JOURNAL OF BACTERIOLOGY, July 2000, p. 3613–3618 Vol. 182, No. 13
0021-9193/00/$04.00+0
Copyright © 2000, American Society for Microbiology. All Rights Reserved.

GUEST COMMENTARY

Ten Commandments: Lessons from the Enzymology of DNA Replication

ARTHUR KORNBERG*

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305-5307

One of the prevailing fashions in bioscience these days is the application of genomics to eukaryotic gene expression. Largely eclipsed are the approaches of a few decades ago in which enzymes derived from microorganisms blazed the trail to much of our current understanding of macromolecular biosynthesis and gene regulation. In this Commentary I will give an anecdotal account of the lessons I learned from my attempts to resolve and reconstitute biological events in DNA replication and reflect on how these lessons may still apply to solving the current problems of growth and development and the aberrations of disease.

The beginning of the 20th century saw the birth of modern biochemistry with the demonstration that alcoholic fermentation could be observed in the juice of yeast cells. This led to the discovery of the dozen enzymes that convert sucrose to alcohol and ultimately to the reconstitution of alcoholic fermentation at a refined molecular level. Along with these early biochemical studies with yeast came the discovery that virtually the same enzymes and pathway were responsible in mammalian cells for the conversion of glycogen to lactic acid, which provides ATP energy for muscle contraction. This astonishing fact, along with many other such examples in metabolic and biosynthetic pathways, made it clear that mechanisms and molecules have been preserved in bacteria, fungi, plants, and animals, essentially intact through billions of years of Darwinian evolution. I regard this insight as one of the great revelations of the 20th century. This and other verities gleaned in our lengthy pursuit of the biochemistry of DNA replications can be presented dogmatically as in the **THOU SHALTS** of the biblical commandments.

I. RELY ON ENZYMOLOGY TO CLARIFY BIOLOGIC QUESTIONS

Based on the conviction that all reactions in the cell are catalyzed and directed by enzymes, the **first commandment** commands that enzymology can be relied on to clarify a biologic question. Chemists once bridled at this. But time and again, spontaneous reactions, such as the melting of DNA and the folding of proteins, are found to be driven and directed by enzymes; in the case of DNA, its melting in a cell is catalyzed by several different helicases. The first and crucial step is to find a way to observe the phenomenon of interest in a cell-free system. Should that succeed, then one should be able to reduce the event to its molecular components by enzyme fractionation. This confidence is derived from the fact that, as mentioned, alcoholic fermentation, which had eluded understanding for centuries, was clarified by fractionation of a cell-free yeast extract, as was glycolysis by fractionation of muscle extracts and in the same vein luminescence in the extracts of a firefly and replication of DNA in microbial cell lysates. Fractionation procedures for these extracts revealed the molecular mechanisms and machines for the catalysis and regulation of many complex reactions and pathways, as recounted here for DNA replication.

With a cell-free system in hand that recreates a biologic event, the biochemist should be able to perform the process as well as the cell does it. Even better! After all, the cell is under great constraints to provide a consensus medium that supports thousands of diverse reactions, only some of which operate under optimal conditions. By contrast, the biochemist enjoys the freedom to saturate each enzyme with its substrate, trap the products, and provide the optimal pH and salt and metal ion concentrations. The biochemist can thus be creative and effective in analyzing the molecular basis of a reaction or pathway.

The refinement of methods to purify proteins by chromatography and to establish their homogeneity by gel electrophoresis, when combined with the power of reverse genetics and genomics, has made the isolation and large-scale preparation of enzymes relatively easy compared to what it was years ago. Despite this, discrete events in the proliferation, differentiation, and adaptations of cells and organisms are almost always analyzed by genetic means. Striking phenotypes are produced by mutations and transfections, but the alterations in enzymes and pathways are generally only inferred. Rarely are they verified by the isolation of proteins with demonstrable functions. To many cell biologists and developmental biologists, the need to examine an event in a cell-free system does not come up on their radar screen.

