

12-2008

# A Fundamental Study of the PCR Amplification of GC-rich DNA Templates

T. G. Mamedov

*University of Nebraska - Lincoln*

Elsje Pienaar

*University of Nebraska - Lincoln, epienaar2@unl.edu*

Scott E. Whitney

*University of Nebraska - Lincoln, swhitney4@unl.edu*

Joel R. TerMaat

*University of Nebraska - Lincoln*

G. Carvill

*University of Cape Town, South Africa*

*See next page for additional authors*

Follow this and additional works at: <http://digitalcommons.unl.edu/cbmeviljoen>



Part of the [Chemical Engineering Commons](#)

---

Mamedov, T. G.; Pienaar, Elsje; Whitney, Scott E.; TerMaat, Joel R.; Carvill, G.; Goliath, R.; Subramanian, A.; and Viljoen, Hendrik J., "A Fundamental Study of the PCR Amplification of GC-rich DNA Templates" (2008). *Hendrik J. Viljoen Publications*. 13.  
<http://digitalcommons.unl.edu/cbmeviljoen/13>

This Article is brought to you for free and open access by the Chemical and Biomolecular Research Papers -- Faculty Authors Series at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Hendrik J. Viljoen Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

---

**Authors**

T. G. Mamedov, Elsje Pienaar, Scott E. Whitney, Joel R. TerMaat, G. Carvill, R. Goliath, A. Subramanian, and Hendrik J. Viljoen

# A fundamental study of the PCR amplification of GC-rich DNA templates

T. G. Mamedov,<sup>a</sup> E. Pienaar,<sup>a</sup> S. E. Whitney,<sup>a</sup> J. R. TerMaat,<sup>a</sup> G. Carvill,<sup>b</sup>  
R. Goliath,<sup>b</sup> A. Subramanian,<sup>a</sup> and H. J. Viljoen<sup>a, c</sup>

<sup>a</sup>Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-0643, U.S.A.

<sup>b</sup>Department of Microbiology, University of Cape Town, South Africa

<sup>c</sup>Department of Internal Medicine, section of Infectious Diseases, 985400 University of Nebraska Medical Center, Omaha, NE 68198, U.S.A.

Corresponding author – H. J. Viljoen, tel 402 472-9318, fax 402 472-6989, email [hviljoen1@unl.edu](mailto:hviljoen1@unl.edu)

## Abstract

A theoretical analysis is presented with experimental confirmation to conclusively demonstrate the critical role that annealing plays in efficient PCR amplification of GC-rich templates. The analysis is focused on the annealing of primers at alternative binding sites (competitive annealing) and the main result is a quantitative expression of the efficiency ( $\eta$ ) of annealing as a function of temperature ( $T_A$ ), annealing period ( $t_A$ ), and template composition. The optimal efficiency lies in a narrow region of  $T_A$  and  $t_A$  for GC-rich templates and a much broader region for normal GC templates. To confirm the theoretical findings, the following genes have been PCR amplified from human cDNA template: ARX and HBB (with 78.72% and 52.99% GC, respectively). Theoretical results are in excellent agreement with the experimental findings. Optimum annealing times for GC-rich genes lie in the range of 3–6 s and depend on annealing temperature. Annealing times greater than 10 s yield smeared PCR amplified products. The non-GC-rich gene did not exhibit this sensitivity to annealing times. Theory and experimental results show that shorter annealing times are not only sufficient but can actually aid in more efficient PCR amplification of GC-rich templates.

**Keywords:** PCR, annealing time, mathematical model, GC-rich

## 1. Introduction

PCR is widely used in either diagnostic or molecular analysis of DNA and RNA. However, amplification of templates with a high-GC content using PCR is usually difficult compared to non-GC-rich targets (McDowell *et al.*, 1998). Although only a small percentage (~3%) of the human DNA is GC-rich, important regulatory domains including promoters, enhancers, and control elements consist of GC-rich cis-elements (Wilson *et al.*, 1997). Most housekeeping genes, tumor-suppressor genes, and approximately 40% of tissue-specific genes contain high-GC sequences in their promoter region, making their DNA less amenable to amplification.

To overcome the problems associated with the amplification of GC-rich genes (and/or using GC-rich primers), several approaches have been developed. Organic molecules such as dimethyl sulfoxide (DMSO), glycerol, polyethylene glycol, formamide, betaine, 7-deaza-dGTP, and dUTP have been included in the reaction mixture and have been shown to improve the amplification of GC-rich DNA sequences (Baskaran *et al.*, 1996; Chakrabarti and Schutt, 2001; Henke *et al.*, 1997; Kang *et al.*, 2005; Musso *et al.*, 2006; Mutter and Boynton, 1995; Pomp and Medrano, 1991; Sidhu *et al.*, 1996; Sun *et al.*, 1993; Turner and Jenkins, 1995; Weissensteiner and Lanchbury, 1996). Recently, a combination of either two (betaine and DMSO) or three additives (betaine, DMSO, and 7-deaza-

dGTP) has been found to enhance amplification of long PCR products (Kang *et al.*, 2005; Musso *et al.*, 2006). These additives resolve the complex secondary structure formation at GC-rich regions (Henke *et al.*, 1997; Musso *et al.*, 2006; Ralser *et al.*, 2006). It has been suggested that betaine affects the extension reaction either by binding to AT pairs in the major groove (Rees *et al.*, 1993) or by increasing the hydration of GC pairs by binding within the minor groove and thus destabilizing GC-rich DNA (Mytelka and Chamberlin, 1996). It was shown that template denaturation with NaOH, hot start PCR, stepdown PCR, and primer modification can also improve PCR amplification of highly GC-rich regions of DNA (Agarwal and Perl, 1993; Sahdev *et al.*, 2007). These efforts to amplify GC-rich regions are useful in some but not all cases and highlight the need to move from an empirical to a more fundamental approach to the problem.

