Polymerase Chain Reaction: A Markov Process Approach

MIKHAIL V. VELIKANOV AND RAYMOND KAPRAL*

Chemical Physics Theory Group, Department of Chemistry, University of Toronto,
Toronto, Ontario M5S 3H6, Canada

(Received on 7 May 1999, Accepted in revised form on 28 September 1999)

A probabilistic approach to the kinetics of the polymerase chain reaction (PCR) is developed. The approach treats the primer extension step of PCR as a microscopic Markov process in which the molecules of deoxy-nucleoside triphosphate (dNTP) are bound to the 3’ end of the primer strand one at a time. The binding probability rates are prescribed by combinatorial rules in accord with the microscopic chemical kinetics. As an example, a simple model based on this approach is proposed and analysed, and an exact solution for the probability distribution of lengths of synthesized DNA strands is found by analytical means. Using this solution, it is demonstrated that the model is able to reproduce the main features of PCR, such as extreme sensitivity to the variation of control parameters and the existence of an amplification plateau. A multidimensional optimization technique is used to find numerically the optimum values of control parameters which maximize the yield of the target sequence for a given PCR run while minimizing the overall run time.

1. Introduction

The polymerase chain reaction (PCR) is an elegant technique for the amplification of specific sequences of DNA. It is currently used in a wide variety of applications in biochemistry and molecular biology (Mullis & Faloona, 1987; Saiki et al., 1988; Mullis et al., 1994); however, despite its general popularity, PCR is not completely free of shortcomings. Many of these shortcomings have their origin in the extreme sensitivity of the amplification efficiency to the reaction conditions, such as temperature, polymerase and Mg²⁺ concentration, primer length, etc. (see e.g. Innis & Gelfand, 1990). As a result, a seemingly insignificant change in the reaction conditions may cause a notable decrease in amplification efficiency and yield of the target DNA segment.

There is a large amount of empirical data on the relationship between amplification efficiency and reaction conditions which has yet to be fully understood and systematized (Innis & Gelfand, 1990; Roux, 1995). Theoretical models of PCR have been constructed to investigate various aspects of its kinetics. In some studies, DNA amplification is viewed as a deterministic process and described mathematically by a geometric series (Raeymaekers, 1993; Hayward et al., 1998; Dimitrov & Apostolova, 1996). A different, albeit still deterministic approach was presented by Schnell & Mendoza (1997a, b) who used the law of mass action to derive kinetic equations for PCR similar to Michaelis–Menten equations of enzyme kinetics. Stochastic models for PCR have also been developed (Nedelman et al., 1992; Sun, 1995; Peccoud & Jacob, 1996; Weiss & von Haeseler,
Alternatively, if the nucleotides are distinct and only the dNTPs complementary to the corresponding nucleotides in the template strand can bind to the primer, the probability rate is proportional to the number of dNTPs of only one type. Assuming equal numbers of dNTPs of each type and equal rate coefficients, the probability rate is given by eqn (2) with \( n \) equal to \( \frac{1}{4} \) of the total number of dNTPs, yielding a model which is mathematically equivalent to that discussed in the text.

2. Markov Process Model for Primer Extension

We consider a primer of initial length \( l_0 \) bound to the template strand and extending by the polymerase-catalysed addition of dNTPs to its 3' end. The addition occurs randomly, with the probability per unit time determined entirely by the present state of the system. Thus, the process of extension is a Markov process (Cox & Miller, 1995). Since the primer extends only in one direction, and the primer binding site is usually offset from the 3' end of the template strand, the limiting length of the template \( L \) which figures in the following derivation may not be identical to the true length of the template strand, \( L_{true} \). The two lengths are trivially related to one another through the offset between the 3' end of the template and the 5' end of the primer, \( L_{off} \):

\[
L = L_{true} - L_{off}. \tag{1}
\]

At the microscopic level, the probability rate of a reaction event is proportional to the number of ways in which the molecules of the reactants available in the system can be combined in order for the reaction event to occur. For simplicity, we assume that all nucleotides are identical (as in, for example, Stolovitzky & Cecchi, 1996) so that only the length of the growing primer and not the sequence is of interest. In this case, the probability rate of addition of a single dNTP is proportional to the total number of dNTPs in the system, i.e.\(^\dagger\)

\[
w = k(t)n, \tag{2}
\]