II. TRUST THE UNIVERSALITY OF BIOCHEMISTRY AND THE POWER OF MICROBIOLOGY

The universality of biochemistry from microbes to humans in basic metabolic and biosynthetic pathways has led to the silly quip: "What's true for *E. coli* is true for elephants, and what's not true for *E. coli* is not true." My faith in this universality encouraged me to focus on how prokaryotes, particularly *Escherichia coli*, replicate their own genomes and those of their phages and plasmids. Microbial generation times, unlike those of eukaryotes, are measured in minutes rather than hours and days. This is where the light on replication shines brightest. I made the choice to work with prokaryotes with the confidence that these systems would be reliable prototypes for how the so-called higher organisms replicate their DNA. In recent years, exciting advances have been made in discovering and characterizing the eukaryotic replication enzymes: the heli-

* Mailing address: Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305-5307. Phone: (650) 723-6167. Fax: (650) 723-6783. E-mail: akornberg@cmgm.stanford.edu.

TABLE 1. Windows opened by viruses: phages M13 and φX174

Window
Provided the DNA substrates that enabled the discovery of RNA priming as a general mechanism for starting a DNA chain
Discovery of the single-strand binding protein needed to identify the origin for priming and to facilitate processive replication
Discovery of the highly mobile primosome
Recognition of a far more complex DNA polymerase, the pol III holoenzyme, a dimeric, 20-subunit, highly processive component of the replisome
Discovery of a helicase that operates at a replication fork

cases, topoisomerases, polymerases, primases, ligases, and other components of the chromosomal replicases. The variations from prokaryotic enzymes are fascinating. Yet virtually all these enzymes and mechanisms were already familiar and adhere to the basic themes discovered earlier in the prokaryotic systems.

A social comment: how ironic that *coli* has become a four-letter word to faculty search committees and granting agencies. The goose that laid so many golden eggs of macromolecular biosynthesis and the awesome networks of signaling and stress responses could lay many more golden eggs if only given the proper nourishment. This attention may finally come with the fresh concerns about antibiotic resistance of pathogens and the roll call of genome sequences that reveal striking homologies and the universality of evolution.

III. DO NOT BELIEVE SOMETHING BECAUSE YOU CAN EXPLAIN IT

In 1950, having found the enzymes that incorporate nucleotides into coenzymes and curious about how they might become part of nucleic acids, I needed first to determine how the purine and pyrimidine bases became substrates for assembly into these polymers. In the course of exploring the biosynthesis of nucleotides, I learned how to use labeled bases and how to tag each of the phosphates of the nucleoside diphosphates (NDPs) and triphosphates (NTPs).

In 1954, we observed an activity in an extract of *E. coli* that incorporated the label of [α-^{32}P]ATP into an acid-insoluble form that we presumed to be RNA. While making progress in purifying this activity, we learned of a discovery in the laboratory of Severo Ochoa in New York. While observing an exchange of [^{32}P]orthophosphate with ADP in extracts of *Azotobacter vinelandii*, they discovered an enzyme that converted the ADP and other NDPs into an RNA-like polymer.

Acting on this information, we substituted [α-^{32}P]ADP for ATP and found that our activity was far greater. Clearly ADP was the preferred substrate over ATP. The *E. coli* enzyme we then purified was the same polynucleotide phosphorylase that the Ochoa laboratory had first identified in *Azotobacter*. As was learned later, the role of the phosphorylase was to degrade RNA rather than effect its synthesis. Had we persisted with ATP as substrate, we would surely have found RNA polymerase, the true synthetic enzyme, a year earlier than we did DNA polymerase and several years before it was discovered in 1961 by the late Sam Weiss.

