

# An *In Silico* Model of DNA Repair for Investigation of Mechanisms in Non-Homologous End Joining

---

**Authors:** John W. Warmenhoven<sup>1</sup>, Nicholas T. Henthorn<sup>1</sup>, Marios Sotiropoulos<sup>1</sup>, Nickolay Korabel<sup>4</sup>, Sergej Fedotov<sup>4</sup>, Randal I. Mackay<sup>3</sup>, Karen J. Kirkby<sup>1,2</sup>, Michael J. Merchant<sup>1,2</sup>

**Affiliations:** <sup>1</sup>Division of Cancer Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, UK; <sup>2</sup>The Christie NHS Foundation Trust, Manchester, UK; <sup>3</sup>Christie Medical Physics and Engineering, The Christie NHS Foundation Trust, Manchester, UK; <sup>4</sup>School of Mathematics, The University of Manchester, UK

## Abstract

In human cells, non-homologous end joining is the preferred process to repair radiation induced DNA double strand breaks. The complex nature of such biological systems involves many individual actions that combine to produce an overall behaviour. As such, experimentally determining the mechanisms involved, their individual roles, and how they interact is challenging. An *in silico* approach to radiobiology is uniquely suited for detailed exploration of these complex interactions and the unknown effects of specific mechanisms on overall behaviour. We detail the construction of a mechanistic model by combination of several, experimentally supported, hypothesised mechanisms. Compatibility of these mechanisms was tested by fitting to results reported in the literature. To avoid over fitting, individual mechanisms within this pathway were sequentially fitted. We demonstrate that using this approach the model is capable of reproducing published protein kinetics and overall repair trends. This process highlighted specific biological mechanisms which are not clearly defined experimentally, and showed that the assumed motion of individual double strand break ends plays a crucial role in determining overall system behaviour.

**Key Words:** DNA repair / Non-homologous End Joining / Mechanistic / Computational Modelling / Monte Carlo

**Characters (including spaces and references):** 62,711

## 1 Introduction

Radiotherapy is a well-established treatment modality for cancers, with substantial clinical experience in delivering palliative and radical photon based procedures. Proton beam therapy is a fast developing alternative modality which offers advantages for specific cancer sites due to the ions stopping a defined depth into the patient. Protons also produce a different biological effect for the same dose delivered, parameterised as the proton Relative Biological Effectiveness (RBE); defined as the ratio between proton and photon doses that are required to cause an equivalent amount of cell killing. Current clinical practice is to scale proton dose by a fixed RBE of 1.1 (DeLuca *et al*, 2007) .

However, experimental data shows a large variance in RBE depending on parameters describing both the tissue environment and the beam quality (Paganetti *et al*, 2002; Paganetti, 2014; IAEA & ICRU, 2008). Studies show that a variable RBE can increase not only the biologically effective range, but also the biologically effective dose at the distal edge of the proton beam (Marshall *et al*, 2016). Consequently, studies based on *in vitro* data have shown the possibility of these biological uncertainties to negatively impact clinical proton treatment plans (Wedenberg & Toma-Dasu, 2014; McNamara *et al*, 2016; Tilly *et al*, 2005; Carabe *et al*, 2012; Wedenberg *et al*, 2013). Tilly *et al*. and McNamara *et al*. demonstrate the possibility of increased normal tissue complication probability (NTCP) whilst Wedenberg *et al*. demonstrate the possibility of a lower tumour control probability (TCP). Carabe *et al*. show that there are clinical situations in which the relatively small biological range uncertainty is relevant. Development of future treatment planning systems should therefore include RBE optimisation to further exploit the inherent physical advantages of proton beam therapy (Giovannini *et al*, 2016; Cao *et al*, 2015). Computational models are an efficient way of probing various biological and physical mechanisms in order to illuminate their possible relationships to RBE.

Several *in silico* models of DNA damage and repair have been proposed that assume a random spread of breaks, homogeneous throughout the cell or centred around a track (Sachs *et al*, 2000; Ballarini & Ottolenghi, 2004; Eidelman *et al*, 2012). An alternative methodology is to analyse radiation track structure interactions within the cell nucleus (Bernal *et al*, 2015). This provides the specific geometrical distributions of double strand breaks within the nucleus caused by radiations of different types and qualities, therefore making it possible to investigate how changes in proximity of breaks affects the repair machinery (Moore *et al*, 2014). Furthermore, it also enables explicit inclusion of repair retarding complexities such as base lesions and additional single strand breaks (Schipler & Iliakis, 2013). Due to the small number of double strand breaks at low doses of ionising radiation, discrete stochastic models are most suited to simulating the repair of these damages (Friedland & Kundrát, 2013; Friedland *et al*, 2010b). Such models also allow for inclusion of specific mechanistic interpretations of the interactions between the repair machinery and the complexity or proximity of double strand breaks; something that is difficult with more top down descriptions. Therefore combining discrete, stochastic, mechanistic models with damage derived from detailed track structure simulations would produce a model suitable for studying the broadest range of radiation doses. Thorough investigation of the emergent behaviour of such a system could give insight into the operation of cellular repair machinery and subsequently into the experimentally observed variations in RBE.

DNA double strand breaks have multiple pathways of repair, the most dominant being Non-Homologous End Joining (NHEJ) (Chiruvella *et al*, 2013) and Homologous Recombination (HR) (Jasin & Rothstein, 2013). The non-homologous end joining pathway is available throughout the cell cycle, and is the only pathway available outside of the replication phases (Thompson, 2012; Brown *et al*, 2017). Furthermore, the most severely radio-sensitive phenotypes have mutations of non-homologous end joining genes (Thompson, 2012) showing that this pathway plays a dominating role in cell survival. Initial focus on this pathway is therefore justifiable and facilitates easy designing of *in vitro* verification experiments that isolates the action of non-homologous end joining. However, the complexity of biological systems, even when reduced to a single pathway, presents a significant challenge to *in vitro/vivo* experiments when attempting to unambiguously define mechanisms responsible for DNA repair. This has resulted in separate publication of many suggested mechanisms within the NHEJ pathway that are not contradicted by experimental evidence. These mechanisms, although

individually plausible, are not necessarily capable of reproducing observed behaviours of the entire cellular system when combined.

In this work we construct a composite system to model non-homologous end joining through the combination of several such published mechanisms. This mechanistic, Monte Carlo based model is developed within the Geant4-DNA tool-kit (Incerti *et al*, 2010; Bernal *et al*, 2015; Karamitros *et al*, 2014) which allows for event-by-event tracking of individual double strand break ends. Initial conditions were set from interfacing with a nano-dosimetric DNA damage model, directly linking radiation track structure to biological effect (Henthorn *et al*, 2017, 2018). The model produces as its output predictions of biologically measurable end points, including repair protein recruitment kinetics and double strand break repair kinetics, which are compared to published experimental data to assess the feasibility of the combined mechanisms.

## 2 Materials and Methods

### 2.1 Construction of Initial Model Conditions

The repair model takes as its input data in a standard DNA damage phase space format developed in collaboration with Queen's University Belfast and Harvard Medical School. Damage models can utilise this standard to populate a phase space that can be passed either in the form of an ASCII file or directly if the damage model is integrated into the Geant4 simulation. In this work we use simulation results from the model described by Henthorn *et al*. (Henthorn *et al*, 2017, 2018) as input. This nano-dosimetric track structure model provides the geometric locations and detailed structure of DNA double strand breaks after irradiation of a nucleus by a specified beam. Individual double strand breaks are processed and separated into two exposed DNA ends at the closest pair of backbone breaks on opposite strands. Remaining backbone and base damages are assigned to individual break ends based on their positions relative to the separation site.

### 2.2 Implementation of Double Strand Break Objects

We have implemented double strand breaks using the existing Geant4-DNA mechanism for handling chemistry objects; that is an object consisting of a track and associated molecule. This allows use of the reaction and diffusion functionality of the Geant4-DNA Chemistry module (Karamitros *et al*, 2014). To monitor the state of each individual double strand break object as the system evolves a new software class is implemented and associated to the track as auxiliary track information. This class is a data structure storing the number of single strand breaks and base lesions (see section 2.1), break ID and correct partner ID (see section 2.4), waiting and diffusion times (see section 2.3), and starting location (see section 2.5) associated with each break end object. There are also empty copies of these variables for use when two break end objects react to form one synaptic object (see section 2.4.3).

### 2.3 Motion

It has been experimentally demonstrated that double strand breaks move by sub-diffusion (Girst *et al*, 2013; Lucas *et al*, 2014; Miné-Hattab *et al*, 2017), and, although not explicitly concluded, there is experimental evidence that this applies to individual double strand break ends (Soutoglou *et al*, 2007). To the best of our knowledge this type of motion has not yet been explicitly included in any mechanistic *in silico* models of DNA repair. There are different microscopic mechanisms that lead to the emergence of sub-diffusive motion (Sokolov, 2012). In this work we have chosen to implement a

continuous time random walk (CTRW) model of sub-diffusion (for a detailed discussion of anomalous diffusion in general see Metzler *et al.* (Metzler & Klafter, 2000)). This represents the double strand break ends undergoing transient states of confined motion, such as would be expected in a system with rapid formation and dissociation of synaptic complexes (Graham *et al.*, 2016). To implement such a continuous time random walk model double strand break objects are assigned a waiting time drawn from:

$$\text{Waiting Time} = \frac{A}{R^{1/\alpha}} \quad (1)$$

Where R is a uniform random number between 0 and 1, A is the minimum waiting time and  $\alpha$  is the anomalous diffusion exponent. In this work the anomalous diffusion exponent was set to 0.5 and the minimum waiting time was set to 1 ms in order to generate the waiting times. Whilst a double strand break object has a waiting time associated it is trapped and cannot diffuse. Waiting times are tracked using the internal clock of the simulation and once reduced to 0 seconds the double strand break object is released. Diffusion then occurs through normal Brownian motion means of stochastic displacement by a distance drawn from a normal distribution. The distribution is described by the diffusion coefficient and the time the object is allowed to diffuse. In our model the diffusion time is set at 1 picosecond, which promotes a single 'jump' event without interference of any chemical processes. The 'jump diffusion coefficient' is taken to be the free fitting parameter to manipulate the overall scale of motion and fit to literature data. In this way we have implemented a mechanism which can lead to sub-diffusive motion. The mechanism is linked to transient confinement of double strand break ends such as could be explained by alternating association/dissociation processes (Saxton, 2007).

