A mathematical model of the Pyrosequencing reaction system

Anna Svantesson[a], Pål O. Westermark[a], Jeanette Hellgren Kotaleski[a], Baback
Gharizadeh[b], Anders Lansner[a], Pål Nyrén[b],*

[a] PSCI/SANS, Department of Numerical Analysis and Computer Science, Royal Institute of Technology (KTH), SE-100 44 Stockholm, Sweden
[b] Department of Biotechnology, AlbaNova University Center, Royal Institute of Technology (KTH), SE-106 91 Stockholm, Sweden

Abstract

The Pyrosequencing™ technology is a newly developed DNA sequencing method that monitors DNA nucleotide incorporation in real-time. A set of coupled enzymatic reactions, together with bioluminescence, detects incorporated nucleotides in the form of light pulses, yielding a characteristic light profile. In this study, a biochemical model of the Pyrosequencing reaction system is suggested and implemented. The model is constructed utilizing an assumption of irreversible Michaelis–Menten rate equations and a constant incorporation efficiency. The kinetic parameters are studied and values are chosen to obtain as reliable simulation results as possible. The results presented here show strong resemblance with real experiments. The model is able to capture the dynamics of a single light pulse with great accuracy, as well as the overall characteristics of a whole pyrogram™. The plus- and minus-shift effects observed in experiments are successfully reconstructed by two constant efficiency factors. Furthermore, pulse broadening can partly be explained by apyrase inhibition and successive dilution.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Biochemical modeling; DNA sequencing; Enzyme kinetics; Michaelis–Menten; Pyrosequencing technology

1. Introduction

With the discovery of the structure of the DNA molecule in the 1950s [1], the biotechnological development started gaining momentum. In the 1970s, the first methods for sequencing DNA were developed [2,3]. This allowed scientists to determine the precise order of the nucleotides in a DNA strand. Today DNA sequencing is a widely used tool in as varying fields as medicine, forensic studies and agriculture. The Pyrosequencing™ technology [4,5] is a highly applicable DNA sequencing technique, where synthesis is monitored by four different enzymes and detected through bioluminescence in the form of a light pulse.

The Pyrosequencing method has emerged as an alternative to the traditional techniques and its strength lies in the sequencing of shorter DNA fragments, typically with lengths up to 100 base pairs [6].

The four enzymes utilized in the Pyrosequencing reaction system are the DNA polymerase I Klenow fragment (KF) without the exonuclease activity (exo⁻), ATP sulfurylase, luciferase and apyrase. The source and enzyme E.C. number of the enzymes are listed in Table 1. The chemical reactions, which are catalyzed by these enzymes, are summarized below.

$$(\text{DNA})_n + \text{dNTP} \xrightarrow{\textit{DNA polymerase}} (\text{DNA})_{n+1} + \text{PP}_i \tag{1}$$

$$\text{PP}_i + \text{APS} \xrightarrow{\textit{ATP sulfurylase}} \text{ATP} + \text{SO}_4^{2-} \tag{2}$$

$$\text{ATP} + \text{D}-\text{luciferin} + \text{O}_2 \xrightarrow{\textit{luciferase}} \text{AMP} + \text{PP}_i + \text{O}-\text{luciferin} + \text{CO}_2 + h\nu \tag{3}$$

$$\text{ATP} \xrightarrow{\textit{apyrase}} \text{ADP} + \text{P}_i \xrightarrow{\textit{apyrase}} \text{AMP} + \text{P}_i \tag{4}$$

$$\text{dNTP} \xrightarrow{\textit{apyrase}} \text{dNDP} + \text{P}_i \xrightarrow{\textit{apyrase}} \text{dNMP} + \text{P}_i \tag{5}$$
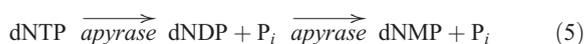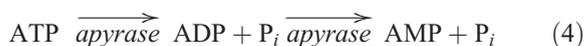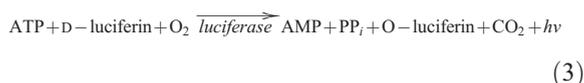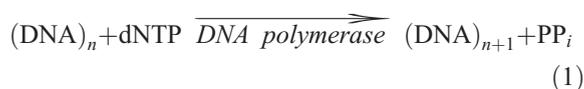
Here N in the nucleotides dNTP, dNDP and dNMP stands for one of the four nitrogenous bases A (adenine), C (cytosine), G (guanine) and T (thymine). In experiments, however, a modified A nucleotide (dATPαS) is utilized. The reason for this is that natural dATP interferes with the luciferase detection reaction (3), whereas dATPαS has been shown not to interfere [4]. In this paper, the special properties of the modified version will be considered, but no distinction in notation between the two A nucleotides will be made.

The single-stranded and primed DNA fragment to be investigated is mixed with the four enzymes and the enzyme substrates adenosine 5′ phosphosulfate (APS) and D-luciferin. The nucleotides dATP, dCTP, dGTP and dTTP are added one at a time in a cyclic manner to the reaction mixture with a 60-s interval. This may, or may not, trigger a chain of reactions (1)–(3), depending on the added nucleotide. If the nucleotide is complementary to the next DNA base, the nucleotide is incorporated by the polymerase and the primer is extended (1). This produces inorganic pyrophosphate PP_i, which is transformed by ATP sulfurylase in reaction (2), producing ATP. The luciferase (3) detects the produced ATP, thereby generating a flash of light, which is then recorded by a CCD camera. The intensity of the light pulse is proportional to the amount of produced ATP. If the nucleotides are not complementary, then no incorporation takes place and hence reactions (1)–(3) do not occur, and the flash of light fails to be emitted. By analyzing the resulting light profile (pyrogram™), the exact DNA sequence can be established.

The fourth enzyme, apyrase, removes the remaining ATP and unincorporated nucleotides (4) and (5) before the new nucleotides are added. This step is highly important since remains would otherwise interfere with the process. This and other aspects concerning the Pyrosequencing technology are described in Section 2.3.1. In this study, a mathematical model of the Pyrosequencing reaction system is suggested. This includes the study of the result from the incorporation of a single nucleotide, and careful considerations of the factors influencing the properties of the resulting light pulse. It also includes a corresponding study of several nucleotide incorporations, which produce a full pyrogram. Once a reliable model has been developed, it can be utilized to find possible ways of improving the Pyrosequencing process itself. This could, for instance, involve modification of the substrate concentrations or use of other enzymes with different kinetic properties. Performing full tests like these in the laboratory is often both expensive and time-consuming, which is

Table 1
The four enzymes in the Pyrosequencing reaction system

Enzyme	Source	E.C.
DNA polymerase I (exo⁻ KF)	*Escherichia coli* (bacteria)	2.7.7.7
ATP sulfurylase	*Saccharomyces cerevisiae* (yeast)	2.7.7.4
Luciferase	*Photinus pyralis* (firefly)	1.13.12.7
Apyrase (Pimpernel type)	*Solanum tuberosum* (potato)	3.6.1.5

why an in silico model may help to isolate especially promising configurations.

2. Model description

Two different models of the Pyrosequencing reaction system, built on the same principle, have been implemented. The first model simulates single nucleotide incorporation and the other, which is a further development of the former one, is used to simulate several nucleotide incorporations. Both models assume irreversible Michaelis–Menten kinetics for all four enzymes, which is motivated by the fact that the equilibrium constants K_{eq} for the enzymatic reactions are known to be large. Measurements and calculations on the polymerase [7] show that the equilibrium constant lies in the interval 1000–3000, and the synthesis of DNA is, in the present study, considered irreversible. ATP sulfurylase has an equilibrium constant of $\sim 10^8$ [8], which is very large. The luciferase reaction, which is photon producing, can also be considered irreversible, since when light has been created and detected it is not likely to reverse. More detailed models of the luciferase reaction have previously been proposed [9,10]. However, the simple irreversible Michaelis–Menten model suggested here was able to reproduce experimental behavior, and hence it was considered sufficient. Apyrase is generally regarded as acting in the forward direction of reactions (4) and (5) [11]. This is due to the high-energy phosphoanhydride bonds formed between the inorganic phosphates P_i and the nucleotide, which makes the reaction direction of (4) and (5) energetically favorable.

2.1. The one-pulse-model

The purpose of the so-called one-pulse-model is to capture the dynamics of the light signal that is produced after a single nucleotide incorporation. In this section, the theoretical background of the one-pulse-model is described in detail.

