MCB137L/237L: Physical Biology of the Cell
Spring 2025
Homework 2
(Due 2/4/24 at 2:00pm)

Hernan G. Garcia

Last updated January 28, 2025

"Trying to understand molecular biology without numbers is like studying History without knowing Geography." - Prof. Marc Kirschner

1 The concentration rule of thumb.

In the last homework, we worked out the rule of thumb that one molecule per *E. coli* cell corresponds to a concentration of ≈ 1 nM.

(a) As an application of this idea, how many H^+ ions are there in a bacterial cell if the pH is 7.0?

(b) It is very useful to have a sense of how far molecules are apart at a given concentration. Work out a formula that relates the spacing between molecules d to the concentration c. Then, make a plot that shows the distance between molecules as a function of the concentration for concentrations ranging from nM to M.

(c) As an application of your thinking from part (b), explain what the concept of the "critical concentration" is for the polymerization of actin filaments. Then, provide a rough estimate of the mean spacing between actin monomers in a solution at the critical concentration.

2 Street fighting the ribosome

One of the most important molecular assemblies in the cell is the ribosome. The number of ribosomes per cell dictates how fast cells can grow. *E. coli* growing with a division time of 24 minutes have 72,000 ribosomes per cell, and slow growing *E. coli* with a division time of

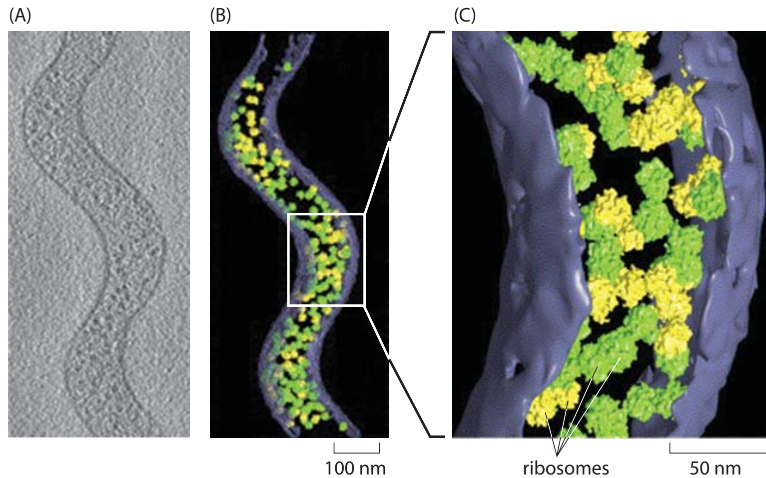

Figure 1: Cryo EM study of a bacterial cell. These images are of the tiny bacterium, *Spiroplasma melliferum*. Using algorithms for pattern recognition and classification, components of the cell such as ribosomes were localized and counted. (A) Single cryo-electron microscopy image. (B) 3D reconstruction showing the ribosomes that were identified. Ribosomes labeled in green were identified with high fidelity while those labeled in yellow were identified with intermediate fidelity. (C) Close up view that you should use to make your count. Adapted from JO Ortiz *et al.*, J. Struct. Biol. 156, 334-341 (2006).

100 minutes have a factor of ten fewer ribosomes with a count of $\approx$ 6800 ribosomes.

(a) In this part of the problem, we will use our street fighting skills to explore the ribosomal density in another organism as shown in Figure 1, and then see how well our results from the electron microscopy study square with the numbers quoted above. By examining the figure, make an estimate of the number of ribosomes per μm^3 and compare that result to the numbers quoted for *E. coli* above.

(b) In a beautiful turn of the millennium paper by Tania Baker and Stephen Bell whose abstract is shown in Figure 2, they imagined a world in which DNA polymerase was the size of a FedEx truck and explored what copying DNA would look like. Write a one-paragraph abstract of your own which carries out a similar analysis, but this time for the ribosome.

3 A feeling for the complete blood count (CBC) test.

Typical results for a complete blood count (CBC) are shown in Table 1. Assume that an adult has roughly 5 L of blood in his or her body. Based on these values estimate:

(a) the number of red blood cells.