IV. DO NOT WASTE CLEAN THINKING ON DIRTY ENZYMES

The late Efraim Racker enunciated this commandment, and I have been one of its ardent disciples. A dramatic example is the discovery of DNA replication. I first observed DNA synthesis in an *E. coli* extract in 1955, when I found that 50 counts out of a million of [^{14}C]thymidine were incorporated into an acid-insoluble form. Those few counts above background seemed real because they were susceptible to DNase. Could we possibly figure out what was going on in so crude a system, let alone in an intact cell?

We identified and purified the first DNA polymerase, but not before our fractionation procedures disclosed a variety of novel enzymes that acted initially on the DNA, the [^{14}C]thymidine, and the ATP in our incubation mixtures. First, the [^{14}C]thymidine we added had to be phosphorylated by ATP via a new enzyme, thymidine kinase, to become thymidylate. The added calf thymus DNA proved to be a substrate for DNases that produced the four hitherto unknown deoxynucleoside 5′-monophosphates. These were then phosphorylated by four distinct nucleotide kinases to the corresponding diphosphates, which were in turn phosphorylated by nucleoside diphosphate kinase to the respective and previously unknown dNTPs. The DNA we had added served three additional functions beyond being a source of the four building blocks. It was a template to direct the precise order of nucleotide assembly, a source of primer termini for chain elongation, and a pool to protect the tiny amount of synthesized DNA from degradation by the nucleases that are abundant in cell extracts.

The *E. coli* extract was thus the source of seven new enzymes in addition to the enzyme we named DNA polymerase. With that, template and primer were introduced into the language of all polymerase actions. We learned all this from fractionating the *E. coli* extract into its multiple activities and finally putting the purified enzymes and their products back together. These studies taught us the basic features of how DNA polymerases act and, incidentally, that the strands of duplex DNA are oriented in opposite directions, not known at the time.

Cell extracts are by their nature "dirty enzymes"; intact cells and organisms are "dirtier" still. F. G. Hopkins, a prescient pioneer in the biochemical basis of nutrition, said it best back in 1931: "(The biochemist's word) may not be the last in the description of life, but without his help the last word will never be said." And so it has been with many cellular events, most recently the fully reconstituted transcription by the 48-subunit yeast RNA polymerase II initiation complex and the sorting of proteins to specific subcellular compartments by fractionated vesicles and enzymes.

Purification of an enzyme to homogeneity now opens the door to reverse genetics, and the enzymes themselves still provide unique reagents, as commandments IX and X will describe. Sometimes, an apparently pure enzyme may be found on further purification to harbor a contaminant of great importance. As one example, an extra step in the purification of DNA polymerase I rendered the enzyme inactive. The reason: the template-primer used in our routine assays was DNA activated by being nicked many times by a DNase. We were not aware at the time that these nicks had been enlarged upon by an exonuclease in our polymerase preparation to create a stretch of exposed template needed by the polymerase. The exonuclease activity we had fractionated away, which we then purified and named exonuclease III, proved to be a crucial reagent in the discovery of recombinant DNA (see commandment X). The importance of purifying an enzyme was such that