## 2.4 Time Evolution

### 2.4.1 Implementation of the Non-Homologous End Joining Process

Time evolution of individual double strand break ends is governed by a series of time constant based state changes (see section 2.4.2). The states are linked according to the scheme outlined in Figure 1. This represents the progressions through the canonical non-homologous end joining pathway.

State A represents an exposed double strand break end which may have associated single strand breaks and base lesions as discussed in section 2.2. Recruitment of the Ku70/80 heterodimer is widely reported in literature to be the first step in the canonical non-homologous end joining pathway (Ma *et al.*, 2004; Lieber, 2010; Chiruvella *et al.*, 2013; Wang & Lees-Miller, 2013; Brown *et al.*, 2017) (State B). Ku70/80 cannot dissociate from DNA ends until phosphorylated (mechanism proposed by Lee *et al.* (Lee *et al.*, 2015)). Double strand break ends may alternatively become temporarily inhibited towards recruitment of Ku70/80 (state X). The inhibition is a catch all state which represents competition from other proteins with affinities for exposed double strand break ends. Following attachment to a double strand break end, Ku70/80 recruits the DNA-Protein Kinase catalytic subunit (DNA-PKcs) to form the DNA-PK complex (state C). DNA-PKcs cannot dissociate from DNA ends until phosphorylated (mechanism proposed by Uematsu *et al.* (Uematsu *et al.*, 2007)). DNA-PK is required for the formation of synaptic complexes between proximal double strand break ends (Graham *et al.*, 2016) (state D, see section 2.4.3).

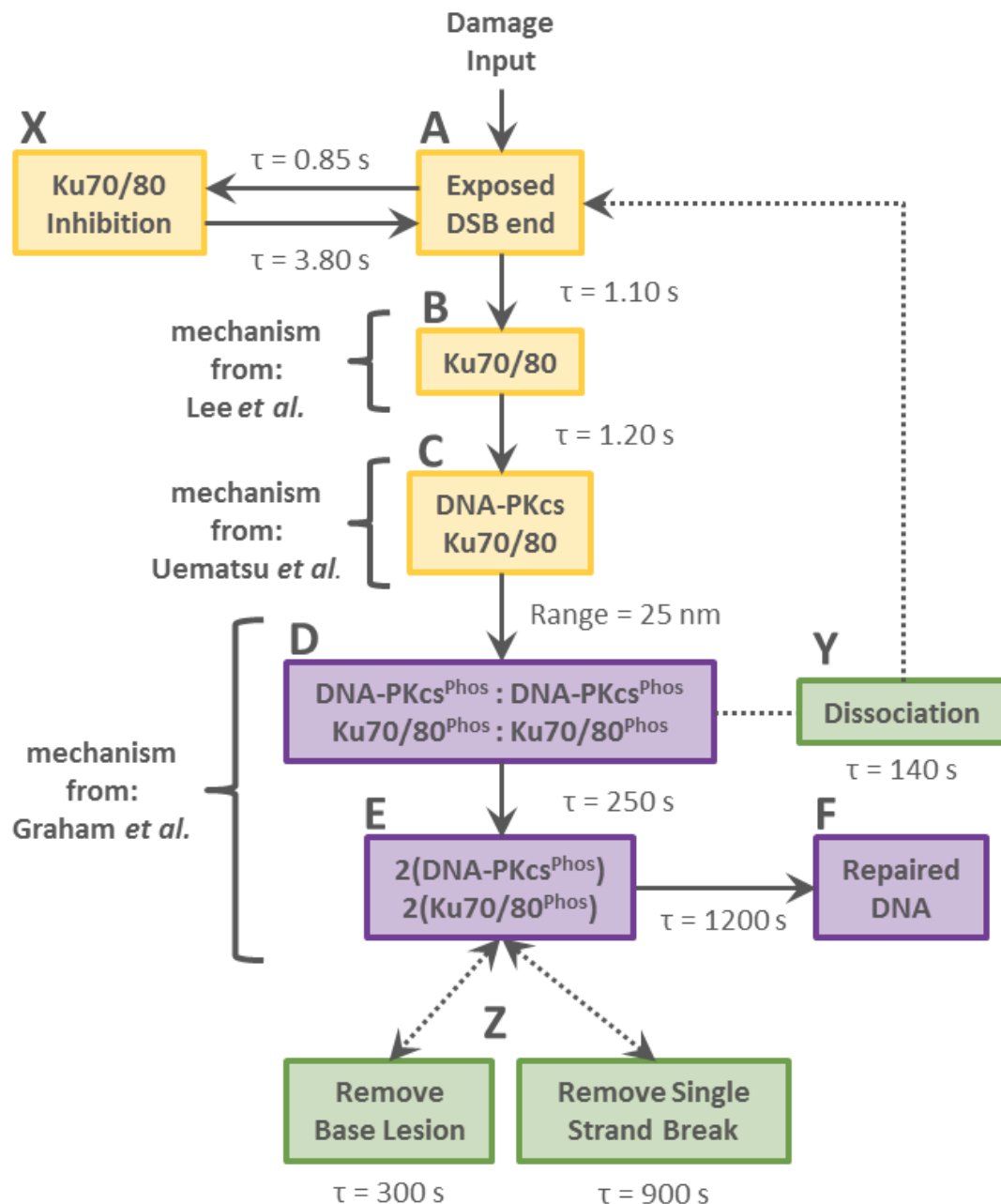


Figure 1: Scheme of the non-homologous end joining pathway implemented in this work with associated time constants. Boxes A-F and X represent the different states double strand break objects can occupy in the simulation, boxes Z and Y and arrows represent the different mechanisms which can change the state of a double strand break object in the simulation. States A-C and X represent a single double strand break end; states D-F represent a synaptic complex of two double strand break ends. Progress between states A-C, A to X, D-F, and D to A is governed by a time constant based stochastic process. Progress from states C to D is governed by a proximity based reaction which produces a single synaptic complex from two double strand break ends. Process Y splits this synaptic complex back into two individual double strand break ends. Also noted in the figure are the areas of influence of the three experimentally supported mechanisms combined by this work. Lee *et al.* suggested that Ku70/80 does not dissociate until phosphorylated, Uematsu *et al.* suggesting DNA-PKcs does not dissociate until phosphorylated, and Graham *et al.* showing formation of transient synaptic complexes between DNA-PK complexes.

Once in a synaptic complex DNA-PKcs autophosphorylates, inducing conformational changes necessary for further processing and allowing it to phosphorylate Ku70/80 (Uematsu *et al*, 2007; Lee *et al*, 2015). This initial "long range" synaptic complex is prone to dissociation (mechanism proposed by Graham *et al*. (Graham *et al*, 2016)) and due to the phosphorylated state of both Ku and DNA-PKcs the entire complex dissociates to state A (Uematsu *et al*, 2007; Lee *et al*, 2015) (process Y, see section 2.4.3). If the complex survives, progression to a stable "short range" synaptic complex follows (state E) (mechanism proposed by Graham *et al*. (Graham *et al*, 2016)). Processing of the break site occurs in state E, repairing all base lesions and extra DNA backbone lesions present (processes Z, see section 2.4.2) before allowing final ligation to state F.

#### 2.4.2 State Change

State changes are implemented as custom Geant4 processes that change the molecule definition associated with the track. Seven different definitions are created to represent the different states a double strand break object can occupy. These are states A-F and X from Figure 1. The time constants used in these processes represent both the attachment and action of specific repair proteins. Cleaning of complexities, state Z in Figure 1, are special cases of the above processes. On execution no change is made to the object state. Instead the number of base lesions or single strand breaks associated with the synaptic complex is reduced. The specific double strand break end this applies to within the complex is determined randomly. Transition from state E to F cannot successfully complete unless all associated base lesions and single strand breaks have been removed.

#### 2.4.3 Reactions and Failure after Synapse

Both DNA-PK mediated and XRCC4 mediated synaptic complex formation for complex and simple double strand breaks respectively have been reported (Reid *et al*, 2015; Graham *et al*, 2016; Li *et al*, 2014). In this work we have chosen to explicitly model only the DNA-PK mediated synaptic complex formation as suggested by Graham *et al*. (Graham *et al*, 2016) due to available experimental data to fit to. Synaptic complex formation is treated as a chemical reaction utilising the provided functionality of Geant4-DNA. Two DNA-PK objects within 25 nm (Friedland *et al*, 2010a) of each other react to form a single synaptic complex object. Outside of this range the reaction probability is governed by the Geant4-DNA diffusion controlled Brownian bridge process (Karamitros *et al*, 2014) modified to accommodate sub-diffusive motion. A reaction replaces two double strand break end objects with a single synaptic complex object. The state information of both double strand break end objects is stored together (see section 2.2) and associated with the new synaptic complex object. This preserves the information required for final data gathering (see section 2.5) and dissociation of the complex. Dissociation replaces the single synaptic complex object with two double strand break end objects and merging of the associated state information is reversed.

### 2.5 Data Gathering

Recruitment kinetics of involved repair proteins and overall double strand break repair kinetics are extracted by real time tracking of object states. Simulations are repeated in order for the stochastic state changes to converge on an average result. Residuals double strand breaks are scored as the initial number of double strand breaks reduced by the number of fully ligated 'fixed' breaks in the system at that time point. This represents all double strand break ends which have become isolated or are still undergoing repair. Data is normalised for each individual simulation and then combined over all repeats to give the average behaviour for a given beam quality. Displacement of all double strand break objects from their initial location is tracked in one second intervals for the first 300

seconds to calculate the mean squared displacement. Misrepair events are scored at the end of each individual simulation by comparing identities of double strand break ends in synaptic complexes.