An overview of the model is presented in Fig. 1, which includes all relevant substrates and products, the four enzymes, reaction fluxes v_i and normalized variable names $\alpha-\varphi$. The normalized variables for the relevant substrates and products (total concen-
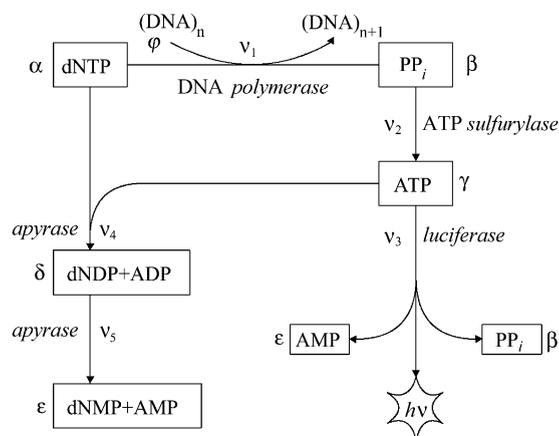

Fig. 1. Schematic overview of the Pyrosequencing reaction system as it is described in the one-pulse-model. The reaction fluxes are denoted by v_i, and Greek variables are utilized for the normalized substrate concentrations.

trations) are, following convention, defined as $\alpha = [dNTP]/K_{m,pol}^{dNTP}$, $\beta = [PP_i]/K_{m,Asul}$, $\gamma = [ATP]/K_{m,luc}$, $\delta = [(d)NDP]/K_{m,apyr}^{(d)NDP}$, $\varepsilon = [(d)NMP]/K_{I,apyr}^{(d)NMP}$, $\varphi = [(DNA)_n]/K_{m,pol}^{DNA}$. Here, K_m are the Michaelis constants of the enzymes with respect to each substrate, and (d)NDP stands for the sum of dNDP and ADP. The equivalent notation is applied for (d)NMP. Note that the definition of ε is made with an inhibition constant K_I in the denominator instead of a Michaelis constant. Next, a mathematical model of the system outlined in Fig. 1 is developed. The treatment follows standard enzyme kinetics, and a description of the methods used here can, for instance, be found in Cornish-Bowden [12].

2.1.1. Equations

The dynamics of the system are captured in a set of differential equations describing the change in the concentrations of each substrate. These equations, which are easily verified by looking at Fig. 1, are:

$$\frac{d\alpha}{dt} = \frac{1}{K_{m,pol}^{dNTP}} \left(-v_1 - fv_4 \right) \tag{6}$$

$$\frac{d\beta}{dt} = \frac{1}{K_{m,Asul}} \left(v_1 - v_2 + v_3 \right) \tag{7}$$

$$\frac{d\gamma}{dt} = \frac{1}{K_{m,luc}}(v_2 - v_3 - (1-f)v_4) \tag{8}$$

$$\frac{d\delta}{dt} = \frac{1}{K_{m,apyr}^{(d)NDP}}(v_4 - v_5) \tag{9}$$

$$\frac{d\varphi}{dt} = -\frac{v_1}{K_{m,pol}^{DNA}} \tag{10}$$

$$\frac{d\varepsilon}{dt} = \frac{1}{K_{l,apyr}^{(d)NMP}}(v_3 + v_5) \tag{11}$$

Here, $f = [dNTP]/([dNTP]+[ATP])$ is the fraction of dNTP of the total amount of triphosphates. Using this expression it is thus assumed that all triphosphates bind equally well to the apyrase.

The reaction fluxes v_i are given by irreversible Michaelis–Menten rate expressions. However, there is a discrepancy from the normal assumptions made for these types of equations. The Pyrosequencing process is operating at nucleotide concentrations lower than the polymerase concentration. As a consequence, a different description of the polymerase is needed, and the one utilized in this study is presented in Fig. 2. At a given time, the polymerase can be found partly as free enzyme E_{free} (E in Fig. 2), and partly bound in different complexes with free DNA, $DNA_{free}(S_1)$, and/or free nucleotide,
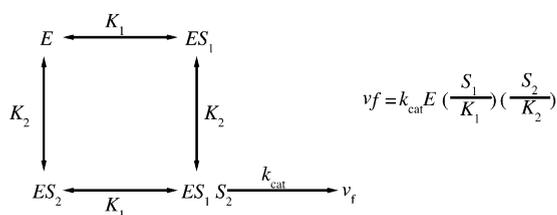
Fig. 2. Enzymatic model of the DNA polymerase at equilibrium. The enzyme E has two substrates S_1 (DNA) and S_2 (dNTP), with which it can form complexes. The final complex ES_1S_2 can be formed by binding either one of the two substrates prior to the other, with equal probability. The dissociation constants are given by K_1 and K_2. The rate expression is given by the forward rate v_f, where k_{cat} is the catalytic constant.

$dNTP_{free}(S_2)$. The complex containing all three units is defined as:

$$E_{free}DNA_{free}dNTP_{free} = \frac{E_{free} \cdot DNA_{free} \cdot dNTP_{free}}{K_{DNA} \cdot K_{dNTP}} \tag{12}$$

where K refers to each compound's dissociation constant. Normalized variables for the available DNA and dNTP concentrations are introduced as follows: $\alpha_{free}=[dNTP_{free}]/K_{dNTP}$ and $\varphi_{free}=[DNA_{free}]/K_{DNA}$.[1] The modified rate expression is then given by the forward reaction flux v_f in Fig. 2, and the expression is directly proportional to the concentration of the triple complex.

Moreover, the total amount of polymerase is consequently equal to the sum of free enzyme and all possible complexes formed between the enzyme and its substrates at a given time. The equivalence is applied for the two substrates. In steady-state, this results in three algebraic Eqs. (13)–(15) that complement the differential Eqs. (6) and (11). Together they form a system of differential-algebraic equations (DAEs). The notation for this model is that variables with the subscript free refer to free available concentrations and variables without the subscript refer to the total amount, i.e. a fixed amount.

$$E_{TOT,pol} = E_{free}(1 + \varphi_{free} + \alpha_{free} + \varphi_{free}\alpha_{free}) \tag{13}$$

$$K_{m,pol}^{DNA}\varphi = \varphi_{free}\left(K_{m,pol}^{DNA} + E_{free} + E_{free}\alpha_{free}\right) \tag{14}$$

$$K_{m,pol}^{dNTP}\alpha = \alpha_{free}\left(K_{m,pol}^{dNTP} + E_{free} + E_{free}\varphi_{free}\right) \tag{15}$$

The different rate expressions can now be written as:

$$v_1 = k_{cat,pol} \cdot E_{free} \cdot \varphi_{free} \cdot \alpha_{free} \tag{16}$$

---

[1] In this model it is assumed that $K \approx K_m$ for DNA and dNTP, and no distinction between the constants will be made in the following treatment.

$$v_2 = \frac{V_{\mathrm{Asul}}\beta}{1 + \beta} \tag{17}$$

$$v_3 = \frac{V_{\mathrm{luc}}\gamma}{1 + \gamma} \tag{18}$$

$$v_4 = \frac{V_{\mathrm{apyr}}^{\mathrm{NTP}}(k_1\alpha + k_2\gamma)}{1 + k_1\alpha + k_2\gamma + \delta + \varepsilon} \tag{19}$$

$$v_5 = \frac{V_{\mathrm{apyr}}^{\mathrm{NDP}}\delta}{1 + k_1\alpha + k_2\gamma + \delta + \varepsilon} \tag{20}$$

where V is the maximum velocity of the enzyme, defined as $k_{\mathrm{cat}}{\cdot}E$, and k_{cat} is the catalytic constant. Moreover, $k_1 = K_{\mathrm{m,pol}}^{\mathrm{dNTP}}/K_{\mathrm{m,apyr}}^{\mathrm{(d)NTP}}$ and $k_2 = K_{\mathrm{m,luc}}/K_{\mathrm{m,apyr}}^{\mathrm{(d)NTP}}$. These factors are required since the triphosphate variables are defined with respect to the polymerase and the luciferase respectively, and in the expressions of v_4 and v_5 they should be with respect to the apyrase.

In this model, apyrase has three different substrates, namely, dNTP, ATP and (d)NDP. These three pools of compounds are treated as competing substrates, which all have the ability to bind the same active site on the enzyme [13]. In addition, the monophosphates are handled as product inhibitors. Both these inhibitory properties are included in the model as extra terms in the denominators of the expressions of v_4 and v_5 [12]. Here it is assumed that only a small amount of dNTP is found in polymerase complexes and that there is an excess of substrates.

The model output, i.e. the luciferase generated light signal, is taken to be proportional to the ATP reaction flux v_3. The motivation is that experiments show that the light response is linear for incorporations of up to 5–6 identical nucleotides [14,15]. For longer multiples the light response is non-linear. However, for the validation of this model, it is sufficient with a maximum multiple length of 5–6. The assumption of proportionality allows for an optional scaling factor. This scaling factor was introduced in order to obtain a light pulse of convenient magnitude.

2.2. Pulse characteristics

In order to examine the light pulse (a typical light pulse can be found in Fig. 4a), and to evaluate the agreement of a simulated pulse with an experimentally determined pulse, some kind of measures of the characteristics of a pulse need to be specified. In this study, three such characteristics are used. They are:

- Pulse height, defined as the maximum value of the pulse.
- Rise time, defined as the time from nucleotide dispensation until the pulse reaches its maximum value.
- Decay time, defined as the time from nucleotide dispensation until the pulse has dropped to 10% of its maximum value. If that value cannot be specified in the given time period (60 s in this study), the value is taken to be the whole cycle time.

2.3. A larger model

Decoding a whole DNA fragment, which involves the sequencing of several bases, introduces some interesting complications. In order to accommodate these complications the model described in the previous section was expanded. First, some aspects regarding full Pyrosequencing are mentioned and then the larger model implementation is described.

