PCR Model 1

Ben Shirt-Ediss

Newcastle University UK: 2016

1 Reaction Set

$$S_1 + S_2 \underset{k_{10}}{\overset{k_1}{\rightleftharpoons}} DNA \tag{1}$$

$$S_1 + P_1 \underset{k_{20}}{\overset{k_2}{\rightleftharpoons}} S_1P_1 \tag{2}$$

$$S_2 + P_2 \underset{k_{20}}{\overset{k_2}{\rightleftharpoons}} S_2P_2 \tag{3}$$

$$S_1P_1 + E \underset{k_{30}}{\overset{k_3}{\rightleftharpoons}} S_1P_1(E) \tag{4}$$

$$S_2P_2 + E \underset{k_{30}}{\overset{k_3}{\rightleftharpoons}} S_2P_2(E) \tag{5}$$

$$S_1P_1(E) + nN \overset{k_e}{\rightarrow} DNA + E \tag{6}$$

$$S_2P_2(E) + nN \overset{k_e}{\rightarrow} DNA + E \tag{7}$$

$$E \overset{k_d}{\rightarrow} E_{denatured} \tag{8}$$

In the initial condition, there are exactly equal concentrations of both strands and primers: $S_1 = S_2 = \frac{1}{2}S_{total}$, $P_1 = P_2 = \frac{1}{2}P_{total}$. Reactions 2 and 3 are considered to have identical kinetics, as are reactions 4 and 5, and 6 and 7.

2 ODE Set

$$\dot{[S1]} = k_{10}[DNA] + k_{20}[S_1P_1] - k_1[S_1][S_2] - k_2[S_1][P_1]$$

$$\dot{[S2]} = k_{10}[DNA] + k_{20}[S_2P_2] - k_1[S_1][S_2] - k_2[S_2][P_2]$$

$$\dot{[DNA]} = k_1[S_1][S_2] + k_e[S_1P_1(E)][N] + k_e[S_2P_2(E)][N] - k_{10}[DNA]$$

$$\dot{[P1]} = k_{20}[S_1P_1] - k_2[S_1][P_1]$$

$$\dot{[P2]} = k_{20}[S_2P_2] - k_2[S_2][P_2]$$

$$\dot{[S_1P_1]} = k_2[S_1][P_1] + k_{30}[S_1P_1(E)] - k_3[S_1P_1][E] - k_{20}[S_1P_1]$$

$$\dot{[S_2P_2]} = k_2[S_2][P_2] + k_{30}[S_2P_2(E)] - k_3[S_2P_2][E] - k_{20}[S_2P_2]$$

$$\dot{[E]} = k_e[S_1P_1(E)][N] + k_e[S_2P_2(E)][N] + k_{30}[S_1P_1(E)] + k_{30}[S_2P_2(E)] - k_3[S_1P_1][E] - k_3[S_2P_2][E] - k_d[E]$$

$$\dot{[S_1P_1(E)]} = k_3[S_1P_1][E] - k_e[S_1P_1(E)][N] - k_{30}[S_1P_1(E)]$$

$$\dot{[S_2P_2(E)]} = k_3[S_2P_2][E] - k_e[S_2P_2(E)][N] - k_{30}[S_2P_2(E)]$$

$$\dot{[N]} = -n\left\{k_e[S_1P_1(E)][N]\right\} - n\left\{k_e[S_2P_2(E)][N]\right\}$$

$$\dot{[ED]} = k_d[E]$$

3 Mass Conservation Equations

PCR is a closed reaction system. Mass conservation equations are useful to check the ODE equations are correct, and that they are being integrated correctly. The following mass conservation equations apply at all times.

- Nucleotide conservation. The total number of nucleotides in the system does not change: only their arrangement changes. If amplicons are 100 base pairs, and primers are 20 bases ($n = 80$):

$$N_0 = N + 20(P_1 + P_2) + 100(S_1 + S_2) + 120(S_1P_1 + S_2P_2 + S_1P_1(E) + S_2P_2(E)) + 200DNA$$

- Polymerase conservation. Polymerase is either free-floating, or in a tenary complex, or denatured:

$$E_0 = E + S_1P_1(E) + S_2P_2(E) + E_{denatured}$$

(In the model, the amount of ssDNA should always increase (S1 and S2). But this does not necessarily mean that the DNA level will increase, as S1 and S2 can be bound up with primers, if these exist in excess.)

4 Melting Behaviour

Temperature changes are considered instantaneous in the model, but the temperature cycle is not fixed: it can be varied.

