Research Article
A Mathematical Model for DNA Damage and Repair

Philip S. Crooke and Fritz F. Parl

Department of Mathematics, Vanderbilt University, Nashville, TN 37240, USA
Department of Pathology, Vanderbilt University, Nashville, TN 37232, USA

Correspondence should be addressed to Fritz F. Parl, fritz.parl@vanderbilt.edu

Received 13 April 2010; Accepted 27 May 2010

Academic Editor: Ashis Basu

Copyright © 2010 P. S. Crooke and F. F. Parl. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In cells, DNA repair has to keep up with DNA damage to maintain the integrity of the genome and prevent mutagenesis and carcinogenesis. While the importance of both DNA damage and repair is clear, the impact of imbalances between both processes has not been studied. In this paper, we created a combined mathematical model for the formation of DNA adducts from oxidative estrogen metabolism followed by base excision repair (BER) of these adducts. The model encompasses a set of differential equations representing the sequence of enzymatic reactions in both damage and repair pathways. By combining both pathways, we can simulate the overall process by starting from a given time-dependent concentration of 17β-estradiol (E2) and 2′-deoxyguanosine, determine the extent of adduct formation and the correction by BER required to preserve the integrity of DNA. The model allows us to examine the effect of phenotypic and genotypic factors such as different concentrations of estrogen and variant enzyme haplotypes on the formation and repair of DNA adducts.

1. Introduction

Estrogens are known carcinogens recognized as prime risk factor for the development of breast cancer [1]. The principal enzyme controlling estrogen metabolism in breast tissue is cytochrome P450 1B1 (CYP1B1), which sequentially oxidizes E2 to 4-OHE2 and the estrogen quinone (E2-3,4-Q, Figure 1). The highly reactive quinone nonenzymatically attacks DNA and forms covalent adducts with bases, such as 4-OHE2-N7-Gua (GUA adduct), which undergoes depurination by hydrolysis of the N-glycosylic bond between base and sugar, leaving an apurinic (AP) site in the double-stranded DNA. AP sites are repaired by the base excision repair (BER) pathway in a sequence of enzymatic reactions. AP endonuclease APE1 cleaves 5’ to the AP site to generate a nick with 3’-OH and 5’-deoxyribose phosphate termini. DNA polymerase (Pol) catalyzes the release of the deoxyribose phosphate residue and DNA synthesis to fill the gap, which is then sealed by DNA ligase (Figure 1). This reaction sequence is carried out in one of two pathways, “short-patch” and “long-patch”, which differ in the number of inserted nucleotides and the protein components.

In this paper, we present a mathematical model for the formation of DNA adducts from oxidative estrogen metabolism followed by BER of these adducts. The modeling is initially done in two stages: (1) formation of estrogen-deoxyribonucleoside adducts, and (2) repair of adducts in the BER. By combining both stages, we can simulate the overall process by starting from a given time-dependent concentration of E2 and 2′-deoxyguanosine, determine the extent of adduct formation and the correction by BER required to preserve the integrity of DNA.

2. Mathematical Model

The mathematical model is constructed in stages. Stage 1 deals with the formation of DNA adducts from estrogen metabolism and Stage 2 encompasses the repair of adducts. In the differential equations that will constitute the model, E2 denotes 17β-estradiol, 4-OHE2 the catechol estrogen, E2-3,4-Q the estrogen quinone, GUA the deoxyribonucleoside 2′-deoxyguanosine, and GUA adduct the estrogen adduct 4-OHE2-N7-Gua over time t. We assume that the starting
Figure 1: Overview of estrogen metabolism-induced DNA adduct formation and subsequent DNA repair by BER. The top sequence of reactions shows the CYP1B1-mediated oxidative estrogen metabolism resulting in estrogen-deoxyribonucleoside adduct formation. The bottom sequence of reactions shows the BER pathway with its two branches, the predominant short-patch pathway (left) and the alternate long-patch pathway (right). Not shown is the formation of the catechol estrogen 2-OHE2, which yields stable adducts, and the redox cycling of catechol estrogens and estrogen quinones, which generates reactive oxygen species capable of producing oxidative DNA adducts.
Figure 2: Comparison of stage 1 mathematical model (DNA damage) with experimental data. The metabolism of E₂ (a), 4-OHE₂ (b), and GUA_{adduct} (c) is shown as a function of time. In each graph the dots represent the experimental data [2], and the curves are derived from the mathematical model.