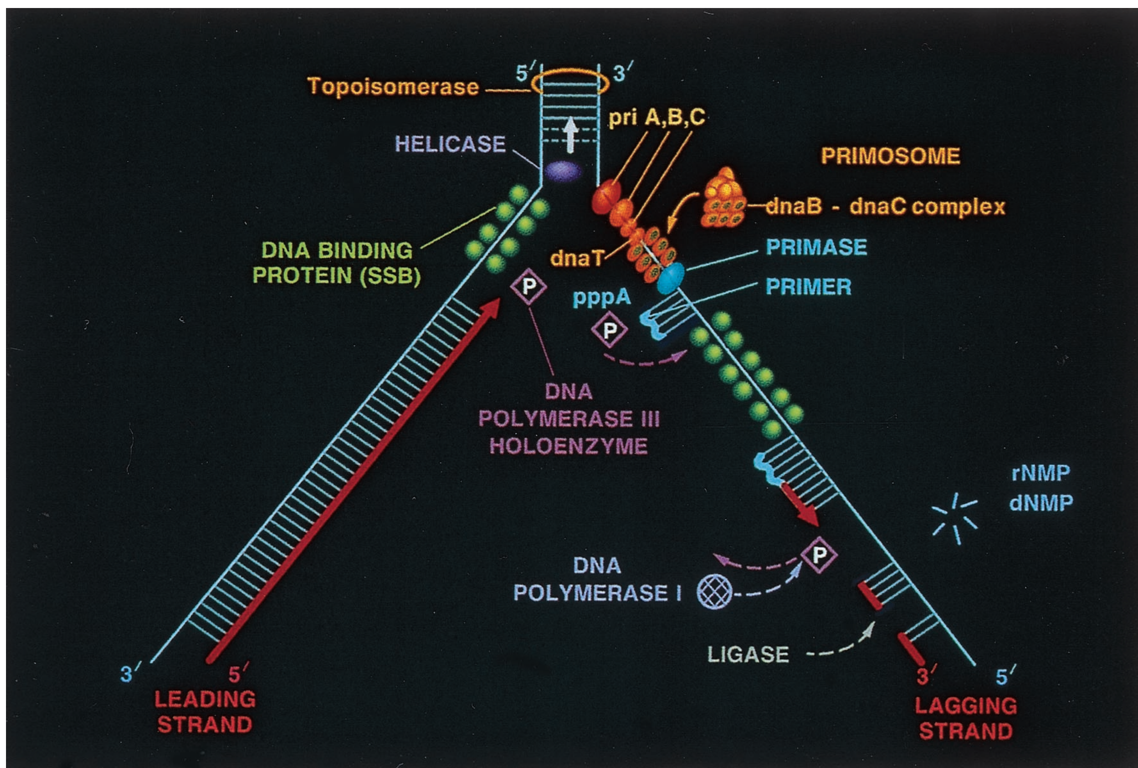

FIG. 1. DNA replication fork. Reprinted from reference 1 with permission of the publisher.

a student of mine had her personalized California license plate read: PURIFY.

In 1967, after 10 years of trying and failing to prove that our enzymatically synthesized DNA was biologically active, we finally succeeded. What made the difference was the use of a single-stranded, circular DNA of a bacterial virus as template and the discovery of DNA ligase that could circularize the linear product. With ϕX174 DNA, we could make quantities of the circularly closed, infectious viral DNA. The sequence of 5,386 nucleotides was correct, and there was no need for novel nucleotides or other components, as had been conjectured. We also pointed out that this in vitro system afforded the means to introduce novel nucleotides for site-directed mutagenesis.

The appearance of our paper announcing the test tube synthesis of infectious DNA generated a huge crush of media attention, a congratulatory phone call from President Lyndon Johnson and headlines worldwide, all based on the belief that we had synthesized a big, hairy virus and "created life in the test tube." I had to explain to the assembled reporters that it was not I who assembled the long DNA chain of the virus but rather it was the awesome enzyme, DNA polymerase, that I had identified and isolated from *E. coli* cells. And further, I had to make it clear that it was these bacteria in a culture flask that imbibed the viral DNA to make the infectious virus particles. As for "creation of life in the test tube," some might dispute that a virus is even a "living" creature.

V. DO NOT WASTE CLEAN ENZYMES ON DIRTY SUBSTRATES

Three years after the hoopla about the synthesis of a viral DNA by our DNA polymerase, serious questions remained. How is a DNA chain started? How is the accumulating genetic evidence for additional polymerases and other factors needed for replication explained? A cartoon at the time showed the apparatus at a replication fork discreetly obscured by a fig leaf, and there were polemical attacks in *Nature New Biology* that dismissed our DNA polymerase as merely a repair enzyme with little relevance to replication, in essence a "red herring."