## 3 Results

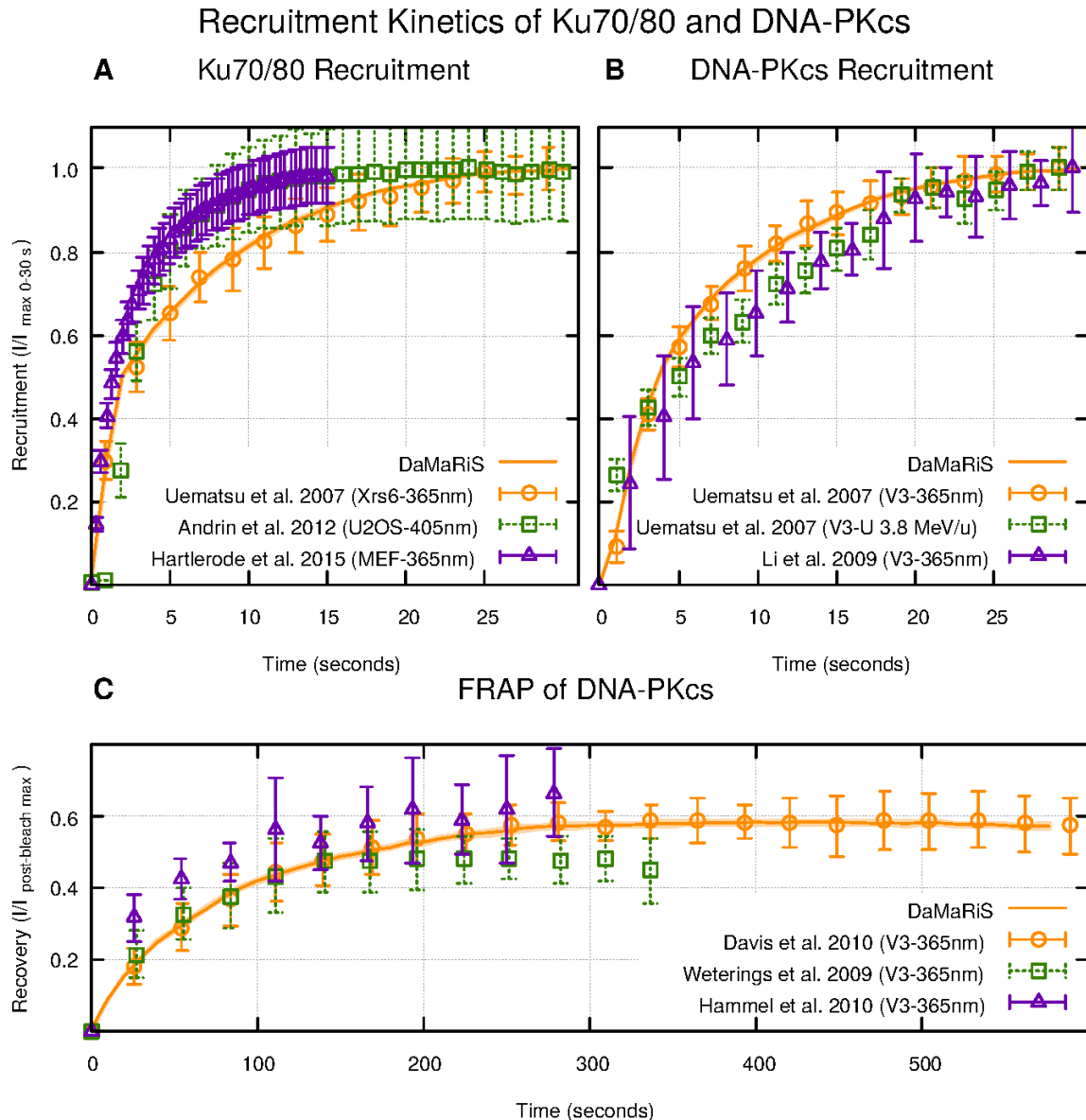
### 3.1 Recruitment Kinetics of DNA-PKcs and Ku70/80

The stoichiometry of Ku70/80 and DNA-PKcs is one each per double strand break end. A high density of breaks is therefore required for these proteins to form visible foci under most microscopy conditions. To meet this requirement Uematsu *et al.* (Uematsu *et al.*, 2007) reported on 365 nm pulsed nitrogen laser generated double strand breaks, stating that they created 2500-3700 per 1.7  $\mu\text{m}^2$  spot. These breaks are also considered to be complex in nature and thus require the DNA-PKcs mediated NHEJ pathway described in this work (Li *et al.*, 2014). Uematsu *et al.* also report experimental data from uranium ions accelerated to 3.8 MeV/u. The high LET (>12,800 keV/ $\mu\text{m}$ ) of these ions is expected to result in similarly highly clustered and complex damage.

Figure 2A and 2B shows good agreement between the simulated behaviour and literature reported experimental data from cell lines derived from Chinese hamster ovary (CHO) cell lines, namely Xrs6 and V-3, regardless of radiation used. Initial rapid recruitment of Ku70/80 and DNA-PKcs reaches 50% of maximum in 2-3 seconds before slowing, reaching 80% at ~10 seconds. Data from non-CHO cell lines show slightly more rapid recruitment of Ku70/80 which is not reproduced as well. Andrin *et al.* (Andrin *et al.*, 2012) use human bone osteosarcoma epithelial cells (U2OS) in their work and Hartlerode *et al.* (Hartlerode *et al.*, 2015) show similar Ku70/80 recruitment in mouse embryonic fibroblast stem cells (MEF). Fitting the DNA-PKcs recruitment to experimental data was trivial; requiring manipulation of only the state B to state C time constant (Figure 1) to produce a slight delay on the Ku70/80 recruitment resulting in the correct behaviour. Obtaining the correct behaviour of Ku70/80 recruitment was more involved, requiring manipulation of the 3 relevant time constants, with inhibition (A to X, Figure 1) playing an important role. Final time constants selected are Ku-inhibition = 0.85 seconds, release from inhibition = 3.8 seconds, Ku70/80 recruitment = 1.1 seconds, DNA-PKcs recruitment = 1.2 seconds.

Authors	Protein	Reduced $\chi^2$	Max / Average / $\sigma$ of Time Deviation (%)
Uematsu <i>et al.</i>	Ku70/80	0.020	0.74 / 0.17 / 0.19
Andrin <i>et al.</i>	Ku70/80	0.562	0.96 / 0.15 / 0.21
Hartlerode <i>et al.</i>	Ku70/80	0.069 (1.336)	0.72 / 0.19 / 0.21
Uematsu <i>et al.</i> (LASER)	DNA-PKcs	0.033	0.95 / 0.51 / 0.28
Uematsu <i>et al.</i> (Uranium)	DNA-PKcs	1.167	0.82 / 0.48 / 0.17
Li <i>et al.</i>	DNA-PKcs	0.577	0.92 / 0.34 / 0.31

**Table 1: Summary of the reduced  $\chi^2$  for the different experimental datasets compared to the simulated behaviour. Data from the different datasets had different measuring time-points and consequently exactly corresponding simulation time points were not available. Experimental data was therefore compared to the nearest simulation point within a 1% time deviation where available. Two values are given for Hartlerode *et al.* data as the method the authors used to calculate SEM was not clear (see supplementary material in Hartlerode *et al.* (Hartlerode *et al.*, 2015)). We back calculated the standard deviation either assuming the 88 cells were used to calculate SEM, or instead that the three experiments were used to calculate SEM, given in brackets.**



**Figure 2: Recruitment of (A) Ku70/80 and (B) DNA-PKcs to exposed double strand break ends; values are normalised to the maximum value achieved during the 30 seconds plotted. Individual simulation repeats were normalised to their initial number of double strand breaks. Final values are the average of these normalised trends. (C) Recovery of fluorescent DNA-PKcs at double strand break ends after a simulated photo-bleaching event; intensities are set to 0 at the photobleaching event and normalised to the maximum value achieved during the 600 seconds plotted. To remove the influence of a variable number of initial double strand breaks, individual simulation results were normalised to themselves before all repeats were averaged to produce final values. The line is the simulation results and empty symbols are the corresponding experimental results. Shaded area around the line is the standard error in the mean of 200 repeats. Empty symbols are experimental values taken from the referenced papers with brackets indicating the cell line followed by the radiation type used. Error-bars associated with Hartlerode *et al.* (Hartlerode *et al.*, 2015) data are standard errors in the mean, the source of error-bars associated with Weterings *et al.* (Weterings *et al.*, 2009) data is unknown, and all other error-bars are the reported standard deviation.**



The simulated fluorescent recovery after photobleaching (FRAP) behaviour in Figure 2C is generated by a customisation of the simulation pathway to include photobleached counterparts of components that involve molecules of DNA-PKcs. At 30 seconds a bleaching event is simulated converting all double strand break ends with attached DNA-PKcs into photobleached versions. Recovery is measured as the rate at which these bleached ends dissociate and acquire new unbleached DNA-PKcs. The behaviour is again compared to literature reported laser generated double strand break data for the same reasons as described above. Weterings *et al.* (Weterings *et al.*, 2009) agrees with the theoretical fit to simulated data up until 140 seconds ( $\chi_v^2 = 0.56$ ). After this the fit worsens ( $\chi_v^2 = 10.47$ ) as the data plateaus earlier in spite of the same cell line being used. The lack of data from Hammel *et al.* (Hammel *et al.*, 2010) for times greater than 300 seconds prohibit identification of a plateau.

The fluorescent recovery of a system can be diffusion limited, reaction limited, or a complex mix of the two. Diffusion controlled systems are limited by the rate at which non-bleached proteins diffuse into the bleached area and replace bleached proteins. The behaviour of these systems tends towards the equation (Soumpasis, 1983):

$$f(t) = e^{-2\tau_D/t} [I_0(2\tau_D/t) + I_1(2\tau_D/t)] \quad (2)$$

where  $f(t)$  is the relative fluorescence,  $t$  is the time, and  $\tau_D$  is the "characteristic" diffusion time. Reaction controlled systems are limited by the rate at which bleached proteins detach from binding sites in the breached region, allowing non-bleached proteins to replace them. The behaviour of these systems tends towards the equation (Bulinski *et al.*, 2001):

$$f(t) = 1 - e^{-k_{off}t} \quad (3)$$

where  $f(t)$  is the relative fluorescence,  $t$  is the time, and  $k_{off}$  is the detachment rate of the protein in question. Fitting these equations to the Davis *et al.* data, Equation 2 with  $\tau_D = 33.46$  seconds gives a reduced  $\chi^2$  of 0.595 whilst Equation 3 with  $k_{off} = 1.26 \times 10^{-2} \text{ s}^{-1}$  gives a reduced  $\chi^2$  of 0.025. From this it can be seen that the Davis *et al.* data is much better fitted by a reaction limited, rather than diffusion limited system. Due to the rapid recruitment of Ku70/80 and DNA-PKcs derived from the data in Figure 2A and 2B, the dissociation time constant is the remaining parameter which has a dominating influence on the initial behaviour of the recovery curve for the first 300 seconds in our model. The final time constant selected for dissociation of transient synaptic complexes is 140 seconds.