\(^\dagger\) Alternatively, if the nucleotides are distinct and only the dNTPs complementary to the corresponding nucleotides in the template strand can bind to the primer, the probability rate is proportional to the number of dNTPs of only one type. Assuming equal numbers of dNTPs of each type and equal rate coefficients, the probability rate is given by eqn (2) with \( n \) equal to \( \frac{1}{4} \) of the total number of dNTPs, yielding a model which is mathematically equivalent to that discussed in the text.
where \( k(t) \) is the rate coefficient which depends on temperature and therefore implicitly depends on time, as the temperature need not remain constant throughout the extension process. Since extension of the primer by one unit requires addition of one dNTP, the sum of the number of dNTPs in the system, \( n \), and the primer length, \( l \), is always equal to a constant, \( m_0 \).

\[
l + n = l_0 + n_0 = m_0,
\]

(3)

where \( n_0 \) is the initial number of dNTPs. Using eqn (3), \( n \) can be expressed in terms of \( l \), which allows one to describe the system in terms of the (time-dependent) probability distribution of a single random variable \( l \)—the length of the extending primer. This distribution contains full information on frequencies of DNA strands of various lengths at any moment of time; given which, one can easily model a multi-cycle PCR run. The evolution of this probability distribution is governed by a master equation (Gardiner, 1985; Cox & Miller, 1965). Below we derive the master equation for the primer extension process and find its exact solution.

The master equation for the primer extension process can be written in terms of the probability rate of dNTP addition as

\[
\frac{\partial}{\partial t} P(l, t) = w_{l-1 \to l} P(l-1, t) - w_{l \to l+1} P(l, t),
\]

(4)

where \( P(l, t) \) is the probability distribution of primer length at time \( t \), and \( w \) is the transition probability rate, with the subscript identifying the transition to which a given rate corresponds.

For a given limiting length of template \( L \) [cf. eqn (1)], there are two different cases of initial condition to consider in solving eqn (4). Let \( n_0 < L - l_0 \), implying that \( L > m_0 \) [cf. eqn (3)]. In this case, the primer extension is clearly limited only by \( n_0 \). Alternatively, if \( L \leq m_0 \), then the extension is limited by \( L \) only. The latter case is of practical interest, since it corresponds to an excess of dNTPs which is standard for all PCR protocols. However, first it is helpful to consider the former case where there is a deficit of dNTPs.

The transition probability rate for addition of a single dNTP can be written as

\[
w_{l \to l+1} = k(t) (m_0 - l).
\]

(5)

We combined eqn (2) and conservation law (3) to remove \( n \) from the expression for the probability rate. We now focus on the uninterrupted extension process corresponding to the extension phase of a single cycle of a PCR run. The results obtained here shall be applied in the following subsection to model complete multi-cycle PCR runs in which consecutive extension phases are separated by denaturation and annealing steps.

Substituting eqn (5) into eqn (4) yields

\[
\frac{\partial}{\partial t} P(l, t) = k(t) [(m_0 - l + 1) P(l-1, t) - (m_0 - l) P(l, t)].
\]

(6)

To solve eqn (6), we introduce a new time variable \( \eta \) which is a nonlinear function of \( t \), such that

\[
\frac{d\eta}{dt} = k(t), \quad \eta(0) = 0.
\]

In terms of \( \eta \), the master eqn (6) can be written as

\[
\frac{\partial}{\partial \eta} P(l, \eta) = (m_0 - l + 1) P(l-1, \eta) - (m_0 - l) P(l, \eta).
\]

(7)

Further, we introduce the generating function \( G(s, \eta) \) related to the probability distribution \( P(l, \eta) \) (Gardiner, 1985),

\[
G(s, \eta) = \sum_{l=0}^{L} s^l P(l, \eta)
\]

(8)

and, using eqn (7), derive the partial differential equation which this generating function satisfies. We find

\[
\frac{\partial}{\partial \eta} G(s, \eta) = (s - 1) \left[ m_0 G(s, \eta) - s \frac{\partial}{\partial s} G(s, \eta) \right].
\]

