The Production and Characterization of Amplified DNA by way of PCR

[redacted]

Abstract

Polymerase chain reaction, or PCR, is the method and machinery by which samples of DNA are cloned exponentially over many cycles. This finished sample can then be run through an electrophoresis gel in order to characterize the lengths of the original DNA sample. We built and gathered the chemicals and machinery necessary for this remarkable procedure, and will now discuss what is necessary for it to function, along with the results of our attempts at functioning it.

Amplifying DNA has a large breadth of uses, ranging from medical diagnosis to forensics.

DNA amplification owes its existence to a number of previous innovations in the field of biology. In 1957 the first polymerases were observed, paving the way to their being harnessed to artificially replicate DNA. Another seminal development was created by Klenow in 1970, when a portion of polymerase that was responsible for breaking down DNA in one direction was successfully inhibited, allowing the polymerase to naturally favor creating new DNA. The first high-temperature DNA was discovered in 1976 in the bacterium *Thermus aquaticus*, which would be important for allowing the PCR process to span multiple thermal cycles. By 1983, PCR was in its infancy, and by 1987 it was being aided by Taq, the polymerase isolated form *T. aquaticus*.

The fundamental principle of the polymerase chain reaction is exponential growth. By utilizing the fact that polymerases duplicate short strands of one-sided DNA, and allowing a single strand to double with every cycle of the PCR process, the amount of DNA grows by a factor of 2 every cycle as long as the necessary conditions for replication are met. A quick calculation shows that a mere 30 cycles could theoretically amplify starting DNA fragments by a factor of roughly 10^9.

The overall PCR process has three main phases: exponential amplification, leveling off, and plateau. In the exponential amplification phase, the process runs as an exponential as described in the previous paragraph. During the leveling off phase, the resources required by the reaction are entirely consumed, and during the plateau phase all the reactants are consumed. This plateau phase is generally included to ensure that the reaction runs its full course, since stopping the PCR process too early can result in greatly reduced yields of amplified DNA.

The PCR reaction has several components that go into producing and replication the DNA. Taq polymerase is needed to replicate the DNA strands at each cycle, and added to this are magnesium ions to improve efficiency. Deoxynucleoside triphosphates (dNTP's) are the components the polymerase uses to produce the DNA. Short DNA primers are needed to allow the polymerase to attach new bases to the DNA fragments, and a buffer solution keeps the system at a stable pH.

Each cycle of the reaction itself consists of three steps, each of which works at a different prerequisite temperature: denaturing, annealing, and replication.

The denaturing process takes place at 95°C, and consists of the DNA being heated to high enough temperatures to break the hydrogen bonds that told the two strands together. Figure 1 shows the proportion of DNA with respect to temperature. As can be seen in the figure, natural DNA will naturally be mostly melted those temperatures. Some PCR processes also require an additional 5-minute period at 95°C in order to prepare the polymerase for reaction.

The annealing step takes place at the significantly reduced temperature of 55°C. In this part of the cycle, the now separated strands of DNA are cooled enough that the primers in solution can attach to the DNA. Each segment of DNA to be amplified has segments on either end called flanking segments. The primers are specifically created to have sequences that match with the flanking structures of the two strands. Since the two strands of DNA run in opposite directions, two different primers are needed in the PCR process. The primers function as a place for the polymerase to attach dNTP to during the replication step.

In the replication step, the polymerase duplicates the strands of DNA already in solution. The process begins with the polymerase attaching to the exposed part of a primer, attached to the DNA strand during the previous step. The polymerase undergoes a conformal change, and dNTP's can enter the active site from solution. A "finger" on the enzyme checks each

