Mathematical Models of Misrejoining DNA Double-Strand Breaks and DNA Fragment Size Distributions

By

Tomas Radivoyevitch

A dissertation submitted to the faculty of the Medical University of South Carolina in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Graduate Studies.

Department of Biostatistics and Epidemiology

1996

Approved by:

_____

_____

_____

_____

_____

_____

Abstract

Recent high dose pulsed field gel electrophoresis (PFGE) data (Löbrich et al. (1995) *Proc. Natl. Acad. Sci.* **92**, pp 12050-12054) of X-irradiated human fibroblasts is used to select the the Sax subset (SS) model as an adequate mathematical representation of misrejoining DNA double-strand breaks (DSBs). The SS model supports the notion that DSB misrepair results from the random recombination of completely disconnected (active) DSB free ends. A low dose extension of the SS model, the Sax-Markov binary eurepair (SMBE) model, predicts a linear-quadratic-linear misrepair dose response that is consistent with chromosomal aberration data.

Two existing induced DNA fragment distribution models, the ethidium bromide (EtBr) model and the specific probe model, are derived in detail using direct and indirect methods. These models describe a cell's DNA fragment distribution immediately after an acute exposure to low LET irradiation, but not after a significant amount of misrepair has occured. To describe DNA fragments in misrepaired cells, this research develops the Sax subset EtBr (SSEB) model, an extension of the EtBr model to include the DSB misrepair dynamics of the SS model. The SSEB model is thoroughly analyzed for exponentially distributed (high dose) initial conditions.

Acknowledgements

I thank my advisor, Prof. David G. Hoel, for guiding me into radiation research and letting me choose my own paths, Prof. Rainer K. Sachs for being generous with his time and for advising me on many mathematical issues, Dr. Björn Cedervall for introducing me to pulsed field gel electrophoresis and fragment distribution models, Dr. Philip J. Hahnfeldt for helping with data interpretation, Dr. Björn Rydberg for many interesting discussions concerning DSB misrejoining mechanisms, the Medical University of South Carolina for providing resources, and the Department of Energy for financial support under grant DE-FG01-92EW50625.

I also thank my professors at Case Western Reserve University for challenging me to my utmost capacity, my father for showing me the wonderful world of microbes in a flower vase, my mother and my grandmother for teaching me kindness, my brothers and friends for instilling in me a no-quit attitude, my in-laws for their support as a family, and most of all, I thank Kirsten and Carlton for their love and patience.

Contents

List of Figures

List of Tables

List of Abbreviations

1. LET = linear energy transfer
2. DSB = double-strand break
3. SSB = single-strand break
4. PFGE = pulsed field gel electrophoresis
5. EtBr = ethidium bromide
6. (M)bp = (million) basepairs
7. FAR = fraction activity released
8. RBM = Revell binary misrepair
9. SA = Sax activation
10. SRA = Sax-Revell activation
11. SS = Sax subset
12. RSSB = Revell single-strand break
13. LQL = linear-quadratic-linear
14. SMBE = Sax-Markov binary eurepair
15. SSEB = Sax subset ethidium bromide

List of Symbols

1. $U(t)$ = the expected number of unrejoined DSBs at time t.
2. $M(t)$ = the expected number of misrejoined DSBs at time t.
3. $U_1(t)$ = the expected number of protein splinted unrejoined DSBs at time t.
4. $U_2(t)$ = the expected number of disconnected DSB free ends at time t.
5. $U_3(t)$ = the number of disconnected DSB free ends destined for misrepair.
6. $U_s(t)$ = the expected pairs of staggered SSBs at time t.
7. λ_r= the linear eurepair rate constant for splinted DSBs.
8. λ_s= the linear eurepair rate constant for SSB pairs.
9. κ or (κ_1) = the binary misrepair rate constant for splinted DSBs.
10. κ or (κ_2) = the binary misrepair rate constant for DSB free ends.
11. G = the yield of DSBs per Gray.
12. D = the dose in Gray.
13. p = the fraction of DSBs that are active.
14. σ = the fraction of PFGE DSBs that are SSB pairs in the cell.
15. f = the fraction of DSBs channeled into misrepairs at steady state.
16. σ^2 = the variance of an arbitrary reactant distribution.
17. i = the number of pairs of DSB free ends destined for misrepair in one cell.
18. j = the number of pairs of DSB free ends with a chance for eurepair in one cell.
19. m = the number of misrepairs in one cell.
20. μ = the expected number of active (disconnected) DSBs.
21. η = the number of isolated nuclear subcompartments.
22. μ_1 = the expected number of active (disconnected) DSBs due to one particle track traversal through one nuclear subcompartment.
23. D_1 = the dose at which the whole nucleus expects one particle track traversal.
24. γ = the yield of active DSBs per Gray.
25. α, β = the coefficients of the linear quadratic dicentric yield $\alpha D + \beta D^2$.
26. $\Delta(t)$ = an arbitrary measure of genetic deviation from normality.
27. $R(\Delta)$ = cancer risk as a function of genetic deviation.
28. d = the dispersion of a distribution (the variance divided by the mean).

29. x= the size in basepairs of an arbitrary DNA fragment.

30. λ= the expected number of DSBs per basepair.

31. $u_0^S(x)$= the unit step function, one if $0 \leq x \leq S$, zero otherwise.

32. $\delta(S-x)$ = the Dirac delta function centered at $x = S$.

33. $n_{at}(x,t)$ = the edge fragment density.

34. $n_{aa}(x,t)$ = the internal fragment density.

35. $n_{\delta}(x,t)$ = the density of unhit chromosomes.

36. $n_r(x,t)$ = the ring fragment density.

37. $n_{tt}(x,t)$ = the bitelomeric fragment density.

38. $n(x)$ = the total fragment *number* density.

39. $m(x)$ = the total fragment *mass* density.

40. h = the hybridization point on a chromosome.

41. S_j = the size of the jth chromosome.

42. S_T = the sum of all chromosome sizes.

43. N = the number of chromosomes in G_0 cells.

44. $N_c(t)$ = the total number of fragments of type $c \in \{at, aa, tt, \delta, r\}$.

45. $\lambda_i(t)$ = the expected number of inactive (splinted) DSBs per basepair.

46. $\lambda_a(t)$ = the expected number of active (disconnected) DSBs per basepair.

47. $\tilde{n}_c(k,t)$ = the Laplace transform of $n_c(x,t)$ with $c \in \{at, aa, tt, \delta, r\}$.

48. $p_c(x,t)$ = the probability a fragment of type $c \in \{at, aa, tt, \delta, r\}$ has size x.

Chapter 1

Introduction

Over the years, the United States has accumulated a large number of hazardous waste sites. Since the overwhelming total cost to remediate these sites is expected to exceed the availability of funds, a strategy is needed to optimize the allocation of our finite resources. One such strategy is to clean the sites that contain the greatest risks first. Another strategy is to clean each site to the point where the additional lives saved per dollar spent equals that for each of the other sites. In either case a quantitative risk assessment is required for each hazardous waste site. Quantitative risk assessments entail two basic steps: 1) an assessment of the exposure levels to the population near the waste site, and 2) an assessment of the biological consequences of the exposures. This research relates to the second of these steps.

An exposure can have several health consequences, and a biological endpoint can be the result of various environmental agents. For a particular exposure level, the risk of an adverse health effect is estimated from a dose-response model. The parameters of a dose-response model are typically estimated from high dose rodent data or from human epidemiological data. The dose-response model is then used to make predictions at low doses. A problem with this approach lies in the fact that many dose-response models can provide equally suitable fits to rodent (or human) data at high doses, but produce drastically different risk predictions at very low doses [11]. To mitigate this problem, the choice of dose-response models must somehow be constrained. For example, chromosomal aberration dose-response data might be used to derive a model structure that could then be enforced upon a tumor dose-response curve. Alternatively, one could develop mathematical models for each cellular process relevant to cancer and merge them to form a risk model (figure 1.1). The latter approach has the potential of transferring large amounts of cancer-relevant information to the dose-response curve.

At least seven cellular processes relevant to cancer can be identified: DNA repair systems [39, 31], signal transduction pathways [43], gene regulatory networks [28], apoptotic pathways [14], mechanisms of cellular differentiation [26, 29], the regulation of cellular migration [42] and cell cycle controls [27], see figure 1.1. Signal transduction pathways transmit information from the cell's environment to the cell's nucleus. Within the nucleus, transcription factors modulate the expression of other transcription factors and thus form a biochemical network which, like a brain, allows a cell to compute its best response given the current conditions. The cell's response options include suicide, a new career, relocation and reproduction. To keep things manageable, this research will focus on the repair of DNA double-strand breaks (DSBs), a process believed to be important in radiation induced carcinogenesis [46]. This work therefore represents one small component of a much larger project to develop whole cell cancer models.

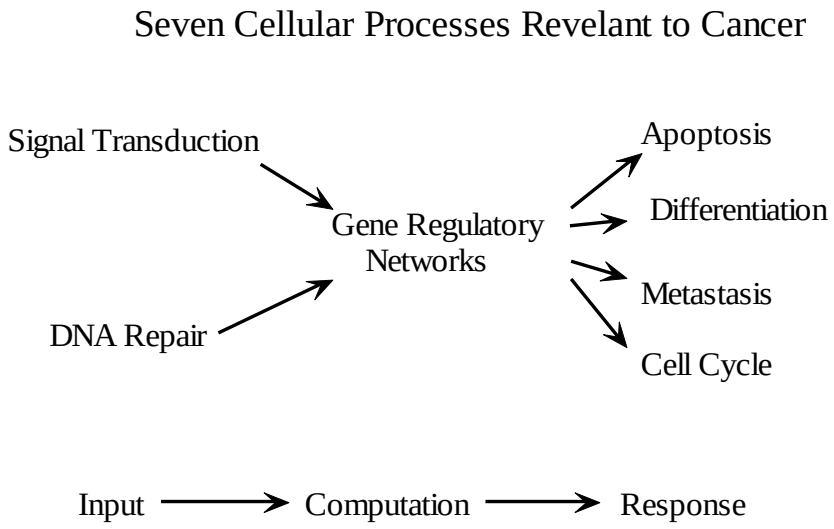

Figure 1.1: The seven cellular processes shown in this figure are all relevant to cancer. The processes fall into three categories: input, computation, and response. This work focuses on DNA double-strand break repair, a subsystem of DNA repair.

Aside from the background material contained in chapter 2, this dissertation consists of two parts: the DSB misrepair models of chapters 3 and 4, and the DNA fragment size distribution models of chapters 5 and 6. The DSB misrepair models are intended to serve as integral components of radiation induced cancer risk models. Fragment distribution models, on the other hand, are developed for the purpose of improving pulsed field gel electrophoretic (PFGE) estimates of unrejoined and misrejoined DSBs.

In chapter 3 several deterministic models [44] of DSB misrepair are compared to recent PFGE time course data [25] (X-rays delivered to human fibroblasts as an acute 80 Gray dose). The Sax subset (SS) model, so named because it applies the breakage-and-reunion model of Karl Sax [40] to a subset of the induced DSBs, is selected from five models as the best representation of the data [25]. The SS model, though adequate at 80 Gray doses, is inadequate at lower doses because it does not include the accidental binary eurepair of correct DSB free ends. In chapter 4 a stochastic extension of the SS model, the Sax-Markov binary eurepair (SMBE) model, is developed for chromosomal aberration applications (doses less than 8 Gray). In conjunction with recent dicentric yield data [15], the SMBE model suggests that the free ends of a DSB are constrained to a nuclear subvolume approximately 100 times smaller than the cell's nucleus. Assuming radiation induced cancer results from chromosomal translocations, cancer risks are reasoned to be linear in misrejoined DSBs for doses less than 1 Gray.

The second part of this research concerns DNA fragment size distributions. Chapter 5 provides detailed derivations [34] of two induced DNA fragment distribution models currently available, the EtBr model [9], and the specific probe model [10]. These models can be used to form distribution-based DSB estimates, but only when the level of misrepairs is insignificant. Chapter 6 introduces the Sax subset EtBr (SSEB) model, an extension of the EtBr model to include the misrepair dynamics of the SS model. The SSEB model is solved analytically for exponentially distributed (high dose) initial conditions. A distribution-based approach to misrepair estimation is outlined. Concluding remarks and ideas for future work are provided in chapter 7.

Chapter 2

Background

The following topics will be discussed: linear energy transfer (LET), the formation of DNA double-strand breaks (DSBs), the DSB yield, and the measurement of DSBs using pulsed field gel electrophoresis (PFGE).

2.1 Linear Energy Transfer

A photon is a wave packet of orthogonal electric and magnetic fields where the frequency of the wave determines the photon's energy and an envelope multiplies the wave amplitude thus localizing it spatially. Charged particles on the other hand, produce constant electric fields in the moving frame and spiked electric fields in the stationary frame. In almost all cases, the interaction between radiation and matter is through the coupling of an electric field with the dipole moment of a target molecule.[1] For a specific particle and energy, the LET (linear energy transfer) is defined as the *average* linear density of energy depositions over the remainder of the particle track [19].[2] It is known that charged particles have a higher LET than photons and that high energy photons have a lower LET than low energy photons [17]. One explanation for both of these facts is that slower moving electric fields couple more strongly to molecular dipole moments than do highly oscillatory fields. This would also explain why protons (being slow) have higher LET than electrons and why slow electrons have higher LET than fast electrons.

2.2 Double-Strand Break Formation

Ionizing radiation creates a wide variety of reactive chemical species, the most important of which are hydroxyl radicals in the neighborhood of DNA [45]. Upon striking the DNA, hydroxyls can cause base-adducts which lead to mutations, or they can react with the ribose moiety to create a single strand break (SSB) in the DNA backbone. If two SSBs occur within about 10 base-pairs (bp) of each other, but on opposite strands of the DNA, a double strand break (DSB) is likely to result. Unlike chemicals that cause single strand breaks uniformly throughout the

[1] The exception is the neutron, but even here, the neutron first undergoes the rare event of actually hitting a nucleus. The secondaries, charged protons and electrons, are then responsible for the biological effects.

[2] A particle track is the set of ionizing event positions produced as a consequence of one incoming high energy particle.

genome, ionizing radiation delivers energy along particle tracks so that SSBs become correlated in space. This spatial correlation of SSBs makes ionizing radiation particularly effective at producing DSBs. It is now well accepted that DSBs are the main intermediate responsible for the biological effects of ionizing radiation [46]. This research will thus focus on mathematical descriptions of rejoining and misrejoining DSBs.

2.3 Double-Strand Break Yield

By definition, one Gray of absorbed dose equals one joule of energy absorbed into one kilogram of mass. The dose itself says nothing about how the energy is spread over time. In particular, it is the dose rate that determines the rate at which DNA damaging radicals are generated within a cell. If it is assumed that the same fraction of these radicals reacts with DNA regardless of their level,[3] then, regardless of the dose rate schedule, the same number of DSBs will be generated for a given dose.[4] Since the amount of damage a DNA molecule receives is proportional to its size, the number of DSBs per base pair should be directly proportional to dose. This proportionality is known as the DSB yield and has been estimated to be approximately 0.5 DSBs per 100 Mbp per Gray [8].

2.4 Double-Strand Break Measurement

Pulsed field gel electrophoresis (PFGE) is a relatively recent technique [4, 41] in which DNA fragments as large as 12 million base-pairs (Mbp) are sorted according to size. Since an increase in dose causes the average DNA fragment length to decrease, higher doses are correlated with larger amounts of DNA entering the gel. If the cut-off fragment size for gel entry is known, induced DNA fragment distribution models can be used to convert the fractions of activity released (FAR)[5] into λ, the number of DSBs per bp created just after an acute exposure [5, 3] (see figure 5.3). Two problems arise with the use of FAR [6]: 1) the cut-off fragment size is rarely known and can vary significantly between labs, and 2) replication forks from S-phase cells leave DNA trapped within the well, artificially lowering the measured FAR value.[6] If one includes a gel lane of size markers (fission and budding yeast chromosomes) to obtain a calibration curve of gel position versus fragment size, a fragment distribution model (see chapter 5) can be fit by least squares to the measured distribution to form estimates of λ [5, 21]. Compared to FAR, advantages of distributional approaches are: 1) the shape of the measured distribution is used, not just the integral, and 2) the estimate is independent of S-phase DNA trapped in the well.

[3] DNA viewed as a radical scavenger can hardly be saturated.

[4] It is assumed that there are a neglible number of DSBs created by SSBs from two separate particle tracks.

[5] "Activity" refers to radiolabel activity or fluorescence activity.

[6] It is not clear why FAR is still being used [5, 6, 25, 48]. Perhaps sensitivity is gained at low DSB levels by "pooling the samples" into one FAR value.

Chapter 3

Deterministic Models of Double-Strand Break Misrepair

This chapter compares five different mathematical models of DSB misrepair to recent experimental data [25]. Corresponding to each model is a proposed DSB misrejoining mechanism. In the Sax [40] mechanism, DSBs are either *inactive*, in which case their ends are held together by protein "splints", or they are *active*, in which case the DSB ends are free to randomly misrejoin with other active DSB free ends. In the Revell [35] mechanism, complete (reciprocal) exchanges between inactive DSBs are the only source of misrepair.

The Revell binary misrepair (RBM) model, a high dose limiting form of the repair misrepair (RMR) model previously presented by Tobias [44], is shown to be inconsistent with recent experimental data [25]. Although the Sax activation (SA) model fits the data better than the RBM model, it fails to capture the early misrepair transients. The Sax-Revell activation (SRA) model correctly fits the misrepair transients but incorrectly predicts an increasing steady state misrepair fraction between 80 and 160 Gray. In the Sax subset (SS) model, a subset of the DSBs are defined as active and the remainder as inactive. The SS model appears to be the best of the models since it has one parameter less than the SRA model and correctly predicts the misrepair transients as well as a constant misrepair fraction from 80 to 160 Gray. A fifth model, the Revell single-strand break (RSSB) model, fits the data as well as the SS model but includes an additional parameter. The scope of this chapter is limited to mammalian cells at doses between 80 and 160 Gray.

3.1 The Data Set

In a recent paper by Löbrich et al. [25], DNA was digested into notI restriction fragments before separation on a PFGE gel. NotI restriction fragments act like chromosomes within a cell because the cut sites, like telomeres, are chemically inactive during DSB dynamics. When notI-somes are stained in the gel with EtBr, the signals of many notI-some "bands" overcrowd the gel. Löbrich et al.[25] masked-out the signal from all but one of the bands by hybridizing a probe to a unique 3.2 Mbp notI-some. As the DSBs rejoined, the probe labeled fragments that returned to the 3.2 Mbp peak were eurepaired while those that remained outside of the band were either misrejoined

or unrejoined. By taking minus the log of the intensity within the band divided by the total signal in the lane, Löbrich estimated the sum of the expected number of unrejoined and misrejoined DSBs within the 3.2 Mb fragment. A conventional FAR estimate of unrejoined DSBs thus enabled Löbrich to predict the expected number of misrejoined DSBs. The asterisks $[U(t)]$, open circles $[M(t)]$ and plus signs $[H(t) = M(t) + U(t)]$ shown in the figures of this chapter as data points are actually samples taken by visual inspection from a curve-fit provided by Löbrich et al. [25] (figure 4 therein).

3.2 The Revell Binary Misrepair (RBM) Model

In the Revell [35] mechanism of DSB misrepair, inactive DSBs either form complete exchanges or eurepairs. This mechanism underlies the following mathematical model due to Tobias [44], hereafter called the Revell binary misrepair (RBM) model,

$$\frac{dU(t)}{dt} = -\lambda_r U(t) - \kappa U^2(t) \tag{3.1}$$

$$\frac{dM(t)}{dt} = \kappa U^2(t), \qquad U(0) = GD, \quad M(0) = 0, \tag{3.2}$$

where $U(t)$ is the expected number of inactive DSBs within the cell at time t, $M(t)$ is the expected number of misrejoinings, λ_r is the linear eurepair rate constant, κ is the binary misrepair rate constant, G is the yield (40 DSBs per Gray) and D is the dose in Gray. The first equation states that inactive DSBs are consumed either by linear eurepair or binary misrepair. The second equation states that the number of misrejoinings at time t equals the sum (integral) of all misrejoining events that have occurred in the time interval $[0, t]$. Acute doses are implicitly assumed through the use of non-zero initial conditions.

Equation 3.1 for $U(t)$ can be integrated [44] by substituting $U(t) = \frac{1}{V(t)}$ to give

$$\begin{aligned}
\frac{-\frac{dV(t)}{dt}}{V^2(t)} = -\lambda_r \frac{1}{V(t)} - \kappa \frac{1}{V^2(t)} \quad &\Rightarrow \quad \frac{dV(t)}{dt} = \lambda_r V(t) + \kappa \\
\frac{d}{dt}(V(t)e^{-\lambda_r t}) = \kappa e^{-\lambda_r t} \quad &\Rightarrow \quad V(t)e^{-\lambda_r t} - V(0) = \frac{\kappa}{\lambda_r}(1 - e^{-\lambda_r t}) \\
V(t) = e^{\lambda_r t}(\frac{\kappa}{\lambda_r}(1 - e^{-\lambda_r t}) + V(0)) \quad &\Rightarrow \quad U(t) = \frac{U(0)e^{-\lambda_r t}}{1 + \frac{U(0)\kappa}{\lambda_r}(1 - e^{-\lambda_r t})}.
\end{aligned} \tag{3.3}$$

Since the number of DSBs repaired is $R(t) = \int_0^t \lambda_r U(\tau) d\tau$ we have

$$R(t) = \frac{\lambda_r}{\kappa} \int_0^t \frac{U(0)\kappa e^{-\lambda_r \tau}}{1 + \frac{U(0)\kappa}{\lambda_r}(1 - e^{-\lambda_r \tau})} d\tau = \frac{\lambda_r}{\kappa} \ln\left(1 + \frac{U(0)\kappa}{\lambda_r}(1 - e^{-\lambda_r t})\right). \tag{3.4}$$

Because $M(t) = [U(0) - U(t)] - R(t)$ we obtain

$$M(t) = U(0) - \frac{U(0)e^{-\lambda_r t}}{1 + \frac{U(0)\kappa}{\lambda_r}(1 - e^{-\lambda_r t})} - \frac{\lambda_r}{\kappa} \ln\left(1 + \frac{U(0)\kappa}{\lambda_r}(1 - e^{-\lambda_r t})\right) \tag{3.5}$$

$$M(\infty) = U(0) - \frac{\lambda_r}{\kappa} \ln(\frac{U(0)\kappa}{\lambda_r} + 1). \tag{3.6}$$

Recent experimental data [25] has shown that $M(\infty) \approx 0.33U(0)$ and that if t^* is defined by $U(t^*) = 0.25U(0)$, then $M(t^*) \approx 0.50M(\infty)$. The constraint $M(\infty) = \frac{U(0)}{3} = U(0) - \frac{\lambda_r}{\kappa} \ln(\frac{U(0)\kappa}{\lambda_r} + 1)$ implies that $\frac{2}{3}\frac{U(0)\kappa}{\lambda_r} = \ln(\frac{U(0)\kappa}{\lambda_r} + 1)$

which has the numerical solution $\frac{U(0)\kappa}{\lambda_r} = 1.14$. By definition of t^*, $0.25 = \frac{e^{-\lambda_r t^*}}{1+1.14(1-e^{-\lambda_r t^*})}$ or $0.416 = e^{-\lambda_r t^*}$ so that

$$\begin{aligned} \frac{M(t^*)}{M(\infty)} &= \frac{0.75U(0) - \frac{\lambda_r}{\kappa}\ln(1+1.14(1-0.416))}{0.33U(0)} \\ &= \frac{0.75 - \frac{1}{1.14}\ln(1+1.14(1-0.416))}{0.33} = 0.916. \end{aligned} \tag{3.7}$$

Since Löbrich et al. [25] concluded that $\frac{M(t^*)}{M(\infty)} \approx 0.5$, the RBM model must be rejected. To further illustrate the inconsistency of the RBM model, suppose we wish to match $M(\infty)$ and $\frac{dU(0)}{dt}$ with the data of Löbrich et al. [25].[1] The corresponding parameter constraints, $1.14 = \frac{U(0)\kappa}{\lambda_r}$ and $\frac{0.55U(0)}{1800} = \lambda_r U(0) + \kappa U^2(0)$, can be solved to give $\kappa = 5\text{x}10^{-8}$ and $\lambda_r = 1.4\text{x}10^{-4}$ (units of sec^{-1} are used throughout). These parameter values were used in the left panel of figure 3.1. In the right panel of figure 3.1 the slow component of $U(t)$ is matched instead of $M(\infty)$. The parameters in this case were obtained by noting that the DSB level drops by a factor of two between four and eight hours of repair time (see figure 4 of Löbrich et al. [25]) so that $e^{-\lambda_r 14400} = 0.5$ (since $U(t) \approx U(0)e^{-\lambda_r t}$ for times greater than four hours) and thus $\lambda_r = 4.8\text{x}10^{-5}$ which implies $\kappa = 7.8\text{x}10^{-8}$.