(9)
Moreover, we have by definition $G(s, \eta) = s^{-m_0} F(s, \eta)$. From eqn (9), one sees that $F(s, \eta)$ obeys

$$\frac{\partial F}{\partial \eta} = s(1-s) \frac{\partial F}{\partial s}. \quad (10)$$

Now, introduce a new variable $\sigma = \ln(s/(s-1))$ and substitute in eqn (10). This gives, after some algebra

$$\frac{\partial F}{\partial \eta} = \frac{\partial F}{\partial \sigma}. \quad (11)$$

The solution of eqn (11) is clearly an arbitrary function of $\eta + \sigma$, $F = F(\eta + \sigma)$. The analytical form of the solution must be determined from the initial conditions. Since the initial length of the primer is $l_0$, the initial probability distribution is $P(l,0) = \delta_{l,l_0}$, where $\delta_{l,l_0}$ is the Kronecker delta. From eqn (8) it follows that $G(s,0) = s^{l_0}$. We can re-write $G(s,0)$ in terms of the variable $\sigma$ as

$$G(\sigma, 0) = (1 - e^{-\sigma})^{-l_0}. \quad (12)$$

Moreover, we have by definition $G(s, \eta) = s^{m_0} F(s, \eta)$, which implies

$$G(\sigma, \eta) = (1 - e^{-\sigma})^{-m_0} F(\eta + \sigma). \quad (13)$$

Comparing eqns (12) and (13), we find for $G(\sigma, \eta)$,

$$G(\sigma, \eta) = (1 - e^{-\sigma})^{-m_0} (1 - e^{-\eta + \sigma})^{m_0 - l_0}. \quad (14)$$

Noting that $m_0 - l_0 = n_0$ [cf. eqn (3)] and converting $\sigma$ into $s$ in eqn (14), we finally obtain

$$G(s, \eta) = s^{m_0} \left(1 - e^{-\eta} + \frac{e^{-\eta}}{s}\right)^{n_0}. \quad (15)$$

To obtain the probability distribution $P(l, \eta)$, we expand eqn (15) in powers of $s$ and read off the coefficients of the series [cf. eqn (8)]. In this way, we arrive at

$$P(l, \eta) = \begin{cases} 0 & \text{for } l < l_0, \\ \left(\frac{n_0}{l - l_0}\right) (1 - e^{-\eta})^{-l_0} e^{-(m_0 - l)\eta} & \text{for } l \geq l_0, \\ 1 - \sum_{k=0}^{L-1} P(k, \eta; l_0, L) & \text{for } l = L. \end{cases} \quad (16)$$

where $\binom{n_0}{l_0}$ stands for a binomial coefficient. The probability distribution (16) is the exact solution of the master equation (7) as can be easily verified by substitution and, by construction of the master equation, it pertains to the case of primer extension limited by $n_0$. Note that under the conditions $n_0(1 - e^{-\eta}) \gg 1$, $n_0 \to \infty$, $n_0 \gg l - l_0$ and $e^{-\eta} \to 1$, distribution (16) can be approximated by a normal distribution with mean $l_0 + n_0(1 - e^{-\eta})$ and half-width $\sqrt{n_0 e^{-\eta}(1 - e^{-\eta})}$.

We now consider the case of the primer extension limited by $L$. One notes that for $l < L$, the probability distribution must be exactly the same as in the case of extension limited by $n_0$. This is due to the fact that, for both cases, evolution of $P(l, \eta)$ is determined only by itself and $P(l-1, \eta)$ [cf. eqn (7)]. Furthermore, the value of the distribution at $l = L$ is given by the normalization condition:

$$P(L, \eta) = 1 - \sum_{l=0}^{L-1} P(l, \eta). \quad (17)$$

This discontinuity at the end of the interval is a natural consequence of the boundary conditions of the problem (i.e. extension cannot proceed beyond length $L$). In the following, we adopt a special notation for such distributions with discontinuity at the boundary where the initial primer length $l_0$ and the location of discontinuity $L$ are denoted as additional parameters. Thus, the distribution for the case of $L$-limited extension is denoted as $P(l, \eta; l_0, L)$ and has the following form:
In Fig. 1, we present graphs of the probability distribution $P(l, \eta; l_0, L)$ for four different values of $\eta$. One sees that the behavior of the distribution in time is similar to that of a “wave” passing through the interval between $l_0$ and $L$, in accord with the nature of the extension process.