This study was motivated by difficulties in amplifying a 660 bp fragment of the human ARX gene from genomic DNA using conventional PCR protocols, with certain additives. It was decided to develop a theoretical model of the annealing step, in the hope that the model may provide some guidance. The theory, included in this study, is based on competitive binding of primers at incorrect sites. Ideally designed primers bind uniquely at the correct binding site. In practice, mispriming occurs and three contributing events are considered. First, primers may anneal at incorrect sites, the probability of such

an occurrence depends on the difference between the melting rates of primers at correct vs. incorrect sites. Secondly, polymerase molecules may bind to annealed primers, including ones at incorrect sites, to further stabilize the complexes. Finally, the complexes begin to elongate, albeit at reduced rates (at the annealing temperature), and further stabilize the ds-DNA. Therefore the formation of primer/template/polymerase complexes at incorrect sites must be minimized. The model results provided a possible solution—the use of shorter annealing times should minimize incorrect product yield.

It is well known that GC content influences both optimal annealing temperatures and primer specificity, and that annealing occurs quickly: “It is critical that the primers anneal stably to the template... For primers with high-GC content, higher annealing temperatures may be necessary... [T]he time for this step is based mainly on the time it takes to reach the proper temperature, because the primers are in such excess that the annealing reaction occurs very quickly” (Kramer and Coen, 2001). However, the terms of this general knowledge have not yet been defined theoretically. The theory developed here shows, and experimental results confirm, that shorter annealing times are not only sufficient but also necessary for efficient PCR amplification when GC-rich templates or primers are being used. The implementation of the theoretical solution and the experimental results are reported in this study.

## 2. Materials and Methods

### 2.1. Instrumentation

PCR was carried out in a PCRJet<sup>®</sup> thermocycler (Megabase Research Products, Lincoln, NE). With the PCRJet<sup>®</sup> thermocycler, there is a ~1 s delay for heat to conduct through the entire PCR sample (unpublished results). Thus, actual hold times can be up to 1 s less than the programmed hold times in the thermocycler. For example, with the PCRJet<sup>®</sup> set at 3 s annealing time, the outer regions of the PCR sample will anneal for 3 s while the inner PCR sample will anneal for ~2 s. All hold times reported in this work are the programmed hold times.

### 2.2. PCR Amplification of GC-Rich ARX and Normal GC $\beta$ -Globin (HBB)

All PCR amplifications were performed with KOD Hot-Start polymerase (Novagen, Madison, WI). PCR was carried out in 25  $\mu$ l reaction volumes containing 100 ng genomic DNA (Roche) templates, 200  $\mu$ M of each dNTP, 4 mM MgSO<sub>4</sub>, 400  $\mu$ g/ml non-acetylated BSA, 0.75  $\mu$ M of each primer for ARX, 0.25  $\mu$ M of each primer for HBB, 0.5 units KOD Hot-Start polymerase and 1 $\times$  manufacturer’s buffer (Novagen, Madison, WI) in presence (ARX) or absence (HBB) of 11% DMSO (v/v). The sequences of both forward and reverse primers, size of the amplified gene product, and the GC content of the gene products amplified in this paper are listed in Table 1. The primers were designed based on the genomic sequence deposited in GenBank under accession number of NC\_000023.

The PCR conditions were: 30 s hot start at 94 °C followed by 35 or 38 cycles of 94 °C for 2 s, 56–64 °C for 3–20 s, and 72 °C for 4 s, and final extension at 72 °C for 30 s.

*Remark:* The KOD Hot-Start polymerase has an average extension rate of near 200 nucleotides per second (Griep *et al.*, 2006b) and the 4 s extension time is more than sufficient to complete extension, as confirmed by experiments (results not shown).

## 3. Results and Discussion

### 3.1. Amplification of ARX Gene

Human genomic DNA was used as initial template for the amplification of a 660 bp ARX gene fragment with a GC content of 78.72%. The theoretical model provided guidelines for annealing conditions. PCR reactions were optimized for two parameters: (1) the annealing times were varied in the range of 1–20 s at constant temperature and (2) the annealing temperatures were varied whilst the annealing time remained constant.

In Figure 1A–C gel electrophoresis results of ARX amplification are shown for varying annealing times at constant temperatures (58, 60, and 62 °C, respectively). The following observations are made. Qualitatively, the outcomes at each of the temperatures are similar; at longer annealing times increasing smear occurs. Quantitatively, at an annealing temperature of 58 °C (Figure 1A), the initial appearance of a smear, albeit it faint, occurs at annealing times as short as 5 s. This observation should be compared to the results at an annealing temperature of 60 °C (Figure 1B) and 62 °C (Figure 1C), where the appearance of a distinguishable smear only starts at  $t_A = 6$  and 9 s, respectively. Also note that at 60 °C, the yield increases between  $t_A = 3$  and 4 s. At higher annealing temperatures (cf. Figure 1C), one notes the formation of fewer incorrect products (compare  $t_A = 9$  s at all three temperatures).