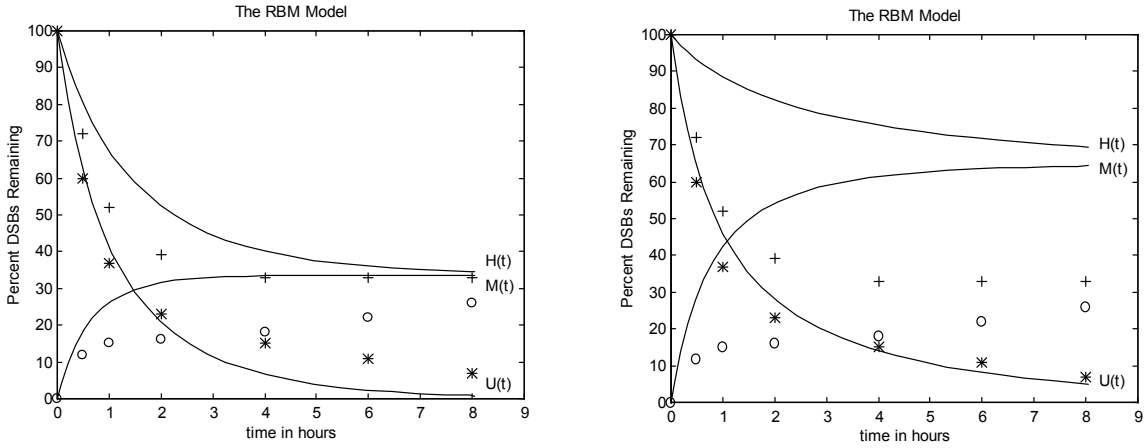

Figure 3.1: Left panel: If $M(\infty)$ and $\frac{dU(0)}{dt}$ are forced to match the data, the RBM model ($\kappa = 5\text{x}10^{-8}$ and $\lambda_r = 1.4\text{x}10^{-4}$) mispredicts the slow component of $U(t)$ and the transients of $M(t)$. Right panel: If both components of $U(t)$ are forced to match the data, the RBM model ($\kappa = 7.8\text{x}10^{-8}$ and $\lambda_r = 4.8\text{x}10^{-5}$) mispredicts $M(\infty)$. Both panels: The asterisks $U(t)$ and plus signs $H(t) = M(t) + U(t)$ are samples taken by visual inspection from a curve-fit provided with the actual data. The open circles $M(t)$ were formed as the difference $H(t) - U(t)$.

[1] In figure 5 of Lobrich et al. [25], 55 percent of the DSBs have disappeared within 1800 seconds.

3.3 The Sax Activation (SA) Model

In the Sax activation (SA) model, inactive DSBs are either activated or repaired and the free ends of active DSBs randomly misrejoin. Let $U_1(t)$ be the mean number of inactive DSBs and $U_2(t)$ the mean number of active DSB *free ends* [note that $U(t) = U_1(t) + \frac{1}{2}U_2(t)$]. The SA model is then

$$\frac{dU_1(t)}{dt} = -\lambda_r U_1(t) - k_1 U_1(t) \tag{3.8}$$

$$\frac{dU_2(t)}{dt} = 2k_1 U_1(t) - \kappa U_2^2(t) \tag{3.9}$$

$$\frac{dM(t)}{dt} = \frac{1}{2}\kappa U_2^2(t), \qquad U_1(0) = GD, \quad U_2(0) = 0, \quad M(0) = 0, \tag{3.10}$$

where k_1 is the DSB activation rate constant, κ is the binary misrepair rate constant of *active* DSB ends, and λ_r, G, D, and $M(t)$ are defined as in the RBM model. The first equation integrates to $U_1(t) = U(0)e^{-(\lambda_r + k_1)t}$ while $U_2(t)$ is solved numerically, see figure 3.2. To match this model to the data of Löbrich et al. [25], the steady state constraint $0.33 = \frac{k_1}{k_1+\lambda_r}$ and the initial slope constraint $\frac{dU(0)}{dt} = -\lambda_r U_1(0) - \frac{\kappa}{2}U_2^2(0) = -\lambda_r U(0) = \frac{-0.55U(0)}{1800}$ together imply that $\lambda_r = 0.0003$ and $k_1 = 0.00015$. At repair times greater than four hours (14400 seconds), since $U_1(14400) = 10$ is negligible compared to the 480 DSBs that were measured [25], equation 3.9 becomes $\frac{dU_2(t)}{dt} = -\kappa U_2^2(t)$, $U_2(14400) = 960$, with solution $U_2(t) = \frac{U_2(14400)}{U_2(14400)\kappa[t-14400]+1}$. At $t = 8$ hours (28800 seconds) approximately 240 DSBs were measured so $0.5 = \frac{1}{960\kappa 14400+1}$ or $\kappa = 7.2\text{x}10^{-8}$.

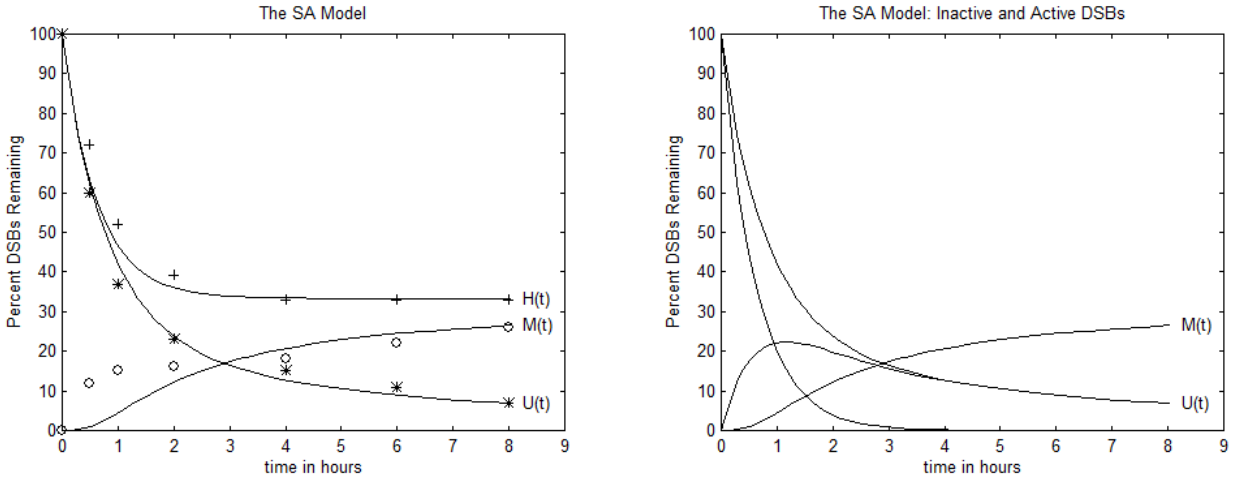

Figure 3.2: Left panel: $U(t)$ and $M(\infty)$ of the SA model match the data. Right panel: Inactive DSBs $[U_1(t)]$ decay rapidly; active DSBs $[\frac{1}{2}U_2(t)]$ rise to a peak and decay slowly. Both panels: $\kappa = 7.2\text{x}10^{-8}$, $\lambda_r = 3.0\text{x}10^{-4}$, and $k_1 = 1.5\text{x}10^{-4}$.

3.4 The Sax-Revell Activation (SRA) Model

The Sax-Revell activation (SRA) model is

$$\frac{dU_1(t)}{dt} = -(\lambda_r + k_1)U_1(t) - \kappa_1 U_1^2(t) \tag{3.11}$$

$$\frac{dU_2(t)}{dt} = 2k_1 U_1(t) - \kappa_2 U_2^2(t) \tag{3.12}$$

$$\frac{dM(t)}{dt} = \frac{1}{2}\kappa_2 U_2^2(t) + \kappa_1 U_1^2(t), \quad U_1(0) = GD,\ U_2(0) = 0,\ M(0) = 0, \tag{3.13}$$

where κ_1 is the misrepair rate constant of inactive DSBs and κ_2 is the misrepair rate constant for active DSB free ends. If $\frac{dM(0)}{dt}$ is matched to the data of Löbrich et al. [25], $\frac{0.12U(0)}{1800} = \kappa_1 U_1^2(0)$ or $\kappa_1 = 2.0\text{x}10^{-8}$. Matching $\frac{dU(0)}{dt}$ to the data gives $\frac{0.55U(0)}{1800} = \lambda_r U_1(0) + \kappa_1 U_1^2(0)$ or $\lambda_r = 0.00024$. Since $U_1(t)$ vanishes after 4 hours, matching the slow component of $U(t)$ requires that $\kappa_2 = 7.2\text{x}10^{-8}$ as in the SA model. To estimate k_1, $M(\infty)$ must be decomposed into a Revell component $M_1(\infty)$ that arises from the κ_1 term, and a Sax component $M_2(\infty)$ that arises from the κ_2 term. If $\lambda_r + k_1$ replaces λ_r in the steady state misrepair equation of the RBM model (equation 3.6), $M_1(\infty)$ in equation 3.14 results. The second term in equation 3.14 is the number of DSBs that did not pass through κ_1 and thus must have been partitioned between k_1 and λ_r with $\frac{k_1}{\lambda_r + k_1}$ of the flux eventually becoming $M_2(\infty)$.

$$M_1(\infty) = U(0) - \frac{\lambda_r + k_1}{\kappa_1}\ln(1 + \frac{U(0)\kappa_1}{\lambda_r + k_1}) \tag{3.14}$$

$$M_2(\infty) = \frac{k_1}{\lambda_r + k_1}\frac{\lambda_r + k_1}{\kappa_1}\ln(1 + \frac{U(0)\kappa_1}{\lambda_r + k_1}) = \frac{k_1}{\kappa_1}\ln(1 + \frac{U(0)\kappa_1}{\lambda_r + k_1}) \tag{3.15}$$

$$M(\infty) = M_1(\infty) + M_2(\infty) = U(0) - \frac{\lambda_r}{\kappa_1}\ln(1 + \frac{U(0)\kappa_1}{\lambda_r + k_1}) \tag{3.16}$$

The steady state constraint $M(\infty) = 0.33U(0)$ implies $e^{0.66U(0)\kappa_1/\lambda_r} = 1 + \frac{U(0)\kappa_1}{\lambda_r + k_1}$ so that $k_1 = \frac{U(0)\kappa_1}{e^{0.66U(0)\kappa_1/\lambda_r} - 1} - \lambda_r$ and hence $k_1 = 0.9\text{x}10^{-4}$ (left panel, figure 3.3). If κ_1 =3.0x10^{-8}, so as to increase $\frac{dM(0)}{dt}$, the time course fit at 80 Gray improves (right panel, figure 3.3) but the misrepair fit at 160 Gray worsens (figure 3.6).

3.5 Misrepairs as a Function of DSBs Consumed

If the previous $M(t)$ plots are redrawn as the percent DSBs misrepaired versus the percent DSBs consumed, see figure 3.4, the SRA model gains further support.[2]

3.6 The Sax Subset (SS) Model

In the Sax subset (SS) model, a subset of the DSBs are defined as active and destined to misrejoin, all other DSBs are inactive and destined for repair. The SS model is

$$\frac{dU_1(t)}{dt} = -\lambda_r U_1(t) \tag{3.17}$$

[2]These plots were suggested by Dr. P.J. Hahnfeldt, Joint Center for Radiation Therapy, Harvard Medical School.

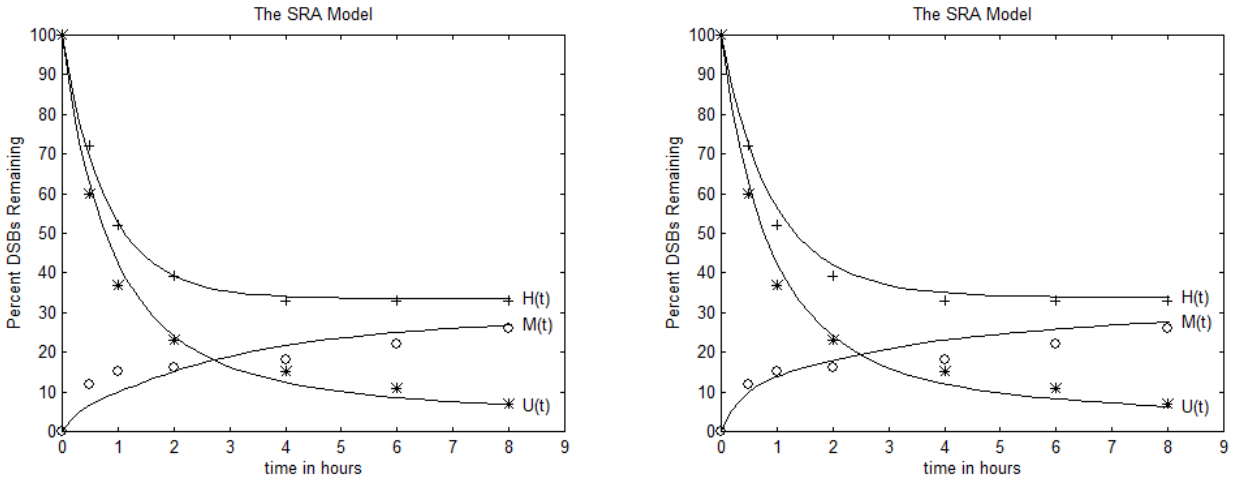

Figure 3.3: The initial misrepair slope of the SRA model in the left panel ($\kappa_1 = 2.0\text{x}10^{-8}$, $\kappa_2 = 7.2\text{x}10^{-8}$, $\lambda_r = 2.4\text{x}10^{-4}$, and $k_1 = 0.9\text{x}10^{-4}$) is increased to provide an improved fit to the data in the right panel ($\kappa_1 = 3.0\text{x}10^{-8}$, $\kappa_2 = 7.2\text{x}10^{-8}$, $\lambda_r = 2.1\text{x}10^{-4}$, and $k_1 = 0.6\text{x}10^{-4}$).

$$\frac{dU_2(t)}{dt} = -\kappa U_2^2(t) \tag{3.18}$$

$$\frac{dM(t)}{dt} = \frac{1}{2}\kappa U_2^2(t) \tag{3.19}$$

$$U_2(0) = 2pU(0),\ U_1(0) = (1-p)U(0),\ M(0) = 0, \tag{3.20}$$

where p is the fraction of DSBs that are active. The solutions, $U_1(t) = U_1(0)e^{-\lambda_r t}$, $U_2(t) = \frac{U_2(0)}{U_2(0)\kappa t+1}$ and $M(t) = \frac{U_2(0)}{2}[1 - \frac{1}{U_2(0)\kappa t+1)}]$ are shown in figure 3.5 using $\kappa = 7.2\text{x}10^{-8}$, $\lambda_r = 3.0\text{x}10^{-4}$, and $p = 0.33$. The agreement between the SS model and the data is excellent.

3.7 The Revell SSB (RSSB) Model

Suppose an induced DSB, as measured by PFGE, has probability $\sigma = 0.66$ of actually being a staggered pair of SSBs which in living cells would have remained intact. Furthermore, let the mean number of such pairs of SSBs be $U_s(t)$ with an exponential decay rate constant of λ_s. The Revell SSB (RSSB) model represents these assumptions as

$$\frac{dU_s(t)}{dt} = -\lambda_s U_s(t) \tag{3.21}$$

$$\frac{dU_1(t)}{dt} = -\lambda_r U_1(t) - \kappa U_1^2(t) \tag{3.22}$$

$$\frac{dM(t)}{dt} = \kappa U_1^2(t) \tag{3.23}$$

$$U_s(0) = \sigma GD, \quad U_1(0) = (1-\sigma)GD, \quad M(0) = 0, \tag{3.24}$$

where $U_1(t)$ is the expected number of inactive DSBs per cell, $U(t) = U_s(t) + U_1(t)$ is the DSB level as measured by PFGE, and $M(t)$, κ, λ_r, G and D, remain as defined in the RBM model. Mathematically, this four parameter model contains the three parameter SS model as a special case. Indeed, the parameter values $\kappa = 2 * 7.2\text{x}10^{-8}$, $\lambda_s = 3.0\text{x}10^{-4}$, $k_1 = 1.5\text{x}10^{-4}$, and $\lambda_r \leq 1.0\text{x}10^{-6}$ produced plots identical to those shown in figure 3.5.

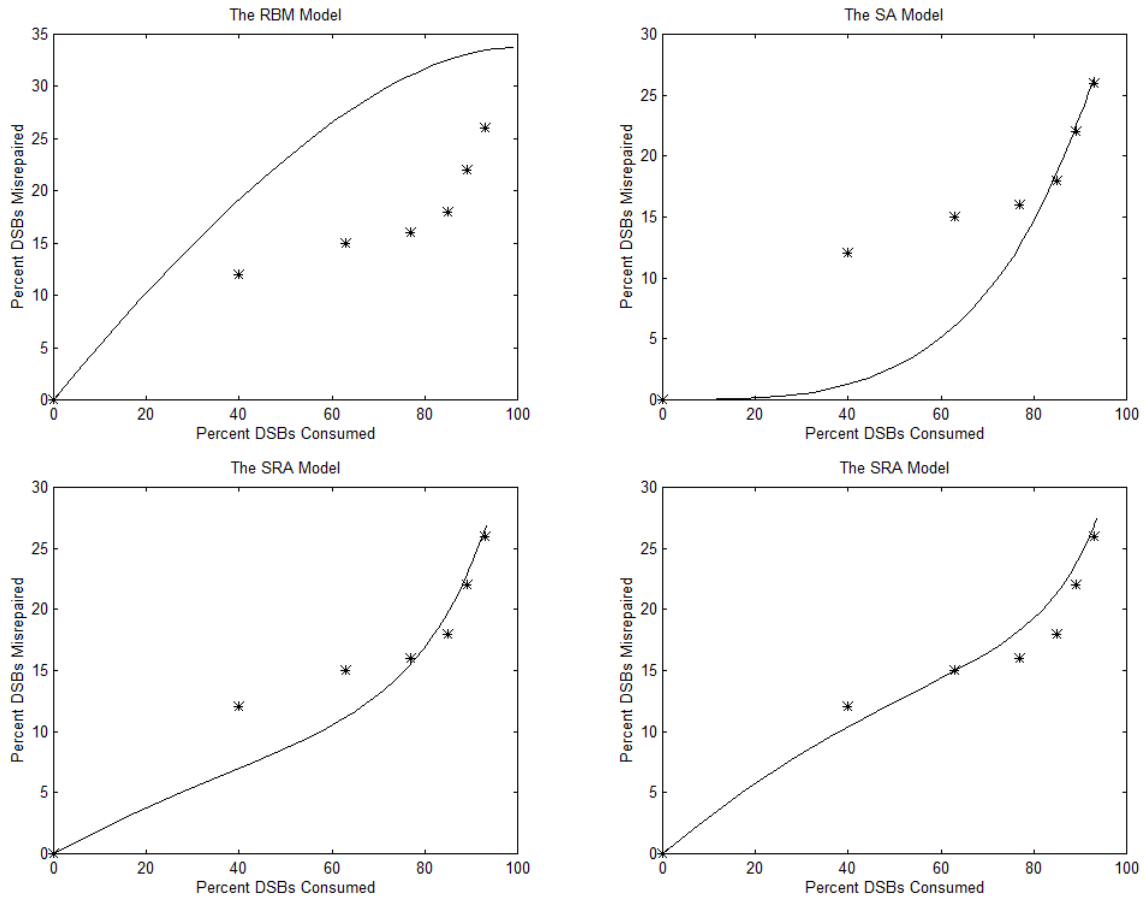

Figure 3.4: Upper left panel: The RBM model, when forced to match $M(\infty) = 0.33U(0)$, is incapable of matching the other data points. Upper right panel: The SA model fits the data better than the RBM model. Lower panels: The SRA model (left and right panels as in figure 3.3) fits the data quite well.

3.8 Misrepair Fractions

In addition to time course data, Lobrich et al. [25] observed a constant steady state misrepair fraction $f = \frac{M(\infty)}{U(0)}$ between 80 and 160 Gray. The consistency of this observation with the models of this chapter is shown in figure 3.6.

3.9 Summary

Five DSB misrepair models were compared to the data of Löbrich et al. [25]. Of these, the RBM model is the least consistent with the data. The SA model improves the fit but fails to capture the early misrepair transients. The SRA model correctly predicts the misrepair transients at the price of one parameter and a new problem, an increasing misrepair fraction between 80 and 160 Gray. The SS model has no problems and uses one parameter less than the SRA model. The RSSB model contains the SS model as a special case and thus matches the data well, but it includes an additional parameter and is therefore not selected. Of the models presented, the SS model appears to be the most appropriate representation of the Löbrich data [25].

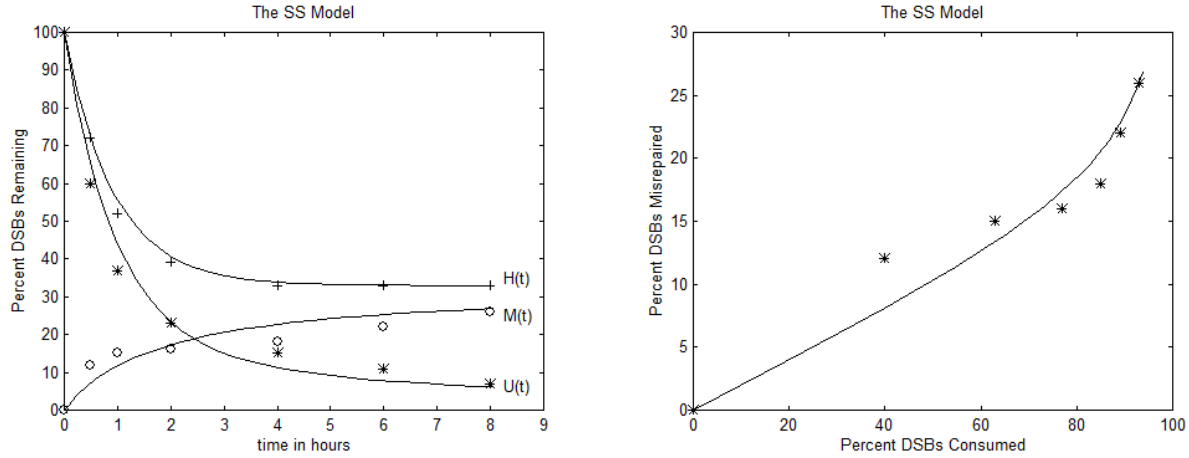

Figure 3.5: Left panel: The SS model differs from the SA model in that the initial slope of $M(t)$ is non-zero, the result is an improved fit to the data. Right panel: In comparison to the four panels shown in figure 3.4, the three parameter SS model provides a fit similar to that of the four parameter SRA model. Both panels: The parameter values were $\kappa = 7.2\text{x}10^{-8}$, $\lambda_r = 3.0\text{x}10^{-4}$, and $p = 0.33$.

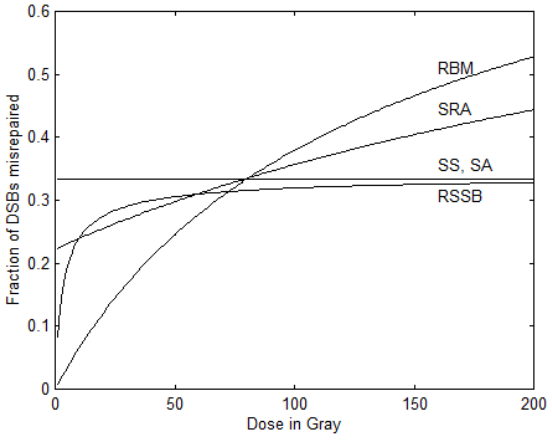

Figure 3.6: The steady state misrepair fractions predicted by each of the five models are plotted as functions of dose. A constant misrepair fraction between 80 and 160 Gray is consistent with the SS, SA, and RSSB models. The parameter values were: RBM model ($\kappa = 5\text{x}10^{-8}$, $\lambda = 1.4\text{x}10^{-4}$), SA model ($\kappa = 7.2\text{x}10^{-8}$, $\lambda_r = 3.0\text{x}10^{-4}$, $k_1 = 1.5\text{x}10^{-4}$), SRA model ($\kappa_1 = 3.0\text{x}10^{-8}$, $\kappa_2 = 7.2\text{x}10^{-8}$, $\lambda_r = 2.1\text{x}10^{-4}$, $k_1 = 0.6\text{x}10^{-4}$), SS model ($\kappa = 7.2\text{x}10^{-8}$, $\lambda_r = 3.0\text{x}10^{-4}$, $p = 0.33$) and RSSB model ($\kappa = 1.4\text{x}10^{-7}$, $\lambda_s = 3.0\text{x}10^{-4}$, $p = 0.33$, $\lambda_r = 1.0\text{x}10^{-6}$).

