Chapters 20 & 21. Weaver, 2/e. Mol Biol X107A.

Ch 20. DNA replication. I. Basic mechanism and enzymology.
Ch 21. DNA replication. II. Detailed mechanism.

Reading note. In general, emphasize the basic material on bacterial replication. For the most part, ignore odd viral systems, and only briefly note the eukaryotic systems.

Knowledge of the Greek alphabet will be very helpful in these Ch. ☺

Clark & Russell: Ch 5. Ch 8 includes plasmids. Fig 11.6 describes telomerase.

A. Introduction

These two chapters are intertwined. It makes for a long "super-chapter" to cover them together, but I think it is hard to really separate them. I suggest that you start by browsing both

chapters, and making an outline of them. Read the introductions and summaries for the two chapters. These should help you establish a framework; then you can read individual topics and fit them in to the big picture. As usual, try to see what's "merely" structure and what's function. There are areas where we have lots of data about a protein (for example), but don't really know what it is for (biologically). And then there are functions we know must occur, but we don't see how. Etc.

The overall story of DNA replication is already clear (Ch 2-3). Each of the two strands of the parental DNA double helix serves as a template for a new complement by the base pairing rules (Fig 2.15). The result is two daughter helices, each of which has one new strand and one old strand; semiconservative replication (Fig 3.21).

Nevertheless, many questions remain:

- The enzymatic machinery.
- The energy problem.
- The polarity problem.
- The initiation problem.
- The termination problem.
- Ends problems.
- The accuracy problem.
- The unwinding problem.
- The processivity problem.
- The regulation problem.
- The segregation problem.

In addition to addressing those "problems", we will look at…

- The concept of the replicon.
- The major players in replication: polymerases, clamps and clamp loaders, nucleases, helicases, primases, SSB, topoisomerases.
- The general strategy… recognition of origin, priming, initiation, semidiscontinuous synthesis.

- How are all the steps of replication coordinated?

- DNA replication: the big picture, the cell cycle.

At the heart, DNA replication is very similar in all organisms. Of course, the integration of replication into the cell cycle varies with the type of organism. We will briefly compare aspects of DNA replication in bacteria and eukaryotes.

B. Overview

We noted the general nature of DNA replication in Ch 2. We generally recognized that DNA replicates by strand separation with each parental strand then serving as a template for a new strand.

Weaver has already noted that DNA replication is semi-conservative (Fig 3.21, which he repeats as Fig 20.1). He now shows the classical experiment that proved this, the Meselson-Stahl experiment; the actual results are shown in Fig 20.4. The method used is still of interest. It measures the density of molecules, by finding their buoyant position in a density gradient. In this case, isotopes of nitrogen were used to "density label" the DNA; a solution of cesium chloride, spinning in a centrifuge, was used to create the density gradient.

Some terms: replication fork (p 649); uni- vs bi-directional replication (p 649); θ (theta) structure (p 649); replicon (the idea; p 652). The term "bubble" (Fig 20.11) or "eye" structure is sometimes used for a θ structure that has been linearized.

C. The energy problem; DNA polymerase and activated nucleotides

The basic biochemistry of making DNA is the same as for making RNA, Ch 6 and Fig 3.13.

The only differences are
- use of the sugar 2′-deoxyribose (instead of ribose);
- use of thymine (= 5-methyluracil, instead of uracil).

Neither has any logical implication for the basic biochemistry.

Polymerizing nucleotides is thermodynamically unlikely. The cell's solution to this problem is to feed energy into the reaction, by using activated nucleotides: nucleoside triphosphates.

An enzyme, DNA polymerase, adds nucleotides to DNA chains:

$DNA_n + dNTP \rightarrow DNA_{n+1} + PP_i$

Compare this with the analogous equation for RNA synthesis, Ch 6 handout Sect L. dNTP is a general notation for any deoxyNucleotide TriPhosphate.

The reaction as shown here liberates PP_i. In practice, the PP_i is then hydrolyzed, to 2 P_i; this PP_i hydrolysis provides the actual driving force for the

reaction. (This was first mentioned in Ch 2 handout Sect F, and discussed further in the context of RNA synthesis, Ch 6 handout Sect L.)

This brief introduction to the basic biochemistry of making DNA raises several issues, which are addressed in the following sections. In fact, Sect D & E deal with problems inherent in the above description of DNA polymerase. We will also return to the nature (and diversity) of DNA polymerases in Sections K, P & Q; the reaction discussed here is basic for all.

See Kornberg (2000) for some insight into the discovery of DNA Pol and the development of its story.

D. The polarity problem; leading and lagging strands

The two strands of DNA in a double helix have opposite polarities. One runs 5′→3′, its complement runs 3′→5′; they are anti-parallel. (Fig 2.14.)

DNA polymerase -- all known DNA polymerases -- can make DNA only 5′→3′.

Thus there is a problem replicating both strands of DS DNA at the same time, since they run in opposite chemical directions.

The clue as to how the cell deals with this polarity problem came from studies where replicating DNA is labeled for very short times -- a pulse (Fig 20.7). Much of the pulse label is found in very short pieces of DNA. Later, after a chase, this same label is in large DNA, the size expected for the DNA being replicated.

From this work evolved the following model… One strand of the DNA replicates as expected, 5′→3′ along the replicating chain. This is called the leading strand. However, the other strand, the lagging strand, replicates "discontinuously". After the leading strand has replicated for a while, there is a stretch of unreplicated DNA on the other strand. At some point, the initiation mechanism initiates on this strand, and lagging strand synthesis begins. This, too, is 5′→3′, the only way DNA polymerase can go, but it is going in the opposite direction (along the helix axis) to the overall replication direction. Eventually, this "wrong way" piece runs into a region that has already replicated. The polymerase falls off, and the new fragment is tied in to the older material (using the enzyme DNA ligase).

Weaver introduces DNA ligase on p 62 (including Fig 4.2), in the context of joining together restriction fragments during "cloning". Nakatani et al (2000) describe an unusual ligase.

The short, wrong way fragments on the lagging strand are called Okazaki fragments.

Fig 20.6b diagrams this scheme of semidiscontinuous replication.

We now understand, generally, that the replication complex, called a replisome, is a dimer. One head of it replicates the leading strand, continuously and highly processively (= without falling off; Sect J); the other head replicates the lagging strand, discontinuously. [If the

lagging strand is folded back, the polymerase is moving in the same physical direction on both strands.] We will elaborate on this model in Sect K.

> Since chain growth requires a 3′-OH, nucleotides without a 3′-OH can be incorporated, but cannot be further extended. The efficiency of incorporation varies among Pols (Hiratsuka & Reha-Krantz, 2000). The use of nucleotides with a blocked 3′ end is central to the Sanger (enzymatic) method for DNA sequencing (Fig 5.18). Some nucleotide analogs used as drugs, such as AZT, work on this principle (Meyer et al, 1999).

E. The initiation problem; RNA primers

DNA polymerase -- all known DNA polymerases -- cannot start a new chain; it can only extend an existing chain. (All? Chiang & Lambowitz, 1997, discuss an exception.)

RNA polymerase, which makes RNA chains (Ch 6), works much like DNA polymerase, except that it polymerizes ribonucleotides rather than deoxyribonucleotides. RNA polymerase can initiate new chains, whereas DNA polymerase cannot.

The mechanical answer to the problem of initiating a DNA chain is (usually) that a short RNA chain -- a primer -- is made at the initiation site. The DNA polymerase then extends the RNA primer. (Other enzymes come back later and clean up the starting sequence; Sect G.)

Weaver introduces priming in Fig 20.8, and expands on the process in Ch 21 Sect 2. In that section, emphasize the E. coli system, and briefly note the eukaryotic system.

> The primosome (p 686) is the complex of proteins that recognizes DNA initiation sites and makes the RNA primer, which effectively initiates DNA synthesis. The specialized RNA polymerase that makes the primer is called a primase. Keck et al (2000) explore the structure of the primase; they also discuss the transition from primase to polymerase (mentioned in Sect L, below).

> There are other mechanisms for initiation, not involving RNA primers. I will briefly mention an example, but it is not important to deal with these. The main lesson is that, one way or another, the cell gets around the inability of DNA polymerase itself to initiate.

More about the details of initiation later, in the context of regulation, Sect M.

F. The termination problem

(Ch 21 Sect 4)

DNA polymerase does not know when to stop.

Some replicons (but not all), contain special sites and special proteins for termination of replication. Some information about the mechanism of this termination system is developing.

Weaver shows the bacterial termination system in Fig 21.27. Note the orientation of the Ter sites. The two replication forks will run into each other before they encounter a Ter site. This suggests that normal termination does not rely on Ter sites and Tus. That suggestion is reinforced by the finding that deletion of the Ter/Tus system has no detectable phenotype.

What then is the significance of the system? It is possible that it serves as a backup system -- that it is active only when one fork progresses more rapidly than the other, so that fork collision would occur outside the "preferred region". In this case, the Ter sites do slow down the speeding fork, and help to trap the termination event into the desired region.

> (There is evidence that the Ter system prevents over-replication of plasmids, and this may be of value. Or it may be a clue.)

The nature of the Tus protein is now known. It is a contrahelicase, which inhibits the replication helicase (Sect I).

See Griffiths & Wake (2000) for recent work. Gerber et al (1997) discuss a barrier to a eukaryotic replication fork.

We will discuss decatenation (disentangling) of the newly replicated chromosomes in Sect O.

We will discuss the special problem of filling in the primer gap at the ends of linear replicons in the next section.

G. Ends problems

DNA polymerase has problems with ends. The problems are a result of its inability to do anything but extend chains 5′→3′. There are two, somewhat distinct, ends problems. One is where each Okazaki fragment begins with a short segment of RNA (Sect E). The other is at the end of a linear chromosome.

In the first case, the solution is fairly straightforward. The RNA primer is removed, the gap is filled in with deoxynucleotides (by extension from the previous DNA) and sealed; Fig 20.29. This is an example of a DNA repair process (more in Sect H). Bae et al (2001) explore the greater complexity of this process in eukaryotes.

For chromosome ends the problem is more serious (Fig 21.31). Logically, it is impossible to lay down the first nucleotide by extension. As a corollary, it is impossible to replace a primer at the 5′ end.

Eukaryotic chromosomes have a special structure at the ends of linear chromosomes, called a telomere. We now understand that (in most eukaryotes) the chromosomal telomeres are maintained by a special enzyme, called telomerase. The basic logic of telomerase is shown in Fig 21.31c, with more detail in Fig 21.33.

Two things make this story particularly interesting or important.

- The telomerase is an RNA-dependent DNA polymerase -- a reverse transcriptase. What makes it unusual is that the template is part of the enzyme, Fig 21.33.