3. Model for Multi-Cycle PCR Runs

A PCR run normally involves a number of consecutive cycles of primer extension, separated by denaturation and primer annealing phases. Thus far, we have considered only the basic unit of such a run, namely extension of a single primer bound to a single template molecule. In order to deal with a complete PCR run, it is necessary to introduce two additional features into the model: the increasing number of DNA molecules involved in the reaction and their complementarity. The first feature is easily taken into account. Because of the statistical independence of the extension process, the probability distribution obtained in the single-template picture corresponds to fractional abundances (i.e. concentration) of DNA strands with different lengths in the multi-template picture. The number of dNTPs also has to be re-interpreted; thus, $n_p$ henceforth will stand for the initial number of dNTPs per template strand.

An additional parameter must also be introduced to indicate the number of primers per template strand available in the system; it will be denoted as $n_p$. It is clear that during a single cycle of a PCR run, one primer per template strand is consumed. Therefore, if $n_p$ is greater than the total number of cycles in the run, all cycles of the run will be completed and the final probability distribution for the DNA strands synthesized in the run will be independent of the actual value of $n_p$. On the other hand, if $n_p$ is smaller than the total number of cycles, only $n_p$ cycles will be completed and all further cycles will not change the probability distribution for synthesized DNA strands. In the following, we assume that $n_p$ is greater than the total number of cycles and omit the notation due to $n_p$.

The second feature, which arises from the fact that each template is initially supplied along with its complementary counterpart, requires two kinds of probability distribution, one for each type of complementary DNA strand. The template molecules in the original DNA sample are in the double-stranded form; we arbitrarily label one as the “+” type and the other as the “−” type. All primers and DNA strands produced by PCR amplification can also be classified as either “+” or “−” type, according to their complementarity; a primer or a synthesized strand which is complementary to the “+” template is considered to be of the “−” type, and vice versa. We use the superscripts “+” and “−” to distinguish parameters as well as probability distributions pertaining to the extension of different types of primer. For the probability distributions, subscripts will also be used to denote the cycle to which a given distribution refers; thus, we have $P^+_1(l, n)$, $P^-_1(l, n)$, $P^+_2(l, n)$, $P^-_2(l, n)$, and so forth. The diagram in Fig. 2 serves to illustrate this notation. The same diagram shows the definition of a number of parameters which are essential for the following discussion, namely the target length $(l_t)$, the limiting “+” and “−” template lengths $(L^+$ and $L^−$), and the “+” and “−” primer lengths $(l^+_0$ and $l^-_0$).

The target length $l_t$ and the true template length $L_{true}$ are the same for both “+” and “−” DNA strands. One can easily see that even though $L_{true}$ is the same for both types of templates, the limiting length which actually affects the shape of the probability distribution can be quite different ($L^+$ and $L^−$ in Fig. 2). In fact, this
is the case for the majority of PCR protocols since the primers usually have different lengths and their binding sites are offset by different distances from the 3′ ends of the corresponding templates. Hence, extension of “+” and “−” primers occurs under different conditions, leading to different “+” and “−” probability distributions for each cycle.

3.1. EVOLUTION OF “+” AND “−” PROBABILITY DISTRIBUTIONS

We denote the duration of the extension phase of each cycle by \( \eta_n \), where \( n \) is the cycle number. The time is reset at the beginning of each extension phase, and therefore each \( \eta_n \) is counted from \( \eta = 0 \). The distributions for the first cycle \( P_1^+ (l, \eta) \) and \( P_1^- (l, \eta) \) at the end of the extension phase are given by eqn (17) with \( \eta = \eta_1 \), and \( L = L^+ \), \( l_0 = l_0^+ \) and \( L = L^- \), \( l_0 = l_0^- \), respectively,

\[
P_1^+ (l, \eta_1) = P(l, \eta_1; l_0^+, L^+),
\]

\[
P_1^- (l, \eta_1) = P(l, \eta_1; l_0^-, L^-).
\]

The extension process consumes dNTPs; moreover, the number of dNTPs consumed in “+” primer extension is different from that consumed in “−” primer extension. One finds

\[
A_1^+ = \sum_{l=0}^{L^+} l P_1^+ (l, \eta_1) - l_0^+,
\]

\[
A_1^- = \sum_{l=0}^{L^-} l P_1^- (l, \eta_1) - l_0^-,
\]

where \( A_1^+ \) stands for the depletion of dNTPs due to “+” primer extension in the first cycle, and \( A_1^- \) is a similar quantity for “−” primer extension. Hence, at the end of the first extension phase the average number of dNTPs per template strand decreased from its initial value of \( n_0 \) by \( (A_1^+ + A_1^-)/2 \). The number of dNTPs available per template strand must be updated in this manner at the end of the extension phase of each cycle.