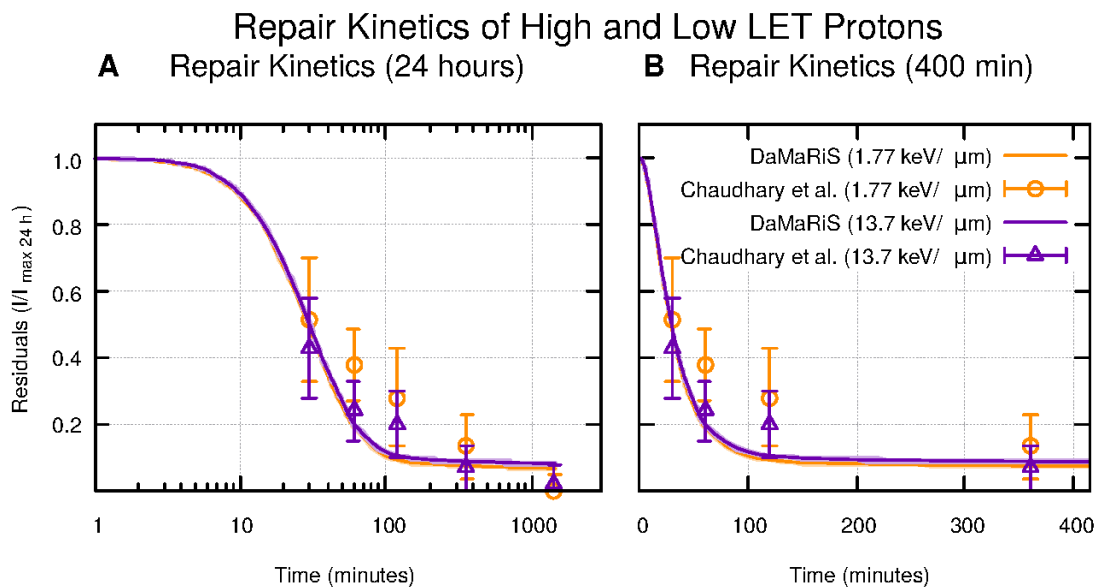
Authors	Diffusion Limited		Reaction Limited	
	$\tau_D$ (s)	$\chi_v^2$	$k_{off}$ ( $\text{s}^{-1}$ )	$\chi_v^2$
<b>This work</b>	37.78	0.263	0.0128	0.008
<b>Davis <i>et al.</i></b>	34.37	0.595	0.0125	0.025
<b>Weterings <i>et al.</i></b>	20.81	0.281	0.0206	0.022
<b>Hammel <i>et al.</i></b>	32.00	0.070	0.0169	0.379

**Table 2: Summary of the reduced  $\chi^2$  for the different experimental datasets compared to theoretical fits assuming diffusion limited or reaction limited recovery. All works are fitted best by assuming reaction limited recovery apart from that of Hammel *et al.*.**

In the simulation, DNA-PKcs is assumed to be retained until the double strand breaks are converted into a fully fixed state. Due to the complex nature of all the breaks, this state is only reached after backbone breaks and base lesions are cleaned. The plateau of the recovery curve from 300-600 seconds in Figure 2C therefore gives a lower bound for the time constants associated with these actions. Final time constants selected are Backbone Clean-up = 300 seconds, Base Lesion Clean-up = 900 seconds in agreement with those of Friedland *et al.* (Friedland *et al.*, 2010a).

The fluorescent recovery plateaus at ~60% of pre-bleaching event intensity as observed by Davis *et al.*. In the simulation DNA-PKcs can only dissociate as part of the entire synaptic complex dissociating. Dissociation of these synaptic complexes can only happen whilst they are in their unstable, transient, states. Therefore, manipulation of the rate at which synaptic complexes stabilise is sufficient to vary the amount of bleached DNA-PKcs trapped at double strand breaks and determine the overall recovery. The final time constant selected for stabilisation of transient synaptic complexes is 250 seconds.

### 3.2 Repair Kinetics



**Figure 3: Resolution of double strand breaks over (A) a 24 hour period and (B) a zoom in of the first 400 minutes, for two separate LET proton radiations. Residuals were calculated as the initial number of double strand breaks reduced by the number of fully ligated 'fixed' breaks in the system at that time point. Individual simulation repeats were normalised to their initial number of double strand breaks. Final values are the average of these normalised trends. Lines are simulation results with the shaded area around the line representing the standard error in the mean of 200 repeats. Empty circles are the corresponding experimental results of 53BP1 foci resolution from Chaudhary *et al.* (Chaudhary *et al.*, 2016), with error-bars representing the reported standard deviations.**

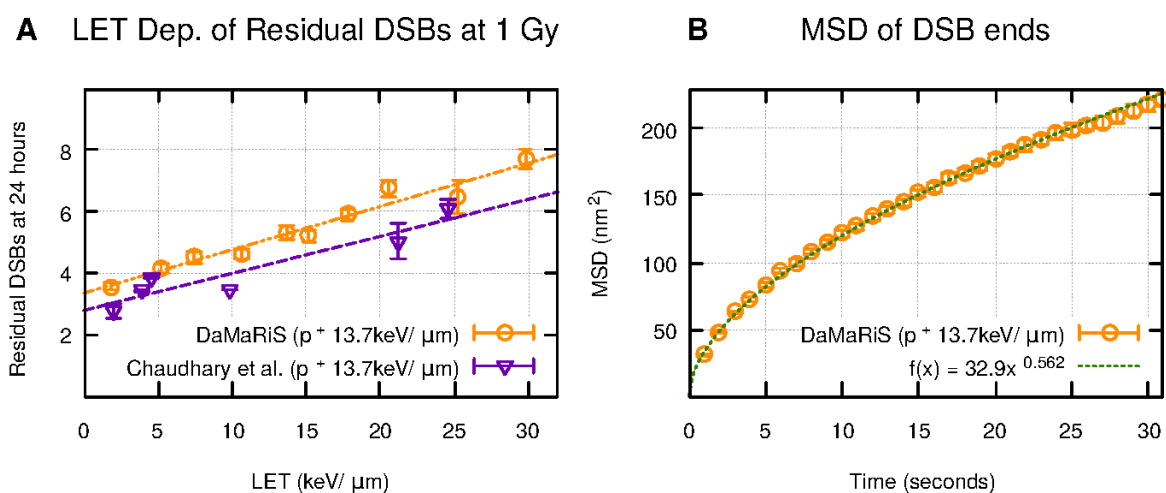
DNA damage patterns were generated using the model published by Henthorn *et al.* (Henthorn *et al.*, 2017, 2018) with beam parameters corresponding to those used experimentally by Chaudhary *et al.* (Chaudhary *et al.*, 2016). This allowed comparison of high (13.7 keV/μm) and low LET (1.77 keV/μm) proton beams directly to experimental data (Figure 3) in order to fit the remaining final ligation time constant and the parameters that govern the mobility of double strand break ends. The primary effect

of the final ligation time constant is to determine the delay in the sharp drop off in residuals observed in the first 100 minutes in Figure 3 but has minimal impact on the slope of this curve. The initial experimental time points at 30 minutes therefore set an upper limit to the final ligation time constant in order to avoid overshooting them. The final time constant selected for final ligation is 1200 seconds.

The mechanisms of the repair machinery have thus far been determined by fitting to experimental data which has resulted in limiting the period of action to the initial two hours. Consequently, to fit our model to the residual breaks at 24 hours we must address the diffusion of double strand break ends.

### 3.3 Diffusion

#### DSB Mobility and Residual Breaks



**Figure 4: (A) Residual breaks at 24 hours for 1 Gy of proton radiation at a range of LET showing the same trend as experimental measurements. Empty circles are simulation results, with error-bars showing the standard error in the mean of 69 repeats for the  $\sim 25.3$  keV/ $\mu\text{m}$  point and 200 for the rest. Experimental values are shown as empty triangles and taken from Chaudhary *et al.* (Chaudhary *et al*, 2016), with error-bars representing the standard error in the mean. Dashed lines are linear fits. Fit to this work has  $m=0.141\pm 0.011$  and  $c=3.34\pm 0.18$ , fit to Chaudhary *et al.* data has  $m=0.121\pm 0.021$  and  $c=2.78\pm 0.30$ . (B) Mean squared displacement of double strand breaks caused by 1 Gy of 13.72 keV/ $\mu\text{m}$  protons over the initial 30 seconds. Empty circles are simulation results with error-bars representing the standard error in the mean of 200 repeats. The dashed line is Equation 4 fitted with values of  $D_{\alpha} = (32.9\pm 0.4) \text{ nm}^2/\text{s}^{0.56}$ , and  $\alpha = 0.562\pm 0.004$ .**

The yield of residual double strand breaks at 24 hours was scored for 1 Gy of 13.72 keV/ $\mu\text{m}$  proton irradiation whilst varying the magnitude of the jump diffusion coefficient assigned to double strand break ends. The average number of residual breaks reported by the simulation was compared with the experimental data of the corresponding time point. The best fit was achieved ( $\sim 4.6$  residual foci (Chaudhary *et al*, 2016)) at a jump diffusion coefficient of  $2.8\times 10^{11} \text{ nm}^2/\text{s}$ . The residual double strand break yield produced by this motion was then compared across a range of LET to residual 53BP1 foci experimentally observed by Chaudhary *et al.* (Chaudhary *et al*, 2016), Figure 4A. Our model reproduces a similar trend in LET dependence with gradients for both fitted lines within error. There is

however a small systematic over-estimation shown by the difference in intercepts of the fitted lines of  $0.56 \pm 0.35$  breaks.

To quantify the sub-diffusive behaviour in the model, the mean squared displacement of double strand break ends were tracked over the initial 300 seconds of the simulation for 200 repeats, Figure 4B. The data was fitted with a power law:

$$MSD = D_{\alpha} t^{\alpha} \quad (4)$$

with  $D_{\alpha}$  in units of  $\text{nm}^2/\text{s}^{\alpha}$  being the general diffusion coefficient, and  $\alpha$  being the anomalous diffusion exponent.  $\alpha = 1$  indicates normal Brownian diffusion following Fick's law,  $\alpha > 1$  indicates super-diffusive properties, and  $\alpha < 1$  indicates sub-diffusion.