  Peng et al (2001) explore telomerase as an example of a reverse transcriptase. Baumann & Cech (2001) reveal one more protein in the telomere story.

- Telomerase has been implicated in human aging and cancer. In humans, most somatic cells lack telomerase, but most cancer cells have the enzyme. These are fascinating and unresolved stories, with complications beyond these simple statements.

  Weaver introduces the story of telomerase and cell mortality, with possible implications for cancer, in the Box on p 712. A caution in interpreting the work with TEL-deficient mice is that the dynamics of TEL shortening are quite different in mice and humans. The TEL-deficient mice he discusses do have high cancer rates (e.g., Artandi et al, 2000). Hackett et al (2001) show that TEL-deficient yeast show high mutation rates.

  Hahn et al (1999) construct tumor cells, with alleviation of TEL-deficiency being one part of the story. Mitchell et al (1999) implicate a telomerase deficiency in a human genetic disease; Rudolph et al (2000) implicate TEL shortening in cirrhosis of the liver in a mouse model. Shay & Wright (2001) emphasize that telomere shortening is not the entire explanation for cellular senescence in the lab. See Sect S for Internet resources on telomeres.

Many (most?) bacteria avoid the problem of chromosome ends by having circular chromosomes.

Chromosomes break from time to time. Breaks can be healed by a rather sloppy but adequate procedure called non-homologous end joining (NHEJ). How are normal ends distinguished from broken ends? Weaver notes the special loop structure of normal ends, p 711. Walker et al (2001) discuss a key protein involved in NHEJ.

Rizki & Lundblad (2001) discuss non-telomerase systems for maintaining the telomeres, and reveal a role for the mismatch repair system (next Section).

Kirkwood & Austad (2000) introduce a series of articles on the broad topic of aging.

H. The accuracy problem; DNA repair

The choice of which base to add is made by the base pairing rules, which depend on small free energy differences. Sometimes the wrong base is inserted -- about once per 10^5 bases (p 667).

So what? After all, changes in the DNA (mutations) are the source of the variation that is fundamental to natural selection. However, the frequency of mistakes is too high. So, the DNA polymerase proofreads its work. The 3′→5′ exonuclease activity that is part of DNA polymerase chews off (most) incorrect insertions just after they are made. Fig 20.28 shows the basic idea.

> All DNA polymerases involved in ordinary replication have a proofreading activity somehow associated with them. Note that E coli Pol I has the proofreading activity as part of the same peptide chain, whereas Pol III has a separate subunit for proofreading.
>
> (Eukaryotic Pol α may be an exception to the rule. It seems to not have ordinary proofreading. On the other hand, DNA made by Pol α may be removed later; Bae et al, 2001.)

The simple view, that the proofreading enzyme removes incorrect bases, suggests that the proofreading enzyme can discriminate incorrect from correct insertions. However, for the E. coli Pol I that is not the case; its proofreading activity apparently has very poor discrimination. How, then, does it achieve effective error removal? The main factor is that the polymerase activity becomes extremely slow at adding the next nucleotide when the previous one is wrong; this allows the proofreading enzyme more time to act on an incorrect nucleotide. Thus there is a kinetic competition between the polymerase and proofreading activities, and it is the polymerase activity that varies when there is an error.

Sometimes the proofreading enzyme removes a correct base. The removed base is replaced by the polymerase, but there is an energy cost. One can imagine a more effective proofreading enzyme, one which competes better with the polymerase. It would presumably result in a lower overall error frequency, but it would also remove more correct bases. The energy cost of these additional unnecessary "corrections" may place an evolutionary constraint on the effectiveness of proofreading.

> This argument implies that error rates, hence mutation rates, themselves are subject to natural selection. Among other things, this opens the possibility that mutation rates vary with time (conditions) for a particular organism. Metzgar & Wills (2000) discuss the idea, and Oliver et al (2000) provide an example showing hypermutability. Also recall Wright (2000; Ch 2 FR).
>
> Recall discussion of kinetic competition in proofreading the choice of aminoacyl-tRNA in protein synthesis, Ch 19 handout Sect G. Yan et al (1999) discuss an example of kinetic proofreading relating to how topoisomerases disentangle chromosomes; more in Sect O.

"Proofreading" is only the first line of defense against DNA errors. The second is "mismatch repair", p 675. This monitors newly synthesized DNA for mismatches, and fixes the "new" strand. How does it know which strand is new? In E. coli it can tell because methylation of the new strand is delayed; thus the new strand appears undermethylated; Fig 20.44. This basic system seems to operate in all organisms. Mutations in the mismatch repair system have been shown to increase the chances of some kinds of cancer (p 675). A similar phenomenon may explain the colonization of cystic fibrosis patients with Pseudomonas (Oliver et al, 2000).

> The role of methylation in mutation prevention (or in anything else) cannot be universal. Some organisms, including yeast and Drosophila, have little or no DNA methylation. One possibility in these cases… the system may recognize SS breaks, left from Okazaki fragments. Mammals, however, do have methylated DNA.

The structure of DNA polymerase seems designed to enforce proper base pairing; Steitz (1998). This works not so much by recognizing any specific matching features, but rather by rejecting any base "pair" that does not fit into the active site correctly.

Proofreading and mismatch repair are only two of many mechanisms that help ensure the integrity of DNA. DNA errors (damage) are introduced not only by replication, but also by a variety of physical and chemical damage agents, as well as by the inherent chemical instability of DNA. Many systems monitor DNA for one or another kind of damage, and attempt to repair it. Weaver introduces this topic in Sect 20.3; I suspect that we will not have time for this.

Most of the "new" DNA polymerases introduced in Sect Q, below, have specialized roles in DNA repair.

The FR contains a sampling of papers on DNA repair, with some emphasis on papers discussing a human disease. These include:

Lindahl & Wood (1999) broadly review DNA repair.

Meyer et al (1999) describe an unusual "error correction" mechanism in reverse transcriptase from AZT-resistant HIV.

Kren et al (1999) exploit the DNA repair system to treat a genetic disease.

Nilsen et al (2000) discuss the removal of uracil from DNA. The major source of U in DNA is spontaneous deamination of C, an event which would lead to mutation if not repaired.

Junop et al (2001) explore the role of ATP in a proofreading step during mismatch repair.

Macintyre et al (2001) explore the roles of methylation in mutation.

Hoeijmakers (2001) broadly reviews genome-cancer issues.

Sinden (2001) discuss how DNA repair may be relevant to triplet repeat expansions. Also recall Sakamoto et al (1999; Ch 2 FR).

Funchain et al (2001) discuss mutator strains.

Also recall Scully & Livingston (2000; Ch 10 FR), on the breast cancer protein BRCA1; and Le Page et al (2000; Ch 10 FR), on the involvement of transcription factor TFIIH in the DNA repair disease Cockayne syndrome. The latter paper deals with DNA damaged by oxidation and with transcription-coupled DNA repair.

I. The unwinding problem; helicase and SSB

DNA replication requires separating the parental strands. It is not obvious that any special proteins should be required to do this. After all, RNA Pol must also separate the parental strands and does not require any additional proteins for this.

However, DNA replication is typically associated with two additional types of protein whose job is to open the DNA -- and hold it open. These are called, generically, helicase and single-stranded binding protein (SSB).

Helicases open up the DNA, and they expend energy to do it. Fig 20.17 shows a lab assay for helicase; note the ATP.

Then, SSB (pp 655 ff) comes along and holds the DNA open, keeping it from renaturing or kinking. Fig 20.18 shows stimulation of apparent helicase activity, presumably by stabilizing the SS product.

Twelve DNA helicases are known in E coli (though Weaver notes only four, p 654). The need for so many helicases may be due to many of them forming specific complexes with other proteins, thus having specialized roles.

Helicases seem to come in two general types, those that function as dimers and those that function as hexamers.

The major E. coli replication helicase is the DnaB protein, an example of a hexameric helicase. Coupling of this helicase to the DNA Pol is a key part of assembly of the replisome (more in Sect K, below).

> The RuvB helicase, which is involved in recombination (e.g., Fig 22.25), is also an example of a hexameric helicase. SSBs are also involved in recombination, Ch 22.
>
> The computer resource listed in Sect S for Ch 22, showing animations of RuvB-mediated branch migration in recombination, may be useful here.

Hingorani & O'Donnell (2000b) include the hexameric helicases in their discussion of proteins that form rings around DNA. They suggest that the hexameric type of helicase may be preferred when the enzyme is to act over long distances, thus should stay on the DNA.

There also are RNA helicases. These include:

- The ρ protein that is involved in the termination of transcription (Fig 6.50); ρ also acts as a hexamer.

- The eukaryotic protein synthesis initiation factor eIF-4A, which unwinds mRNA during scanning for a start codon (p 552).

von Hippel & Delagouette (2001) broadly review how helicases work.

Cantor et al (2001) implicate a helicase in breast cancer.

J. The processivity problem

The degree to which polymerase continues along a chain without falling off is called processivity. (pp 668 & 694, but also recall Fig 8.21)

DNA polymerases vary widely in their processivity. Replication polymerases have high processivity.

The mechanism for maintaining high processivity is now clear for E. coli Pol III. It wears a safety belt. The β subunit of Pol III forms a "sliding clamp" that loosely fits around the template (Figs 21.14 & 15).

PCNA plays a similar role for eukaryotic DNA polymerases; p 668 and Fig 21.16.

See Sect S for pdb files of sliding clamps, which you can view and manipulate in RasMol. And see the homework for more.

More about these sliding clamps in the next section.

Hingorani & O'Donnell (two articles, both 2000) review sliding clamps, and discuss how they may allow various proteins to track DNA; they also generalize on the idea of proteins that form rings around DNA.

Zuccola et al (2000) discuss a DNA Pol processivity factor that does not form a belt around the DNA.

K. Structure and function of polymerases

E. coli Pol I (p 661) is a single chain enzyme, but the chain carries three activities (polymerase and two nuclease activities). The 3′→5′ exonuclease is the proofreading activity; Sect H. The (relatively uncommon) 5′→3′ nuclease, can replace RNA primers with DNA; Fig 20.29. This activity is popularly called "nick translation", and is useful in the lab for labeling DNA with radioactivity (p 662).

The replication polymerase, Pol III, has long been recognized as a much more complex enzyme. Clarification of the Pol III story was complicated by the fact that preparations with different subunit compositions were active in vitro. All subassemblies shown in Table 20.3 that include the α subunit are active as polymerases.

> Weaver introduces Pol III in Ch 20 (e.g., p 664), but completes the story of the Pol III complexity and its function in Ch 21, mainly in Sect 3.

With the development of the β story (previous section), we can now provide a reasonably complete functional description of Pol III. The subunit composition of the complete enzyme probably is $(\alpha\varepsilon\theta)_2\tau_2(\gamma_2\delta\delta'\chi\psi)(\beta_2)_2$. The α subunit is responsible for polymerization, ε is for proofreading; Taft-Benz & Schaaper (1999) discuss how these two subunits interact. β is for processivity (as noted above, Sect J).