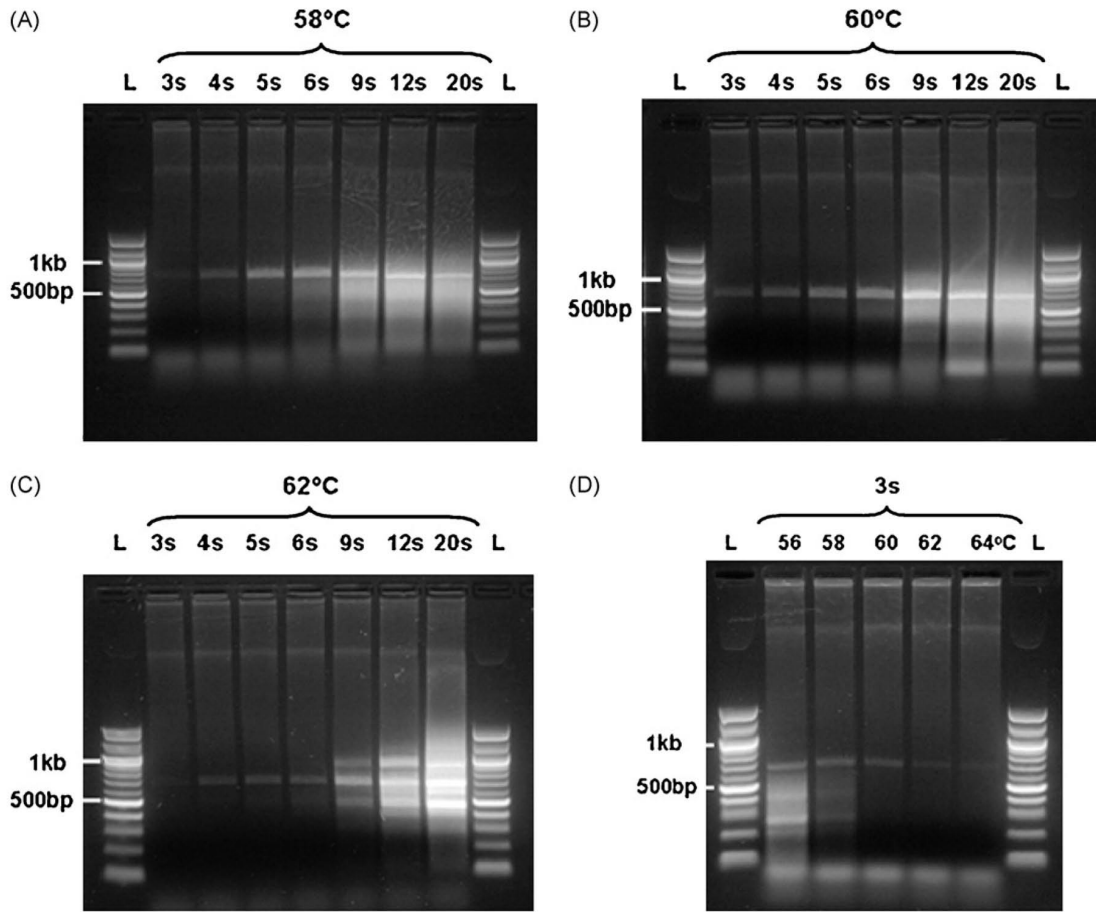
Figure 1D shows the results for amplification of the ARX gene at a constant annealing time of 3 s and varying annealing temperatures between 56 and 64 °C. A unique, specific band with the expected size of 660 bp, confirmed by DNA sequencing, is present at 60 °C annealing temperature and 3 s annealing time. At temperatures below 60 °C the gel products contain smears—typical of competitive binding. At temperatures above 60 °C the yield of the 660 bp product decreases.

### 3.2. Amplification of $\beta$ -Globin Gene

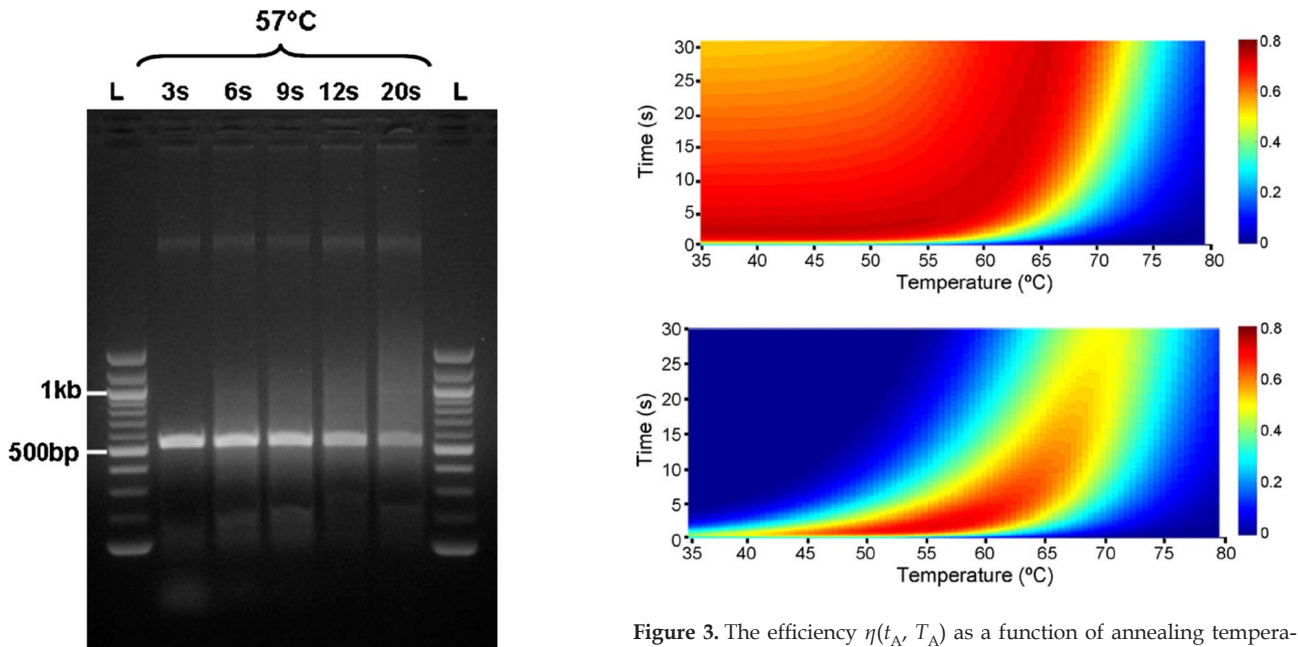
Figure 2 shows PCR amplification results of a 52.99% GC, 536 bp, fragment of the human  $\beta$ -globin gene. Compared to the ARX amplification results (Figure 1A–C), one notes the similar trend of increased smear with longer annealing times. The  $\beta$ -globin gene, however, appears less sensitive to annealing times—even at  $t_A = 20$  s the background smear is much less prominent than the cases shown in Figure 1. These observations are consistent with the theoretical predictions for low-competitive binding (cf. Figure 3A in Section 4.2 below).