0006-3495/08/09/2624/12 $2.00 N/A

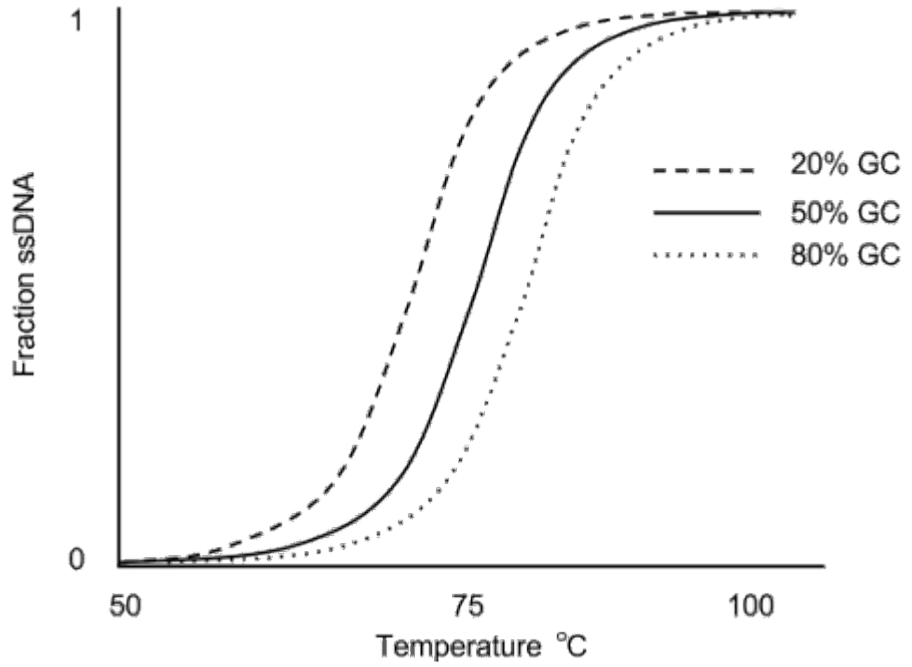

Figure 1: The proportion of melted DNA with respect to temperature. This figure shows the melting curves for diffrent concentrations of guanine-cytosine (GC) bonds in the DNA. Because a GC bond contains three hydrogen bonds instead of two, DNA high in GC bonds has a higner melting point. https://ka-perseus-images.s3.amazonaws.com/4ede5a3218abe1404fddd1d998c33db42aebe2c5.png

dNTP on the active site to see if it fits the base on the strand. Upon a match, the enzyme creates a bond between the first phosphorus atom in the dNTP and a free hydroxyl group on the previous nucleotide, extending the DNA backbone and adding a base to the strand.

Methods

We sought to build a basic PCR machine using fairly elementary components. The machine consisted of a transfer arm actuated by two servos, and three thermoses each containing a water bath of a different temperature. The arm was controlled by Arduino, and was used to cycle between the three baths to cycle between the reaction steps. The thermoses were filled with water at 95°C, 55°C, and 72°C, and were insulated with 11mL of canola oil, to allow the arm to enter and exit without issue. Temperatures in the thermoses were calculated by mixing boiling water and room-temperature (21-25°C) water using basic calorimetry.

The PCR machine spent 5 minutes in the hot bath before the cycling process began. After that, 35 complete cycles were executed. Each cycle consisted of 30 seconds in each bath, with time the sample spent in the air between the baths minimized.

Two variations of the reaction mixture were used, as recommended by Dr. Goodisman. One had a quadrupled concentration of taq polymerase, with the associated amount of water excluded to compensate. The reaction mixture used consisted of:

- water (96 μL)
- buffer solution (16 μL)
- $MgCl_2$ (aq) (16 μL)
- dNTP's (12 μL)
- primers (8 μL)
- taq polymerase (4 μL)*
- DNA (1 μL)