τ is involved in dimer formation (Fig 21.18). It also couples the Pol and helicase; more about this in Sect L, below.

The γ complex ($\gamma_2\delta\delta'\chi\psi$) is the "clamp loader"; Fig 21.18. It promotes addition of β to the active core; this requires ATP hydrolysis (see Sect L). One can consider the γ complex as a "matchmaker", a protein whose job is to deliver another protein (β) to DNA. The cartoons at the top of Figs 21.23 & 24 show the idea. Fig 21.28 summarizes the clamp loading cycle, including unloading. Jeruzalmi et al (2001) update how the γ complex works.

> Interestingly, τ and γ are made from the same gene; synthesis of γ requires a frameshift (a "rule-break"; recall Ch 18 handout Sect M). Actually, there is some ambiguity about the requirement for γ; τ can substitute for it. Also, the number of γ subunits in the γ complex, listed above as 2, is not entirely clear; you will see numbers from 1-4. Walker et al (2000) review all these points in their introduction.

The role of θ remains unclear; Taft-Benz & Schaaper (1999).

Fig 21.18 summarizes a reasonable view of the organization of Pol III (actually, it is Pol III*, lacking β; recall Table 20.3).

Steitz (1998) discusses recent information on DNA polymerase structures. Clover & McHenry (2001) explore the basis for the asymmetry of the Pol III complex.

L. Coordination of replication functions

At this point we have generally established what the functions are in replication, and which proteins carry them out. A striking feature is how complex the process is; the complexity calls for a careful coordination of the functions. For example, think about what happens during each Okazaki fragment cycle -- and in the homework you will calculate that this entire cycle takes only 1-2 seconds. One direction, then, of current work is to try to figure out what coordinates all the steps. Some ideas are summarized here.

Keck et al (2000) discuss how the χ subunit of the γ complex mediates the transition from primase to polymerase. Yuzhakov et al (1999) show a similar transition in eukaryotic replication.

Bertram et al (2000) show that ATP binding and hydrolysis control the clamp loader cycle. Binding of ATP to the γ complex allows it to bind a β_2 ring, and to open up the ring so that it can be fitted around the DNA. ATP hydrolysis causes the γ complex to release its cargo. Note the use of an NTP-NDP cycle to drive a cycle of protein conformational changes, as introduced in Ch 17-18.

Levine & Marians (1998) suggest that the γ subunit may also interact with the topoisomerase that decatenates the daughter duplexes (Sect O).

As noted above (Sect K), τ brings the two halves of the replication complex together, and also is involved in coupling the Pol and helicase. See “Fig 7”, attached p 38, and Yuzhakov et al (1996).

Katayama et al (1998) show that the β subunit is also involved in regulating initiation, Sect M.

M. Initiation of a round of replication

We know that cellular DNA synthesis is regulated; it occurs exactly once per generation. Further, in many cases (including for most cellular chromosomes) it begins at a fairly precise time in the cell cycle. In this Section and the following Section we will try to develop some ideas about this regulation. The first issue -- the topic of this Section -- is to describe the initiation event in more detail. After all, the key step that is regulated is initiation.

The E. coli origin of replication, *oriC*, has been cloned. It can be put into plasmids, which now replicate like the chromosome (Fig 21.3). These *oriC* plasmids are called minichromosomes. Minichromosomes are a good system for studying the initiation of bacterial DNA synthesis. They have the experimental conveniences of a plasmid, but the relevance of “the real thing.”

The DnaA protein is a key player in the initiation step. The DnaA protein binds to *oriC* sites, then promotes opening of the DNA. With the DNA opened by the DnaA protein, the DnaB helicase can enter into the putative replication fork, as a prelude to the priming step discussed earlier. DnaB entry requires DnaC; DnaC is another “matchmaker” (delivery protein; recall

discussion of the γ complex in the previous Sections). These steps are collectively termed prepriming. Fig 21.5 describes this.

Some of the in vitro work leading to this model was done with supercoiled CCC *oriC* minichromosomes (Fig 21.3). With these supercoiled templates, DnaA protein alone can open the helix.

> (This can be shown because the protein-promoted opening exposes the DNA to nucleases that are specific for SS DNA. The resulting cleavages are detected by gel electrophoretic analysis.)

With non-supercoiled templates, DnaA protein (alone) cannot open the helix. However, if the template is being transcribed at or near *oriC*, then DnaA protein can provide the specific opening that is needed. This is presumably what happens in vivo, and explains the requirement for RNA polymerase for initiation of a round of replication (p 689).

Fang et al (1999) have studied the interaction between helicase action and actual initiation by priming. They show that priming sites per se are not well defined, because the primase is "chasing" the moving helicase. Their "Fig 7" is attached, p 39. (This work reminds us that the "origin" as defined by where synthesis begins may not correspond to the critical recognition site for the event.)

N. The regulation problem; how to start a round of replication

[not in book]

The story above gives us a rather complete picture of the initiation process at the DNA level. However, we still want to know the key events in regulation. How is the action of the DnaA protein regulated? There are some models, but we really don't know. Weaver does not discuss this issue; if time permits, I may discuss it some.

It may be useful to dissect what roles we expect for an origin. We can recognize three, logically distinct, roles for *ori*:

1. determining the timing of initiation;
2. determining the site of initiation;
3. opening the DNA to allow polymerase (etc.) to get onto the DNA.

Briefly, we might call these when, where and how.

From the preceding Section, we seem to have a reasonable understanding of #2 & 3. DnaA protein recognizes the site of initiation, and the prepriming model gives a reasonable description of the opening process. Further, these two steps are logically similar for other replication systems.

What's left? #1, determining the timing of replication. We can divide this into two parts:

1.a. determining when a round of replication is initiated;
1.b. restricting that event to only one new round at a time.

Timing of initiation (1.a.)

The common view is that DnaA controls the timing of initiation. A simple model might be that the level of DnaA protein controls initiation. However, measurements of DnaA level do not show any variation that would explain initiation. Therefore, the discussion moves to the level of "active" DnaA protein. It now seems clear that "active" DnaA protein is the form bound to ATP. How the level of active DnaA protein -- ATP-bound -- is regulated is being explored. Further, it seems that some other proteins are involved. Lipids are also involved, and the key events may occur at the membrane (Newman & Crooke, 2000). The presence of both positive and negative control proteins would allow for much sharper switches, such as that normally observed for the timing of *oriC* initiation.

Preventing multiple initiations (1.b.)

Two components of this now seem evident. One serves to make the initiation site unavailable, and the other serves to make the initiator protein unavailable.

Role of methylation and sequestration

In E. coli a DNA methylation system prevents reinitiation. The key enzyme is DNA adenine methyltransferase ("dam methylase"), which methylates the A in the sequence GATC. (Fig 20.44; this is the same enzyme involved in marking strands for mismatch repair, a distinct function, Sect H.) The target sequence is a palindrome, so there is a methylation site on each strand. *oriC* contains several target sites for the dam methylase.

> [Also see Heithoff et al (1999) for a medically relevant role of Dam in gene expression, and Reisenauer et al (1999) for a broader discussion of DNA methylation.]

Two results are clear.

1. DNA with a hemimethylated *oriC* cannot replicate -- in vivo. (Hemimethylated means that the sequence is methylated on one strand.) Fully methylated or nonmethylated *oriC* origins do replicate.
2. In *dam* mutants, the usual precise timing of initiation is lost (see hw).

When fully methylated DNA replicates, it becomes hemimethylated -- transiently, until the methylase acts on the new strand. Methylation of the origin region is slow (compared to methylation of bulk DNA). This seems to relate to the binding of the origin to the membrane -- oddly, the outer membrane. The SeqA protein and the HobH protein are involved in this membrane binding, although the details remain unclear. Regardless of these details, this "sequestration" of the new origin serves to prevent unwanted multiple initiations at the time of

an initiation event. Sequestration also prevents (delays) methylation of the origin. It now seems likely that the relationship between hemimethylation and initiation failure is that hemimethylation causes sequestration, which then causes initiation failure.

Role of ATP

One more piece of the puzzle has recently been uncovered. One aspect of DnaA function is that only DnaA with ATP bound is active. Most DnaA has ADP bound. The significance of this intriguing result has long been unclear. But Katayama et al (1998) now show that the Pol III β subunit -- when assembled on the DNA as a sliding clamp -- stimulates hydrolysis of ATP bound to DnaA. (It is quite likely that it does this by stimulating an intrinsic ATPase activity of the DnaA protein.) Another protein, called IdaB and not yet characterized, is also required. This ATP hydrolysis is further stimulated by DNA replication. It is clear that this action couples the initiation event (formation of a sliding clamp and then actual replication) to prevention of further initiations.

> (Newman & Crooke, 2000, discuss how to get back to the ATP form, for the next round, but without a clear picture of what the key control is.)

Katayama et al argue that the effect of the sliding clamp in inactivating the initiator protein along with the *oriC* sequestration discussed earlier combine to prevent re-initiations. How proper initiation occurs at the proper time remains an open question.

del Solar et al (1998) review the variety of replication systems found in plasmids.

O. The segregation problem

The DNA replicates, and the two copies go to the two daughter cells. In eukaryotic cells the mitotic process (p 6; Ch 1 handout Sect I) ensures this precise segregation of daughter chromosomes. The same result occurs in bacteria, but it is less clear how. However, we can clarify what the question is. For example, we can distinguish distinct steps:

1. decatenation of the daughter chromosomes;
2. resolution of multimers;
3. segregation of daughter chromosomes into separate nucleoids;
4. segregation of nucleoids, one per daughter cell.

Weaver discusses only #1, but we will note all of them briefly.

1. Decatenation

For any DNA to replicate, the parental strands must be separated. If the DNA is closed circular, then bonds must be broken to allow this to occur. Topoisomerases (pp 658 & 706;

also recall Ch 2 handout Sect H and Ch 6 handout Sect M-N) are involved in separating the daughter chromosomes.

The major decatenation enzyme in E. coli is Topo IV, a type II enzyme (which makes DS breaks). Topo IV mutants make large nucleoids. Levine & Marians (1998) suggest that Topo IV may interact with the replisome, through the γ subunit.

Topoisomerases are also likely to be involved in dealing with the stresses of winding and unwinding near the replication fork.

Pouliot et al (1999) discuss an enzyme that repairs aborted topoisomerase events.

Strick et al (2000) observe the action of single topoisomerase molecules.

Li et al (2001) discuss how a type I topoisomerase works.

2. Resolution of multimers

There is another way in which daughter chromosomes may be combined: multiple chromosomes may recombine with each other. A single crossover between two circles leads to one large circle, which contains two genomes worth of DNA (a "dimer").

