A kinetic model for equilibrium binding of intercalation dye with DNA during its amplification

A A Fedorov¹ and A V Murav’ev²

¹ Institute for Analytical Instrumentation, Russian Academy of Sciences, 31-33 lit. A Ivana Chernyh st., St. Petersburg 198095, Russia
² Peter The Great St. Petersburg Polytechnic University, 29 Politekhnicheskaya st., St. Petersburg 195251, Russia

E-mail: f_aa@mail.ru

Abstract. Recently, a number of models of polymerase chain reaction (PCR) have been proposed. In such models, the kinetics of a reaction product and a fluorescence reporter used are usually identified. An intercalation dye is one of the two ways for “real time” detection of reaction product accumulation. In this work a PCR model is proposed, with using of intercalation dye to obtain a PCR curve. The concentration of bound intercalation dye was calculated at the end of each PCR cycle. This calculation is based on an equilibrium equation for intercalation dye binding with DNA. An analysis of simulation results demonstrates a significant difference between the accumulation kinetics of the reaction product and the fluorescent signal of the bound intercalation dye.

1. Introduction

For more than thirty years, the method of polymerase chain reaction (PCR) serves as an analytical tool in genetic research. In the process of real-time PCR, the synthesis of DNA chains is accompanied by an increase in the fluorescence signal of a fluorescent reporter which allows to register reaction product accumulation at each reaction cycle. The effect of reaction components to the resulting PCR curve is well studied at theoretical and practical levels. A number of PCR models have been proposed that can analyze some aspects of the reaction at semi quantitative level [1-7]. However, a comprehensive quantitative model of the PCR process has not been proposed yet. It is the result of PCR being multicomponent and multistage reaction with complex kinetics and many parameters. Such a model can help development of both PCR reagents and PCR equipment.

Establishing correlation between the fluorescent reporter signal and the concentration of the reaction product is one of the problems that should be solved. Most of the existing models are based on explicit or implicit assumption on the identity of the experimental and model kinetic PCR curves (the former are the curves of the fluorescence signal accumulation, and the latter are the curves of the amplified DNA fragment copy accumulation). In our opinion, such identity can be questioned. In [8] we have demonstrated that the PCR product kinetics and fluorescent signal of an oligonucleotide probe can have significant differences.

In this paper we propose a PCR model in which fluorescent signal is generated by an intercalation dye (ID). ID has a completely different way to generate signal compared to an oligonucleotide probe, as a result the kinetics of ID fluorescence signal is different. This difference must be taken into account when developing mathematical methods of PCR analysis, as well as mathematical models of PCR.
Intercalating dyes are a family of fluorescent organic compounds capable of non-covalent binding to double-stranded DNA with dramatic increase of its fluorescence response. The dye SYBR Green I with an emission wavelength of 520 nm which is currently widely used in PCR is capable of enhancing fluorescence by more than 100 times after binding to double-stranded DNA [9]. The main disadvantage of using ID for real-time PCR is their ability to bind to any double-stranded DNA in the reaction mixture, as a result fluorescent signal can be generated by both the target PCR product and non-specific reaction byproducts. Therefore, to validate the results, the presence of the target amplification product is usually additionally confirmed using temperature denaturation (“melting”) of the amplicons.

2. The proposed model
The PCR model proposed is based on chemical kinetics equations. The model reactions of each PCR cycle are shown in Fig. 1.

A

\[ F + Pr \xleftrightarrow{k_i} S + F' + Pr' \xleftrightarrow{k_i} S', \]
\[ F + F' \xrightarrow{k_f} P_2, \]

B

\[ F + S' \xrightarrow{k_i} P_2 + Pr', F' + S \xrightarrow{k_i} P_2 + Pr. \]
\[ S + S' \xrightarrow{k_f} P_2 + Pr', Pr \rightarrow P_1. \]

C

\[ S \xrightarrow{k_i} P_1, S' \xrightarrow{k_i} P_1. \]

Figure 1. The model reactions. (A) reversible hybridization processes, (B) fragment synthesis by Taq polymerase, (C) irreversible complementary chains hybridization.

The model considers the following reaction components:
- \( F \) – amplified single stranded DNA fragment, \( F' \) – complementary amplified single stranded DNA fragment;
- \( Pr \) – primer to DNA strand \( F \), \( Pr' \) – primer to DNA strand \( F' \);
- \( S \) – complex of fragment \( F \) and primer \( Pr \), \( S' \) – complex of fragment \( F' \) and primer \( Pr' \);
- \( P_1 \) – double-strand fragment produced by synthesis;
- \( P_2 \) – duplex (double-strand fragment produced by hybridization of initial complementary chains).