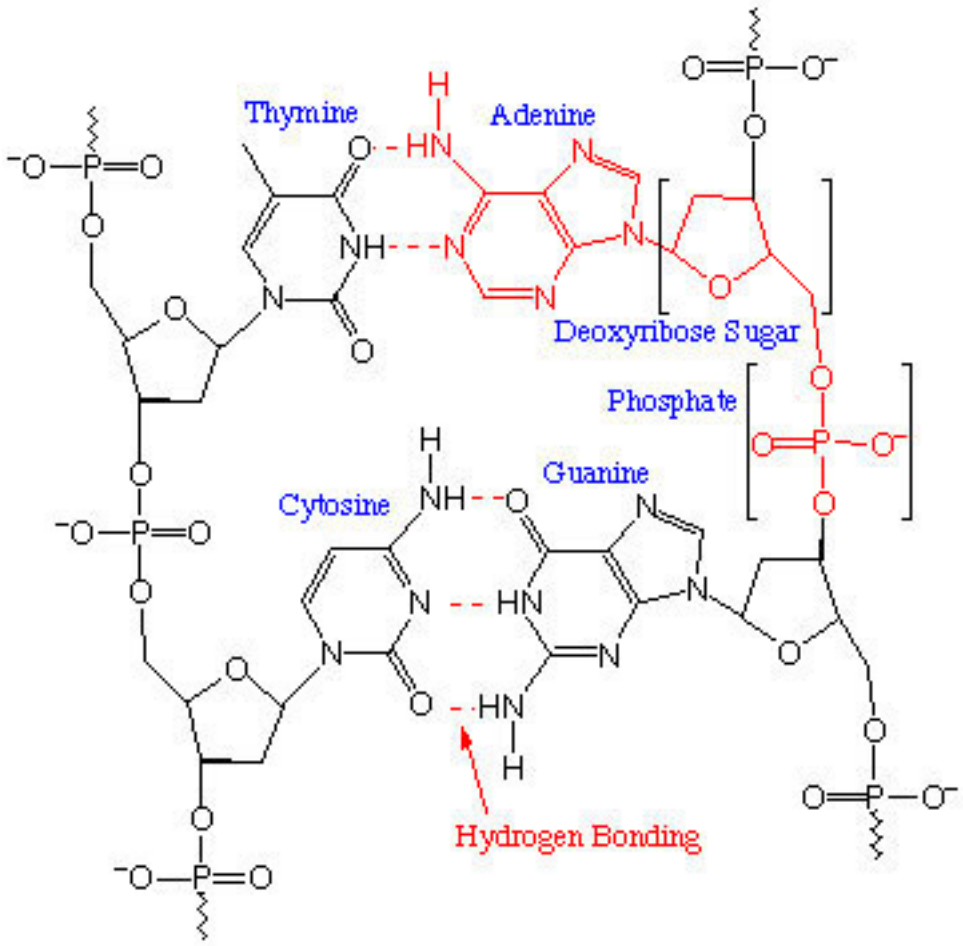

Figure 2: Here we see the bonding structure of DNA, where the phosphate molecules form the backbone of the DNA, where one can clearly see the single bonded oxygen as the source of negative charge

Characterization

The characterization of the DNA lengths of our sample was done by gel electrophoresis. This is the method in which we suspend amplified DNA samples in a thick gel, then run a voltage through it to cause the negatively charged DNA to move through it at a velocity dependent on its mass, allowing us to characterize its length.

The purpose of the gel provide a solid resistive force to the DNA being pushed along by the electric field, and slow the larger DNA. The gel uses a thick helical polymers bounded together by hydrogen bonds to create a porous medium for the DNA to move through.

We had multiple options for the gel. Namely, we could have used the polysaccharide agarose, or polyacrylamide as our helical polymers. The choice between the two is decided by the size of the DNA one wishes to characterize. Polyacrylamide binds at much smaller pore sizes, allowing for a more specific characterization by size of smaller DNA, but can restrict completely the movement of large DNA. This is in stark contrast to the agarose gel, which creates very large pores, thus allowing much larger DNA to pass, but poorly resolves size differences between small DNA. Ultimately, our test was to detect whether or not we successfully amplified DNA, so we wanted as much permeability to our DNA as possible. Thus, agarose was chosen.

As for the application of a voltage, we had two options as well. By a direct, linear electric field running from the cathode of the gel to the anode, or from the well to the opposite wall of the gel, we can very simply and intuitively reach our goal of separating DNA. The issue with this method comes with larger scales of DNA. Even with agarose, DNA on the length scale of tens of thousands of nucleotides in length have some trouble being resolved. The solution here is the method of pulsed field gel electrophoresis. By this method, the field is switched from angles of zero degrees, negative sixty, and positive sixty degrees, where the zero degree point is along the direct path to the opposite wall. This allows the much larger DNA to travel

over a greater distance, thus resolving their size differences more clearly. Ultimately, we expected our DNA to be on the length scales of hundreds of nucleotides, so we opted for the simple, linear method.

