RADIATION RESEARCH **161**, 451–457 (2004)
0033-7587/04 $15.00

A Fast Monte Carlo Algorithm to Simulate the Spectrum of DNA Damages Formed by Ionizing Radiation

V. A. Semenenko and R. D. Stewart[1]

Purdue University, School of Health Sciences, West Lafayette, Indiana 47907-2051

Semenenko, V. A. and Stewart, R. D. A Fast Monte Carlo Algorithm to Simulate the Spectrum of DNA Damages Formed by Ionizing Radiation. *Radiat. Res.* 161, 451–457 (2004).

Ionizing radiation produces both singly and multiply damaged DNA sites. Multiply damaged sites (MDS) have been implicated in radiation-induced cell killing and mutagenesis. The spatial distribution of elementary damages (strand breaks and base damages) that constitute MDS is of special interest, since the complexity of MDS has an impact on damage repair. A fast and easy-to-implement algorithm to simulate the local clustering of elementary damages produced by ionizing radiation is proposed. This algorithm captures the major trends in the DNA damage spectrum predicted using detailed track-structure simulations. An attractive feature of the proposed algorithm is that only four adjustable parameters need to be identified to simulate the formation of DNA damage. A convenient recipe to determine the parameters used in the fast Monte Carlo damage simulation algorithm is provided for selected low- and high-LET radiations. The good agreement among the damage yields predicted by the fast and detailed damage formation algorithms suggests that the small-scale spatial distribution of damage sites is determined primarily by independent and purely stochastic events and processes.

INTRODUCTION

Low-LET radiation creates approximately 1,000 single-strand breaks (SSBs) and 40 double-strand breaks (DSBs) per gray in a typical mammalian cell (*1*). In addition, ionizing radiation causes massive amounts of damage to nucleobases. The level of base damage is estimated at 2,500 to 25,000 Gy^{-1} $cell^{-1}$ (*2*). Besides isolated damages, ionizing radiation is known to produce multiply damaged sites (MDS) consisting of two or more elementary damages within a few helical turns of the DNA. Here the term elementary damage means either a strand break or a base damage. Isolated strand breaks and multiple strand breaks on the same DNA strand within a few turns of the DNA are detected as an SSB in most experimental assays. DSBs are formed when at least two strand breaks are formed in proximity on opposite strands of the DNA. DSBs and other classes of MDS may be the primary cause of radiation-induced cell killing (*3*) and mutagenesis (*2*). The biological significance of MDS is attributed to the fact that they are likely to present a special challenge to DNA repair systems (*1, 4, 5*).

Knowledge of the spatial configuration of elementary damages within an MDS is important. For example, the experimental studies show that the outcome from repair depends on the types and relative positions of the elementary damages (see refs. *6* and *7* for reviews). Most of the experimental techniques used to detect radiation damage to the DNA provide only limited information about the exact number and spatial configuration of elementary damages within one or two turns of the DNA. Instead, detailed information about the spectrum of possible damage configurations produced by ionizing radiation is often obtained using Monte Carlo track structure simulations in combination with geometrical models of the DNA and chromatin (*8–12*). Although detailed simulations such as these will most likely remain the ''gold standard'' for predicting the spectrum of damage configurations produced by radiation, this approach is very expensive computationally and may be impractical for some applications, e.g. estimating the absolute yield of some rare but possibly difficult to repair damage configurations.

In this article, we propose a simple and fast Monte Carlo scheme to simulate the formation of singly and multiply damaged DNA sites by radiations of different quality. The Monte Carlo damage simulation (MCDS) algorithm requires the determination of only four adjustable parameters. Two of these parameters can be constrained to a reasonable (and rather narrow) range of values using measured data. Comparisons of results from the fast and detailed (track structure) Monte Carlo damage formation algorithms suggest the possibility that the small-scale spatial distribution of elementary damages is determined primarily by independent and purely stochastic events and processes. That is, all of the nucleotides within one or two turns of the DNA are equally likely to be damaged by radiation.

[1] Author to whom correspondence should be addressed at Purdue University, School of Health Sciences, 550 Stadium Mall Drive, West Lafayette, IN 47907-2051; e-mail: trebor@purdue.edu.

METHODS

Algorithm to Simulate DNA Damage

To simulate the formation of DNA damage, four adjustable parameters must be specified:

1. Number of strand breaks Gy^{-1} $cell^{-1}$, σ_{Sb}.
2. Number of base damages Gy^{-1} $cell^{-1}$, σ_{Bd}. The yield of base damages may conveniently be specified in terms of the base damage to strand break ratio, i.e. $f \equiv \sigma_{Bd}/\sigma_{Sb}$.
3. DNA segment length in base pairs (bp) Gy^{-1} $cell^{-1}$, n_{seg}. This DNA segment length is an *ad hoc* parameter and should not be considered equivalent to the DNA content of a specific chromosome or cell.
4. Minimum length of undamaged DNA (bp) between neighboring elementary damages such that these elementary damages are said to belong to two different lesions, N_{min}. Here the term lesion means both isolated elementary damages and MDS.

