# Geant4-DNA simulation of DNA damage caused by direct and indirect radiation effects and comparison with biological data.

Carmen Villagrasa[1,a], Sylvain Meylan[1], Geraldine Gonon[1], Gaëtan Gruel[1], Ulrich Giesen[2], Marta Bueno[1] and Hans Rabus[2]

[1] *Institut de Radioprotection et Sûreté nucléaire (IRSN), BP-17, 92262-Fontenay-aux-Roses, France;*
[2] *Physikalisch-Technische Bundesanstalt (PTB), D-38116 Braunschweig, Germany*

**Abstract.** In this work we present results obtained in the frame of the BioQuaRT project. The objective of the study was the correlation between the number of radiation-induced double strand breaks (DSB) of the DNA molecule and the probability of detecting nuclear foci after targeted microbeam irradiation of cells with protons and alpha particles of different LET. The former were obtained by simulation with new methods integrated into Geant4-DNA that permit calculating the number of DSB in a DNA target model induced by direct and indirect radiation effects. A particular focus was laid in this work on evaluating the influence of different criteria applied to the simulated results for predicting the formation of a direct SSB. Indeed, these criteria have an important impact on the predicted number of DSB per particle track and its dependence with LET. Among the criteria tested in this work, the case that a direct radiation interaction leads to a strand break if the cumulative energy deposited in the backbone part of one nucleotide exceeds a threshold of 17.5 eV leads to the best agreement with the relative LET dependence of number of radiation induced foci. Further calculations and experimental data are nevertheless needed in order to fix the simulation parameters and to help interpreting the biological experimental data observed by immunofluorescence in terms of the DSB complexity.

## 1 Introduction

In cancer treatment using ionising radiation (radiotherapy), dosage is quantified by the absorbed dose to water expressed in the derived SI unit gray (Gy). Several radiotherapy modalities such as proton and ion beams require an additional weighting factor to account for increased tumor cell mortality for a given absorbed dose (as compared to a high-energy photon beam). The growing use of radiotherapy modalities whose biological effectiveness differ from that of conventional high-energy photon beams raises the need for establishing a new dosimetric concept expressed in terms of new quantities, which allow a transparent separation of the physical processes (dependent on modality) from the biological ones (independent of modality).

The joint research project "Biologically weighted quantities in radiotherapy" (BioQuaRT [1], 2012-2015) aimed at laying the foundation for these biologically-weighted radiation quantities based on a multi-scale approach. This approach used measurement and simulation techniques to determine the physical properties of ionising particle tracks on different length scales (from about 2 nm to 10 μm), which were then correlated with biological effects of radiation.

Presented here are the development of simulation tools and the quantitative comparison of the simulation results with experimental radiobiological data. The aim of this study was to evaluate the probability of structural DNA damage originating from particle irradiation as a function of the particle energy. From the simulations, the pattern at nanometric scale of the energy deposited by protons and alpha particles of different energies was obtained and different criteria for these energy depositions to break the DNA backbone were tested. The corresponding experimental biological data on early DNA damages were produced by targeted irradiation of cells at an ion microbeam and subsequent analysis of the formation of radiation-induced nuclear foci, such as γ-H2AX and 53BP1, observed by immunofluorescence.

## 2 Materials and methods

The evaluation of radiation-induced DNA damage in this work considers single strand breaks (SSB) and double strand breaks (DSB) of the DNA molecule. Monte Carlo simulations were used to compute both directly and indirectly induced breaks. An important aspect in this evaluation is the geometrical description of the DNA target. Indeed, the separation between neighbouring ionisations in the track of a densely ionising particle (such as low-energetic secondary electrons or ions in the Bragg peak region) is in the order of some nanometers and corresponds to the same order of magnitude as the size of the different molecules constituting the DNA. Hence, it is necessary to use a realistic description of the

---

[a] Corresponding author: Carmen.villagrasa@irsn.fr

target at nanometric scale in order to properly determinethe energy depositions that are located in the backbone region of the DNA which can give rise to direct single strand breaks (direct SSB).