Since the reaction products of both complementary chains \( F \) and \( F' \) are identical, the equations describing kinetics of these reactions can be simplified by replacing the concentration of \( F \) by \( F' \), \( Pr \) by \( Pr' \) and \( S \) by \( S' \). The enzymatic synthesis is considered to follow Michaelis-Menten kinetics. Michaelis-Menten parameters (the maximal reaction rate \( V_{\text{max}} \) and Michaelis constant \( K_M \) for Taq polymerase) can be estimated as \( 1 \times 10^{-8} \text{ M} \cdot \text{c}^{-1} \) and \( 2 \times 10^{-9} \text{ M} \), respectively [8].

The system of ordinary differential equations describing the reaction kinetics are as follows:

\[
\frac{dC_F}{dt} = -k_1 C_F C_{Pr} + k_{-1} C_S + k_2 C_F C_F - k_2 C_S C_F 
\]  
(1)

\[
\frac{dC_{Pr}}{dt} = -k_1 C_F C_{Pr} + k_{-1} C_S + k_2 C_S C_F + k_2 C_S C_S 
\]  
(2)

\[
\frac{dC_F'}{dt} = k_1 C_F C_{Pr} - k_{-1} C_S - k_2 C_S C_F - k_2 C_S C_S - k_m C_S 
\]  
(3)
\[
\begin{align*}
\frac{dC_{F2}}{dt} &= k_2 C_F C_F + 2k_3 C_S C_F + k_4 C_S C_S \\
\frac{dC_{F1}}{dt} &= \frac{V_{max} C_S}{K_M + C_S}
\end{align*}
\]  
(4)

where \( C_F \) – the concentration of reaction component \( F \).

The detectable fluorescent signal was generated by bound \( ID \). The concentration of bound \( ID \) was calculated based on equilibrium assumption of \( ID \) binding. The ID dissociation constant \( k_d^{ID} \), according to [9], can be estimated as 45 nM. This constant can be expressed as:

\[
k_d^{\text{free}} = \frac{C_{ID}^{\text{free}} \cdot C_{site}^{\text{bound}}}{C_{ID}^{\text{bound}}}
\]  
(5)

where \( C_{ID}^{\text{free}} \) – the concentration of free \( ID \), \( C_{ID}^{\text{bound}} \) – the concentration of bound \( IC \), \( C_{site} \) – the total concentration of free binding sites for \( IC \).

The concentration of \( IC \) binding sites \( C_{site} \) depends on the size of the DNA fragment amplified in the reaction. One bound \( ID \) molecule occupies a DNA segment of about five nucleotide length, so the parameter \( C_{site} \) can be expressed as \( 0.2 \cdot N \cdot C_S \), where \( N \) – the nucleotide length of amplified DNA, \( C_S \) – the total concentration all double stranded DNA molecules in reaction mixture. The \( C_S \) can be calculated as \( C_S = 2 \cdot C_{Pr1} + C_{Pr2} \).

The current concentration of free \( ID \) and the concentration of free \( ID \) binding sites are calculated as follows:

\[
C_{site} = 0.2 \cdot N \cdot C_S - C_{IC}^{\text{bound}}
\]  
(7)

\[
C_{IC}^{\text{free}} = C_{IC}^{\text{full}} - C_{IC}^{\text{bound}}
\]  
(8)

Substituting Eqs.7,8 into Eq.6 we get a quadratic equation for \( C_{IC}^{\text{bound}} \). This equation has the following solution for calculating \( C_{IC}^{\text{bound}} \):

\[
C_{IC}^{\text{bound}} = \frac{(0.2 NC_S + k_d^{IC} + C_{IC}^{\text{full}}) - \sqrt{(0.2 NC_S + k_d^{IC} + C_{IC}^{\text{full}})^2 - 0.8 NC_S C_{IC}^{\text{full}}}}{2}
\]  
(9)

The simulation process for Eqs.1 was performed using an original C++ computer program. Eqs. (1-5) were solved by the fourth-order Runge–Kutta procedure. The program performs serial calculations to simulate the required number of cycles. For the first PCR cycle, we entered the initial concentrations of chains \( C_F(0) \), primers \( C_{Pr}(0) \), and \( ID C_{ID}(0) \), and elongation stage duration \( T \). At the end of the cycle, the program calculated concentrations of synthesized DNA fragments and the concentration of bound \( ID \) by Eq.(4) which allowed to obtain one point at the simulated PCR curve. For the next cycle, the amount of initial DNA is increased by the number of DNA copies synthesized at previous cycle.