Table 1. Forward and reverse primer specifications for human genes.

Gene	Size (bp)	GC content (%)	forward primer	reverse primer
ARX	660	78.72	5-CCAAGGCGTCGAAGTCTG-3	5-TCATCTTCTTCGTCCTCCAG-3
HBB	536	52.99	5-GCTCACTCAGTGTGGCAAAG-3	5-GGTTGGCCAATCTACTCCCAGG-3



**Figure 1.** Gel electrophoresis results of the PCR amplification of the ARX gene. (A)–(C) Varying annealing times at a constant annealing temperatures of 58, 60, and 62 °C, respectively. (D) Varying annealing temperatures at constant annealing time of 3 s. DNA markers shown in L lanes are QuickLoad 100 bp DNA ladders (Ipswich, MA).



**Figure 2.** Gel electrophoresis of PCR product of  $\beta$ -globin gene. Lanes marked 'L' are 100 bp QuickLoad DNA ladder. Other lane labels indicate varying annealing times at annealing temperature of 57 °C.

**Figure 3.** The efficiency  $\eta(t_A, T_A)$  as a function of annealing temperature and time for (A) a case of low competition and (B) the ARX gene (high competition). The primer and polymerase binding rates are  $p = 72 \text{ s}^{-1}$ , and  $s = 2.4 \text{ s}^{-1}$ , respectively. Color bars indicate values of annealing efficiency,  $\eta$ .



#### 4. Mathematical Model of Annealing

At the onset of the annealing stage in PCR, the target DNA molecules are separated into 3'-5' and 5'-3' single strands. Complementary forward and reverse primers are designed to bind uniquely to the respective single strands. The efficiency of annealing, and hence the overall PCR process, is compromised by competitive annealing events. Annealing is a dynamic process; primers bind and melt, but the average time a template site is annealed to a primer depends on the stability of the primer/template bond at the annealing temperature. Ideally the primer/template bonds at the target sites are very stable and all other binding sites are very unstable. GC-rich regions in either the primers and/or the DNA template increase the stability of undesirable binding sites; therefore, non-ideal primer/template bonds play a greater role in GC-rich PCR. If a primer is annealed to the template, DNA polymerase may bind to its 3' end to form a ternary complex and further stabilize the complex.

##### 4.1. Formulating the Model

Let the fraction of templates which have primers annealed at the  $i$ th binding site be denoted by  $X_i$ . Also define  $X_{pi}$  as the fraction of the templates that have formed a ternary complex with the primer and DNA polymerase. Assuming excess primer and polymerase, the balancing equations for  $X_i$  and  $X_{pi}$  are

$$\frac{dX_i}{dt} = p[1 - X_i - X_{pi}] - m_i X_i - sX_i, \quad (1)$$

$$\frac{dX_{pi}}{dt} = sX_i. \quad (2)$$

The rate of primer annealing is  $p$  and the rate constant of melting at the  $i$ th binding site is  $m_i$ . The ternary complex forms at a rate  $sX_i$ . The ternary complex is stable at the annealing conditions; in other words, the reaction is assumed to be irreversible. The solution of Eqs. (1) and (2) is trivial, noting that the initial conditions are  $X_i(0) = X_{pi}(0) = 0$ :

$$X_{pi}(t) = 1 + \frac{n_{2i} e^{n_{1i}t}}{n_{1i} - n_{2i}} - \frac{n_{1i} e^{n_{2i}t}}{n_{1i} - n_{2i}}, \quad (3)$$

where

$$n_{1i} = \frac{-(p + m_i + s) - \sqrt{(p + m_i + s)^2 - 4ps}}{2},$$

$$n_{2i} = \frac{-(p + m_i + s) + \sqrt{(p + m_i + s)^2 - 4ps}}{2} \quad n_{1i} < n_{2i} < 0.$$

Note: The rate constants  $n_{1i}$ ,  $n_{2i}$  depend on annealing temperature ( $T_A$ ) (see discussion on melting rates in Section 4.1.3).