Mobility of individual double strand break ends are difficult to investigate, however, there are numerous groups that have investigated the mobility of double strand break loci or chromosomal loci. Cabal *et al.* (Cabal *et al.*, 2006) measured the motion of fluorescently labelled GAL genes in *Saccharomyces cerevisiae*, measuring  $\alpha = 0.42 \pm 0.01$  for gene foci not confined to the nuclear envelope. Weber *et al.* (Weber *et al.*, 2010) studied fluorescently labelled chromosomal loci in *E. coli* and *Vibrio cholerae*, measuring  $\alpha$  to range from  $0.35 \pm 0.02$  to  $0.44 \pm 0.06$  for various gene foci. Bronstein *et al.* (Bronstein *et al.*, 2009) used GFP-tagged TRF2 in U2OS cells to measure  $\alpha$  for telomeres, observing growth from  $0.32 \pm 0.12$  to  $0.51 \pm 0.20$  during initial anomalous motion. Girst *et al.* (Girst *et al.*, 2013) used GFP-tagged MDC1 in U2OS cells to measure  $\alpha = 0.49 \pm 0.05$  for double strand break foci after carbon irradiation and  $\alpha = 0.50 \pm 0.05$  after proton irradiation. Girst *et al.* did not observe splitting of foci and therefore concluded that individual double strand break ends must have a similar scaling of MSD with time. The measured value of the anomalous diffusion exponent,  $\alpha$ , for individual double strand break ends in this work is  $0.562 \pm 0.004$  (Figure 4B). This agrees well with the work of Girst *et al.*.

In contrast to the relatively similar measurements of  $\alpha$ , measurements of the general diffusion coefficient by the same groups differ considerably. Cabal *et al.* reported coefficients on the order of  $\times 10^{-2} \mu\text{m}^2/\text{s}^{\alpha}$  for chromosomal foci in yeast, Bronstein *et al.* reported coefficients on the order of  $\times 10^{-3} \mu\text{m}^2/\text{s}^{\alpha}$  for telomeres in bacteria, and Girst *et al.* reported coefficients on the order of  $\times 10^{-4} \mu\text{m}^2/\text{s}^{\alpha}$  for double strand break foci in human cells. The measured value of the general diffusion coefficient,  $D_{\alpha}$ , for individual double strand break ends in this work is  $(32.9 \pm 0.4) \text{nm}^2/\text{s}^{0.56}$ , on the order of  $\times 10^{-5} \mu\text{m}^2/\text{s}^{\alpha}$ . This is an order of magnitude smaller than the motion measured by Girst *et al.*, indicating tighter confinement of the constituent ends than the parent double strand break.

## 4 Discussion

The complex nature of biological systems involves many individual actions that combine to produce an overall behaviour. To model these systems *in silico*, either a top down or bottom up approach can be used. Top down approaches, such as phenomenological models, have the advantage of limiting the number of variables used to recreate the observed behaviour of a system. Such an approach is useful for parameterising dependencies of the system on known variables, but does not provide clear insight into the mechanisms responsible. Bottom up approaches have the advantage of being useful in exploring the unknown effects of specific mechanisms on overall behaviour. As such they are useful

tools for determining the compatibility of separately proposed, experimentally supported mechanisms. In this work we have combined several such mechanisms into one system and tested the subsequent overall behaviour against experimentally observed trends. We show that, with reasonable parameters for individual mechanistic steps, experimental results can be reproduced for relevant time points when combining mechanisms proposed by Graham *et al.* (Graham *et al.*, 2016), Lee *et al.* (Lee *et al.*, 2015), and Uematsu *et al.* (Uematsu *et al.*, 2007). However, we show that a CTRW interpretation of the sub-diffusive motion of double strand break ends, as suggested by Bronstein *et al.* (Bronstein *et al.*, 2009) and Girst *et al.* (Girst *et al.*, 2013), is likely insufficient to explain overall repair kinetics.

Mechanistic, bottom up models must assign parameters to each mechanism which increases the danger of over-fitting. To guard against over-fitting, we have chosen to fit the model detailed in this work progressively, starting from fitting the earliest mechanisms in isolation and building from there. Table 3 summarises these results and other parameters used in the model.

Process	Value	Fitted
Ku70/80 recruitment	1.1 s	Figure 2A
Ku70/80 inhibition	0.85 s	
Release from Ku70/80 inhibition	3.8 s	
DNA-PKcs recruitment	1.2 s	Figure 2B
Reaction range to form synaptic complex	Set at 25 nm	Section 2.4.3
Dissociation of synaptic complex	140 s	Figure 2C
Clean Base Lesion	300 s	Figure 2C
Clean Backbone Break	900 s	
Stabilisation of synaptic complex	250 s	Figure 2C
Final ligation steps	1200 s	Figure 3
Jump diffusion coefficient	$2.8 \times 10^{11} \text{ nm}^2/\text{s}$	Section 3.3
Minimum waiting time	Set at 1 ms	Section 2.3
'Jump' time	Set at 1 ps	

**Table 3: Table summarising the parameters used in this model. Values preceded by 'Set at' are not fitted parameters. The table is divided such that parameters manipulated concurrently to fit literature data are within the same cell. From this it can be seen that at most 3 parameters were used to reproduce any one experimentally observed behaviour.**

The initial pre-synaptic recruitment of Ku70/80 and DNA-PKcs is well reproduced by the combination of mechanisms in our model. The observed behaviour of DNA-PKcs recruitment can be well explained simply as a sequential recruitment of Ku70/80 followed by DNA-PKcs. This makes sense structurally as it has been shown that DNA-PKcs binds to the C-terminal motif of Ku80 (Hammel *et al.*, 2010), and Ku70/80 is commonly reported to be the first NHEJ protein recruited to break sites (Ma *et al.*, 2004; Lieber, 2010; Chiruvella *et al.*, 2013; Wang & Lees-Miller, 2013; Brown *et al.*, 2017). The inhibition of Ku70/80 in our model represents competition by alternative pathways for exposed double strand break ends. This means that the rate of Ku70/80 phosphorylation, and subsequent dissociation of the synaptic complex, governs the rate at which double strand breaks have the opportunity to progress down alternative pathways, as suggested by Lee *et al.* (Lee *et al.*, 2015).

Due to the self-normalised nature of the experimental data reported, determining the mechanisms governing Ku70/80 recruitment by fitting parameters in our model is challenging and ambiguous. There are conceivably different combinations of the 3 time constants used which would lead to the same overall kinetics. Furthermore, the difference in Ku70/80 recruitment reported by Andrin *et al.* (Andrin *et al.*, 2012) and Hartlerode *et al.* (Hartlerode *et al.*, 2015) could be explained by either a decrease or substantial increase in the time constant for release from inhibition. Decreasing the time constant would logically result in an increase in the overall recruitment rate of Ku70/80. A substantial increase in the time constant would result in the inhibited state not being released during the time frame investigated and thus the self-normalised Ku70/80 recruitment would represent only the rapid kinetics of Ku70/80 recruitment, unperturbed by inhibition. More rigorous fitting therefore requires a better understanding of the initial mechanisms governing competition by more directly comparable experiments, such as those by Mao *et al.* (Mao *et al.*, 2009, 2008).

The mechanisms included in this work from Graham *et al.* (G), Lee *et al.* (L), and Uematsu *et al.* (U) combine to produce a constant exchange of DNA-PK at double strand break sites. Specifically these mechanisms combine as follows: Ku70/80 (L) and DNA-PKcs (U) needs to be phosphorylated in order to dissociate. DNA-PKcs, once autophosphorylated in a synaptic complex, is responsible for phosphorylation of Ku70/80 (L). These synaptic complexes are prone to dissociation, but can alternatively stabilise retaining the phosphorylated DNA-PK on the double strand break (G). In this work we have demonstrated that such a combined system is capable of reproducing the observed DNA-PKcs foci kinetics during FRAP experiments as well as the reported long term recovery of ~60% pre-bleach intensity (Uematsu *et al.*, 2007; Davis *et al.*, 2010). Sequential fitting of time constants governing DNA-PKcs recovery was possible due to individual parameters dominating largely separate areas of the FRAP curve, again avoiding over-fitting. Long term recovery was controlled by the rate at which unstable phosphorylated DNA-PK synaptic complexes were stabilised. This could be the result of XRCC4 recruitment by Ku70/80, which together with XLF has been shown to form bridging filaments resistant to dissociation (Reid *et al.*, 2015; Brouwer *et al.*, 2016). Experiments investigating co-localisation of fluorescently labelled XRCC4 and DNA-PKcs foci would enable implementation and verification of a more detailed mechanistic description of this process in our model.

Having fitted the explicit mechanisms of Ku70/80 and DNA-PKcs recruitment kinetics we show that in order to subsequently fit the general mechanism of repair kinetics for the initial 30 minutes, all fitted time constant are set such that their influence is restricted to the first two hours. In our model therefore, the yield of residual breaks at 24 hours is predominantly determined by the mobility of the individual double strand break ends. This leads to the conclusion that the determining factor which leads to a residual break is the capability of the individual ends to escape the local volume and end up isolated. This mechanism alone is enough to reproduce the LET response of residual breaks at 24 hours observed by Chaudhary *et al.* (Chaudhary *et al.*, 2016). However, fitting in this manner causes the model to deviate from experimental data for 1 Gy of 13.7 keV/ $\mu\text{m}$  proton irradiation over the time period of 30-300 minutes post irradiation. At 30 minutes and at 360 minutes post irradiation the difference between the Chaudhary *et al.* (Chaudhary *et al.*, 2016) data and our simulation is not statistically significant, with  $p=0.27$  and  $0.24$  respectively via Welch's t-test. At the intervening points of 60 and 120 minutes, however, the difference is significant with  $p=0.00001$  in both cases. Furthermore, when comparing to a lower LET of 1.77 keV/ $\mu\text{m}$ , Chaudhary *et al.* report that repair kinetics show a decreased rate of resolving breaks over that of 13.7 keV/ $\mu\text{m}$  radiation during this period, which is the inverse of the simulated behaviour. The behaviour shown in experiment cannot

be explained by the differing repair kinetics of complex double strand breaks (Li *et al*, 2014) as high-LET radiation is expected to cause more repair retarding complexities (Henthorn *et al*, 2017). This suggests that there is a missing mechanism behind this behaviour.

Our earlier work showed that the proximity of breaks increases as LET is increased for the same particle type (Henthorn *et al*, 2018). We propose that this LET dependent density could explain the inverse relationship of residual break yield with LET at time points between 30 and 300 minutes. At lower densities, a break that escapes the local area is less likely to encounter another break end and be repaired. This increase in separation from initial partner can be achieved by increasing the mobility of the break end. However, as shown earlier, this would increase the residual yield reported at 24 hours which is currently fitted to experiment. Therefore some mechanism is needed by which breaks have the high mobility necessary to delay formation of synaptic complexes, but remain local enough that given sufficient time meeting another break is probable.