After the extension phase of the first cycle is completed, the DNA strands synthesized by primer extension are separated from the templates by thermal denaturation. New primers are then annealed onto templates from original DNA sample, as well as onto a fraction of the newly synthesized strands. What determines whether a newly synthesized strand may potentially bind to a primer is its length: binding is not possible for those strands which are not sufficiently long, i.e. have no or little overlap with the primers at the 3′ end. Extension of primers annealed onto the synthesized strands is limited by the target length \( l_t \) as a consequence of complementarity, while extension of primers annealed onto the templates from the original sample proceeds as in the first cycle and is limited by the template length, \( L^+ \) or \( L^- \). Given the duration of the extension phase of the second cycle, \( \eta_2 \), the probability distribution for extension on synthesized strands is either \( P(l, \eta_2; l_0^+, l_t) \), if a “+” primer bound to a “−” strand is considered, or \( P(l, \eta_2; l_0^-, l_t) \) in the opposite case. The distributions appropriate for extension on the templates from the original sample are \( P(l, \eta_2; l_0^+, L^+) \) or \( P(l, \eta_2; l_0^-, L^-) \). As a result, the overall probability distributions for the second cycle \( P_2^+ (l, \eta_2) \) and \( P_2^- (l, \eta_2) \) can be written as linear combinations of \( P(l, \eta_2; l_0^+, l_t) \) and \( P(l, \eta_2; l_0^-, l_t) \) and \( P(l, \eta_2; l_0^+, L^+) \) and \( P(l, \eta_2; l_0^-, L^-) \).
The extension phase of the second cycle causes further depletion of dNTPs in the system. Therefore, once this extension phase is completed, the number of dNTPs has to be updated using the procedure outlined above [see eqn (20) and following discussion].

The four quantities \( \omega_{11}^+, \omega_{21}^+, \omega_{22}^+, \omega_{22}^- \) in eqn (21) are the weights of component distributions in the linear combinations. The meaning of the subscripts is as follows: the first subscript is the cycle number, and the second subscript is 1 for weights corresponding to the primer extension limited by \( L^- \) or \( L^+ \), and 2 for weights corresponding to \( l^- \)-limited extension. Note that eqn (21) formally applies to \( P_1^-(l, \eta_1) \) and \( P_1^+(l, \eta_1) \) as well if one sets \( \omega_{11}^+ = \omega_{11}^- = 1 \) and \( \omega_{12}^+ = \omega_{12}^- = 0 \). As the denaturation–annealing–extension cycle repeats, extension will again be limited partly by \( l \), and partly by \( L^+ \) or \( L^- \), so that the distributions for all subsequent cycles are given by the same general formula (21). However, the weights change from cycle to cycle, reflecting variation in the amplification rate.

### 3.2. CALCULATION OF WEIGHTS

The weights are just the fractions of certain kinds of DNA molecules out of the entire DNA matter accumulated. Recall that the distributions \( P_1^\pm (l, \eta_1) \), \( P_2^\pm (l, \eta_2) \), etc. which figure in the above discussion are for the extending strands only and therefore do not contain sufficient information to compute these fractions. One must combine the distributions for all cycles including the present to obtain the necessary information. A different type of distribution must therefore be introduced to monitor composition of all of the DNA in the system, and not just the strands extending during the present cycle. Apart from the length variable \( l \), such a distribution depends on the duration of the extension phases of all cycles. For simplicity, distributions of this type will be denoted in the following by \( \Pi_n^+ (l) \) or \( \Pi_n^- (l) \) (\( n \) being the cycle number) with dependence on \( \eta_1, \eta_2, \ldots, \eta_n \) implied. These distributions can be obtained using recursive update rules:

\[
\Pi_n^+(l) = (1 + \omega_{n1}^+ + \omega_{n2}^+)^{-1} \left[ \Pi_{n-1}^+(l) + P_n^+(l, \eta_n) \right],
\]

\[
\Pi_n^-(l) = (1 + \omega_{n1}^- + \omega_{n2}^-)^{-1} \left[ \Pi_{n-1}^-(l) + P_n^-(l, \eta_n) \right].
\]