##### 4.1.1. Annealing Efficiency

Let  $i = 1$  be the location of the intended primer binding site. If no competitive binding occurs, then the efficiency of annealing equals the fraction of templates which are annealed after an annealing period  $t_A$ :

$$\eta = X_{p1}(T_A, t_A) = 1 + \frac{n_{21} e^{n_{11}t_A} - n_{11} e^{n_{21}t_A}}{n_{11} - n_{21}}. \quad (4)$$

At constant annealing temperature, the yield increases monotonically with time (i.e.  $d\eta/dt > 0$ ). The belief that longer annealing times are better is widely held, regardless of competition from other possible annealing sites. However, it is only true if no competition is present. If there are  $k$  possible anneal-

ing sites where the primers may bind, with  $i = 2, 3, \dots, k$  representing undesired competitive sites, then the efficiency of annealing is given by the following expression that is also our main theoretical result:

$$\eta = X_{p1}(T_A, t_A) \prod_{i=2}^k [1 - X_{pi}(T_A, t_A)]. \quad (5)$$

Thus the efficiency  $\eta$  is defined as the probability that the ternary complex has formed at the correct site and not at any competitive site. As in any serial process, the overall PCR efficiency is the product of the efficiency of each step. Hence the overall PCR efficiency is directly proportional to  $\eta$ . The annealing efficiency  $\eta$  is itself affected by many factors, and annealing temperatures and times are only two of them. For the definition of  $\eta$  (Equation (5)), it is tacitly assumed that the polymerase cannot process beyond ternary complexes that have formed downstream at competitive sites  $i > 1$ .

##### 4.1.2. Arrival Rates

In order to calculate the rate constants  $n_{1i}$  and  $n_{2i}$ , the primer arrival and melting rates, and polymerase arrival rate must be known. The primers (and polymerase) diffuse in the bulk medium towards the templates, and the average arrival time at the binding sites depends on a measure of the size ( $\lambda$ ) of a primer (polymerase) and the bulk diffusion coefficient:

$$\tau = \frac{\lambda^2}{6D}. \quad (6)$$

If there are  $R$  primers in the solution, then  $w=R/N$  is the probability that a primer arrives at a binding site in time interval  $\tau$ . The primer arrival frequency (rate constant) is

$$p = \frac{w}{\tau}. \quad (7)$$

Similar arguments are used to estimate the polymerase binding rate constant:

$$s = \frac{w_{pol}}{\tau_{pol}}. \quad (8)$$

A 20 bp primer has an estimated diffusion coefficient of  $D \approx 10^{-10} \text{ m}^2/\text{s}$  (Tirado *et al.*, 1984; Lukacs *et al.*, 2000). The polymerase diffusion coefficient is calculated from the Stokes-Einstein equation (Krouglova *et al.*, 2004) as  $D \approx 4 \times 10^{-11} \text{ m}^2/\text{s}$ .

If the amount of primer in a 25  $\mu\text{l}$  reaction volume is 10 picomoles, and the characteristic length of the primer is taken as  $\lambda = 5 \times 10^{-10}$ , then  $\tau \approx 4.17 \times 10^{-10} \text{ s}$ ,  $N \approx 2 \times 10^{20}$ , and  $w \approx 3 \times 10^{-8}$ . The arrival rate is estimated to be

$$p = \frac{w}{\tau} \approx 72 \text{ s}^{-1}. \quad (9)$$

In this study 0.5 EU of KOD polymerase has been used per experiment. The supplier of KOD (Novagen, Madison, WI) uses the following definition: "1 EU is the amount of enzyme that catalytically inserts 10 nanomoles of dNTPs in 30 min". The amount of polymerase is calculated as 0.5 EU  $\equiv 3.5 \times 10^{-14} \text{ mol}$ . The activity of *Taq* pol. has been used (Innis *et al.*, 1988; Gelfand and White, 1990) although higher rates have been observed for KOD (Griep *et al.*, 2006a; Griep *et al.*, 2006b). The diffusion coefficient and characteristic length are calculated as  $D = 4 \times 10^{-11} \text{ m}^2/\text{s}$  and  $\lambda = 1.2 \times 10^{-8} \text{ m}$ . Based on these values, the arrival rate of the polymerase is

$$s = \frac{w_{pol}}{\tau_{pol}} = 2.4 \text{ s}^{-1}. \quad (10)$$

### 4.1.3. Melting Rates

The melting rate of a primer-template complex depends on its excess free energy ( $\Delta G_p$ ) at a specific temperature  $T$ . SantaLucia and Hicks (2004) provide a comprehensive review of thermodynamic variables for nearest neighbors of base pairs. The nearest-neighbor method is used to calculate the melting probability  $\theta_i$  of a primer at a specific template site as a function of temperature. The melting rates  $m_i$  follow from the relationship:

$$m_i(T) = \frac{\theta_i(T)p}{1 - \theta_i(T)}. \quad (11)$$

### 4.2. Application of the Model

Equation (5) describes the efficiency of annealing when there are  $k - 1$  competitive primer binding sites on the template. For the sake of simplicity, one might write Equation (5) for  $k = 2$ , representing a single competitive site:

$$\eta = X_{P1}(t_A)[1 - X_{P2}(t_A)] \\ = \left[ 1 + \frac{n_{21} e^{n_{11}t_A} - n_{11} e^{n_{21}t_A}}{n_{11} - n_{21}} \right] \left[ \frac{n_{12} e^{n_{22}t_A} - n_{22} e^{n_{12}t_A}}{n_{12} - n_{22}} \right]. \quad (12)$$

In Figure 3A efficiency of annealing,  $\eta(t_A, T_A)$ , is plotted as a function of annealing time and annealing temperature, for the case where competitive binding is low. This situation is expected to represent templates with average GC content (< 50–60%). High efficiency exists over a wide range of annealing times and the efficiency is not adversely affected by the longer annealing times of typical PCR protocols, i.e. annealing times between 10 and 30 s.