Cell, Vol. 92, 295–305, February 6, 1998, Copyright ©1998 by Cell Press

Polymerases and the Replisome: Machines within Machines

Review

Tania A. Baker* and Stephen P. Bell
Department of Biology
*Howard Hughes Medical Institute
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Synthesis of all genomic DNA involves the highly coordinated action of multiple polypeptides. These proteins assemble two new DNA chains at a remarkable pace, approaching 1000 nucleotides (nt) per second in *E. coli*. If the DNA duplex were 1 m in diameter, then the following statements would roughly describe *E. coli* replication. The fork would move at approximately 600 km/hr (375 mph), and the replication machinery would be about the size of a FedEx delivery truck. Replicating the *E. coli* genome would be a 40 min, 400 km (250 mile) trip for two such machines, which would, on average make an error only once every 170 km (106 miles). The mechanical prowess of this complex is even more impressive given that it synthesizes two chains simultaneously as it moves. Although one strand is synthesized in the same direction as the fork is moving, the other chain (the lagging strand) is synthesized in a piecemeal fashion (as Okazaki fragments) and in the opposite direction of overall fork movement. As a result, about once a second one delivery person (i.e., polymerase active site) associated with the truck must take a detour, coming off and then rejoining its template DNA strand, to synthesize the 0.2 km (0.13 mile) fragments.

In this review we describe our current understanding of the organization and function of the proteins of the replication fork and how these complexes are assembled at origins of replication. Understanding the architecture of DNA polymerases is relevant to RNA polymerases as well, as the core of the polynucleotide polymerization machine appears to be similar for all such enzymes. In the discussion of the replisome, we particularly focus on features shared by the machinery from different organisms.

Polymerases: Template-Directed Phosphoryl Transfer Machines

Synthesis of the new DNA strands occurs as a result of a collaboration between the synthetic capacities of multiple polymerases. Two types of polymerases are required: primases, which start chains, and replicative polymerases, which synthesize the majority of the DNA (Kornberg and Baker, 1992). The replication fork, however, contains at least three distinct polymerase activities: a primase and a replicative polymerase for each of the two template strands. In *E. coli*, primase is a single polypeptide, and the replicative polymerase is a dimer of DNA polymerase (pol) III core and several accessory proteins that together form the pol III holoenzyme (reviewed in Marians, 1992; Kelman and O′Donnell, 1995). Similarly, phage T4 has one primase and one replicative polymerase that appears to function as a dimer (Alberts, 1987; Munn and Alberts, 1991). The situation in eukaryotic cells is slightly different (Stillman, 1994). The primase is in a tight complex with a DNA polymerase (pol α) and eukaryotic cells have two distinct replicative polymerases: polymerase δ (pol δ) and polymerase ϵ (pol ϵ).

All the replicative polymerases have one large subunit that contains the polymerase active site and, with the exception of pol α–primase, the same subunit or an associated polypeptide carries a proofreading $3' \rightarrow 5'$ exonuclease. The polymerase subunits also interact with proteins that dramatically influence their association with DNA. In *E. coli*, the replicative polymerase is found in a complex with proteins that control polymerase processivity; this holoenzyme, consists of 10 distinct polypeptides (Kelman and O′Donnell, 1995). In contrast, neither the T4 nor the eukaryotic polymerases copurify in a complex with the processivity factors (Alberts, 1987; Stillman, 1994). Therefore, these proteins are called accessory proteins rather than subunits (see Table 1).

Polymerase Architecture. The central feature of all the known polymerase structures is the existence of a large cleft comprised of three subdomains referred to as the fingers, palm, and thumb by virtue of the similarity of the structures to a half-opened right hand (Figure 1; polymerase structures are reviewed in Joyce and Steitz, 1994, 1995; Sousa, 1996). A diverse set of polymerases—

Figure 2: Abstract of a paper from Tania Baker where she maps the action of DNA polymerase onto human length scales to give a sense of its amazing properties. This parable is the basis of your own analysis of the ribosome. Adapted from Baker TA and Bell, SP Cell, Vol. 92, 295-305, February 6, (1998).

Test	Value
Red blood cell count (RBC)	Men: $\approx$(4.3–5.7) $\times$ 10^6 cells/μL
	Women: $\approx$(3.8–5.1) $\times$ 10^6 cells/μL
Hematocrit (HCT)	Men: $\approx$(39–49)%
	Women: $\approx$(35–45)%
Hemoglobin (HGB)	Men: $\approx$(13.5–17.5) g/dL
	Women: $\approx$(12.0–16.0) g/dL
Mean corpuscular hemoglobin (MCH)	$\approx$(26–34) pg/cell
MCH concentration (MCHC)	$\approx$(31–37)%
Mean corpuscular volume (MCV)	$\approx$(80–100) fL
White blood cell count (WBC)	$\approx$(4.5–11) $\times$ 10^3 cells/μL
Differential (% of WBC):	
Neutrophils	$\approx$(57–67)
Lymphocytes	$\approx$(23–33)
Monocytes	$\approx$(3–7)
Eosinophils	$\approx$(1–3)
Basophils	$\approx$(0–1)
Platelets	$\approx$(150–450) $\times$ 10^3 cell/μL

Table 1: Typical values from a CBC. (Adapted from R. W. Maxwell, Maxwell Quick Medical Reference, Tulsa, Maxwell Publishing Company, 2002.)