The proposed Monte Carlo damage generation algorithm has two major steps: (1) randomly distribute in a DNA segment (n_{seg} parameter) the expected number of elementary damages produced in a cell by a specified amount of radiation and (2) subdivide the distribution of elementary damages in the segment into lesions. The grouping of elementary damages into lesions is determined primarily by the N_{min} parameter. The step-by-step procedure to distribute elementary damages in the DNA segment is as follows:

1. Compute the parameters for a specific absorbed dose, D (Gy). The segment length $cell^{-1}$ is $N_{seg} = gn_{seg}D$. The total number of strand breaks $cell^{-1}$ is $\Sigma_{Sb} = g\sigma_{Sb}D$. Here g is a dimensionless scale factor that can be used to adjust the absolute yield of DNA lesions to better mimic experimental observations for specific cell types (see the Discussion). When cell-specific information is not available, g should be set to unity. The number of base damages $cell^{-1}$ can be calculated from the number of strand breaks $cell^{-1}$, i.e. $\Sigma_{Bd} = f\,\Sigma_{Sb}$.
2. Select a nucleotide pair at random from the DNA segment; i.e., select a uniformly distributed integer in the range $[1, N_{seg}]$.
3. Select a DNA strand (1 or 2) at random. If the selected nucleotide is not already damaged, record the strand break at the location. Otherwise, go to step 2.
4. Set $\Sigma_{Sb} = \Sigma_{Sb} - 1$. If $\Sigma_{Sb} > 0$, go to step 2.
5. Repeat steps 2 through 4 for base damages.

The second major step in the damage simulation algorithm is the grouping of elementary damages into lesions. This elementary damage grouping process is the somewhat arbitrary procedure of identifying a subset of the elementary damages in the DNA segment that are expected to behave as a single entity. Such entities—DNA lesions—arise as a result of energy deposits created by radiation tracks in small, of the order of 1–4 nm (*13*), isolated regions of DNA that are separated from each other by long segments of undamaged DNA. In the proposed damage simulation algorithm, the N_{min} parameter determines the manner in which elementary damages are grouped into lesions. That is, two elementary damages separated by at least N_{min} base pairs are treated as different lesions. Elementary damages separated by less than N_{min} base pairs are considered part of the same lesion (an MDS in this case). Figure 1 shows an idealized schematic illustrating the manner in which elementary damages are grouped into lesions. The value specified for N_{min} results in a unique grouping of elementary damages into lesions. Moreover, the proposed definition of a lesion guarantees that any elementary damage in a lesion is within N_{min} bp of another elementary damage. The proposed lesion identification (elementary damage grouping) algorithm is as follows:

1. Start at one end of the DNA segment and locate the first elementary damage on either or both strands. Set the start of the DNA lesion to the location of the elementary damage(s).
2. Starting with the base pair after the last identified elementary damage (upstream elementary damage), move along the DNA segment in the same direction, and count the number of undamaged base pairs present before the next (downstream) elementary damage is encountered. If the end of the DNA segment is reached before encountering another elementary damage, set the end of the DNA lesion to the location of the last detected elementary damage and quit.
3. If the number of undamaged base pairs is $\geq N_{min}$, set the end position of the lesion to the location of the upstream elementary damage. Then set the start position of the next lesion to the location of the downstream elementary damage.
4. Go to step 2.

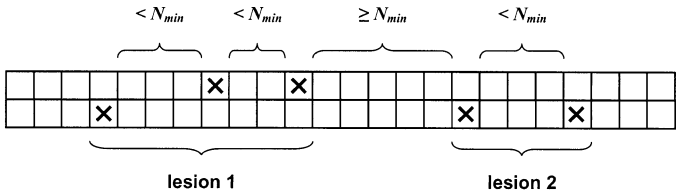

FIG. 1. Diagram illustrating how elementary damages are grouped into lesions. Each square cell represents a nucleotide. Crosses indicate the locations of elementary damages (strand breaks or base damages).

After all of the elementary damages in the DNA segment have been grouped into lesions, the properties of the lesions can be analyzed further in terms of the nature and spatial distribution of the elementary damages forming each lesion. For example, lesions can be grouped into various categories of simple and complex SSBs and DSBs according to the scheme proposed by Charlton and Humm (*14*).

Parameter Estimation

The Monte Carlo damage generation algorithm requires the specification of four adjustable parameters (σ_{Sb}, f, n_{seg} and N_{min}). To produce 1,000 SSBs and 40 DSBs Gy^{-1} $cell^{-1}$ (*1*), the number of strand breaks must be at least 1,080 Gy^{-1} $cell^{-1}$ (1,000 SSBs Gy^{-1} $cell^{-1}$ + 2 strand breaks per DSB × 40 DSBs Gy^{-1} $cell^{-1}$). Therefore, σ_{Sb} is $\sim 10^3$ Gy^{-1} $cell^{-1}$ for low-LET radiation. Ward (*15*) estimated that the base damage to strand break ratio, f, is 2.7. More recent measured data suggest that this ratio may be as high as 25 (*2*). The n_{seg} parameter and, to a lesser extent, the N_{min} parameter are used to adjust the amount of spatial clustering among adjacent elementary damages. For example, the amount of elementary damage clustering tends to increase as the value of n_{seg} decreases. On theoretical grounds, $n_{seg}D > N_{min}$ for all absorbed doses, even doses that approach 0 Gy. Moreover, practical considerations suggest that $n_{seg}D \gg N_{min}$. Otherwise, elementary damages that are many thousands or millions of base pairs apart will be scored as a single lesion.