Over the years, we had tried a variety of DNA samples to demonstrate the start of a DNA chain. The results were negative or equivocal. Then it dawned on me that we were violat-

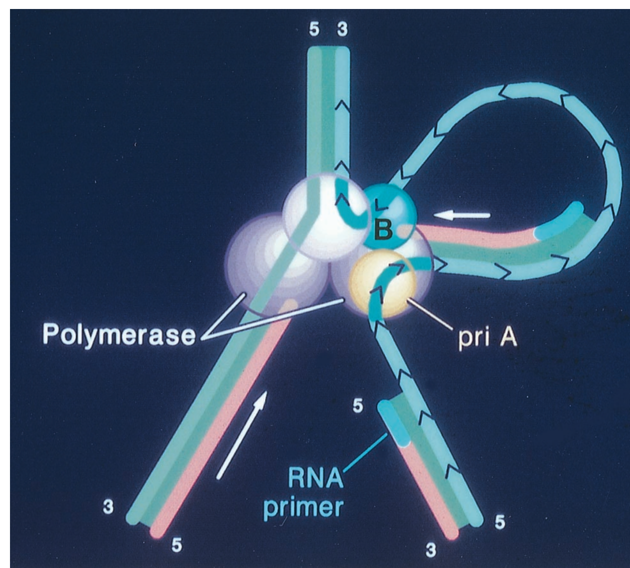

FIG. 2. "Replisome" at a fork: locomotive or sewing machine?