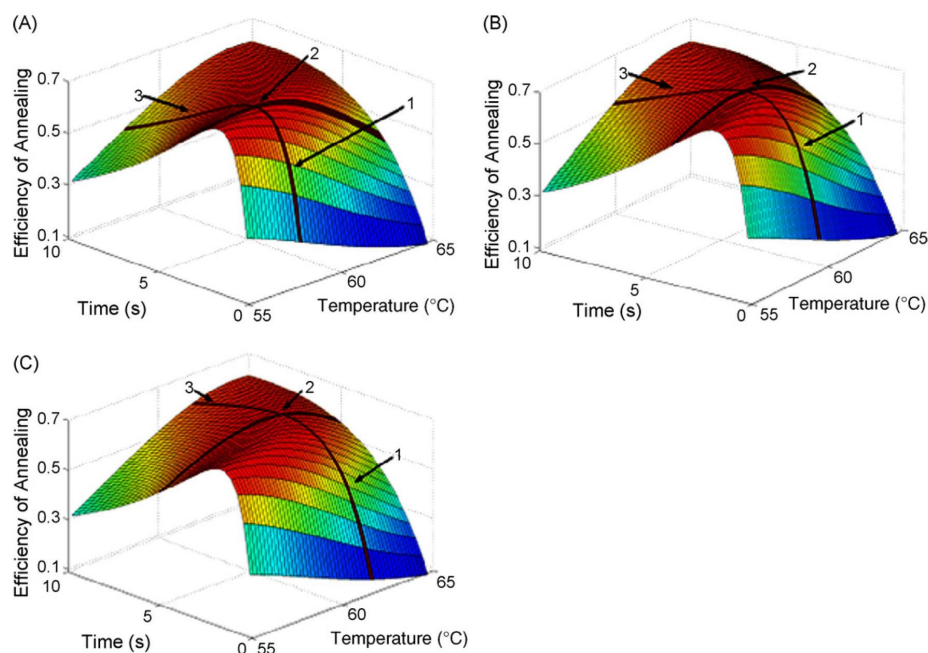
In comparison, Figure 3B shows the efficiency of annealing,  $\eta(t_A, T_A)$ , for the case where competitive binding is high. The figure represents the annealing efficiency profile for the ARX gene, with consideration of competition (defined by Equation (12)). The most important difference between Figure 3A and B is the appearance of a ridge of optimal temperature–time values, and a prominent local optimum – the optimum annealing times lie in

a narrow region between 3 and 6 s at temperatures below 60 °C. Low-annealing efficiencies exist on both sides of the optimum. At  $t_A \ll t_{\text{optimum}}$  the time is insufficient to form ternary complexes at the correct site. Neither the desired nor the undesired competitive PCR products are produced. At  $t_A \gg t_{\text{optimum}}$  the efficiency is low because incorrect competitive binding (mispriming) becomes significant; the result is a smear of PCR products.

### 4.3. Connecting the Experimental Results with the Theoretical Model

In Figure 4A–C the efficiency of annealing  $\eta$  (zoomed in from Figure 3B) is plotted as a function of annealing temperature and annealing time. Figure 4A–C correspond to the experimental results of Figure 1A–C. In each figure the loci at constant temperature (58, 60 and 62 °C) extends over ten seconds and traces the changes in efficiency of annealing (which is proportional to yield) as a function of annealing time. Three points are marked on each loci: point 1 shows lower efficiency than the optimum since insufficient time has passed to form any significant amount of the polymerase-stabilized complex at the correct site, point 2 lies at or near the optimum; sufficient time has passed to form ternary complexes at the correct site, but no significant amount of complexes have yet formed at the competitive sites, point 3 shows the reduction in efficiency due to too long annealing period – more competitive product has now formed.

Consider Figure 1 and Figure 4. The theory predicts a narrow optimal hold time of about 3 s for annealing at 58 °C (point 2 in Figure 4A). This is consistent with the experimental results showing optimal results at 4 s and an increasing smear from as low as 5 s annealing, with a considerable drop in efficiency by 9 s (Figure 1A). Next, consider Figure 1 and Figure 4. The theory predicts an optimal  $t_A$  of about 4 s for annealing at 60 °C (point 2 in Figure 4B) that corresponds well with the experimental results. Although not shown, experimental results at an annealing time of 1 s did not show any significant product – consistent with the theory, as indicated by point 1 in Figure 4B. Finally, consider Figure 1 and Figure 4. At 62 °C



**Figure 4.** The efficiency of annealing  $\eta$  is shown as a function of annealing time and annealing temperature. Loci at constant time and constant temperature are shown in bold on the surface. (A) Loci at 3 s and 58 °C, (B) loci at 4 s and 60 °C, and (C) loci at 5 s and 62 °C. See text for discussion. Color scales correspond to color bars in Figure 3A and B.

annealing temperature, an optimal annealing time of 5 s is predicted (point 2 in Figure 4C), but the surface in the proximity of the optimum is flatter compared to the topography near the optima in the case of the lower annealing temperatures. The near-similar results at 4, 5 and 6 s annealing times in Figure 1C are consistent with the model predictions.