On the other hand, a molecular description of the different components of the DNA was also needed in this work to extend the simulation beyond the physical interactions, i.e. in order to include the chemical reactions between the radicals produced by radiolysis of water in the vicinity of the target and the DNA. These chemical reactions could also break the structure of DNA and are accounted as indirectly induced breaks (indirect SSB). Both types of damages, direct and indirect SSB were analysed using the DBSCAN clustering algorithm [2] in order to determine the relative distance between them and to calculate the number of predicted DSBs to be compared to the biological experimental data. More precisely, the simulations modeled the ionising particle interactions on a virtual phantom of a chromatin fiber. The results obtained in the simulation were then scaled to the whole DNA (~ 6 Gbp) contained in the nucleus of a HUVEC (Human umbilical vein endothelial) cell of similar shape and volume as those HUVEC cells used in the radiobiological microbeam experiments. Then, the expected number of radiation-induced foci (RIF) was calculated from the number of DSB per track by taking into account the characteristics of the microbeam experiment. Finally the results predicted for the different radiation qualities were compared to the respective experimental radiobiological data.

## 2.1 Cell irradiation using a microbeam and DNA damage detection

Primary HUVEC cells cultures, predominantly in G0/G1 phase of cell cycle, were exposed to ion microbeam irradiation using the facility at the Physikalisch-Technische Bundesanstalt (PTB) [3]. The microbeam cell irradiation system enables hitting individual cell nuclei with a pre-determined number of particles in a given pattern of irradiation. In this experiment, each cell nucleus was hit by 5 particle tracks placed at the corners and the center of a square of 4 µm side length. The middle of the square was positioned at the barycenter of the targeted cell nucleus. Ten or thirteen minutes after the irradiations, the cells were fixed and the immunofluorescence protocol was followed.

Four types of ion/energy combinations were used in the experiment: α particles of 8 MeV, 10 MeV and 20 MeV (nominal beam energy) with an LET in the center of the cell nucleus of about 160 keV/µm, 90 keV/µm and 37 keV/µm, respectively, as well as 3 MeV protons with an LET of about 23 keV/µm. The karyotype and the genomic stability of cells were evaluated to ensure an equivalent DNA content in all cells of a population.

The foci analysis was based on a platform of high-throughput microscopy, with a powerful and robust infrastructure necessary for a massive image analysis (Scan-R software, Olympus). It enables the efficient

measurement of numerous topological parameters on foci such as area, shape (i.e. circularity factor, elongation factor, etc.), relative positions in the cell nuclei, density, etc. With automated high-speed microscopy, the statistical evaluation of these measurements can be done in a large population of cell nuclei and provides an objective way to estimate a probability of foci formation related to a particle traversal.

## 2.2 Monte Carlo simulation of the early DNA damage for proton and alpha particle irradiation

### 2.2.1 Geometrical model for DNA target

The simulations were performed within a DNA fibre [4] of 34 nm diameter composed of 90 linked nucleosomes as depicted in figure 1. For each nucleosome, the histone proteins were represented by a cylinder of 2.4 nm radius and were surrounded by a DNA double helix composed of 200 pairs of nucleotides. Each of these nucleotides was composed of 3 adjacent spherical volumes of 0.091 $nm^3$, 0.060 $nm^3$ and 0.093 $nm^3$ representing the sugar, phosphate and base molecules, respectively. A liquid water shell of 12 molecules per nucleotide around the DNA structure was also taken into account as a part of the target for the direct effects calculations. Throughout the whole geometrical structure, liquid water cross-sections were used for the physical radiation interactions.
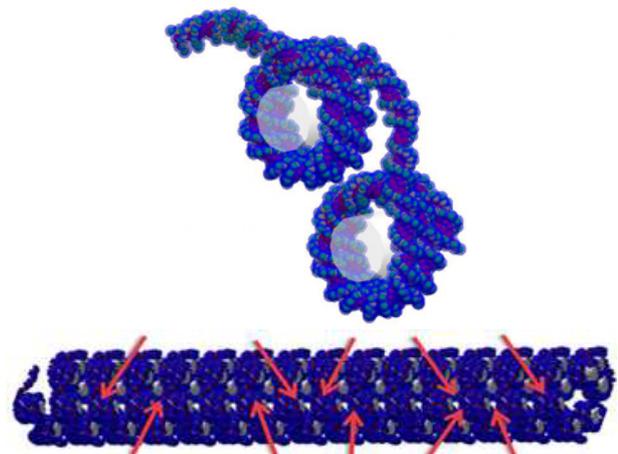


**Fig. 1.** Geometrical description of the DNA target used in this work. Upper part: (Two adjacent nucleosomes linked with the DNA double helix. Each nucleosome consists of a part of the DNA double helix wound around histone proteins (white cylinders). The DNA double helix is described at the molecular level. Lower part: 90 nucleosomes forming a 34 nm diameter chromatin fibre. Red arrows represent the random directions of the projectiles.