(b) the percentage in volume they represent in blood.

(c) their mean spacing.

(d) the total amount of hemoglobin in the blood.

(e) the number of hemoglobin molecules per cell.

(f) the number of white blood cells in the blood.

4 RNA Polymerase and Rate of Transcription.

One of the ways in which we are trying to cultivate a "feeling for the organism" is by exploring the processes of the central dogma. Specifically, I want you to have a sense of the number of copies of the key molecular players in the central dogma as well as the rates at which they operate. Further, I argue that it is critical you have a sense of *how* we know these numbers.

(a) If RNA polymerase subunits β and β' together constitute approximately 0.5% of the total mass of protein in an *E. coli* cell, how many RNA polymerase molecules are there per

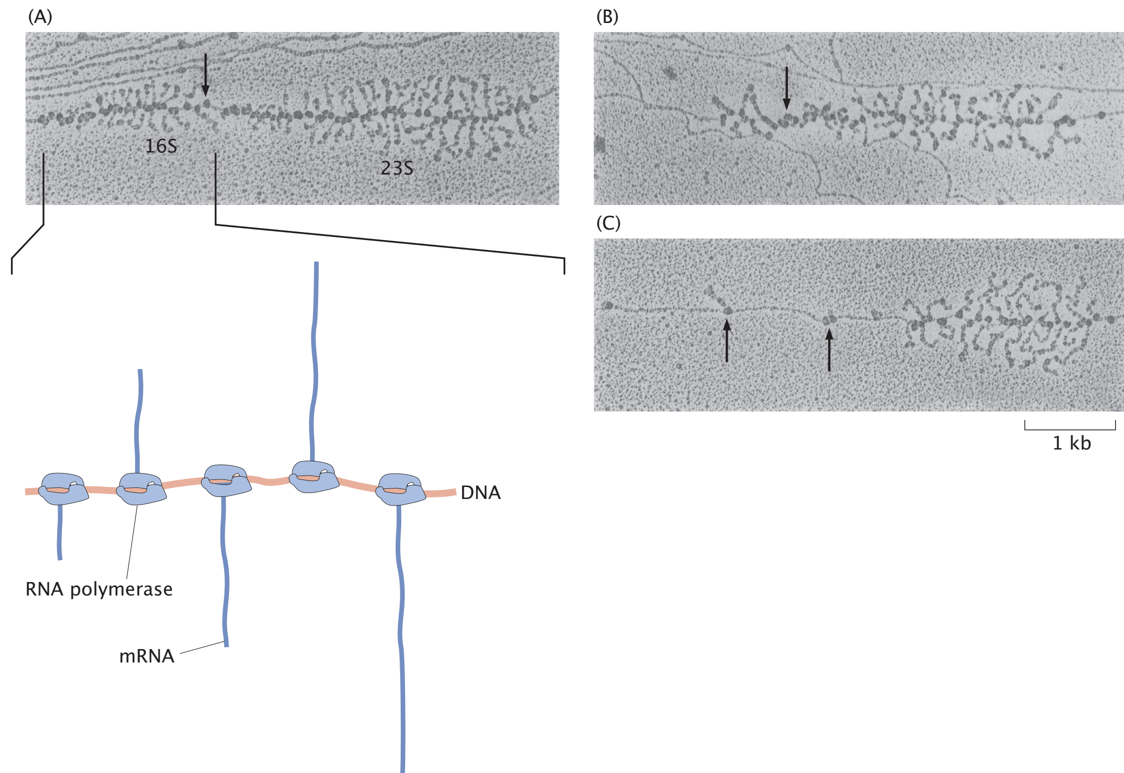

Figure 3: Effect of rifampin on transcription initiation. Electron micrographs of *E. coli* rRNA operons: (A) before adding rifampin, (B) 40 s after addition of rifampin, and (C) 70 s after exposure. No new transcripts have been initiated, but those already initiated are carrying on elongation. In parts (A) and (B) the arrow signifies the site where RNaseIII cleaves the nascent RNA molecule producing 16S and 23S ribosomal subunits. RNA polymerase molecules that have not been affected by the antibiotic are marked by the arrows in part (C). (Adapted from L. S. Gotta et al., *J. Bacteriol.* 20:6647, 1991.)

cell, assuming each β and β' subunit within the cell is found in a complete RNA polymerase molecule? The subunits have a mass of 150 kDa each. (Adapted from problem 4.1 of Schleif, 1993.)

