US20006143536 (2000).
- [117] Copeland, W.C., Wang, T.S.F.: US20036670161 (2003).
- [118] Friedberg, E.C., Gerlach, V., Feaver, W.J.: US20030017573A1 (2003).
- [119] Wang, Z., Lin, W., Xin, H., Wu, X.: US20046677442 (2004).
- [120] Mahajan, P.B., Kannan, P.: US20030084476 (2003).
- [121] Derbyshire, V, Grindley, NDF, Joyce CM. The 3'-5' exonuclease of DNA polymerase I of *Escherichia coli*: contribution of each amino acid at the active site to the reaction. *EMBO J* 1991; 10(1): 17-24.
- [122] Reha-Krantz LJ. In: DNA Replication and Mutagenesis, Ch. 4: Collection of informative bacteriophage T4 DNA polymerase mutants. Moses RE, Summers WC, Eds. American Society for Microbiology, Washington DC1988: 34-40.
- [123] Reha-Krantz LJ. Amino acid changes coded by bacteriophage T4 DNA polymerase mutator mutants. *J Mol Biol* 1988; 202(4): 711-724.
- [124] Blanco L, Bernda A, Blasco MA, Salas M. A general structure for DNA-dependent DNA polymerases. *Gene* 1991; 100: 27-38.
- [125] Bambara RA, Uyemara D, Choi T. On the processive mechanism of *Escherichia coli* DNA polymerase I. Quantitative assessment of processivity. *J Biol Chem* 1978; 253(2): 413-423.
- [126] Atkinson MR, Deutscher MP, Kornberg A, Russell AF, Moffatt JG. Enzymatic synthesis of deoxyribonucleic acid. XXXIV.

- Termination of chain growth by a 2',3'-dideoxyribonucleotide. *Biochem* 1969; 8(12): 4897-4904.
- [127] Tabor S, Richardson SS. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc Natl Acad Sci USA* 1987; 84(14): 4767-4771.
- [128] Tabor S, Richardson CC. Effect of manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and *Escherichia coli* DNA polymerase I. *Proc Natl Acad Sci USA* 1989; 86(11): 4076-4080.
- [129] Tabor S, Richardson CC. A single residue in DNA polymerases of the *Escherichia coli* DNA polymerase I family is critical for distinguishing between deoxy- and dideoxyribonucleotides. *Proc Natl Acad Sci USA* 1995; 92(14): 6339-6343.
- [130] Engler MJ, Lechner RL, Richardson CC. Two forms of the DNA polymerase of bacteriophage T7. *J Biol Chem* 1983; 258(18): 11165-11173.
- [131] Tabor S, Richardson C.C.: US4942130 (1990).
- [132] Tabor S, Richardson CC. Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by *in vitro* mutagenesis. *J Biol Chem* 1989; 264(11): 6447-6458.
- [133] Tabor S, Richardson C.C.: US5266466 (1993).
- [134] Tabor S, Richardson C.C.: US5409811 (1995).
- [135] Tabor S, Richardson C.C.: US5674716 (1997).
- [136] Tabor S, Richardson C.C.: US5639608 (1997).
- [137] Chatterjee, D.K.: US5270179 (1993).
- [138] Chatterjee, D.K.: US5541099 (1996).
- [139] Chatterjee, D.K.: US5716819 (1998).
- [140] Innis MA, Myambo KB, Gelfand DH, Brow MAD. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc Natl Acad Sci USA* 1988; 85(24): 9436-9440.
- [141] Innis, M.A., Myambo, K. B., Gelfand, D.H., Brow, M.A.D.: US5075216 (1991).
- [142] Tabor S, Richardson C.C.: US5614365 (1997).
- [143] Reeve MA, Fuller CW. A novel thermostable polymerase for DNA sequencing. *Nature* 1995; 376(6542): 796-797.
- [144] Li Y, Mitaxov V, Waksman G. Structure-based design of *Taq* DNA polymerases with improved properties of dideoxynucleotide incorporation. *Proc Natl. Acad Sci USA* 1999; 96(17): 9491-9496.
- [145] Yang S-W, Astatke M, Potter J, Chatterjee DK. Mutant *Thermotoga neapolitana* DNA polymerase I: altered catalytic properties for non-templated nucleotide addition and incorporation of correct nucleotides. *Nucleic Acids Res* 2002; 30(19): 4314-4320.
- [146] Fuller, C.W.: US5741640 (1998).
- [147] Davis, M., Moffett, B.R., Fuller, C.W., Cunniff, J.J.: US5885813 (1999).
