



UvA-DARE (Digital Academic Repository)

Understanding gene expression variability in its biological context using theoretical and experimental analyses of single cells

Kempe, H.

Publication date

2017

Document Version

Other version

License

Other

[Link to publication](#)

Citation for published version (APA):

Kempe, H. (2017). *Understanding gene expression variability in its biological context using theoretical and experimental analyses of single cells*.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 5

Transcription recovery after UV-C-induced DNA damage: the relationship between UV-C dose, gene size and DNA repair at the single gene level

In collaboration with:
Ilona Vuist, Adriaan B. Houtsmuller and Pernette J. Verschure
Adapted from: manuscript in preparation.

Introduction

Maintenance of genome integrity and gene expression patterning is essential for proper cell functioning and identity. The genome is constantly threatened by DNA damaging agents including endogenous metabolic products and by-products (reactive oxygen species and alkylating agents), environmental chemicals, and UV irradiation. UV-induced DNA damages such as cyclobutane pyrimidine dimers (CPDs), 6-4 pyrimidine-pyrimidone photoproducts (6-4PPs), and bulky lesions are repaired by nucleotide excision repair (NER). Defects in DNA repair pathways are associated with mutations and DNA rearrangements as noted in a variety of diseases such as Xeroderma pigmentosum, Cockayne syndrome, skin cancer predisposition and aging [180]. DNA damages can be detected by two mechanisms, global genome detection (GG-NER) and transcription-coupled detection (TC-NER). GG-NER is performed by proteins such as Xeroderma pigmentosum complementation group C (XPC) and Damage-Specific DNA Binding Protein 2 (DDB2) that are constantly scanning the entire genome for helix distortions. Upon DNA damage recognition, core NER proteins are recruited to repair the lesion. TC-NER detects DNA damages by stalled RNA polymerase II at the site of lesions [181, 182]. This stalled polymerase mediates the recruitment of TC-NER-specific proteins such as CSB and UVSSA and eventually allows binding of core NER proteins [182, 183]. GG-NER and TC-NER use the same core NER proteins to repair the lesion. TFIIH unwinds the DNA surrounding the lesion, XPA confirms the lesion, and RPA binds to single-stranded DNA. After RPA binding, XPF-ERCC1 and XPG make a dual incision neighboring the lesion leading to an approximately 30 nt DNA fragment that is bound by TFIIH. The gap is filled by DNA polymerases copying the template strand and sealed by DNA ligases resulting in error-free double-stranded DNA [184].

UV-irradiation is known to trigger a dramatic reduction in RNA synthesis and triggers changes in mRNA splicing [185]. The factors, mechanisms, and cellular consequences of UV-induced transcription inhibition are poorly understood. UV-induced damage has been shown, using EU incorporation, to induce global inhibition of transcription [186]. Transcription recovery at the global level is reported to take at least six hours [186].

Recently, it has been shown that GG-NER detection is suppressed upon providing DNA damage induction at UV doses below $5 J/m^2$ UV-C. Suppression of GG-NER detection indicates a priority to repair transcriptionally active genes using only TC-NER as DNA damage recognition mechanism [187]. These observations might suggest that transcription restart has a different response depending on the specific UV dosage given. The mechanisms underlying transcription-related DNA damage responses at UV-C dosages lower than $5 J/m^2$ are largely unexplored. Insight in TC-NER mechanisms upon low UV-C dosage exposure opens a fascinating field of study of relevance for UV exposure at physiological levels.

Single-molecule mRNA-FISH (smRNA-FISH) enables measuring single molecule transcription levels of single genes within single cells [29]. This method represents a robust and sensitive

technique to measure the number of transcripts in nuclei of individual cells [85]. We employ smRNA-FISH to measure gene transcription upon low-dose UV-C-induced DNA damage. This approach enables us to determine DNA damage-induced transcription stalling and reinitiation in a gene specific manner.

It has been shown that the probability to inflict DNA damages is random and linear up to 40 J/m^2 UV-C, and that UV-C induces 0.12 damages per 10 kb per J/m^2 [188]. This relation indicates that relatively long genes acquire more UV-C-induced lesions compared to shorter genes.

We use primary fibroblasts and low UV-C doses to study transcription inhibition and DNA damage responsiveness. Two genes are studied: a relatively long gene *PLOD2* (92 kbp) and a relatively short gene *GAPDH* (4.4 kbp). Both genes are not involved in stress responses or repair pathways, suggesting that their expression is not differentially regulated after UV-C-induced DNA damage. *PLOD2* is a protein involved in the formation of the extracellular matrix and *GAPDH* in glycolysis. We observe transcription inhibition of *PLOD2* one hour after exposure of the cells to a UV-C dose range of 3-10 J/m^2 whereas transcription of *GAPDH* under similar conditions remains unaffected. Our data indicate that transcription inhibition by exposure to low UV doses is regulated at the gene level and suggest that genes that have acquired at least a single damage are not transcribed. Combining our experimental data with a coarse-grained transcription-related DNA damage model we are able to determine the transcription recovery and hence repair rate of DNA damages upon exposure to relatively low UV-C doses. On average, DNA lesions are calculated to have a half-life of approximately two hours.