In the simulation, the fibre was surrounded by liquid water and was irradiated with helium or hydrogen ions impinging with random directions in order to simulate random positions of the chromatin fibre within the cell

nucleus. The results obtained in this target of 18,000 base pairs (bp) were then extrapolated to the total DNA content of a eukaryotic cell ($6\cdot10^9$ bp) by following the method described in Ref. [5] and taking into account the shape and mean-size characteristics of the irradiated HUVEC cell nuclei: cylinder with an ellipsoidal basis with semi-axes a=9.5 μm and b=5.5 μm, and a height h=2 μm.

### 2.2.2 Physical interactions and evaluation of direct SSB

The Monte Carlo track structure simulation was computed with the Geant4 code [6]. In particular, the Geant4-DNA [7,8] processes included in the Low Energy Package were used. These processes allow the transport of protons, alpha particles and all other charged states of hydrogen and helium as well as of electrons down to a few eV. All interactions with the target medium leading to either an energy loss or a non negligible change in the particle direction were taken into account.

The simulation of the interactions was purely discrete, calculating all the elementary interactions on an event-by-event basis down to the eV scale and not using a condensed history approximation. This method allows obtaining a nanometric description of the energy deposited by radiation in a target of liquid water material within a complex geometry as detailed below.

Each physical interaction leading to an energy deposition in one of the DNA elements during the Monte Carlo simulation was registered for further analysis.

In order to calculate from these data the number of strand breaks (SSB) that these interactions could cause, the following scenarios were considered:

1. Each ionisation located in a sugar or a phosphate volume is directly converted to a direct SSB ("all ionisations")
2. The energy deposited by different interactions in the sugar and the phosphate volumes of the same nucleotide is added up. Then a threshold value of 17.5 eV is applied. If the cumulated energy is higher that the threshold, a direct SSB is registered at that location of the DNA molecule [9] ("threshold")
3. The energy deposited in the nucleotide backbone is obtained as before but linear probability between 5 eV and 37.5 eV is applied to determine the SSB formation. No direct SSB is formed for an energy lower than 5 eV and an SSB is always formed for energies higher than 37.5 eV [10] ("linear probability")

### 2.2.3 Chemical interactions in the DNA target

The simulation of the radicals created after the ionisation or excitation of the water molecules around the DNA target was performed using the processes implemented in Geant4-DNA since December 2014 [8, 11]. The simulation takes into account the creation of 7 different radical species, their diffusion in liquid water and their chemical reactions. The implementation of the DNA target within this simulation was made by taking into account the position of each molecule composing the target: deoxyribose, phosphate, base and histone volumes. These positions were maintained fixed during the simulation of the chemical stage and new chemical reactions were included in the list of possible reactions given by default in the code (see Table 1). Therefore, in this simulation, an OH˙ radical was able to chemically interact with a deoxyribose molecule of the target giving raise to an indirect SSB in ~40% of the cases. The reason of this 40% probability comes from the fact that the deoxyribose molecular structure only leads to a break of the DNA backbone when a chemical attack occurs in certain parts of the molecule. Since our geometrical DNA model represents the whole deoxyribose structure as a sphere, we took that fact into account by applying a 40 % uniform random selection of the interactions leading effectively to a strand break. Chemical reactions between the OH˙ radical and the DNA bases were also simulated but they were never considered to lead to a strand break. Instead, they were considered as base damages and were not included in the comparison with biological data.