2.3.1. Some aspects on the Pyrosequencing technology

During the sequencing process, it is common that template bases appear in multiples. The kinetics of the polymerase is constructed so that the whole multiple is read practically at the same time, resulting in a light pulse that is of multiple intensity [4,15]. This is motivated by the fact that a multiple amount of PP$_i$ is formed by reaction (2) and propagated to reaction (3), in that way causing more light to be produced. The CCD camera then detects the multiple incorporations as light of increased intensity integrated each second.

Furthermore, the incorporation efficiency of the polymerase is not 100%, but perhaps instead 99%. This leads to DNA remains that can act as substrates the next time a matching nucleotide is added to the reaction mixture. This effect is called minus-shift and is most prominent after the incorporation of consecutive identical bases.

Additionally, there is an effect called plus-shift. Just as the minus-shifts are caused by templates lagging behind, the plus-shifts are caused by primers extended beyond the active template base. The plus-shifts are mainly due to the apyrase's incomplete removal of the unincorporated nucleotides. Remaining nucleotides are able to react with the following template if they match, independent of the nucleotide type added.

Another effect that should be considered is the dilution that occurs, as more and more nucleotide drops are dispensed as sequencing proceeds. Every 60 s, approximately 0.2 μl liquid is added to the reaction volume (initially 50 μl). The dilution effect results in a decrease in enzyme activity of all four enzymes and a decrease in substrate concentrations as well. The dilution effect might to some extent be counteracted by evaporation.

Pyrosequencing experiments reveal a broadening of the light pulses as time evolves. This can be explained by a continuously decreasing activity of the apyrase [6]. The apyrase activity is lowered by dilution and product accumulations. Product accumulation occurs for all given products, but is most noticeable for the monophosphates (since these are never degraded). In addition, it is believed that dADP is harder to hydrolyze than the other diphosphates, because a different kind of nucleotide (dATPαS instead of dATP) is used, and that (the modified) dADP has an inhibitory effect on the apyrase [6].

2.3.2. The model

In this subsection, the one-pulse-model is expanded. This involves separating the nucleotide pool dNTP into its four different types dATP, dCTP, dGTP and dTTP. The same is done with $(DNA)_n$ based on active template type. The pools of di- and monophosphates are also separated, but the C, G and T contributions are brought together. Hence, the three nucleotide types are with this given similar properties. A schematic overview of this larger model is given in Fig. 3, with all relevant substrates, enzymes, variables and reaction fluxes shown. All variable definitions, differential equations, algebraic equations and reaction fluxes are equivalent to those for the smaller model and are found in Appendix A.

The main difference in the two models is thus the number of nucleotide variables. Including more variables makes it possible to distinguish between the different types of nucleotides with respect to their kinetic properties and inhibitory effects. It also allows for more substrates to compete for the active site of the apyrase.

The other aspects described above have all, to some extent, been considered in the implementation of this larger model. Adding the effect of dilution and possible evaporation is straight-forward, and dilution gradually decreases the enzyme and substrate concentrations. The incorporation efficiency is included as a somewhat artificial factor, controlling the amount of
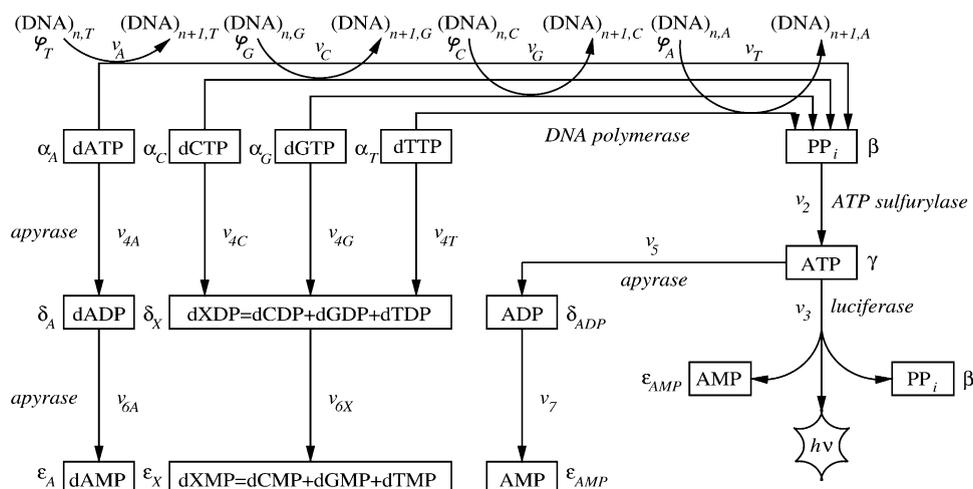

Fig. 3. Schematic overview of the Pyrosequencing reaction system as it is described in the larger model. The reaction fluxes are denoted by v_i, and Greek variables are utilized for the normalized substrate concentrations.

matching DNA available for the corresponding nucle-otide. This means that, for instance, 99% of the total amount of matching DNA is spent during incorpora-tion, and that 1% is not utilized and instead becomes a candidate for future minus-shifts. In a similar manner, the plus-shifts are simulated by an artificial factor for incorporation of accumulated nucleotides. The factor is in the order of a few percent (revealed in the simulations), and is only applied if the amount of accumulated nucleotide exceeds some threshold con-centration. Multiples in the template strand are mod-eled as an increased amount of PP_i produced. The incorporation efficiency and plus-shifts will of course affect this factor.

3. Materials and methods

3.1. Experimental methods

An experimental reference case was taken from Gharizadeh et al. [16], where human papilloma viruses (HPV) were sequenced using the Pyrose-quencing technology. For a description of the exper-imental methods and conditions, see that article. The HPV-31 amplicon will be used for comparisons be-tween experiment and simulations.

3.2. Computer simulations

The simulations were performed on a Pentium 4 workstation running under the Unix (Red Hat Linux) operating system. The systems of differential-algebra-ic equations were implemented in MATLAB [26] and

Table 2
Enzyme and substrate concentrations

Compound	Concentration (μM)
DNA polymerase	0.25
ATP sulfurylase	0.12
Luciferase	0.26
Apyrase	0.030
DNA	0.008
dATP	3.2
dCTP	0.92
dGTP	0.62
dTTP	3.2

The values were obtained from Gharizadeh et al. [16].

Table 3
Kinetic constants

Parameter	Value	Reported value	Reference
$k_{cat,pol}$ (s^{-1})	2	2–4	[17]
$k_{cat,Asul}$ (s^{-1})	33	38	[18]
$k_{cat,luc}$ (s^{-1})	0.015	0.015	[19]
$k_{cat,apyr}^{(d)NTP}$ (s^{-1})	1340	10^4	[20]
$k_{cat,apyr}^{(d)NDP}$ (s^{-1})	134	10^3	[21]
$K_{m,pol}^{dNTP}$ (μM)	0.1	0.1–0.2	[17]
$K_{m,pol}^{DNA}$ (nM)	5.1	5	[22,23]
$K_{m,Asul}$ (μM)	7	7	[18]
$K_{m,luc}$ (μM)	50	48.5±2.6	[24]
$K_{m,apyr}^{(d)NTP}$ (μM)	114	120	[25]
$K_{m,apyr}^{(d)NDP}$ (μM)	260	260	[25]
$K_{I,apyr}^{(d)NMP}$ (mM)	40	40	[25][a]

[a] The inhibition constant was extrapolated using the reported values.

solved with a stiff DAE solver. Consistent initial values for the DAEs were calculated using a non-linear least squares method.

The kinetic parameters, along with the initial substrate and enzyme concentrations utilized in the simulations are tabulated in Tables 2 and 3. The substrate and enzyme concentrations were taken to be the same as in the experimentally studied reference case. This allowed comparisons between the results of the simulations and the experiments. The kinetic constants reported in literature [17–25], listed in Table 3, were used as a starting-point for the simu-lations. These values were later adjusted with the intention of receiving a light pulse that bears as much resemblance to the experimentally determined one as possible. The values in Table 3 are a result of one such optimization.

The parameters utilized for the larger model were chosen to be the same as for the one-pulse-model, since it is desired that a single nucleotide incorpora-tion with the larger model, should produce the same result as the smaller model.

Both models were run with a nucleotide dispensa-tion time interval of 60 s, and the dilution constituted an addition of 0.2 μl for each time interval to an initial volume of 50 μl. The amount of evaporation in each simulation was chosen as a constant value from the interval of 0–0.2 μl per dispensation.

4. Results and discussion

In this section, results from simulations with the two models are presented. First, the characteristics of a single light pulse are examined and discussed with respect to the different parameters. Then the properties of a pyrogram and several nucleotide incorporations are considered.

