

Good improvements in edits.

Polymerase Chain Reaction

Anonymous

April 23, 2005

Polymerase chain reaction (PCR) is a method of amplifying a "template" DNA sequence (usually known beforehand) *in vitro*, typically for the purposes of detection or measurement. I will begin by giving an overview of the PCR method itself, and then discuss a variation known as "Touchdown" PCR, designed to improve the results of PCR. I will then discuss PCR's limitations: it's accuracy is in part limited by the size of the sequence to be copied. As such, it is more easily used in scientific applications, though it has found several applications in domains with similar requirements to engineering.

1 Method

Basic PCR consists of three major steps, which are repeated 20-30 times. Each cycle roughly doubles the number of DNA strands with the sequence between the primers on the template; the overall effect is that the number of template strands is increased by a factor somewhere between a million and a billion.

PCR requires several components: a DNA template (to be copied), two primers which "bracket" the desired region of the template to be copied (one for the 3' end and one for the 5' end), DNA polymerase, dNTPs to build the new DNA from, and (obviously) a buffer in which this can all take place. The three steps to PCR are:

1. If the initial DNA template is double-stranded, it must be denatured by heating to 94-96 degrees Celsius. This takes approximately 1-2 minutes.
2. The temperature must then be lowered below the melting temperature of the primers (usually between 45 and 60 degrees), to allow the primers to anneal to the ssDNA. This also takes approximately 1-2 minutes.
3. Finally, the temperature is lowered to 25 degrees to allow the DNA polymerase to extend the primers along the template strands. The time for this step depends on which polymerase is used, and how long the template DNA strand is. A good rule of thumb is approximately 1 minute per kilobase[3].

After the appropriate number of cycles, the PCR product can be identified by its length using gel electrophoresis.

language in this part ₁ a bit too informal.

Because the temperature must be controlled accurately, PCR is performed in a thermal cycler. To prevent evaporation of the buffer and its contents, either a heated lid is placed over the mixture tubes, or a layer of oil is put on the surface of the mixture.

The first hurdle biologists had to overcome in PCR was finding polymerase which did not denature at the melting temperature of DNA. *Taq* polymerase, from the bacterium *Thermus aquaticus* was the first discovered thermostable DNA polymerase. Subsequently, *Pwo* and *Pfu*, from *Archaea*, and *Vent* or *Tli*, from *Thermococcus litoralis*, were discovered.

The next hurdle in PCR is that, due to the exponential nature of the amplification, an overwhelming number of products may reflect early errors. *Taq* has a 5' → 3' exonuclease activity, but does not have any 3' → 5' exonuclease, resulting in an error rate of 285×10^{-6} errors per base[4]. *Pwo*, *Pfu*, and *Vent* all have 3' → 5' exonuclease activity, resulting in much lower error rates (about five times lower for *Vent* than for *Taq*[6]). *Vent* or *Pfu* are often used for exactly this reason.

Vent also has a half-life about 5 times as long as *Taq*[1], and so is more useful for very long target sequences. Additionally, *Vent* comes in a *Vent(exo-)* form, without the 3' → 5' exonuclease, which allows for greater speed.

1.1 Primer selection

Selecting the wrong primers can have disastrous effects on PCR amplification, and the process of selecting the right primers is not as simple as one might imagine. There are two main factors to take into consideration: binding specificity and melting point¹. Again, because of the exponential amplification, nonspecific binding can ruin entire experiments.

In general, the longer the primer, the higher the binding specificity: a short sequence (or a small variation thereof) may appear many times in a sequence of DNA, but (usually) only one of these will be the actual target template.

Similarly, the longer the primer, the higher the melting point: a long sequence makes more hydrogen bonds (in base pairs) which must be broken for the primer to come unbound from the target DNA. However, above approximately 80 degrees Celsius, the polymerase is less effective, so the melting point should be less than 80 degrees.

These two concerns typically mean choosing a primer of an optimum length of 20-40 nucleotides, with a melting temperature between 60 and 75 degrees Celsius[3].

Finally, there are two other concerns which should be taken into account. The first is that the two primers should not bind to themselves (e.g. in a hairpin turn or a dimer) or to one another. The second is that the difference in melting temperatures between the two primers, and between the primers and the amplified product, should not be significant (say, less than 5-10 degrees).

Degenerate primers (a set of primers with variation in a few nucleotides) are sometimes used to amplify genes where the exact sequence of the template is not known. This

¹The melting point (for our purposes) is not the point at which the molecule disassociates, but rather the point at which only half of the primers in solution are bound to a complementary strand of DNA.

can happen when copying a gene based on its protein output (as several codons encode the same amino acid), or when amplifying the same gene from a number of species. However, this can have the obvious effect of decreasing binding specificity, so often another technique, called "Touchdown" PCR, is used to recover this loss of specificity.