If  $t_A$  is prolonged beyond the optimum time, a decrease in efficiency occurs, as is evident from the increasing smear (Figure 1A–C). The decrease in efficiency beyond the optimal annealing times is less pronounced at higher temperatures (compare the slopes of the loci between points 2 and 3 in Figure 4A–C)—the experimental findings support this. The experimental results at 62 °C show a less intense smear at 9 s compared to the same annealing time at 58 °C (Figure 1A and C). The gel results at 62 °C for annealing times of 9 s and longer for 62 °C maintain a distinct band pattern, despite the increase in the background smear. However, no distinguishable bands are present at the lower annealing temperatures (Figure 1A and B). These results are consistent with the theory that the number of viable alternative primer binding sites becomes fewer at higher annealing temperatures. However, if the annealing times are extended beyond 20 s at 62 °C, one expects that smears will develop similar to those seen for 58 and 60 °C.

Juxtapose the experimental results of Figure 1D with the theoretical results (the constant time locus) in Figure 4A. This locus traces the effect of different annealing temperatures on the efficiency—the annealing time is constant at  $t_A = 3$  s. Theory predicts an optimum annealing temperature of around 59 °C for 3 s annealing times. Points which lie before the optimum, mark lower temperatures than 59 °C and correspond to the lanes marked '56' and '58' in Figure 1D. These two lanes exhibit smears that are typical of competitive product formation. At temperatures above the optimum, a sharp reduction in efficiency is predicted by the model, and that is consistent with experimental results in the lanes marked '62' and '64' in Figure 1D.

In conclusion, refer to Figure 3A, which is representative of minimal competition for binding. Note that the region near optimality is more flat compared to Figure 3B. In other words, the system is much less sensitive to annealing time if competitive binding is absent or minimal. The gel results in Figure 2 confirm that, for normal GC templates with minimal competitive primer binding sites,  $t_A$  can be varied over a broad range without too much variation in yield.

## 5. Conclusions

The most important conclusions that can be drawn from the study are:

1. The theoretical model predicts that, in the absence of competitive binding the efficiency of annealing (definition—Equation (5)) is a monotone function of annealing time. Thus the theory suggests that, in the complete absence of competitive binding, prolonged annealing improves yield.
2. In the case of low-competitive binding (as is typical of genes with average GC content), theory predicts a very broad optimum in annealing times (Figure 3A). This relative insensitivity to annealing times is experimentally confirmed (Figure 2).
3. If there is strong competition to bind primers at alternative sites on the template (as is often the case for GC-rich templates), then the annealing time has a local optimum. The annealing time must be sufficiently long to form the ternary complexes at the correct template site, but too long

annealing time creates the opportunity for ternary complexes to form at incorrect binding sites.

4. The model as applied to the ARX gene is consistent with the experimental results (cf. Figure 1 and Figure 4).
5. Although it is a well-documented fact that the use of DMSO improves the PCR yield, its role in the context of competitive binding is worth mentioning again. DMSO ameliorates the inhibitive effect of competitive binding on the yield function by increasing the melting rates at those sites.

## Acknowledgements

This work was performed through a grant from the National Institutes of Health (1 R21 RR022860-01).