Reactions 1-5

The van't Hoff equation is used to set the equilibrium constants for Reactions 1-5, based on the Kelvin temperature T:

$$ln\ K_{eq} = -\frac{\triangle H^0}{RT} + \frac{\triangle S^0}{R}$$

$$K_{eq} = e^{\left[-\frac{\triangle H^0}{RT} + \frac{\triangle S^0}{R}\right]}$$

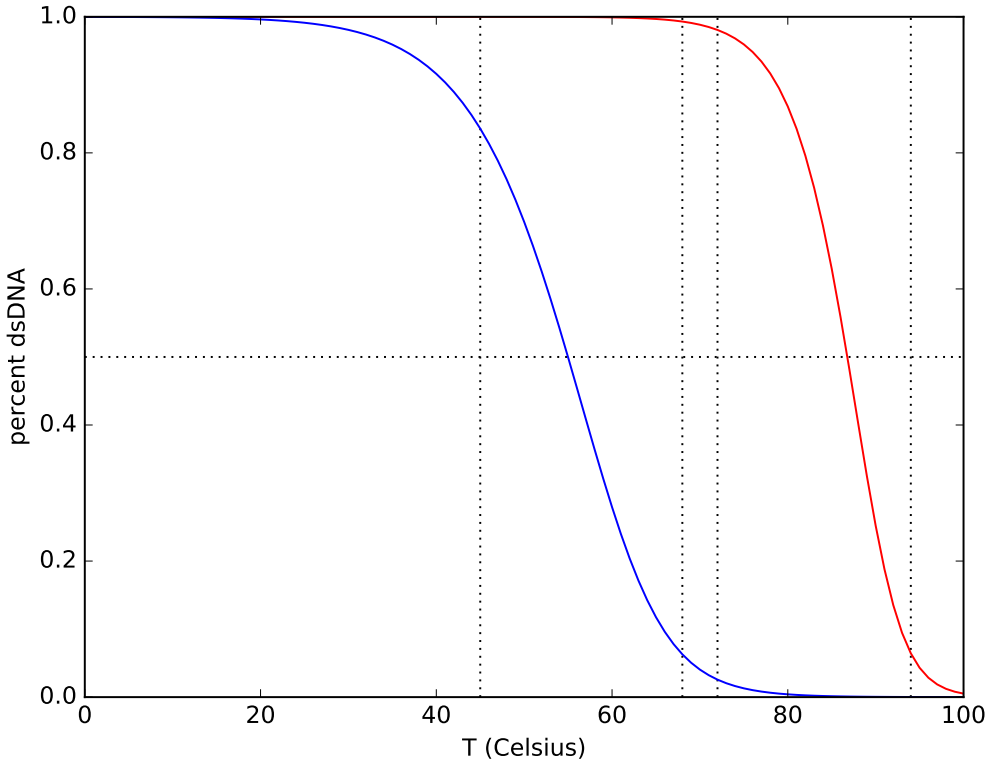
$\triangle H^0$ and $\triangle S^0$ are set for each reaction, such that a melt curve with the required melting temperature T_m is produced (see document on melting curve derivation). Here, it is assumed that a two-part model is sufficient to model DNA melting. The total single strand concentration used to calculate the below melt curves is 200nM, the default used on the NEB Tm calculator: http://tmcalculator.neb.com/#!/ when the product group is "Taq DNA polymerase".

The melting thermodynamic parameters assumed for DNA and primers are shown below:

Reaction		ΔH^0(cal/mole)	ΔS^0(cal/mole K)	T_m	Curve
1	DNA Melting	-120000	-300	86C	Red
2	Primer Melting	-56900	-140	55C	Blue
3	SP(E) Complex Melting	-56900	-140	55C	Blue

These values give (1) a high melt temperature for DNA, and a steep melt curve. This ensures that amplification cannot occur unless a high temperature denaturation phase is present (50C or 72C denaturation temperature is too low, and results in zero DNA amplification). Also (2) the primer melt curve is more gradual, so that SP complexes still exist in the elongation phase, and thus this phase is not useless. In Model 1.1 I made elongation 2-part, and so a shallow primer melt curve isn't needed.

Also Note: The melt behaviour of the SP(E) complex is assumed the same as the melt behaviour of the SP complex. It may be possible to specify more accurate melting behaviour for the SP(E) complex by using data in this publication: Datta & LiCata 2003, "Thermodynamics of the binding of Thermus aquaticus DNA polymerase to primed-template DNA".