5.1 Results and discussion

5.1.1 Measurement of transcription inhibition with smRNA-FISH

To study the effect of UV-C-induced DNA damage on transcription stalling and transcription reinitiation, we measured nuclear transcript abundance in primary fibroblasts in response to DNA damage induction by exposing fibroblasts to UV-C (Figure 5.1A). We analyzed the number of *PLOD2* transcripts in individual cells using smRNA-FISH. We measured an average transcript distribution of 8.5 transcripts per nucleus (Figure 5.1B, C). A significant decrease in nuclear *PLOD2* transcripts was observed after treatment with actinomycin D, DRB, and exposure to 10 J/m^2 UV-C (Figure 5.1B). One hour after actinomycin D treatment, 4.9 transcripts per nucleus were measured representing a 42% reduction in the average number of transcripts. A more pronounced transcription reduction is noted upon DRB treatment, i.e. 4.6 transcripts per nucleus, representing a 46% reduction in the average number of transcripts. UV-C exposure (10 J/m^2) caused a 34% reduced average transcription level in fibroblasts. The transcript distribution one hour after 10 J/m^2 UV-C exposure is shown in Figure 5.1C. The effect of DNA

damage on the number of cytoplasmic transcripts was less profound than on nuclear transcripts (data not shown). This difference between cytoplasmic and nuclear transcript abundance may be caused by the time mRNA molecules spend in the nucleus compared their presence in the cytoplasm, i.e. the retention time [40, 189].

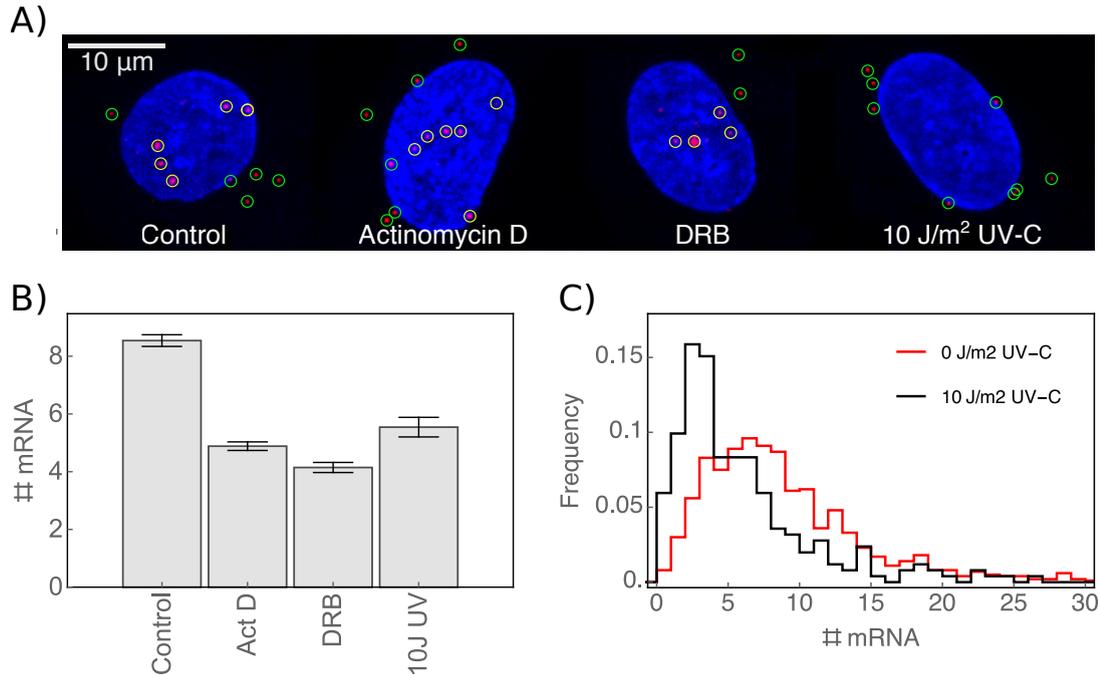


FIGURE 5.1: *PLOD2* nuclear transcripts upon drug induced transcription inhibition and UV-C exposure. Nuclear *PLOD2* transcript numbers in primary fibroblasts were determined in untreated control fibroblasts, in fibroblasts treated with DRB or actinomycin D for one hour, and in fibroblasts exposed to 10 J/m² UV-C. A) Images of all experimental conditions are shown. The detected transcripts are shown in red and DAPI stained nucleus is shown in blue. Nuclear transcripts are annotated with a yellow circle and cytoplasmic transcripts with a green circle. B) The average nuclear transcript numbers of cells treated with DRB or actinomycin D and cells exposed to 10 J/m² UV-C differ significantly from untreated control cells (Mann-Whitney U test, $p < 0.05$, $n > 250$). The error bars represent the standard error of the mean. C) The full nuclear transcript distribution of untreated and treated cells are shown, respectively 0 J/m² UV-C ($n=1000$) and 10 J/m² UV-C ($n=252$).