Points for the dynamic model are the initial concentration of GUA(0) = GUA₀ and a prescribed time series is given for E₂.

The first stage of the model is based on the first three steps of the reaction sequence shown in Figure 1. As chemical reactions, these steps are

\[
\begin{align*}
E₂ + E_{CYP1B1} & \rightleftharpoons 4-OHE₂ + E_{CYP1B1}, \\
4-OHE₂ & \rightarrow E₂-3,4-Q, \\
GUA + E₂-3,4-Q & \rightarrow GUA_{adduct}.
\end{align*}
\]

In the first reaction, the production of 4-OHE₂ from E₂ is an enzymatic reaction mediated by CYP1B1 that follows Michaelis-Menten kinetics with parameters \(K_m\) and \(k_{cat}\) [3]. In the second reaction, E₂-3,4-Q is produced from 4-OHE₂ according to Hill-type kinetics. It has been shown that E₂-3,4-Q is highly reactive and will spontaneously form adducts with DNA without enzyme involvement [2]. Thus, in the third reaction, the formation of adducts is proportional to the levels of the quinone and the deoxyribonucleoside with a rate constant denoted by \(k_Q\). Estrogen-DNA adducts have been detected in normal and malignant human breast tissues [4, 5], and we have recently provided direct experimental evidence that oxidative metabolism of the parent hormone E₂ leads to the formation of 4-OHE₂ and deoxyribonucleoside adducts, such as 4-OHE₂-N7-Gua [2]. Using a compartmental analysis we created a mathematical model for reaction sequence (1). Figure 2 shows superimposed the experimental data (dots) and the model simulations (curves) for E₂, 4-OHE₂, and GUA_{adduct}.

The second stage of the model is based on the BER pathway (Figure 1) and uses a compartmental analysis with Michaelis-Menten kinetics to simulate the enzymatic reactions in the pathway [6]. The enzymes include the endonucleases APE1 and FEN1, DNA polymerases β and δ (Polβ, Polδ), and ligases 1 and 3 (Lig1, Lig3). The dynamic process is modeled by a system of nonlinear differential equations. \(S(t)\) denotes the concentration of DNA adducts. We define the source and removal function for the repair cycle functions, \(S(t)\) and \(R(t)\), as \(S(t) = k_s GUA_{adduct}\) and \(R(t) = k_r (E₂-3,4-Q) y_9\). Here, \(k_s\) is a modeling parameter expressed in [1/min] which reflects the rate at which GUA_{adduct} enters the repair cycle. The parameter \(k_r\) is the rate at which \(y_9(t)\) interacts with E₂-3, 4-Q to form GUA_{adduct}. The two parameters, \(k_s\) and \(k_r\), can be used to control the rates of entering and exiting of adducts and
Table 1: Parameters for System of Differential Equations [6].

| Reaction subscript (i) | $k_{cat}$ [1/min] | $K_m$ [mmol/L] |
|------------------------|-------------------|----------------|
| 5                      | 192.168           | 34.7           |
| 6                      | 4.500             | 500            |
| 7                      | 49.020            | 210            |
| 8,11                   | 1.278             | 56.7           |
| 9                      | 36.000            | 100            |
| 10                     | 8.130             | 39             |

There are several Michaelis-Menten parameters, $k_{cat}$ and $K_m$, for which experimental values have been estimated or measured (Table 1) [6]. The enzyme levels are denoted by $E_{APE1}$, $E_{FEN1}$, $E_{Polβ}$, $E_{Polδ}$, $E_{Lig1}$, and $E_{Lig3}$. The variable, $y_9(t)$, represents the repaired GUA$_{adduct}$, and if we set its initial condition in the system of differential equations as $y_9(0) =$ GUA$_0$, it has the same meaning as GUA$(t)$. AP endonucleases, such as AP endonuclease 1 (APE1), are quite abundant in most cells, ranging from 200,000 to 7,000,000 APE1 molecules per cell in fibroblasts and HeLa cells, respectively [7, 8]. Thus, the minimal concentration of APE1 in nuclear extract can be assumed to be 2000 nM [6]. The equivalent estimated concentrations of Polβ, Polδ, FEN1, Lig1, and Lig3 are 419, 600, 450, 254, and 254 nM, respectively [6, 9, 10].