4.1. Single nucleotide incorporation

4.1.1. Towards a realistic pulse

In Fig. 4a, a typical light pulse, obtained from the real Pyrosequencing experiment [16], can be viewed.
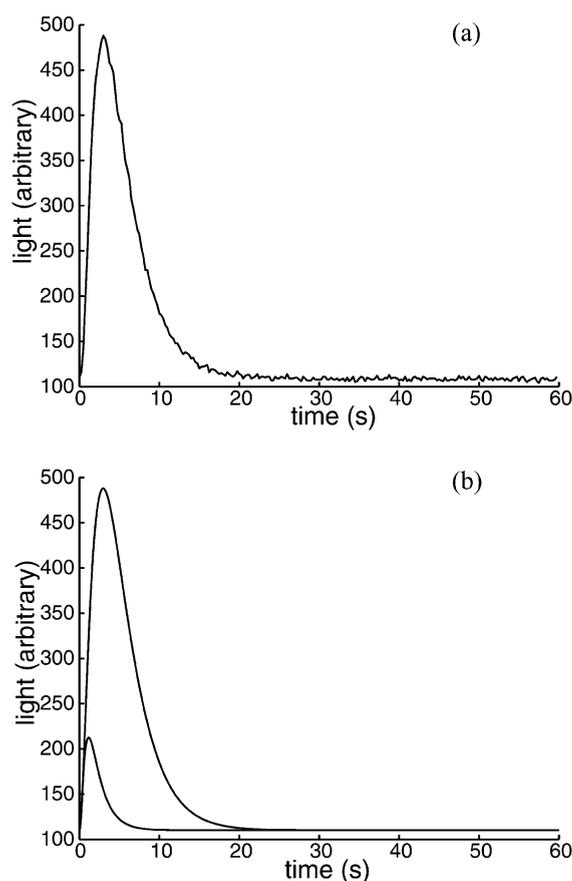

Fig. 4. (a) Experimental reference pulse, which is the second pulse in the HPV-31 pyrogram taken from Gharizadeh et al. [16]. (b) Light pulses generated with the one-pulse-model. The smaller pulse corresponds to the literature parameters and the larger to the optimized parameters in Table 3.

This pulse has been selected to act as a reference case, with which the simulated pulses are compared and fitted.

The model was first implemented with the kinetic parameters found in the literature, and the result of that simulation is depicted in Fig. 4b (smaller pulse). The resulting response is pulse-shaped but with a much faster time course compared to the experimental pulse.

To obtain a pulse more similar to the reference pulse, the parameters related to $k_{cat,pol}$, $k_{cat,apyr}$, $K_{m,pol}^{dNTP}$, $K_{m,Asul}$ and $K_{m,apyr}$ were adjusted. Largest modifications were needed for the catalytic constant of apyrase, which was almost decreased by a factor 8, see Table 3. This is not necessarily displeasing, since the uncertainty of this parameter may very well be this significant. In general, the uncertainties in the measured values are considered large, mainly due to measuring difficulties and other experimental conditions than those used in the Pyrosequencing process. For instance, the Pyrosequencing temperature, pH and Mg/Ca concentrations are different from the prevailing conditions of the experiments where the reported data were extracted. The true Pyrosequencing values may differ as much as 50–100%, or more in some cases, from the reported values. Furthermore, the system's special sensitivity of the $k_{cat,apyr}$ parameter is partly explained by the fact that the catalytic constant affects the maximum velocity V_{apyr}. Hence it is to some extent equivalent to the apyrase concentration. More apyrase results in a more efficient degradation of nucleotide and ATP, and thereby the light producing reaction (3) is shut off earlier. Less apyrase, on the other hand, results in a decreased hydrolysis rate, and a slower and more extended light pulse. If no apyrase is included at all, the light response will continue until either one of the substrates are consumed.

Using the values in Table 3, the larger light pulse in Fig. 4b was obtained, and the resemblance with the experiment is excellent. The numeric values of the pulse characteristics for the three pulses are collected in Table 4, and the similarity between the reference case and the optimized pulse is apparent. It should be emphasized that the light pulse height in the model has been scaled to better agree with the reference pulse, but since the signal is based on proportionality with the ATP flux, the unit is arbitrary. The pulse height measure is hence somewhat arbitrary and dependent on scaling.

Table 4
Pulse characteristics the three simulated pulses in Fig. 4

Pulse characteristic	Ref. case[a]	Literature pulse[b]	Optimized pulse[c]
Pulse height (arbitrary)	488	212	488
Rise time (s)	3.00	1.16	2.98
Decay time (s)	12.10	5.32	12.17

[a] Reference pulse taken from Gharizadeh et al. [16].
[b] Pulse generated with the parameters reported in Refs. [17–25].
[c] Pulse generated with the optimized values in Table 3.

4.1.2. Varying enzyme activity

How do the kinetic parameters influence the shape of the light pulse? The simulations reveal very little dependence on the luciferase properties. This is partly explained by the slow dynamics of the enzyme ($k_{cat,luc}=0.015$ s^{-1}). However, smaller $K_{m,luc}$ values produce higher pulses as expected. The three additional enzymes, on the other hand, are all involved in determining the properties of the pulse.

The effects of varying the catalytic constant of the polymerase $k_{cat,pol}$ are shown in Fig. 5a,b. As is seen,

the pulse height increases and the decay time decreases with increasing $k_{cat,pol}$ until a saturation level is reached, which is expected. However, increasing the polymerase concentration E_{pol}, causes the pulse maximum to decrease for high values of E_{pol}, but lowers the maximum if E_{pol} is small, see Fig. 5c,d. This implies that there is an interval where an optimal pulse height can be found. The values utilized in the experiments lie within this interval and are marked with an asterisk. The corresponding graph for the decay time shows an optimal E_{pol} region as well. In fact, this region coincides with that found for the pulse maxi-
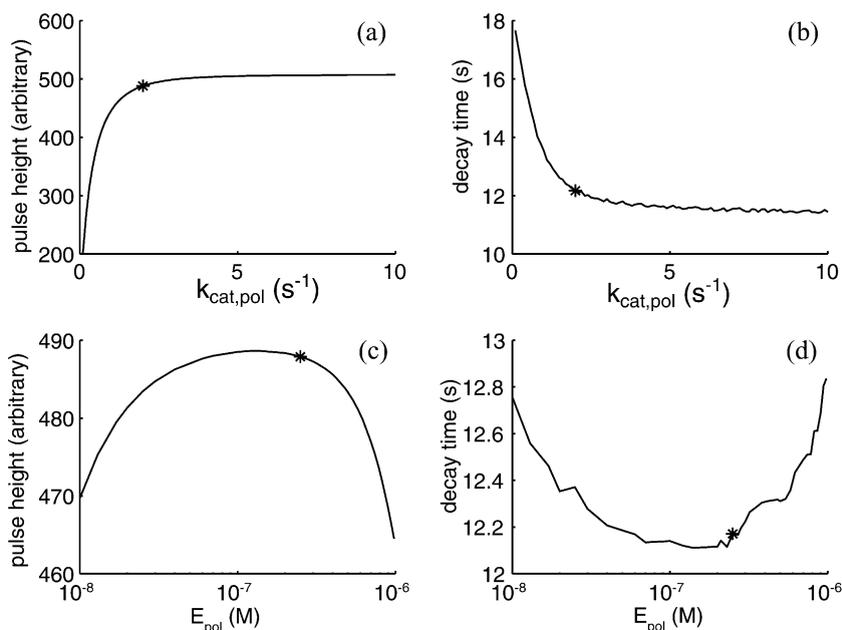
Fig. 5. (a) Maximum pulse height as function of the catalytic constant of polymerase $k_{cat,pol}$. (b) Decay time as function of the catalytic constant of polymerase $k_{cat,pol}$. (c) Maximum pulse height as function of the polymerase concentration E_{pol} (logarithmic scale). (d) Decay time as function of the polymerase concentration E_{pol} (logarithmic scale). The values utilized in the simulations are marked with an asterisk.

mum. The corresponding graph for the rise time shows an equivalent behavior (not shown). Comparing the effects of varying $k_{cat,pol}$ and E_{pol}, it is concluded that the system is more sensitive to changes in the catalytic constant, than to changes in the enzyme concentration. Moreover, the optimal region is only observed when varying E_{pol}. Future experiments will reveal whether this phenomenon is an artifact arising only in the model, caused by a poor polymerase description, or whether it is a true prediction.

Fig. 6 shows the pulse height (a), decay time (b) and rise time (c) as functions of the ATP sulfurylase maximum velocity V_{Asul} for different values of the apyrase velocity V_{apyr} [curves (1)–(7)]. For this enzyme, varying V_{Asul} can be done either by altering E_{Asul} or by altering $k_{cat,Asul}$. The observation is that the pulse maximum increases with increasing V_{Asul} and decreasing V_{apyr}. This seems likely since more, or a faster, sulfurylase will bring the ATP production forward in time and thus also the light pulse. This is observed in Fig. 6c. As a consequence, more light is produced before the apyrase starts degrading the

nucleotides to any greater extent. The same argument can be applied to the apyrase. More apyrase will give a more efficient hydrolysis of nucleotides, which results in an earlier decay of the light pulse, and hence the peak value becomes lower.

Analyzing Fig. 6a, it also becomes apparent that at very high V_{Asul} values the increase in pulse height is less pronounced. Thus the enzyme activity has reached a saturation level, above which the cost of using more enzyme (or a faster) is no longer compensated by the increase in pulse height and the decrease in decay time.