(b) Rifampin is an antibiotic used to treat *Mycobacterium* infections such as tuberculosis. It inhibits the initiation of transcription, but not the elongation of RNA transcripts. The time evolution of an *E. coli* ribosomal RNA (rRNA) operon after addition of rifampin is shown in Figures 3(A)–(C). An operon is a collection of genes transcribed as a single unit. Use the figure to estimate the rate of transcript elongation. Use the beginning of the "Christmas-tree" morphology on the left of Figure 3(A) as the starting point for transcription.

(c) Using the calculated elongation rate estimate the frequency of initiation off of the rRNA operon. These genes are amongst the most transcribed in *E. coli*.

5 Real Estate for the Factories of ATP Synthesis

We are captivated by the tension between those things about living organisms that are universal and those things that are baroque and specific to a given organism. One of the nearly universal features of living organisms on our planet is their use of ATP hydrolysis as an energy source for a huge variety of processes. Further, as explained below, there is a nearly constant power density to run cellular life across the entire span of the tree of life. Where does all of this ATP come from? Cells have tiny molecular machines known as ATP synthase in the membrane which use an ion gradient to drive the 6000 rpm rotation of these machines to produce a few ATPs each rotation. However, the ATP is consumed within the volume of cells, but is produced on membranes. This leads to the possibility that as cells get bigger, there may be a point at which the surface area is insufficient to keep up with the demands of the cytoplasmic volume. Indeed, this problem explores the hypothesis that for cells above a certain size, the synthesis of ATP at the plasma membrane (such as in bacteria) no longer sufficed and that a new specialized energy factory (i.e. the mitochondria) was required.

(a) By considering the cost of protein synthesis for a dividing bacterium with a 1000 s division time, justify the assertion that the power usage is

$$\text{power density} = 10^6 \frac{\text{ATP}}{\mu\text{m}^3\ \text{s}}. \tag{1}$$

As a reminder, one way to do this estimate is to figure out how many proteins there are in an *E. coli* cell and how many amino acids there are per protein, noting that it costs 4 ATP equivalents for every polypeptide bond.

(b) As shown in Figure 4, compute the maximum radius of a spherical cell that could sustain the demands of ATP synthesis (i.e. the $10^6 \text{ATP}/(\mu\text{m}^3\ \text{s})$ required to run the cellular economy) by the presence of ATP synthases on its surface. Use your results to comment on the way prokaryotes and eukaryotes generate ATP and how large eukaryotes get around this conundrum.

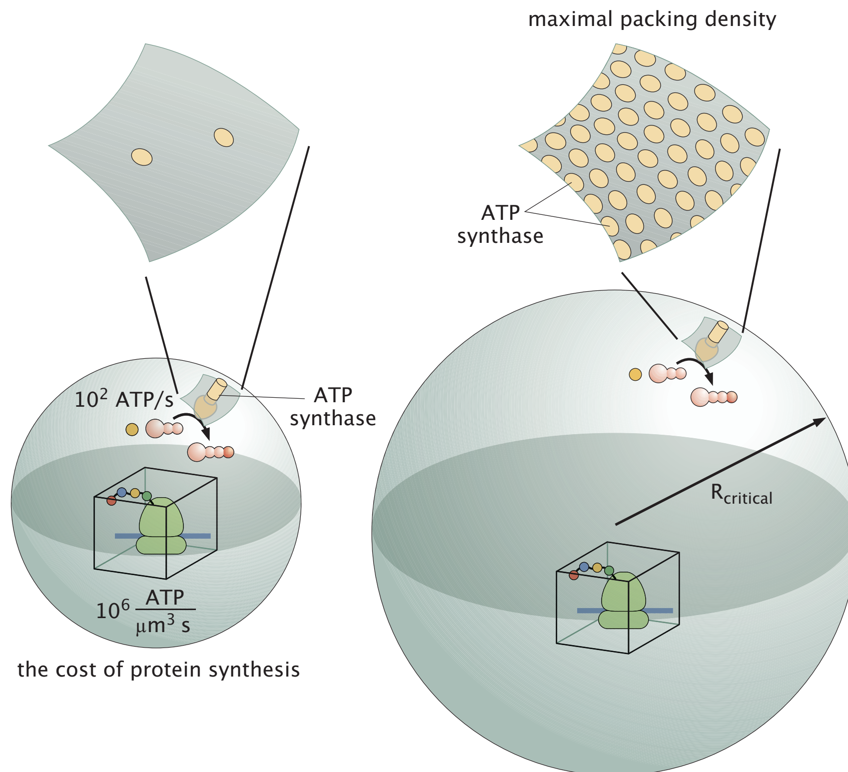

Figure 4: Surface coverage of bacterial cells with ATP synthase. For small cells, the demands of the cytoplasmic power consumption can be met by ATP synthases on the plasma membrane. However, for larger cells, there is not enough surface area to keep up with the demands of the power needs of the cellular interior.