To refine the estimates of σ_{Sb} and f and estimate appropriate values for the quasi-phenomenological parameters n_{seg} and N_{min}, we conducted a series of studies to identify the optimal model inputs for radiations of different quality. Nikjoo *et al.* (*10, 11*) reported a battery of values that describe different aspects of DNA damage complexity with respect to both strand breakage and base damage. Both low-LET (electrons) and high-LET (α particles and protons) radiations were considered. To obtain the best agreement between the values predicted using the proposed Monte Carlo scheme and the damage yields reported by Nikjoo and colleagues, we minimized a criterion defined as

$$C = \sum_i \frac{(O_i - E_i)^2}{E_i}, \tag{1}$$

where O_i is the yield of the ith type of DNA lesion obtained with the fast Monte Carlo algorithm and E_i is the yield of the ith type of DNA lesion obtained from the detailed Monte Carlo algorithm (*10, 11*).

The fitted parameters reported in this work are based on the absolute SSB and DSB yields (converted to units of Gy^{-1} $cell^{-1}$) reported in Table 2 of ref. (*11*) and the relative yields of different types of breaks (including information about base damage) reported in Table 1 of ref. (*11*): SSB, SSB+, 2SSB, DSB, DSB+, DSB++, SSB_c, SSB_{cb}, DSB_c and DSB_{cb}.

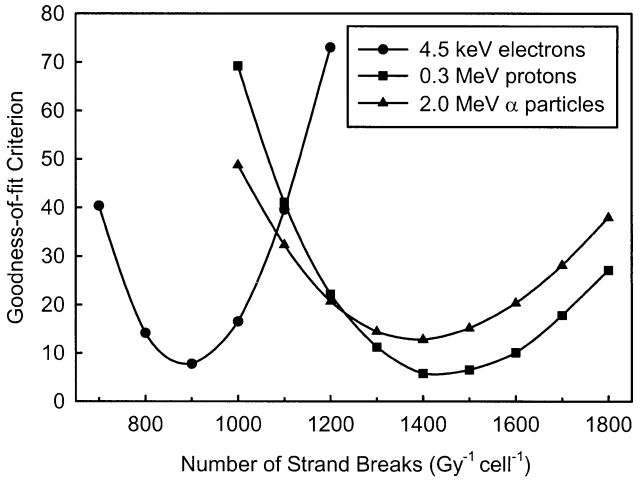

FIG. 2. Dependence of the goodness-of-fit criterion, *C*, on the number of strand breaks, σ_{Sb}. Ratio of base damages to strand breaks, *f*, is 3.0 for all radiations. For each data point, n_{seg} and N_{min} are such that *C* is minimized.

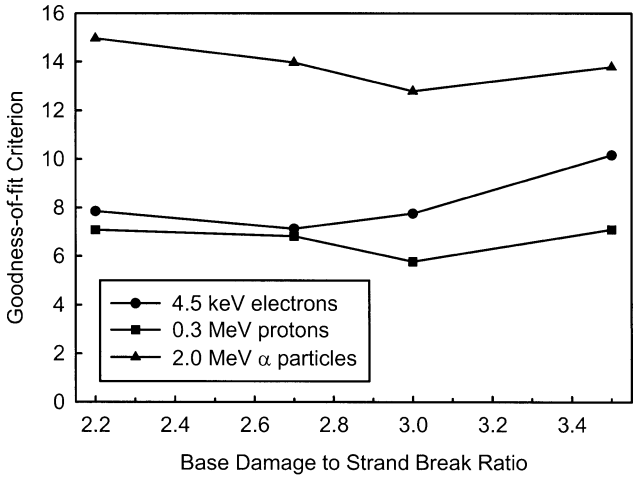

FIG. 3. Dependence of the goodness-of-fit criterion, *C*, on the ratio of base damages to strand breaks, *f*. The number of strand breaks is 900 for electrons and 1,400 for protons and α particles. For each data point, n_{seg} and N_{min} are such that *C* is minimized.

For 4.5 keV electrons, there appears to be an inconsistency in the SSB and DSB yields reported by Nikjoo *et al.* (*9–12*). Since we expect the SSB and DSB yields to decrease with increasing electron energy, we adopted the values reported in Table 4 of ref. (*10*) for 4.5 keV electrons instead of the corresponding values reported in ref. (*11*).

For a specific set of input parameters (σ_{Sb}, *f*, n_{seg} and N_{min}), lesions formed in one cell by a dose of 1 Gy were generated as described in the previous section. Each lesion was then categorized according to the break classification scheme of Nikjoo *et al.* (*9*). Absolute SSB and DSB yields (Gy^{-1} $cell^{-1}$), as well as percentage relative damage yields, were determined. A DSB was registered if at least one strand break was formed on each DNA strand within 10 base pairs, i.e. the same DSB identification scheme used by other investigators (*8–12*). The entire procedure was repeated for a large number of cells, and the damage yields were averaged.

To find the optimal combination of model parameters, the goodness-of-fit criterion (Eq. 1) was computed for a range of model inputs. The model inputs considered optimal for the simulation of DNA damage are the ones that minimize the goodness-of-fit criterion.