E. coli contains a site-specific recombination system whose "purpose" seems to be resolving multimeric chromosomes. In this system, the "Xer" enzyme acts at "*dif*" sites. Interestingly, the recombination site is near the terminus region (Sect F). Many plasmids have similar systems. Sciochetti et al (2001) explore the similar system in Bacillus subtilis.

> Weaver introduces site-specific recombination in Ch 23 Sect 1, but does not include dimer resolution systems.

An interesting question arises for both decatenation and multimer resolution systems. Such enzyme systems should catalyze both directions of the equilibrium reaction between monomers and dimers. Why, then, do they appear to specifically resolve dimers into monomers? This is not clear. One possibility is that they really do carry out both reactions, but that chromosome separation occurs when the chromosomes are in the monomer form, thus driving the reaction in the "desired" direction. On the other hand, it is possible that assembly of the apparatus for the events is efficient only if the sites are on the same DNA molecule, thus directing the reaction towards resolving multimers. Yan et al (1999) discuss this issue for decatenation.

3 & 4. Segregation

The two steps above deal merely with making sure that we have the number of copies we thought we have. Now the issue is actually getting those copies to two daughter cells. We can recognize, in principle, two stages here: making two daughter nucleoids (if we are talking

about chromosomes); getting the nucleoids into separate cells. We actually have very little information to distinguish these two steps, but should keep them in mind.

We can make a distinction between “active” and “helper” segregation mechanisms. Start by assuming that segregation might be random. (For high copy number plasmids this might be sufficient to achieve good survival of the plasmid.) However, recombination to produce multimers results in “less than random” segregation. The multimer resolution system, discussed above, helps to restore random segregation (by restoring the number of copies we thought we had); such a system is called a “helper” system. On the other hand, there are “active” systems -- which help to achieve better than random segregation. These are obviously important when the copy number is very low. Classical mitosis is an example of an active system. Bacterial chromosome segregation presumably must also involve active segregation, to get two cells each with one chromosome. In studying plasmid systems, we must be careful to distinguish which class of segregation system we are studying.

In some plasmids there are DNA sequences that promote accurate partitioning of the plasmid copies to the daughter cells. These are recognized by mutations that cause poor segregation, and are sometimes called *par* sequences (for par titioning). Addition of a *par* sequence to a plasmid that is poorly segregated leads to good, controlled segregation. The *par* region contains gene(s) for 1 or 2 proteins and a (cis-acting) site at which the protein acts. However, the function of the protein and the site are not known. It seems reasonable that the system involves membrane attachment.

> Weaver mentions some *par* genes on p 706. The genes he mentions turn out to be Topo genes, involved in the earlier steps discussed above. In earlier work, it was not easy to tell whether segregation-defective mutants were defective at decatenation or segregation per se.

Analogous proteins have been found in bacteria, and recently the site at which some act has been found. Lemon & Grossman (2001) review bacterial chromosome partitioning.

Bear in mind that studying segregation of the bacterial chromosome is more difficult (than studying plasmids), since complete failure to segregate would be lethal. Some of the work mentioned above is based on sophisticated cytological observations, using, for example, fluorescence-tagged proteins.

Le Dantec et al (2001) examine maintenance of linear plasmids in Mycobacteria.

P. DNA replication in eukaryotes

DNA is DNA. And DNA replication probably is DNA replication. The basic processes -- prepriming, priming, chain elongation, leading and lagging strand synthesis -- appear to be functionally about the same in prokaryotes and eukaryotes. The details of which functions are in which proteins are different in different systems, but the functional equivalence of the systems is the main take home lesson for now.

A "standard model" for eukaryotic replication has emerged. It postulates a replication complex similar to that of Fig 21.19 (the trombone model), with two essential polymerases along with some accessory proteins. As Weaver discusses (pp 668), the emerging view is that Pol α works along with the primase, and Pol δ actually does the bulk of the polymerization, on both leading and lagging strands. In fact, the part made by Pol α may even be removed during primer removal; Bae et al, 2001.

There is still no clear picture of what Pol ε does. In yeast, Pol ε is essential -- apparently for replication (as are Pol α and δ). So we now have three polymerases all essential for replication, even though a two polymerase system seems adequate. Kesti et al (1999) suggest that the essential nature of Pol ε is not due to its polymerase activity. Weaver lists Pol ε as a repair polymerase, Table 20.4; that is a tentative but plausible assignment.

PCNA is the processivity factor for Pol δ (p 668 and then Fig 21.16). There is little sequence homology between PCNA and the E. coli Pol III processivity subunit (β, Sect J above). However, they do have very similar structures and -- presumably -- function (Figs 21.15 & 16). Yuzhakov et al (1999) show a transition from primase to Pol in eukaryotic replication, involving PCNA, similar to that found in bacteria.

Prokaryotic chromosomes typically consist of one replicon, whereas eukaryotic chromosomes typically consist of many replicons (Fig 20.12 and p 652). These multiple replicons are replicated in an ordered sequence during the DNA synthesis part of the cell cycle ("S phase"). Thus we must deal with two regulation issues in eukaryotic DNA synthesis:

1. timing of the S phase;
2. timing of the replication of individual replicons, within S phase.

#1, the timing of S phase, is part of the cell cycle story.

#2 is poorly understood. A developing idea is that an origin needs a "license" in order to actually function. Licenses are issued during cell division, and are good for only one round.

The licensing model has stimulated much thinking, but can at most be only part of the story. Licensing per se does not explain how the timing of replication is actually determined for individual replicons. Current work focuses on analysis of the "origin recognition complex" (ORC) -- a large protein complex which, in one form or another, is situated on most origins most of the time. Protein kinases that control both assembly and firing of initiation complexes are under study.

Lipford & Bell (2001) discuss the role of nucleosomes in initiation of replication.

Nguyen et al (2001) discuss the problem of preventing reinitiations.

Gilbert (2001) reviews the complex story of initiation sites for replication in higher eukaryotes.

van Brabant et al (2001) explore the checkpoints that ensure that DNA replication is completed before mitosis proceeds.

Cimbora & Groudine (2001) generally review the regulation issues.

Cook (1999) considers spatial aspects of replication (and transcription).

Tyler et al (1999) discuss assembly of chromatin from newly made DNA.

Q. How many DNA polymerases?

Weaver tells you about bacterial DNA Pol up to III (p 661), and eukaryotic DNA Pol up to ε (Table 20.4, p 668). An interesting recent development is the identification of more DNA polymerases -- about a doubling -- in both systems. E coli Pols are now up to V, and eukaryotic Pol up to μ (mu, the twelfth Greek letter). All of the new polymerases fall into a general class: they are repair polymerases with specialized functions. They all help the replication machinery bypass one or another type of difficult lesions, due either to external damage (such as UV) or an unrepaired replication error.

Some of these Pols are "error-prone", thus lead to increased mutation. One can think of this error-prone replication (or repair) as a cost of replicating otherwise unreadable lesions, but it also may be a feature to promote adaptation. Radman (1999) discusses the role of error-prone replication, and Metzgar & Wills (2000) more broadly discuss the heritability and control of mutation rates.

Weaver introduces some of these new polymerases in his sub-section on Error-prone bypass, pp 677 ff. (Interestingly, he does not include them in his main count of polymerases, presumably reflecting how new all this is.)

See McKenzie et al (2001) for recent work on E coli Pol IV. See Friedberg et al (2000) for a general update on the eukaryotic DNA Pols -- now up through μ. Johnson et al (2000) show that Pol ζ is more complicated than Weaver tells you (p 679). Bemark et al (2000) show that Pol ζ is essential in mice. Haracska et al (2001) explore Pol η.

An unidentified error-prone DNA Pol has been found in E coli carrying the *mutA* mutation, which results in a tRNA that misreads certain Gly codons (Ren et al, 2000). Strange story!

R. Further reading

A Yuzhakov et al, Replisome assembly reveals the basis for asymmetric function in leading and lagging strand replication. Cell 86:877, 9/20/96. We have long understood that the replication apparatus must be functionally asymmetric, because of the "polarity problem". Now, Yuzhakov et al show the basis of this asymmetry: the DnaB helicase, positioned on the lagging strand just ahead of the fork. The helicase interacts with Pol via the τ subunit. This interaction is necessary for the helicase to function efficiently. See Clover & McHenry (2001) for more.

J-K Gerber et al, Termination of mammalian rDNA replication: polar arrest of replication fork movement by transcription termination factor TTF-1. Cell 90:559, 8/8/97. The same protein serves to terminate transcription and replication. Note that this deals specifically with rRNA genes, thus the transcription is by Pol I. They discuss the problem of collisions between RNA Pol and DNA Pol, at highly expressed genes.

C-C Chiang & A M Lambowitz, The Mauriceville retroplasmid reverse transcriptase initiates cDNA synthesis de novo at the 3′ end of tRNAs. Mol Cell Biol 17:4526, 8/97. "We show that [this] reverse transcriptase … has the unprecedented ability for a DNA polymerase to initiate DNA synthesis at a specific site in a natural template without a primer." (From the abstract of an earlier paper.) Thus another common generality (Sect E) falls. This new result is interesting from an evolutionary perspective. RNA polymerase can initiate, and the first DNA Pol presumably evolved from an RNA Pol. Thus it is reasonable (even expected) that a primitive DNA Pol might be found that initiates. The failure of common modern DNA Pols to initiate would seem to reflect the loss of a feature.

T A Steitz, Structural biology: A mechanism for all polymerases. Nature 391:231, 1/15/98. News item to accompany two articles in this issue. The articles present crystal structures of DNA Pols, and provide some insight into mechanistic details. For example, it seems that the binding site for the nascent DNA quite strictly enforces Watson-Crick base pairing. Improperly paired bases just won't fit well.

C Levine & K J Marians, Identification of dnaX as a high-copy suppressor of the conditional lethal and partition phenotypes of the parE10 allele. J Bact 180:1232, 3/98. The implication is that the γ subunit of Pol III may interact with Topo IV, which decatenates replicated chromosomes.

G del Solar et al, Replication and control of circular bacterial plasmids. Microbiol & Molec Biol Rev 62:434, 6/98. Major review.

T Katayama et al, The initiator function of DnaA protein is negatively regulated by the sliding clamp of the E. coli chromosomal replicase. Cell 94:61, 7/10/98. The β subunit does more than just hold Pol on the DNA; it also inactivates DnaA protein, thus preventing multiple initiations.

D M Heithoff et al, An essential role for DNA adenine methylation in bacterial virulence. Science 284:967, 5/7/99. Dam^- mutants of Salmonella are not virulent, probably due to altered binding of regulatory proteins. They suggest that Dam might be a good target for antibacterials. Also see Reisenauer et al (1999), below.