For both the decay and rise times, the time courses are decreased by increasing sulfurylase and apyrase activities. The sulfurylase saturation effect is perceptible as well. The decay time is thus fairly insensitive to changes in V_{Asul}, at least for higher velocities, which is seen in Fig. 6b. However, changes in apyrase activity substantially alter the decay time, which of course is part of the apyrase's main purpose. However, looking at Fig. 6c, the apyrase's influence on the rise time is much smaller, and furthermore it is
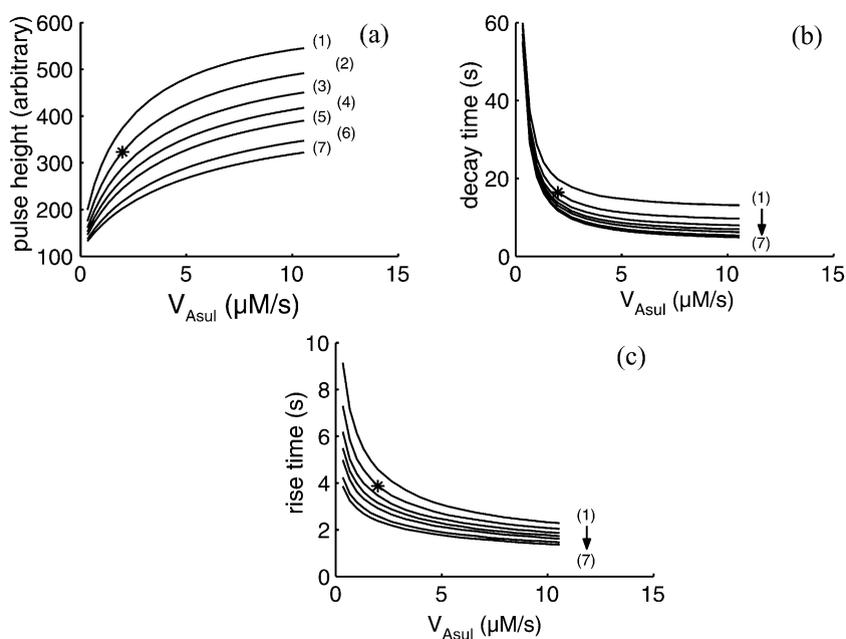
Fig. 6. (a) Maximum pulse height as function of the ATP sulfurylase velocity V_{Asul} for different values on apyrase velocity V_{apyr}. (b) Decay time as function of V_{Asul} for different V_{apyr}. (c) Rise time as function of V_{Asul} for different V_{apyr}. The V_{apyr} curves correspond to: (1) 200 μM/s; (2) 300 μM/s; (3) 400 μM/s; (4) 500 μM/s; (5) 600 μM/s; (6) 800 μM/s; and (7) 950 μM/s. The values utilized in the simulations are marked with an asterisk.

dependent on the sulfurylase velocity (larger differences between the curves for lower V_{Asul} values).

4.2. A full pyrogram

The reference light pulse in Fig. 4a is the second pulse in a full experimental pyrogram extracted from a DNA fragment run [16] of totally 34 nucleotides. The second pulse was used instead of the first in order to avoid initial calibration effects and the fact that the first pulse corresponded to an A nucleotide (the modified nucleotide) addition. The pyrogram is shown in its full length in Fig. 7a, and acted as a reference pyrogram.
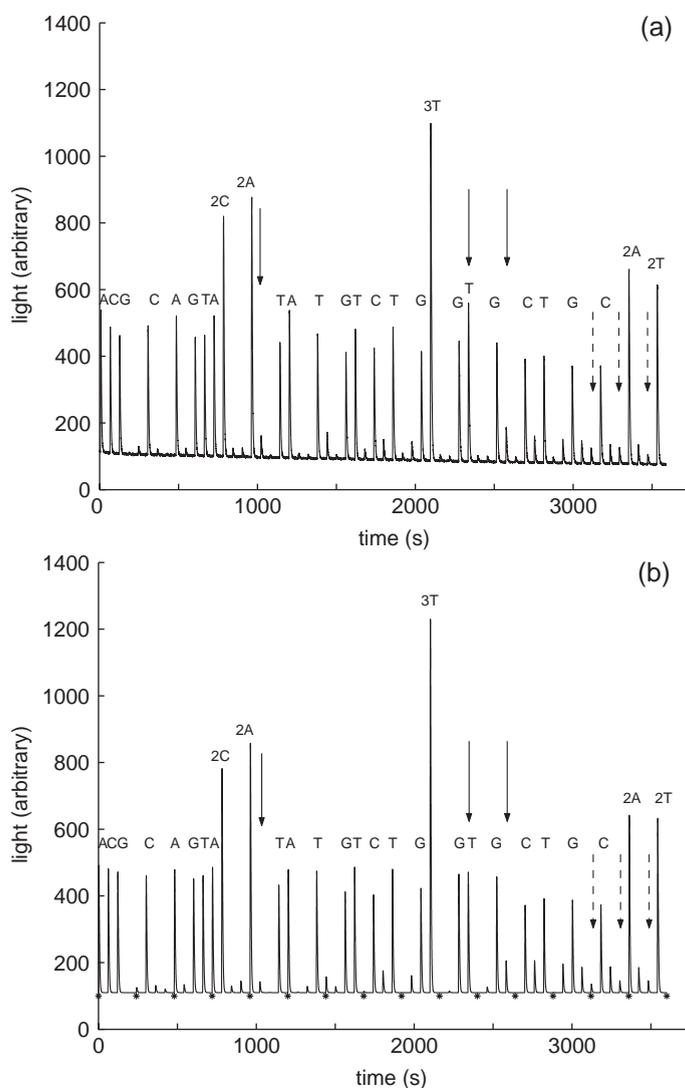
Fig. 7. (a) Experimentally determined reference pyrogram (HPV-31) taken from Gharizadeh et al. [16]. The nucleotide dispensation order is ACGT and the time between additions is 60 s. Interesting peaks discussed in the text are indicated by arrows—minus-shift, solid, and plus-shift, dashed. (b) Pyrogram simulated using an incorporation efficiency of 99% and a plus-shift effect of 3%. The nucleotide dispensation order is ACGT, and the time between additions is 60 s. The beginning of a new dispensation cycle is marked with an asterisk.
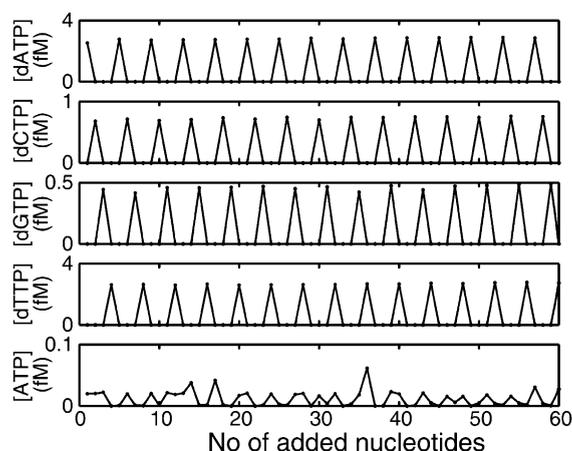
Fig. 8. Accumulated (d)NTP concentrations during a simulation with an incorporation efficiency of 99% and a plus-shift of 3%. Each point corresponds to the concentration remaining 60 s after the nucleotide addition.

4.2.1. Plus- and minus-shifts

Fig. 7b displays a simulated pyrogram corresponding to the reference case in Fig. 7a. The simulation was carried out with the incorporation efficiency set to 99%, and the plus-shift effect set to 3%. These values were chosen since they produced a quite realistic pyrogram. Comparisons between the two pyrograms reveal a rather strong resemblance between the two, with all the larger pulses placed in the right positions and of proportionally accurate heights. The fact that the maximum values do not match perfectly is considered less relevant as argued in Section 4.1.

The gradual decrease in pulse height is also captured by the model. Choosing a slightly smaller value on the incorporation efficiency will produce a more accentuated fall-off, since more and more template will lag behind and hence give less matching template. Moreover, a lower incorporation efficiency results in higher, and in an increasing number, of small pulses that eventually [after >100 dispensations (not shown)] become inseparable from the true light responses. These smaller peaks are a result of both minus- and plus-shifts. Typical minus-shifts can, for instance, be observed in Fig. 7b after the incorporation of the three T:s. The following T-dispensation (at 2340 s) is higher than it would have been without the minus-shift (cf. the next matching T at 2820 s).

Looking at the next following T-dispensation (at 2580 s), there is a pronounced, relatively large, false pulse, and the same is true directly after the first double C:s (at 1020 s). These phenomena are present in the reference pyrogram as well.

The plus-shift effect acts in the other direction and is harder to detect in the pyrograms. Two distinct examples are the preceding false pulses corresponding to the last double A:s (at 3120 s) and double T:s (at 3300 s). A plus-shift pulse can also be observed at 3480 s. This pulse originates from the triple A:s in the extra sequence GCAAA that continues the HVP-31 sequence in Fig. 7 (not shown). As time passes and more nucleotides are dispensed, the effects of plus- and minus-shifts are mixed, and serve as a general loss in true pulse heights and an increase in false pulse heights. Utilizing a plus-shift of more than 3% accentuates this effect.