Reactions 6, 7, 8

These reactions are irreversible, and their temperature dependence is modelled qualitatively as detailed below.

5 Kinetics

Reactions 1-5

Reactions 1-5 have approximate dynamics. As a first rough approximation, in each case, the forward constant is set to a fixed value (and is not assumed to changed with temperature):

$$k_{forward} = 10^6 \qquad M^{-1}s^{-1}$$

This value is consistent with that quoted in Mehra 2005, Gevertz 2005 and Dannenberg 2015. And then reverse constant is calculated as so that:

$$\frac{k_{forward}}{k_{reverse}} = K_{eq}$$

So, for Reaction 1, the reverse constant is calculated as follows:

$$k_{10} = \frac{k_1}{e^{\left[-\frac{\Delta H_0^1}{RT} + \frac{\Delta S_0^1}{R}\right]}}$$

Reactions 6, 7

This reaction is considered as a single irreversible step for simplicity. Making the reaction multi-step requires an advanced kinetic model that can also model how e.g. the partly-elongated SP(E) complexes break up as the temperature changes. Having a single reaction has the advantage that the SP(E) complex is either in its initial state, or fully elongated, and melt curves for each of these two states exist (i.e. melt curves of Reaction 3 and Reaction 1, respectively).

Considering Reaction 4 as a single-step reaction raises the question of rate dependence on the concentration of SP(E) and N, as the reaction is clearly not elementary (as up to 80 dNTPs are consumed over the course of the reaction!). As a rough approximation[1], reaction rate is assumed to depend in the first order on both SP(E) and dNTP concentrations:

$$rate = k_e[SP(E)][N]$$

Constant k_e is calculated by reading off from a qualitative Arrhenius plot, to give the elongation reaction not only a concentration dependence but also a temperature dependence. A log-normal distribution reflected in the y-axis is (arbitrarily) chosen to model the dependence of the rate constant k_e with temperature:

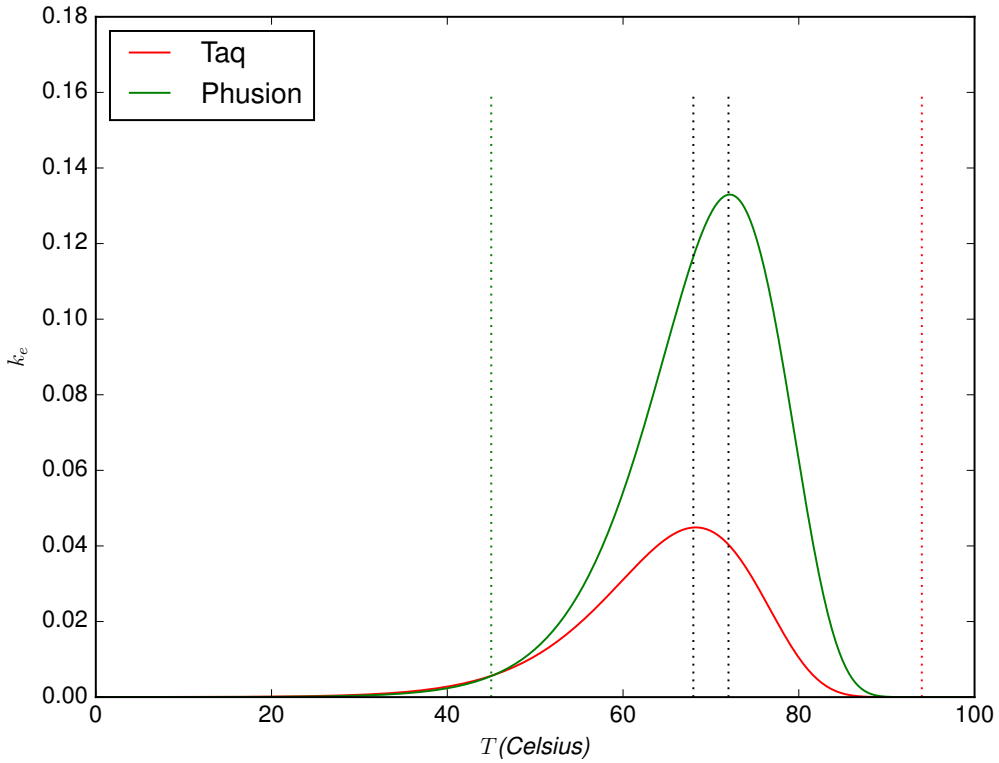
$$k_e = HV\frac{1}{S\sqrt{2\pi}(100 - T_C)}e^{-\frac{(ln(100-T_C)-M)^2}{2S^2}}$$

Where T_C is the temperature in Celsius. The V, M and S parameters are set differently for the Taq and Phusion polymerases such that:

- Taq has an optimal elongation temperature of 68C
- Phusion has an optimal elongation temperature of 72C
- Phusion elongates 3 times faster than Taq, when the enzymes are at their optimal temperature. This is based on the rough statement that "Taq elongates at 1kb per min, and Phusion elongates at 1kb every 20 seconds" (about 3kb per minute). The constant V maintains the 3x ratio between curve heights, and then the constant H scales the height of both curves at this ratio.