The prefactors \( (1 + \omega_{n1}^+ + \omega_{n2}^+)^{-1} \) and \( (1 + \omega_{n1}^- + \omega_{n2}^-)^{-1} \) are the normalization constants. One can easily see from eqns (21) and (22) that, given the normalized initial distributions \( \Pi_0^+(l) \) and \( \Pi_0^-(l) \), these prefactors ensure that the distributions obtained by iterations of eqn (22) are always normalized. The initial distributions which we use are \( \Pi_0^+(l) = \Pi_0^-(l) = \delta_{l_0, l_{true}} \), since there are only templates with length \( L_{true} \) in the system before the reaction run is started.

The weights for the \( n \)-th cycle can now be calculated from the distributions \( \Pi_{n-1}^+(l) \) and \( \Pi_{n-1}^-(l) \). Thus, \( \omega_{n1}^+ \) represents the fraction of the “+” primers whose extension in the \( n \)-th cycle is limited by \( L^+ \). One can argue, based on the foregoing discussion and the primer–template complementarity, that \( \omega_{n1}^+ \) equals the fraction of the “−” templates from the original DNA sample (with length \( L_{true} \)) out of the entire DNA content of the system accumulated up to the \( n \)-the cycle. This is just the value of \( \Pi_{n-1}^+(l) \) at \( l = L_{true} \). Similarly, \( \omega_{n1}^- \) is equal to the value of \( \Pi_{n-1}^- (l) \) at \( l = L_{true} \), i.e.

\[
\omega_{n1}^+ = \Pi_{n-1}^+(L_{true}),
\]

\[
\omega_{n1}^- = \Pi_{n-1}^-(L_{true}).
\]
must have to bind a “−” primer is \( l_i - a^- \). Similarly, the smallest length of a “+” strand that allows binding of a “+” primer is \( l_i - a^+ \). Now, \( \omega_{n_2}^+ \) is equal to the fraction of the “+” primers whose extension in the \( n \)-th cycle is limited by the target length \( l_i \). In this case, one finds that it is just the fraction of the “−” strands synthesized in all previous cycles which are sufficiently long to bind a “+” primer; again, this fraction is also out of the entire DNA content of the system. Since the length of the synthesized “−” strands is limited by \( l_i \) or \( L^- \), and \( L^- > l_i \), this fraction is given by the partial sum of \( \Pi_{n-1}(l) \) over the interval between \( l_i - a^- \) and \( L^- \). Similar considerations show that \( \omega_{n_2}^- \) equals a partial sum of \( \Pi_{n-1}(l) \) over the interval between \( l_i - a^- \) and \( L^+ \). Hence, we have

\[
\omega_{n_2}^+ = \sum_{k=l_i-a^+}^{L^+} \Pi_{n-1}(k),
\]

\[
\omega_{n_2}^- = \sum_{k=l_i-a^-}^{L^-} \Pi_{n-1}(k).
\]

(24)

Obviously, this is a slightly idealized picture of the primer–template interactions. If necessary, it can be replaced with a more sophisticated model.

Equations (20)–(24) are the basic relations of the algorithm which we use to model PCR runs. Starting with the initial conditions of a run, the algorithm repeatedly applies these equations until the desired number of cycles is executed. The yield of the run \( \phi \) is measured by the fraction (i.e. concentration) of the target strands of both “+” and “−” type out of the entire amount of DNA content of the system. Since \( \sum_{n=0}^{N} \omega_{n_2}^+ + \omega_{n_2}^- = 1 \), this fraction is given as the sum of the probability distributions for the newly synthesized strands [cf. eqn (21)] shift left, i.e. into the region of smaller lengths, resulting in a decrease in the fraction of synthesized strands which can serve as templates. This in turn leads to a decrease in the amplification efficiency as the run progresses through the cycles.