It is estimated that AP sites are also generated spontaneously at an estimated rate of 2,000 to 10,000 per cell per day [11, 12]. We used (2) to determine the time necessary to repair these AP sites. In Figure 3 we demonstrate that the repair of lesions is a nonlinear process. For small values of the ratio (AP lesions/10$^5$ Nucleotides), the slope of the repair time function is relatively large as compared to larger values of the ratio. Figure 3 captures the dynamics of only the repair part of the model, that is, it shows the repair times for a fixed amount of AP lesions.

In the experiments addressed mathematically in Stage 1 of the model, we demonstrated that the CYP1B1-mediated oxidation of E2 resulted in the formation of approximately 10 nmol/L 4-OHE$_2$-N7-Gua, which translates into 10 adducts per 10$^5$ nucleotides per hour [2]. Since the formation of 4-OHE$_2$-N7-Gua leads to spontaneous depurination [13], we can assume an equal number of AP sites formed as a result of oxidative estrogen metabolism. This is one order of magnitude higher than the number of AP sites formed in normal leukocytes and fibroblasts, 10 AP sites per 10$^6$ nucleotides per hour [13]. The living cell depends on efficient enzymatic repair of the continuous chemical damage inflicted on genomic DNA. Therefore, in the last stage of the model we combine the adduct formation model (Stage 1) and the adduct repair model (Stage 2). The mathematical model used in Stage 1 was chosen to fit batch data (a fixed amount of E2). However, we consider the case when E2 is continuously available and can vary over time and hence, there is continuous adduct formation and repair. In our modification of the Stage 1 model, we assume that the total amount of 2’-deoxyguanosine is fixed, and it exists in two states: damaged (GUA$_{adduct}$) and undamaged or repaired (y9). The modified Stage 1 model, which accounts for the continuous nature of damage and repair, is then given by the following system of ordinary differential equations:

$$
\frac{d(4\text{-OHE}_2)}{dt} = \frac{k_{cat, E_{CYP1B1}} E_2}{K_m + E_2} \frac{V_{max, 4\text{-OHE}_2}}{K_m + 4\text{-OHE}_2},
$$

$$
\frac{d(E_2-3, 4\text{-Q})}{dt} = \frac{V_{max, 4\text{-OHE}_2}}{K_m + 4\text{-OHE}_2} - k_Q E_2-3, 4\text{-Q},
$$

$$
\frac{d(\text{GUA}_{adduct})}{dt} = k_r(E_2-3, 4\text{-Q}) y_7 - k_s \text{GUA}_{adduct}.
$$

FIGURE 3: Stage 2 mathematical model of BER. The plot shows the time required to repair a fixed number of AP sites per 10$^5$ nucleotides with wild-type BER enzymes. The horizontal axis represents the starting number of AP sites per 10$^5$ nucleotides and the vertical axis the time for this number of sites to be repaired.
Figure 4: Effects of different concentrations of estrogen and variant haplotypes of CYP1B1 and APE1 on the formation and repair of DNA adducts. (a) Shows the effect of doubling the concentration of E₂ on the number of AP lesions per 10⁵ nucleotides over time using E₂ = 1 (red curve) and E₂ = 2 (blue curve). (b) Shows the effect of variant haplotypes of CYP1B1 and APE1 leading to an increase in the number of AP lesions per 10⁵ nucleotides relative to wild-type enzymes. Each curve uses different assumptions about the kinetic constant, \( k_{\text{cat}} \), for the CYP1B1 and APE1 enzymes. The red curve depicts the wild-types with \( k_{\text{cat}} = 1.6 \) (CYP1B1) and \( k_{\text{cat}} = 192.168 \) (APE1). The blue curve uses a 30% increase in \( k_{\text{cat}} \), keeping \( k_{\text{cat}} = 192.168 \). The green curve employs a 30% reduction in \( k_{\text{cat}} \), keeping \( k_{\text{cat}} = 192.168 \). Finally, the purple curve uses both a 30% increase in \( k_{\text{cat}} \) and a 30% reduction in \( k_{\text{cat}} \). (c) Shows the effect of different variant haplotypes of CYP1B1 and APE1 that cause a decrease in the number of AP lesions per 10⁵ nucleotides relative to wild-type enzymes. The wild-type is shown in red, a 30% decrease of CYP1B1 activity with APE1 activity held at its wild-type level in blue, 30% increase of APE1 activity with CYP1B1 activity held at its wild-type level in green, and a 30% decrease in CYP1B1 activity with a 30% increase of APE1 activity in purple.