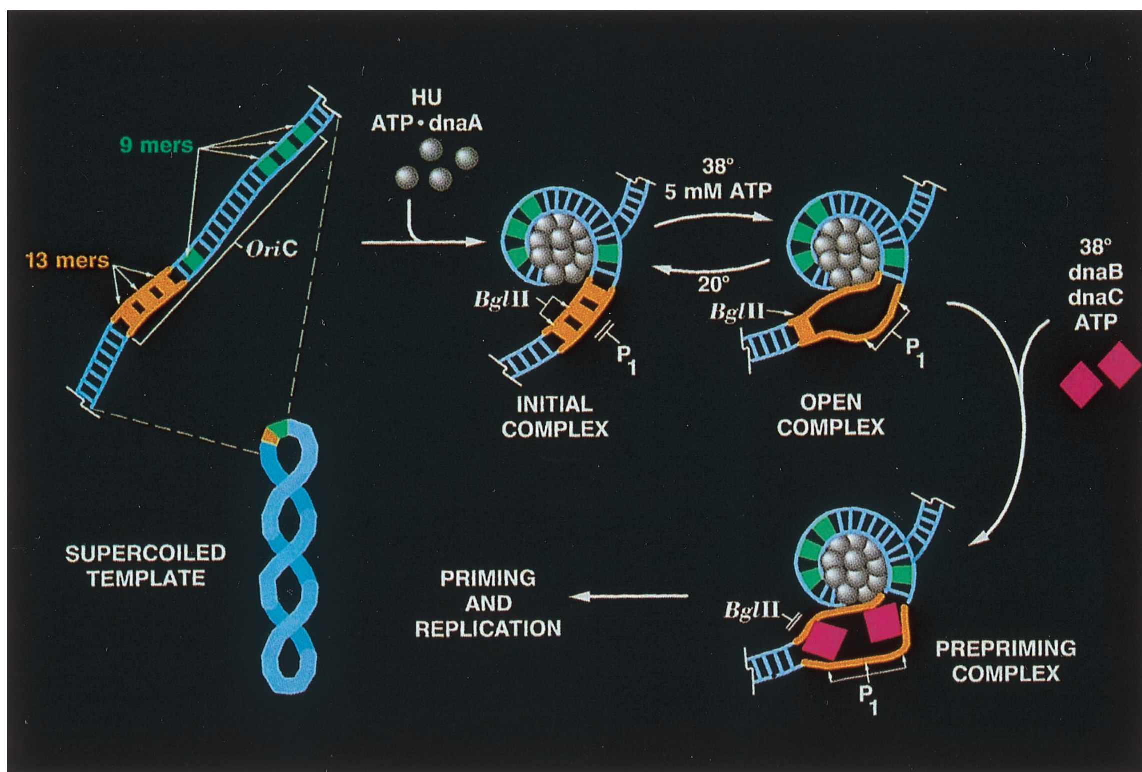

FIG. 3. Initiation of replication at the *E. coli* chromosomal origin. Reprinted from reference 1 with permission of the publisher.

ing the fifth commandment. We were using a pure DNA polymerase on a dirty DNA substrate: frayed, gapped, fragmented, denatured and heterogeneous. When we finally switched to the intact, single-stranded, circular DNA of a small bacteriophage, we discovered how a DNA chain is started: priming with RNA.

VI. DEPEND ON VIRUSES TO OPEN WINDOWS

Single-stranded phages provided not only the DNA substrate with which we could discover the RNA priming of new chains, but also the enzyme systems responsible for the priming and subsequent replication (Table 1). The filamentous phage M13 depends on the host RNA polymerase to make a short transcript of an origin region, whereas the icosahedral phage ϕX174 appropriates a complex primosome, used by the host to prime the start of chains on the lagging strand at the replication fork. Whereas the conversion of the single M13 viral strand to the duplex replicative form was readily resolved and reconstituted, the conversion of the ϕX174 single-stranded circle was far more complex and required the discovery of 15 new proteins, which constitute the apparatus at the host chromosomal replicating fork (Fig. 1 and 2). With these many proteins in hand we could attempt to discover how replication was initiated at the origin of the intact *E. coli* chromosome.

VII. CORRECT FOR EXTRACT DILUTION WITH MOLECULAR CROWDING

Cells are gels. Half of the cell dry weight is made up of proteins packed in highly organized communities. That some of their functions, individually and collectively, can be observed despite great dilution (20-fold or more) is a fortunate break for biochemistry. But there is an absolute need in some cases to restore the crowded molecular state, as we learned from our attempts to observe initiation of replication at the origin of an intact chromosome.

We were given a 5-kb plasmid containing the origin of the 4,000-kb *E. coli* chromosome that is replicated in the cell with the physiological and genetic features of the host chromosome, in effect a minichromosome. When Seichi Yasuda came from Japan with this *oriC* plasmid, I thought we would soon resolve and reconstitute its replication much as we had done with phage ϕX174. But it took 10 man-years of utter frustration before we finally succeeded in making a cell-free system work.

Success in achieving *oriC* plasmid replication in a cell-free state depended on two strange maneuvers. One was to include a high concentration of polyethylene glycol (PEG) (10% [wt/vol]), 10,000 Da) in the incubation mixture. As is true of such hydrophilic polymers, the PEG gel occupies most of the aqueous volume and excludes a small volume into which large molecules are crowded. This concentration is essential when several proteins are needed in the consecutive steps of a pathway. The other maneuver repeated an earlier experience in which we proceeded to fractionate an inactive lysate with ammonium sulfate. Progressive additions of the salt yielded precipitates, in one of which the active proteins were present and concentrated when dissolved in a small volume. Just as important, the supernatant fraction we discarded contained a potent inhibitor, a nuclease that relaxed the plasmid DNA from its essential supercoiled state.

Along with a purified, origin-binding DnaA protein, we provided primosomal, replication, and ligase proteins with other factors to obtain rapid, origin-specific, extensive replication of the *oriC* minichromosome (Fig. 3). With these many proteins added in sufficient amounts, PEG was no longer needed. The mechanisms we discovered in *oriC* replication were found to

FIG. 4. Ten commandments. Lessons from the enzymology of DNA replication.

apply to the replication of many microbial plasmids and phages and to some eukaryotic viruses and episomes.