4. Numerical Results

In this section, we report numerical simulations performed on a number of PCR runs under different reaction conditions using the algorithm developed above. In our numerical implementation of the algorithm, the threshold lengths \( a^+ \) and \( a^- \) [cf. eqn (24)] are defined as 

\[
a^+ = \frac{l_0^+}{2}, \quad a^- = \frac{l_0^-}{2}.
\]

4.1. SIMULATION OF PCR RUNS

Figure 3 is a plot of the target strand yield \( \phi \) against the cycle number for one of the runs considered. One can see that the yield tends to an asymptotic value which is less than 100% as the cycle number increases. This is a manifestation of the amplification plateau effect which is well-known from PCR experiments.

One can easily see why the asymptotic yield is less than 100%. As pointed out earlier (see discussion in Section 3), only a fraction of strands synthesized in each cycle can serve as templates in the following cycle. Furthermore, with the number of dNTPs \( n_0 \) decreasing from cycle to cycle and the cycle duration remaining constant, the maxima of the probability distributions for the newly synthesized strands [cf. eqn (21)] shift left, i.e. into the region of smaller lengths, resulting in a decrease in the fraction of synthesized strands which can serve as templates. This in turn leads to a decrease in the amplification efficiency as the run progresses through the cycles.

![Graph of the yield of target sequence \( \phi \) vs. the cycle number \( n \) in a PCR run.](image)
In order to probe sensitivity to the reaction conditions, we consider a series of runs with the total number of cycles $N$ fixed and the duration of the extension phase $\eta$ the same for all cycles in each run and increasing monotonously from run to run. In Fig. 4, we compare the final yield of these runs as a function of duration of the extension phase. We observe a very steep rise in the yield as the duration passes through a certain interval. Clearly, in the vicinity of this interval, a small change in the value of the duration, primer length, or any other parameter of the process will result in a large deviation in the final yield of the run.

These results indicate that our Markov process model, as implemented in the above algorithm, is able to reproduce the major qualitative features of the PCR amplification process. Namely, we found that given favourable reaction conditions the target sequence is amplified so that it rapidly outnumbers all side products. Furthermore, we observed that the yield of target sequence in a PCR run always tends to an asymptotic limit which is less than 100% as the cycle number increases. Finally, the amplification process was found to be highly sensitive to the parameters of PCR run.

4.2. OPTIMIZATION OF PCR RUNS

Next, we consider a number of PCR runs with different template length ($L_{\text{true}}$), primer lengths ($l_0^\text{f}$ and $l_0^\text{r}$), and primer binding site offsets ($L_{\text{off}}^\text{f}$ and $L_{\text{off}}^\text{r}$), and numerically optimize the duration of the extension phases for all cycles of each of these runs. The optimization procedure uses a simple multidimensional direction set method (see, e.g., Press et al., 1989) to search for the local maxima of the overall yield of the run as a function of duration sequence $\eta_1, \eta_2, \ldots, \eta_N$. The optimum duration sequence is chosen from these local maxima according to a trade-off between the overall yield and the total run time. Specifically, the optimum sequence is chosen so that it provides the shortest possible run time and the overall yield is within 0.1% of the highest yield found (i.e., global maximum).

In Fig. 5, the optimum duration sequences are shown for three different PCR runs. All durations are given in the usual time units (seconds), not in units of $\eta$. Conversion was done numerically using the definition of the $\eta$ variable [see also the discussion following eqn (6) in Section 2):

$$\eta(t) = \int_{0}^{t} k(t) \, dt,$$

(26)

where $k(t)$ is the rate coefficient which depends implicitly on time through temperature $T$. This dependence is specified by Arrhenius' law:

$$k(t) = k_0 e^{-E_a/R T(t)},$$

(27)
where \( k_0 \) is a constant prefactor, \( E_a \) is the molar activation energy for the dNTP addition process, and \( R \) is the universal gas constant. We assume that during the extension phase the temperature rises linearly from 315 to 345 K with the heating rate \( q \), and then remains constant until the onset of the denaturation phase,

\[
T(t) = \begin{cases} 
315 + qt & \text{for } 0 \leq t \leq 30/q, \\
345 & \text{for } t > 30/q.
\end{cases}
\] (28)

The values of the parameters \( k_0, E_a, \) and \( q \) used in the conversion from \( \eta \) to \( t \) for each of the three runs are listed in Table 1.