Chapter 4

A Stochastic Model of Double-Strand Break Misrepair

In the previous chapter the SS model was selected as an adequate representation of recent high dose PFGE data [25]. In this chapter the SS model is extended to lower doses by including the process of accidental binary eurepair between the free ends of completely disconnected (active) DNA double-strand breaks (DSBs). The resulting Sax-Markov binary eurepair (SMBE) model predicts that the expected number of steady state misrepairs is quadratic for low to moderate doses, becoming linear with a non-zero intercept (threshold-like) at higher doses. Since single track effects guarantee linear misrepairs as a function of very low doses, the overall misrepair dose-response is predicted to be piece-wise linear-quadratic-linear (LQL). The LQ portion of the SMBE model is consistent with dicentric aberration yields if the free ends of a DSB are restricted to a nuclear subvolume roughly 100 times smaller than the nucleus.

4.1 The Sax Subset (SS) Model Revisited

In a *single cell*, the reaction rate of $A + A \rightarrow P$ is proportional to the number of possible reactant interactions, i.e. $\frac{dA}{dt} = -2k\binom{A}{2} = -kA(A-1)$. The average reaction rate of an *ensemble of cells* is thus

$$\frac{d < A >}{dt} = -k(< A^2 > - < A >) \tag{4.1}$$

$$= -k(< A >^2 + \sigma^2 - < A >) \tag{4.2}$$

where $< A >$ and σ^2 are the mean and variance of the reactant distribution, respectively. Thus, if the dispersion of the reactant distribution (the variance divided by the mean) equals one (normal dispersion), as with a Poisson distribution, we have

$$\frac{d < A >}{dt} = -k < A >^2 . \tag{4.3}$$

Normal dispersion will be assumed for DSBs; this is consistent with low LET irradiation [36]. Since each active DSB consists of two DSB free ends, the initial distribution of DSB *free ends* is overdispersed with a variance equal to

twice its mean.[1] Letting $U_2(t)$ be the average number of active DSB *free ends* at time t and $U_1(t)$ the mean number of inactive DSBs, the SS model becomes

$$\frac{dU_1(t)}{dt} = -\lambda_r U_1(t) \tag{4.4}$$
$$\frac{dU_2(t)}{dt} = -\kappa[U_2^2(t) + U_2(t)] \tag{4.5}$$
$$\frac{dM(t)}{dt} = \frac{1}{2}\kappa[U_2^2(t) + U_2(t)] \tag{4.6}$$
$$U_2(0) = 2pU(0), \quad U_1(0) = (1-p)U(0), \quad M(0) = 0 \tag{4.7}$$

with solutions $U_2(t) = \frac{U_2(0)e^{-\kappa t}}{1+U_2(0)(1-e^{-\kappa t})}$ and $M(t) = \frac{U_2(0)}{2}\left(1 - \frac{e^{-\kappa t}}{1+U_2(0)(1-e^{-\kappa t})}\right)$ and parameters as previously defined in section 3.6. At the high doses of Löbrich et al. [25], the SS model defined in section 3.6 is adequate since $U_2^2(t) + U_2(t) \approx U_2^2(t)$. Note that all active DSBs in the SS model eventually form misrepairs (accidental binary eurepair is not included in the SS model) and thus $M(\infty) = pU(0)$ irrespective of dose.

4.2 The Sax-Markov Binary Eurepair Model

The SS model does not account for the possibility that two active DSB free ends can rejoin with their correct mates, i.e. form an accidental binary eurepair. At low doses, if only one active DSB is created, the free ends have no choice but to eventually rejoin correctly. At higher doses, binary eurepair becomes less significant because correct partners are less likely to find each other. To model accidental binary eurepair, we shall distinguish between two types of active DSB *free ends*: those that are destined to become misrepairs because their true mates have already remarried (let there be $U_3(t)$ of these) and those that still have a chance of finding their correct mates (there are $U_2(t) - U_3(t)$ of these). The reaction rates for *one* cell are then

$$\begin{aligned}\frac{dU_2(t)}{dt} &= -2\kappa\frac{[U_2(t)-U_3(t)][U_2(t)-U_3(t)-2]}{2} - 2\kappa\frac{U_3(t)[U_3(t)-1]}{2} \\ &\quad - 2\kappa U_3(t)[U_2(t)-U_3(t)] - 2\kappa\frac{U_2(t)-U_3(t)}{2} \qquad (4.8)\\ &= -2\kappa\frac{U_2(t)[U_2(t)-1]}{2} \qquad (4.9)\end{aligned}$$
$$\frac{dU_3(t)}{dt} = 2\kappa\frac{[U_2(t)-U_3(t)][U_2(t)-U_3(t)-2]}{2} - 2\kappa\frac{U_3(t)[U_3(t)-1]}{2} \tag{4.10}$$
$$\begin{aligned}\frac{dM(t)}{dt} &= \kappa\frac{[U_2(t)-U_3(t)][U_2(t)-U_3(t)-2]}{2} + \kappa\frac{U_3(t)[U_3(t)-1]}{2} \\ &\quad + \kappa U_3(t)[U_2(t)-U_3(t)] \qquad (4.11)\\ &= \kappa\frac{U_2(t)[U_2(t)-1]}{2} - \kappa\frac{U_2(t)-U_3(t)}{2} \qquad (4.12)\end{aligned}$$
$$U_2(0) = 2pU_1(0), \quad U_3(0) = 0, \quad M(0) = 0,$$

where $M(t)$, $U(0)$, κ, and p are defined as in the SS model and $U_1(t)$ is not modeled since it does not affect misrepair. To understand equations 4.8 to 4.12, note that the total number of free end interactions within our single cell, i.e. $\binom{U_2(t)}{2}$, equals the sum of four components:

1. The binary eurepair interactions: There are $\frac{U_2(t)-U_3(t)}{2}$ of these interactions. Since they do not lead to misrepairs, we subtract them from the total interactions to arrive at equation 4.12 for $\frac{dM(t)}{dt}$.

[1] If X is a random variable with mean μ, variance σ^2, and thus a dispersion $\frac{\sigma^2}{\mu}$, the random variable $Y = nX$ has mean $n\mu$, variance $n^2\sigma^2$, and dispersion $n\frac{\sigma^2}{\mu}$.

2. The misrepair interactions within the $U_2 - U_3$ pool: The number of such interactions is $\frac{1}{2}[U_2(t) - U_3(t)][U_2(t) - U_3(t) - 2]$. Productive interactions create one misrepair (the first term in equation 4.11) and two new elements of the U_3 pool (the synthesis term in equation 4.10).

3. The misrepair interactions within the U_3 pool: There are $\frac{1}{2}U_3(t)[U_3(t) - 1]$ of these. Misrepairs by this route (second term in equation 4.11) simultaneously degrade two members of the U_3 pool (degradation term in equation 4.10).

4. The misrepair interactions between the U_3 and $U_2 - U_3$ pools: The number of such interactions is $U_3(t)[U_2(t) - U_3(t)]$. Since this reaction cannot occur without creating a new U_3 member from the $U_2 - U_3$ pool, it does not enter into equation 4.10. Productive interactions remove two members of the U_2 pool (third term in equation 4.8) while creating one misrepair (third term in equation 4.11).

Note that members of the U_3 and $U_2 - U_3$ pools are synthesized and degraded in pairs, so $U_2(t) - U_3(t)$ is restricted to non-negative *even* integers and thus $\frac{1}{2}[U_2(t) - U_3(t)][U_2(t) - U_3(t) - 2] \geq 0$.

Although active DSB free ends are initially distributed with a variance equal to twice their mean (low LET is assumed throughout), cells that have many DSBs are expected to rejoin more rapidly than those with few DSBs, so the overdispersion should decrease with time. Furthermore, an analytic expression for the variance of the $U_3(t)$ distribution is not readily available. Since there is no guarantee that the variances of $U_2(t)$ and $U_3(t)$ can even be expressed as functions of the means, no attempt will be made to derive a set of ordinary differential equations for the mean value dynamics of $U_2(t)$ and $U_3(t)$ in terms of only the mean values themselves (this was done implicitly in the SS model). Instead, to represent accidental binary eurepair we shall use Markov models [1, 38, 36], a stochastic formulation with the flexibiliy needed for handling arbitrary reactant distributions.

The Sax-Markov binary eurepair (SMBE) model, depicted as a Markov graph in the left panel of figure 4.1, is obtained from equations 4.8 to 4.12 by letting $U_2(t) = 2j$, $U_3(t) = 2i$ and $M(t) = m$ (note that i and j count DSBs, i.e. pairs of DSB free ends). A typical node (j, i, m) in this graph [e.g. (3,1,1)] has four arrows emanating from it corresponding to the four types of active DSB interactions: 1) the eurepair arrow [rate constant $\kappa(j - i)$] terminates on the node $(j - 1, i, m)$; 2) the "within $U_2 - U_3$" misrepair arrow [rate constant $\kappa(j - i)(2(j - i) - 2)$] terminates on the node $(j - 1, i + 1, m + 1)$; 3) the "within U_3" misrepair arrow [rate constant $\kappa(2i - 1)i$] terminates on the node $(j - 1, i - 1, m + 1)$; and 4) the cross-reaction misrepair arrow [rate constant $4\kappa(j - i)i$] terminates on the node $(j - 1, i, m + 1)$. The node (j, i, m) contains the probability $P(j, i, m, t)$ that a cell at time t has j active DSBs, i active DSBs destined for misrepair, and m misrepairs. The arrows are labeled with the first order rate constants of probability mass transfer between the nodes.

Just after an acute dose, the graph in figure 4.1 has probability mass only on the j axis. As time evolves, the probability mass migrates to larger m and different levels of i as it moves toward the m-axis. At steady state all of the probability mass lies on the m-axis, the mass in node $(0, 0, m)$ being the probability that a cell incurred m misrepairs as a result of the exposure.

4.3 Steady State SMBE Misrepairs

The steady state solution of Markov models can be found without computing the transient distributions [18]. This is done by replacing the rate constants in the Markov graph with partitioning fractions equal to the original rate

constant divided by the sum of all rate constants emanating from a particular node. The SMBE model (see the right panel of figure 4.1) has the following flux partitioning fractions emanating from the node (j, i, m):

1. $\frac{j-i}{j(2j-1)}$ is the eurepair flux fraction to i, i.e. the fraction leading to the node $(j-1, i, m)$;
2. $\frac{(j-i)[2(j-i)-2]}{j(2j-1)}$ is the misrepair flux fraction to $i+1$, i.e. the fraction leading to the node $(j-1, i+1, m+1)$;
3. $\frac{i(2i-1)}{j(2j-1)}$ is the misrepair flux fraction to $i-1$, i.e. the fraction leading to the node $(j-1, i-1, m+1)$; and
4. $\frac{4i(j-i)}{j(2j-1)}$ is the misrepair flux fraction to i, i.e. the fraction leading to the node $(j-1, i, m+1)$.

Negative fractions are set to zero. Note that the denominator in each of these fractions, $j(2j-1)=\binom{2j}{2}$, equals the total number of free end interactions. Since the probability mass in any node is eventually divided among daughter nodes according to these partitioning constants, the steady state distribution is reached more efficiently if the node masses are divided in one step. The algorithm begins with the j axis initialized to a Poisson distribution of active DSBs with mean μ. Starting with a j value large enough that a negligible amount of probability mass lies at larger values, we partition the mass on the j axis toward the m-axis. During the first iteration ($m = 0$), $i = 0$ implies that two of the flux fractions are zero so only the row $i = 1$ receives mass at $m = 1$. The second iteration leaves a stack of three rows at $m = 2$ (see figure 4.1), and in general, an iteration at m pushes a stack of m rows toward $j = 0$ and into the stack at $m + 1$, similar to a person shoveling snow on a driveway. The front of the stack (i.e. the end closer to the plane $j = 0$) ramps toward $i = 0$ since i cannot be greater than j. At each iteration on m, the stack increases in height, but decreases in length. The algorithm stops when the length of the stack reaches unity. Mean values of the final misrepair distribution (the m axis) shown in table 4.1 indicate that $M(\infty)$ approaches $\mu - \frac{1}{2}$ at high doses and $\frac{2}{3}\mu^2$ at low doses. This constrasts with the SS model which ignores accidental binary eurepair and predicts $M(\infty) = \mu$ for all doses.

Table 4.1: Simulation results of the SMBE model.

Gray	0.00075	0.0075	0.0150	0.0375	0.075	0.375	0.75	7.5
μ	0.01	0.1	0.2	0.5	1	5	10	100
$M(\infty)$	0.000067	0.0064	0.0247	0.1376	0.462	4.42	9.469	99.5

4.4 Asymptotic Analytic SMBE Solutions

The low dose expression $M(\infty) = \frac{2}{3}\mu^2$ (see table 4.1) can be understood by noting that for small values of μ, the Poisson probability of zero or one active DSB is

$$e^{-\mu} + \mu e^{-\mu} \quad \approx \quad 1 - \mu + \frac{\mu^2}{2} \quad + \quad \mu - \mu^2 \quad = \quad 1 - \frac{\mu^2}{2}. \tag{4.13}$$

Thus, the probability mass at $j = 2$ active DSBs can be approximated as $\frac{\mu^2}{2}$. One third of this mass is expected to return to $m = 0$ through binary eurepair because two out of the $\binom{4}{2}$=6 free end interactions will force both DSBs to eurepair. The remaining two thirds of the mass will end up at $m = 2$ so

$$M(\infty) = \frac{2\mu^2}{3}. \tag{4.14}$$

At moderate doses and higher, the simulation results in table 4.1 indicate that the expected number of misrepairs approaches $\mu - 0.5$. Approximating the Poisson distribution at high doses with a zero variance distribution $P(j = \mu) = 1$, and noting that each free end has a $\frac{1}{2j-1}$ chance of finding its correct mate, since there are j pairs of free ends we expect $\frac{j}{2j-1}$ eurepairs, or 0.5, as μ becomes large.[2]

4.5 Compartmental Nuclei

Suppose the nucleus is divided such that a DSB is equally likely to find itself in any of η isolated subcompartments. The expected number of misrepairs and active DSBs in the cell is then η times the expected number of misrepairs and active DSBs in one compartment, respectively. The result is that the low dose asymptote becomes

$$M(\infty) = \eta \frac{2}{3}(\frac{\mu}{\eta})^2 = \frac{2}{3\eta}\mu^2 \tag{4.15}$$

and the high dose asymptote

$$M(\infty) = \eta(\frac{\mu}{\eta} - \frac{1}{2}) = \mu - \frac{\eta}{2}. \tag{4.16}$$

The same arguments can be applied to the SS model in section 4.1 to yield the Sax subset compartmental (SSC) model

$$\begin{aligned}
\frac{dU_1(t)}{dt} &= -\lambda_r U_1(t) && (4.17)\\
\frac{dU_2(t)}{dt} &= -\kappa[\frac{1}{\eta}U_2^2(t) + U_2(t)] && (4.18)\\
\frac{dM(t)}{dt} &= \frac{1}{2}\kappa[\frac{1}{\eta}U_2^2(t) + U_2(t)] && (4.19)\\
& U_2(0) = 2pU(0), \quad U_1(0) = (1-p)U(0), \quad M(0) = 0 && (4.20)
\end{aligned}$$

with solutions

$$U_2(t) = \frac{U_2(0)e^{-\kappa t}}{1 + \frac{U_2(0)}{\eta}(1 - e^{-\kappa t})}, \qquad M(t) = \frac{U_2(0)}{2}\left(1 - \frac{e^{-\kappa t}}{1 + \frac{U_2(0)}{\eta}(1 - e^{-\kappa t})}\right). \tag{4.21}$$

Note that estimates of κ made in chapter 3 using $\eta = 1$ will change with $\eta > 1$. Although κ does not affect $M(\infty)$, it is important that its value be reasonable, i.e. a single DSB should not require a year to rejoin.

The number of subcompartments is important for determining the transition dose below which single track effects dominate the dose-response. For example, if the whole nucleus expects the traversal of only one particle track at 0.005 Gray, a compartmental nucleus will have its transition dose at approximately 0.005η Gray.

4.6 Single Track Effects

At doses for which less than one particle track is expected to traverse a nuclear subcompartment, decrements in dose result in commensurate decrements in the number of compartments receiving typically just one track, and thus commensurate decrements in the expected number of misrepairs. Since the SMBE model does not consider particle

[2]This high dose argument was provided by Prof. R.K. Sachs, Department of Mathematics, University of California, Berkeley.

track features, its prediction that $M(\infty) = \frac{2}{3\eta}\mu^2$ can only be expected to hold at doses high enough to deliver several particle tracks to each subcompartment. Let us assume that a single track through a subcompartment produces a Poisson distributed number of active DSBs with mean μ_1 and that this is the single track transition μ between linear and quadratic behavior. If D_1 is the dose at which the whole nucleus expects one particle track traversal, $M(\infty)$ can be modified to be

$$\begin{aligned} M(\infty) &= \frac{2}{3\eta}\mu_1\mu(\frac{\mu}{\mu_1}+1) \\ &= \frac{2}{3\eta}\mu(\mu+\mu_1) \\ &= \frac{2\gamma^2}{3\eta}D(D+\eta D_1), \end{aligned} \tag{4.22}$$

where $\mu_1 = \gamma\eta D_1$ and γ is the yield of active DSBs per Gray. Since a compartment with two active DSBs is equally likely to produce two translocations as it is one dicentric and one acentric, the number of misrepairs should roughly follow four times the yield of dicentrics. From Finnon et al. [15] the dicentric yield was fit to the model $\alpha D + \beta D^2 = \beta D(D + \frac{\alpha}{\beta})$ with $\alpha = 0.03$ and $\beta = 0.08$. We thus have $\frac{2\gamma^2}{3\eta} = 4\beta = 0.32$ and $\eta D_1 = \frac{\alpha}{\beta} = 0.375$ Gray.

A crude estimate of D_1 can be obtained by assuming: 1) the particle track is a straight line, 2) energy depositions are uniformly distributed along the track, 3) a column of cells is aligned with the track, and 4) cells and cell nuclei are 10 μm and 5 μm cubes, respectively. For 80 keV electrons (225 kV X-rays [30]) a track range of 100 μm, or 10 cubic cells in our model, is predicted [17]. Thus, using the four assumptions, one 80 keV electron will deposit 8 keV in each cell or 4 keV in each nucleus. We assume that a nucleus has a volume of 125 μm^3 with unit density so that one Gray delivers 781.25 keV to each nucleus and thus $D_1 \approx \frac{4}{800} = 0.005$ Gray. This value of D_1, in conjunction with $\alpha = 0.03$ and $\beta = 0.08$, implies $\eta = \frac{0.375}{0.005} = 75$ and $\gamma = \sqrt{0.32 * 75 * 1.5} = 6$ active DSBs per Gray.

The analysis of Löbrich's data [25] using the SS model in the previous chapter suggests that $\gamma = \frac{40}{3} = 13 \neq 6$. Assuming $\gamma = 13$ is correct and D_1 undetermined, $\eta = \frac{2\gamma^2}{3*4\beta} = \frac{2*169}{3*0.32} = 352$, and $D_1 = \frac{0.375}{352} = 0.0011$ Gy. With η=352, $\kappa = 2\text{x}10^{-5}$ and $\lambda_r = 3.0\text{x}10^{-4}$, the SSC model (section 4.5) fits the Löbrich [25] data no better than the SS model. However, whereas the rejoining time constant of a *single* active DSB is 14 hours in the SSC model, in the SS model ($\kappa = 7.2\text{x}10^{-8}$ and $\eta = 1$) the half-life is on the order of 6 months! Compartmentalized nuclei thus provide: 1) a single track transition dose consistent with aberration data, and 2) a realistic rejoining time constant for a cell with a single active DSB.

4.7 Cancer Risks and Misrepairs

To the extent that radiation causes cancer through translocation events, two arguments are provided that cancer risks are linear in misrepairs, i.e. that cancer risks follow the same dose-response (within a scaling constant) as dicentric chromosomal aberrations. The first argument only applies for doses less than that required to provide one tissue averaged misrepair (less than 2 Gray) and is analogous to the single track effects argument; namely, for low LET irradiation, decrements in $M(\infty) < 2$ are attained through commensurate decrements in the fraction of target cells having exactly two misrepairs (a negligible number of cells will have more than two) and thus commensurate reductions in cancer risk (cancer is assumed here to be a single cell phenomenon).

The second argument is based on cancer incidence being a smooth function of age and thus a smooth function of the lifetime accumulated permanent genetic damage as quantified by some measure $\Delta(t)$ of the total genetic

deviation an individual's cancer related genes have from normality. Since $M(\infty)$ adds to $\Delta(t)$ and cancer risk $R(\Delta(t))$ is smooth, the excess risk induced by an exposure should be $\frac{dR(\Delta(t))}{d\Delta(t)}M(\infty)$ for sufficiently small $M(\infty)$, see Crump et al. [12]. How large $M(\infty)$ can be while maintaining this first order approximation is not known and will depend on the individual, the individual's age, and the cancer type. Since cancer incidence increases with age, extending this concept to a second order approximation should introduce a positive quadratic term. Cancer risks at doses greater than two Gray are therefore predicted to have a larger quadratic component than the misrepair dose-response.

4.8 Overdispersed Initial Conditions

At high doses the probability mass of the initial active DSB distribution must traverse large portions of the SMBE Markov graph before reaching the m-axis. During its travels, the distribution is likely to gain system attributes as it loses initial condition attributes. The high dose SMBE prediction that $M(\infty) = \mu - \frac{1}{2}$ is therefore expected to be insensitive to the initial distribution shape. To test this hypothesis, overdispersed distributions were formed by inserting zero probability "spacers" between the Poisson probabilities. Indeed, the left panel of figure 4.2 shows that $M(\infty) = \mu - 0.5$ is approached at high doses regardless of the initial distribution shape.

At low doses the initial distribution shape is expected to influence $M(\infty)$. The right panel of figure 4.2 shows that $M(\infty)$ becomes a linear function of μ if the dispersion of the initial distribution is greater than or equal to two, quadratic if the dispresion equals one, and threshold-like if the dispersion is less than one (this situation may not be realizable). Thus, for high LET (overdispersed initial conditions) the SMBE model loses its quadratic traits and instead becomes linear-linear (LL) with the first linearity having zero intercept and the second linearity taking the form $M(\infty) = \mu - \frac{\eta}{2}$.

4.9 Summary

The main point of this chapter is to show that although the SS model of chapter 3 was selected based on high dose PFGE data, it is also consistent (when extended to the SMBE model) with aberration data at much lower doses. Furthermore, the SMBE model predicts a linearization of the misrepair dose response for $\mu > \frac{\eta}{2} = 176$, i.e. doses greater than $\frac{176}{13} = 13.5$ Gray. This linearization results from a saturation of the expected number of accidental binary eurepairs at $\frac{\eta}{2}$.

The intended usage of the SS model is with high dose PFGE time course measurements of $U(t)$ and $M(t)$. The SMBE model, on the other hand, is meant for chromosomal aberration derived measurements of $M(\infty)$. Note that the SS model uses $\gamma = pG$ and κ to match the PFGE *time* course of $M(t)$ at 80 Gray [λ_r is used strictly to match $U(t)$] while the SMBE model uses γ and η to match only $M(\infty)$, but at many different *doses* between 0.05 and 8 Gray. The parameter κ serves only to scale the time axis with respect to misrepair formation and does not affect $M(\infty)$.