VIII. RESPECT THE PERSONALITY OF DNA

For years, DNA was regarded as a rigid rod devoid of personality and plasticity. Only upon heating did DNA change shape, melting into a random coil of its single strands. Then we came to realize that the shape of DNA is dynamic in ways essential for its multiple functions. Chromosome organization, replication, transcription, recombination, and repair have revealed that DNA can bend, twist, and writhe, can be knotted, catenated, and supercoiled (positive and negative), can be in A, B, and Z helical forms, and can breathe.

Especially noteworthy is breathing, the transient thermodynamic-driven opening (melting) of the duplex that facilitates the binding of specific proteins such as the helicase responsible for priming and the onset of replication. Certain DNA sequences are also predisposed to a more extensive form of melting ("heavy breathing") that creates a relatively large opening for transcription. The resulting RNA-DNA duplex (R-loop) can activate an inert origin of replication by altering its structure, even hundreds of base pairs away, which facilitates its opening by origin-binding and replication proteins. Negative supercoiling supplies the energy for the breathing and other features that direct the shape and movements of DNA at the *oriC* origin of replication. These DNA responses have led to an appreciation of the role of transcriptional activation of replication origins near primers in large chromosomes, both prokaryotic and eukaryotic.

IX. USE REVERSE GENETICS AND GENOMICS

Direct genetics, in which a randomly mutated gene can ultimately be linked to a deficiency in a single enzyme, was a landmark discovery in biologic science. This approach served well by providing *E. coli* mutants defective in replication, some in initiation of a chromosome (e.g., *dnaA*) and others in elongation (e.g., *dnaB*, *dnaC*, *dnaE*, and *dnaG*). But randomly generated mutants do not readily disclose the products of their genes nor their particular functions. Nevertheless, these replication mutants were crucial in validating our assays because DNA synthesis was absent in the extracts of mutant cells and restored when extracts or purified fractions from wild-type cells were added.

Reverse genetics and genomics have now made the enzymologic approach even more powerful. Unlike direct genetics, enzymology starts with a defined function, after which finding the responsible genes has become relatively easy. With even a picomole of a purified enzyme or a band on a gel, a peptide sequence can be determined and the encoding gene identified, cloned, and overexpressed; genomics facilitates the process by providing the complete genome sequences of *E. coli*, yeast, and many other microbes. Profound insights into the physiologic role of an enzyme or pathway emerge from the behavior of cells with a null, point, or truncated mutation of a gene or modulated levels of its overexpression.

The ease with which large quantities of pure enzymes can be produced by overexpression of a cloned gene has made their use as reagents even more attractive (commandment X). "DNA shuffling", a new technique, has made enzyme reagents compelling. By creating a very large number of random rearrangements of a gene or genome, a particular gene product can be selected for a desired property (e.g., heat resistance) with wide applications in industry and biomedical science.

X. EMPLOY ENZYMES AS UNIQUE REAGENTS

Biochemistry is replete with examples in which enzymes have been employed as analytic and preparative reagents. From basic research to industrial processes, proteases, amylases, phospholipases, kinases, and phosphatases, etc., have been crucial in operations that were beyond the capabilities of available chemical technology. I will mention just a few exam-

ples of applications to DNA and its replication and one from my recent research on inorganic polyphosphate (poly P).

The key discovery in 1944 that identified DNA as the genetic substance was based on the destruction by crystalline pancreatic DNase of the factor that transformed one strain of *Pneumococcus* sp. to another. It was the action of this DNase again, as mentioned in commandment IV, which in 1955 made me believe that the few counts of [^{14}C]thymidine incorporated by an *E. coli* extract into an acid-insoluble form signaled the synthesis of DNA. Many more examples can be cited in which an enzyme reagent was decisive: the circularization of linear DNAs by ligases, the creation of "sticky" tails by specific exonucleases used to prepare the first recombinant DNAs, the innumerable uses of restriction nucleases, and on and on.