1.2 Touchdown PCR[2]

Touchdown PCR is a method that was first reported in Nucleic Acids Research in 1991. It is a very simple modification on the standard PCR technique, and is designed to minimize erroneous amplification by effectively increasing binding specificity.

In typical PCR, the primers are allowed to bond to the target by lowering the temperature approximately 5 degrees below their melting point, which may allow for nonspecific binding. However, nonspecific bindings will have lower melting temperatures than exact matches. In touchdown PCR (after denaturing the target), the temperature begins at or above the melting temperature of the primers. In each successive PCR cycle, the temperature is lowered by a small amount, until the usual temperature at which PCR is performed is reached.

primer annealing in the early cycle is set

For example, in [2], they begin with an annealing temperature of 65 degrees. After every second cycle of PCR, they decrease the annealing temperature by 1 degree, until they reach a "touchdown" temperature of 55 degrees.

Each degree of difference between the melting temperature of a specific and a nonspecific binding gives approximately a 4-fold advantage to the specific binding when using touchdown PCR. So a difference of 5 degrees C gives roughly a 1000-fold advantage. The goal is for the nonspecific binding products to no longer overrun the correct products.

2 Applications and Limitations

PCR can be used in any situation in which something is to be determined based on a very small amount of DNA. Genetic fingerprinting is a classic example of this. PCR is also used in detecting hereditary disease and cloning genes. These are, essentially, engineering uses of the technique: they must be accurate, quick, and (relatively) cheap. The degree of accuracy required in these applications is incredibly high: often times, someones life depends on it. Because PCR is so fast, it would work well for large-scale engineering as well.

PCR is also widely used for scientific applications. For example, it can be used to help determine evolutionary pathways, or find a gene based on its protein product. In these cases, the final error tolerance is often much greater: often, as long as a clearly dominant band shows up in a gel, the experimenter has "succesfully" discovered what she was looking for.

Let us compare the scientific and engineering applications more concretely. Suppose we perform 20 cycles of PCR. If 95% accuracy is required, as might be the case in an engineering application, then each cycle must be at least 99.75% accurate. If 50%

accuracy is required, as might be the case in a scientific application, then each cycle must be at least 96% accurate.

But recall that the error rate of *Vent* polymerase is 57×10^{-6} errors *per base* [6]. Since the error rate of the polymerase is measured in errors per base, the percent accuracy of the PCR depends both on the size of the target and on the number of sequences being copied at each stage. Thus new errors are more likely to be introduced at later cycles of the PCR, but they will have less effect on the product. Let us examine how this affects the overall accuracy.

If we begin with 100 target strands and the target is one kilobase long, then we expect to end up with 190 good copies strands and 10-12 erroneous copies (i.e. about 95% accuracy). Of course, at this point, these "erroneous" copies only have one error per strand. For simplicity we will count any single error as making a strand "erroneous," and we will ignore the possibility that errors might get corrected by future errors. After the second cycle, we expect to have about 360 good copies and 40 erroneous copies (approximately 90% accuracy). Continuing this crude analysis for 20 cycles, we get approximately 56% final accuracy. (As I said, this analysis is very crude.)

So high fidelity application such as engineering are limited by both the size of the target and the degree of accuracy required. Scientific applications are much less sensitive to these variables. PCR is more easily used in scientific applications, but methods have been developed to achieve sufficient accuracy for some engineering applications, as mentioned above. The length of time required to achieve greater accuracy may limit the utility of PCR in mass-production of high fidelity product.

References

- [1] Austgen, L., Bowen, R. A., Rouge, M., contributors. "Thermostable DNA Polymerases," in *Biotechnology and Genetic Engineering*. Colorado State University, Colorado (Dec 1999): <http://arbl.cvmbs.colostate.edu/hbooks/genetics/biotech/enzymes/hotpolys.html>; (accessed 4/23/05).
- [2] Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., Mattick, J. S. " 'Touchdown' PCR to circumvent spurious priming during gene amplification." *Nucleic Acids Research* Vol. 19 (14): 4008 (Apr 1991).
- [3] "Polymerase Chain Reaction." *Wikipedia* (Apr 2005): http://en.wikipedia.org/wiki/Polymerase_chain_reaction; (accessed 4/23/05)
- [4] *Taq* DNA Polymerase with ThermoPol Buffer. *New England Biolabs, Inc.* <http://www.neb.com/nebecomm/products/productM0267.asp>; (accessed 4/23/05).
- [5] "Touchdown PCR." *Wikipedia* (Mar 2005): http://en.wikipedia.org/wiki/Touchdown_PCR; (accessed 4/23/05).

[6] Vent DNA Polymerase. *New England Biolabs, Inc.*
;http://www.neb.com/nebecomm/products/productM0254.asp; (accessed 4/23/05).