5.1.2 UV-C-induced transcription inhibition is dependent on gene size and UV dose

The expected number of DNA damages per gene depends on gene size and the administered UV-C dose. Previous studies showed that UV-C induces 0.065 damages per 10 kb per J/m² in the coding strand [188]. Considering this damage per base relationship we calculated that exposure to 3 J/m² UV-C imposes at least 1 DNA damage in 80% of all *PLOD2* genes (equation 5.2, with parameters $p=1.8 \cdot 10^5$ and $n=92000$). We calculated that at least one DNA damage in *PLOD2* is induced upon 5 J/m² or 10 J/m² UV-C exposure in 94% and 99.5% of all cells, respectively. The number of *PLOD2* nuclear transcripts 1 hour after 3, 5, and 10 J/m² UV-C exposure

is shown in Figure 5.2A. We estimated (using equation 5.1) that 92%, 88% and 77% of the cells exposed to 3, 5 and 10 J/m^2 UV-C, have not acquired a single DNA damage in *GAPDH*. Exposure to all three UV-C doses induced only a minor decrease in *GAPDH* transcripts (Figure 5.2B). A striking difference in the number of nuclear transcripts upon treating the cells with UV-C was noted when comparing *PLOD2* and *GAPDH* transcript numbers. These data suggest that gene size is an important parameter for UV-C DNA damage-induced transcription stalling.