RESULTS

Figure 2 shows an example of the goodness-of-fit criterion, *C*, as a function of the number of strand breaks (Gy^{-1} $cell^{-1}$), σ_{Sb}. The criterion is very sensitive to the value of σ_{Sb}, and marked minima are observed for both low- and high-LET radiation. For low-LET radiation, the optimal value of σ_{Sb} falls within the range 800 to 1,000 strand breaks Gy^{-1} $cell^{-1}$. For both protons and α particles, the optimal value of σ_{Sb} is within the range from about 1,300 to 1,500 strand breaks Gy^{-1} $cell^{-1}$. Figure 3 shows trends in *C* for base damage to strand break ratios in the range from 2.2 to 3.5. Slightly better fits (smaller values of *C*) to the data of Nikjoo *et al.* (*10, 11*) are obtained with a base damage to strand break ratio of 2.7–3.0. In the Monte Carlo simulations of Nikjoo *et al.* (*10, 11*), the base damage to strand break ratio is ~2.2. A base damage to strand break ratio in the range from 2.7 to 3.0 is consistent with the lower bound of 2.7 estimated by Ward (*2, 15*). The trends in the goodness-of-fit criterion for other particle energies (not shown) are similar to those in Figs. 2 and 3.

Fixing the number of strand breaks (σ_{Sb}) at 1,400 Gy^{-1} $cell^{-1}$ and the ratio of base damages to strand breaks (*f*) at 3.0 for protons and α particles, the other two algorithm parameters (n_{seg} and N_{min}) were adjusted to minimize *C*. Figure 4 shows the optimal value of the DNA segment length (n_{seg}) for a range of proton and α-particle energies. The optimal DNA segment length increases in an approximately linear fashion as the particle energy increases. Figure 5 shows the dependence of n_{seg} on LET for protons and α particles. The optimal value of n_{seg} tends to decrease as LET, and correspondingly the amount of spatial clustering among elementary damages, increases. Figure 6 shows the best-fit value of the N_{min} parameter as a function of proton and α-particle energy. As the data shown in Fig. 6 illustrate,

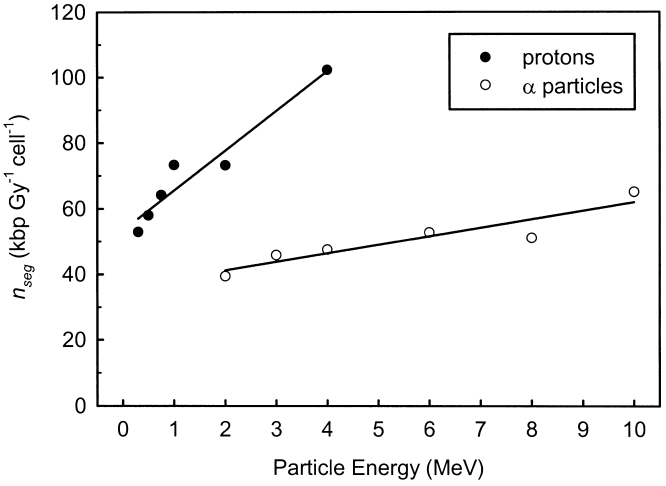

FIG. 4. Optimal values of the segment length parameter, n_{seg}, as a function of proton and α-particle energy. Error bars expressed as one standard deviation are generally smaller than the symbols (not shown). Linear regression lines for each data set are shown.

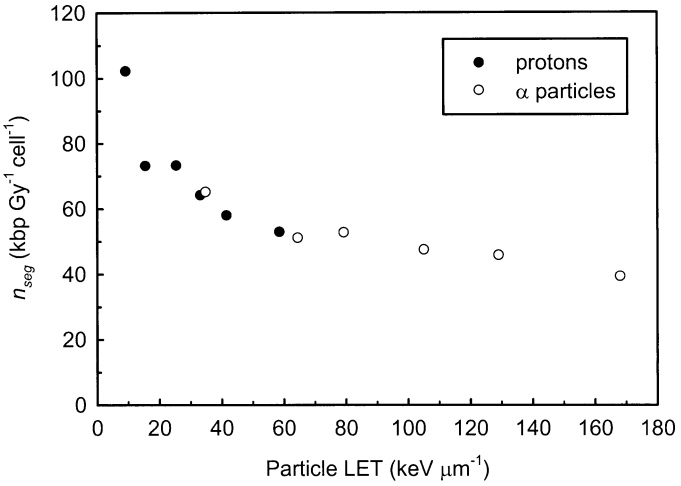

FIG. 5. n_{seg} parameter for protons and α particles as a function of particle LET.

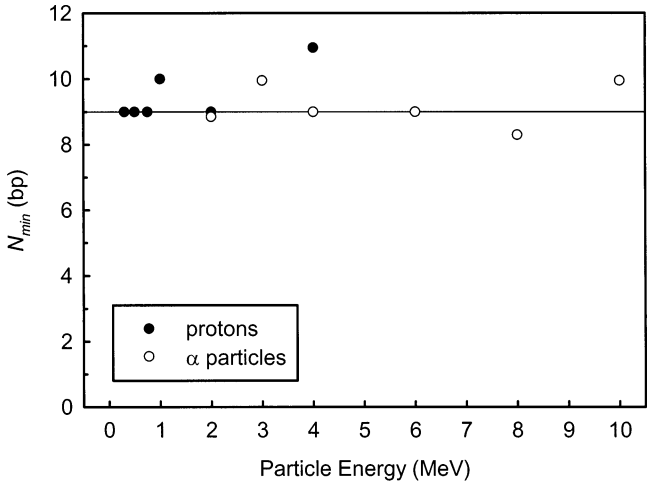

FIG. 6. Optimal values of the N_{min} parameter as a function of proton and α-particle energy. Error bars expressed as one standard deviation are generally smaller than the symbols (not shown). The line at 9 bp is shown for reference.

the same energy-independent value can be used to group elementary damages into lesions for protons and α particles. The optimal value of N_{min} for high-LET radiation is 9 bp (see reference line in Fig. 6). The data shown in Figs. 4–6 suggest that the n_{seg} parameter is the key model input that determines how elementary damages are clustered together in the DNA.