**Table 1.** Chemical reactions included in the simulation of the chemical stage performed with Geant4-DNA [11]. Reaction rate values were taken from [12]

| Chemical reaction | Reaction rate ($m^3$ $mol^{-1}\cdot s^{-1}$) |
|---|---|
| Deoxyribose + OH˙ | $2.5\cdot10^6$ |
| Adenine + OH˙ | $6.1\cdot10^6$ |
| Guanine + OH˙ | $9.2\cdot10^6$ |
| Thymine + OH˙ | $6.4\cdot10^6$ |
| Cytosine+ OH˙ | $6.1\cdot10^6$ |

### 2.2.4 Double strand breaks (DSB) evaluation and RIF simulation

Double strand breaks were calculated from computed single strand breaks obtained either by direct or indirect effect in the chemical stage. The simplest definition of a DSB is a DNA damage formed by two SSB separated by less than 10 bp. Nevertheless, in this work, the DBSCAN algorithm [2] was used to identify clustered DNA damages within the simulated SSB results. Therefore, the DSB presented here can be composed by two or more SSBs (also called complex DSB) only if all of them were separated by less than 10 bp from another SSB within the cluster.

In order to be able to compare the simulated results on DSB per track of a given projectile to the experimental biological data of probability of RIF formation per particle track, a simple assumption was made based on the characteristics of the microbeam irradiation. The

particles irradiating the cell nucleus crossed perpendicularly to the ellipsoidal surface (x-y plane) of the cell nucleus as shown in figure 2. While traversing the cells, the hydrogen and helium ions are only slightly deviating from this initial linear trajectory.

The image of experimentally detected RIF is obtained in the x-y plane and is not able to resolve two possible RIFs formed in very close (x,y) values along the z axis. For this reason, whatever the number of DSB created by the projectile (and their complexity) along the z axis and even though each of them could potentially lead to a visible RIF, the experimental image in the x-y plane will only show one single RIF. Therefore, the number of DSB per track obtained by Monte Carlo calculation was translated in a probability going from 0 to a maximum of 1 by considering the probability that at least one DSB is formed.
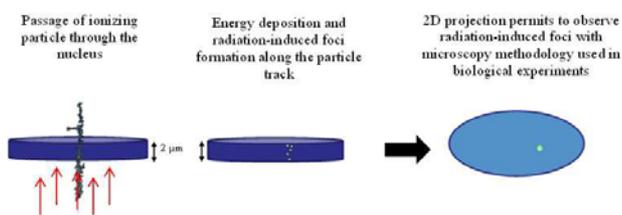


**Fig. 2.** Representation of the microbeam irradiation with ions traversing an endothelium cell nucleus. The experimental foci are obtained in a 2D projection losing information on the number of foci or DSB formation along the track.

# 3 Results and discussion

The mean number of DSBs per particle track traversing the endothelium cell nucleus was calculated using the three different criteria for the direct SSB formation. The method described in Ref. [5] to extrapolate the simulated data on a chromatin fibre to the whole trajectory within the cell nucleus results in a probability distribution on the number of traversed fibres with a mean number of 7.77. From this assumption, the results for the 4 radiation qualities in terms of mean number of DSB per track are given in table 2 with the uncertainty corresponding to the standard deviation of the distribution.

As it can be seen from these values, the criteria used for determining the formation of a direct SSB in the DNA backbone has a strong influence on both the absolute value of the number of DSB per track and also the dependence of DSB formation on LET. Moreover, these different criteria also have a high impact on the proportion between direct and indirect damages. Indeed, in the case that all ionisations in the backbone region are taken as direct breaks, the number of direct damages is higher than the number of indirect damages for all the LET values considered here: $SSB_{indirect}$ over $SSB_{direct}$ decreases with increasing LET from 0.9 at 23 keV $\mu m^{-1}$ to 0.5 at 160 keV $\mu m^{-1}$. On the contrary, for the linear probability or the energy threshold, the indirect damages are more frequent than the direct ones and the ratio $SSB_{indirect}$ over $SSB_{direct}$ decreases with increasing LET:

from 3.4 to 1.7 for the linear probability and from 16.5 to 3.2 for the threshold case.

**Table 2.** Simulation results on the mean number of DSB per track produced by the ion projectile when traversing the 2 µm thickness of the endothelium cell nucleus.

| | All ionisations | Threshold 17.5 eV | Linear probability (5 eV-37.5 eV) |
|---|---|---|---|
| Protons (23 keV $\mu m^{-1}$) | 0.73±0.11 | 0.14±0.05 | 0.20±0.08 |
| Alpha particles (37 keV $\mu m^{-1}$) | 1.01±0.11 | 0.30±0.09 | 0.45±0.01 |
| Alpha particles (90 keV $\mu m^{-1}$) | 4.6±0.3 | 1.17±0.13 | 1.88±0.17 |
| Alpha particles (160 keV $\mu m^{-1}$) | 14.8±0.5 | 5.3±0.2 | 6.9±0.2 |

Nevertheless, in order to calculate the probability of RIF formation in the cell nucleus per track, the important simulated result is the cumulative probability of obtaining at least 1 DSB in the track. Those results are presented in table 3 for the different criteria used in this work and compared to the experimentally observed 53BP1 foci in figure 3.