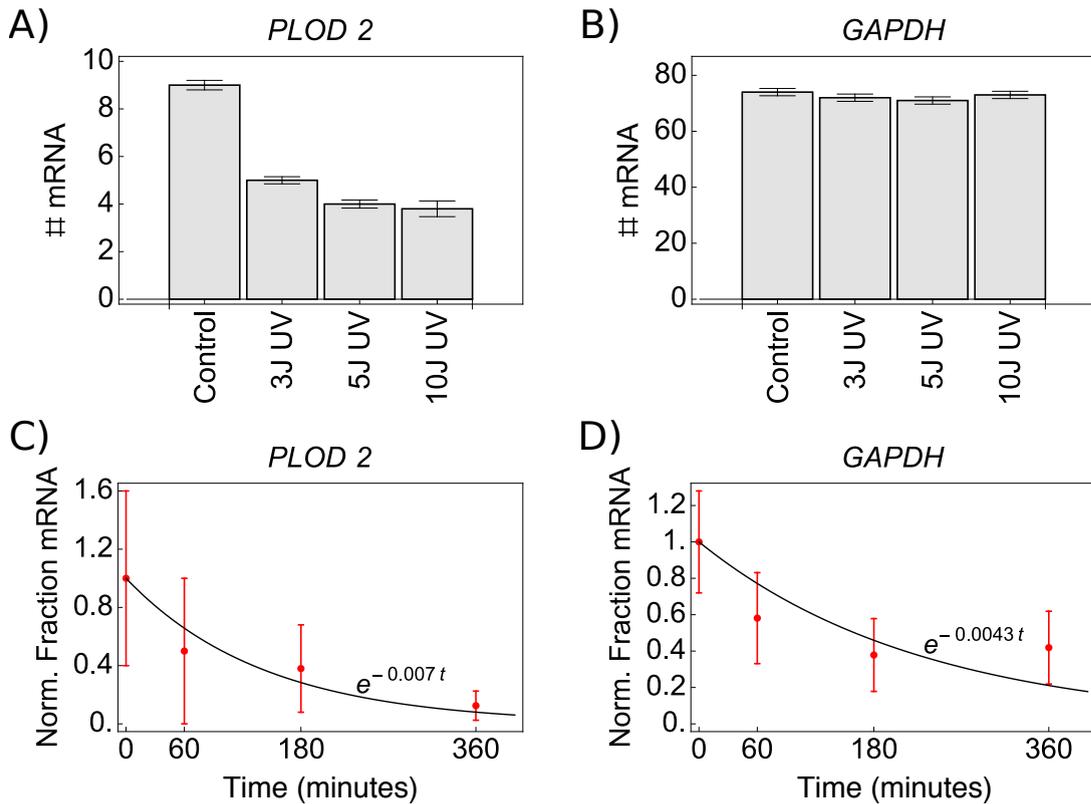


FIGURE 5.2: **The effect of UV-C exposure on gene expression in a relatively long and short gene.** A) The nuclear transcript numbers of a relatively long gene (*PLOD2*, 94kb) were determined in untreated control cells and in cells exposed to UV-C (3-10 J/m^2). One hour after 3 J/m^2 , 5 J/m^2 or 10 J/m^2 UV-C-exposure, *PLOD2* nuclear transcript numbers are decreased compared to transcript numbers in untreated cells. B) Nuclear transcript numbers of a relatively short gene (*GAPDH*, 4.4kb) after UV-C-exposure are not altered compared with transcript numbers in untreated cells. C-D) Nuclear transcripts of both *PLOD2* and *GAPDH* genes were determined at different time points after actinomycin D treatment. We observed a decrease in *PLOD2* and *GAPDH* nuclear transcripts. The data is fitted with an exponential function providing the export kinetics (k_e) of *PLOD2* and *GAPDH* transcripts. The error bars show the standard deviation of the transcript number distribution at each time point after actinomycin D treatment.

5.1.3 mRNA export kinetics of *PLOD2* and *GAPDH*

The amount of transcripts in the nucleus depends on the dynamics of mRNA production (k_p) and export (k_e). We determined mRNA export dynamics by measuring the decrease in *PLOD2* nuclear transcripts at subsequent time points after actinomycin D treatment (Figure 5.2C-D).

Six hours after actinomycin D exposure the transcript number is reduced to 12%. Based on a first-order mRNA export model with no active mRNA degradation in the nucleus, we estimated that the k_e of *PLOD2* is $0.007 \text{ mRNA}^{-1}\text{min}^{-1}$ (Figure 5.2C). The corresponding half-life of *PLOD2* mRNA in the nucleus is 100 minutes. Similarly, we determined the mRNA export dynamics of *GAPDH* providing a k_e of $0.0065 \text{ mRNA}^{-1}\text{min}^{-1}$ yielding a mean half-life of 113 minutes (Figure 5.2D).

5.1.4 Transcription recovery within six hours

Our data show that the average number of DNA damages in individual genes depends on gene size and UV-C dose. We predicted that 3, 5, and 10 J/m^2 UV-C exposure causes on average 1.7, 2.8, and 5.7 damages in *PLOD2* (Figure 5.4). Figure 5.3 shows the *PLOD2* nuclear transcript number 60, 180, and 360 minutes after exposing the cells to 3 or 5 J/m^2 UV-C. We observed a drop in *PLOD2* nuclear transcripts already 1 to 3 hours after 3 J/m^2 and 5 J/m^2 UV-C. After 360 minutes we measured an increase in mRNA numbers both for 3 J/m^2 and 5 J/m^2 UV-C (Figure 5.3). We observed that transcription is recovered to the original level after 12 hours (Figure 5.5). In addition, we determined the *PLOD2* transcription dynamics when exposing NER-deficient primary fibroblasts (XP25RO) carrying a mutation leading to a non-functional truncated XPA protein to a low UV-C dose ($3\text{-}5 \text{ J/m}^2$). XPA acts as a NER core protein for damage repair for either GG-NER or TC-NER. We observed that the *PLOD2* transcript number remains low after exposure to 5 J/m^2 UV-C does in NER-deficient XPA mutant fibroblasts (Figure 5.3, black dots). It has been reported that after UV-C-induced transcription stalling, the mRNA level is shortly increased before returning to steady state expression levels [186]. We did not observe such an overshoot in transcript levels. This suggests that either such an overshoot does not occur in the *PLOD2* gene or that an overshoot occurs at a later time point. We noted that the transcript numbers of *PLOD2* and *GAPDH* in primary fibroblasts varies in different experiments. Such variability can be explained by variation in passage number or differences in experimental conditions such as the time to allow cells to adhere to the microscopy slides or a combination of these conditions. For this reason, we only compared transcription in terms of transcript numbers from experiments performed during the same culture period and condition. In all other cases the expression levels were normalized to steady state expression levels.