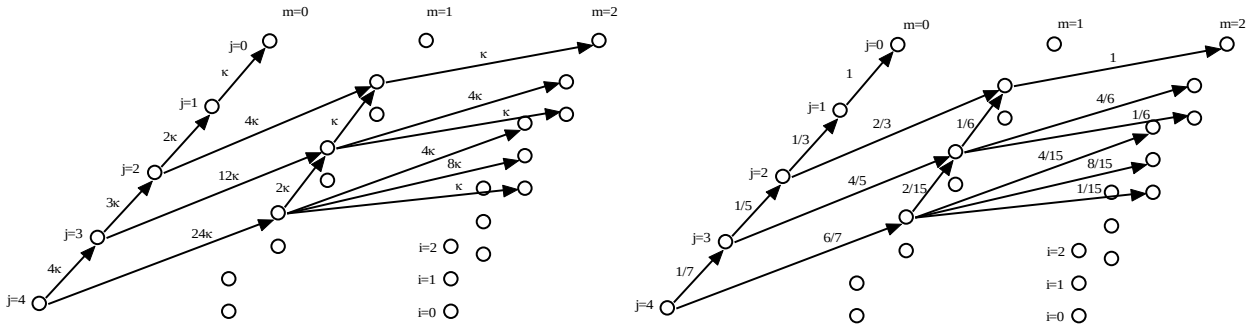

Figure 4.1: Left panel: The SMBE model is shown as a three dimensional Markov graph in which the node (j, i, m) contains the probability that a cell at time t has j active DSBs, i active DSBs destined for misrepair, and m misrepairs. The rate constants of probability mass transfer between nodes label the arrows. At $t = 0$ the j-axis is set to an initial distribution of active DSBs. As time evolves, the probability mass migrates toward the m-axis. Note that the state $(j, i, m) = (0, 0, 1)$ is not reachable since if two ends misrepair, their correct mates must also eventually misrepair. Right panel: In the steady state algorithm, each rate constant is replaced by the original rate constant divided by the sum of the rate constants leaving the node.

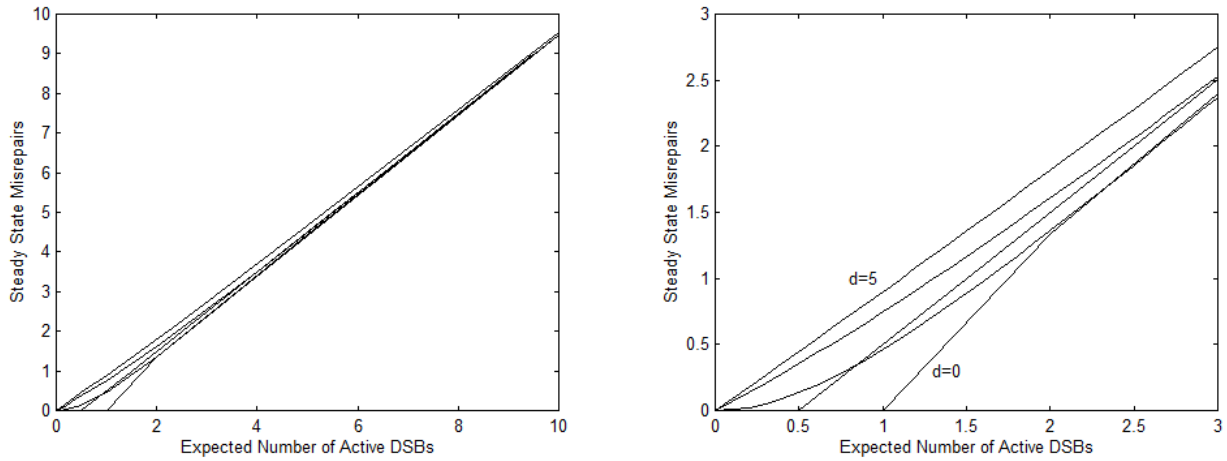

Figure 4.2: Left panel: The SMBE curve approaches the high dose asymptote $M(\infty) = \mu - 0.5$ for active DSB dispersions (the variance divided by the mean) of 5, 2, 1, or 0. Right panel: As the dispersion decreases from five to one to zero, the low dose region changes qualitatively from linear to quadratic to threshold-like. Dispersions of one, two and five were created by adding zero, one, or four, respectively, zero probability "spacers" between the non-zero Poisson probabilities. Zero dispersion implies all of the probability mass lies at $j = \mu$. The line $\mu - 0.5$ is also plotted.

Chapter 5

Induced DNA Fragment Distributions

This chapter provides derivations of two *induced* DNA fragment distribution models corresponding to two basic techniques by which DNA fragment distributions are measured. The first technique uses a uniform DNA label such as ethidium bromide (EtBr) or full radiolabeling. Uniform labeling techniques are described by one theoretical model [9], referred to here as the EtBr model. The second technique uses sequence specific probes (Southern blots) and requires a different model [10] which will be referred to as the specific probe model. The fundamental difference between these techniques is that EtBr detects how many fragments are generated of size x regardless of DNA sequence while specific probes detect only those fragments of size x which contain the hybridization sequence. There is also a difference in measurement intensity because uniform labels generate a signal proportional to the fragment size whereas a probe produces the same amount of signal regardless of the fragment size. The EtBr and specific probe models can be used with PFGE data to obtain estimates of λ [22, 8, 7, 6], the number of DSBs per base pair (bp) induced just after an acute dose.

The EtBr model [9] and the specific probe model [10] are derived, first using an indirect approach, and then using a direct approach [23]. The indirect approach finds the fragment density (the model) as a Poisson weighted sum of the fragment densities expected when a known number of DSBs, j, has struck the chromosome. The direct approach arrives at the final equations directly from the assumption that the cells have a Poisson distributed number of DSBs spread uniformly over the chromosome.

Induced DNA fragment distribution models describe three different types of fragments: edge fragments (subscript at) have one chemically *active* end and one *telomeric* end; internal fragments (subscript aa) have two chemically *active* ends, i.e. both ends were generated by DSBs; and delta fragments (subscript δ) are the original chromosomes, named after the Dirac delta function that describes their density of sizes. The functions $n_{at}(x)$, $n_{aa}(x)$, and $n_\delta(x)$ represent the expected *number* density of fragments of size x within a cell just after an acute dose.[1] Later we shall see that *partially-rejoined* fragment models have two additional fragment density functions, one for ring fragments, $n_r(x,t)$, and the other for misrejoined fragments in which both ends are telomeric, $n_{tt}(x,t)$.[2]

[1] In the EtBr model, the measured PFGE signal is proportional to the fragment *mass* density $m(x) = xn(x)$ where $n(x) = n_{at}(x) + n_{aa}(x) + n_\delta(x)$. Note that $n(x)\Delta x$ is the expected number of DNA fragments with sizes between $x - \Delta x/2$ and $x + \Delta x/2$.

[2] In the next chapter, all of the fragment densities will be functions of time.

5.1 The EtBr Model

5.1.1 Fixed Numbers of DSBs

Assume that a cell has one chromosome of size S and j DSBs striking it. Index the DNA base pairs (bp) as 1 through S and let each base pair incur a DSB with equal probability $\frac{1}{S}$.[3] The probability that the first fragment, the one containing base pair 1, is of size x, given that there are j DSBs, is written as $P(x_1 = x|j)$. This event occurs when one DSB strikes at the point x and the remaining $j-1$ DSBs strike beyond the point x. There are j ways to choose the DSB which lands at the point x. The remaining $j-1$ DSBs will hit outside of x with probability $(\frac{S-x}{S})^{j-1} = (1-\frac{x}{S})^{j-1}$ so that

$$P(x_1 = x|j) = j\left(\frac{1}{S}\right)\left(1-\frac{x}{S}\right)^{j-1}. \tag{5.1}$$

By symmetry, the $(j+1)$th fragment has the same probability of being size x as the first fragment. Thus the edge contribution to the fragment density is

$$n_{at}(x|j) = 2j\left(\frac{1}{S}\right)\left(1-\frac{x}{S}\right)^{j-1}. \tag{5.2}$$

The density of internal fragments equals the probability of forming an internal fragment at a particular chromosomal site, i.e. $\frac{j}{S}\frac{j-1}{S}(\frac{S-x}{S})^{j-2}$, multiplied by $S-x$, the number of potential fragment sites on the chromosome. Note that $\frac{j}{S}\frac{j-1}{S}$ is the probability of having DSBs strike at the two specific points on the chromosome and $\left(\frac{S-x}{S}\right)^{j-2} = \left(1-\frac{x}{S}\right)^{j-2}$ is the probability that the remaining $j-2$ DSBs strike outside of these two points. The internal fragment density is thus

$$\begin{aligned} n_{aa}(x|j) &= \frac{j(j-1)}{S^2}\left(1-\frac{x}{S}\right)^{j-2}(S-x) &(5.3)\\ &= \frac{j(j-1)}{S}\left(1-\frac{x}{S}\right)^{j-1} &(5.4)\end{aligned}$$

and, for $j > 0$, the total fragment density is

$$\begin{aligned} n(x|j) \;=\; n_{at}(x|j) + n_{aa}(x|j) &= \frac{(j-1)j}{S}\left(1-\frac{x}{S}\right)^{j-1} + \frac{2j}{S}\left(1-\frac{x}{S}\right)^{j-1} &(5.5)\\ &= \frac{(j+1)j}{S}\left(1-\frac{x}{S}\right)^{j-1} \qquad j>0. &(5.6)\end{aligned}$$

The left panel of figure 5.1 shows $n(x|j)$ for $j = 1, 2, 3$, and 4.

5.1.2 Poisson Distributed DSBs

A Poisson process $P(j, \lambda t) = \frac{(\lambda t)^j e^{-\lambda t}}{j!}$ is typically viewed as describing the probability that j events occurred within time t, given that the events have an exponentially distributed inter-arrival time τ with mean $\frac{1}{\lambda}$, i.e. $P(\tau) = \lambda e^{-\lambda\tau}$. If we now think of base pairs as ticks of a clock, the assumption that a chromosome of size S has a poisson distributed number of DSBs, $P(j, \lambda S) = \frac{(\lambda S)^j e^{-\lambda S}}{j!}$, with mean $\mu = \lambda S$, implies that the distances between DSBs, the fragment sizes x, are exponentially distributed with mean $\frac{1}{\lambda}$ base pairs, i.e.

$$P(x) = \lambda e^{-\lambda x} u_0^S(x) + e^{-\lambda S}\delta(S-x) \tag{5.7}$$

[3] A base pair x is defined as incurring a DSB if the DNA backbone is broken between base pair x and base pair $x+1$ on the indexing strand concurrent with a neighboring break on the opposite strand.

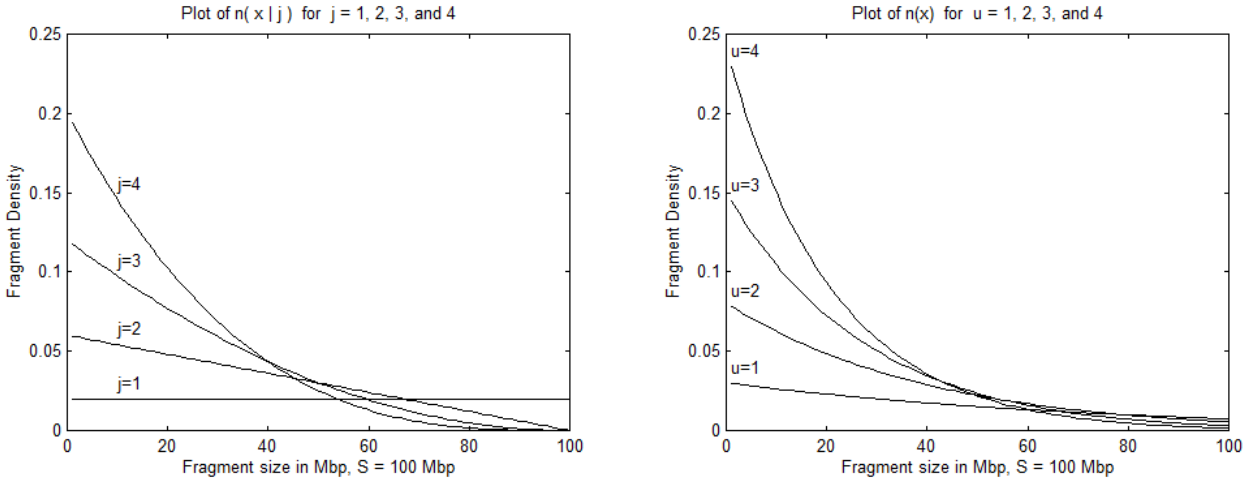

Figure 5.1: Plots of the fragment number density $n(x|j)$ for $j = 1, 2, 3, 4$ and $n(x)$ for $\mu = 1, 2, 3, 4$. Left panel: As j (the number of DSBs) increases, the fragment density shifts toward smaller fragment sizes. Note that the integral over x equals $j+1$, the number of fragments generated by j DSBs. As fragments rejoin, the density will dynamically shift toward higher x while simultaneously decreasing the area under the curve. Right panel: When the number of DSBs is Poisson, rather than known, the density is formed as a weighted sum of the curves in the left panel, the weights being the Poisson probabilities that j DSBs occured given that we expect μ. At low doses ($\mu < 1$), $n(x)$ is the weighted sum of mostly a delta function ($j = 0$), partly a square wave ($j = 1$), and a small amount of triangle wave ($j = 2$). As the dose increases, smaller weights are given to these j curves and larger weights to higher order polynomials with density focused closer to $x = 0$. A Dirac delta function (not shown) exists at $x = S$.

where $\delta(S-x)$ is a Dirac delta function centered at $x = S$ and $u_0^S(x)$ equals one if $0 \le x \le S$ and zero otherwise.[4] If S is expressed in base pairs (bp), λ represents the expected number of DSBs per bp or, equivalently, the probability that a DSB occurs at a particular base pair on the chromosome.

Assuming a Poisson distributed number of DSBs, $P(j,\mu) = \frac{\mu^j e^{-\mu}}{j!}$, the density of fragments of size x becomes

$$\begin{aligned} n(x) &= \sum_{j=0}^{\infty} n(x|j)P(j,\mu) = \sum_{j=1}^{\infty}[n_{at}(x|j) + n_{aa}(x|j)]P(j,\mu) + e^{-\mu}\delta(x-S) \\ &= n_{at}(x) + n_{aa}(x) + n_{\delta}(x) \end{aligned} \tag{5.8}$$

where the delta function arises because $j = 0$ forces $x = S$.

The contribution from edge fragments is found using equation 5.2 as

$$\begin{aligned} n_{at}(x) &= \sum_{j=1}^{\infty} n_{at}(x|j)P(j,\mu) = \sum_{j=1}^{\infty} 2j\left(\frac{1}{S}\right)\left(1-\frac{x}{S}\right)^{j-1}\frac{\mu^j e^{-\mu}}{j!} \\ &= 2\frac{\mu e^{-\mu}}{S}\sum_{j=1}^{\infty}\left(1-\frac{x}{S}\right)^{j-1}\frac{\mu^{j-1}}{(j-1)!} = 2\frac{\mu e^{-\mu}}{S}e^{\left(1-\frac{x}{S}\right)\mu} \\ &= 2\frac{\mu e^{-\mu\frac{x}{S}}}{S} = 2\lambda e^{-\lambda x}. \end{aligned} \tag{5.9}$$

[4] Since $\int_S^{\infty} \lambda e^{-\lambda x}dx = e^{-\lambda S}$, the probability mass of the exponential's tail is contained in the delta function at $x = S$.

Similarly, using equation 5.4, the contribution from internal fragments is

$$\begin{aligned} n_{aa}(x) &= \sum_{j=1}^{\infty} n_{aa}(x|j)P(j,\mu) = \sum_{j=1}^{\infty} j(j-1)\left(\frac{1}{S}\right)\left(1-\frac{x}{S}\right)^{j-1}\frac{\mu^j e^{-\mu}}{j!} \\ &= \sum_{j=2}^{\infty} j(j-1)\left(\frac{1}{S}\right)\left(1-\frac{x}{S}\right)^{j-1}\frac{\mu^j e^{-\mu}}{j!} \\ &= \left(1-\frac{x}{S}\right)\frac{\mu^2 e^{-\mu}}{S}\sum_{j=2}^{\infty}\left(1-\frac{x}{S}\right)^{j-2}\frac{\mu^{j-2}}{(j-2)!} \\ &= \left(1-\frac{x}{S}\right)\frac{\mu^2 e^{-\mu}}{S}e^{(1-\frac{x}{S})\mu} = \left(1-\frac{x}{S}\right)\frac{\mu^2 e^{-\mu\frac{x}{S}}}{S} \\ &= \lambda^2(S-x)e^{-\lambda x}. \end{aligned} \tag{5.10}$$

The total fragment *number* density is thus

$$\begin{aligned} n(x) &= n_{at}(x) + n_{aa}(x) + n_{\delta}(x) \\ &= 2\lambda e^{-\lambda x} + \lambda^2(S-x)e^{-\lambda x} + e^{-\lambda x}\delta(x-S) \end{aligned} \tag{5.11}$$

as plotted in the right panel of figure 5.1. The fragment *mass* density[5]

$$m(x) = xn(x) = x[2\lambda e^{-\lambda x} + \lambda^2(S-x)e^{-\lambda x} + e^{-\lambda x}\delta(x-S)] \tag{5.12}$$

is compared to $n(x)$ in figure 5.2.

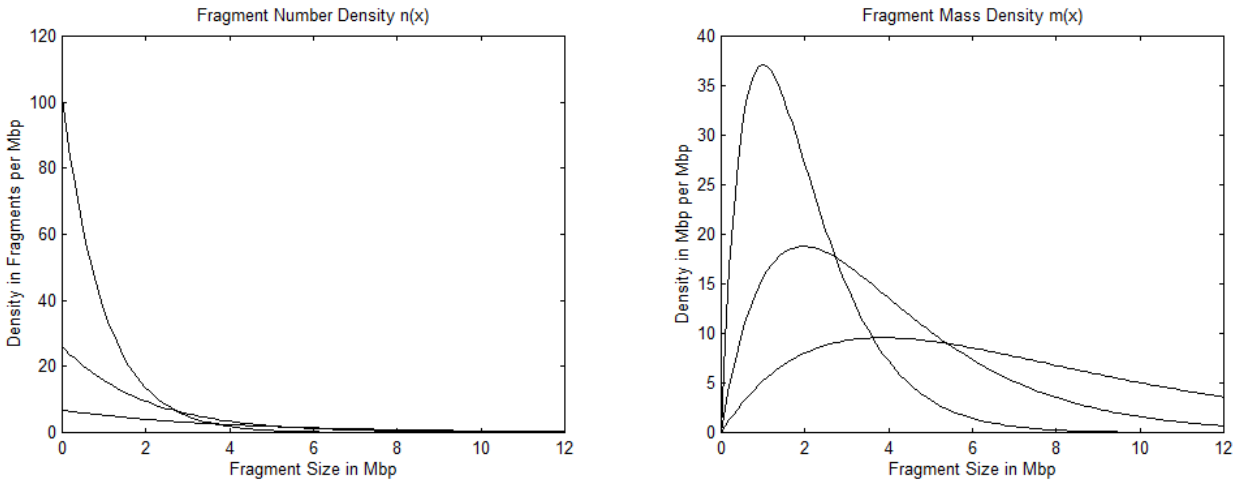

Figure 5.2: The fragment number density $n(x)$ and the fragment mass density $m(x) = xn(x)$ are shown in the left and right panels, respectively. Note that the area under the $m(x)$ curves equals the chromosome size S and that the area under the $n(x)$ curves equals the expected number of fragments $\lambda S + 1$. Both plots were made with $S = 100$ Mbp and $\lambda = 0.25$, 0.5, and 1 DSB/Mbp.

Since EtBr produces a signal proportional to fragment length (mass), the measured PFGE signal intensity at gel positions corresponding to x should in theory be proportional to $m(x)\frac{dx}{dz}$, where z is the position along the gel and $\frac{dx}{dz}$ is the slope of the calibration curve for molecular length versus distance migrated. In practice, thermal diffusion causes band broadening, a process represented mathematically as the convolution of $m(x)$ with Gauss [16]

[5]Mass is measured in base-pairs for convenience. The function $m(x)$ was previously presented [9] as $I(x)$ for "intensity."

or Lorentz [21] distributions. The resulting intensity is then fit to PFGE data by least squares to obtain estimates of a thermal dispersion parameter and λ.[6] A model of FAR (section 2.4) can be obtained from $m(x)$ as

$$\begin{aligned} \text{FAR} \quad = \quad Q(X,\lambda) \quad &= \quad \frac{\int_0^X m(x)dx}{\int_0^\infty m(x)dx} \quad = \quad \frac{1}{S}\int_0^X m(x)dx \\ &= \quad 1 - e^{-\lambda X}\left(1 + \lambda\left(X - \frac{X^2}{S}\right)\right) \end{aligned} \tag{5.13}$$

where X is the cut-off fragment size for entry into the gel [3, 5, 8]. A plot of $Q(X,\lambda)$ as a function of λ is shown in figure 5.3 for cut-off sizes $X = 6$ and 12 Mbp.

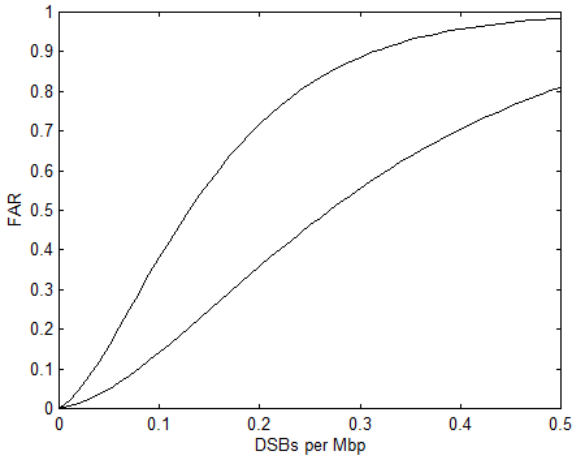

Figure 5.3: A plot of FAR=$Q(X,\lambda)$ as a function of λ for $S = 100$ Mbp and cut-off sizes $X = 6$ Mbp (lower curve) and $X = 12$ Mbp (upper curve).

5.2 The Specific Probe Model

If the DNA fragments are hybridized with a sequence specific probe unique to one site on the chromosome, a new model must be formed to describe the measured fragment distribution [10]. Approximate the hybridization sequence as a single point h and choose the base pair ordering such that $h < S/2$. Let x be the fragment size and y the starting base pair position of the fragment $\mathbf{x}$. The fragment $\mathbf{x}$ of size x can be either an internal fragment, in which case its boundaries are defined by DSBs, or an edge fragment, in which case one of the chromosomal ends serves as a delimiter. The ability of the fragment to utilize the chromosomal ends as a delimiter depends on the location of the hybridization point h, the fragment size x, and the chromosome size S. The number of potential chromosomal sites of size x that contain h is called the shift factor and also depends on h, x and S. As a result, three regions of x values must be considered separately, see figure 5.4.