For high-LET radiations, our analyses suggest that the optimal model inputs are:

$$\begin{aligned} \sigma_{Sb}\ (\mathrm{Gy}^{-1}\ \mathrm{cell}^{-1}) &= 1{,}400 \\ f &= 3.0 \\ n_{seg}\ (\mathrm{bp\ Gy}^{-1}\ \mathrm{cell}^{-1}) &= \begin{cases} 12{,}125 \cdot T + 53{,}372 \\ \quad \text{for 0.3 to 4 MeV protons} \\ 2{,}602 \cdot T + 36{,}005 \\ \quad \text{for 2 to 10 MeV } \alpha \text{ particles} \end{cases} \\ N_{min}\ (\mathrm{bp}) &= 9 \end{aligned} \tag{2}$$

where T is the particle energy (MeV). Nikjoo *et al.* (*10, 11*) reported damage configurations generated by electrons with energies from 0.3 to 4.5 keV. Because very low-energy electrons exhibit effects similar to those of high-LET radiation (*16*), we consider the optimal model inputs for 4.5 keV electrons (the highest energy for which the full damage spectrum is available) to be representative of those expected from low-LET radiation, including X rays and γ rays. For sparsely ionizing radiation, the optimal model inputs are:

$$\begin{aligned} \sigma_{Sb}\ (\mathrm{Gy}^{-1}\ \mathrm{cell}^{-1}) &= 900 \\ f &= 3.0 \\ n_{seg}\ (\mathrm{bp\ Gy}^{-1}\ \mathrm{cell}^{-1}) &= 70{,}000 \\ N_{min}\ (\mathrm{bp}) &= 8 \end{aligned} \tag{3}$$

Table 1 shows a comparison of damage configurations predicted using the fast and detailed Monte Carlo damage generation algorithms. The proposed (fast) Monte Carlo algorithm captures the major trends in the DNA damage spectra predicted by the detailed track structure simulations (*10, 11*). Although the overall agreement between the two data sets is quite good, some systematic differences do exist. For example, the fast Monte Carlo scheme overestimates the percentage yield of simple SSBs by ~25–40% and underestimates the DSB+ yields by ~40–55%. The reasons for these differences are not entirely clear. These differences may be an indication that the spatial distribution of damage sites has some non-random characteristics (e.g. associated with the structure of the DNA or chromatin). Some of these differences may also be attributed to uncertainties associated with the detailed Monte Carlo simulations of Nikjoo *et al.* (*10, 11*). That is, accurate estimates of low-frequency events (e.g. the yield of SSB+, DSB+ and DSB++ lesions) are notoriously difficult to obtain using analog Monte Carlo simulations.

To gain some additional insight into the nature of the damage configurations generated using the fast Monte Carlo algorithm, we investigated the characteristics of the damage sites produced by selected radiations. Figure 7 shows representative distributions of the number of elementary damages per lesion. The mean number of elementary damages per lesion increases as the particle LET increases [4.5 keV electrons (~0.2 keV μm^{-1}), 1.5; 0.3 MeV protons (58.5 keV μm^{-1}), 2.5; 2.0 MeV α particles (168 keV μm^{-1}), 3.7]. Furthermore, the ratio of MDS (≥2 elementary damage per lesion) to singly damaged sites (1 elementary damage per lesion) also increases with increasing LET (electrons, 0.5; protons, 1.5; α particles, 2.7).

In Fig. 8, distributions of lesion lengths are shown for the same types of radiation as in Fig. 7. Lesion length is defined as a distance (in bp) between the first and the last

TABLE 1
Comparison of DNA Damage Yields Obtained Using the Fast and Detailed Monte Carlo Algorithms

	4.5 keV electrons		0.3 MeV protons		2.0 MeV α particles	
End point	Track structure[a]	This work[b]	Track structure	This work	Track structure	This work
SSB yield (Gy^{-1} $cell^{-1}$)[c]	741.0	747.2	858.0	879.0	663.0	675.8
DSB yield (Gy^{-1} $cell^{-1}$)[c]	46.8	45.9	136.5	130.6	156.0	158.5
SSB (%)	24.1	29.9	26.5	32.9	23.0	32.9
SSB+ (%)	2.3	1.8	7.1	5.2	7.0	8.2
2 SSB (%)	0.3	0.2	1.5	1.7	2.0	3.6
DSB (%)	1.5	1.6	4.8	3.5	4.8	4.6
DSB+ (%)	0.5	0.3	3.6	1.6	6.2	3.0
DSB++ (%)	0.04	0.05	1.3	0.8	5.7	2.8
SSB_c (%)	9	6	24	17	28	26
SSB_{cb} (%)	37	49	70	75	75	85
DSB_c (%)	29	17	51	40	73	56
DSB_{cb} (%)	62	62	90	87	96	94