In this work we have chosen to use a CTRW description of sub-diffusion as observed experimentally by Bronstein *et al*. (Bronstein *et al*, 2009) and Girst *et al*. (Girst *et al*, 2013). We show that this leads to sub-diffusive behaviour similar to that reported by others (Bronstein *et al*, 2009; Weber *et al*, 2010; Girst *et al*, 2013; Cabal *et al*, 2006). Having fitted motion to residual double strand breaks from work by Chaudhary *et al.*, it is interesting to note that we also reproduce the ~100 nm displacements of individual double strand break ends observed by Soutoglou *et al*. (Soutoglou *et al*, 2007). Soutoglou *et al*. reported that when Ku80 was missing this displacement increased to >500 nm, which could be explained by a dependency of the constrained motion on synapsis formation. If this is the case then it lends weight to the CTRW description of sub-diffusion, which could arise due to the mechanism of transient synaptic complex formation proposed by Graham *et al*. (Graham *et al*, 2016).

In contrast to the CTRW model of sub-diffusive motion, Lucas *et al*. and Weber *et al*. observe behaviours indicating a fractional Langevin motion (fLm) form of sub-diffusion (Lucas *et al*, 2014; Weber *et al*, 2010). This form of motion arises when an object is trapped by a visco-elastic boundary; collisions with which results in the frequent reversals of direction observed by these authors. Lucas *et al*. propose that the surrounding nuclear structures provide this visco-elastic boundary, confining double strand breaks to a local volume. The general diffusion coefficient was found to be  $2.4 \times 10^{-3} \mu\text{m}^2/\text{t}^{0.5}$ , two orders of magnitude larger than in this work, and the anomalous diffusion exponent of 0.5, similar to that of this work, whilst the confinement radius was found to reproduce experimental data best at 500 nm. We propose then that this mechanism could explain the inverse relationship of residual break yield with LET at time points between 30 and 300 minutes. The increase in diffusion coefficient would increase the separation of initial partner double strand break ends, whilst the boundary radius would maintain proximity such that the ends are likely to encounter each other again.

## 5 Conclusion

In this work we use an *in silico* model to test the compatibility of several experimentally supported mechanisms hypothesised to operate at distinct stages along the pathway. We determine that the mechanisms proposed by Graham *et al*. (Graham *et al*, 2016), Lee *et al*. (Lee *et al*, 2015), and Uematsu *et al*. (Uematsu *et al*, 2007) can be incorporated into a theory of sequential joining of the major canonical NHEJ proteins (Lieber, 2010; Yang *et al*, 2016; Li *et al*, 2014) to reproduce

experimentally observed fluorescent foci kinetics. However, we highlight specific biological mechanisms which are not clearly defined experimentally. Furthermore, we investigate the impact of an assumed CTRW mode of sub-diffusion on the pathway. We show that the motion of individual double strand break ends has a determining role in the repair kinetics of our model for mid (30 min) to late (24 h) time points. Although the motion implemented produces similar displacements as reported by Soutoglou *et al.* (Soutoglou *et al.*, 2007), and reproduces the LET dependence of residual breaks reported by Chaudhary *et al.* (Chaudhary *et al.*, 2016) we show that CTRW sub-diffusion alone is insufficient to reproduce repair kinetics observed between 60-400 minutes post irradiation. This wide time frame over which double strand break motion can enact substantial influence over the repair kinetic curve highlights an important need for scientific investigation. Improper modelling of this motion could lead to behaviours which need to be compensated for by fitting of the recruitment and activity of repair proteins. This in turn would lead to erroneous conclusions about the kinetics of these proteins and the biological mechanisms they affect.

The model described in this work is incorporated into the existing Geant4 Monte Carlo software familiar to the field of radiation research. As such it can be readily combined with simulations defining the irradiating apparatus and extending these physics models into the biological realm. The model has been constructed in such a way that it can be easily expanded to include relevant biological processes on the scale of individual cell nuclei. Due to following the conventions and structure of Geant4 it can be readily modified and incorporated into existing code by researchers familiar with the tool-kit. As an example we have, in collaboration with Massachusetts General Hospital and the Harvard Medical School, demonstrated the successful implementation of this model into the TOPAS-nBio software (Perl *et al.*, 2012; Schuemann, 2012; McNamara *et al.*, 2017). These properties make this model a good foundation for continued development by the community towards a multi-scale system capable of clinically relevant simulations.

**Acknowledgements:** JWW, NTH, and KJK would like to acknowledge financial support from EPSRC (grant No.: EP/J500094/1). JWW, NTH, KJK, MJM, and RIM would like to acknowledge financial support from the European Commission (grant No.:EC/730983). This research was funded by the STFC Global Challenge Network+ in Advanced Radiotherapy and EPSRC Grand Challenge Network+ in Proton Therapy and supported by the NIHR Manchester Biomedical Research Centre; KJK and MJM would like to acknowledge financial support from the NIHR Manchester Biomedical Research Centre: Radiotherapy Stream (grant No.:BRC-1215-20007). NK and SF would like to acknowledge financial support from the EPSRC (grant No.:EP/N018060/1). MS, MJM, and KJK would like to acknowledge financial support from the European Commission (grant No.:EC/317169). The computational simulation of DNA damage by NTH was achieved using the Condor High Throughput Computing facility at the University of Manchester.

**Author Contributions:** JWW developed the DNA repair model, generated, fitted, and analysed the data presented in this work. NTH generated the DNA damage data used in this work. JWW, NK, and SF developed the sub-diffusive motion model. JWW drafted the manuscript based on discussions with MS, RIM, KJK, and MJM. All authors reviewed and approved the manuscript.

**Conflict of Interest:** The authors declare that they have no conflict of interest.



## 6 References

- Andrin C, McDonald D, Attwood KM, Rodrigue A, Ghosh S, Mirzayans R, Masson JY, Dellaire G & Hendzel MJ (2012) A requirement for polymerized actin in DNA double-strand break repair. *Nucl. (United States)* **3**: 14
- Ballarini F & Ottolenghi A (2004) Models of chromosome aberration induction: an example based on radiation track structure. *Cytogenet. Genome Res.* **104**: 149–156
- Bernal MA, Bordage MC, Brown JMC, Davidková M, Delage E, Bitar Z El, Enger SA, Francis Z, Guatelli S, Ivanchenko VN, Karamitros M, Kyriakou I, Maigne L, Meylan S, Murakami K, Okada S, Payno H, Perrot Y, Petrovic I, Pham QT, et al (2015) Track structure modeling in liquid water: A review of the Geant4-DNA very low energy extension of the Geant4 Monte Carlo simulation toolkit. *Phys. Medica* **31**: 861–874 Available at: <http://www.sciencedirect.com/science/article/pii/S1120179715010042>
- Bronstein I, Israel Y, Kepten E, Mai S, Shav-Tal Y, Barkai E & Garini Y (2009) Transient anomalous diffusion of telomeres in the nucleus of mammalian cells. *Phys. Rev. Lett.* **103**: 1–4
- Brouwer I, Sitters G, Candelli A, Heerema SJ, Heller I, Melo De AJ, Zhang H, Normanno D, Modesti M, Peterman EJJ & Wuite GJL (2016) Sliding sleeves of XRCC4-XLF bridge DNA and connect fragments of broken DNA. *Nature* **535**: 566–569 Available at: <http://dx.doi.org/10.1038/nature18643>
- Brown JS, O’Carrigan B, Jackson SP & Yap TA (2017) Targeting DNA Repair in Cancer: Beyond PARP Inhibitors. *Cancer Discov.* **7**: 20–37
- Bulinski JC, Odde DJ, Howell BJ, Salmon TD & Waterman-Storer CM (2001) Rapid dynamics of the microtubule binding of ensconsin in vivo. *J. Cell Sci.* **114**: 3885–3897
- Cabal GG, Genovesio A, Rodriguez-Navarro S, Zimmer C, Gadal O, Lesne A, Buc H, Feuerbach-Fournier F, Olivo-Marin JC, Hurt EC & Nehrbass U (2006) SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature* **441**: 770–773
- Cao W, Zaghian M, Randeniya K, Lim G, Titt U & Mohan R (2015) SU-ET-109: An Investigation of Including Variable Relative Biological Effectiveness in Intensity Modulated Proton Therapy Planning Optimization for Head and Neck Cancer Patients. *Med. Phys.* **42**: 3356
- Carabe A, Moteabbed M, Depauw N, Schuemann J & Paganetti H (2012) Range uncertainty in proton therapy due to variable biological effectiveness. *Phys. Med. Biol.* **57**: 1159
- Chaudhary P, Marshall TI, Currell FJ, Kacperek A, Schettino G & Prise KM (2016) Variations in the processing of DNA double-strand breaks along 60-MeV therapeutic proton beams. *Int. J. Radiat. Oncol. Biol. Phys.* **95**: 86–94
- Chiruvella KK, Liang Z, Wilson E, Webb CJ, Wu Y, Virginia A & Wilson TE (2013) Repair of Double-Strand Breaks by End Joining Repair of Double-Strand Breaks by End Joining. *Cold Spring Harb. Perspect. Biol.*: 1–22
- Davis AJ, So S & Chen DJ (2010) Dynamics of the PI3K-like protein kinase members ATM and DNA-PKcs at DNA double strand breaks. *Cell Cycle* **9**: 2529–2536
- DeLuca PM, Wambersie A & Whitmore G (2007) Prescribing, recording, and reporting proton-beam therapy. *J. ICRU* **7**: 1–210
- Eidelman YA, Slanina S V, Salnikov I V & Andreev SG (2012) Mechanistic modelling allows to assess

pathways of DNA lesion interactions underlying chromosome aberration formation. *Russ. J. Genet.* **48**: 1247–1256

Friedland W, Jacob P & Kunderat P (2010a) Mechanistic simulation of radiation damage to DNA and its repair: on the track towards systems radiation biology modelling. *Radiat. Prot. Dosimetry*: ncq383

Friedland W, Jacob P & Kunderat P (2010b) Stochastic simulation of DNA double-strand break repair by non-homologous end joining based on track structure calculations. *Radiat. Res.* **173**: 677–688

Friedland W & Kunderat P (2013) Track structure based modelling of chromosome aberrations after photon and alpha-particle irradiation. *Mutat. Res. Toxicol. Environ. Mutagen.* **756**: 213–223

Giovannini G, Böhlen T, Cabal G, Bauer J, Tessonier T, Frey K, Debus J, Mairani A & Parodi K (2016) Variable RBE in proton therapy: comparison of different model predictions and their influence on clinical-like scenarios. *Radiat. Oncol.* **11**: 1

Girst S, Hable V, Drexler GA, Greubel C, Siebenwirth C, Haum M, Friedl AA & Dollinger G (2013) Subdiffusion supports joining of correct ends during repair of DNA double-strand breaks. *Sci. Rep.* **3**:

Graham TGW, Walter JC & Loparo JJ (2016) Two-Stage Synapsis of {DNA} Ends during Non-homologous End Joining. *Mol. Cell* **61**: 850–858 Available at: <http://www.sciencedirect.com/science/article/pii/S1097276516000940>

Hammel M, Yu Y, Mahaney BL, Cai B, Ye R, Phipps BM, Rambo RP, Hura GL, Pelikan M, So S & others (2010) Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *J. Biol. Chem.* **285**: 1414–1423

Hartlerode AJ, Morgan MJ, Wu Y, Buis J & Ferguson DO (2015) Recruitment and activation of the ATM kinase in the absence of DNA-damage sensors. *Nat. Struct. Mol. Biol.* **22**: 736–743

Henthorn NT, Warmenhoven JW, Sotiropoulos M, Mackay RI, Kirkby KJ & Merchant MJ (2017) Nanodosimetric Simulation of Direct Ion-Induced DNA Damage Using Different Chromatin Geometry Models. *Radiat. Res.* **188**: 690–703

Henthorn NT, Warmenhoven JW, Sotiropoulos M, Mackay RI, Kirkby NF, Kirkby KJ & Merchant MJ (2018) In Silico Non-Homologous End Joining Following Ion Induced DNA Double Strand Breaks Predicts That Repair Fidelity Depends on Break Density. *Sci. Rep.* **8**: 2654 Available at: <http://www.nature.com/articles/s41598-018-21111-8>

IAEA & ICRU (2008) Relative Biological Effectiveness in Ion Beam Therapy. *IAEA Tech. Reports*: 1–165 Available at: [http://www-pub.iaea.org/MTCD/publications/PDF/trs461\\_web.pdf](http://www-pub.iaea.org/MTCD/publications/PDF/trs461_web.pdf)

Incerti S, Ivanchenko A, Karamitros M, Mantero A, Moretto P, Tran HN, Mascialino B, Champion C, Ivanchenko VN, Bernal MA & others (2010) Comparison of GEANT4 very low energy cross section models with experimental data in water. *Med. Phys.* **37**: 4692–4708

Jasin M & Rothstein R (2013) Repair of strand breaks by homologous recombination. *Cold Spring Harb. Perspect. Biol.* **5**: a012740

Karamitros M, Luan S, Bernal MA, Allison J, Baldacchino G, Davidkova M, Francis Z, Friedland W, Ivantchenko V, Ivantchenko A & others (2014) Diffusion-controlled reactions modeling in

Geant4-DNA. *J. Comput. Phys.* **274**: 841–882

Lee K-J, Saha J, Sun J, Fattah KR, Wang S-C, Jakob B, Chi L, Wang S-Y, Taucher-Scholz G, Davis AJ & others (2015) Phosphorylation of Ku dictates DNA double-strand break (DSB) repair pathway choice in S phase. *Nucleic Acids Res.*: gkv1499

Li Y, Reynolds P, O'Neill P & Cucinotta FA (2014) Modeling damage complexity-dependent non-homologous end-joining repair pathway. *PLoS One* **9**: e85816

Lieber MR (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**: 181–211

Lucas JS, Zhang Y, Dudko OK & Murre C (2014) 3D trajectories adopted by coding and regulatory DNA elements: first-passage times for genomic interactions. *Cell* **158**: 339–352

Ma Y, Lu H, Tippin B, Goodman MF, Shimazaki N, Koiwai O, Hsieh C-L, Schwarz K & Lieber MR (2004) A biochemically defined system for mammalian nonhomologous DNA end joining. *Mol. Cell* **16**: 701–713

Mao Z, Bozzella M, Seluanov A & Gorbunova V (2008) DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell cycle* **7**: 2902–2906

Mao Z, Bozzella M, Seluanov A & Gorbunova V (2009) Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair (Amst)*. **7**: 1765–1771 Available at: <http://www.biochemj.org/bj/imps/abs/BJ20141236.htm>

Marshall TI, Chaudhary P, Michaelidesová A, Vachelová J, Davidková M, Vondráček V, Schettino G & Prise KM (2016) Investigating the Implications of a Variable RBE on Proton Dose Fractionation Across a Clinical Pencil Beam Scanned Spread-Out Bragg Peak. *Int. J. Radiat. Oncol. Biol. Phys.* **95**: 70–77 Available at: <http://dx.doi.org/10.1016/j.ijrobp.2016.02.029>

McNamara A, Geng C, Turner R, Mendez JR, Perl J, Held K, Faddegon B, Paganetti H & Schuemann J (2017) Validation of the radiobiology toolkit TOPAS-nBio in simple DNA geometries. *Phys. Medica* **33**: 207–215

McNamara A, Underwood T, Wo J & Paganetti H (2016) SU-FT-132: Variable RBE Models Predict Possible Underestimation of Vaginal Dose for Anal Cancer Patients Treated Using Single-Field Proton Treatments. *Med. Phys.* **43**: 3492

Metzler R & Klafter J (2000) The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.* **339**: 1–77

Miné-Hattab J, Recamier V, Izeddin I, Rothstein R & Darzacq X (2017) Multi-scale tracking reveals scale-dependent chromatin dynamics after DNA damage. *Mol. Biol. Cell* **28**: 3323–3332 Available at: <http://www.molbiolcell.org/lookup/doi/10.1091/mbc.E17-05-0317>

Moore S, Stanley FKT & Goodarzi AA (2014) The repair of environmentally relevant DNA double strand breaks caused by high linear energy transfer irradiation--no simple task. *DNA Repair (Amst)*. **17**: 64–73

Paganetti H (2014) Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. *Phys. Med. Biol.* **59**: R419

Paganetti H, Niemierko A, Ancukiewicz M, Gerweck LE, Goitein M, Loeffler JS & Suit HD (2002) Relative biological effectiveness (RBE) values for proton beam therapy. *Int. J. Radiat. Oncol.* **53**:

407–421

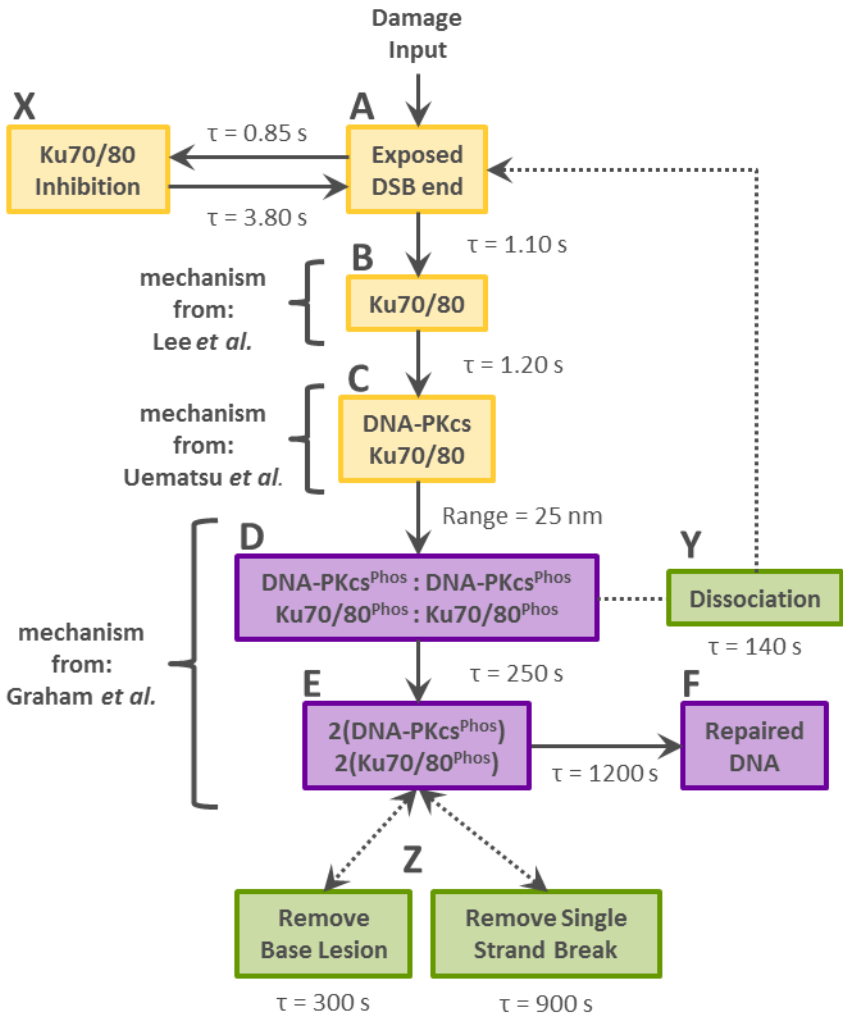
- Perl J, Shin J, Schümann J, Faddegon B & Paganetti H (2012) TOPAS: An innovative proton Monte Carlo platform for research and clinical applications. *Med. Phys.* **39**: 6818–6837
- Reid DA, Keegan S, Leo-Macias A, Watanabe G, Strande NT, Chang HH, Oksuz BA, Fenyo D, Lieber MR, Ramsden DA & others (2015) Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair. *Proc. Natl. Acad. Sci.* **112**: E2575–E2584
- Sachs K, Levy D, Chen AM, Ingerman EA, Simpson PJ, Cornforth MN, Hahnfeldt P & Hlatky LR (2000) Random breakage and reunion chromosome aberration formation model; an interaction distance version based on chromatin geometry. *Int. J. Radiat. Biol.* **76**: 1579–1588
- Saxton MJ (2007) A Biological Interpretation of Transient Anomalous Subdiffusion. I. Qualitative Model. *Biophys. J.* **92**: 1178–1191 Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1783867/>
- Schipler A & Iliakis G (2013) DNA double-strand-break complexity levels and their possible contributions to the probability for error-prone processing and repair pathway choice. *Nucleic Acids Res.:* gkt556
- Schuemann J (2012) SU-E-T-475: Nano-Dosimetric Track Structure Scoring including Biological Modeling with TOPAS-NBio. *Med. Phys.* **39**: 3814 Available at: <https://doi.org/10.1118/1.4735564>
- Sokolov IM (2012) Models of anomalous diffusion in crowded environments. *Soft Matter* **8**: 9043–9052
- Soumpasis DM (1983) Theoretical analysis of fluorescence photobleaching recovery experiments. *Biophys. J.* **41**: 95–97
- Soutoglou E, Dorn JF, Sengupta K, Jasin M, Nussenzweig A, Ried T, Danuser G & Misteli T (2007) Positional stability of single double-strand breaks in mammalian cells. *Nat. Cell Biol.* **9**: 675–682
- Thompson LH (2012) Recognition, signaling, and repair of DNA double-strand breaks produced by ionizing radiation in mammalian cells: the molecular choreography. *Mutat. Res. Mutat. Res.* **751**: 158–246
- Tilly N, Johansson J, Isacson U, Medin J, Blomquist E, Grusell E & Glimelius B (2005) The influence of RBE variations in a clinical proton treatment plan for a hypopharynx cancer. *Phys. Med. Biol.* **50**: 2765
- Uematsu N, Weterings E, Yano K, Morotomi-Yano K, Jakob B, Taucher-Scholz G, Mari P-O, Van Gent DC, Chen BPC & Chen DJ (2007) Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. *J. Cell Biol.* **177**: 219–229
- Wang C & Lees-Miller SP (2013) Detection and repair of ionizing radiation-induced DNA double strand breaks: new developments in nonhomologous end joining. *Int. J. Radiat. Oncol. Biol. Phys.* **86**: 440–449
- Weber SC, Spakowitz AJ & Theriot JA (2010) Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm. *Phys. Rev. Lett.* **104**: 238102
- Wedenberg M, Lind BK & Hårdemark B (2013) A model for the relative biological effectiveness of

protons: The tissue specific parameter  $\alpha/\beta$  of photons is a predictor for the sensitivity to LET changes. *Acta Oncol. (Madr)*. **52**: 580–588