If $0 < x < h$, $\mathbf{x}$ must be an internal fragment, see figure 5.4. The probability that one of the j DSBs strikes at y is $\frac{j}{S}$. The probability that a second DSB (out of the remaining $j-1$ DSBs) strikes at $y+x$ equals $\frac{j-1}{S}$. Finally, the probability that the remaining $j-2$ DSBs land outside of the fragment $\mathbf{x}$, is $(1-\frac{x}{S})^{j-2}$. The product of these probabilities equals the probability of forming a fragment of size x starting at location y. In contrast to the EtBr

[6] Some researchers have ignored thermal diffusion [8, 22, 7, 6].

case, the fragment **x** is no longer free to be positioned anywhere across the chromosome. Instead, the location of the hybridization point h is limited to one of the x base pair locations of the fragment **x**. The shift factor is therefore x rather than $S - x$ as in the EtBr model, see equation 5.4. The density of labeled fragments of size x less than h is thus

$$n(x|j) = x\frac{j(j-1)}{S^2}\left(1-\frac{x}{S}\right)^{j-2} \quad \text{for} \quad 0 < x < h. \tag{5.14}$$

This can be substituted into

$$n(x) = \sum_{j=0}^{\infty} n(x|j)P(j,\mu) = \sum_{j=1}^{\infty} n(x|j)P(j,\mu) \quad \text{for} \quad 0 < x < h \tag{5.15}$$

to yield

$$\begin{aligned} n(x) &= \sum_{j=1}^{\infty} \frac{j(j-1)x}{S^2}\left(1-\frac{x}{S}\right)^{j-2}\frac{\mu^j e^{-\mu}}{j!} = \frac{\mu^2 x e^{-\mu}}{S^2}\sum_{j=2}^{\infty} j(j-1)\left(1-\frac{x}{S}\right)^{j-2}\frac{\mu^{j-2}}{j!} \\ &= \frac{\mu^2 x e^{-\mu}}{S^2}\sum_{j=2}^{\infty}\left(1-\frac{x}{S}\right)^{j-2}\frac{\mu^{j-2}}{(j-2)!} = \frac{\mu^2 x e^{-\mu}}{S^2}e^{(1-\frac{x}{S})\mu} = \frac{\mu^2 x e^{-\mu\frac{x}{S}}}{S^2} \\ &= x\lambda^2 e^{-\lambda x} \quad \text{for} \quad 0 < x < h. \end{aligned} \tag{5.16}$$

If $h < x < S - h$, the fragment **x** is large enough to be the first fragment, but not large enough to be the last, see figure 5.4. If **x** is the first fragment,

$$n_{at}(x|j) = \frac{j}{S}\left(1-\frac{x}{S}\right)^{j-1} \quad \text{for} \quad h < x < S - h \tag{5.17}$$

which implies

$$\begin{aligned} n_{at}(x) &= \sum_{j=1}^{\infty}\frac{j}{S}\left(1-\frac{x}{S}\right)^{j-1}\frac{\mu^j e^{-\mu}}{j!} = \frac{\mu e^{-\mu}}{S}\sum_{j=1}^{\infty}\left(1-\frac{x}{S}\right)^{j-1}\frac{\mu^{j-1}}{(j-1)!} \\ &= \frac{\mu e^{-\mu}}{S}e^{(1-\frac{x}{S})\mu} = \frac{\mu e^{-\mu\frac{x}{S}}}{S} = \lambda e^{-\lambda x} \quad \text{for} \quad h < x < S - h. \end{aligned} \tag{5.18}$$

If **x** is not the first fragment (the case when the left edge of the chromosome is not used by the fragment), the starting point y of the fragment **x** can be anywhere in the interval $(1, h)$. Thus, rather than multiplying by a shift factor of x

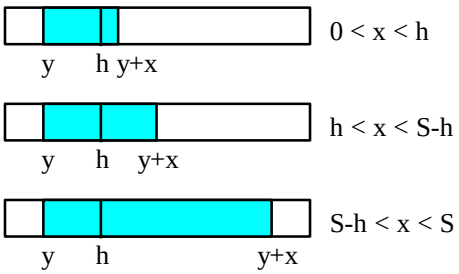

Figure 5.4: The specific probe model involves three distinct fragment size regions. The shaded area represents the fragment **x**. In the top bar, the hybridization point h can be on any of the x base pairs of the fragment; the fragment **x** cannot be an edge fragment. In the middle bar, the point y can be on any of the first h bases of the chromosome; the fragment **x** can, but need not, be the first fragment. In the bottom bar, the point $y + x$ can be on any of the $S - x$ base pairs to the right of base pair x; the fragment **x** can be either the first or the last fragment.

as in equation 5.14, we multiply by the shift factor h.[7] The derivation continues as in equation 5.16 with h replacing x so that

$$n(x) \quad = \quad n_{at}(x) + n_{aa}(x) \; = \; \lambda e^{-\lambda x} + h\lambda^2 e^{-\lambda x}, \quad \text{for} \quad h < x < S - h. \tag{5.19}$$

Finally, when $(S - h) < x \leq S$, $\mathbf{x}$ can use either the left edge, the right edge, neither edge, or both edges (if $x = S$) of the chromosome. By symmetry, the edge cases have identical probabilites, equal to the probability derived in equation 5.18, i.e. $n_{at}(x) = 2\lambda e^{-\lambda x}$. If $\mathbf{x}$ is an internal fragment, the shift factor becomes $S - x$ and $n_{aa}(x) = (S - x)\lambda^2 e^{-\lambda x}$. The total density is then

$$n(x) = 2\lambda e^{-\lambda x} + (S - x)\lambda^2 e^{-\lambda x} + e^{-\lambda x}\delta(x - S) \qquad \text{for} \quad S - h < x \leq S. \tag{5.20}$$

This equation is equivalent to the EtBr model since, when $x > (S - h)$, the fragment $\mathbf{x}$ is guaranteed to be tagged with a probe.

It is interesting to note that the fragment shifts are constrained by either the chromosome size S (for $x > S - h$), the probe site h (for $h < x < S - h$), or the fragment size x (for $0 < x < h$). The specific probe model is summarized in in table 5.1 and compared to the EtBr model in figure 5.5. Observe that as h approaches zero the central region of the specific probe model expands and the density becomes exponential.

Table 5.1: A summary of the specific probe model.

x value	$n_{at}(x)$	shift factor	$n_{aa}(x)$	$n(x)$
$x < h$	0	x	$x\lambda^2 e^{-\lambda x}$	$x\lambda^2 e^{-\lambda x}$
$h < x < S - h$	$\lambda e^{-\lambda x}$	h	$h\lambda^2 e^{-\lambda x}$	$\lambda e^{-\lambda x} + h\lambda^2 e^{-\lambda x}$
$S - h < x < S$	$2\lambda e^{-\lambda x}$	$S - x$	$(S - x)\lambda^2 e^{-\lambda x}$	$2\lambda e^{-\lambda x} + (S - x)\lambda^2 e^{-\lambda x}$

5.3 The EtBr Model (Direct Approach)

Since λ is the probability of incurring a DSB at a particular site and $e^{-\lambda x}$ is the probability that no DSBs occur on a particular DNA stretch of length x, each edge has a probability $\lambda e^{-\lambda x}$ of being size x, i.e. $n_{at}(x) = 2\lambda e^{-\lambda x}$. For internal fragments, $\lambda^2 e^{-\lambda x}$ is the probability that two DSBs occur at two specific sites while the x base pairs in between remain untouched. As there are $S - x$ chromosomal sites from which an internal can arise, $n_{aa}(x) = \lambda^2 e^{-\lambda x}(S - x)$. The total density can thus be written as

$$n(x) = e^{-\lambda x}\lambda + (S - x)\lambda e^{-\lambda x}\lambda + \lambda e^{-\lambda x} + e^{-\lambda S}\delta(x - S) \tag{5.21}$$

where λ represents a DSB, $e^{-\lambda x}$ an unbroken DNA stretch of length x, and $S - x$ the number of potential internal fragments. This direct approach was explained quite clearly by Lehmann and Ormerod [23] and more recently by Kraxenberger et al. [21]. It allows one to write the EtBr model by inspection.

[7] The condition $x < S - h$ is used here because otherwise there would not be enough "room" on the right hand side of the chromosome to allow all of the h fragment shifts, i.e. the right end of the fragment would hit the right end of the chromosome (the point S) before the starting point y reached the hybridization point h (see figure 5.4).

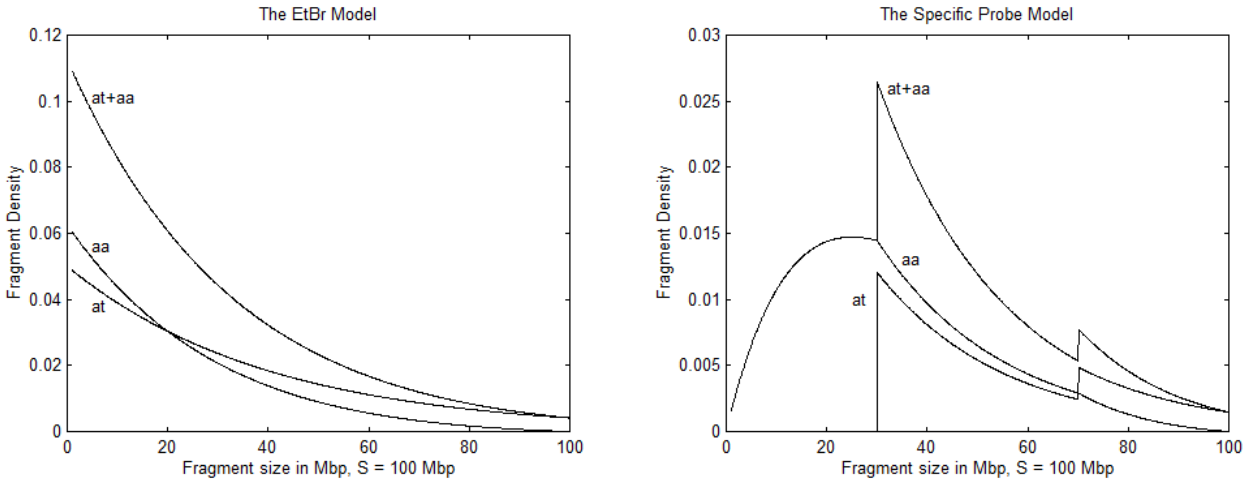

Figure 5.5: The edge (at) and internal (aa) fragment densities of the EtBr model and specific probe model are compared in these plots. Left panel: In the EtBr model, the edge density is exponential on $[0, S]$ and the internal fragment density subexponential ($S = 100$ Mbp and $\lambda = 0.025$ DSB/Mbp). Right panel: In the specific probe model, the edge density is "saw-tooth" exponential while the density of internals has three behaviors: modal on $[0, h]$, exponential on $[h, S-h]$, and subexponential on $[S-h, S]$ ($S = 100$ Mbp, $h = 30$ Mbp and $\lambda = 0.04$ DSB/Mbp).

5.4 The Specific Probe Model (Direct Approach)

The direct approach can also be used to derive the specific probe model as summarized in table 5.1. The term $\lambda^2 e^{-\lambda x}$ is the probability of creating an internal fragment of size x at a particular site. The shift factor is the number of potential sites for such a fragment, subject to the constraint that the fragment includes the hybdridization point. When $x < h$ the shift factor is x, since there are x points on the fragment which can contain h. If $h < x < S - h$, the shift factor becomes h because the starting point of the fragment $\mathbf{x}$ must be one of the first h base pairs of S. Finally, if $S - h < x < S$, we have a shift factor of $S - x$, just like in the EtBr model, since in this case, the fragment must contain the hybridization point h. The edge density includes $\lambda e^{-\lambda x}$ for each chromosomal end (telomere) capable of acting as the edge of fragment $\mathbf{x}$. Figure 5.4 can thus be used to derive the specific probe model directly.

5.5 Pool Sizes and Conditional Probabilities

The expected number of fragments in each of the induced pools can be obtained through integration over x.[8]

$$\begin{aligned} N_{at}(0) &= \int_0^\infty n_{at}(x,0)dx = \int_0^\infty 2u_0^S(x)\lambda e^{-\lambda x}dx = 2(1-e^{-\lambda S}) \qquad (5.22) \\ N_{aa}(0) &= \int_0^\infty n_{aa}(x,0)dx = \int_0^\infty \lambda(S-x)u_0^S(x)\lambda e^{-\lambda x}dx \\ &= \lambda S(1-e^{-\lambda S}) - \int_0^S \lambda x \lambda e^{-\lambda x}dx \end{aligned}$$

[8]The argument for $t = 0$ is included in order to be consistent with later chapters. The EtBr model is assumed hereafter.

$$
\begin{aligned}
&= \lambda S(1-e^{-\lambda S}) + \lambda^2 \frac{d}{d\lambda}\int_0^S e^{-\lambda x}dx \quad = \quad \lambda S(1-e^{-\lambda S}) + \lambda^2 \frac{d}{d\lambda}\frac{1-e^{-\lambda S}}{\lambda} \\
&= \lambda S(1-e^{-\lambda S}) + \lambda^2 \frac{\lambda S e^{-\lambda S} - (1-e^{-\lambda S})}{\lambda^2} \\
&= \lambda S(1-e^{-\lambda S}) + \lambda S e^{-\lambda S} - (1-e^{-\lambda S}) \quad = \quad \lambda S - (1-e^{-\lambda S}) \qquad (5.23)
\end{aligned}
$$

$$
N_\delta(0) = \int_0^\infty n_\delta(x,0)dx \quad = \quad \int_0^\infty \delta(x-S)e^{-\lambda x}dx \quad = \quad e^{-\lambda S} \qquad (5.24)
$$

$$
\begin{aligned}
N(0) &= N_{at}(0) + N_{aa}(0) + N_\delta(0) \quad = \quad 2(1-e^{-\lambda S}) + \lambda S - (1-e^{-\lambda S}) + e^{-\lambda S} \\
&= \lambda S + 1, \qquad (5.25)
\end{aligned}
$$

where $u_0^S(x)$ equals one if $0 < x < S$ and zero otherwise. Since we expect λS DSBs we expect $\lambda S + 1$ fragments. The conditional probability that a fragment has size x, given that it is in a specific pool,[9] is equal to the fragment density divided by the total number of fragments expected in that pool, i.e.

$$
p_{at}(x,0) = \frac{2u_0^S(x)\lambda e^{-\lambda x}}{2(1-e^{-\lambda S})} \qquad (5.26)
$$

$$
p_{aa}(x,0) = \frac{\lambda(S-x)u_0^S(x)\lambda e^{-\lambda x}}{\lambda S - (1-e^{-\lambda S})} \qquad (5.27)
$$

$$
p_\delta(x,0) = \frac{\lambda(S-x)u_0^S(x)\lambda e^{-\lambda x}}{\delta}(x-S). \qquad (5.28)
$$

If there are N chromosomes, the total density is

$$
\begin{aligned}
n(x) &= n_{at}(x) + n_{aa}(x) + n_\delta(x) \\
&= \sum_{j=1}^N 2u_0^{S_j}(x)\lambda e^{-\lambda x} + \sum_{j=1}^N \lambda(S_j - x)u_0^{S_j}(x)\lambda e^{-\lambda x} + \sum_{j=1}^N \delta(x-S_j)e^{-\lambda S_j} \qquad (5.29)
\end{aligned}
$$

and the expected pool sizes become

$$
N_{at}(0) = 2\sum_{j=1}^N (1-e^{-\lambda S_j}) \qquad (5.30)
$$

$$
N_{aa}(0) = \sum_{j=1}^N \lambda S_j - (1-e^{-\lambda S_j}) \quad = \quad \lambda S_T - \sum_{j=1}^N (1-e^{-\lambda S_j}) \qquad (5.31)
$$

$$
N_\delta(0) = \sum_{j=1}^N e^{-\lambda S_j} \qquad (5.32)
$$

$$
\begin{aligned}
N(0) &= 2\sum_{j=1}^N (1-e^{-\lambda S_j}) + \lambda S_T - \sum_{j=1}^N (1-e^{-\lambda S_j}) + \sum_{j=1}^N e^{-\lambda S_j} \\
&= \lambda S_T + N, \qquad (5.33)
\end{aligned}
$$

where $S_T = \sum_{j=0}^N S_j$. The conditional probabilities are then

$$
p_{at}(x,0) = \frac{\sum_{j=1}^N 2u_0^{S_j}(x)\lambda e^{-\lambda x}}{2\sum_{j=1}^N (1-e^{-\lambda S_j})} \qquad (5.34)
$$

$$
p_{aa}(x,0) = \frac{\sum_{j=1}^N \lambda(S_j - x)u_0^{S_j}(x)\lambda e^{-\lambda x}}{\lambda S_T - \sum_{j=1}^N (1-e^{-\lambda S_j})} \qquad (5.35)
$$

$$
p_\delta(x,0) = \frac{\sum_{j=1}^N \delta(S_j - x)e^{-\lambda S_j}}{\sum_{j=1}^N e^{-\lambda S_j}}. \qquad (5.36)
$$

[9] The conditionalities are denoted by subscripts.

The fragment densities have thus been decomposed into a shape factor $p_c(x,0)$ and an amplitude factor $N_c(0)$, i.e. $n_c(x,0) = N_c(0)p_c(x,0)$ where $c \in \{at, aa, \delta\}$. Note that the expected number of unhit chromosomes is $\sum_{j=1}^{N} e^{-\lambda S_j}$ and that the expected number of hit chromosomes is $\sum_{j=1}^{N}(1-e^{-\lambda S_j})$. Equation 5.30 thus states that there are twice as many edges as hit chromosomes and equation 5.31 states that after one DSB is consumed for each chromosome hit, the remaining DSBs create internal fragments in a one-to-one fashion.

5.6 Summary

Two existing DNA fragment distribution models, the EtBr model and the specific probe model, were derived in this chapter using direct and indirect methods. The models can be used to obtain distribution-based (least squares) estimates of DSB levels from PFGE data [8], but only in cells with negligible amounts of misrepair (immediately after the exposure); the models do not consider the products of misrejoining reactions. The next chapter provides a dynamical description of misrejoining DNA fragment distributions.

Chapter 6

Misrejoined DNA Fragment Distributions

Although DNA fragment models exist (chapter 5) for cells with negligible amounts of misrepair, no models have yet been developed to describe the DNA fragment distribution of cells with substantial amounts of misrepair. This chapter attempts to fill this void by introducing the Sax subset EtBr (SSEB) model, an extension of the EtBr model to include the DSB misrepair dynamics of the Sax subset (SS) model (section 3.6).

6.1 The Sax Subset EtBr (SSEB) Model

To derive the SSEB model, fragments will be defined in terms of active DSBs; inactive DSBs are viewed as markers that keep the DNA intact unless measurement is made by PFGE. The fragment distribution defined with complete disregard to inactive DSBs is *non-measurable* during repair times for which significant levels of inactive DSBs exist. This *non-measurable* fragment distribution consists of five components corresponding to five types of DNA fragments: 1) $n_{at}(x,t)$ – the density of "edge" fragments in which one end is active and the other telomeric, 2) $n_{aa}(x,t)$ – the density of "internal" fragments in which both ends are active, 3) $n_{tt}(x,t)$ – the density of "bitelomeric" fragments in which both ends of a misrejoined fragment are telomeric, 4) $n_r(x,t)$ – the density of ring fragments, and 5) $n_\delta(x,t)$ – the density of chromosomes that have never been activated. These non-measurable fragment densities represent the "internal state" of the cell's DNA fragments. The SSEB model describes the dynamics of this internal state. A means of converting this state into a measurable PFGE output (at arbitrary repair times) is provided in section 6.2.

Just after an acute dose, all of the fragment density lies within the edge, internal, and delta fragment pools. As time evolves, DNA leaves the edge and internal pools and moves into the ring and bitelomere pools, see figure 6.1. Edge fragments can merge with other edge fragments to form bitelomeric fragments or they can rejoin with internal fragments to form bigger edge fragments. Internal fragments can rejoin with other internals to form larger internals or react with themselves to form rings. The delta pool remains constant since inactive DSBs can only be repaired and accidental binary eurepair is not modeled. At moderate times post-irradiation ($t > 4$ hrs), nearly all inactive DSBs will have been repaired and the total fragment density thus becomes measurable. When the system reaches

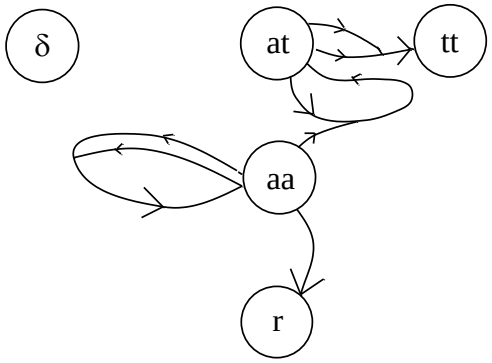

Figure 6.1: The SSEB model, depicted here as a fragment flow graph, is based on five pools of non-measurable fragments. The pools interact through misrejoining events (converging arrows) and ring closures.

steady state all active DSBs will have found partners, hence, the edge and internal fragment pools will be empty, the eurepaired fragments all lie in the delta pool, and the misrepaired fragments will either be ring fragments or bitelomeric fragments. The dynamics that take place between the initial distribution and the final distribution are described mathematically by the SSEB model as:

$$\begin{aligned}
\frac{\partial n_{at}(x,t)}{\partial t} &= -\kappa n_{at}(x,t)N_{at}(t) \quad - \quad 2\kappa n_{at}(x,t)N_{aa}(t) \\
&\quad +2\kappa \int_0^x n_{at}(x',t)n_{aa}(x-x',t)dx' && (6.1)\\
\frac{\partial n_{aa}(x,t)}{\partial t} &= -4\kappa n_{aa}(x,t)N_{aa}(t) \quad - \quad 2\kappa n_{aa}(x,t)N_{at}(t) \\
&\quad +2\kappa \int_0^x n_{aa}(x',t)n_{aa}(x-x',t)dx' && (6.2)\\
\frac{\partial n_{tt}(x,t)}{\partial t} &= \frac{\kappa}{2} \int_0^x n_{at}(x',t)n_{at}(x-x',t)dx' && (6.3)\\
\frac{\partial n_{r}(x,t)}{\partial t} &= \kappa n_{aa}(x,t), && (6.4)\\
n_{at}(x,0) &= 2\lambda e^{-\lambda x} \sum_{j=1}^{N} u_0^{S_j}(x) && (6.5)\\
n_{aa}(x,0) &= \lambda^2 e^{-\lambda x} \sum_{j=1}^{N} (S_j - x) u_0^{S_j}(x) && (6.6)\\
n_{\delta}(x,0) &= \sum_{j=1}^{N} \delta(x-S_j) e^{-\lambda S_j} && (6.7)
\end{aligned}$$

$$\frac{\partial n_{\delta}(x,t)}{\partial t} = 0, \qquad n_r(x,0) = 0, \qquad n_{tt}(x,0) = 0, \tag{6.8}$$

where κ is the binary misrepair rate constant for active DSB free ends, $N_c(t) = \int_0^\infty n_c(x,t)dx$ is the total number of fragments in pool c ($c \in \{at, aa, tt, r, \delta\}$), λ is the induced level of active DSBs per basepair, S_j is the size of the jth chromosome, $u_0^{S_j}(x)$ equals one on the interval $[0, S_j]$ and zero elsewhere, $\delta(x - S_j)$ is a Dirac delta function centered at $x = S_j$ and N is the total number of chromosomes within the cell.

The terms of the SSEB model fall into two categories: source terms and target terms. The source terms are

negative because they withdraw fragment density to supply the target terms which are positive. The source terms include pool size factors because an edge or internal of size x can react with any other edge or internal. The target terms are convolution integrals[1] that include all possible fragment interactions leading to a misrejoined fragment of size x. Specific explanations for each of the terms are now given. Keep in mind that fragments are defined with respect to *active* DSBs.

Equation 6.1 describes the rate of size x edge fragment formation. The first and second terms are the rates at which density leaves $n_{at}(x,t)$ due to misrejoining reactions. It is assumed that an edge of size x can react with any other edge or internal with equal probability except that internals are twice as likely since they have two active ends. The third term describes how density moves into $n_{at}(x,t)$ through misrejoining events between all possible combinations of fragments of size x' and $x-x'$. This convolution term, when integrated over all x, is equal in magnitude to the second term, i.e.

$$\begin{aligned}
\int_0^\infty 2\kappa \int_0^x n_{at}(x',t)n_{aa}(x-x',t)dx'dx &= \int_0^\infty 2\kappa \int_0^\infty n_{at}(x',t)n_{aa}(x-x',t)dx'dx \\
&= 2\kappa \int_0^\infty n_{at}(x',t) \int_0^\infty n_{aa}(x-x',t)dxdx' \\
&= 2\kappa \int_0^\infty n_{at}(x',t) \int_{-x'}^\infty n_{aa}(l,t)dldx' \\
&= 2\kappa \int_0^\infty n_{at}(x',t) \int_0^\infty n_{aa}(l,t)dldx' \\
&= 2\kappa \int_0^\infty n_{at}(x',t)dx' \int_0^\infty n_{aa}(l,t)dl \\
&= 2\kappa N_{at}(t)N_{aa}(t). \qquad (6.9)
\end{aligned}$$

This result is intuitive because density lost from $n_{at}(x,t)$ by reaction with an internal should be regained by some other edge at some other x.

Equation 6.2 describes the rate at which internal fragments of size x are formed. The first two terms are the rates at which fragment density leaves $n_{aa}(x,t)$ due to misrejoining events. Because internal fragments have two active ends, the self-reaction rate is multpilied by a factor of 4 and the cross-reaction rate by a factor of 2. The third term is a auto-convolution integral that represents the formation of $n_{aa}(x,t)$ through misrejoining events between two smaller internal fragments. The convolution integral is multiplied by a factor of two, rather than four, because fragment interactions on the interval $x' < \frac{x}{2}$ are identical with those on the interval $x' > \frac{x}{2}$. The total flux (integral over x) of the first term equals twice the total flux of the third term since the rejoining of *two* internal fragments produces *one* larger internal fragment. The loss of internal fragments due to ring closure is neglected because the SS model, and thus the SSEB model, assumes that doses are high enough that linear terms can be neglected when adding to quadratic terms.

Equation 6.3 states that bitelomeric fragments are synthesized by misrejoining edges. The synthesis term is an auto-convolution and therefore includes a factor of $\frac{1}{2}$. Note that the total flux of this term equals one-half the total flux of the first term in equation 6.1 since *two* misrejoining edges form *one* larger bitelomere.

In equation 6.4, rings are synthesized at a rate equal to κ times the density of internals. There is no change in size so there are no interactions with densities at other x values. Note that the SSEB model does not conserve mass since the flux entering the ring pool is neglected as a flux leaving the internal pool.

[1] Rings are an exception to this and other general statements.

The intial conditions are those of the EtBr model since the subset of basepairs not in association with core subsets will have Poisson distributed DSBs with mean λS_T. Note that the delta pool remains constant over time so $n_\delta(x,t) = n_\delta(x,0)$.