[a] Calculations of Nikjoo *et al.* (*10, 11*).
[b] Damage yields are based on the optimal model inputs reported in the main text (Eqs. 2 and 3).
[c] SSB and DSB yields were converted into the units of Gy^{-1} $cell^{-1}$ using the factor 3.9×10^{12} Da $cell^{-1}$ (650 Da $bp^{-1} \cdot 6 \times 10^9$ bp $cell^{-1}$).

elementary damage in the lesion, inclusive. For 4.5 keV electrons, the average lesion length is 3.2 bp. The average lesion lengths for protons (0.3 MeV) and α particles (2.0 MeV) are 7.4 bp and 11.4 bp, respectively. The apparent plateaus at about 10 bp are due to the lesions (MDS in this case) with two elementary damages per lesion. Such lesions are the most abundant among MDS (refer to Fig. 7) and their lengths are distributed uniformly between 1 and (N_{min} + 1) bp (the latter is the maximum lesion length for MDS consisting of two elementary damages), thus forming a fine structure in the lesion length distributions. The data shown in Figs. 7 and 8 indicate that the fast Monte Carlo algorithm predicts that DNA lesions become more complex as particle LET increases, as expected.

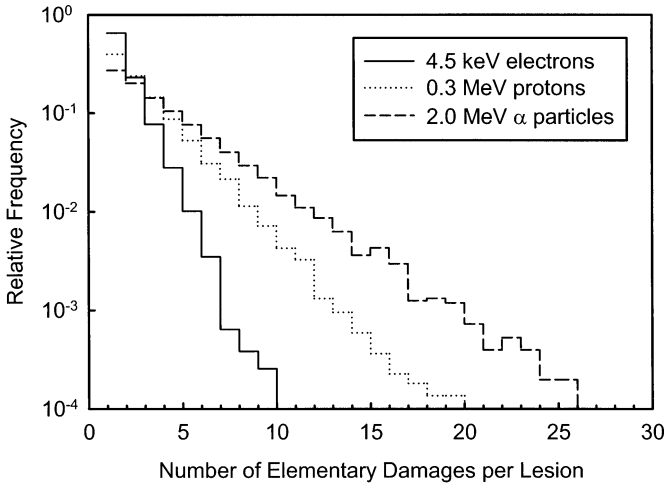

FIG. 7. Distributions of number of elementary damages per lesion for three radiations of different quality. Areas under each histogram are normalized to unity.

DISCUSSION

Small-scale clustering of DNA damage sites is one of the hallmarks of ionizing radiation. In contrast to ionizing radiation, for example, endogenous processes create mainly isolated damage sites. The proposed Monte Carlo scheme provides a simple and fast algorithm to simulate the formation of singly and multiply damaged DNA sites, including various classes of SSBs and DSBs. The success of this algorithm (Table 1) indicates that small-scale damage clustering effects can be simulated without explicitly considering the structure of the DNA, the higher-order folding of the chromatin, or the spatial distribution of energy deposits created along a radiation track. This observation is consistent with the hypothesis that the physical and chemical processes responsible for creating local clusters of elementary

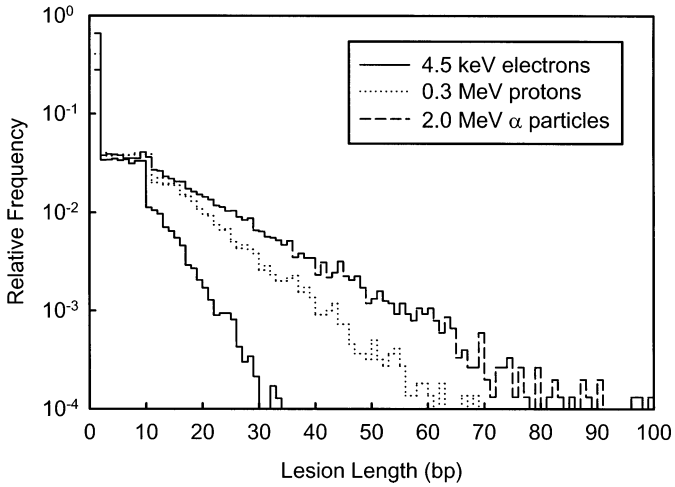

FIG. 8. Distributions of lesion length for three radiations of different quality. Areas under each histogram are normalized to unity.

damages (lesions) are initiated mainly by energy deposits created within a few nanometers of the DNA (*13*). However, even in these nanometer-sized regions, different types of radiation produce slightly different spatial distributions of energy deposits. These effects are reflected in the radiation-specific values of the σ_{Sb}, f, n_{seg} and N_{min} parameters (see Eqs. 2 and 3).

Although the proposed Monte Carlo algorithm successfully reproduces many of the small-scale damage clustering effects predicted by track structure codes, the algorithm provides no information about where the lesions are located within the DNA of a cell. To mimic the spatial distribution of lesions expected in a specific chromosome or cell, a lesion placement algorithm needs to be specified. For example, the lesions generated using the proposed Monte Carlo scheme could be placed at random locations in the DNA. For low-LET radiation, distributing lesions at random within the DNA of a cell is quite reasonable. As a first approximation, the random placement of lesions within the DNA could also be used for high-LET radiation. However, high-LET radiation produces non-random break distributions (see ref. *17* for a review), which implies that the distribution of all classes of lesions may also be non-random. To mimic the non-random spatial distribution of lesions produced by high-LET radiation will most likely require formulating a higher-level Monte Carlo scheme that accounts for the organization and structure of the chromatin as well as the overall structure of radiation tracks.