As for seeing the DNA, our gel contains a compound called ethidium bromide, which forms a complex with DNA that fluoresces under UV light. Running the electrophoresis over a UV light then allows us to see our DNA samples.

We will now examine the source of the velocity disparity between larger and smaller DNA. Examining figure 2, we can clearly discern the source of charge on the DNA. For every phosphate molecule that is bonded to a nucleoside through a sugar molecule, there is a free double bonded oxygen and a single bonded oxygen. This single bonded oxygen has a free electron that offers the DNA a charge quanta of +1 per nucleotide. The counter-intuition here is that larger DNA receive more resistive force, as they will have more nucleotides, and thus more mass, but also more Coulomb force. It is simple to show the model by which this functions.

The DNA will flow through the gel by way of Stokes law, calculated as

$$F = 6\pi\eta R_s v = fv$$

where η is the viscosity of the gel, R_s is the radius of the DNA modeled as a sphere, and v is the velocity of the DNA, and the Coulomb force supplied by the electric field we run along the gel, given as

$$F = qE$$

where q is the charge of the DNA, and E is our constant electric field. We expect these forces to be equal, as we do not expect acceleration from our samples, so, setting them equal

$$qE = fv$$

we define a quantity

$$U \equiv v/E = q/f$$

where U is the electrophoretic mobility. We express this as

$$U = \frac{q}{6\pi\eta R_s v}$$

We have many interdependent quantities here, where radius corresponds to length of the DNA which corresponds to increased nucleotide count which gives us our increased charge and mass. Ultimately, we model the interdependence of these quantities in this force equation as

$$logU = logU_o - K_r T$$

where U_o is the electrophoretic mobility of the DNA in a sucrose solution (experimentally determined),T is the total concentration of gel material (agarose or polyacrylamide), and K_r is the slope of a $logU$ vs T plot for the given DNA sample. K_r scales with the length of the DNA, so we see a formula for velocity of the DNA strands that is *inversely proportional* to the length, and thus mass, of the DNA. Thus, we expect the heavier DNA strands to stay at the back, and the lighter DNA strands to advance farther along the gel.

Now we just need a repeatable scale to run next to our samples to tell us what the lengths of the DNA are based on how far they've moved along the gel. For this, we will run a ladder in the gel made of specific lengths of DNA that we know in advance. We can do this by selecting a very small sample of 100 nucleotide long DNA, amplifying it, then keeping those samples in high concentration. These strands bind end to end with some probability, making a sample of DNA strands at steps of 100 nucleotides. With expected behaviour for our samples, and a scale for our results, we ran the amplified DNA in our gel to characterize our results.

Results

Unfortunately, after running the DNA through the PCR machine, and running electrophoresis for 10 minutes, we were unable to see our samples for DNA amplified under UV light, even though we saw the ladder. There are many possible explanations for what happened.

The cocktail of taq polymerase, primers, nucleotides, and DNA samples was used successfully by Dr. Goodisman, but could have been mixed incorrectly prior to amplification. During the amplification process, the thermoses were left open, and

not monitored for temperature. For the multiple steps, DNA that is not heated to the proper temperature will not denature, and thus not split to be primed, which must happen at a specific temperature, which allows the taq to add nucleotides to the DNA, which operates at an optimum temperature. This gives ~~enormous~~ room for error, and must be more thoroughly monitored.

Supplementary Material

The arduino code can be found at https://docs.google.com/document/d/1GVNttD8CCxJxihl8zzl0GXAv3$_zthg43tQA8iqVfVEA/edit?us$
$sharing$

References

1. Jakubowski, H., 2006. Experimental Techniques Electrophoresis
2. Wong, G., Wong I., Chan, K., Hsieh, K., Wong, S. 2015. A Rapid and Low-Cost PCR Thermal Cycler for Low Resource Settings
3. Schwartz, D. 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell 37 (1): 6775