To relate the SSEB model to the SS model, let us define $\lambda_i(t)$ as the number of *inactive* DSBs per bp and $\lambda_a(t)$ as the number of *active* DSBs per bp (note that $\lambda_a(0) = \lambda$ of the SSEB model initial conditions). If we divide the SS model solutions (see section 3.6) by the total DNA length S_T and define $\lambda_0 \equiv \frac{U(0)}{S_T}$ we obtain $\lambda_i(t) = (1-p)\lambda_0 e^{-\lambda_r t}$ and $\lambda_a(t) = \frac{p\lambda_0}{U_2(0)\kappa t+1}$, where p is the fraction of induced DSBs that are active, λ_r is the linear repair rate constant for inactive DSBs, κ is the misrepair constant for active DSB free ends, $U(0)$ is the expected number of induced DSBs, and $U_2(0) = 2p\lambda_0 S_T$.

6.2 Measurable SSEB Density Functions

As t becomes large, the fragment density eventually becomes part of either $n_\delta(x,\infty)$, $n_{tt}(x,\infty)$, or $n_r(x,\infty)$. Since inactive DSBs vanish at large times, the process of measurement does not change these final distributions and, aside from the effects of thermal diffusion [21, 16], the bitelomeric fragment density is measurable on a PFGE gel (rings larger than approximately 50 kbp are not expected to enter the gel [24]).

At intermediate times t, inactive DSBs still exist so a mathematical mapping is needed to convert non-measurable distributions into PFGE measurable eurepair and misrepair distributions. This mapping parallels the physical process of digesting the protein "splints" of inactive DSBs at the time of PFGE measurement. Mathematically, all fragments defined by active DSB ends will instantaneously be converted into subfragment distributions as determined by the induced EtBr model of equation 5.11. In this case, x acts as the chromosome size, $\lambda_i(t)$ the "induced" level of DSBs per bp, and x' a particular subfragment size. Since it is the x''s that will actually be measured on the gel, the distributions must be integrated over x to get the total density at x' at time t. We thus obtain the density measured at time t at position x' as

$$n(x',t) = \int_{x'}^{\infty} n_T(x,t)\left(\lambda_i^2 e^{-\lambda_i x'}(x-x') + 2\lambda_i e^{-\lambda_i x'} + \delta(x-x')e^{-\lambda_i x}\right)dx \tag{6.10}$$

where $n_T(x,t) = n_{at}(x,t) + n_{aa}(x,t) + n_{tt}(x,t) + [1-e^{-\lambda_i(t)x}]n_r(x,t) + n_\delta(x,t)$ and $[1-e^{-\lambda_i(t)x}]$ is the probability that a size x ring has at least one inactive DSB. Although the Dirac function is under the integral in equation 6.10, one must keep in mind that $n_\delta(x,t)$ is a train of Dirac delta functions

$$n_\delta(x,t) = \sum_{j=1}^{N} n_\delta(S_j,t)\delta(x-S_j). \tag{6.11}$$

Thus one of the cross products will be a train of delta functions squared, which when integrated over x, produces delta functions themselves. If we pull the terms involving delta functions outside of the integral, the measurable density becomes

$$\begin{aligned} n(x',t) &= \int_{x'}^{\infty} (n_T(x,t) - n_\delta(x,t))\left(\lambda_i^2(t)e^{-\lambda_i(t)x'}(x-x') + 2\lambda_i(t)e^{-\lambda_i(t)x'}\right)dx \\ &+ (n_T(x',t) - n_\delta(x',t))e^{-\lambda_i(t)x'} \\ &+ \sum_{j=1}^{N} n_\delta(S_j,t)[\left(\lambda_i^2(t)e^{-\lambda_i(t)x'}(S_j-x') + 2\lambda_i(t)e^{-\lambda_i(t)x'}\right)u_0^{S_j}(x') \\ &\quad + e^{-\lambda_i(t)x'}\delta(S_j-x')]. \end{aligned} \tag{6.12}$$

This density is the sum of a misrepair component (the first two terms) and a eurepair component (the summation over j).

6.3 SSEB Pool Size Equations and Solutions

If the SSEB model is integrated over x, the following ordinary differential equations (ODEs) result:

$$\frac{dN_{at}(t)}{dt} = -\kappa N_{at}^2(t) \tag{6.13}$$

$$\frac{dN_{aa}(t)}{dt} = -2\kappa N_{aa}^2(t) - 2\kappa N_{at}(t)N_{aa}(t) \tag{6.14}$$

$$\frac{dN_{tt}(x,t)}{dt} = \frac{\kappa}{2}N_{at}^2(t) \tag{6.15}$$

$$\frac{dN_r(t)}{dt} = \kappa N_{aa}(t). \tag{6.16}$$

The first of these equations integrates to

$$N_{at}(t) = \frac{N_{at}(0)}{N_{at}(0)\kappa t + 1}. \tag{6.17}$$

Letting $V(t) = \frac{1}{N_{aa}(t)}$ and noting that $\kappa \int_0^t N_{at}(\tau)d\tau = \ln(N_{at}(0)\kappa t + 1)$, the second equation integrates to

$$\frac{dN_{aa}(t)}{dt} = -2\kappa N_{aa}^2(t) - 2\kappa N_{at}(t)N_{aa}(t) \tag{6.18}$$

$$\frac{dV(t)}{dt} - 2\kappa N_{at}(t)V(t) = 2\kappa \tag{6.19}$$

$$\frac{d}{dt}(V(t)e^{-2\kappa\int_0^t N_{at}(\tau)d\tau}) = 2\kappa e^{-2\kappa\int_0^t N_{at}(\tau)d\tau} \tag{6.20}$$

$$\frac{d}{dt}\left(\frac{V(t)}{(N_{at}(0)\kappa t+1)^2}\right) = 2\kappa\frac{1}{(N_{at}(0)\kappa t+1)^2} \tag{6.21}$$

$$\frac{V(t)}{(N_{at}(0)\kappa t+1)^2} - V(0) = \frac{2}{N_{at}(0)}(1-(N_{at}(0)\kappa\tau+1)^{-1}) \tag{6.22}$$

$$V(t) = (N_{at}(0)\kappa t+1)\left(V(0)(N_{at}(0)\kappa t+1) + \frac{2}{N_{at}(0)}(N_{at}(0)\kappa\tau+1-1)\right)$$

$$N_{aa}(t) = \frac{\frac{N_{aa}(0)}{(N_{at}(0)\kappa t+1)}}{(N_{at}(0)\kappa t+1) + \frac{2N_{aa}(0)}{N_{at}(0)}N_{at}(0)\kappa t} \tag{6.23}$$

$$= \frac{\frac{N_{aa}(0)}{N_{at}(0)\kappa t+1)}}{(N_{at}(0)\kappa t+1)+2N_{aa}(0)\kappa t} = \frac{\frac{N_{aa}(0)}{(N_{at}(0)\kappa t+1)}}{(U_2(0)\kappa t+1)}. \tag{6.24}$$

Thus $U_2(t) = 2N_{aa}(t) + N_{at}(t)$ becomes

$$U_2(t) = \frac{\frac{2N_{aa}(0)}{(N_{at}(0)\kappa t+1)}}{(U_2(0)\kappa t+1)} + \frac{N_{at}(0)}{N_{at}(0)\kappa t+1} \tag{6.25}$$

$$= \frac{\frac{2N_{aa}(0)}{(N_{at}(0)\kappa t+1)} + (U_2(0)\kappa t+1)\frac{N_{at}(0)}{N_{at}(0)\kappa t+1}}{(U_2(0)\kappa t+1)}$$

$$= \frac{\frac{U_2(0)+U_2(0)N_{at}(0)\kappa t}{N_{at}(0)\kappa t+1}}{(U_2(0)\kappa t+1)} = \frac{U_2(0)}{(U_2(0)\kappa t+1)}. \tag{6.26}$$

As expected, this is the solution to $\frac{dU_2(t)}{dt} = -\kappa U_2^2(t)$, the equation that results from adding equation 6.13 to twice equation 6.14.[2]

The solution for $N_{aa}(t)$ can be partial fractioned as,

$$\begin{aligned} N_{aa}(t) &= \frac{N_{aa}(0)}{[N_{at}(0)\kappa t+1][U_2(0)\kappa t+1]} \\ &= \frac{\frac{1}{2}U_2(0)}{U_2(0)\kappa t+1} - \frac{\frac{1}{2}N_{at}(0)}{N_{at}(0)\kappa t+1} \quad = \quad \frac{1}{2}U_2(t) - \frac{1}{2}N_{at}(t), \end{aligned} \tag{6.27}$$

which restates that

$$U_2(t) = 2N_{aa}(t) + N_{at}(t),$$

i.e. the integration used to obtain $N_{aa}(t)$ was not necessary.

The bitelomeric fragment pool size is solved as

$$N_{tt}(t) = \frac{\kappa}{2}\int_0^t N_{at}^2(\tau)d\tau \tag{6.28}$$

$$\begin{aligned} &= \frac{\kappa}{2}\int_0^t \frac{N_{at}^2(0)}{(N_{at}\kappa\tau+1)^2}d\tau \quad = \quad \frac{N_{at}(0)}{2}\left(1 - \frac{1}{N_{at}(0)\kappa t+1}\right) \\ &= \frac{N_{at}(0)}{2}\frac{N_{at}(0)\kappa t}{(N_{at}(0)\kappa t+1)} \end{aligned} \tag{6.29}$$

$$N_{tt}(\infty) = \frac{N_{at}(0)}{2}. \tag{6.30}$$

Since edges continue to grab internals until they grab another edge, the final number of bitelomeric fragments equals half the initial number of edges.

The ring pool size $N_r(t)$ is obtained by integrating the partial fractioned form of $N_{aa}(t)$ shown in equation 6.27.

$$\begin{aligned} N_r(t) &= \kappa\int_0^t N_{aa}(\tau)d\tau \\ &= \kappa\frac{1}{2\kappa}\ln\left(\frac{U_2(0)\kappa t+1}{N_{at}(0)\kappa t+1}\right) \quad = \quad \frac{1}{2}\ln\left(\frac{U_2(0)\kappa t+1}{N_{at}(0)\kappa t+1}\right) \end{aligned} \tag{6.31}$$

$$N_r(\infty) = \frac{1}{2}\ln\left(\frac{U_2(0)}{N_{at}(0)}\right) \quad = \quad \frac{1}{2}\ln\left(1+\frac{2N_{aa}(0)}{N_{at}(0)}\right). \tag{6.32}$$

The qualitative form of equation 6.32 is reasonable because $N_r(\infty)$ decreases with increasing $N_{at}(0)$ and edge fragments are expected to pull internals away from ring formation. Cells with many chromosomes (or notI-somes) are therefore expected to allow the formation of only a few rings since such cells, when irradiated, contain many edges. For male human cells in G_0 (left panel of figure 6.2), $N_{at}(0) = 2\sum_{j=1}^{N}(1-e^{-\lambda S_j})$ and $N_{aa}(0) = \lambda S_T - \sum_{j=1}^{N}(1-e^{-\lambda S_j})$ were calculated to form $N_r(\infty) = \frac{1}{2}\ln\left(1+\frac{2N_{aa}(0)}{N_{at}(0)}\right)$ and $N_{tt}(\infty) = \frac{N_{at}(0)}{2}$ as functions of dose, see table 6.1 and figure 6.2.[3] Although the results predict that rings are negligible, it must be kept in mind that the model ingores proximity effects and thus may underestimate the true number of rings.

[2]The sum of the three misrepair flux terms is $2\kappa N_{at}(t)N_{aa}(t) + 2\kappa N_{aa}^2(t) + \frac{\kappa}{2}N_{at}^2(t) = \frac{\kappa}{2}[2N_{aa}(t)+N_{at}(t)]^2 = \frac{\kappa}{2}U_2^2(t) = \frac{dM(t)}{dt}$.

[3]The initial pool sizes were computed with $\lambda = p\lambda_0$ where $p = 0.33$ and $\lambda_0 = \frac{40}{6349}D$.

Table 6.1: The ratio $N_r(\infty)/N_{tt}(\infty)$ is small for a wide range of doses. Note that almost every human chromosome is hit at least once for doses greater than 10 Gray.

Dose in Gray	0.1	1	10	100	1000	10000
$N_{at}(0)$	2.5955	22.4231	82.3920	91.9998	92.0000	92.0000
$N_{aa}(0)$	0.022	1.988	90.799	1273.957	13153.571	131949.7
$N_r(\infty)$	0.0085	0.0816	0.5822	1.6784	2.8296	3.9809
$N_{tt}(\infty)$	1.2977	11.2116	41.1960	45.9999	46.0000	46.0000
$N_r(\infty)/N_{tt}(\infty)$	0.0065	0.0073	0.0141	0.0365	0.0615	0.0865

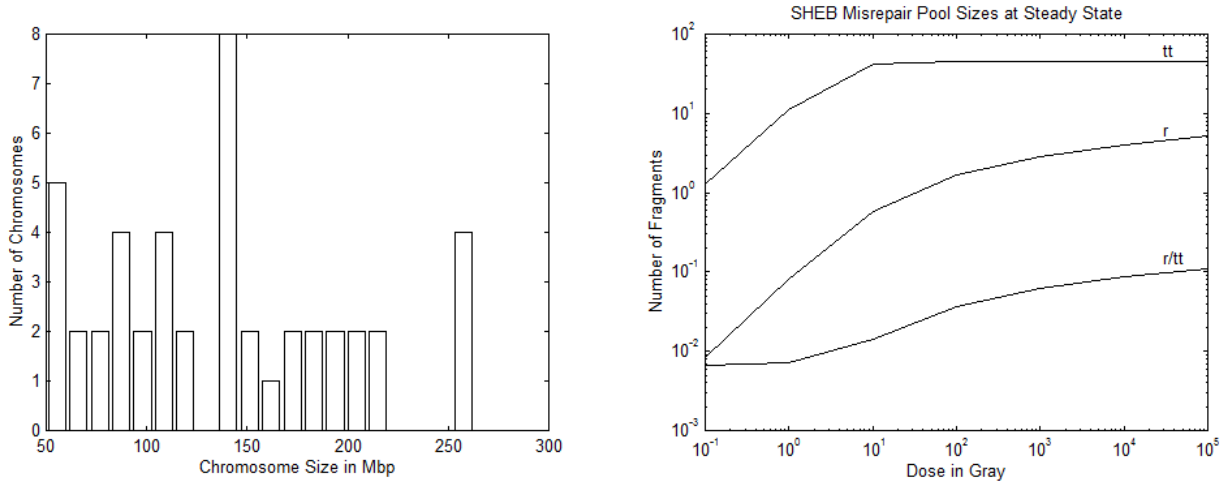

Figure 6.2: Left panel: A histogram of $n_\delta(x,0)$ for human (male) cells resting in G_0. Right panel: The steady state misrepair pool sizes and their ratio are plotted as functions of dose (see table 6.1).

6.4 The SSEB Model in Fourier Coordinates

The Fourier transform [33] of the SSEB model is[4]

$$\frac{\partial \tilde{n}_{at}(k,t)}{\partial t} = -\kappa \tilde{n}_{at}(k,t)U_2(t) + 2\kappa \tilde{n}_{at}(k,t)\tilde{n}_{aa}(k,t) \tag{6.33}$$

$$\frac{\partial \tilde{n}_{aa}(k,t)}{\partial t} = -2\kappa \tilde{n}_{aa}(k,t)U_2(t) + 2\kappa \tilde{n}^2_{aa}(k,t) \tag{6.34}$$

$$\frac{\partial \tilde{n}_{tt}(k,t)}{\partial t} = \frac{\kappa}{2}\tilde{n}^2_{at}(k,t) \tag{6.35}$$

$$\frac{\partial \tilde{n}_{r}(k,t)}{\partial t} = \kappa \tilde{n}_{aa}(k,t). \tag{6.36}$$

Since equation 6.34 is uncoupled from the others in k-space, we solve it first. Let $\tilde{v}(k,t) = \frac{1}{\tilde{n}_{aa}(k,t)}$, and note that $\kappa \int_0^t U_2(\tau)d\tau = \ln(U_2(0)\kappa t + 1)$, the second equation then becomes

$$-\frac{\dot{\tilde{v}}(k,t)}{\tilde{v}^2(k,t)} = -\frac{2\kappa U_2(t)}{\tilde{v}(k,t)} + \frac{2\kappa}{\tilde{v}^2(k,t)} \tag{6.37}$$

[4]The Fourier transform of $n_c(x,t)$ is defined as $\tilde{n}_c(k,t) = \int_0^\infty n_c(x,t)e^{-ikx}dx$ where $i = \sqrt{-1}$.

$$\dot{\tilde{v}}(k,t) = 2\kappa U_2(t)\tilde{v}(k,t) - 2\kappa \tag{6.38}$$

$$\frac{\partial}{\partial t}[\tilde{v}(k,t)e^{-2\kappa\int_0^t U_2(\tau)d\tau}] = -2\kappa e^{-2\kappa\int_0^t U_2(\tau)d\tau} \tag{6.39}$$

$$\frac{\partial}{\partial t}[\tilde{v}(k,t)(U_2(0)\kappa t+1)^{-2} = -2\kappa(U_2(0)\kappa t+1)^{-2} \tag{6.40}$$

$$\tilde{v}(k,t)(U_2(0)\kappa t+1)^{-2} - \tilde{v}(k,0) = \frac{2}{U_2(0)}[(U_2(0)\kappa t+1)^{-1}-1] \tag{6.41}$$

$$\begin{aligned}\tilde{v}(k,t) &= (U_2(0)\kappa t+1)\left(\tilde{v}(k,0)(U_2(0)\kappa t+1) + \frac{2}{U_2(0)}[1-(U_2(0)\kappa t+1)]\right)\\ &= (U_2(0)\kappa t+1)\left(\tilde{v}(k,0)(U_2(0)\kappa t+1) - 2\kappa t\right)\end{aligned} \tag{6.42}$$

$$\begin{aligned}\tilde{n}_{aa}(k,t) &= \frac{\frac{\tilde{n}_{aa}(k,0)}{U_2(0)\kappa t+1}}{U_2(0)\kappa t+1-\tilde{n}_{aa}(k,0)2\kappa t}\\ &= \frac{\tilde{n}_{aa}(k,0)}{[U_2(0)\kappa t+1][(U_2(0)-2\tilde{n}_{aa}(k,0))\kappa t+1]}\\ &= \frac{\frac{1}{2}U_2(0)}{U_2(0)\kappa t+1} - \frac{\frac{1}{2}(U_2(0)-2\tilde{n}_{aa}(k,0))}{[U_2(0)-2\tilde{n}_{aa}(k,0)]\kappa t+1}.\end{aligned} \tag{6.43}$$

This expression for the internal fragment density can be integrated to form the ring density

$$\tilde{n}_r(k,t) = \kappa\int_0^t \tilde{n}_{aa}(k,\tau)d\tau = \frac{1}{2}\ln\left(\frac{U_2(0)\kappa t+1}{[U_2(0)-2\tilde{n}_{aa}(k,0)]\kappa t+1}\right) \tag{6.44}$$

$$\tilde{n}_r(k,\infty) = \frac{1}{2}\ln\left(\frac{U_2(0)}{U_2(0)-2\tilde{n}_{aa}(k,0)}\right). \tag{6.45}$$

The edge density is thus

$$\begin{aligned}\frac{\partial\tilde{n}_{at}(k,t)}{\partial t} &= [2\kappa\tilde{n}_{aa}(k,t)-\kappa U_2(t)]\tilde{n}_{at}(k,t)\\ \frac{\partial\tilde{n}_{at}(k,t)}{\partial t} - [2\kappa\tilde{n}_{aa}(k,t)-\kappa U_2(t)]\tilde{n}_{at}(k,t) &= 0\\ \frac{\partial}{\partial t}(\tilde{n}_{at}(k,t)e^{-\int_0^t[2\kappa\tilde{n}_{aa}(k,\tau)-\kappa U_2(\tau)]d\tau}) &= 0\\ \tilde{n}_{at}(k,t)e^{-\int_0^t[2\kappa\tilde{n}_{aa}(k,\tau)-\kappa U_2(\tau)]d\tau} &= \tilde{n}_{at}(k,0)\end{aligned}$$

$$\begin{aligned}\tilde{n}_{at}(k,t) &= \tilde{n}_{at}(k,0)e^{\int_0^t[2\kappa\tilde{n}_{aa}(k,\tau)-\kappa U_2(\tau)]d\tau}\\ &= \frac{\tilde{n}_{at}(k,0)}{U_2(0)\kappa t+1}e^{\int_0^t 2\kappa\tilde{n}_{aa}(k,\tau)d\tau}\\ &= \frac{\tilde{n}_{at}(k,0)}{U_2(0)\kappa t+1}\frac{U_2(0)\kappa t+1}{[U_2(0)-2\tilde{n}_{at}(k,0)]\kappa t+1}\\ &= \frac{\tilde{n}_{at}(k,0)}{[U_2(0)-2\tilde{n}_{aa}(k,0)]\kappa t+1}.\end{aligned} \tag{6.46}$$

This solution can be squared and integrated to obtain the bitelomeric density

$$\begin{aligned}\frac{\partial\tilde{n}_{tt}(k,t)}{\partial t} &= \frac{\kappa}{2}\frac{\tilde{n}_{at}^2(k,0)}{[(U_2(0)-2\tilde{n}_{aa}(k,0))\kappa t+1]^2}\\ \tilde{n}_{tt}(k,t) &= \frac{1}{2}\frac{\tilde{n}_{at}^2(k,0)}{U_2(0)-2\tilde{n}_{aa}(k,0)}\left(1-\frac{1}{(U_2(0)-2\tilde{n}_{aa}(k,0))\kappa t+1}\right)\end{aligned} \tag{6.47}$$

$$
\begin{aligned}
&= \frac{1}{2}\frac{\tilde{n}_{at}^2(k,0)}{U_2(0)-2\tilde{n}_{aa}(k,0)}\left(\frac{(U_2(0)-2\tilde{n}_{aa}(k,0))\kappa t}{(U_2(0)-2\tilde{n}_{aa}(k,0))\kappa t+1}\right)\\
&= \frac{1}{2}\frac{\tilde{n}_{at}^2(k,0)\kappa t}{(U_2(0)-2\tilde{n}_{aa}(k,0))\kappa t+1} \qquad (6.48)\\
\tilde{n}_{tt}(k,\infty) &= \frac{1}{2}\frac{\tilde{n}_{at}^2(k,0)}{U_2(0)-2\tilde{n}_{aa}(k,0)}. \qquad (6.49)
\end{aligned}
$$

Note that, in contrast to the $n(x,t)$ dependence on $n(x',0)$ for all $x' \le x$, the $\tilde{n}(k,t)$'s depend on the $\tilde{n}(k,0)$'s only for the same value of k. The SSEB model in Fourier coordinates is thus decoupled in the sense that density at different k values do not interact. This is similar to the way exponentials diagonalize, and thus decouple, linear differential operators with constant coefficients.[5]

The Fourier transform of the initial edge density, with $\lambda = p\lambda_0$, is

$$
\begin{aligned}
\tilde{n}_{at}(k,0) = \int_0^\infty \sum_{j=1}^{N} 2e^{-ikx}u_0^{S_j}(x)\lambda e^{-\lambda x}dx &= 2\sum_{j=1}^{N}\int_0^{S_j} e^{-(ik+\lambda)x}\lambda dx\\
&= \frac{2\lambda}{\lambda+ik}\sum_{j=1}^{N}(1-e^{-(ik+\lambda)S_j}). \qquad (6.50)
\end{aligned}
$$

If the dose is large enough that each chromosome is hit at least once (the *weak* condition), the initial edges sizes become exponentially distributed since

$$
\sum_{j=1}^{N}(1-e^{-(ik+\lambda)S_j}) \approx N \qquad \Rightarrow \qquad \tilde{n}_{at}(k,0) = 2N\frac{\lambda}{\lambda+ik}. \qquad (6.51)
$$

The weak condition implies that there are $2N$ edge fragments and $\lambda S_T - N$ internal fragments at $t = 0$.