Small-scale (non-random) damage clustering effects are introduced into the damage simulation algorithm using two *ad hoc* parameters (i.e. the n_{seg} and N_{min} parameters). Although n_{seg} and N_{min} are specified in units of base pairs, no biophysical significance should be assigned to either of these parameters. For protons and α particles with a wide range of energies, the optimal value of N_{min} is 9 bp (see Fig. 6). For 4.5 keV electrons, the optimal value is 8 bp (see Eq. 3). These data suggest that the optimal value of N_{min} may be independent or nearly independent of the type and energy of the radiation. The optimal value of the n_{seg} tends to increase with increasing particle energy (Fig. 4). The same value for the n_{seg} parameter can be used for protons and α particles that have the same LET (Fig. 5). The readily discernible trends shown in Figs. 4–6 suggest that the optimal values of the n_{seg} and N_{min} parameters may be related to the nanometer-scale structure of radiation tracks. The exact nature of this relationship remains to be elucidated, and damage configurations for proton and α-particle energies outside the range indicated in Eq. (2) should be considered provisional.

The production of strand breaks and base damages is linear with absorbed dose up to at least a few hundreds of grays (*2, 18*). The proposed algorithm is premised on this linear relationship. Implicit in the damage generation algorithm is that (1) the spectra of produced damages do not change with absorbed dose and (2) the SSB and DSB yields per cell increase in direct proportion to the absorbed dose. The latter is well substantiated by data reported in the literature. For mammalian cells with a DNA content of about 6,000 Mbp per cell, the initial number of DSBs per cell increases linearly with increasing dose up to at least 50 Gy (*19*). In yeast cells, which have a DNA content about 250-fold smaller than that of a mammalian cell, the initial DSB yield per cell increases linearly with dose up to at least 2400 Gy (*20*). For mammalian cells, the proposed Monte Carlo damage generation algorithm is thus valid up to doses of at least a few tens of grays.

The proposed algorithm also has a lower dose limit. To give meaningful results, the algorithm requires that $n_{seg}D \gg N_{min}$. It follows that dose, D, should be much larger than the ratio N_{min}/n_{seg}. The maximum value of this ratio occurs for 2.0 MeV α particles, i.e. $N_{min}/n_{seg} \approx 2 \times 10^{-4}$ Gy. Therefore, we expect that the damage generation algorithm will be accurate for doses as low as $\sim$1 cGy.

The reported best-estimate parameters (Eqs. 2 and 3) are considered appropriate for an average mammalian cell. For low-LET radiation, the reported parameters result in about 46 DSBs Gy^{-1} $cell^{-1}$ (see Table 1). The dimensionless scale factor, g, can be used to adjust the DNA damage yields to better reflect the spectrum of DNA damage expected in a particular type of cell (see Methods). If, for example, the DSB yield (Gy^{-1} $cell^{-1}$) is known for a specific cell line, the factor g can be estimated as the ratio of the experimental DSB yield to 46 Gy^{-1} $cell^{-1}$. The DSB yield is the most frequently encountered piece of information that can be used to estimate the scaling factor g. However, other information, such as SSB yield, can also be used to estimate g. Among eukaryotes, DSB formation is approximately proportional to the DNA content of a cell. This empirical observation becomes readily apparent if DSB yields among the lower and higher eukaryotes are expressed in units of Gy^{-1} bp^{-1} (*21*). To approximate damage yields among different types of eukaryotic cells, g could be set equal to the ratio of the DNA content of the cell of interest to the DNA content of an average mammalian cell ($\sim 6 \times 10^9$ bp). In the absence of any cell-specific information, the factor g should be set to unity.

Although the proposed algorithm reproduces the general trends in the DNA damage spectrum predicted using track structure simulations (*10, 11*), questions remain about the spatial arrangement of elementary damages within one or two turns of the DNA. Ward *et al.* (*22*) estimated that, in the case of low-LET radiation, elementary damages within an MDS can be up to 30 bp apart. Taking into account nucleosome periodicity, MDS up to 80 bp long are possible (*1*). These lesion lengths are in good agreement with the distributions shown in Fig. 8, which provides further confidence in the proposed algorithm and parameter estimates.

CONCLUSIONS

A fast and easy-to-implement method of generating random DNA damage configurations has been shown to re-

produce the major trends in the spectrum of DNA damage that is expected from more detailed, but computationally expensive, Monte Carlo track structure simulations. The good agreement between the fast and detailed Monte Carlo damage generation algorithms suggests that the small-scale spatial distribution of elementary damages is determined primarily by independent and purely stochastic events and processes. That is, all of the nucleotides within a few turns of the DNA are equally likely to be damaged by radiation. The generation of DNA damage configurations for 100,000 cells exposed to 1 Gy of radiation only takes about 1.5 min on a 2.8 GHz Intel® Pentium III Xeon computer. The amount of computer memory required to perform the simulations is nominal, i.e. less than 1.3 megabytes.[2]

Parameters suitable for radiations of different quality are reported in Eqs. (2) and (3). An attractive feature of the proposed algorithm is that only four adjustable parameters need to be identified to simulate the formation of DNA damage by radiation. Two of these parameters (σ_{Sb} and f) have a clear physical meaning, and their values can be restricted to a narrow range of values using data from the literature. The other two model inputs (n_{seg} and N_{min}) are *ad hoc* parameters that relate to the small-scale spatial clustering of elementary damages within lesions. Reasonable estimates for all parameters have been obtained as a function of radiation quality using previously published data on DNA damage spectra (*10, 11*). Additional work to identify appropriate parameters for other types of low-, intermediate- and high-LET radiations is desirable.