S A Taft-Benz & R M Schaaper, The C-terminal domain of DnaQ contains the polymerase binding site. J Bact 181:2963, 5/99. For E. coli Pol III, the DnaQ protein is the proofreading subunit, ε, which binds the polymerizing subunit α and the mysterious θ.

T Kesti et al, DNA polymerase ε catalytic domains are dispensable for DNA replication, DNA repair, and cell viability. Mol Cell 3:679, 5/99. Weaver classifies the eukaryotic DNA Pol ε as a repair enzyme (Table 20.4), but this is not clear. The basic confusion comes from two facts:

1) Yeast mutants lacking Pol ε are inviable, thus showing that Pol ε is essential; 2) Complete DNA replication in systems totally lacking Pol ε is well characterized. Here, Kesti et al use Pol ε mutants. They show that mutants lacking all known catalytic domains are sufficient for yeast viability. However, a mutant lacking the C-terminal domain that has no recognized function is inviable. The results support the essential role of this protein, but indicate that its essential role is not as a polymerase per se.

P R Cook, The organization of replication and transcription. Science 284:1790, 6/11/99. Review. Recent work suggests that polymerases may have fixed locations, and that the DNA threads through polymerization factories. Cook discusses both prokaryotic and eukaryotic cells.

W C Hahn et al, Creation of human tumour cells with defined genetic elements. Nature 400:464, 7/29/99. (+ News, Weitzman & Yaniv, p 401.) Of immediate relevance is that one of the three mutations required was to restore telomerase.

P R Meyer et al, A mechanism of AZT resistance: an increase in nucleotide-dependent primer unblocking by mutant HIV-1 reverse transcriptase. Mol Cell 4:35, 7/99. The anti-HIV drug AZT works because it has a blocked 3′ position, thus further chain extension cannot occur. (This is the same principle as the use of dideoxynucleotides in DNA sequencing; Fig 5.18.) Here they show that viral mutants with AZT resistance have a polymerase that now can remove the terminal AZT, thus allowing chain extension. The polymerase in this case is a reverse transcriptase (RT); RTs do not normally proofread, and this seems to be an unusual activity.

B T Kren et al, Correction of the UDP-glucuronosyltransferase gene defect in the Gunn rat model of Crigler–Najjar syndrome type I with a chimeric oligonucleotide. Proc Natl Acad Sci USA 96:10349, 8/99. (Also see The Scientist 1/10/00, p 13.) The idea is to treat a genetic disease by exploiting the DNA repair system of the organism. The treatment uses a specially designed oligonucleotide ("oligo") that will hybridize to the defective gene, but carries the "corrected" sequence. The oligo directs the DNA repair system to change the chromosomal allele to what is on the oligo. The technique is nicknamed chimeraplasty, since the oligo is part DNA and part RNA. This paper reports a test of the method in rats.

A Reisenauer et al, Bacterial DNA methylation: a cell cycle regulator? J Bacteriol 181(17):5135-9, 9/99. Minireview. They review the roles of methylation, beyond that in modification-restriction systems. These roles are diverse and variable; some are essential, some not. Also see Heithoff et al (1999), above.

J J Pouliot et al, Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes. Science 286:552, 10/15/99. The topoisomerase reactions involve covalent protein-DNA intermediates, which preserve the phosphate ester linkages. If the reaction is interrupted, these intermediates may accumulate, and that is "bad". In fact, some anticancer drugs work by inhibiting the topo at such an intermediate stage. Here, they uncover an enzyme that can hydrolyze such protein-DNA adducts. Yeast mutants lacking this enzyme are hypersensitive to accumulation of topo-DNA adducts. The enzyme is also present in humans; its presence may have implications for cancer therapy using topo inhibitors.

M Radman, Mutation: Enzymes of evolutionary change. Nature 401:866, 10/28/99. News. Discusses new DNA polymerases, especially in context of causing mutations.

J Yan et al, A kinetic proofreading mechanism for disentanglement of DNA by topoisomerases. Nature 401:932, 10/28/99. They discuss how disentangling enzymes may achieve better disentangling than predicted by a simple equilibrium between tangled and untangled forms. They invoke an ATP-dependent kinetic proofreading model.

L Fang et al, Replisome assembly at oriC, the replication origin of E. coli, reveals an explanation for initiation sites outside an origin. Mol Cell 4:541, 10/99. They analyze the detailed sequence of events at the origin region. In part, they make use of a mutant helicase that can bind but not move. They show that primase can initiate only after helicase has moved 65 bases or so. Since the primase seems to be chasing a moving target (the helicase), priming sites per se are not well defined. Their "Fig 7" is attached, p 39.

A Yuzhakov et al, Multiple competition reactions for RPA order the assembly of the DNA polymerase δ holoenzyme. EMBO J 18:6189, 11/1/99. They show that a similar series of hand-offs occur in the assembly of a eukaryotic replication complex as with a prokaryotic replication complex. The primase, associated with Pol α, is held to the primer terminus by the SSB called RPA. RFC then joins, competing off Pol α. A PCNA ring is then assembled on the DNA, and Pol δ then displaces RFC, thus joining with PCNA. They note the importance of the proteins changing their affinity for the next protein in the series, as the series of hand-offs proceeds.

J R Mitchell et al, A telomerase component is defective in the human disease dyskeratosis congenita. Nature 402:551, 12/2/99. (Also see News item, Shay & Wright, Science 286:2284, 12/17/99.) They implicate a telomerase deficiency in a human genetic disease.

J K Tyler et al, The RCAF complex mediates chromatin assembly during DNA replication and repair. Nature 402:555, 12/2/99. RCAF = replication-coupling assembly factor. In fact, RCAF contains histones H3 and H4, the first to be laid down on newly made DNA. They are present in the RCAF in an acetylated form, then deacetylated after binding to DNA. Recall discussion of histone acetylation, to modify the charge, in the context of transcriptional effects, Ch 10 handout Sect P.

T Lindahl & R D Wood, Quality control by DNA repair. Science 286:1897, 12/3/99. Review. They discuss types of damage and types of repair systems. They include the multitude of DNA polymerases specialized for repair, and the role of defective repair in cancer. This article is another in the "Quality Control" set introduced with Ibba & Söll (1999; Ch 19 FR).

M M Hingorani & M O'Donnell (2000a), Sliding clamps: A (tail)ored fit. Current Biology 10:R25–R29, 1/1/00. Update, based on recent structural information. Includes discussion of how sliding clamps help proteins other than DNA Pol track DNA. These other proteins include transcriptional and repair proteins. In some cases, proteins may compete for a single binding site on the clamp. They also note that the clamp may allow Pol to let go of the DNA momentarily, to relieve torsional stresses.

K L Rudolph et al, Inhibition of experimental liver cirrhosis in mice by telomerase gene delivery. Science 287:1253, 2/18/00. (+ related article, Kobayashi et al, p 1258. + News, Hagmann, p 1185.) In a mouse model, cirrhosis of the liver seems to be due to excessive TEL shortening; remember that the liver is an organ with high cell proliferation. Providing the missing telomerase by gene therapy alleviated the condition.

H J Zuccola et al, The crystal structure of an unusual processivity factor, herpes simplex virus UL42, bound to the C-terminus of its cognate polymerase. Mol Cell 5:267, 2/00. The Herpes virus processivity factor does not form a belt, but rather binds tightly to DNA, via electrostatic interactions with the phosphate backbone. How it actually moves is unclear. Despite the different mechanism, much of the protein is similar to belt-forming processivity factors. The differences might be suitable targets for anti-herpes drugs.

J L Keck et al, Structure of the RNA polymerase domain of E coli primase. Science 287:2482, 3/31/00. (+ News, von Hippel & Jing, p 2435.) Includes discussion of the coordination issues in replication. The news article is particularly good for this. In particular, they discuss the handoff from primase to polymerase, mediated by the sliding clamp loader subunit χ, interacting with SSB just ahead of the primer.

L Ren et al, Requirement for homologous recombination functions for expression of the mutA mistranslator tRNA-induced mutator phenotype in Escherichia coli. J Bacteriol 182:1427-1431, 3/00. Mutator strains are strains that show high mutation rates; there are tricks to find such mutants in lab strains. The mutA mutation was shown to be due to an anticodon mutation in a Gly tRNA; the tRNA now inserts Gly at certain Asp codons. The strains also have an error-prone DNA Pol activity, and hence a high mutation rate. To my knowledge, it is not known whether the Pol is a novel Pol or “merely” a mistranslated version of a known Pol.

A A Griffiths & R G Wake, Utilization of subsidiary chromosomal replication terminators in Bacillus subtilis. J Bacteriol 182:1448-1451, 3/00. They study strains with chromosomal rearrangements, and show that which Ter sites are used depends on their positioning.

J J Blow & S Tada, Cell cycle: A new check on issuing the licence. Nature 404:560, 4/6/00. News. Discusses two papers in this issue, dealing with how licensing and initiation per se are separated. It is logically required that licensing activity cease before initiation begins, or else some regions might replicate twice.

T R Strick et al, Single-molecule analysis of DNA uncoiling by a type II topoisomerase. Nature 404:901, 4/20/00. They follow the topoisomerase reaction by measuring the change in extension of a single DNA molecule as the enzyme changes the topology.

A Oliver et al, High frequency of hypermutable Pseudomonas aeruginosa in cystic fibrosis lung infection. Science 288:1251, 5/19/00. (+ News, Rainey & Moxon, p 1186.) They show that the Pseudomonas isolated from patients are variable, and many isolates are hypermutable. They show that some of the hypermutability is due to mutations in mismatch repair system.

E C Friedberg et al, The many faces of DNA polymerases: Strategies for mutagenesis and for mutational avoidance. PNAS 97(11):5681–5683, 5/23/00. Commentary article on the zoo of eukaryotic polymerase, now up to Pol μ.

G Newman & E Crooke, DnaA, the initiator of Escherichia coli chromosomal replication, is located at the cell membrane. J Bacteriol 182:2604-2610, 5/00. As the title suggests, they show, by immunostaining, that the major location of DnaA protein is at the membrane. This fits with the known role of anionic phospholipids in activating DnaA, by promoting removal of the bound ADP. However, the method is low resolution, thus does not identify potentially important smaller amounts of DnaA at other sites, and the paper offers little new about the key controlling events. Nevertheless, the article is well written, and is a useful and readable update about the DnaA mystery.

D Metzgar & C Wills, Evidence for the adaptive evolution of mutation rates. Cell 101(6):581, 6/9/00. Minireview. A broad discussion of mutation rates, and how they may vary with conditions. This includes a brief mention of error-prone DNA polymerases.

