Chemical Engineering Science 66 (2011) 1783–1789

Contents lists available at ScienceDirect

Chemical Engineering Science

journal homepage: www.elsevier.com/locate/ces

Experimental validation of a fundamental model for PCR efficiency

Tobias M. Louw, Christine S. Booth, Elsje Pienaar, Joel R. TerMaat, Scott E. Whitney, Hendrik J. Viljoen *

Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588-0643, USA

ARTICLE INFO

Article history:
Received 28 September 2010
Received in revised form
22 December 2010
Accepted 11 January 2011
Available online 19 January 2011

Keywords:
Biological and biomolecular engineering
Enzyme
Kinetics
Mathematical modeling
Molecular biology
PCR efficiency

ABSTRACT

Recently a theoretical analysis of PCR efficiency has been published by Booth et al. (2010). The PCR yield is the product of three efficiencies: (i) the annealing efficiency is the fraction of templates that form binary complexes with primers during annealing, (ii) the polymerase binding efficiency is the fraction of binary complexes that bind to polymerase to form ternary complexes and (iii) the elongation efficiency is the fraction of ternary complexes that extend fully. Yield is controlled by the smallest of the three efficiencies and control could shift from one type of efficiency to another over the course of a PCR experiment. Experiments have been designed that are specifically controlled by each one of the efficiencies and the results are consistent with the mathematical model. The experimental data has also been used to quantify six key parameters of the theoretical model. An important application of the fully characterized model is to calculate initial template concentration from real-time PCR data. Given the PCR protocol, the midpoint cycle number (where the template concentration is half that of the final concentration) can be theoretically determined and graphed for a variety of initial DNA concentrations. Real-time results can be used to calculate the midpoint cycle number and consequently the initial DNA concentration, using this graph. The application becomes particularly simple if a conservative PCR protocol is followed where only the annealing efficiency is controlling.

1. Introduction

The polymerase chain reaction (PCR) has become a major technology in microbiology, molecular biology and related fields. Whereas PCR still has a lot of qualitative applications, it is increasingly used as a quantitative tool. The sensitivity of PCR permits amplification from a small number of starting templates. However, the exponential increase in product makes the inverse problem difficult—i.e. to infer the starting concentration from a large number of amplicons. Real-time PCR provides a proportional measure of the number of templates at each cycle.

Several methods have been proposed to calculate the initial template concentration from the real-time curves. Traditionally, standard calibration curves were used (Higuchi et al., 1993) to compare real-time results to reference samples, but this technique requires DNA standard plasmids. Samples with known DNA concentrations are used to construct linear functions relating the initial DNA concentration to some cross-over cycle number. While these functions generally correlate extremely well to experimental data, they are purely empirical in nature.

More recently, investigations of the plateau phase of the real-time PCR curve have revealed methods to calibrate the measurements "internally", using the initial primer or probe concentrations (Swillens et al., 2004). These methods rely on mathematical models to determine the ratio between the primer- and DNA-concentrations (Smith et al., 2007). These models are usually formulated in terms of cycle efficiency.

The DNA yield depends on the efficiency of the reaction during each cycle (Saiki et al., 1992). The cycle efficiency is the product of the individual efficiencies of the denaturing, annealing, polymerase binding and elongation steps (Booth et al. 2010). As the reaction progresses the efficiency decreases resulting in the characteristic sigmoidal real-time curve (Kainz, 2000; Schnell, 1997; Schnell and Mendoza, 1997; Stolovitzky and Cecchi, 1996).

Numerous mathematical models of varying complexity have been published describing the reaction. The most general models assume constant efficiencies across all of the PCR cycles: the $\Delta\Delta C_T$ method assumes 100% efficiency while methods by Pfaffl (2001) and Liu and Saint (2002a) calculate reaction specific efficiencies. When these methods are applied for quantitative real-time PCR, they are only applied to the early phase of the reaction when efficiency is assumed to be nearly constant. More complex models account for per cycle variation in efficiency, but still combine the efficiencies of each step (denaturing, annealing and elongation) into an overall efficiency for each cycle (Liu and Saint, 2002b; Platts et al., 2008). While some models account for the decrease in cycle efficiency using empirical estimates (Alvarez et al., 2007), even more complex models consider the efficiency of the steps of

* Corresponding author. Tel.: +1 402 472 9318; fax: +1 402 472 6989.
E-mail address: hviljoen1@unl.edu (H.J. Viljoen).

0009-2509/$ - see front matter © 2011 Published by Elsevier Ltd.
doi:10.1016/j.ces.2011.01.029

each cycle independently, but require numerical solution, making them difficult to apply (Gevertz et al., 2005; Mehra and Hu, 2005; Rubin and Levy, 1996; Smith et al., 2007).

The mathematical model described by Booth et al. (2010) presents an analytical model that can be used to better understand the PCR process. The model provides explicit expressions for the efficiencies of each individual PCR cycle. These efficiencies are combined into an easily implementable expression for the yield per cycle.