Enzyme reagents have been decisive in my approach to determine the functions of poly P, an inorganic polymer of hundreds of phosphate residues linked by "high-energy" anhydride bonds. Likely present on prebiotic earth, poly P is now found in every living cell, but for lack of any known functions, was earlier regarded as a "molecular fossil." True to the first commandment, I have sought and isolated enzymes that make and act upon poly P. With these enzymes we developed assays that are definitive, facile, and sensitive in place of those that are ambiguous, laborious, and insensitive. Together with the use of reverse genetics, we have learned that many microbes need poly P to adapt to adverse conditions and to survive in the stationary phase. The kinase that makes poly P from ATP is highly conserved in some of the major pathogenic bacteria, and mutants lacking the kinase are defective in motility and virulence. Thus, this enzyme, absent from eukaryotes, may prove to be an attractive target for antimicrobial drugs.

In this Commentary, I have suggested commandments based largely on the classic pursuit of enzymology in the 20th century. Selected for relevance to DNA replication (Fig. 4), these commandments surely need many amendments for exploration of other key events in the physiology of cells and organisms. Recall the Mel Brooks movie *History of the World, Part I*, in which he portrays Moses descending from Mt. Sinai carrying three tablets with the Lord's commandments. He stumbles, dropping one tablet which shatters. He sighs: "Oh well, 10 commandments are enough!" Not really! Among the commandments on the lost tablet was the most important one of all: "Thou shalt respect basic research."

ACKNOWLEDGMENTS

In the lessons I have learned from the enzymology of DNA replication, I depended at every turn on colleagues near and far for orientation and guidance. Most of all, I learned from the efforts and contributions of my students, too numerous to be mentioned individually. Without them there would be no story for me to tell. I am deeply indebted to them and as we all are to the National Institutes of Health, without which none of us could have pursued our dreams.

SUGGESTED READING

Goulian, M., A. Kornberg, and R. L. Sinsheimer. 1967. Enzymatic synthesis of DNA XXIV. Synthesis of infectious phage ϕX174. Proc. Natl. Acad. Sci. USA **58:**2321–2328.

Kornberg, A. 1989. For the love of enzymes. Harvard University Press, Cambridge, Mass.

Kornberg, A. 1997. Centenary of the birth of modern biochemistry. FASEB J. **11:**1209–1214; Trends Biochem. Sci. **22:**281–283.

Kornberg, A., and T. A. Baker. 1992. DNA replication, 2nd ed. W. H. Freeman & Co., New York, N.Y.

Kornberg, A., N. N. Rao, and D. Ault-Riché. 1999. Inorganic polyphosphate: a molecule of many functions. Annu. Rev. Biochem. **68:**89–125.

Lehman, I. R., and P. E. Boehmer. 1999. Replication of herpes simplex virus DNA. J. Biol. Chem. **274:**28059–28062.

Racker, E. 1985. Reconstitutions of transporters, receptors and pathologic states. Academic Press, New York, N.Y.

Steitz, T. A. 1999. DNA polymerases: structural diversity and common mechanisms. J. Biol. Chem. **274:**17395–17398.

Waga, S., and B. Stillman. 1998. The DNA replication fork in eukaryotic cells. Annu. Rev. Biochem. **67:**721–751.

Zimmerman, S. B., and B. H. Phieffer. 1983. Macromolecular crowding allows blunt-end ligation by DNA ligases from rat liver or *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80:**5852–5856.

Also recommended is the prefatory autobiographical chapter in each volume of the Annu. Rev. Biochem., including that in vol. 68, 1999 (p. 1–32).

REFERENCE

1. **Kornberg, A., and T. A. Baker.** 1992. DNA replication, 2nd ed. W. H. Freeman & Co., New York, N.Y.

The views expressed in this Commentary do not necessarily reflect the views of the journal or of ASM.