The endpoint of the repair process is \( y_9(t) \) which is the concentration of GUA(t). The term, \( R(t) = k_r y_9(E_2 - 3, 4 - Q) \) where \( k_r \) is a constant, is the term that connects (2) and (3). In case the AP sites are not generated spontaneously but produced by DNA glycosylases, we can easily incorporate the enzymatic reaction catalyzed by glycosylases, which is not rate limiting for BER [9].

The combined Stage 3 model is composed of the differential in (2) and (3) along with their initial conditions. One can easily show that the system has a conservation law

\[
\frac{\text{GUA}_{\text{adduct}}(t)}{\text{GUA}_0} = 1 - \frac{1}{\text{GUA}_0} \sum_{i=0}^{11} y_i(t).
\]  

(4)

Since this equation holds for all times, it calculates, at least in theory, the steady-state value of the unrepaired ratio.

3. Discussion

Numerous epidemiological studies have implicated estrogens in the development of breast cancer [1]. A pooled analysis of nine prospective studies of serum estrogen levels and breast cancer in over 2,400 postmenopausal women revealed a strong association of serum E₂ concentrations with breast cancer risk [14]. In Figure 4(a), we show the results of the combined model for a constant source of estrogen, \( E_2(t) \equiv 1 \), on the number of AP sites per 10⁵ nucleotides as well as the effect of doubling the E₂ level, \( E_2(t) \equiv 2 \). Clinical studies estimate that the doubling of serum E₂ levels confers a 1.3-fold increase in the risk of breast cancer [14]. Our model predicts a larger increase in the number of AP sites, indicating the involvement of other factors, such as enzyme activities. Enzymes involved in the production of
of different variants of CYP1B1 and APE1, which cause a decrease in number of AP sites relative to the wild-type enzymes. The panels in Figure 4 demonstrate the number of DNA adducts, which are constantly being created and repaired as modeled in (2) and (3), approaching steady values in the integrated process of adduct formation and repair.

A limitation of the model is its focus on BER and the omission of other DNA repair pathways. While the majority of adducts derived from 4-OHE2 and E2-3, 4-Q are depurinating N7 guanine adducts susceptible to BER, other stable, bulky DNA adducts that do not depurate are also formed. However, the level of the stable adducts is three to four orders of magnitude lower than the level of depurinating adducts justifying the emphasis on BER [13, 18]. A vast number of depurination events occur under normal conditions involving not just guanine but other bases as well [19]. The importance of depurinating adducts derived from oxidative estrogen metabolism relative to the total spectrum of depurination events is presently unknown.

A critical strength of the model is that it can readily incorporate additional enzymes involved in the metabolism, as we have shown for CYP1A1, catechol-O-methyl transferase, and glutathione S-transferase P1 in the estrogen metabolism pathway [20]. Moreover, the model incorporates both phenotypic and genotypic measures of estrogen exposure and genotypic data. This is schematically shown in Figure 5, which displays the interaction of estrogens, enzyme genotypes, and resulting AP sites per 10^5 nucleotides as a three-dimensional graph. In designing Figure 5, we assumed that the difference in estrogen exposure between individual women is no more than twofold, with the tertiles 1.0, 1.5, and 2.0 used for the x-axis. This range is conservative since up to fivefold differences in breast tissue concentrations have been reported in [21]. Regardless of the scale used for the x-axis, the production of AP sites would be expected to be greater in women with more endogenous (more ovulatory cycles) or exogenous (hormone replacement therapy, oral contraceptives) estrogen exposure. The effect of wild-type and variant enzyme haplotypes on AP sites is displayed on the y-axis. It is evident from Figure 5 that the combined phenotypic and genotypic factors exert not just an additive but multiplicative effect on AP site levels, which are shown on the z-axis.