H Nilsen et al, Uracil-DNA glycosylase (UNG)-deficient mice reveal a primary role of the enzyme during DNA replication. Molecular Cell 5(6):1059, 6/00. UNG is an enzyme to remove uracil from DNA. In microbes, reduction of this enzyme activity leads to higher mutation rates. The mutations occur when C in DNA deaminates to produce U, thus making a G-U pair; replication results in a GC pair and an AT pair. Here, they make mice deficient in the major UNG enzyme, and show little if any effect on mutation rate. From this and other results, they suggest that the enzyme they dealt with is specialized for removing newly incorporated U from DNA, and that another enzyme deals with U resulting from C deamination. (Weaver briefly discusses the UNG enzyme, p 647, in the context of the complications it causes for observing Okazaki fragments. In the lab, UNG is used in some PCR procedures as part of a process to minimize cross contamination.)

A Kornberg, Ten commandments: Lessons from the enzymology of DNA replication. J Bacteriol 182:3613-3618, 7/00. A “Commentary” from the original discoverer of DNA polymerase. Fun and useful. (# VIII: Respect the personality of DNA.)

K Hiratsuka & L J Reha-Krantz, Identification of Escherichia coli dnaE (polC) mutants with altered sensitivity to 2′,3′-dideoxyadenosine. J Bacteriol 182:3942-3947, 7/00. “Dideoxy” nucleotides are like regular DNA nucleotides, except that they lack the 3′-OH group. Thus, they can (in principle) be incorporated into DNA, but cannot be added onto. They are used as chain terminators in DNA sequencing reactions, and some are use as drugs. Polymerases vary in their ability to actually incorporate these nucleotide analogs. In this work, they isolate mutants of E. coli Pol III with an altered ability to use them.

S E Artandi et al, Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. Nature 406:641, 8/10/00. Telomerase-deficient mice show increased cancer rates. In the work here, the mice were also P53-deficient. They note that mice and humans typically get different kinds of cancer; part of the reason may be the different TEL dynamics.

R E Johnson et al, Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. Nature 406:1015, 8/31/00. They describe in some detail how these two Pols work to allow bypass of a lesion. Pol ι incorporates nucleotides opposite non-coding lesions, and Pol ζ can extend from such nucleotides.

J G Bertram et al, Molecular mechanism and energetics of clamp assembly in Escherichia coli - The role of ATP hydrolysis when γ complex loads β on DNA. J Biol Chem 275(37):28413–28420, 9/15/00. Analysis of the role of ATP and its hydrolysis in the clamp loading function. Much of the paper is on the technical issues of such measurements.

M Bemark et al, Disruption of mouse polymerase ζ (Rev3) leads to embryonic lethality and impairs blastocyst development in vitro. Current Biology 10(19):1213, 10/1/00. (+ related papers, p 1217, 1221.) This set of three papers shows some intriguing results for Pol ζ. Absence of this Pol in mice leads to uniform lethality in utero, with numerous embryonic defects. Thus Pol ζ is essential in mice -- though not essential in yeast. The reason for this, and whether it relates to its activity as a DNA Pol, are not known.

M M Hingorani & M O'Donnell (2000b), A tale of toroids in DNA metabolism. Nature Reviews Molecular Cell Biology 1:22, 10/00. A broad overview of a variety of proteins involved in DNA metabolism that share the common structural feature of forming a ring around the DNA. These include sliding clamps, the hexameric helicases, topoisomerases, and others. These proteins share no other functional features, and probably are evolutionarily unrelated. In some cases they note that members of the family in which holding on to the DNA for long times is more important are toroidal, whereas other members of the family with less need to stay on the DNA for extended periods are not toroidal. This is from the inaugural issue of a new journal from Nature. (I'm not sure how readily available this is; if you want a copy, bring me a disk and I will give you the pdf file, which is about 1 Mb.)

T B L Kirkwood & S N Austad, Why do we age? Nature 408:233, 11/9/00. One of a set of six articles in a special feature section on Ageing (as Nature spells it).

J R Walker et al, Escherichia coli DNA polymerase III τ- and γ-subunit conserved residues required for activity in vivo and in vitro. J Bacteriol 182:6106-6113, 11/00. The τ- and γ-subunits are made from the same gene; the τ product is the full length translation product, but a frameshift results in the shorter γ subunit. Although we normally show distinct roles for these two proteins, there is actually some doubt. The introduction to this article is a good overview of the τ-γ story. The article itself helps to identify key amino acids, but does not resolve the τ-γ ambiguities.

M Nakatani et al, A DNA ligase from a hyperthermophilic Archaeon with unique cofactor specificity. J Bacteriol 182:6424-6433, 11/00. ATP-dependent and NAD-dependent ligases were thought to be entirely distinct. Here they describe a ligase that will use either cofactor.

H Hemmi et al, A Toll-like receptor recognizes bacterial DNA. Nature 408:740, 12/7/00. (+ News, Modlin, p 659.) How do you recognize a generic bacterial infection? In part, by characteristic features of bacterial DNA, in particular unmethylated CpG sequences. Here they identify a receptor responsible for this recognition.

P H von Hippel & E Delagouette, A general model for nucleic acid helicases and their "coupling" within macromolecular machines. Cell 104(2):177-190, 1/26/01. Major review of how helicases work. Includes both dimeric and hexameric helicases, acting on either DNA or RNA.

M S Junop et al, Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. Molecular Cell 7(1):1-12, 1/01. They show that ATP is involved in activating a mismatch repair event. Exactly how isn't clear. But they suggest that MutS must bind both a mismatch and ATP at the same time in order to activate an event.

J R Lipford & S P Bell, Nucleosomes positioned by ORC facilitate the initiation of DNA replication. Molecular Cell 7(1):21-30, 1/01. They show that the yeast ORC causes nucleosome phasing, and that some of the phased nucleosomes play a role in initiation of replication.

J W Shay & W E Wright, Aging: When do telomeres matter? Science 291:839, 2/2/01. News, to accompany two articles in this issue. Mammalian cells have a finite lifespan in culture, but the details vary for different organisms. Here they emphasize the role of "culture shock", the response of the cells to the changing conditions in culture, as an important part of the story. The articles here show indefinite replication of rat cells in vitro.

G Macintyre et al, Lowering S-adenosylmethionine levels in Escherichia coli modulates C-to-T transition mutations. J Bacteriol 183:921-927, 2/01. S-Adenosylmethionine (SAM) is a physiological methyl donor. Given the importance of methyl groups in DNA (for example, in distinguishing T vs U), there are questions about the effect of SAM on mutation rates. Effects in both directions have been demonstrated. This paper explores the effect of lowered SAM levels in E. coli; depending on the specific situation, an increase or decrease of the mutation rate is seen.

S A Sciochetti et al, Identification and characterization of the dif site from Bacillus subtilis. J Bacteriol 183:1058-1068, 2/01. The dimer-resolution system of Bacillus subtilis.

Z Li et al, The mechanism of type IA topoisomerase-mediated DNA topological transformations. Molecular Cell 7(2):301, 2/01. They engineer a modified form of the enzyme that allows them to trap proposed intermediates, with DNA bound within the enzyme cavity.

D M Cimbora & M Groudine, The control of mammalian DNA replication: A brief history of space and timing. Cell 104(5):643-646, 3/9/01. Minireview. Good introduction to the issues of control of eukaryotic replication.

1) B Sat et al, Programmed cell death in Escherichia coli: Some antibiotics can trigger mazEF lethality. J Bacteriol 183:2041-2045, 3/01. 2) R Hazan et al, Postsegregational killing mediated by the P1 Phage "Addiction Module" phd-doc requires the Escherichia coli programmed cell death system mazEF. J Bacteriol 183:2046-2050, 3/01. One system for plasmid maintenance is the addiction system, where the plasmid makes something that protects the cell from its own toxin. However, it now seems that such toxin-antitoxin systems are more common, and play a part in cell metabolism, unrelated to plasmids. They suggest that such systems are a primitive form of programmed cell death, resulting in an intentional cell killing under adverse conditions. See hw.

G J McKenzie et al, SOS mutator DNA polymerase IV functions in adaptive mutation and not adaptive amplification. Molecular Cell 7(3):571-579, 3/01. An introduction to the least known DNA Pol of E. coli, which Weaver briefly notes on p 677.

S B Cantor et al, BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. Cell 105(1):149-160, 4/6/01. Another part of the (still murky) BRCA1 story.

C Le Dantec et al, Genomic sequence and transcriptional analysis of a 23-kilobase Mycobacterial linear plasmid: Evidence for horizontal transfer and identification of plasmid maintenance systems. J Bacteriol 183:2157-2164, 4/01. Exploration of plasmid maintenance in a system that is novel in a couple of ways… linear plasmid, and a less studied host. Their tentative conclusion is that the plasmid maintenance functions are similar to those more commonly characterized in circular plasmids.

A J van Brabant et al, An origin-deficient yeast artificial chromosome triggers a cell cycle checkpoint. Molecular Cell 7(4):705-713, 4/01. Checkpoints are controls that ensure that necessary precursor steps are completed before doing following steps. For example, logically, replication must be completed before mitosis can occur. What is the molecular basis of such checkpoints? Here they show that a YAC that lacks an origin of replication for one region triggers the checkpoint. This suggests that this checkpoint is not just measuring the "damage" that occurs during replication, such as SS breaks, but is somehow measuring lack of replication. Taking into account other work, they suggest that it is not the DNA content per se that is measured, but rather the existence of unfired initiation complexes, which serve to indicate that replication is incomplete. What unfired initiation complexes? Weak ones, too weak to actually be used; they would be removed if an active fork moved through the region. (Hm. Does this imply that having extra, weak and unusable, origins is actually good?)

P Baumann & T R Cech, Pot1, the putative telomere end-binding protein in fission yeast and humans. Science 292:1171, 5/11/01. (+ News, de Lange, p 1075.) Another protein involved with telomeres. This one caps the overhanging end. Thus it is the protein that Weaver says is unknown (p 711). How this relates to the story of t-loops (Fig 21.35) remains to be sorted out.

J H J Hoeijmakers, Genome maintenance mechanisms for preventing cancer. Nature 411:366, 5/17/01. Review. Part of a set of articles in a feature section on cancer.

A Rizki & V Lundblad, Defects in mismatch repair promote telomerase-independent proliferation. Nature 411:713, 6/7/01. (+ News, Kucherlapati & DePinho, p 647.) A complication of the TEL story is that cells can maintain their telomeres without telomerase (p 712). This is done by recombination between TELs. Here they show that mutations in the mismatch repair system enhance TEL maintenance. The reason is probably that the mismatch repair system serves as a barrier to recombination between "similar" (but non-identical) sequences.

R R Sinden, Neurodegenerative diseases: Origins of instability. Nature 411:757, 6/14/01. News. Update on the issue of how triplet expansions may occur.