We observe that the optimum duration sequences computed group around straight lines, each characterized by a different slope and intercept.

5. Discussion

The Markov process approach to the PCR amplification kinetics described in this paper is specifically designed to account for the microscopic nature of the processes comprising a PCR run. The approach is based primarily on the assumption that DNA synthesis (i.e. primer extension) occurs independently (in the statistical sense) on each template strand in the system. The range of validity of this assumption is determined by the mobility of template molecules relative to that of other PCR reactants such as dNTPs and primers. In realistic PCR runs, the diffusive movement of massive, bulky templates is always much slower than that of dNTPs and primers, hence the assumption remains valid. By means of simulation of various PCR runs with the use of a simple model we were able to explain a number of well-known features of PCR amplification, of which the most important is the sensitivity of amplification efficiency to the reaction conditions.

The microscopic level of description which forms the basis of the Markov process approach provides our model with a number of important advantages. First, the model can be solved exactly by analytical means, which considerably simplifies all calculations. Second, it accounts for the fluctuations inherent in PCR kinetics through a description of their natural microscopic source—stochasticity of the primer extension process. Third, the model is easy to modify and can be used as the basis for constructing dedicated algorithms for numerical simulations of PCR.

Extensions of the model should incorporate the various factors controlling the amplification efficiency. The most obvious of such factors is the variation of temperature within each cycle of a PCR run. In order to account for temperature variation in the model the reaction mechanism for the primer extension process must include, in addition to dNTP binding, the template–template and template–primer equilibria, DNA polymerase binding and dissociation, \( \text{Mg}^{2+} \) binding, hydrolysis, and so forth. It is also necessary to specify the rate constants and activation energies for all microscopic processes involved. At present, there are novel experimental techniques, such as real-time monitoring with fluorescent probes (Wittwer et al., 1998), single-molecule fluorescence measurements (Lu et al., 1998), and optical trapping (Wang et al., 1998) which can possibly be used to obtain the necessary information on kinetic and thermodynamic parameters.

The Markov process approach can also be employed in simulating the quantitative PCR (Q-PCR), with obvious application to gene quantification problems. The necessary modifications

---

### Table 1

**Parameters used in computing optimum duration sequences**

<table>
<thead>
<tr>
<th>Datapoint symbol in Fig. 5</th>
<th>( L_{true} )</th>
<th>( L_{off}^{+} )</th>
<th>( L_{off}^{-} )</th>
<th>( l_1 )</th>
<th>( l_0^+ )</th>
<th>( l_0^- )</th>
<th>( n_0 )</th>
<th>( k_0 ), sec(^{-1} )</th>
<th>( E_{acce} ), kJ/mol</th>
<th>( q ), K/sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>○</td>
<td>3000</td>
<td>200</td>
<td>2650</td>
<td>150</td>
<td>10</td>
<td>17</td>
<td>10(^4)</td>
<td>0.0012</td>
<td>2.5</td>
<td>2.0</td>
</tr>
<tr>
<td>□</td>
<td>2200</td>
<td>500</td>
<td>725</td>
<td>975</td>
<td>20</td>
<td>22</td>
<td>5( \cdot 10^4 )</td>
<td>0.0012</td>
<td>2.5</td>
<td>2.0</td>
</tr>
<tr>
<td>△</td>
<td>4400</td>
<td>4020</td>
<td>130</td>
<td>250</td>
<td>16</td>
<td>31</td>
<td>5( \cdot 10^4 )</td>
<td>0.0012</td>
<td>2.5</td>
<td>2.0</td>
</tr>
</tbody>
</table>
involve “inverting” the approach so that, given the reaction conditions and the final distribution of products after the run, one can obtain information about the initial state of the reaction system, namely the number of copies of certain sequences in the initial DNA sample. For Markov processes, such an “inversion” can be achieved by simple manipulation of the master equations which determine their evolution.

Funding for this work was provided in part by Natural Sciences and Engineering Research Council of Canada. M.V. was supported by a Connaught Scholarship.

REFERENCES