- [148] Brandis, J., Bloom, C., Richards, J.H.: US20016265193 (2001).
- [149] Davis, M., Moffett, B.R., Fuller, C.W., Cunniff, J.J.: US5885813 (1999).
- [150] Gelfand, D.H., Lawyer, F.C., Stoffel, S.: US5420029 (1995).
- [151] Abramson, R.D., Gelfand, D.H.: US5466591 (1995).
- [152] Abramson, R.D., Gelfand, D.H.: US5795762 (1998).
- [153] Barnes, W.M.: US5616494 (1997).
- [154] Riedl, W.A., Fly, S.J., Kaboord, B.F., Nye, S.H., Ting, E.Y.: US5882904 (1999).
- [155] Hughes, J.A., Chatterjee, D.K.: US20006015668 (2000).
- [156] Schoenbrunner, N.J., Myers, T.W., Gelfand, D.H.: US20067148049 (2006).
- [157] Myers TW, Gelfand DH. Reverse transcription and DNA amplification by a *Thermus thermophilus* DNA polymerase. *Biochem* 1991; 30(31): 7661-7666.
- [158] Gelfand, D.H., Myers, T.W.: US5310652 (1994).
- [159] Smith, E.S., Elfstrom, C.M., Gelfand, D.H., Higuchi, R.G., Myers, T.W., Schonbrunner, N.J., Wang, A.M.: US 20020012970 (2002).
- [160] Blanco, L., Bernad, A., Salas, M.: US5198543 (1993).
- [161] Blanco, L., Bernad, A., Salas, M.: US5576204 (1996).
- [162] Goodman, M.F., Reha-Krantz, L.J.: US5547859 (1996).
- [163] Reha-Krantz, L.J., Goodman, M.F.: US5928919 (1999).
- [164] Reha-Krantz LJ, Nonay RL. Genetic and biochemical studies of bacteriophage T4 DNA polymerase 3'→5' exonuclease activity. *J Biol Chem* 1993; 268(36): 27100-27108.
- [165] Elisseeva E, Mandal SS, Reha-Krantz LJ. Mutational and pH studies of the 3'→5' exonuclease activity of bacteriophage T4 DNA polymerase. *J Biol Chem* 1999; 274(35): 25151-25158.
- [166] Nossal NG. Protein-protein interactions at a DNA replication fork: bacteriophage T4 as a model. *FASEB J* 1992; 6: 871-878.
- [167] Bittner M., Burke RL, Alberts BM. Purification of the T4 gene 32 protein free from detectable deoxyribonuclease activities. *J Biol Chem* 1979; (19): 9565-9572.
- [168] Wang TS-F, Wong, SW, Korn D. Human DNA polymerase α : predicted functional domains and relationships with viral DNA polymerases. *FASEB J* 1989; 3: 14-21.
- [169] Delarue M, Pick O, Tordo N, Moras Argos P. An attempt to unify the structure of polymerases. *Protein Eng* 1990; 3: 461-467.
- [170] Joyce CM. Choosing the right sugar: How polymerases select a nucleotide substrate. *Proc Natl Acad Sci USA* 1997; 94(5): 1619-1622.
- [171] Reha-Krantz LJ, Nonay RL, Stocki S. Bacteriophage T4 DNA polymerase mutations that confer sensitivity to the PPi analog phosphonoacetic acid. *J Virol* 1993; 67(1): 60-66.
- [172] Fidalgo da Silva E, Mandal SS, Reha-Krantz LJ. Using 2-aminopurine fluorescence to measure incorporation of incorrect nucleotides by wild type and mutant bacteriophage T4 DNA polymerases. *J Biol Chem* 2002; (43): 40640-40649.
- [173] Tullius MV, Harth G, Horwitz MA. High extracellular levels of *Mycobacterium tuberculosis* glutamine synthetase and superoxide dismutase in actively growing cultures are due to high expression and extracellular stability rather than to a protein-specific export mechanism. *Infect Immun* 2001; 69(10): 6348-6363.
- [174] Goodman, M.F., Reha-Krantz, L.J.: US5945312 (1999).
- [175] Reha-Krantz LJ, Nonay RL. Motif A of bacteriophage T4 DNA polymerase: role in primer extension and DNA replication fidelity. Isolation of new antimutator and mutator DNA polymerases. *J Biol Chem* 1994; 269(8): 5635-5643.
- [176] Dong A, Copeland WC, Wang T,S-F. Mutational studies of human DNA polymerase α . Identification of residues critical for deoxynucleotide binding and misinsertion fidelity of DNA synthesis. *J Biol Chem* 1993; 268(15): 24163-24174.