V Q Nguyen et al, Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. Nature 411:1068, 6/28/01. They show that multiple mechanisms are involved in preventing reinitiations in yeast. All must be disrupted before an origin will reinitiate.

B P Clover & C S McHenry, The DNA polymerase III holoenzyme: An asymmetric dimeric replicative complex with leading and lagging strand polymerases. Cell 105(7):925-934, 6/29/01. They show that there is a functional asymmetry in the Pol III complex.

P Funchain et al, Amplification of mutator cells in a population as a result of horizontal transfer. J Bacteriol 183:3737-3741, 6/01. Mutator cells are cells with a high mutation rate. They may be of considerable significance in how a population adapts. Cells with a defective mismatch repair system are one type of mutator cell, and they may arise during interspecific matings. Why? Because the mismatch repair system tends to inhibit interspecific recombination (due to mismatches between the two species); as a result, the relatively infrequent recombinants that do occur are enriched for mutators.

Y Peng et al, Analysis of telomerase processivity: Mechanistic similarity to HIV-1 reverse transcriptase and role in telomere maintenance. Molecular Cell 7(6):1201-1211, 6/01. Telomerase acts as a reverse transcriptase (RT). In fact, it is quite similar is many ways to a typical RT. However, it is not very processive.

S-H Bae et al, RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes. Nature 412:456, 7/26/01. The details of primer removal and fragment joining have remained remarkably elusive in eukaryotes. Here they propose a new model, involving several enzymes. One interesting point is the possibility that not only the RNA primer, but also the region made by Pol α may be removed.

J R Walker et al, Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature 412:607, 8/9/01. We have briefly noted how the telomere aids in replication maintenance of chromosome ends in eukaryotes. But the TEL plays another role -- to distinguish a “normal” end from a “broken” end. The latter must be repaired. A key protein in repairing double-strand breaks (DSB) is Ku, which brings together ends and aligns them for a substantially non-specific end-joining reaction. (Ku is also required for immunoglobulin gene rearrangements, repairing the DSB that are a part of that process.)

J A Hackett et al, Telomere dysfunction increases mutation rate and genomic instability. Cell 106(3):275-286, 8/10/01. Yeast.

K P Lemon & A D Grossman, The extrusion-capture model for chromosome partitioning in bacteria. Genes & Development 15:2031–2041, 8/15/01. Review. They broadly review the area of bacterial chromosome segregation. They emphasize the coordination of the steps, especially how the movement of the chromosomes during replication is a precursor to chromosome segregation.

1) D Jeruzalmi et al, Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of E. coli DNA polymerase III. Cell 106(4):417-428, 8/24/01. 2) D Jeruzalmi et al, Crystal structure of the processivity clamp loader gamma (γ) complex of E.

coli DNA polymerase III. Cell 106(4):429-441, 8/24/01. A pair of papers updating the story of how the γ complex works.

1) L Haracska et al, Interaction with PCNA is essential for yeast DNA polymerase η function. Molecular Cell 8(2):407-415, 8/01. 2) J Trincao et al, Structure of the catalytic core of S. cerevisiae DNA polymerase η: implications for translesion DNA synthesis. Molecular Cell 8(2):417-426, 8/01. Details of Pol η (p 679).

D M Gilbert, Making sense of eukaryotic DNA replication origins. Science 294:96, 10/5/01. Review. Although specific origin sequences have been well characterized for bacteria and for yeast, the situation is much less clear for higher eukaryotes. Gilbert argues that much of the control of replication is done through the protein complexes, with sequence specificity less important.

S. Computer resources

The Protein Data Bank now features one protein (or other biological structure) each month, with pictures and extensive discussion. They call this feature "Molecule of the Month". Features relevant to X107A material include: DNA polymerase (March 2000), nucleosomes (July 2000), restriction enzymes (August 2000), ribosomes (October 2000), transfer RNA (March 2001), aminoacyl-tRNA synthetases (April 2001), DNA (November 2001). For the list of featured Molecules of the Month:
http://www.rcsb.org/pdb/molecules/molecule_list.html

PDB files for **sliding clamps** are available, from the Protein Data Bank. This will let you visualize and manipulate these as 3D structures. See my RasMol web page for getting and working with protein files. File 2pol shows the structure of the E coli Pol III β subunit, and file 1axc shows the structure of PCNA. [Small problem: Two of the three chains of the PCNA structure display in colors that are either identical or very similar, depending on your system.]

A comprehensive web site about **telomeres** and **telomerase** (Sect G):
http://resolution.colorado.edu/~nakamut/telomere/telomere.html

While there, you might jump to the site for the Aging Research Centre (ARC). That site is developed by a recent X107 student -- who also suggests the following related sites:
http://www.pbs.org/newshour/forum/october98/glenn.html
http://nihlibrary.nih.gov/internet/Biomedicine_Science/bsphysproc.htm#aging

Ch 22

Although we will not discuss Ch 22, the following site seems well worth listing. It has animated gif pictures showing some of the steps in genetic recombination, especially how DNA is threaded through the Ruv protein complex during the process of branch migration.

They can be thought of as animated versions of Weaver's Figs 22.25 and 27. Branch migrations are similar to the action of a DNA helicase.
http://www.shef.ac.uk/uni/academic/I-M/mbb/ruva/ruva.html

or the US mirror at
http://www.sdsc.edu/journals/mbb/ruva.html

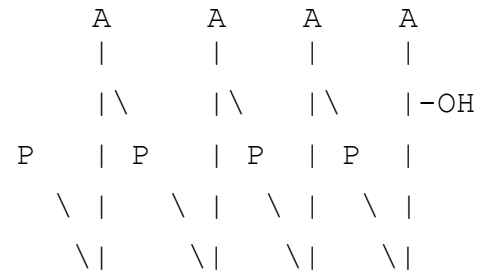
T. Homework

1. Mutations in the *polA* or *lig* genes lead to the accumulation of unusually large numbers of unsealed Okazaki fragments (e.g., Fig 20.7b). Why?

2. On p 686 (Summary) Weaver gives the rate of DNA replication. Given this rate, calculate how long it takes to replicate the E. coli genome. If your result substantially disagrees with the given time of 40 min, think about why.

3. How often do new Okazaki fragments initiate in E. coli? Assume average size of Okazaki fragments is 1.5 kilobasepair (kbp) (p 646).

4. What is unusual about the linkages in the following oligonucleotide? (Oligo means "a few"; an oligonucleotide is a polynucleotide with "a few" subunits.) Which common nucleic acid sugar is the sugar in this oligonucleotide?



5. Consider the following situation: You are walking down an ordinary street. Each house has its number in front, printed horizontally in the usual way. Make an analogy between how you read the house numbers as you proceed down the street and the polarity problem of DNA synthesis.

6. Is the reaction of an exonuclease (which removes one nucleotide at a time from the end of a chain, by hydrolysis) the reverse of the reaction of polymerase (which adds one nucleotide at a time to the end of a chain)? What are the energetic considerations?

7. It is often said that enzymes are not involved in making or breaking weak bonds. An X107A student challenged that point based on the discussion of DNA replication. What is the basis of the challenge? Do you agree?

8. The bacterial chromosome replicates from an origin, called *oriC* in E. coli, outward in both directions until the two replication forks meet, about half way around.

For simplicity here, let's consider bi-directional replication of a linear replicon. Consider the following map, with equally spaced genes…

A B C D oriC E F G H

In a cell with a complete chromosome, not in the process of replicating, all genes are present in equal amounts.

– a. In what order do the genes listed above replicate?

– b. Half way through a replication cycle, which of these genes have replicated, which have not? What are the relative numbers of copies of these genes in the cell?

– c. Under conditions of very rapid growth (very rich medium), E. coli will initiate a second round of DNA replication while a first round is still in progress, about half completed. In this case, what is the relative number of copies of a gene that replicates early compared to a gene that replicates late? Give examples of such genes.

– d. Suggest a way that you might measure the relative numbers of copies of various genes.

9. The replication DNA polymerase of E. coli, Pol III, is called an asymmetric dimer. Why is it logically necessary that Pol III is asymmetric? Two reasons???

10. Primers for DNA replication are (usually) RNA, and are made by an RNA polymerase, which does not proofread. Does this mean that the mutation frequency is higher at origin regions (i.e., where primers are made)? Explain.

11. Consider synthesis of the complementary strand of the following section of template DNA: 5′-------TTAGCTAG----------
[There is no special significance to the particular sequence.]

a. Sketch the partially formed new strand, including three bases into the specific sequence shown. (That is, show specific bases in your new strand for those three positions.) Show the polarity (direction) of the new strand.

> Parts b and c are independent situations. For both, the following additional information is important. The replication system is severely deficient in dCTP. [This could be in vitro, with a low level of dCTP being supplied, or in vivo with a mutant that has a serious defect in making dCTP.]

b. The next template base is a C. By mistake, an A is incorporated into the new strand here (instead of a G). [We are not concerned with why the A is incorporated here.] What is the effect of the dCTP-deficiency on the mutation rate at the C position, where the incorrect base A has already been incorporated? Explain. [You can describe the effect as increase/no change/decrease, compared to the normal situation with adequate dCTP.]

c. What is the effect of the dCTP-deficiency on the mutation rate at the G position (i.e., when replicating the site with a G in the template)? Why? [As noted above, this is independent of part b. You can assume that the preceding bases are replicated properly.]

12. Consider a particular E. coli gene, say X. It replicates. When will it replicate again? (Assume that the chromosomes replicate once per cell generation.)

13. This question addresses the issue raised in the previous question experimentally. It involves an experiment similar to the Meselson-Stahl experiment. "Heavy" DNA is DNA containing ^{13}C and ^{15}N; it is distinguished from light or hybrid DNA in a CsCl density gradient, just as Meselson & Stahl did.

Consider a population of E. coli cells growing in "heavy" medium for many generations, so that all the DNA is "heavy". At time = 0 a radioactively labeled base is added for a very short time (a "pulse" label). This will label each chromosome at its own replication fork. (Presumably these replication forks are at random positions along the chromosomes.)

After the radioactive pulse label, the cells are transferred to light medium. At various times (say, 0.5, 1.0 and 1.5 generations) after the transfer, the DNA is isolated and analyzed in a CsCl density gradient. Specifically, we look at where the radioactive label is. (Remember, it is at the "heavy" or HH position when we start.) What do you expect? Why? (Assume that the chromosomes replicate once per cell generation.)

14. Consider an ordinary E. coli with a large number of copies of *oriC* plasmids (minichromosomes). What effect would you expect all these extra origins to have on the E. coli, for example, on the timing of replication? Why?