ACKNOWLEDGMENT

The work was supported by the Office of Science (BER), U.S. Department of Energy, Grant No. DE-FG02-03ER63541.

Received: September 10, 2003; accepted: November 17, 2003

REFERENCES

1. J. F. Ward, DNA damage produced by ionizing radiation in mammalian cells: Identities, mechanisms of formation, and reparability. *Prog. Nucleic Acid Res. Mol. Biol.* **35**, 95–125 (1988).
2. J. F. Ward, Radiation mutagenesis: The initial DNA lesions responsible. *Radiat. Res.* **142**, 362–368 (1995); Errata, *Radiat. Res.* **143,** 355 (1995).
3. J. F. Ward, W. F. Blakely and E. I. Joner, Mammalian cells are not killed by DNA single-strand breaks caused by hydroxyl radicals from hydrogen peroxide. *Radiat. Res.* **103**, 383–392 (1985).
4. D. T. Goodhead, Initial events in the cellular effects of ionizing radiations: Clustered damage in DNA. *Int. J. Radiat. Biol.* **65**, 7–17 (1994).
5. P. A. Jeggo, The fidelity of repair of radiation damage. *Radiat. Prot. Dosim.* **99**, 117–122 (2002).
6. M. Weinfeld, A. Rasouli-Nia, M. A. Chaudhry and R. A. Britten, Response of base excision repair enzymes to complex DNA lesions. *Radiat. Res.* **156**, 584–589 (2001).
7. M. E. Lomax, M. K. Gulston and P. O'Neill, Chemical aspects of clustered DNA damage induction by ionising radiation. *Radiat. Prot. Dosim.* **99**, 63–68 (2002).
8. W. R. Holley and A. Chatterjee, Clusters of DNA damage induced by ionizing radiation: Formation of short DNA fragments. I. Theoretical modeling. *Radiat. Res.* **145**, 188–199 (1996).
9. H. Nikjoo, P. O'Neill, D. T. Goodhead and M. Terrissol, Computational modelling of low-energy electron-induced DNA damage by early physical and chemical events. *Int. J. Radiat. Biol.* **71**, 467–483 (1997).
10. H. Nikjoo, P. O'Neill, M. Terrissol and D. T. Goodhead, Quantitative modelling of DNA damage using Monte Carlo track structure method. *Radiat. Environ. Biophys.* **38**, 31–38 (1999).
11. H. Nikjoo, P. O'Neill, W. E. Wilson and D. T. Goodhead, Computational approach for determining the spectrum of DNA damage induced by ionizing radiation. *Radiat. Res.* **156**, 577–583 (2001).
12. H. Nikjoo, C. E. Bolton, R. Watanabe, M. Terrissol, P. O'Neill and D. T. Goodhead, Modelling of DNA damage induced by energetic electrons (100 eV to 100 keV). *Radiat. Prot. Dosim.* **99**, 77–80 (2002).
13. D. J. Brenner and J. F. Ward, Constraints on energy deposition and target size of multiply damaged sites associated with DNA double-strand breaks. *Int. J. Radiat. Biol.* **61**, 737–748 (1992).
14. D. E. Charlton and J. L. Humm, A method of calculating initial DNA strand breakage following the decay of incorporated ^{125}I. *Int. J. Radiat. Biol.* **53**, 353–365 (1988).
15. J. F. Ward, Biochemistry of DNA lesions. *Radiat. Res.* **104** (Suppl.), S103–S111 (1985).
16. D. E. Charlton, The range of high LET effects from ^{125}I decays. *Radiat. Res.* **107**, 163–171 (1986).
17. K. M. Prise, M. Pinto, H. C. Newman and B. D. Michael, A review of studies of ionizing radiation-induced double-strand break clustering. *Radiat. Res.* **156**, 572–576 (2001).
18. M. Frankenberg-Schwager, Induction, repair and biological relevance of radiation-induced DNA lesions in eukaryotic cells. *Radiat. Environ. Biophys.* **29**, 273–292 (1990).
19. K. Erixon and B. Cedervall, Linear induction of DNA double-strand breakage with X-ray dose, as determined from DNA fragment size distribution. *Radiat. Res.* **142**, 153–162 (1995).
20. M. Frankenberg-Schwager, D. Frankenberg, D. Blöcher and C. Adamczyk, Effect of dose rate on the induction of DNA double-strand breaks in eukaryotic cells. *Radiat. Res.* **87**, 710–717 (1981).
21. K. M. Prise, G. Ahnström, M. Belli, J. Carlsson, D. Frankenberg, J. Kiefer, M. Löbrich, B. D. Michael, J. Nygren and B. Stenerlöw, A review of dsb induction data for varying quality radiations. *Int. J. Radiat. Biol.* **74**, 173–184 (1998).
22. J. F. Ward, G. D. D. Jones and J. R. Milligan, Biological consequences of non-homogenous energy deposition by ionising radiation. *Radiat. Prot. Dosim.* **52**, 271–276 (1994).

[2] A sample computer code implementing the MCDS (Monte Carlo damage simulation) algorithm is available at http://healthsciences.purdue.edu/faculty/rds/mcds.html.