The model shows that different mechanisms may control the efficiency. A decrease in polymerase concentration and/or elongation time reduce the cycle efficiency, but do not affect the final template concentration (the sigmoidal concentration curve shifts laterally). Decreasing the primer concentration not only decreases the efficiency, but also decreases the final template concentration. Some model parameters, such as reaction rate constants, are unknown and must be determined by matching the model with experimental results. A short review of the model and the key parameters is given in the next section.

Experimental validation of the mathematical model presented by Booth et al. (2010) is presented in this work. Various real-time experiments have been designed to explore reactions that are limited by the annealing-, polymerase binding- and elongation efficiencies. These results have been used to determine the unknown model parameters. Finally, it is shown that this model provides an elegant method to determine initial DNA concentrations, using real-time data and the PCR protocol.

2. Mathematical model

An analytical model was used to calculate the template concentration S_j for each PCR cycle j. The template is the region of the sample DNA flanked by the sense- and anti-sense primers for replication; thus the initial DNA concentration is equal to the initial template concentration. For a complete derivation of the model, see Booth et al. (2010). The model is based on the following assumptions:

- There are equal numbers of forward and reverse primers and they anneal to equal numbers of sense and anti-sense single-stranded DNA.
- All of the double-stranded DNA denatures completely to form single-stranded DNA.
- No primer-dimers are formed, nor does non-specific primer-template annealing occur.
- Primer-template annealing does not occur during the elongation phase.
- The annealing and elongation reactions are irreversible at the relevant temperatures.
- Partial elongation is not considered. Strands that are not fully extended by the end of the elongation cycle are treated as primers in subsequent cycles.
- The extension rate remains constant, i.e. no slow-down due to pyro-phosphorolysis or dNTP depletion.

The model calculates an overall per cycle efficiency (η_j), which is the product of three individual efficiencies. The annealing efficiency ($\eta_{j,a}$) is the fraction of available templates that anneal to primers. The polymerase binding efficiency ($\eta_{j,E}$) is the fraction of template-primer (binary) complexes that bind to polymerase to form ternary complexes. Finally, the elongation efficiency ($\eta_{j,e}$) is the fraction of ternary complexes that are fully extended by the end of the elongation step:

$$\eta_j = \eta_{j,a}\eta_{j,E}\eta_{j,e} \tag{1}$$

$$\eta_{j,a} = (P_j - P_{j,a})/S_j \tag{2}$$

$$\eta_{j,E} = C_{j,e}/(B_{j,a} + C_{j,a}) \tag{3}$$

$$\eta_{j,e} = C_{j,c}/C_{j,e} \tag{4}$$

The variables are defined in Tables 1 and 2. The subscript j identifies the cycle and the subscripts a and e denote values at the end of the annealing and elongation stages respectively. For example, there are S_j templates and P_j primers at the start of cycle j, but at the end of the annealing stage there are $P_{j,a}$ primers left. Thus the number of binary and ternary complexes that have formed during the annealing stage is $(P_j - P_{j,a})$ and the ratio $(P_j - P_{j,a})/S_j$ defines the annealing efficiency. Eqs. (5)–(7) give the primer, ternary and binary complex values at the end of the annealing stage. The number of ternary complexes at the end of the elongation stage is given by Eq. (8). The ternary complex concentration at the cut-off time ($C_{j,c}$) is the amount of primer–template–polymerase complexes that have formed after $t_c = t_e - l/V$ time has passed in the elongation phase. The value l/V is the time it takes the polymerase to extend the primer to full length DNA. Thus, $C_{j,C}$ is the concentration of ternary complexes that will fully

Table 1
Experimental and model parameters used in analytic model.

Experimental parameters	Description	Model parameters	Description
t_a; t_e	Annealing/elongation phase duration	k_p	Rate of primer annealing
S_0	Initial template concentration	k_c	Rate of polymerase binding at the annealing temperature
P_0	Initial primer concentration	k_c^*	Rate of polymerase binding at the elongation temperature
E_0	Initial polymerase concentration	β	Ratio of template annealing rate to primer annealing rate
V	Polymerase extension rate	η_d	Template denaturing damage
l	Template length	η_{dE}	Polymerase denaturing damage

Table 2
Variables used in analytic model.

Variable	Description	Variable	Description
S_j	Template concentration at the beginning of annealing	E_j	Polymerase concentration at the beginning of annealing
P_j;$P_{j,a}$	Primer concentration at the beginning and end of annealing	$B_{j,a}$	Binary complex concentration at the end of annealing
γ_j	Ratio of template to primer concentration (S_j/P_j)	$C_{j,a}$; $C_{j,e}$; $C_{j,c}$	Ternary complex concentration at the end of annealing, elongation and at the cut-off time, respectively