Wedenberg M & Toma-Dasu I (2014) Disregarding RBE variation in treatment plan comparison may lead to bias in favor of proton plans. *Med. Phys.* **41**: 91706

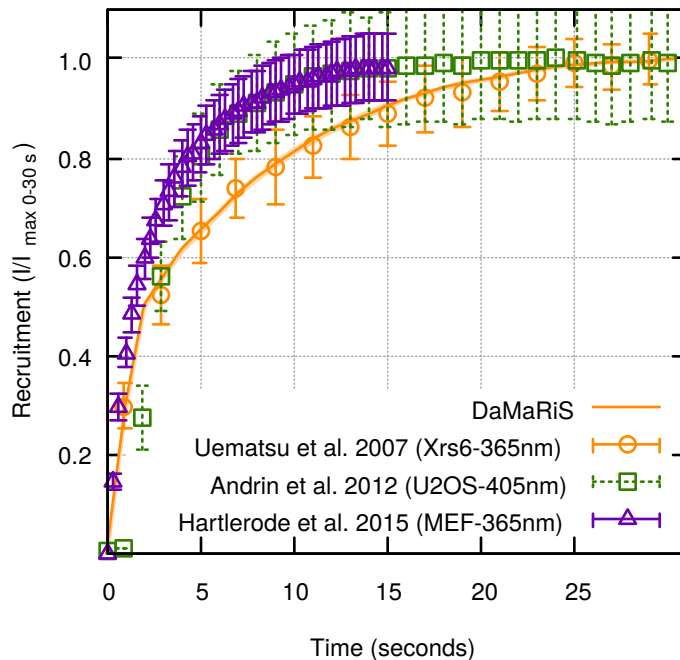
Weterings E, Verkaik NS, Keijzers G, Florea BI, Wang S-Y, Ortega LG, Uematsu N, Chen DJ & van Gent DC (2009) The Ku80 carboxy terminus stimulates joining and artemis-mediated processing of DNA ends. *Mol. Cell. Biol.* **29**: 1134–42 Available at:  
<http://www.scopus.com/inward/record.url?eid=2-s2.0-61749092799&partnerID=tZOtx3y1>

Yang K, Guo R & Xu D (2016) Non-homologous end joining: Advances and frontiers. *Acta Biochim. Biophys. Sin. (Shanghai)*. **48**: 632–640

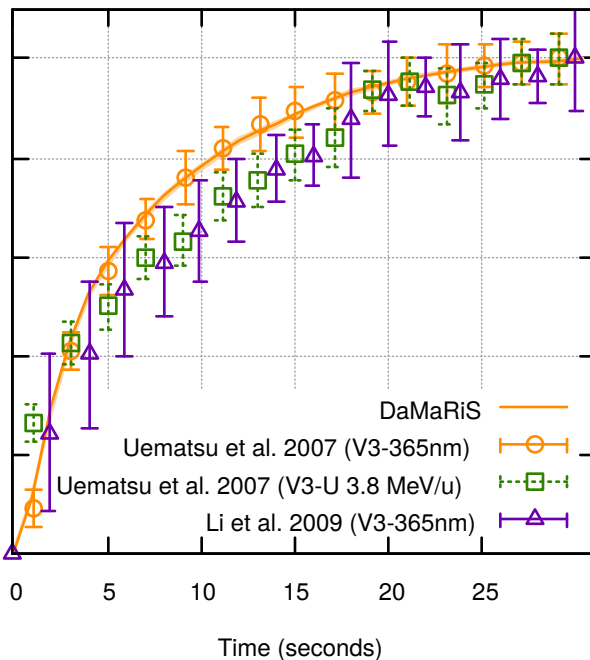


# Recruitment Kinetics of Ku70/80 and DNA-PKcs

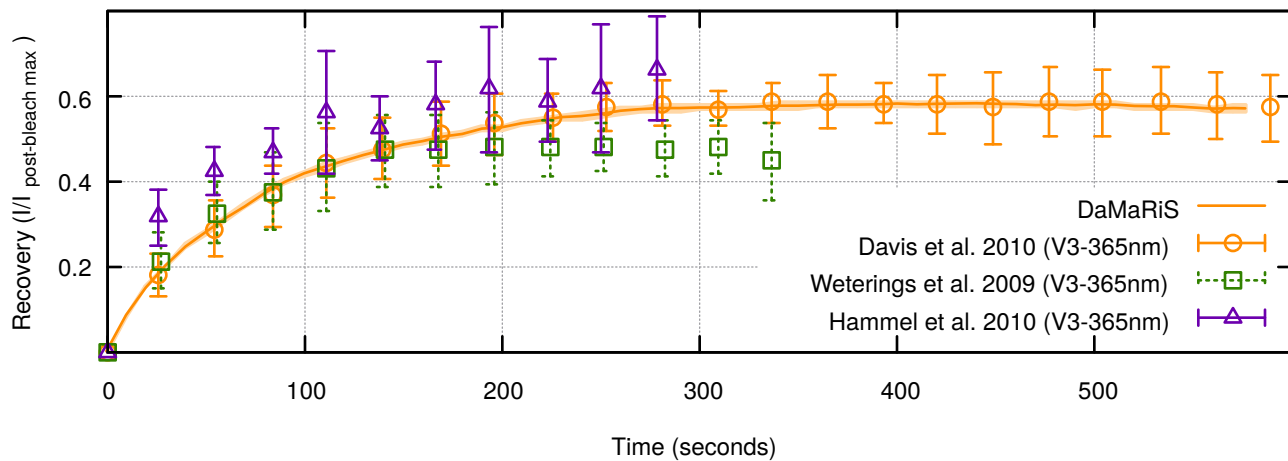
## A Ku70/80 Recruitment



## B DNA-PKcs Recruitment

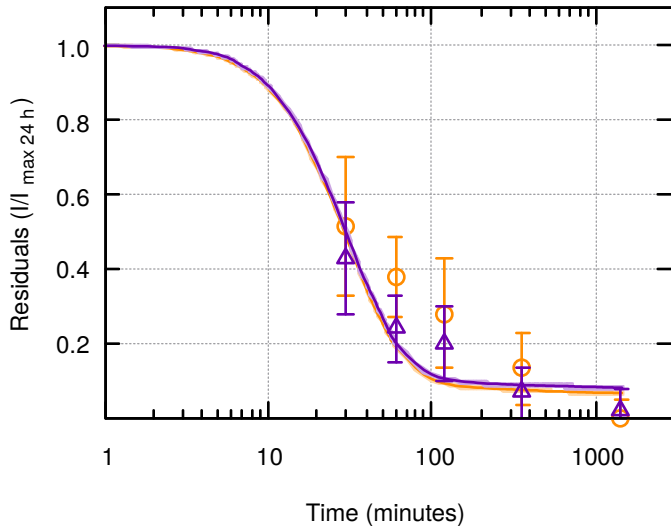


## C FRAP of DNA-PKcs

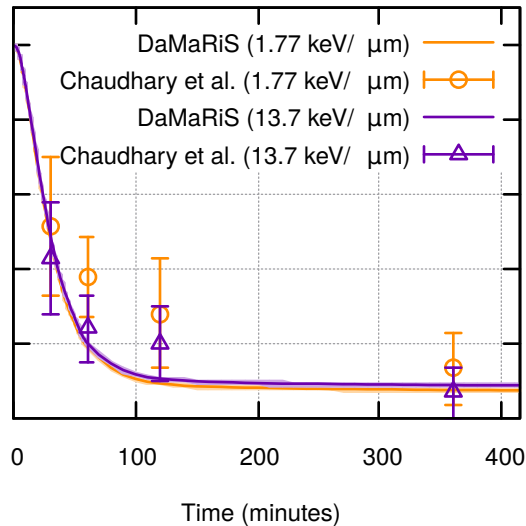


# Repair Kinetics of High and Low LET Protons

## A Repair Kinetics (24 hours)



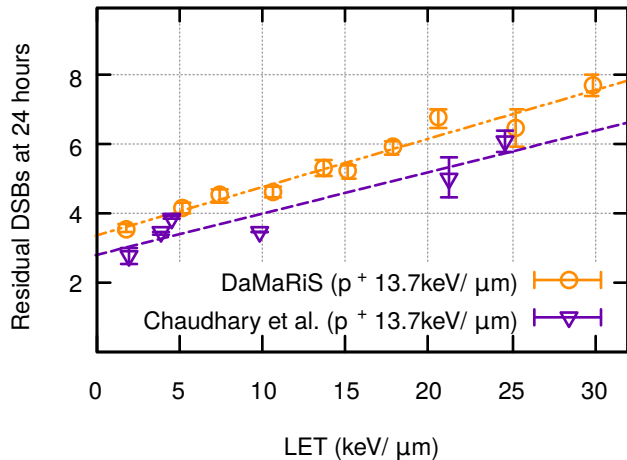
## B Repair Kinetics (400 min)





# DSB Mobility and Residual Breaks

**A** LET Dep. of Residual DSBs at 1 Gy



**B** MSD of DSB ends