**Table 3.** Simulation results on the probability of RIF formation per particle track for the three different direct SSB formation criteria.

| | All ionisations | Threshold 17.5 eV | Linear probability (5 eV - 37.5 eV) |
|---|---|---|---|
| Protons (23 keV $\mu m^{-1}$) | 0.50 | 0.13 | 0.19 |
| Alpha particles (37 keV $\mu m^{-1}$) | 0.68 | 0.26 | 0.35 |
| Alpha particles (90 keV $\mu m^{-1}$) | 0.94 | 0.63 | 0.79 |
| Alpha particles (160 keV $\mu m^{-1}$) | 0.99 | 0.48 | 0.69 |

The different slopes in the increasing number of DSB with LET in the results presented in table 2 translate into different probabilities of RIF formation with LET. If all the ionisations are considered as responsible of a direct SSB, the probability of RIF formation increases rapidly with LET and for projectiles of LET of 90 keV/µm or higher, this probability is very close to unity. On the contrary, in the case of the threshold criterium or the linear probability criterium, the resulting probability of RIF formation shows a maximum around 100 keV/µm that is lower than 1 and that corresponds to what is also seen in the biological experimental data.

In particular, the threshold criterium of 17.5 eV [9] seems to be the one that better reproduces the relative dependence of the biological data on LET even though the absolute values deviate significantly.
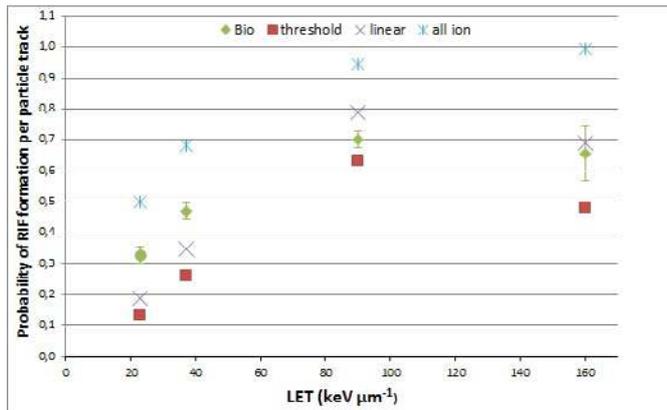


**Fig. 3.** Comparison of the probability of RIF formation simulated in this work for the three presented criteria on the direct SSB formation (blue stars, purple crosses and red squares) compared to the mean number of 53BP1 foci detected per particle track (green circle for protons and green diamonds for alpha particles).

The simulated results in this work are not given with associated uncertainty estimates. The reason for this is the difficulty of estimating these uncertainties owing to the simulation procedure being composed of several stages, each of them based on hypotheses and parameters that are still under investigation and do not offer a simple approach for estimating their associated uncertainty. Nevertheless, the main contributors to the uncertainty budget of the results presented here can be highlighted.

Concerning the absolute value of the number of DSB produced in the endothelium nucleus for a given track, the most important uncertainty comes from the scaling procedure adopted in this work to extend the simulated results in the chromatin fibre to the whole nucleus [5]. The chromatin domains size or the compaction of the chromatin fibre within the domains used in the procedure, can change the mean value (and the distribution) of the number of traversed fibres per track by more than 50%. For this reason, the relevant feature of the results presented in figure 3 is the dependency on LET and not their absolute value.

Besides the scaling procedure, the total uncertainty associated with the simulation performed in this work is the result of combining the uncertainties related to several factors: First, the calculation of physical interactions using Monte Carlo transport and liquid water cross-sections (especially for very low energy electrons). Second, the geometrical description of the DNA target at G0/G1 phase and the potential geometrical variability of the whole nucleus that is not taken into account here. Third, the limited number of chemical reactions taken into account and their reaction rates and probabilities. Fourth, the maximum time simulated for the chemical stage. Here the simulation time was limited to 2.5 ns in order to take into account the scavenging of the OH• radicals by other entities in the nucleus which are not

explicitly simulated in the present work. Finally, the statistical uncertainty related to the number of simulated projectiles (here 900 per energy), to cite the most important uncertainty sources.