## References

- Agarwal and Perl, 1993 ▶ R. K. Agarwal and A. Perl, PCR amplification of highly GC-rich DNA template after denaturation by NaOH, *Nucl. Acids Res.* **21** (1993), pp. 5283–5284.
- Baskaran *et al.*, 1996 ▶ N. Baskaran, R. P. Kandpal, A. K. Bhargava, M. W. Glynn, A. Bale and S. M. Weissman, Uniform amplification of a mixture of deoxyribonucleic acids with varying GC content, *Genome Res.* **6** (1996), pp. 633–638.
- Chakrabarti and Schutt, 2001 ▶ R. Chakrabarti and C. E. Schutt, The enhancement of PCR amplification by low molecular-weight sulfones, *Nucl. Acids Res.* **29** (2001), pp. 2377–2381.
- Gelfand and White, 1990 ▶ D. H. Gelfand and T. J. White, Thermostable DNA polymerases. In: M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White, Editors, *PCR Protocols: A Guide to Methods and Applications*, Academic Press Inc., San Diego (1990).
- Griep *et al.*, 2006a ▶ M. A. Griep, S. E. Whitney, R. M. Nelson and H. J. Viljoen, DNA polymerase chain reaction: a model of error frequencies and extension rates, *AIChE J.* **52** (2006), pp. 384–392.
- Griep *et al.*, 2006b ▶ M. A. Griep, C. A. Kotera, R. M. Nelson and H. J. Viljoen, Kinetics of the DNA polymerase *Pyrococcus kodakaraensis*, *Chem. Eng. Sci.* **61** (2006), pp. 3885–3892.
- Henke *et al.*, 1997 ▶ W. Henke, K. Herdel, K. Jung, D. Schnorr and S. A. Loening, Betaine improves the PCR amplification of GC-rich DNA sequences, *Nucl. Acids Res.* **25** (1997), pp. 3957–3958.
- Innis *et al.*, 1988 ▶ M. A. Innis, K. B. Myambo, D. H. Gelfand and M. A. D. Brow, DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain-reaction amplified DNA, *Proc. Natl. Acad. Sci. U.S.A.* **85** (1988), pp. 9436–9440.
- Kang *et al.*, 2005 ▶ J. Kang, M. S. Lee and D. G. Gorenstein, The enhancement of PCR amplification of a random sequence DNA library by DMSO and betaine: application to in vitro combinatorial selection of aptamers, *J. Biochem. Biophys. Meth.* **64** (2005), pp. 147–151.
- Kramer and Coen, 2001 ▶ Kramer, M. F., Coen, D. M., 2001. Enzymatic amplification of DNA by PCR: standard procedures and optimization. *Curr. Protoc. Mol. Biol.* (Chapter 15: Unit 15. 1).
- Krouglova *et al.*, 2004 ▶ T. Krouglova, J. Vercammen and Y. Engelborghs, Correct diffusion coefficients of proteins in fluorescence correlation spectroscopy. Application to tubulin oligomers induced by Mg<sup>2+</sup> and Paclitaxel, *Biophys. J.* **87** (2004), pp. 2635–2646.
- Lukacs *et al.*, 2000 ▶ G. L. Lukacs, P. Haggie, O. Seksek, D. Lechardeur, N. Freedman and A. S. Verkman, Size-dependent DNA mobility in cytoplasm and nucleus, *J. Biol. Chem.* **275** (2000), pp. 1625–1629.
- McDowell *et al.*, 1998 ▶ D. G. McDowell, N. A. Burns and H. C. Parkes, Localized sequence regions possessing high melting temperature prevent the amplification of a DNA mimic in competitive PCR, *Nucl. Acids Res.* **26** (1998), pp. 3340–3347.
- Musso *et al.*, 2006 ▶ M. Musso, R. Bocciardi, S. Parodi, R. Ravazzolo and I. Ceccherini, Betaine, dimethyl sulfoxide, and 7-deaza-dGTP, a powerful mixture for amplification of GC-rich DNA sequences, *J. Mol. Diagn.* **8** (2006), pp. 544–550.
- Mutter and Boynton, 1995 ▶ G. L. Mutter and K. A. Boynton, PCR bias in amplification of androgen receptor alleles a trinucleotide



- repeat marker used in clonality studies, *Nucl. Acids Res.* **23** (1995), pp. 1411-1418.
- Mytelka and Chamberlin, 1996** ► D. S. Mytelka and M. J. Chamberlin, Analysis and suppression of DNA polymerase pauses associated with a trinucleotide consensus, *Nucl. Acids Res.* **24** (1996), pp. 2774-2781.
- Pomp and Medrano, 1991** ► D. Pomp and J. F. Medrano, Organic solvents as facilitators of polymerase chain reaction, *Biotechniques* **10** (1991), pp. 58-59.
- Ralser et al., 2006** ► M. Ralser, R. Querfurth, H. J. Warnatz, H. Leh-rach, M. L. Yaspo and S. Krobitsch, An efficient and economic en-hancer mix for PCR, *Biochem. Biophys. Res. Commun.* **347** (2006), pp. 747-751.
- Rees et al., 1993** ► Rees, W. A., Yager, T. D., Korte, J., von Hippel, P. H., 1993. Betaine can eliminate the base pair composition depen-dence of DNA melting. *Biochemistry* **12**, 137-144.
- Sahdev et al., 2007** ► S. Sahdev, S. Saini, P. Tiwari, S. Saxena and K. Singh Saini, Amplification of GC-rich genes by following a combi-nation strategy of primer design, enhancers and modified PCR cy-cle conditions, *Mol. Cell Probes* **21** (2007), pp. 303-307.
- SantaLucia and Hicks, 2004** ► J. SantaLucia Jr. and D. Hicks, The ther-modynamics of DNA structural motifs, *Annu. Rev. Biophys. Biomol. Struct.* **33** (2004), pp. 415-440.
- Sidhu et al., 1996** ► M. K. Sidhu, M. J. Liao and A. Rashidbaigi, Di-methyl sulfoxide improves RNA amplification, *Biotechniques* **21** (1996), pp. 44-47.
- Sun et al., 1993** ► Y. Sun, G. Hegamyer and N. H. Colburn, PCR-di-rect sequencing of a GC-rich region by inclusion of 10% DMSO: application to c-jun, *Biotechniques* **15** (1993), pp. 372-374.
- Tirado et al., 1984** ► M. M. Tirado, C. L. Martinez and J. G. de la Torte, Comparison of theories for translational and rotational diffusion coefficients of rod-like macromolecules. Application to short DNA fragments, *J. Chem. Phys.* **81** (1984), pp. 2047-2052.
- Turner and Jenkins, 1995** ► S. L. Turner and F. J. Jenkins, Use of de-oxyinosine in PCR to improve amplification of GC-rich DNA, *Bi-otechniques* **19** (1995), pp. 48-52.
- Weissensteiner and Lanchbury, 1996** ► T. Weissensteiner and J. S. Lanchbury, Strategy for controlling preferential amplification and avoiding false negatives in PCR typing, *Biotechniques* **21** (1996), pp. 1102-1108.
- Wilson et al., 1997** ► A. G. Wilson, J. A. Symons, T. L. McDowell, H. O. McDevitt and G. W. Duff, Effects of a polymorphism in the hu-man tumor necrosis factor alpha promoter on transcriptional acti-vation, *Proc. Natl. Acad. Sci. U.S.A.* **94** (1997), pp. 3195-3199.