	Taq	Phusion
H	same	same
V	1	2.6
M	3.53	3.4
S	0.27	0.27

[1] In reality, the rate of DNA production depends on the SP(E) and dNTP concentration, but also on the distribution of intermediate elongation states of SP(E).



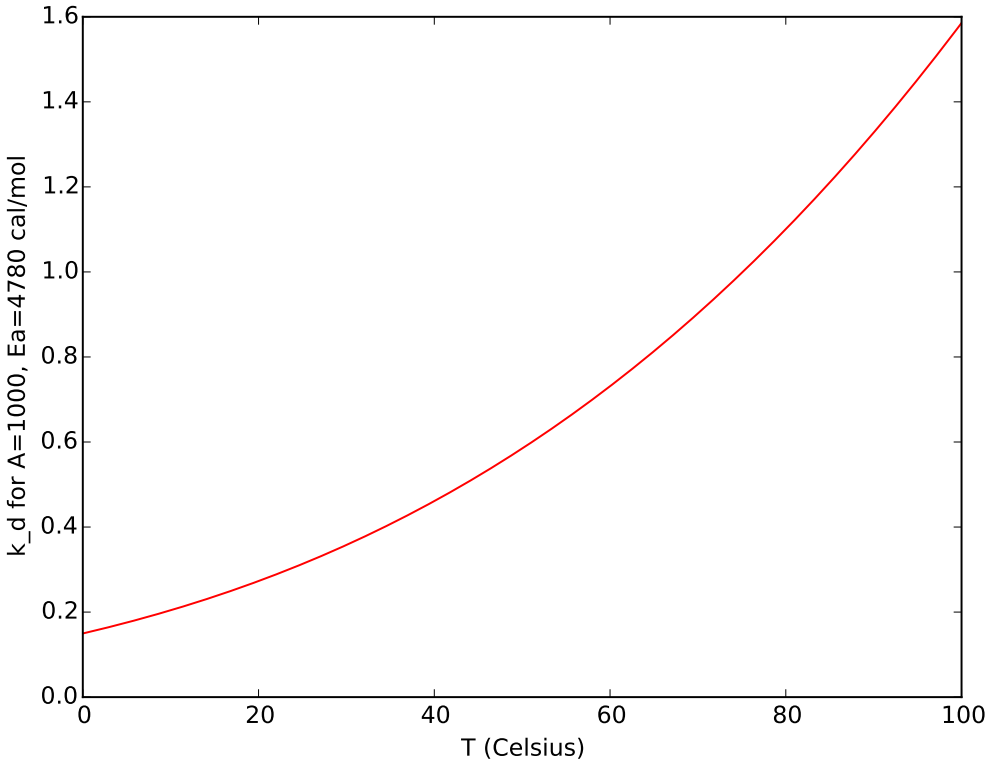
Note: The speed of the polymerase has little or no effect on amplification in Model 1. This is because many SPE complexes are elongated to 100% (in 1 step) in the annealing phase. In the elongation phase, there are alot fewer SPE complexes around, too few that enzyme elongation speed has an effect.

Reaction 8

There is a basal decay rate of the polymerase, which increases in the high temperature phases. Therefore, constant k_d has Arrhenius-like behaviour. It is calculated by

$$k_d = Ae^{-\frac{E_a}{RT}} \qquad s^{-1}$$

Where E_a is the activation energy barrier from a functioning polymerase state, to a denatured polymerase state. Both the Taq and Phusion polymerases are assumed to have the same denaturation curve.



On a forum (http://www.bio.net/bionet/mm/methods/1995-June/029361.html) I found a rough estimate of how a polymerase decays (it does not say which polymerase):

"Given that the normal half life for the enzyme is of the region of >40 cycles of 1min 96C, 1min 55C, 1min 72C"

If parameter A is set to some preset value, the value of E_a can be varied until after 40 cycles of the above, we end up with half the amount of enzyme we started with.