For internal fragments, we have

$$
\begin{aligned}
\tilde{n}_{aa}(k,0) &= \int_0^\infty \sum_{j=1}^{N}\lambda(S_j-x)e^{-ikx}u_0^{S_j}(x)\lambda e^{-\lambda x}dx\\
&= \lambda^2\sum_{j=1}^{N}\int_0^{S_j}(S_j-x)e^{-(ik+\lambda)x}dx\\
&= \lambda^2\sum_{j=1}^{N}\frac{S_j}{\lambda+ik}-\frac{1}{\lambda+ik}\int_0^{S_j}e^{-(ik+\lambda)x}dx\\
&= \lambda^2\sum_{j=1}^{N}\frac{S_j}{\lambda+ik}-\frac{1}{(\lambda+ik)^2}(1-e^{-(ik+\lambda)S_j})\\
&= \lambda^2\frac{S_T}{\lambda+ik}-\lambda^2\left(\frac{(N-\sum_{j=1}^{N}e^{-(ik+\lambda)S_j}}{(\lambda+ik)^2}\right). \qquad (6.52)
\end{aligned}
$$

[5]The linear operator $\sum_{j=0}^{n} a_j \frac{d^j}{dt^j}$ can be viewed as a weighted sum of matrices in Hilbert space where, for example $\frac{d}{dt}$ has ones below the main diagonal and minus ones above the main diagonal (higher order operators can be formed as powers of this matrix). This operator is coupled in the time domain because it has non-zero off-diagonal matrix elements. The Laplace transform of the linear operator is $\sum_{j=0}^{n} a_j s^j$. Since this transform multiplies $Y(s)$ point-wise for each s, it is a diagonal matrix operator in Hilbert space. Just as matrices are diagonal in the coordinates of their eigenvectors, the linear operator $\sum_{j=0}^{n} a_j \frac{d^j}{dt^j}$ is diagonalized (decoupled) in the coordinates of its eigenfunctions, the basis set of exponential functions.

Assuming the weak condition,

$$\tilde{n}_{aa}(k,0) = \frac{\lambda^2 S_T}{\lambda+ik} - \frac{N\lambda^2}{(\lambda+ik)^2}. \tag{6.53}$$

For internals to be exponentially distributed, the *strong* condition, $\lambda S_T - N \approx \lambda S_T$, is also needed so that

$$\begin{aligned}\tilde{n}_{aa}(k,0) &= \lambda^2\left(\frac{S_T}{\lambda+ik} - \frac{N}{(\lambda+ik)^2}\right) = \lambda^2\left(\frac{S_T(\lambda+ik)-N}{(\lambda+ik)^2}\right)\\ &\approx \lambda^2\left(\frac{S_T(\lambda+ik)}{(\lambda+ik)^2}\right) = \lambda S_T\frac{\lambda}{(\lambda+ik)}.\end{aligned} \tag{6.54}$$

Since the strong condition does not necessarily imply the weak condition,[6] the assumption of exponentially distributed internal fragments requires both the strong and weak conditions. The strong condition implies that $N_{aa}(0) = \lambda S_T$ while the weak condition implies that $N_{aa}(0) = \lambda S_T - N$. When both conditions hold, the strong condition implies that these differences in $N_{aa}(0)$ are negligible. Furthermore, since the strong condition implies $\lambda \gg N/S_T$, we can add N/S_T to λ and redefine it as the number of breaks (active DSBs plus pairs of natural telomeres) per base pair.

6.5 Exponential Initial Conditions

6.5.1 Internal Fragments

Suppose we have a hypothetical cell with no edge fragments and exponentially distributed internal fragments. The solution to the SSEB model with $U_2(t) = 2N_{aa}(t)$, $N_{at}(t) = 0$ and $\tilde{n}_{aa}(k,0) = N_{aa}(0)\frac{\lambda}{\lambda+ik}$ is

$$N_{aa}(t) = \frac{N_{aa}(0)}{[N_{at}(0)\kappa t+1][U_2(0)\kappa t+1]} = \frac{N_{aa}(0)}{[2N_{aa}(0)\kappa t+1]} \tag{6.55}$$

and

$$\begin{aligned}\tilde{n}_{aa}(k,t) &= \frac{\tilde{n}_{aa}(k,0)}{[U_2(0)\kappa t+1][(U_2(0)-2\tilde{n}_{aa}(k,0))\kappa t+1]}\\ &= \frac{N_{aa}(0)\frac{\lambda}{\lambda+ik}}{[2N_{aa}(0)\kappa t+1][(2N_{aa}(0)-2\tilde{n}_{aa}(k,0))\kappa t+1]}\\ &= N_{aa}(t)\frac{\frac{\lambda}{\lambda+ik}}{[(1-\frac{\lambda}{\lambda+ik})2N_{aa}(0)\kappa t+1]} = N_{aa}(t)\frac{\frac{\lambda}{\lambda+ik}}{[\frac{ik}{\lambda+ik}2N_{aa}(0)\kappa t+1]}\\ &= N_{aa}(t)\frac{\lambda}{2N_{aa}(0)\kappa tik+\lambda+ik}\\ &= N_{aa}(t)\frac{\lambda}{(2N_{aa}(0)\kappa t+1)ik+\lambda} = N_{aa}(t)\frac{\frac{\lambda}{2N_{aa}(0)\kappa t+1}}{ik+\frac{\lambda}{2N_{aa}(0)\kappa t+1}}\\ &= N_{aa}(t)\frac{\lambda(t)}{ik+\lambda(t)}\end{aligned} \tag{6.56}$$

$$n_{aa}(x,t) = N_{aa}(t)\lambda(t)e^{-\lambda(t)x}, \quad \lambda(t) \equiv \frac{\lambda}{2N_{aa}(0)\kappa t+1}, \tag{6.57}$$

where $\lambda = \frac{N_{aa}(0)}{S_T}$. Internal fragments thus remain exponential if they start exponential. The shape parameter $\lambda(t)$ and the amplitude factor $N_{aa}(t)$ decay with the same hyperbolic rate because every time a fragment disappears, a

[6] If one of the chromosomes is very small compared to the others, high doses may create many more DSBs than the total number of chromosomes (i.e. satisfy the strong condition) and still leave the small chromosome unhit (i.e. not satisfy the weak condition).

DSB disappears. Note that the mean fragment length $\frac{1}{\lambda(t)} = \frac{1}{\lambda}(2N_{aa}(0)\kappa t + 1)$ becomes infinite as $t \to \infty$. Since a fragment cannot be larger than S_T, the model is unrealistic at very large times.[7]

The distribution of ring fragments is

$$\begin{aligned}
\tilde{n}_r(k,t) &= \frac{1}{2}\ln\left(\frac{U_2(0)\kappa t + 1}{[U_2(0) - 2\tilde{n}_{aa}(k,0)]\kappa t + 1}\right) = \frac{1}{2}\ln\left(\frac{2N_{aa}(0)\kappa t + 1}{\frac{ik}{\lambda+ik}2N_{aa}(0)\kappa t + 1}\right) \\
&= \frac{1}{2}\ln\left(\frac{(2N_{aa}(0)\kappa t + 1)(\lambda + ik)}{ik2N_{aa}(0)\kappa t + \lambda + ik}\right) = \frac{1}{2}\ln\left(\frac{\lambda + ik}{ik + \frac{\lambda}{2N_{aa}(0)\kappa t+1}}\right) \\
&= \frac{1}{2}\ln\left(\frac{ik + \lambda}{ik + \lambda(t)}\right) \qquad (6.58) \\
n_r(x,t) &= \frac{1}{2x}(e^{-\lambda(t)x} - e^{-\lambda x}) \quad \text{and} \quad n_r(x,\infty) = \frac{1}{2x}(1 - e^{-\lambda x}). \qquad (6.59)
\end{aligned}$$

The shape of the ring distribution is focused toward small x because early times in the rejoining reaction are characterized by both high levels of internals (high rates of ring production) and short average fragment lengths.

6.5.2 Edge Fragments

Suppose the initial conditions are such that only exponentially distributed edge fragments exist. Setting $N_{aa}(0) = 0$, $U_2(0) = N_{at}(0)$, $\tilde{n}_{aa}(k,0) = 0$ and $\tilde{n}_{at}(k,0) = N_{at}(0)\frac{\lambda}{\lambda+ik}$ yields

$$\begin{aligned}
N_{at}(t) &= \frac{N_{at}(0)}{[N_{at}(0)\kappa t + 1]} \qquad (6.60) \\
\tilde{n}_{at}(k,t) &= \frac{\tilde{n}_{at}(k,0)}{[U_2(0) - 2\tilde{n}_{aa}(k,0)]\kappa t + 1} = \frac{\tilde{n}_{at}(k,0)}{N_{at}(0)\kappa t + 1} \\
&= \frac{N_{at}(0)\frac{\lambda}{\lambda+ik}}{N_{at}(0)\kappa t + 1} = N_{at}(t)\frac{\lambda}{\lambda + ik} \qquad (6.61) \\
n_{at}(x,t) &= N_{at}(t)\lambda e^{-\lambda x}, \qquad (6.62)
\end{aligned}$$

i.e. the edge density remains exponential. As the amplitude of the edge density decays to zero, the mean edge size remains constant. In the previous section, it was the feedback of internals to internals that caused the mean value of the internal distribution to increase with time. Here, we have no internals to map edges into larger edges, so the shape of the edge density does not change.

For bitelomeric fragments we have,

$$\begin{aligned}
N_{tt}(t) &= \frac{N_{at}(0)}{2}\left(1 - \frac{1}{N_{at}(0)\kappa t + 1]}\right) = \frac{N_{at}(0)}{2}\frac{N_{at}(0)\kappa t}{N_{at}(0)\kappa t + 1} \\
&= \frac{N_{at}(0)}{2}N_{at}(t)\kappa t \qquad (6.63) \\
\tilde{n}_{tt}(k,t) &= \frac{1}{2}\frac{\tilde{n}_{at}^2(k,0)\kappa t}{(U_2(0) - 2\tilde{n}_{aa}(k,0))\kappa t + 1} = \frac{1}{2}\frac{N_{at}^2(0)\kappa t}{N_{at}(0)\kappa t + 1}\frac{\lambda^2}{(\lambda + ik)^2} \\
&= N_{tt}(t)\frac{\lambda^2}{(\lambda + ik)^2} \qquad (6.64) \\
n_{tt}(x,t) &= N_{tt}(t)\lambda^2 x e^{-\lambda x} \quad \text{and} \quad n_{tt}(x,\infty) = \frac{N_{at}(0)}{2}\lambda^2 x e^{-\lambda x}. \qquad (6.65)
\end{aligned}$$

[7] This should not alarm the reader, since a cell without edge fragments is also unrealistic.

Bitelomeric fragments thus have a distribution that, like the edge fragment distribution, maintains a constant shape as the amplitude changes. The shape factor $p_{tt}(x,t) = \lambda^2 x e^{-\lambda x}$ is the probability density function for the sum of two exponentially distributed edges.[8] This results because, in the absence of internals, bitelomeres can only be formed through recombinations between two original edges. In the full model, an undetermined number of internal fragments can insert themselves between the edges so a different $p_{tt}(x,t)$ results.

6.5.3 Full Model (Edges and Internals)

The full SSEB model will now be solved for exponentially distributed initial fragment distributions (weak and strong conditions). The pool size solutions remain exactly as given in section 6.3, namely

$$N_{at}(t) = \frac{N_{at}(0)}{N_{at}(0)\kappa t + 1} \tag{6.66}$$

$$N_{aa}(t) = \frac{N_{aa}(0)}{[N_{at}(0)\kappa t + 1][U_2(0)\kappa t + 1]} \tag{6.67}$$

$$N_{tt}(t) = \frac{N_{at}(0)}{2}\left(1 - \frac{1}{N_{at}(0)\kappa t + 1}\right) = \frac{N_{at}(0)}{2} N_{at}(t)\kappa t \tag{6.68}$$

$$N_r(t) = \frac{1}{2}\ln\left(\frac{U_2(0)\kappa t + 1}{N_{at}(0)\kappa t + 1}\right) \tag{6.69}$$

The initial conditions $\tilde{n}_{at}(k,0) = N_{at}(0)\frac{\lambda}{ik+\lambda}$ and $\tilde{n}_{aa}(k,0) = N_{aa}(0)\frac{\lambda}{ik+\lambda}$ yield

$$\begin{aligned}\tilde{n}_{at}(k,t) &= \frac{\tilde{n}_{at}(k,0)}{[U_2(0) - 2\tilde{n}_{aa}(k,0)]\kappa t + 1} = \frac{N_{at}(0)\frac{\lambda}{ik+\lambda}}{[2N_{aa}(0)\frac{ik}{ik+\lambda} + N_{at}(0)]\kappa t + 1} \\ &= \frac{N_{at}(0)\lambda}{(U_2(0)\kappa t + 1)ik + \lambda(1 + N_{at}(0)\kappa t)} = \frac{N_{at}(t)\lambda(t)}{ik + \lambda(t)} \qquad (6.70)\end{aligned}$$

$$n_{at}(x,t) = N_{at}(t)\lambda(t)e^{-\lambda(t)x}, \tag{6.71}$$

where

$$\lambda(t) = \lambda\frac{N_{at}(0)\kappa t + 1}{U_2(0)\kappa t + 1} \qquad \text{and} \qquad \lambda(\infty) = \lambda\frac{N_{at}(0)}{U_2(0)} = \lambda\frac{2N}{2\lambda S_T} = \frac{N}{S_T}. \tag{6.72}$$

The edge density is thus exponentially distributed with a shape parameter $\lambda(t)$ that decays hyperbolically at early times but then converges at steady state to $\frac{N}{S_T}$, the natural break frequency of the organism. Thus, at very large times, the average edge fragment length approaches, but does not exceed, the organisms average chromosome length.[9] This result should be compared to the previous two sections where the mean fragment length approaches infinity (no edges) or remains constant (no internals).

For internal fragments,

$$\tilde{n}_{aa}(k,t) = \frac{\tilde{n}_{aa}(k,0)}{[U_2(0)\kappa t + 1][(U_2(0) - 2\tilde{n}_{aa}(k,0))\kappa t + 1]}$$

[8] Recall from section 5.5 that $p_c(x,t)$ is the probability that a fragment has size x given it is in pool c. Thus $p_c(x,t)$ defines the shape of the fragment density and $N_c(t)$ the amplitude, i.e. $n_c(x,t) = N_c(t)p_c(x,t)$ and $\tilde{n}_c(k,t) = N_c(t)\tilde{p}_c(k,t)$.

[9] Recall that the mean of an exponential $p(x,t) = \lambda e^{-\lambda x}$ is $\frac{1}{\lambda}$.

$$
\begin{aligned}
&= \frac{N_{aa}(0)\frac{\lambda}{ik+\lambda}}{[U_2(0)\kappa t+1][(2N_{aa}(0)\frac{ik}{ik+\lambda}+N_{at}(0))\kappa t+1]} \\
&= \frac{N_{aa}(0)\lambda}{[U_2(0)\kappa t+1][(U_2(0)\kappa t+1)ik+\lambda(1+N_{at}(0)\kappa t)]} \\
&= \frac{N_{aa}(0)}{[U_2(0)\kappa t+1][1+N_{at}(0)\kappa t]}\left(\frac{\lambda[1+N_{at}(0)\kappa t]}{[(U_2(0)\kappa t+1)ik+\lambda(1+N_{at}(0)\kappa t)]}\right) \\
&= N_{aa}(t)\frac{\lambda[1+N_{at}(0)\kappa t]}{[(U_2(0)\kappa t+1)ik+\lambda(1+N_{at}(0)\kappa t)]} \\
&= N_{aa}(t)\frac{\lambda(t)}{ik+\lambda(t)} \qquad (6.73)\\
n_{aa}(x,t) &= N_{aa}(t)\lambda(t)e^{-\lambda(t)x}, \qquad (6.74)
\end{aligned}
$$

i.e. the internal fragment density also remains exponential. Internals and edges have the same $\lambda(t)$ dependence. Why is this? Note that when internal fragments react with edges or internals they act to drive the fragment densities toward larger x values and thus smaller $\lambda(t)$ values. Although internals drive themselves toward large x twice as fast as they drive edges, internals are also anihilated by edges twice as fast as edges anihilate themselves. These effects appear to cancel so that edges and internals have the same $\lambda(t)$.

Since $\lambda_a(t)=\frac{p\lambda_0}{U_2(0)\kappa t+1}$ is *hyperbolic*, the *bilinear* density shape parameter $\lambda(t)$ is not equivalent to $\lambda_a(t)$. From its form as $t\to\infty$, it appears that $\lambda(t)$ represents the number of breaks per base pair regardless of whether the break is an active DSB or two natural telomeres. This interpretation of $\lambda(t)$ is consistent with our definition of λ as the initial number of active DSBs per base pair, since, by the strong condition, the initial number of breaks approximately equals the initial number of active DSBs. However, at large times, when the majority of active DSBs have rejoined, the difference between $\lambda(t)$ and $\lambda_a(t)$ becomes significant.

Bitelomeric fragments have the following distribution

$$
\begin{aligned}
\tilde{n}_{tt}(k,t) &= \frac{1}{2}\frac{\tilde{n}_{at}^2(k,0)\kappa t}{(U_2(0)-2\tilde{n}_{aa}(k,0))\kappa t+1} \\
&= \frac{1}{2}\frac{N_{at}^2(0)\kappa t\frac{\lambda^2}{(\lambda+ik)^2}}{[2N_{aa}(0)\frac{ik}{ik+\lambda}+N_{at}(0)]\kappa t+1} \\
&= \frac{1}{2}\frac{N_{at}^2(0)\kappa t\frac{\lambda^2}{(\lambda+ik)}}{[(U_2(0)\kappa t+1)ik+\lambda(1+N_{at}(0)\kappa t)]} \\
&= \frac{1}{2}\frac{N_{at}^2(0)\kappa t}{1+N_{at}(0)\kappa t}\left(\frac{\lambda}{\lambda+ik}\right)\left(\frac{\lambda(1+N_{at}(0)\kappa t)}{(U_2(0)\kappa t+1)ik+\lambda(1+N_{at}(0)\kappa t)}\right) \\
&= N_{tt}(t)\left(\frac{\lambda}{\lambda+ik}\right)\left(\frac{\lambda(t)}{ik+\lambda(t)}\right) \qquad (6.75)\\
&= N_{tt}(t)\frac{\lambda\lambda(t)}{\lambda-\lambda(t)}\left(\frac{1}{ik+\lambda(t)}-\frac{1}{ik+\lambda}\right) \qquad (6.76)\\
n_{tt}(x,t) &= N_{tt}(t)\frac{\lambda\lambda(t)}{\lambda-\lambda(t)}\left(e^{-\lambda(t)x}-e^{-\lambda x}\right) \qquad (6.77)\\
n_{tt}(x,\infty) &= \frac{N_{at}(0)}{2}\frac{\lambda\frac{N}{S_T}}{\lambda-\frac{N}{S_T}}\left(e^{-\frac{N}{S_T}x}-e^{-\lambda x}\right). \qquad (6.78)
\end{aligned}
$$

The density $n_{tt}(x,\infty)$ has a mode at $x=\frac{1}{\lambda}\ln(\frac{\lambda S_T}{N})$ that can be found by solving $\frac{\partial n_{tt}(x,\infty)}{\partial x}=0$. Since the mode was located at $x=\frac{1}{\lambda}$ when there were only edges, the presence of internal fragments augments the mode's position by the factor $\ln(\frac{\lambda S_T}{N})$, a measure of the number of internals available for insertion between edges.

Equation 6.75 is a k-space statement that the shape factor $p_{tt}(x,t)$ equals the cross-convolution of $p_{at}(x,0)$ and $p_{at}(x,t)$. This result is interesting because, in view of equations 6.3 and 6.35, $n_{tt}(x,t)$ is also equal to the integral of many differential density contributions, each an *auto*-convolution of $n_{at}(x,\tau)$. The number of these contributions is counted by $N_{tt}(t)$ while, curiously, the average shape of the density contributions $p_{tt}(x,t)$ equals the *cross*-convolution $p_{at}(x,0) * p_{at}(x,t)$. Since $\lambda(\infty) = N/S_T$, $p_{at}(x,\infty)$ depends only on the organism while, by the weak condition, $p_{at}(x,0)$ depends only on the dose. The bitelomeric probability density $p_{tt}(x,\infty)$ is thus a product, in k-space, of a pure dose factor and a pure cell factor.

The ring solutions are

$$\begin{aligned}
\tilde{n}_r(k,t) &= \frac{1}{2}\ln\left(\frac{U_2(0)\kappa t + 1}{[U_2(0) - 2\tilde{n}_{aa}(k,0)]\kappa t + 1}\right) \\
&= \frac{1}{2}\ln\left(\frac{U_2(0)\kappa t + 1}{[2N_{aa}(0)\frac{ik}{ik+\lambda} + N_{at}(0)]\kappa t + 1}\right) \\
&= \frac{1}{2}\ln\left(\frac{(U_2(0)\kappa t + 1)(ik+\lambda)}{[2N_{aa}(0)ik + (ik+\lambda)N_{at}(0)]\kappa t + (ik+\lambda)}\right) \\
&= \frac{1}{2}\ln\left(\frac{(U_2(0)\kappa t + 1)(ik+\lambda)}{(U_2(0)\kappa t + 1)ik + \lambda(1 + N_{at}(0)\kappa t)}\right) \\
&= \frac{1}{2}\ln\left(\frac{ik+\lambda}{ik+\lambda(t)}\right) \qquad (6.79) \\
n_r(x,t) &= \frac{1}{2x}\left(e^{-\lambda(t)x} - e^{-\lambda x}\right) \qquad (6.80) \\
n_r(x,\infty) &= \frac{1}{2x}\left(e^{-\frac{N}{S_T}x} - e^{-\lambda x}\right), \qquad (6.81)
\end{aligned}$$

where the inverse transform was inferred from a table of Laplace transforms [2].

The steady state *number* densities $n_r(x,\infty)$ and $n_{tt}(x,\infty)$ are plotted in figure 6.3 for human cells at doses of 10, 100, and 1000 Gray. The bitelomeric distributions have a pool size which converges to 46 with increasing dose, see table 6.1. If we let the dose get very large, the bitelomeric fragment density approaches $n_{tt}(x,\infty) = 46\frac{N}{S_T}e^{-\frac{N}{S_T}x}$ which, when integrated over x, equals 46. This limiting bitelomeric density is apparent in the left panel of figure 6.3. The *total* ring pool size does not have a high dose limit. However, the right panel of figure 6.3 shows that in going from 100 Gray to 1000 Gray, the increase in total rings is strictly due to the generation of ring fragments less than 25 Mbp.

Since $m_c(x,t) = x n_c(x,t)$, equations 6.78 and 6.81 imply that the misrepaired fragment *mass* densities at very large doses converge to

$$m_{tt}(x,\infty) \approx x\frac{N_{at}(0)}{2}\frac{N}{S_T}e^{-\frac{N}{S_T}x} \quad \text{and} \quad m_r(x,\infty) \approx \frac{1}{2}e^{-\frac{N}{S_T}x}. \qquad (6.82)$$

In the upper panels of figure 6.4, plots of the mass densities with increasing doses illustrate this high dose convergence. Integrating these limiting expressions from $x = 0$ to ∞ yields $M_{tt}(\infty) = S_T$ and $M_r(\infty) = \frac{S_T}{2N}$, i.e. in the limit of high doses, the ring mass approaches one half the average chromosome mass. Note that mass is not conserved ($S_T + \frac{S_T}{2N} > S_T$) because, in forming the SSEB model, the ring term was dropped as an internal fragment degradation flux, but kept as a ring synthesis flux. If we include the term in both locations, $S_T - \frac{S_T}{2N}$ should be the amount of DNA that ends up in bitelomeric fragments at very high doses.