5.1.5 Modeling of DNA damage repair

We used the transcription recovery data combined with a DNA damage and repair model to estimate the DNA damage repair rate. We assumed a linear transcription model of zero order mRNA production and first order mRNA export. A Bernoulli distribution with a likelihood parameter of $6 \cdot 10^{-6} (\text{J}^{-1}\text{kb}^{-1}\text{UV} - C)$ provides the expected distribution of the number of damages inflicted per gene. Only genes with no damages are assumed to stay transcriptionally

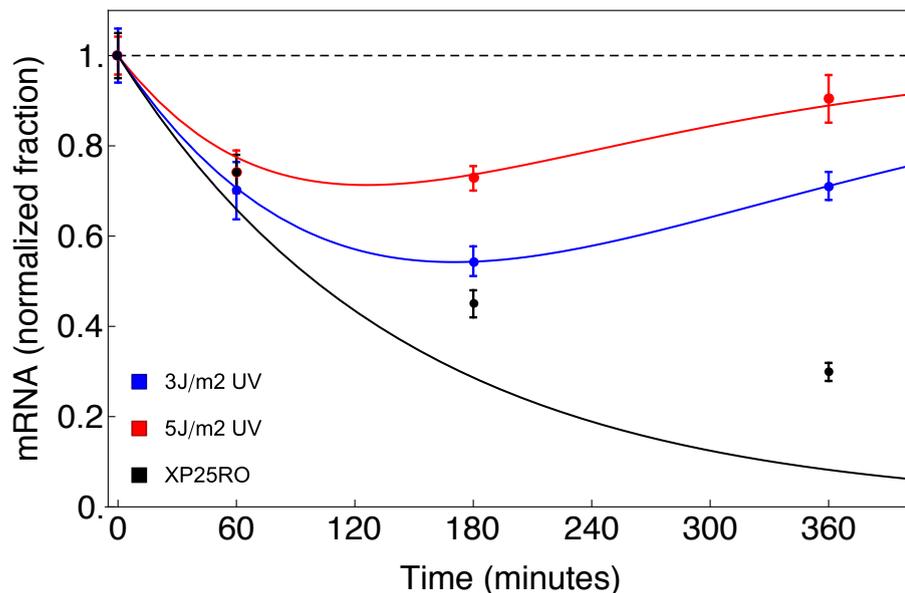


FIGURE 5.3: ***PLOD2* transcription recovery after UV-C exposure in normal and NER-deficient fibroblasts.** The observed *PLOD2* transcription recovery in primary fibroblasts after UV-C exposure ($3 J/m^2$ and $5 J/m^2$) is fitted with a DNA damage repair model. The figure illustrates *PLOD2* nuclear transcript numbers in time after exposure to $3 J/m^2$ (red dots) and $5 J/m^2$ (blue dots) UV-C. The solid lines show the model fit of the experimental data. *PLOD2* transcription recovery after $5 J/m^2$ UV-C exposure in NER-deficient fibroblast carrying an XPA mutation is shown in black. The model simulation describes the experimental data well. The dashed line provides the steady state situation before DNA damage as a transcription recovery end-point.

active. Due to repair events, the likelihood parameter of the Bernoulli distribution changes over time with e^{krt} . We consider a gene to be repressed when one or more damages occur within the body of the gene. A more elaborate description of the DNA damage repair model is described in the Materials and methods section. The used model is highly simplified and lumps an array of sequential TC-NER reactions into a single rate (Figure 5.6). Moreover, in practice, engaged mRNA polymerase II can continue transcription downstream of the damaged site. Considering the size of *PLOD2* (92kb) and the elongation rate of transcription [50], we calculated that transcription proceeds on average for 18 minutes after UV-C-induced damage. Similarly, transcription restart after repair might have a delayed response when elongating polymerases are stalled before the site of damage. The position of engaged polymerases with respect to sites of DNA damage are not taken into account in our model.

5.1.6 Estimation of the DNA damage repair rate

It has been shown that distinct repair pathways are active when exposing the cells to a defined UV dose [187]. At $3 J/m^2$ UV-C, the repair mechanism depends on DNA damage recognition by stalled RNA polymerase II (TC-NER) whereas at $5 J/m^2$ UV-C, DNA damages will also be recognized by GG-NER [187]. In our model we assume that all damages have the same

likelihood to be discovered and repaired. This assumption can be made for TC-NER when the gene involved is actively transcribed and RNA polymerase II is present in excess on the UV-C exposed gene. It should be noted that a gene with a minor amount of RNA polymerase II will behave differently compared to the described model. It can be expected that when the number of RNA polymerase II molecules at the UV-C exposed gene is low, the probability that RNA polymerase II detects the lesion and initializes repair is not uniformly distributed. This will result in a more time-consuming repair, especially in the case of genes with multiple DNA damages. The model is expected to represent GG-NER transcription recovery irrespective of the transcriptional activity of the gene involved, since damage detection depends on factors that are not involved in transcription. Our model enables us to estimate k_r by fitting this model to the observed data. Figure 5.3 shows the model fit with an estimated k_r of 0.013 min^{-1} . This repair rate after 3 J/m^2 UV-C corresponds to the repair rate of TC-NER [187]. In cells exposed to 5 J/m^2 UV-C we measured a slower repair rate of 0.009 min^{-1} . This difference might indicate the efficiency of the different DNA damage detection and repair mechanisms. The observation that repair after 5 J/m^2 UV-C is more time consuming could be explained by the involvement of NER proteins in GC-NER, which reduces the concentration of available core NER proteins for TC-NER.