- [177] Li L, Murphy KM, Kanevets U, Reha-Krantz LJ. Sensitivity to phosphonoacetic acid: a new phenotype to probe DNA polymerase δ in *Saccharomyces cerevisiae*. *Genetics* 2005; (2): 569-580.
- [178] Venkatesan RN, Hsu JJ, Lawrence NA, Preston BD. Mutator phenotypes caused by substitution at a conserved motif A residue in eukaryotic DNA polymerase δ . *J Biol Chem* 2006; 281(7): 4486-4494.
- [179] Jett, J.H., Keller, R.A., Martin, J.C., Moyzis, R.K., Ratliff, R.L., Shera, B.E., Stewart, C.C.: US4962037 (1990).
- [180] Hariharan C, Bloom LB, Helquist SA, Kool ET, Reha-Krantz, LJ. Dynamics of nucleotide incorporation: snapshots revealed by 2-aminopurine fluorescence studies. *Biochemistry* 2006; 45(9): 2836-2844.
- [181] Goodwin PM, Affleck RL, Ambrose, *et al.* Progress toward DNA sequencing at the single molecule level. *Exp Tech Phys* 1995; 41: 279-294.
- [182] Werner JH, Cai H, Jett JH, Reha-Krantz, L Keller RA, Goodwin PM. Progress towards single-molecule DNA sequencing: a one color demonstration. *J Biotech* 2003; 102(1): 1-14.
- [183] Aksyonov SA, Bittner M, Bloom LB, *et al.* Multiplexed DNA sequencing-by-synthesis. *Anal Biochem* 2006; 348(1): 127-138.
- [184] Williams, P., Hayes, M.A., Rose, S.D., Bloom, L.B., Reha-Krantz, L.J., Pizziconi, V.B.: US20067037687 (2006).
- [185] Seo TS, Bai X, Kim DH, *et al.* Four-color DNA sequencing by synthesis on a chip using photocleavable fluorescent nucleotides. *Proc Natl Acad Sci USA* 2005; 102(17): 5926-5931.
- [186] Meng Q, Kin DH, Bai X, Bi L., Turros NJ, Ju, J. Design and synthesis of a photocleavable fluorescent nucleotide 3'-O-allyl-dGTP-PC-bodipy-FL-510 as a reversible terminator for DNA sequencing by synthesis. *J Org Chem* 2006; 71(8): 3248-3252.
- [187] Ju J, Kim DH, Bi L, *et al.* Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators. *Proc Natl Acad Sci USA* 2006; 103(52): 19635-19640.
- [188] Gardner AF, Jack WE. Determinants of nucleotide sugar recognition in an archaeon DNA polymerase. *Nucleic Acids Res* 1999; 27(12): 2545-2553.
- [189] Gardner AF, Jack WE. Acyclic and dideoxy terminator preferences denote divergent sugar recognition by archaeon and *Taq* DNA polymerases. *Nucleic Acids Res* 2002; 30(2): 605-613.
- [190] Ichida JK, Horhota A, Zou K, McLaughlin LW, Szostak JW. High fidelity TNA synthesis by Terminator polymerase. *Nucleic Acids Res* 2005; 33(16): 5219-5225.
- [191] Sorge, J.A., Hogrefe, H.H., Hansen, C.J.: US2005 6946273 (2005).
- [192] Mathur, E.J.: US5489523 (1996).

- [193] Komatsubara, H., Kitabayashi, M., Kamimura, H., Kawakami, B., Kawamura, Y., Takagi, M., Imanaka, T.: US6008025 (1999).
- [194] Muehlegger, K., Angerer, B., Seela, F., Ankenbauer, W., Augustin, M., Gumbiowski, K., Zulauf, M.: US20040214221 (2004).
- [195] Mamone, J.A.: US5827716 (1998).
- [196] Evans, S., Mamone, J.A., Davis, M., Connolly, B.A.: US2001633183 (2001).
- [197] Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* 2003; 299(5605): 682-686.
- [198] Korlach J, Marks PJ, Cicero RL, *et al.* Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures. *Proc Natl Acad Sci* 2008; 105(4): 1176-1181.
- [199] Korlach, J., Webb, W., Levene, M., Craighead, H.G., Foquet, M.: US20067056661B2 (2006).
- [200] Hanzel, D.K., Otto, G., Murphy, D., Peluso, P., Pham, T., Rank, D.R., Mitsis, P.: US20070196846A1 (2007).
- [201] Hanzel, D.K., Otto, G., Murphy, D., Peluso, P., Pham, T., Rank, D.R., Mitsis, P.: WO07076057A2 (2007).