After the end of each cycle, the amount of \( ID \) is returned to its original value, thus assuming that the \( ID \) did not degrade during PCR.

3. The model experimental validation

Experiment data for validation of the proposed model was obtained using real-time PCR kit for analysis of human IgG receptor IIIa gene segment coding Fc fragment (FCGR3A, GenBank NG_009066.1) manufactured by Syntol (Moscow, Russia). The direct primer sequence: CACATATTACAGAATGGCAAAGG, the reverse primer sequence: CTGAAGACACATTTTTACTCCCAA. The concentration of each primer in the reaction mixture \( C_{Pr}(0) = 200 \) nM, the enzyme concentration was 5 units per reaction, the initial amount of DNA in model samples was \( 10^7 \) DNA molecules per reaction. The reaction volume was 25 µl.
SYBR GREEN I at a concentration of 200 nM was used as an intercalation dye. All PCR experiments were performed using CFX96 PCR machine (Bio-Rad, USA). Termocycling protocol consisted of preliminary stage when the reaction mixture was kept at a temperature of 95 °C for 129 sec, followed by cycles of cooling to 62 °C for 30 sec and heating to 95°C for 15 sec. Number of cycles was 50. For convenience of analysis, all experimental curves expressed in arbitrary fluorescence units by the device, as well as the simulated curves were normalized to the maximum.

In Figure 2 experimental and theoretical PCR curves for samples of $10^7$ DNA molecules are compared. It can be seen that the model curve approximates the experimental PCR curve quite well.

Figure 2. Experimental (solid line) and model (squares) real-time PCR curves for $10^7$ DNA molecules in the analyzed sample. The model curve corresponds to the amount of bound ID at a given reaction cycle. The vertical axis shows the values normalized to the maximum.

In Figure 3 the model kinetic curves for the fluorescent signal accumulation and the accumulation of amplified DNA fragments are compared. As can be seen from the data presented, the kinetics of
these curves is different. The curve of the fluorescent signal lags behind the curve of amplicon synthesis by almost 1 cycle.

Thus, if we assume that the kinetics of accumulation of the reaction product from cycle to cycle corresponds to the registered fluorescent curve, the concentration of initial DNA estimated will be almost two times higher than the real one. In this regard, when approximating the experimental data obtained using ID, a special model should be used that reflects the dynamics of the ID fluorescent signal.

4. The results and conclusion
The presented results demonstrate a high accuracy of approximation of the experimental PCR curves, which indicates the adequacy of the proposed model. An important result obtained using the model is the demonstration that the accumulation curves of DNA and the signal of bound ID are not identical and cannot be obtained by simple linear recalculation for the entire range of the PCR curve. It should be noted that the accuracy of the approximation is somewhat lower than for the previously published model with a fluorescent probe [8], which can be explained by a number of effects not taken into account in the present model, such as the cooperativity of ID binding, inhibition of the reaction by ID and photobleaching of the dye.

Acknowledgements
The work was supported by the Ministry of Science and Higher Education of the Russian Federation, state assignment 075-00780-19-02, project No. 0074-2019-0013.

References
[1] Mehra S and Hu W S 2005 Biotechnol. Bioeng. 91 848
[2] Cobbs G 2012 BMC bioinformatics 13 203
[3] Marimuthu K, Jing C and Chakrabarti R 2014 Biophys. J. 107 1731
[4] Belov D A, Belov Yu V and Kurochkin V E 2019 J. Phys. Conf. Ser. 1400 033021
[5] Lievens A, Van Aelst S, Van den Bulcke M and Goetghebeur E 2012 PLoS One 11 e47112
[6] Rutledge R G and Stewart D 2010 PLoS One 3 e9731
[7] Liu W and Saint D A 2002 Anal. Biochem. 302 52
[8] Fedorov A A, Sochivko D G, Varlamov D A and Kurochkin V E 2018 Tech. Phys. 63 1430
[9] Dragan A I, Pavlovic R, McGivney J B, Casas-Finet J R, Bishop E S, Strouse R J, Schenerman M A and Geddes C D 2012 J. Fluoresc. 22 1189