5.2 Conclusion

We analyzed the effect of low dose UV-C-induced DNA damage on transcription and used this data to calculate the transcription recovery and DNA repair rates. Our model represents a coarse-grained view on transcription-related DNA damage and repair. The calculated repair rate is a lumped parameter of all molecular steps in the repair mechanism including damage recognition of TC-NER and GG-NER, attraction of DNA repair enzymes, cofactors, and all processes related to the restart of transcription. The mean half-life of a DNA damage is 53 minutes when the TC-NER pathway is active and repair of DNA damage is slower when GC-NER is involved (mean half-life is 77 minutes). Our data show that transcription of a short gene (*GAPDH*) is relatively unaffected by UV-C irradiation, probably due to the low probability of obtaining DNA damage. Genes encoding for proteins that are involved in GG-NER and/or TC-NER repair might on average be short and therefore not observe an altered transcription upon being exposed by low UV-C exposure. As a next step it would be interesting to make a genome map of the relationship between transcription stalling and UV-exposure paying special attention to genes encoding NER proteins.

5.3 Materials and Methods

5.3.1 Cells and cell culture

Primary human fibroblasts 1093-SK (ATCC® CRL-2115™) were cultured in DMEM (Gibco® 31965023) supplemented with 10% (v/v) fetal calf serum (Gibco® 16140) and 100 *U/ml* penicillin-streptomycin (Gibco® 15140). The cells were incubated at 37°C in a humidified 5% CO₂ atmosphere.

5.3.2 DNA Damage and transcription inhibition

Cells were cultured in tissue culture chambers on coverslips (Sarstedt, 94.6190.402). The cells were washed with PBS and exposed to different doses of UV-C light at 1.3 *J/m²/s*. After UV-C exposure, fresh medium was added to the cultured cells during the remaining culture period. Actinomycin D (A1410 Sigma) and DRB (D1910 Sigma) were added to the fresh medium and supplied one hour before fixation.

5.3.3 Single Molecule RNA FISH

Samples were treated according to the Protocol for Adherent Mammalian Cell Lines for Custom Stellaris FISH probes (Stellaris). The samples were fixed with 4% formaldehyde. Cells were permeabilized with ethanol overnight at 4°C. For hybridization, the cells were exposed to 125 nM probe in hybridization buffer and incubated for 4 hours at 37°C. The cells were counterstained with 5 ng/ml DAPI. The sequence of the probe targeting *PLOD2* and *GAPDH* mRNA can be found in Supplemental Tables 1 and 2. The CAL Fluor[®] Red 590 coupled probe set was ordered via Biosearch Technologies, Inc..

5.3.4 Image acquisition

Samples were imaged using a Nikon Ti-E scanning laser confocal inverted microscope (A1) with 60x oil objective. Excitation was performed with a 561.5 nm diode-pumped solid state and a 402.1 nm diode laser. For signal detection, 595-50 nm and 450-50 nm filters were used, respectively. Optical sections were captured at 0.300 μm intervals. To reduce camera noise, we used four times image averaging.

5.3.5 Image analysis

Image analysis was performed as described previously [85]. In short, images are filtered with a Laplacian of Gaussian (LoG) filter. The number of mRNA spots in the cells was determined by

applying a threshold for which the number of mRNA transcripts was least sensitive to changes in this threshold. Nuclear mRNA was identified by overlap of the mRNA signals with the DAPI counterstain.

5.3.6 DNA damage and repair model

We modeled DNA damage after UV-C exposure with a Bernoulli process for each single base of the gene studied. We propose that each base can be either damaged (=1) or undamaged (=0) after UV-C exposure and that the likelihood to become damaged is independent of all other nucleotides. The likelihood parameter of this process is calculated from literature (see Supplementary Information). The number of successes (k) of a Bernoulli process in n trials (i.e. the gene length) is given by:

$$P(X = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \quad (5.1)$$

with k as the number of successes. This is also known as the binomial distribution ($B(n, p)$). Here we are only interested in genes that have not obtained damages (i.e. $k = 0$) since these genes remain transcriptional active. When $k=0$, we can simplify equation 5.1 to:

$$P(X = 0) = (1-p)^n \quad (5.2)$$

We assume that repair is independent of the site of the lesion on a gene and of the amount of lesions on the gene. In that case, the number of damages in a population of cells in time is given by a binomial distribution with a variable likelihood parameter. The likelihood parameter decreases in time due to repair and is provided by:

$$p_t = p_0 \cdot e^{-k_r \cdot t} \quad (5.3)$$

Combining equations 5.2 and 5.3 provides the fraction of undamaged genes in time:

$$f(t) = (1 - p_0 \cdot e^{-k_r \cdot t})^n \quad (5.4)$$

5.3.7 Combining DNA damage and gene transcription

We combined the damage and repair function (equation 5.4) with a linear transcription model of which the production depends on the fraction of undamaged DNA:

$$v_p = k_p \cdot f(t) \quad (5.5)$$

The export of mRNA to the cytoplasm is modeled as a first order reaction:

$$v_e = k_e \cdot mRNA \quad (5.6)$$

We fitted this model to the mRNA expression data using the least squares difference to obtain the repair rate (k_r) of DNA damages. The rate constant for mRNA export (k_e) was obtained by fitting the actinomycin transcription inhibition data with an exponential decay function. The steady state measurements before UV-C exposure were used as the initial mRNA number and provided the rate constant for mRNA production (k_p). The likelihood parameter of UV-C-induced damage of $6 \cdot 10^{-6}$ per J/m^2 was taken from literature [188].

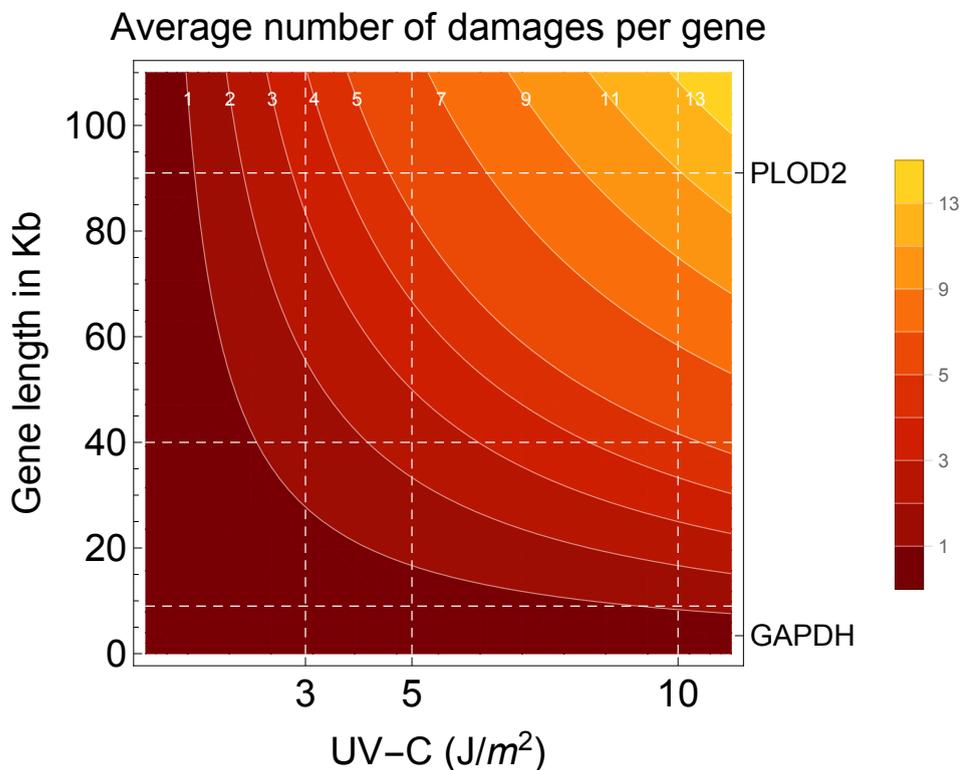


FIGURE 5.4: **The relation between DNA damage, gene size and UV-C dose.** The average number of damages per gene depends on gene size and UV-C dose. The horizontal dashed lines show genes used in this study and the vertical dashed lines the UV-C dosage used. The predicted average number of damages for *PLOD2* are 1.7, 2.8, and 5.7 upon exposing *PLOD2* to $3 J/m^2$, $5 J/m^2$ or $10 J/m^2$ UV-C, respectively. For *GAPDH* the average number of damages is below 1 for all three UV-C doses.