- [202] Kuroita, T., Kitabayashi, M., Ishida, Y., Komatsubara, H., Nishiya, Y., Kawakami, B., Kawamura, Y., Imanaka, T.: US20020076768A1 (2002).
- [203] Garmendia C, Bernad A, Esteban JA, Blanco L, Salas M. The bacteriophage ϕ 29 DNA polymerase, a proofreading enzyme. *J Biol Chem* 1992; 267(4): 2594-2599.
- [204] Greagg MA, Fogg MJ, Panayotou G, *et al.* A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil. *Proc Natl Acad Sci USA* 1999; 96(16): 9045-9050.
- [205] Connolly, B., Fogg, M., Pearl, L.: US20060057682 (2006).
- [206] Sorge, J.A., Hogrefe, H., Hansen, C.J., Arezi, B., Mullinax, R.L.: US20050123940 (2005).
- [207] Connolly, B., Biles, B.: US20070269877 (2007).
- [208] Peters, L.-E.: WO06108179A2 (2006).
- [209] Goodman, M.F., Reha-Krantz, L.J.: US5945312 (1999).
- [210] Baker RP, Reha-Krantz LJ. Identification of a transient excision intermediate at the crossroads between DNA polymerase extension and proofreading pathways. *Proc Natl Acad Sci USA* 1998; 95(7): 3507-3512.
- [211] Marquez L, Reha-Krantz LJ. Using 2-aminopurine fluorescence and mutational analysis to demonstrate an active role of bacteriophage T4 DNA polymerase in strand separation required for 3'→5' exonuclease activity. *J Biol Chem* 1996; 271(46): 28903-28911.
- [212] Reha-Krantz LJ, Marquez LA, Elisseva E, *et al.* The proofreading pathway of bacteriophage T4 DNA polymerase. *J Biol Chem* 1998; 273(36): 22969-22976.
- [213] Murphy K, Darmawan H, Schultz A, *et al.* A method to select for mutator DNA polymerase δ s in *Saccharomyces cerevisiae*. *Genome* 2006; 49: 403-410.
- [214] Stocki SA, Nonay RL, Reha-Krantz LJ. Dynamics of bacteriophage T4 DNA polymerase function: identification of amino acid residues that affect switching between polymerase and 3'→5' exonuclease activities. *J Mol Biol* 1995; 254(1): 15-28.
- [215] Reha-Krantz LJ, Wong C. Selection of bacteriophage T4 antimutator DNA polymerases: a link between proofreading and sensitivity to phosphonoacetic acid. *Mutat Res* 1996; 350: 9-16.
- [216] Reha-Krantz LJ. Use of genetic analyses to probe structure, function, and dynamics of bacteriophage T4 DNA polymerase. *Methods in Enzymology*. Academic Press. 1995; 262: 323-331.
- [217] Hiratatsuka K, Reha-Krantz LJ. Identification of *Escherichia coli dnaE* (*polC*) mutants with altered sensitivity to 2',3'-dideoxyadenosine. *J Bacteriol* 2000; 182(14): 3942-3947.
- [218] Bailey S, Wing RA, Steitz TA. The structure of *T. aquaticus* DNA polymerase III is distinct from eukaryotic DNA polymerases. *Cell* 2006; 126(5): 893-904.
- [219] Rashbaum SA, Cozzarelli NR. Mechanism of DNA synthesis inhibition by arabinosyl cytosine and arabinosyl adenine. *Nature* 1976; 264(5587): 679-680.
- [220] Sweasy JB, Loeb LA. Mammalian DNA polymerase β can substitute for DNA polymerase I during DNA replication in *Escherichia coli*. *J Biol Chem* 1992; 267(3): 1407-1410.
- [221] Kim B, Loeb LA. Human immunodeficiency virus reverse transcriptase substitutes for DNA polymerase I in *Escherichia coli*. *Proc Natl Acad Sci USA* 1995; 92(3): 684-688.
- [222] Patel PH, Loeb LA. DNA polymerase active site is highly mutable: evolutionary consequences. *Proc Natl Acad Sci USA* 2000; 97(10): 5095-5100.
- [223] Patel PH, Loeb LA. Multiple amino acid substitutions allow DNA polymerases to synthesize RNA. *J Biol Chem* 2000; 275(51): 40266-40272.
- [224] Patel PH, Kawate H, Adman E, Ashbach M, Loeb LA. A single highly mutable catalytic site amino acid is critical for DNA polymerase fidelity. *J Biol Chem* 2001; 276(7): 5044-5051.