In the lower panels of figure 6.4 we see that in the high dose limit, the total DNA mass in fragments with $x \leq 10$ Mbp approaches $\int_0^{10} m_r(x,\infty)dx \approx$5 Mbp and $\int_0^{10} m_{tt}(x,\infty)dx \approx$15 Mbp. As rings do not enter PFGE gels [24], the

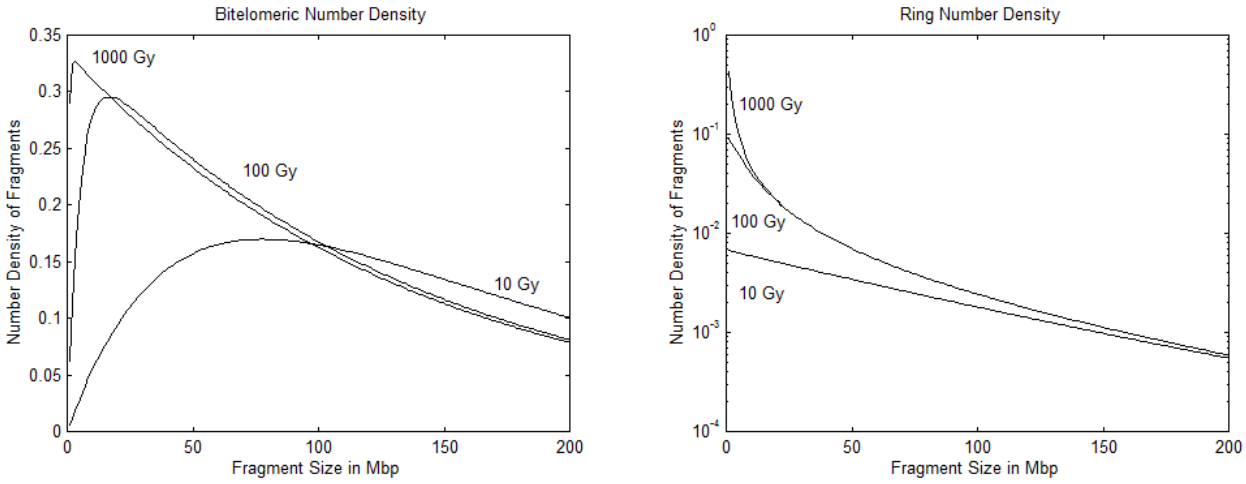

Figure 6.3: The misrepaired fragment *number* densities $n_{tt}(x, \infty)$ (left panel) and $n_r(x, \infty)$ (right panel) are plotted here for doses of 10, 100, and 1000 Gray. At 10 Gray the weak condition is marginally valid (see table 6.1) and the strong condition does not hold. However, for $D \geq 100$ Gray both conditions can be assumed.

steady state FAR value predicted for very high doses is $\frac{15}{138*46} = 0.0024$. Since this FAR value is below the noise level, the induced DNA fragment models of chapter 5 should be adequate for DSB estimation in *undigested* misrejoined cells.

6.6 Misrepair Estimation

If weak and strong conditions hold (human cells at 80 Gray), the solutions to the full SSEB model can be inserted into equation 6.12 to obtain an expression for the measurable density $n(x', t)$ at arbitrary repair times. In theory, the resulting expression could be used to provide least squares estimates of the parameters λ, $\lambda(t)$, and $\lambda_i(t)$ from a single gel lane.[10] These estimates could then be converted into $U_1(t) = \lambda_i(t)S_T$, $U_2(t) = \lambda(t)S_T - N$, and $M(t) = \lambda S_T - U_2(t)$. In practice, however, the amount of misrepaired fragment signal in the PFGE detectable size range (0 to 10 Mbp) appears to be inadequate for misrepair estimation in undigested cells.

To increase the amount of PFGE detectable fragment signal, DNA can be digested with the *notI* restriction enzyme before running the gel [25]. Since notI-somes (section 3.1) are much smaller than human chromosomes (most are less than 3 Mbp), weak and strong conditions cannot be assumed at 80 Gray. In the Löbrich experiment [25] (see section 3.1), an 80 Gray dose was chosen so that an average size notI-some receives approximately one DSB. Since only a fraction of these DSBs are expected to be active, it is reasonable to assume (for 80 Gray) that the number of internal fragments after notI digestion is negligible, i.e. that a notI-some receives at most one active DSB.

Suppose a probe is hybridized to an end of the 3.2 Mbp notI-some so that hybridized edge fragments have the induced distribution $\lambda e^{-\lambda x} u_0^{3.2}(x)$. Furthermore, let a be the size of the notI-some that produces the edge that quenches the hybridized edge. Assume $a < 3.2 = b$ Mbp. The hybridized bitelomere (final product) will then have a

[10] Recall that weak and strong conditions imply $N_{at}(0) = 2N$ and $N_{aa}(0) = \lambda S_T$.

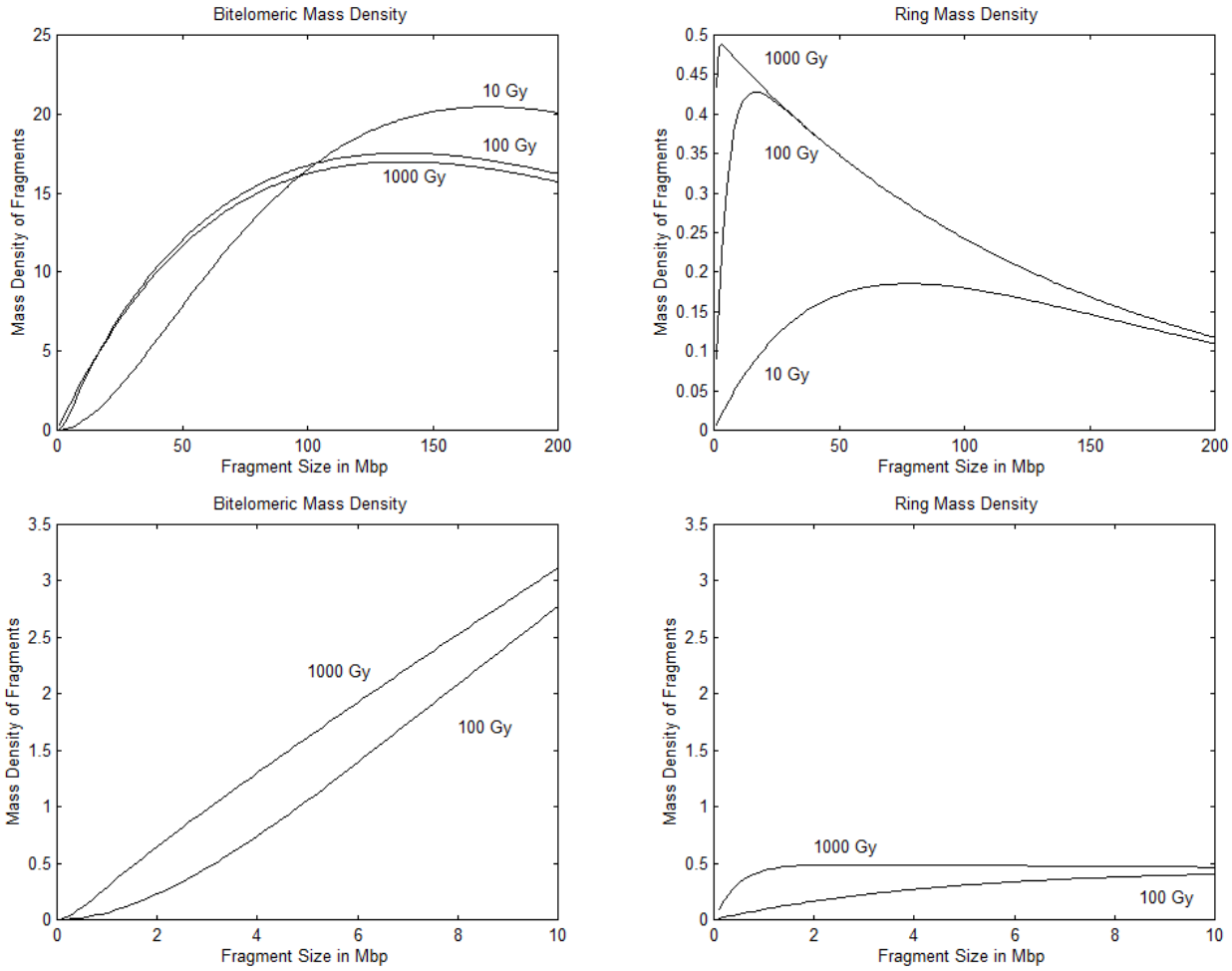

Figure 6.4: Upper panels: The misrepaired fragment *mass* densities $m_{tt}(x,\infty)$ and $m_r(x,\infty)$ are plotted for human cells at doses of 10, 100, and 1000 Gray. Lower panels: An enlarged view of the region x=0 to 10 Mbp above.

size distribution equal to

$$\lambda^2 x e^{-\lambda x} u_0^a(x) + \lambda^2 a e^{-\lambda x} u_a^b(x) + \lambda^2 (a+b-x) e^{-\lambda x} u_b^{a+b}(x). \tag{6.83}$$

If the density of notI-somes before irradiation $[n_\delta(x,0^-)]$ is estimated from PFGE unirradiated control cells, the final misrepair density can be written as

$$\begin{aligned} n_m(x,\infty) &= \int_0^b n_\delta(a,0^-)[\lambda^2 x e^{-\lambda x} u_0^a(x) + \lambda^2 a e^{-\lambda x} u_a^b(x) \\ &+ \lambda^2 (a+b-x) e^{-\lambda x} u_b^{a+b}(x)]da. \end{aligned} \tag{6.84}$$

This misrepair distribution, when added to the eurepair density $e^{-\lambda b}\delta(x-b)$, is equal to the expected hybridization signal at very large times. If the integral over a is approximated as a finite sum, the resulting expression could be fit to PFGE data by least squares to obtain a distribution-based estimate of λ and thus $M(\infty) = \lambda S_T$. The advantage of distribution-based estimators is that density *shape* information is extracted from the gels.

6.7 Summary

The SSEB model is an extension of the EtBr model to include the DSB misrepair dynamics of the SS model. It has elegant closed form solutions for doses high enough to satisfy both the weak and strong conditions. In the hypothetical case that a cell has no edge fragments, the internal fragment density has a mean value that increase with time while the amplitude, the pool size, decays hyperbolically. Hypothetical cells without internal fragments (but at doses high enough that edges are exponential, impossible!) have fragment densities (edges and bitelomeres) with constant shapes but changing amplitudes. The full SSEB model includes edges and internals and is therefore *not* hypothetical. Its solutions, the fragment densities, have shapes that change during a transient period before becoming constant at steady state. In the limit of high doses, the steady state ring mass is predicted to approach one-half the average chromosome mass. The SSEB model possesses the interesting property that exponential densities remain exponential. A strategy for distribution-based misrepair estimation was outlined but its development is left for future work.

Chapter 7

Concluding Remarks

7.1 Summary and Conclusions

Recent high dose PFGE data [25] was shown to support the SS model as the best of five deterministic DSB misrepair models. The Sax-Markov binary eurepair (SMBE) model was then presented as a stochastic extension of the SS model that includes the accidental binary eurepair of correctly matched DSB free ends. The SMBE model predicts a linear-quadratic-linear misrepair dose-response which when fit to chromosomal aberration data [15] implies that the free ends of a DSB are kept within a nuclear subvolume approximately 100 times smaller than the nucleus. Arguments that translocation mediated cancer risks are linear in misrepairs for doses less than 1 Gray were then given. Finally, the quadratic region of the SMBE dose response was shown to be eliminated under high LET conditions.

The research focus then shifted from models of double-strand breaks to models of DNA fragment length distributions. Fragment models are needed to derive least squares estimates of DSBs and misrepairs. Least squares techniques (i.e. distribution-based estimation approaches) increase the amount of information extracted from a PFGE gel because they transform experimental data directly into a best estimate of the model parameters. Thus distributional approaches bypass the informational losses associated with single statistic "middle men" such as FAR, or the amount of signal outside of a specific gel band. Chapter 5 provided detailed derivations of the EtBr model and the specific probe model. These models describe a cell's DNA fragment distribution immediately after an acute exposure, but not after a significant amount of misrepair has occured. Chapter 6 filled this void by presenting a model for misrepaired DNA fragments, the SSEB model. The SSEB model is an extension of the EtBr model to include the DSB misrepair dynamics of the SS model. Solutions to the SSEB model were thoroughly examined for doses high enough to satisfy both the weak and the strong conditions. A low steady state FAR value in the limit of high doses suggests that DSB estimation using induced fragment models is acceptable in *undigested* misrepaired cells.

7.2 Model Limitations

The SS model was selected based on PFGE data from quiescent human fibroblasts X-irradiated with acute doses of 80 and 160 Gray. The SS and SSEB models are not applicable to lower doses because they do not include the accidental binary eurepair of DSB free ends. The parameter values of the SMBE model are lymphocyte specific. The models do not apply to yeast or to other lower organisms that repair DSBs via homologous recombination.

If asynchronous cell populations are exposed to an arbitrary dose-rate schedule, the early part of the schedule can perturb the cell cycle phase distribution and thus alter the cell's response to the latter part of the schedule. Since a tumor is an asynchronous cell population, the DSB repair models of this work cannot be applied to radiation therapy without the inclusion of cell cycle effects [13, 20, 32, 49]. Cell cycle effects do not appear to be critical for environmental risk applications; at very low doses (the predictive end of the model) cell cycle delays are not observed [47], and at high acute doses (the data fitting end of the model) there is no latter part of the dose-rate schedule and thus no detection of the disrupted cell cycle phase distribution. Environmental applications may, however, require a representation of the adaptive response [47].

7.3 Future Work

Aberration Meta-Analysis: We had originally believed that quadratic SMBE misrepair risks would continue to doses as low as 0.005 Gray, the single track transition dose obtained using microdosimetric considerations that assume one well-mixed nucleus. Had this been true, the cancer risks predicted by a linear fit to A-bomb data would overestimate the actual risks (at occupational exposure levels) by a factor close to 100. However, chromosomal aberration data appears to support a single track linear-quadratic transition dose on the order of 0.5 Gray. Assuming this to be correct, cancer risks are not overestimated and the cell nucleus consists of approximately 100 isolated subvolumes rather than one well-mixed chamber. If data can be found to support a single track transition dose less than 0.3 Gray (the average dose to A-bomb survivors), occupational risks would again be overestimated to some degree. A meta-analysis of the chromosomal aberration literature should therefore be performed to determine a confidence interval for the single track transition dose.

Dose-Rate Models: Risk extrapolations from acute to chronic doses are important because experimental exposures tend to be acute while human environmental exposures tend to be chronic. Chronic dose extrapolations can be made by adding dose-rate forcing functions to the DSB repair models and setting the initial conditions to zero.[1] Simulations using chronic dose-rate schedules can then provide chronic dose misrepair predictions. The SS model is not expected to display a dose-rate effect since it assumes that a constant fraction of the DSBs misrepair regardless of their formation rate.[2] The SMBE model, on the other hand, is expected to have a dose-rate effect since accidental binary eurepair becomes more likely as the dose-rate is lowered. To characterize the SMBE dose-rate effect, a continuous-time SMBE simulation is needed. Furthermore, the stochastic nature of particle track inter-arrival times may require the use of Monte-Carlo simulations [37].

The SSEB Model: To relax weak and strong conditions, the Fourier tranform of the EtBr model (as an initial

[1] The parameter values remain as identified previously for acute doses.

[2] A dose-rate effect is said to occur if an acute dose of size D is more potent than a chronic dose of size D, i.e. if spreading the dose over time helps the cell overcome the assault.

condition to the SSEB model) can be inserted into the SSEB k-space solutions and the resulting expression inverse transformed. Alternatively, the initial condition could be sampled, fast fourier transformed (FFT), inserted into the SSEB solutions, and inverse transformed. The advantage of a numerical approach is that it can also be applied to high LET initial conditions.

Misrepair Estimation: The strategy outlined in section 6.6 should be implemented in order to compare sensitivities of both distribution-based and conventional misrepair estimation techniques. A sensitive misrepair assay could have applications in radiation oncology as a predictive assay, i.e. an assay that predicts an individual's radiobiological response before therapeutic doses are admistered.

Bibliography

[1] N. A. Albright. A Markov formulation of the repair-misrepair model of cell survival. *Radiat. Res.*, 118:1–20, 1989.

[2] W. H. Beyer. *CRC Standard Mathematical Tables.* CRC Press, Boca Raton, 1983.

[3] D. Blöcher. In CHEF electrophoresis a linear induction of DSB corresponds to a non-linear fraction of extracted DNA with dose. *Int. J. of Radiat. Biol.*, 57:7–12, 1990.

[4] G. F. Carle and M. V. Olson. An electrophoretic karyotype for yeast. *Proc. Natl. Acad. Sci.*, 82:3756–3760, 1985.

[5] B. Cedervall, P. Källman, and W. C. Dewey. Repair of DNA double-strand breaks: errors encountered in the determination of half-life times in pulsed-field gel electrophoresis and neutral filter elution. *Radiat. Res.*, 142:23–8, 1995.

[6] B. Cedervall and T. Radivoyevitch. Methods for analysis of DNA fragment distributions on pulsed field gel electrophoretic gels. *Electrophor.*, 17:1080–1086, 1996.

[7] B. Cedervall, R. Wong, N. Albright, J. Dynlacht, P. Lambin, and W. C. Dewey. Methods for the quantification of DNA double-strand breaks determined from the distribution of DNA fragment sizes measured by pulsed-field gel electrophoresis. *Radiat. Res.*, 144:122, 1995.

[8] B. Cedervall, R. Wong, N. Albright, J. Dynlacht, P. Lambin, and W. C. Dewey. Methods for the quantification of DNA double-strand breaks determined from the distribution of DNA fragment sizes measured by pulsed-field gel electrophoresis. *Radiat. Res.*, 143:8–16, 1995.

[9] C. R. Contopoulou, V. E. Cook, and R. K. Mortimer. Analysis of DNA double strand breakage and repair using orthogonal field alternation gel electrophoresis. *Yeast*, 3:71–6, 1987.

[10] V. E. Cook and R. K. Mortimer. A quantitative model of DNA fragments generated by ionizing radiation, and possible experimental applications. *Radiat. Res.*, 125:102–6, 1991.

[11] V. T. Covello and M. W. Merkhofer. *Risk Assessment Methods: approaches for assessing health and environmental risks.* Plenum Press, New York, 1993.

[12] K. S. Crump, D. G. Hoel, C. H. Langley, and R. Peto. Fundamental carcinogenic processes and their implications for low dose risk assessment. *Cancer Res.*, 36:2973–2979, 1976.

[13] L. E. Dillehay. A model of cell killing by low-dose-rate radiation including repair of sublethal damage, G_2 block, and cell division. *Radiat. Res.*, 124:201–207, 1990.

[14] G. I. Evan, L. Brown, M. Whyte, and E. Harrington. Apoptosis and the cell cycle. *Current Opinion in Cell Biology*, 7:825–834, 1995.

[15] P. Finnon, D. C. Lloyd, and A. A. Edwards. Fluorescence *in situ* hybridization detection of chromosomal aberrations in human lymphocytes: applicability to biological dosimetry. *Int. J. Radiat. Biol.*, 68:429–435, 1995.

[16] A. A. Friedl, W. Beisker, K. Hahn, F. Eckardt-Schupp, and A. M. Kellerer. Application of pulsed field gel electrophoresis to determine gamma-ray-induced double-strand breaks in yeast chromosomal molecules. *Int. J. Radiat. Biol.*, 63:173–81, 1993.

[17] Harley N. H. Toxic effects of radiation and radioactive materials. In M. O. Amdur, J. Doull, and C. D. Klaassen, editors, *Casarett and Doull's Toxicology: The basic science of poisons*, pages 723–752, McGraw-Hill Inc, New York, 1993.

[18] P. J. Hahnfeldt, R. K. Sachs, and L. R. Hlatky. Evolution of DNA damage in irradiated cells. *J. Math. Biol.*, 30:493–511, 1992.

[19] A. M. Kellerer. *The dosimetry of ionizing radiation*, chapter Fundamentals of Dosimetry. Volume 1, Academic Press, 1985.

[20] C. R. King, R. Nath, and S. Rockwell. Effects of continuous low-dose-rate irradiation: computer simulations. *Cell Tissue Kinetics*, 21:339–351, 1988.

[21] A. Kraxenberger, A. A. Friedl, and A. M. Kellerer. Computer simulation of pulsed field gel runs allows the quantitation of radiation-induced double-strand breaks in yeast. *Electrophoresis*, 15:128–36, 1994.

[22] T. S. Lawrence, D. P. Normolle, M. A. Davis, and J. Maybaum. The use of biphasic linear ramped pulsed field gel electrophoresis to quantify DNA damage based on fragment size distribution. *Int. J. Radiat. Oncol., Biol., Phys.*, 27:659–63, 1993.

[23] A. R. Lehmann and M. G. Ormerod. The replication of DNA in murine lymphoma cells (l5178y). i. rate of replication. *Biochimica et Biophysica Acta*, 204:128–43, 1970.

[24] S. D. Levene and B. H. Zimm. Separations of open circular DNA using pulsed-field electrophoresis. *Proc. Natl. Acad. Sci.*, 84:4054–4057, 1987.

[25] M. Löbrich, B. Rydberg, and P. K. Cooper. Repair of X-ray induced DNA double-strand breaks in specific *NotI* restriction fragments in human fibroblasts: joining of correct and incorrect pairs. *Proc. Natl. Acad. Sci.*, 92:12050–12054, 1995.

[26] J. M. Love and L. J. Gudas. Vitamin A, differentiation and cancer. *Current Opinion in Cell Biology*, 6:825–31, 1994.

[27] T. K. MacLachlan, N. Sang, and A. Giordano. Cyclins, cyclin-dependent kinases and cdk inhibitors: implications in cell cycle control and cancer. [review]. *Critical Reviews in Eukaryotic Gene Expression*, 5:127–56, 1995.

[28] K. B. Marcu, S. A. Bossone, and A. J. Patel. Myc function and regulation. *Annu. Rev. Biochem.*, 61:809 – 860, 1992.

[29] M. D. Mason. Differentiation and cancer. *Eur. J. Cancer*, 28:255–8, 1992.

[30] V. Michalik. Model of DNA damage induced by radiations of various qualities. *Int. J. Radiat. Biol.*, 62:9–20, 1992.

[31] P. Modrich. Mismatch repair, genetic stability, and cancer. *Science*, 266:1959–1960, 1994.

[32] B. Novak and J. J. Tyson. Quantitative analysis of a molecular model of mitotic control in fission yeast. *J. Theoret. Biol.*, 173:283–305, 1995.

[33] A. V. Oppenheim, A. S. Wilsky, and I. T. Young. *Signals and Systems*. Prentice-Hall Inc., Englewood Cliffs, NJ, 1983.

[34] T. Radivoyevitch and B. Cedervall. Mathematical analysis of DNA fragment distribution models used with pulsed-field gel electrophoresis. *Electrophor.*, 17:1087–1093, 1996.

[35] S. H. Revell. The breakage-and-reunion theory and the exchange theory for chromosomal aberrations induced by ionizing radiations: a short history. In J.T. Lett, H. Adler, and M. Zelle, editors, *Adavnces in Radiation Biology*, pages 367–416, Academic Press, New York, 1974.

[36] R. K. Sachs, P. L. Chen, P. J. Hahnfeldt, and L. R. Hlatky. DNA damage caused by ionizing radiation. *Math. Biosci.*, 112:271–303, 1992.

[37] R. K. Sachs and L. R. Hlatky. Dose-rate dependent stochastic effects in radiation cell-survival models. *Radiat. Enivron. Biophys.*, 29:169–184, 1990.

[38] R. K. Sachs, L. R. Hlatky, P. J. Hahnfeldt, and P. L. Chen. Incorporating dose-rate effects in Markov radiation cell survival models. *Radiat. Res.*, 124:216–226, 1990.

[39] A. Sancar. Mechanisms of DNA excision repair. *Science*, 266:1954–1956, 1994.

[40] K. Sax. An analysis of X-ray induced chromosomal aberrations in *tradescantia*. *Genetics*, 25:41–68, 1940.

[41] D. C. Schwartz and C. R. Cantor. Separation of yeast chromosome-sized DNA molecules by pulsed field gradient gel electrophoresis. *Cell*, 37:67–75, 1984.

[42] W. G. Stetler-Stevenson, L. A. Liotta, and D. E. Kleiner. Extracellular matrix 6: role of martix metalloproteinases in tumor invasion and metastasis. *FASEB Journal*, 7:1434–1441, 1993.

[43] J. S. Strobl, W. F. Wonderlin, and D. C. Flynn. Mitogenic signal transduction in human breast cancer cells. [review]. *General Pharmacology*, 26:1643–9, 1995.

[44] C. A. Tobias. The repair-misrepair model in radiobiology: comparison to other models. *Radiat. Res.*, 104:S77–S95, 1985.

[45] C. von Sonntag. *The Chemical Basis of Radiation Biology*. Taylor & Francis, London, 1987.

[46] J. F. Ward. The yield of DNA double-strand breaks produced intracellarly by ionizing radiation: a review. *Int. J. Radiat. Biol.*, 57:1141–1150, 1990.

[47] S. Wolff. Aspects of the adaptive response to very low doses of radiation and other agents. *Mutation Res.*, 358:135–142, 1996.

[48] L. Y. Xue, L. R. Friedman, N. L. Oleinick, and S. M. Chiu. Induction of DNA damage in γ-irradiated nuclei stripped of nuclear protein classes: differential modeulation of double-strand break and DNA-protein crosslink formation. *Int. J. Radiat. Biol.*, 66:11–21, 1994.

[49] M. Zaider, C. S. Wu, and G. N. Minerbo. The combined effects of sublethal damage repair, cellular repopulation and redistribution in the mitotic cycle. i. survival probabilities after exposure to radiation. *Radiat. Res.*, 145:457–466, 1996.