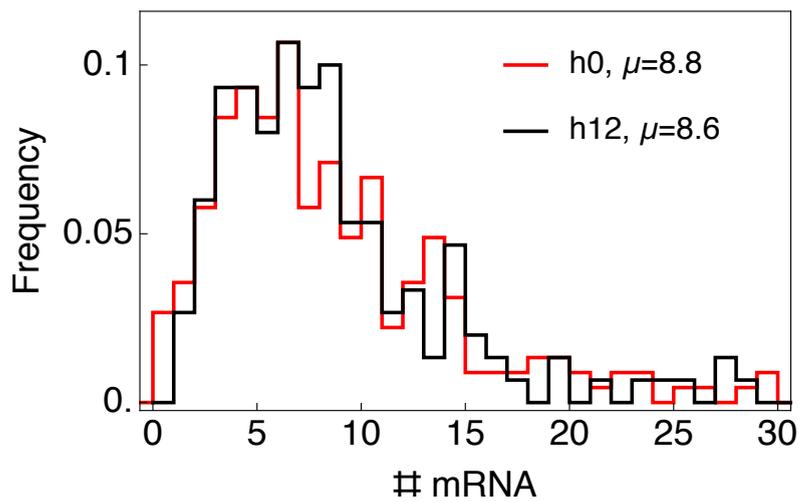


FIGURE 5.5: **Full gene expression recovery.** *PLOD2* transcript number distribution 12 hours after UV-C exposure ($3 J/m^2$) is restored to the initial transcript number distribution in untreated cells. The *PLOD2* transcript number distribution in untreated control cells (h0, $n=225$) and the distribution 12 hours (h12, $n=150$) after $3 J/m^2$ UV-C exposure is similar (Mann-Whitney U test, $p=0.86$), indicating that transcription is fully restored 12 hours after exposure to $3 J/m^2$ UV-C.

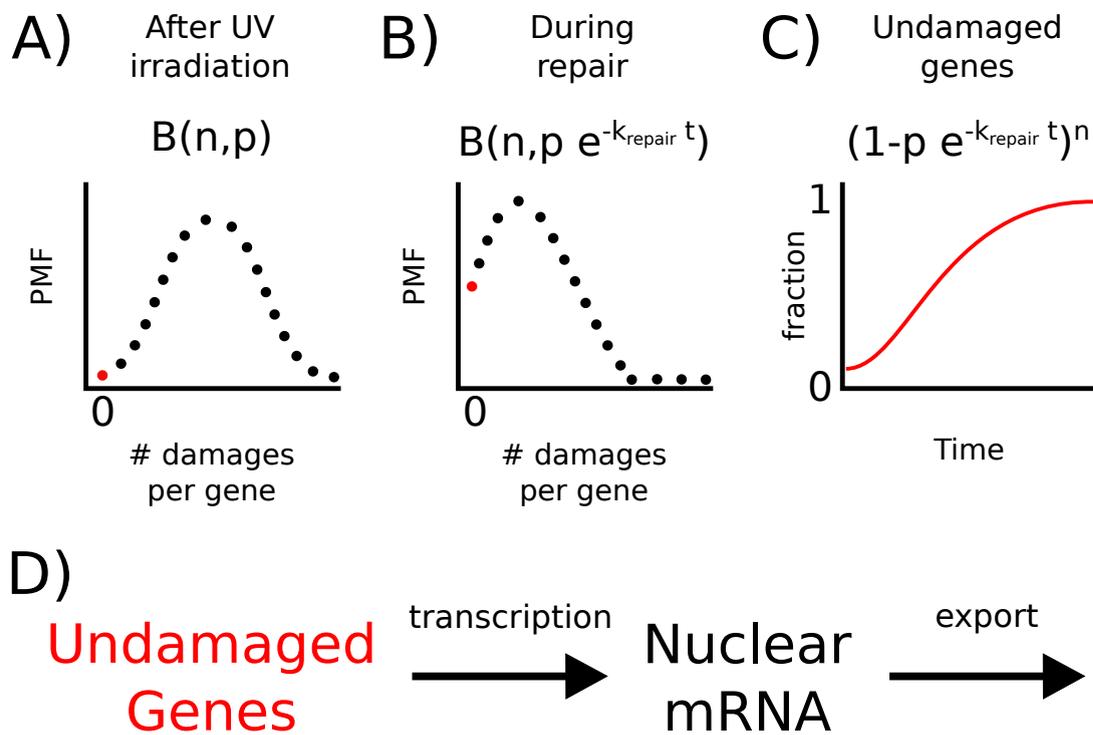


FIGURE 5.6: **Overview of the DNA damage and repair model to estimate the repair rate of UV-C exposure-induced DNA damages.** A) The distribution of the number of damages per gene is given by a binomial distribution $B(n, p)$ in which n stands for gene size and p for the likelihood to acquire a damaged nucleotide. The red dot represents a cell with no damages and is considered to have an unaltered transcriptional activity. B) During DNA repair the probability parameter of the binomial distribution decreases with $e^{-k_{\text{repair}} t}$. During DNA repair the likelihood to remain undamaged increases. C) Given the binomial distribution of DNA damage, the likelihood of no damages is given by $(1 - p \cdot e^{-k_{\text{repair}} t})^n$. This means that the fraction of transcriptionally active cells in a population follows this function. D) The function for the fraction of undamaged genes is coupled to a linear transcription and export model (equation 5.5). We fitted the measured mRNA transcripts in the nucleus in time to obtain the repair rate of DNA damages with the model.