The preceding is not an exhaustive list of parameters and hypotheses in the simulation that will influence the outcome of the calculated energy deposited in the DNA molecule during the physical stage and the mean number of chemical reactions between the radiation-induced radicals and the DNA.

In particular, in order to obtain from the simulated results a number of strand breaks for comparison with the experimental biological data, it is still necessary to determine which of these physical or chemical interactions will lead to a break. In this frame, and more specifically concerning the physical interactions, there is still debate within the scientific community, which is the reason why different criteria have been tested in this work. From the obtained results, using an energy threshold around 17.5 eV for the energy deposited in the backbone region of a nucleosome of our geometrical DNA model seems to give the best agreement with the experimental biological data obtained in this project.

# 4 Conclusions

Monte Carlo simulations using Geant4-DNA have been performed in order to calculate the number of DSB produced in a chromatin fibre target described with molecular detail. In these calculations, both, the physical stage and the chemical stage have been taken into account resulting in direct and indirect DNA damage. Concerning the direct effects, different criteria have been tested for the induction of a single strand break, based on the energy deposited in the backbone or the type of physical interaction. From these results, an evaluation of DSB in a HUVEC cell and the probability of detecting nuclear foci after irradiation with protons or alpha particles of different energy has been calculated. Besides all the hypotheses and parameters used in our simulation, the aforementioned criteria used to determine the formation of direct breaks of the DNA molecule have a great impact on the number of DSB and its dependence on the LET of the projectile. From the results presented in this work, the use of a threshold of 17.5 eV for the cumulated energy deposited in a single nucleosome in the DNA backbone seems to reproduce quite well the LET dependence of the biological data obtained in this work. Nevertheless, absolute values of the RIF formation probability simulated in this work do not coincide with the experimental data.

There are potentially two main reasons for explanation of this discrepancy: On the one hand, in this work a scaling procedure was used to extrapolate from the simulated results in a chromatin fibre to the whole cell nucleus. This could give rise to a significant bias in the results. In order to fix this potential problem, calculations are ongoing that simulate proton and alpha particle tracks in a geometrical model describing the complete DNA genome in the endothelium nucleus. The results of the very computer time consuming simulations

will eventually also be analysed using the criteria studied in this work.

On the other hand, one should also keep in mind that the measured 53BP1 foci, may not represent 1:1 all the DSB produced by the radiation. Therefore, the simulation method developed in this work also offers the possibility of investigating the complexity and position in the chromatin of the DNA damages that can correspond to those detectable by an immunofluorescence technique.

## Acknowledgements

## References

1. URL: http://www.ptb.de/emrp/bioquart.html
2. M. Ester, H-P. Kriegel, J. Sander, X. Xu. In: E. Simoudis, J. Han, U. M. Fayyad (Eds.). *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, AAAI Press, pp. 226–231 (1996)
3. K. D. Greif, H.J. Brede, D. Frankenberg, U. Giesen, Nucl Instrum Methods Phys Res **B217**, 505-512(2004)
4. S. Meylan, U. Vimont, S. Incerti, I. Clairand and C. Villagrasa. Comp. Phys and Comm. In press.
5. B. Lee, PhD thesis, Georgia Institute of Technology, 2014. URL: http://hdl.handle.net/1853/53106
6. Agostinelli, S. et al., Nucl. Instrum. Meth. Phys. Res. A **506** pp. 250-303(2003)
7. S. Incerti et al, Med. Phys. **37** pp. 4692-4708 (2010)
8. M. A. Bernal et al., Phys. Medica **31,** pp. 861-874 (2015)
9. D.T. Goodhead and H. Nikjoo, Int. J. of Rad. Biol. **55**, pp. 513-529 (1989)
10. W. Friedland, P. Jacob, P. Bernhardt, H. G. Paretzke and M. Dingfelder, Rad. Res. **159**, pp 401-410 (2003)
11. M. Karamitros et al., J. Comput. Phys. **274**, pp. 841-882 (2014)
12. B. Aydogan, W.E. Bolch, S.G. Swarts, J.E. Turner and D.T. Marshall, Rad. Res. **169**, pp. 223-231 (2008)