- [225] Patel, P.H., Loeb, L.A.: US20016329178 (2001).
- [226] Patel, P.H., Loeb, L.A.: US20036602695 (2003).
- [227] Suzuki M, Yoshida S, Adman ET, Blank A, Loeb LA. *Thermus aquaticus* DNA polymerase I mutants with altered fidelity. *J Biol Chem* 2000; 275(42): 32728-32735.
- [228] Loeb, L.A., Hood, L., Suzuki, M.: US20026395524 (2002).
- [229] Loeb, L.A., Hood, L., Suzuki, M.: US20066982144 (2006).
- [230] Loeb, L.A., Hood, L., Suzuki, M.: US20077312059 (2007).
- [231] Jestin J-L, Kristensen P, Winter G. A method for the selection of catalytic activity using phage display and proximity coupling. *Angew Chem Int Ed Engl* 1999; 38(8): 1124-1127.
- [232] Jestin, J.-L., Vichier-Guerre, S., Ferris, S.: US20080014609 (2008).
- [233] Xia G, Chen L, Sers T, Fa M, Schultz PG, Romesberg FE. Directed evolution of novel polymerase activities: mutation of a DNA polymerase into an efficient RNA polymerase. *Proc Natl Acad Sci USA* 2002; 99(10): 6597-6602.
- [234] Holliger, P., Ghadessy, F., Ong, J.L.: US20040005594 (2004).
- [235] Holliger, P., Ghadessy, F., d'Abbadie, M.: US20070020653 (2007).
- [236] Holliger, P., Oliynyk, Z.: US20070134682 (2007).
- [237] Ghadessy FJ, Ong JL, Holliger P. Directed evolution of polymerase function by compartmentalized self-replication. *Proc Natl Acad Sci USA* 2001; 98(8): 4552-4557.
- [238] Ghadessy FJ, Ramsay N, Boudsocq, *et al.* Genetic expansion of the substrate spectrum of a DNA polymerase by directed evolution. *Nat Biotech* 2004; 22(6): 755-759.
- [239] Ong JL, Loakes D, Jaroslowski S, Too K, Holliger P. Directed evolution of DNA polymerase, RNA polymerase and reverse transcriptase activity in a single polypeptide. *J Mol Biol* 2006; 361(3): 537-550.
- [240] d'Abbadie M, Hofreiter M, Vaisman A, *et al.* Molecular breeding of polymerases for amplification of ancient DNA. *Nat Biotech* 2007; 25(8): 939-943.
- [241] Thomas M, Gilbert P, Willerslev E. Rescuing ancient DNA. *Nat Biotech* 2007; 25(8): 872-874.
- [242] Andersen, M.R.: US20050196392 (2005).
- [243] Walker, C.L., Mueller, E.J., Kayser, K.J., Milligan, J.S., Eastlund, E.R.: US20050026147 (2005).
- [244] Walker, C.L., Mueller, E.J., Kayser, K.J., Milligan, J.S., Eastlund, E.R.: US20080044883 (2008).
- [245] Evans, T.C., Slatko, B., Chen, L., Vaisvila, R., Guan, C.: US20060088868 (2006).
- [246] Evans, T.C., Slatko, B., Chen, L., Vaisvila, R., Guan, C., Kucera, R.: US20060177867 (2006).
- [247] Gerard, G.F., Qiu, Z., Gleeson, M.A.: US20070184482 (2007).
- [248] Livneh, Z., Reuven, N.B., Tomer, G.: US20016333178 (2001).
- [249] Boudsocq F, Iwai S, Hanaoka F, Woodgate R. *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol η . *Nucleic Acids Res* 2001; 29(22): 4607-4616.
- [250] Napoli A, Zivanovic Y, Bocs C, *et al.* DNA bending, compaction and negative supercoiling by the architectural protein Sso7d of *Sulfolobus solfataricus*. *Nucleic Acids Res* 2002; 30(12): 2656-2662.
- [251] Wang, Y.: US20036627424 (2003).
- [252] Pavlov AR, Belova GI, Kozyavkin SA, Slesarev AI. Helix-hairpin-helix motifs confer salt resistance and processivity on chimeric DNA polymerases. *Proc Natl Acad Sci USA* 2002; 99(21): 13510-13515.
- [253] Frey, B., Villbrandt, B., Schomburg, D., Sobek, H., Ankenbauer, W.: US20077244602 (2007).
- [254] Saus, J.: US20030027165 (2003).
- [255] Saus, J.: US20060068441 (2006).
- [256] Strobel, H., Schoenfeld, W.: US20070184520 (2007).