4.2.2. Pulse broadening

A progressed pulse broadening has been observed in experiments. It is believed, with experimental support, that this is to a large extent caused by product accumulations and subsequent apyrase inhibition. Figs. 8–10 display the accumulated concentrations, 60 s after the nucleotide dispensation, of all nucleotide variables in the simulation as function of the number of added nucleotides.
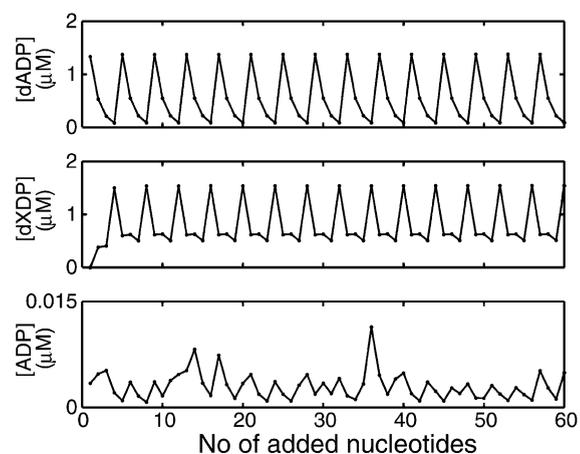

Fig. 9. Accumulated (d)NDP concentrations during a simulation with an incorporation efficiency of 99% and a plus-shift of 3%. Each point corresponds to the concentration remaining 60 s after the nucleotide addition.
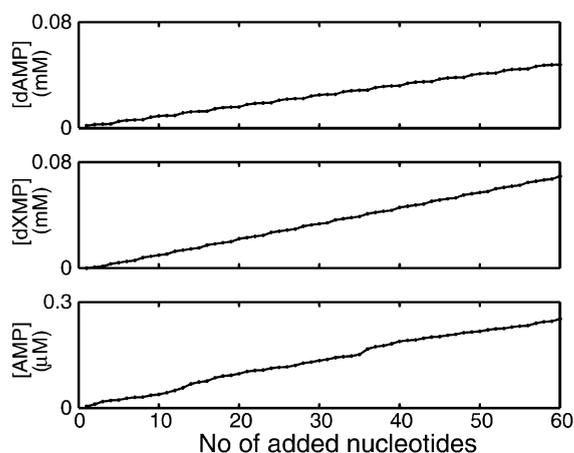
Fig. 10. Accumulated (d)NMP concentrations during a simulation with an incorporation efficiency of 99% and a plus-shift of 3%. Each point corresponds to the concentration remaining 60 s after the nucleotide addition.

The dNTP concentrations in Fig. 8 show oscillating behavior, and almost all remaining free nucleotides are incorporated by matching lagging templates at the next addition. As a consequence, no progressive accumulation occurs. The difference in the values between the nucleotide types are explained by the different amounts used for dispensation. The ATP concentration is kept at a relatively even level and no accumulation occurs there either. The small variations in the ATP graph follow the structure of the pyrogram. This means that it displays somewhat higher values at multiple incorporations, e.g. the triple T:s at the 36th nucleotide addition. Furthermore, all five triphosphate concentrations are in the order of 10^{-16} to 10^{-15} M, which is considered quite low. These pools of substrates are too small to account for the plus-shift effect, since the amount of template is in the order of 10^{-9} M. Moreover, with a Michaelis constant for the polymerase of 5.1 nM, the incorporation of such small amounts of accumulated nucleotides will be extremely slow and will not have any significant impact on the plus-shift. Hence, the accumulated dNTP is in reality most likely to be several orders of magnitude larger than that obtained in the simulations. Alternatively, there is another explanation for the plus-shift effect, but this requires further investigation.

Fig. 9 shows the corresponding plots for the diphosphate variables. Again, the oscillatory behav-

ior is present for the d-types, but everything is not decomposed during the 60 s following the addition. Despite this, no real accumulation takes place. Instead, there is a balance between the in-flux and out-flux of dNDP, which keeps the concentration at a specific level. The only exception occurs during the first four nucleotide additions where there is a short initial transient. The balanced state lies at an even level of $\sim 10^{-16}$ M. This means that the amount of dNDP produced is equal to that hydrolyzed, and since there is almost no accumulation of triphosphates, this amount should be approximately equal to the dispensed amount of dNTP, which it is. Thus the values obtained seem likely. Looking at the ADP graph, the concentration is lower and varies, again following the pyrogram structure, although now with more fluctuations. It is somewhat displeasing that the system does not show any diphosphate accumulations. The apyrase inhibition is hence not large enough in the simulations compared to experimental observations [6].

The monophosphate concentrations are plotted in Fig. 10, and unlike for the previously discussed nucleotides, these show a very distinct accumulation. In fact, the curves are quite linear, although the increase is slightly larger for every fourth nucleotide addition. This has to do with the delay effect observed for the diphosphates. If the minimum value of the dNDP concentration is reached three additions later, the production of dNMP is more even, since the hydrolysis is distributed throughout the whole time cycle. Alternatively, if the minimum concentration is reached practically after the next addition, the dNMP production makes a leap every fourth dispensation. Another observation regards the numerical values of the accumulated concentrations. The total amount of accumulated dAMP originates from all dATP added so far, since no accumulation of dATP or dADP was registered. After 60 dispensations, which correspond to 15 A dispensations, the amount of dAMP should approach $15 \times 3.2 \times 10^{-6} = 4.8 \times 10^{-5}$ M. Looking at Fig. 10, this is in fact the case, which is reassuring. The slight irregularities observed in the AMP graph coincide with multiples in the pyrogram (see for instance, the little twitch at the 36th addition corresponding to the triple T:s).

Finally, the effect of maximal dilution, i.e. dilution with no allowed evaporation, can be viewed in Fig. 11,
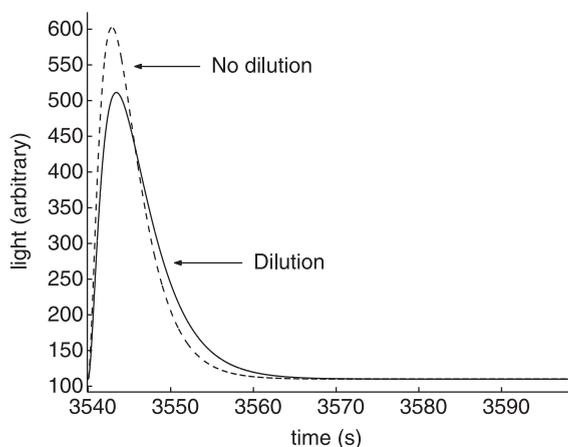
Fig. 11. Last pulse in a 34 nucleotide pyrogram (60 dispensations) calculated with (solid) and without (dashed) the dilution effect. The simulations were performed with an incorporation efficiency of 99% and a plus-shift of 3%.

where the last pulse (double T:s) in the pyrogram in Fig. 7b (dashed) can be compared to the corresponding pulse generated with dilution (solid). The dilution causes a decrease in pulse height together with a pulse broadening. Both effects are primarily explained by a decrease in apyrase activity. The other enzymes are also, to some extent, involved in the process, along with the continued decrease in substrate concentrations. In real Pyrosequencing conditions, evaporation occurs, and the dilution effect is reduced. Pulse broadening due to dilution is in fact regarded as a rather small part of the total broadening observed in experiments. This means that the evaporation is quite significant. From a process optimization point of view, the dilution is an unwanted effect, and thus minimizing the evaporation is unnecessary, but the rate of evaporation should instead be kept at the same level as the rate of dilution, to obtain stable conditions.

5. Conclusions and further research

A model describing single nucleotide incorporation with the Pyrosequencing reaction system was developed. The simulated light pulse was compared to experiments, and the resemblance was excellent. Interesting aspects concerning the influence of changes in the enzyme activities were illustrated. In particular, for a certain range of polymerase concentrations E_{pol}, including that utilized in experiments, an optimal pulse shape was obtained. Future experiments will reveal whether this behavior is a true prediction or not, which will help to validate the model. Furthermore, a model of successive nucleotide additions was implemented and the resulting pyrogram showed good similarity with experiments. The effects of minus- and plus-shifts were well captured by two constant efficiency factors, and the broadening effect could partly be explained by dilution. A weakness of the existing model is that the accumulations of tri- and diphosphates are too small or even non-existing. The possibilities of increasing the nucleotide inhibition of the apyrase in some way are being investigated.

The simulation results of the mathematical model are very promising, and hence it is motivated to further develop the model and make it more detailed and accurate. The approaching task is to find a way to replace the artificial factors of incorporation efficiency and plus-shifts with a non-fixed variable. The incorporation efficiency should, for instance, be dependent on the polymerase velocity (i.e. $k_{cat,pol}$ and E_{pol}), nucleotide concentration, time, and amount of matching DNA. In order to include this in the model, it is likely that a more sophisticated description of the polymerase's kinetics is needed. There is also the notion of the polymerase's processivity. The processivity is the number of nucleotides the enzyme can incorporate before it lets them go off the DNA strand. For the Klenow fragment it is said to be approximately 20 [27], when the enzyme is saturated by nucleotides. The present model utilizes a processivity of 1.

With these features added to the model, information about the benefits and drawbacks of using a different polymerase, with other kinetic properties, can be investigated. This is particularly important in examining the effects of a possible reduction in nucleotide dispensation time. Moreover, once a satisfactory model of the whole system is developed and carefully examined, it is interesting to try and optimize the process with respect to substrate and enzyme concentrations, as well as the kinetic parameters. Given a certain input the model could find the best parameters producing an optimal sequencing result. As DNA sequencing is quite

time-consuming today, there is a lot to gain by optimization.