In conclusion, we modeled the dynamic interaction between the estrogen-mediated DNA damage pathway and the DNA BER pathway in order to assess the overall impact on carcinogenesis. The model encompasses a set of differential equations representing the sequence of enzymatic reactions in both damage and repair pathways. The model allows us to examine the effect of phenotypic and genotypic factors such as different concentrations of estrogen and variant enzyme haplotypes on the formation and repair of DNA adducts. As a practical application, the model can be used to explore genetic factors in the damage-repair cycle (through the rate constants and estrogen levels) as a first step in understanding why some women develop breast cancer while others with similar circumstances do not. As better information about the kinetics in the model (e.g., rate constants for different haplotypes) becomes available,
predictions about breast cancer risk will improve. Thus, the model may be useful for the construction of breast cancer risk models.

**Acknowledgments**

This work was supported by NIH (U54CA113007) and the Vanderbilt Integrative Cancer Biology Center.

**References**

[1] B. E. Henderson and H. S. Feigelson, “Hormonal carcinogenesis,” *Carcinogenesis*, vol. 21, no. 3, pp. 427–433, 2000.

[2] A. R. Belous, D. L. Hachey, S. Dawling, N. Roodi, and F. F. Parl, “Cytochrome P450 1B1-mediated estrogen metabolism results in estrogen-deoxyribonucleoside adduct formation,” *Cancer Research*, vol. 67, no. 2, pp. 812–817, 2007.

[3] I. H. Hanna, S. Dawling, N. Roodi, F. P. Guengerich, and F. F. Parl, “Cytochrome P450 1B1 (CYP1B1) pharmacogenetics association of polymorphisms with functional differences in estrogen hydroxylation activity,” *Cancer Research*, vol. 60, no. 13, pp. 3440–3444, 2000.

[4] J. Embrechts, F. Lemi`ere, W. Van Dongen et al., “Detection of estrogen DNA-adducts in human breast tumor tissue and healthy tissue by combined nano LC-nano ES tandem mass spectrometry,” *Journal of the American Society for Mass Spectrometry*, vol. 14, no. 5, pp. 482–491, 2003.

[5] Y. Markushin, W. Zhong, E. L. Cavaliere et al., “Spectral characterization of catechol estrogen quinone (CEQ)-derived DNA adducts and their identification in human breast tissue extract,” *Chemical Research in Toxicology*, vol. 16, no. 9, pp. 1107–1117, 2003.

[6] B. A. Sokhansanj and D. M. Wilson III, “Estimating the effect of human base excision repair protein variants on the repair of oxidative DNA base damage,” *Cancer Epidemiology Biomarkers and Prevention*, vol. 15, no. 5, pp. 1000–1008, 2006.

[7] E. Cappelli, T. Hazra, J. W. Hill, G. Slupphaug, M. Bogliolo, and G. Frosina, “Rates of base excision repair are not solely dependent on levels of initiating enzymes,” *Carcinogenesis*, vol. 22, no. 3, pp. 387–393, 2001.

[8] D. S. Chen, T. Herman, and B. Demple, “Two distinct human DNA diesterases that hydrolyze 3′-blocking deoxyribose fragments from oxidized DNA,” *Nucleic Acids Research*, vol. 19, no. 21, pp. 5907–5914, 1991.

[9] D. K. Srivastava, I. Husain, C. L. Arteaga, and S. H. Wilson, “DNA polymerase β expression differences in selected human tumors and cell lines,” *Carcinogenesis*, vol. 20, no. 6, pp. 1049–1054, 1999.

[10] A. E. Tomkinson, D. D. Lasko, G. Daly, and T. Lindahl, “Mammalian DNA ligases. Catalytic domain and size of DNA ligase I,” *Journal of Biological Chemistry*, vol. 265, no. 21, pp. 12611–12617, 1990.

[11] H. Atamna, I. Cheung, and B. N. Ames, “A method for detecting abasic sites in living cells: age-dependent changes in base excision repair,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 2, pp. 686–691, 2000.

[12] T. Lindahl and B. Nyberg, “Rate of depurination of native deoxyribonucleic acid,” *Biochemistry*, vol. 11, no. 19, pp. 3610–3618, 1972.