15. Consider a bacterial strain containing a plasmid. The plasmid carries the genes for a RM (restriction-modification) system (Ch 4 Sect 1). That is, it carries a DNA-modifying enzyme (methylase) and a nuclease that degrades unmodified DNA. For simplicity for the moment, assume that the methylase is quite unstable and the nuclease is stable.

a. During growth, a plasmid-free cell is produced (by a failure in the segregation system). What will happen as this plasmid-free cell grows? Explain, based on the properties of the RM system.

b. This is an example of an "addiction system" for plasmid survival. It can be thought of as a part of the segregation system. Such systems are said to consist of a "poison" and an "antidote". Identify the poison and the antidote.

16. Computer question. Load a file for a sliding clamp into RasMol. I encourage you to do these parts with both the E coli Pol III β subunit and the eukaryotic PCNA -- but at least try one. Source files for these proteins are described in Sect S, above.

a. Measure the inner diameter of the clamp -- the hole through which the DNA goes. (If you are not quite sure where to measure, take a few measurements.) The procedure for measuring distances between atoms is described in my RasMol web page, both for regular RasMol and

for Berkeley RasMol. Note that you may need to switch to ball-and-stick display before taking these measurements.

b. Look at the structure of the clamp. Is the part that interacts with DNA mainly α-helix or β-sheet? What about the rest of the protein?

> Note that Fig 21.14 shows you the answer to this. However, you should be able to produce a better view to focus on this particular question. (Anyway, the main purpose of the question is to get you to play with the viewer.) See answer section for a hint. But before you do, experiment with various settings for Display and Colours.

17. Comments on the chapter summaries?

U. Partial answers

1. The corresponding enzymes, DNA polymerase I and DNA ligase, are key to filling the gap left between Okazaki fragments. (Fig 20.29)

2. 4×10^6 ~~bp~~ $\times \frac{1\ \text{~~s~~}}{1{,}000\ \text{~~bp~~}} \times \frac{1\ \text{min}}{60\ \text{~~s~~}} = 67\ \text{min}$

It is probably more like 75-80 min, if we used a more precise rate number. That is about twice the observed time. Why? Replication uses two forks, so each fork only replicates half the chromosome. (If you already took this into account in your original calculation, that is fine.)

3. Approximately once every two seconds. (Genome size is 4×10^3 kbp, with two replication forks providing complete replication in 40 min. Thus each fork replicates 2000 kbp in ~2400 sec.)

4. The linkages are 2′-5′ (rather than the usual 3′-5′). This implies the presence of a 2′-OH group, thus implies ribose.

5. If you can't visualize it, try it. (Some molecular biology homework is best done during daylight hours.)

6. The basic equations:

$DNA_{n+1} + H_2O \rightarrow DNA_n + dNMP$ $\qquad$ $dNTP + DNA_n \rightarrow DNA_{n+1} + PP_i$

These clearly are <u>not</u> the reverse of each other. The normal polymerization reaction uses dNTPs as precursors -- the activated precursors. Nuclease action produces dNMPs. Both polymerization and degradation reactions are designed to be favorable, but they are not direct reverse reactions.

> Recall that most of the driving force for polymerization comes from PP_i hydrolysis (p 3). In vitro -- with purified polymerase, high [PP_i], and no pyrophosphatase -- the reverse reaction can be shown.

7. Recall Purich (2001; Ch 17 FR).

8. a. Replication is bi-directional from *oriC*. Thus one fork will replicate D, C, B, A in that order; the other fork will replicate E, F, G, H. The expected order of replication for these genes is: D&E, C&F, etc.

b. The C-F region has replicated; the ends have not. Thus there is a 2:1 ratio of copy number for "early" genes such as E to "late" genes such as H.

c. There will be a 4:2:1 ratio of gene copy number between very early, middle, and late genes.

d. There are a number of possible approaches. For class discussion.

9. The requirement that the two strands be made in opposite physical directions, resulting from strand polarity and the inability of Pol to make DNA in both chemical directions. But even if this were not true (that is, even if Pol could replicate both strands in the same physical direction), what other consideration(s) might lead to functional asymmetry of the replisome (if not the Pol per se)?

Yuzhakov et al (1996) and Clover & McHenry (2001) examine aspects of the asymmetry.

10. No, at least not for the stated reason. It is possible that primers are less accurate, but they are removed by a nuclease; the primer gap is then filled with DNA, made against the original template. So primer accuracy is not directly reflected in the mutation rate.

Many investigators have looked for differences in mutation rate between the leading and lagging strand. Various results have been obtained. So far as I know, there is no definitive result, and no definitive reason for any alleged positive results.

11. a. 5′-------TTAGCTAG----------
ATC----------5′

b. The next template base is a G, so C is the proper next base to incorporate. But C (i.e., dCTP) is in short supply. Therefore elongation will be slow, thus allowing further time for the mispair to be repaired. The result: dCTP deficiency lowers this mutation rate.

c. In this case, we need the C and it is in short supply. Because of the C deficiency, there is a greater chance of a wrong base getting incorporated. Higher mutation rate.

12. One generation later. That is, a given gene replicates at a particular time in the cell cycle. This follows if replication is always from the same origin and in the same order, and starts at a fixed time in the cell cycle.

13. The radioactive label is in a heavy strand. Originally it is in an HH duplex. Upon replication in light medium it will go into HL duplex DNA. The label can never appear in LL DNA, since the label itself is in an H strand. So the question reduces to: What are the kinetics of shift of the label from the HH band to HL?

The shift will occur the next time the labeled segment replicates. Although the population is asynchronous, this doesn't matter. In a given cell the label is at a specific point. One replication cycle later the same point will replicate, shifting the label to hybrid density. (Since replication is bi-directional, with two active replication forks per replication cycle, two sites are labeled. But the logic is still the same, so long as the replication pattern is consistent.)

Let's start with simple assumptions… the DNA replicates once per cell generation, and replication starts at a specific time in the cell cycle. That is, replication events are precisely timed. In this case, each newly replicated site (which is pulse labeled) will next be replicated

precisely one generation later. Thus we should see no transfer of the label from the heavy position to hybrid before one generation, at which time it should all shift.

This question is based on Bakker & Smith, Methylation of GATC sites is required for precise timing between rounds of DNA replication in Escherichia coli. J Bact 171:5738, 10/89. Their goal was to test whether there is precise timing between rounds of DNA replication. For wild type E. coli they observed the expected result. However, for *dam* mutants, defective in DNA adenine methylation, the label shifted gradually to the hybrid position. They conclude, then, that methylation is necessary for precise timing of DNA replication cycles. Interestingly, the E. coli with random replication timing appear to grow fine.

Reisenauer et al (1999) review the diverse roles for DNA methylation in bacteria.

14. The important point for now is your model. How do you envision the system, and what are the implications of your model?

The observed result, surprising to many observers, is that there is no effect. Apparently E. coli knows how many origins it has, and adjusts accordingly. Adjusts what? Good question. Or maybe it neither knows nor cares how many origins there are.

15. a. It will die (or, more precisely, its progeny will die).

The plasmid-free cell cannot make more of the RM enzymes. Since the methylase is lost rapidly (by assumption), upon replication, the progeny will contain unmodified DNA, and will be killed by the restriction nuclease. Thus the RM system serves to promote apparent stability of the plasmid, by killing cells that don't contain the plasmid.

I gave the assumption that the methylase is unstable, to make the logic clear. That assumption is not necessarily very valid. The effect must depend on the relative activities of the two enzymes (as well as the capacity of the host to repair the damage created by the nuclease).

> This question has nothing to do with *dam* methylation. There are many methylases, each with its own specificity (and role).

b. restriction enzyme = poison; methylase = antidote. Sat et al (2001) discuss such a system, but propose a broader role.

16. a. You should get measurements on the order of 25-30 Å. Remember that the DNA helix has a diameter of 20 Å. The Figures in the textbook give you an idea of the fit.

b. Gross structural features of proteins are usually best viewed with Display = Ribbon (or Strands). Set Colours = Structure to focus on this particular question. Try Colours = Chain to see how many chains there are.

You should see that the inner part of the rings (of both clamps) -- the part that interacts with DNA -- is mostly α-helix. The rest of the protein is mainly β-sheet.

(End of answer section; attached figures on next pages.)

"Fig 7" "One possible…" from S Kim et al, Coupling of a replicative polymerase and helicase: A τ-DnaB interaction mediates rapid replication fork movement. Cell 84:643, 2/23/96.

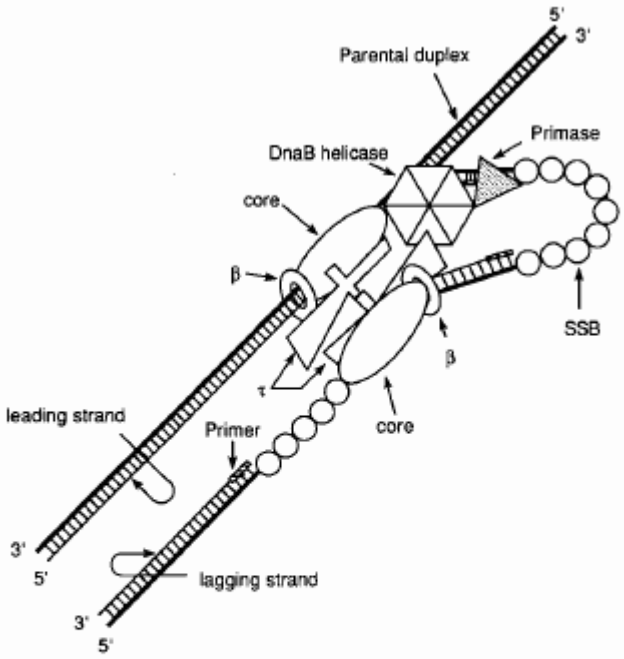


Figure 7. One Possible Arrangement of Proteins at the Replication Fork

In this diagram, the leading- and lagging-strand polymerases are shown in an anti-parallel orientation. With this symmetry, only the τ subunit associated with the lagging-strand polymerase would likely be close enough to DnaB to establish a protein–protein connection. See text for details. The drawing is not to scale.

"Fig 7" from Fang et al (1999).

Their figure legend:

Figure 7. Stages in Assembly of Two Opposed Replication Forks at oriC

(A) Stages in replisome assembly. Stage I, two DnaB hexamers are assembled onto the DnaA-activated open complex through the action of DnaC. Stage II, DnaB helicases pass each other creating ssDNA for primase action. The passing action ensures that the region between the helicases is melted and remains so upon being coated with SSB. Primase must interact with DnaB to initiate RNA synthesis resulting in RNA primers in cis to DnaB. Stage III, two replicases assemble onto the two primed sites. Stage IV, the two molecules of Pol III holoenzyme extend DNA opposite the motion of DnaB assembled on the same strand. Hence, the polymerases pass one another to reach the DnaB helicases on the opposite strand, which move in the same direction as DNA polymerization.

(B) The factory model for replication indicates that the polymerases remain fixed while the DNA moves.