Appendix A. Equations for the larger model

Variable definitions

$$\alpha_A = \frac{[\text{dATP}]}{K_{m,pol}^{\text{dATP}}} \quad \alpha_C = \frac{[\text{dCTP}]}{K_{m,pol}^{\text{dCTP}}}$$

$$\alpha_G = \frac{[\text{dGTP}]}{K_{m,pol}^{\text{dGTP}}} \quad \alpha_T = \frac{[\text{dTTP}]}{K_{m,pol}^{\text{dTTP}}}$$

$$\beta = \frac{[\text{PP}_i]}{K_{m,Asul}} \quad \gamma = \frac{[\text{ATP}]}{K_{m,luc}}$$

$$\delta_A = \frac{[\text{dADP}]}{K_{m,apyr}^{\text{dADP}}} \quad \delta_{ADP} = \frac{[\text{ADP}]}{K_{m,apyr}^{\text{ADP}}}$$

$$\delta_X = \frac{[\text{dXDP}]}{K_{m,apyr}^{\text{dXDP}}} \quad X = \{C, G, T\}$$

$$\varepsilon_A = \frac{[\text{dAMP}]}{K_{I,apyr}^{\text{dAMP}}} \quad \varepsilon_{AMP} = \frac{[\text{AMP}]}{K_{I,apyr}^{\text{AMP}}}$$

$$\varepsilon_X = \frac{[\text{dXMP}]}{K_{I,apyr}^{\text{dXMP}}} \quad X = \{C, G, T\}$$

$$\varphi_A = \frac{[\text{DNA}_A]}{K_{m,pol}^{\text{DNA}}} \quad \varphi_C = \frac{[\text{DNA}_C]}{K_{m,pol}^{\text{DNA}}}$$

$$\varphi_G = \frac{[\text{DNA}_G]}{K_{m,pol}^{\text{DNA}}} \quad \varphi_T = \frac{[\text{DNA}_T]}{K_{m,pol}^{\text{DNA}}}$$

Differential equations

$$\frac{d\alpha_A}{dt} = \frac{1}{K_{m,pol}^{\text{dATP}}} (-v_{1A} - v_{4A})$$

$$\frac{d\alpha_C}{dt} = \frac{1}{K_{m,pol}^{\text{dCTP}}} (-v_{1C} - v_{4C})$$

$$\frac{d\alpha_G}{dt} = \frac{1}{K_{m,pol}^{\text{dGTP}}} (-v_{1G} - v_{4G})$$

$$\frac{d\alpha_T}{dt} = \frac{1}{K_{m,pol}^{\text{dTTP}}} (-v_{1T} - v_{4T})$$

$$\frac{d\beta}{dt} = \frac{1}{K_{m,Asul}} (v_{1A} + v_{1C} + v_{1G} + v_{1T} - v_2 + v_3)$$

$$\frac{d\gamma}{dt} = \frac{1}{K_{m,luc}} (v_2 - v_3 - v_5)$$

$$\frac{d\delta_A}{dt} = \frac{1}{K_{m,apyr}^{\text{dADP}}} (v_{4A} - v_{6A})$$

$$\frac{d\delta_X}{dt} = \frac{1}{K_{m,apyr}^{\text{dXDP}}} (v_{4C} + v_{4G} + v_{4T} - v_{6X})$$

$$\frac{d\delta_{ADP}}{dt} = \frac{1}{K_{m,apyr}^{\text{ADP}}} (v_5 - v_7)$$

$$\frac{d\varepsilon_A}{dt} = \frac{1}{K_{I,apyr}^{\text{dAMP}}} \cdot v_{6A}$$

$$\frac{d\varepsilon_X}{dt} = \frac{1}{K_{I,apyr}^{\text{dXMP}}} \cdot v_{6X}$$

$$\frac{d\varepsilon_{AMP}}{dt} = \frac{1}{K_{I,apyr}^{\text{AMP}}} (v_7 + v_3)$$

$$\frac{d\varphi_A}{dt} = -\frac{1}{K_{m,pol}^{\text{DNA}}} \cdot v_{1T}$$

$$\frac{d\varphi_C}{dt} = -\frac{1}{K_{m,pol}^{\text{DNA}}} \cdot v_{1G}$$

$$\frac{\mathrm{d}\varphi_{\mathrm{G}}}{\mathrm{d}t} = -\frac{1}{K_{\mathrm{m,pol}}^{\mathrm{DNA}}} \cdot v_{1\mathrm{C}}$$

$$\frac{\mathrm{d}\varphi_{\mathrm{T}}}{\mathrm{d}t} = -\frac{1}{K_{\mathrm{m,pol}}^{\mathrm{DNA}}} \cdot v_{1\mathrm{A}}$$

Algebraic equations

$$0 = \varphi_{\mathrm{Tfree}}\left(K_{\mathrm{m,pol}}^{\mathrm{DNA}} + E_{\mathrm{free}} + E_{\mathrm{free}}\alpha_{\mathrm{Afree}}\right) - K_{\mathrm{m,pol}}^{\mathrm{DNA}}\varphi_{\mathrm{T}}$$

$$0 = \alpha_{\mathrm{Afree}}\left(K_{\mathrm{m,pol}}^{\mathrm{dATP}} + E_{\mathrm{free}} + E_{\mathrm{free}}\varphi_{\mathrm{Tfree}}\right) - K_{\mathrm{m,pol}}^{\mathrm{dATP}}\alpha_{\mathrm{A}}$$

$$0 = \varphi_{\mathrm{Gfree}}\left(K_{\mathrm{m,pol}}^{\mathrm{DNA}} + E_{\mathrm{free}} + E_{\mathrm{free}}\alpha_{\mathrm{Cfree}}\right) - K_{\mathrm{m,pol}}^{\mathrm{DNA}}\varphi_{\mathrm{G}}$$

$$0 = \alpha_{\mathrm{Cfree}}\left(K_{\mathrm{m,pol}}^{\mathrm{dCTP}} + E_{\mathrm{free}} + E_{\mathrm{free}}\varphi_{\mathrm{Gfree}}\right) - K_{\mathrm{m,pol}}^{\mathrm{dCTP}}\alpha_{\mathrm{C}}$$

$$0 = \varphi_{\mathrm{Cfree}}\left(K_{\mathrm{m,pol}}^{\mathrm{DNA}} + E_{\mathrm{free}} + E_{\mathrm{free}}\alpha_{\mathrm{Gfree}}\right) - K_{\mathrm{m,pol}}^{\mathrm{DNA}}\varphi_{\mathrm{C}}$$

$$0 = \alpha_{\mathrm{Gfree}}\left(K_{\mathrm{m,pol}}^{\mathrm{dGTP}} + E_{\mathrm{free}} + E_{\mathrm{free}}\varphi_{\mathrm{Cfree}}\right) - K_{\mathrm{m,pol}}^{\mathrm{dGTP}}\alpha_{\mathrm{G}}$$

$$0 = \varphi_{\mathrm{Afree}}\left(K_{\mathrm{m,pol}}^{\mathrm{DNA}} + E_{\mathrm{free}} + E_{\mathrm{free}}\alpha_{\mathrm{Tfree}}\right) - K_{\mathrm{m,pol}}^{\mathrm{DNA}}\varphi_{\mathrm{A}}$$

$$0 = \alpha_{\mathrm{Tfree}}\left(K_{\mathrm{m,pol}}^{\mathrm{dTTP}} + E_{\mathrm{free}} + E_{\mathrm{free}}\varphi_{\mathrm{Afree}}\right) - K_{\mathrm{m,pol}}^{\mathrm{dTTP}}\alpha_{\mathrm{T}}$$

$$0 = E_{\mathrm{free}}(1 + \varphi_{\mathrm{Tfree}} + \alpha_{\mathrm{Afree}} + \varphi_{\mathrm{Tfree}}\alpha_{\mathrm{Afree}} + \ldots$$

$$+ \varphi_{\mathrm{Gfree}} + \alpha_{\mathrm{C}}\mathrm{free} + \varphi_{\mathrm{Gfree}}\alpha_{\mathrm{Cfree}} + \varphi_{\mathrm{Cfree}} + \ldots$$

$$+ \alpha_{\mathrm{Gfree}} + \varphi_{\mathrm{Cfree}}\alpha_{\mathrm{Gfree}} + \varphi_{\mathrm{Afree}} + \alpha_{\mathrm{Tfree}}$$

$$+ \ldots + \varphi_{\mathrm{Afree}}\alpha_{\mathrm{Tfree}}) - E_{\mathrm{TOT,pol}}$$

Rate expressions

$$v_{1\mathrm{A}} = k_{\mathrm{cat,pol}}^{\mathrm{dATP}} \cdot E_{\mathrm{free}} \cdot \varphi_{\mathrm{Tfree}} \cdot \alpha_{\mathrm{Afree}}$$

$$v_{1\mathrm{C}} = k_{\mathrm{cat,pol}}^{\mathrm{dCTP}} \cdot E_{\mathrm{free}} \cdot \varphi_{\mathrm{Gfree}} \cdot \alpha_{\mathrm{Cfree}}$$

$$v_{1\mathrm{G}} = k_{\mathrm{cat,pol}}^{\mathrm{dGTP}} \cdot E_{\mathrm{free}} \cdot \varphi_{\mathrm{Cfree}} \cdot \alpha_{\mathrm{Gfree}}$$

$$v_{1\mathrm{T}} = k_{\mathrm{cat,pol}}^{\mathrm{dTTP}} \cdot E_{\mathrm{free}} \cdot \varphi_{\mathrm{Afree}} \cdot \alpha_{\mathrm{Tfree}}$$

$$v_2 = \frac{V_{\mathrm{Asul}}\beta}{1 + \beta}$$

$$v_3 = \frac{V_{\mathrm{luc}}\gamma}{1 + \gamma}$$

$$v_{4\mathrm{A}} = \frac{V_{\mathrm{apyr}}^{\mathrm{dATP}} k_{\mathrm{A}}\alpha_{\mathrm{A}}}{1 + \Sigma}$$

$$v_{4\mathrm{C}} = \frac{V_{\mathrm{apyr}}^{\mathrm{dCTP}} k_{\mathrm{C}}\alpha_{\mathrm{C}}}{1 + \Sigma}$$

$$v_{4\mathrm{G}} = \frac{V_{\mathrm{apyr}}^{\mathrm{dGTP}} k_{\mathrm{G}}\alpha_{\mathrm{G}}}{1 + \Sigma}$$

$$v_{4\mathrm{T}} = \frac{V_{\mathrm{apyr}}^{\mathrm{dTTP}} k_{\mathrm{T}}\alpha_{\mathrm{T}}}{1 + \Sigma}$$

$$v_5 = \frac{V_{\mathrm{apyr}}^{\mathrm{ATP}} k_{\mathrm{ATP}}\gamma}{1 + \Sigma}$$

$$v_{6\mathrm{A}} = \frac{V_{\mathrm{apyr}}^{\mathrm{dADP}} \delta_{\mathrm{A}}}{1 + \Sigma}$$

$$v_{6X} = \frac{V_{\mathrm{apyr}}^{\mathrm{dXDP}} \alpha_{X}}{1 + S}$$

$$v_7 = \frac{V_{\mathrm{apyr}}^{\mathrm{ADP}} \delta_{\mathrm{ADP}}}{1 + \Sigma}$$

where

$$\Sigma = k_{\mathrm{A}}\alpha_{\mathrm{A}} + k_{\mathrm{C}}\alpha_{\mathrm{C}} + k_{\mathrm{G}}\alpha_{\mathrm{G}} + k_{\mathrm{T}}\alpha_{\mathrm{T}} + \delta_{\mathrm{A}} + \delta_{X} + \delta_{\mathrm{ADP}}$$

$$+ k_{\mathrm{ATP}}\gamma + \varepsilon_{\mathrm{A}} + \varepsilon_{X} + \varepsilon_{\mathrm{AMP}}$$

and

$$k_{\mathrm{A}} = K_{\mathrm{m,pol}}^{\mathrm{dATP}} \Big/ K_{\mathrm{m,apyr}}^{\mathrm{dATP}}$$

$$k_{\mathrm{C}} = K_{\mathrm{m,pol}}^{\mathrm{dCTP}} \Big/ K_{\mathrm{m,apyr}}^{\mathrm{dCTP}}$$

$$k_{\mathrm{G}} = K_{\mathrm{m,pol}}^{\mathrm{dGTP}} \Big/ K_{\mathrm{m,apyr}}^{\mathrm{dGTP}}$$

$$k_{\mathrm{T}} = K_{\mathrm{m,pol}}^{\mathrm{dTTP}} \Big/ K_{\mathrm{m,apyr}}^{\mathrm{dTTP}}$$

$$k_{\mathrm{ATP}} = K_{\mathrm{m,luc}} \Big/ K_{\mathrm{m,apyr}}^{\mathrm{ATP}}$$

References

[1] J.D. Watson, F.H. Crick, Molecular structure of nucleic acids, Nature, (1953) 4356.

[2] F. Sanger, S. Nicklen, A.R. Coulson, DNA sequencing with chain-terminating inhibitors, Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.

[3] A. Maxam, W. Gilbert, A new method of sequencing DNA, Proc. Natl. Acad. Sci. USA 74 (1977) 560–564.

[4] M. Ronaghi, S. Karamohamed, B. Pettersson, M. Uhlén, P. Nyrén, Real-time DNA sequencing using detection of pyrophosphate release, Anal. Biochem. 242 (1996) 84–89.

[5] M. Ronaghi, M. Uhlén, P. Nyrén, Real-time pyrophosphate detection for DNA sequencing, Science 281 (1998) 363–365.

[6] B. Gharizadeh, T. Nordström, A. Ahmadian, M. Ronaghi, P. Nyrén, Long-read pyrosequencing using pure 2′-deoxyadenosine-5′-O-(1-thiotriphosphate) Sp-isomer, Anal. Biochem. 301 (2002) 82–90.

[7] M.E. Dahlberg, S.J. Benkovic, Kinetic mechanism of DNA polymerase I (Klenow fragment): identification of a second conformational change and evaluation of the internal equilibrium constant, Biochemistry 30 (1991) 4835–4843.

[8] P.W. Robbins, F. Lipmann, Enzymatic synthesis of adenosine-5′-phosphosulfate, J. Biol. Chem. 233 (1958) 686–690.

[9] L.Yu. Brovko, O.A. Gandelman, T.E. Polenova, N.N. Ugarova, Kinetics of bioluminescence in the firefly luciferin–luciferase system, Biochemistry (Moscow) 59 (1994) 195–201.

[10] B.R. Branchini, R.A. Magyar, M.H. Murtiashaw, N.C. Portier, The role of active site residue arginine 218 in firefly luciferase bioluminescence, Biochemistry 40 (2001) 2410–2418.

[11] M. Handa, G. Guidotti, Purification and cloning of a soluble ATP-diphosphohydrolase (apyrase) from potato tubers (*Solaanum tuberosum*), Biochem. Biophys. Res. Commun. 218 (1996) 916–923.

[12] A. Cornish-Bowden, Fundamentals of Enzyme Kinetics, Portland Press, London, 1995.

[13] M. Komoszynski, A. Wojtczak, Apyrases (ATP diphosphohydrolases, EC 3.6.1.5): function and relationship to ATPases, Biochim. Biophys. Acta 1310 (1996) 233–241.

[14] M. Ronaghi, Improved performance of Pyrosequencing using single-stranded DNA-binding protein, Anal. Biochem. 286 (2000) 282–288.

[15] M. Ronaghi, Pyrosequencing shreds light on DNA sequencing, Genome Res. 11 (2001) 3–11.

[16] B. Gharizadeh, M. Kalantari, C.A. Garcia, B. Johansson, P. Nyrén, Typing of human papillomavirus by Pyrosequencing, Lab. Invest. 81 (2001) 673–679.

[17] N.A. Van Draanen, S.C. Tucker, F.L. Boyd, B.W. Trotter, J.E. Reardon, β-L-thymidine 5′-triphosphate analogs as DNA polymerase substrates, J. Biol. Chem. 267 (1992) 25 019–25 024.

[18] B.A. Foster, S.M. Thomas, J.A. Mahr, F. Renosto, H.C. Patel, I.H. Segel, Cloning and sequencing of ATP sulfurylase from *Penicillium chrysogenum*, J. Biol. Chem. 269 (1994) 19 777–19 786.

[19] S.R. Ford, K.H. Bunton, G.J. Hampton, J. McCarthy, M.S. Hall, S.J. Pangburn, et al., Use of firefly luciferase for ATP measurement: other nucleotides enhance turnover, J. Biolumin. Chemilumin. 11 (1996) 149–167.

[20] G. Del Campo, J. Puente, M.A. Valenzuela, A. Traverso-Cori, O. Cori, Hydrolysis of synthetic pyrophosphoric esters by an isoenzyme of apyrase from *Solanum tubersosum*, Biochem. J. 167 (1977) 525–529.

[21] A.M. Kettlun, L. Ulribe, V. Calvo, S. Silva, J. Rivera, M. Mancilla, et al., Properties of two apyrases from *Solanum tuberosum*, Phytochemistry 21 (1982) 551–558.

[22] W.R. McClure, T.M. Jovin, The steady state kinetic parameters and non-processivity of *Escherichia coli* deoxyribonucleic acid polymerase I, J. Biol. Chem. 250 (1975) 4073–4080.

[23] R.D. Kuchta, V. Mizrahi, P.A. Benkovic, K.A. Johnson, S.J. Benkovic, Kinetic mechanism of DNA polymerase I (Klenow), Biochemistry 26 (1987) 8410–8417.

[24] N. Lembert, L.-Å. Idahl, Regulatory effects of ATP and luciferin on firefly luciferase activity, Biochem. J. 305 (1995) 929–933.

[25] A. Traverso-Cori, H. Chaimovich, O. Cori, Kinetic studies and properties of potato apyrase, Arch. Biochem. Biophys. 109 (1965) 173–184.

[26] The MathWorks Inc., MATLAB 6.0.0.88 R12, http://www.mathworks.com.

[27] R.A. Bambara, D. Uyemura, T. Choi, On the processive mechanism of *Escherichia coli* DNA polymerase I, J. Biol. Chem. 253 (1978) 413–423.