Robust DNA Repair through Collective Rate Control

Inaugural - Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by

Tim Heinemann
born in: Frankfurt (Oder)

December 2014

Referees: **Prof. Dr. Thomas Höfer**
Theoretical Systems Biology
German Cancer Research Center, Heidelberg

Prof. Dr. Ursula Kummer
Modellierung Biologischer Prozesse
Ruperto-Carola University, Heidelberg

Die vorliegende Arbeit wurde in der Abteilung Theoretische System Biologie am Deutschen Krebsforschungszentrum (DKFZ) in Heidelberg zwischen Februar 2011 und Dezember 2014 durchgeführt.

Declaration

The applicant, Tim Heinemann, declares that he is the sole author of the submitted dissertation and no other sources or materials from those specifically referred to have been used.

In addition, the applicant declares that he has not applied for permission to enter examination procedure at another institution and this dissertation has not been presented to other faculties and not used in its current or in any other form in another examination.

Heidelberg, November 12, 2014	______________________
Place, Date	Tim Heinemann

Contents

Zusammenfassung i

Summary iii

1 Introduction 1

- 1.1 DNA damage . . . 1
  - 1.1.1 Sources of DNA damage . . . 1
  - 1.1.2 Types of DNA damage . . . 1
  - 1.1.3 Variety of DNA damage repair mechanisms . . . 1
- 1.2 Experimental research on NER . . . 1
  - 1.2.1 Key components of the NER pathway . . . 1
  - 1.2.2 Expression of fluorescent NER factors . . . 1
- 1.3 Mathematical models of NER . . . 1
  - 1.3.1 Metabolic control analysis . . . 1

2 Mathematical Modelling of Nucleotide Excision Repair (NER) 3

- 2.1 Fluorescent time-laps imaging of the NER process . . . 3
  - 2.1.1 Locally inflicted UV-induced DNA damage . . . 3
  - 2.1.2 Repair factor accumulation and dissociation occur on different time scales . . . 5
  - 2.1.3 Direct measurement of DNA resynthesis . . . 9
  - 2.1.4 Repair rate follows first order rate kinetic . . . 9
- 2.2 Kinetic model of NER . . . 12
  - 2.2.1 Slow first-order reaction kinetic due to many fast interacting components 12
  - 2.2.2 Model structure and parametrization . . . 13
- 2.3 Maximum likelihood approach together with PLE analysis identifies realistic model of NER . . . 18
  - 2.3.1 A maximum likelihood approach for efficient model fitting . . . 18
  - 2.3.2 NER model fits accumulation, FLIP and repair synthesis measurements 22

2.3.3 Identifiability analysis . . . 22

3 Control analysis of the DNA repair rate 27

3.1 Kinetic NER model predicts collective rate control . . . 27

3.1.1 Response coefficients . . . 27

3.1.2 Exploiting natural variability in protein expression to quantify rate control 31

3.1.3 NER Variability comparable between complemented patient cell lines and native repair . . . 35

3.2 Sources for repair rate variability . . . 37

3.2.1 Variable NER factor expression and inflicted lesions account for the distribution of repair rates . . . 37

3.2.2 Imbalanced contribution . . . 37

4 Co-expression of NER repair factors 41

4.1 Nuclear expression of NER factors is strongly correlated . . . 41

4.1.1 Microscopy analysis of co-staining experiments . . . 41

4.1.2 Flow cytometry . . . 44

4.1.3 Calibration of the response coefficients . . . 44

5 Discussion 47

6 Bibliography 49

7 Appendix 53

Zusammenfassung

The nucleotide-excision repair pathway removes mutagen-inflicted DNA lesions from the genome. Repair proteins recognize DNA lesions and form multi-protein complexes that catalyze the excision of the lesion and the re-synthesis of the excised part. Imaging the dynamics of fluorescently labeled repair proteins in living human cells has revealed that all factors continuously and rapidly exchange at repair sites. We asked how this dynamic mode of protein-complex assembly shapes the repair process. Measuring repair DNA synthesis in intact cells, we obtained a surprisingly simple result. Over the entire process, the rate is proportional to the amount of DNA lesions, where the proportionality factor is a single 'slow' rate constant. Such kinetic behavior is often regarded as evidence for a rate-limiting step, but we show here that it is an emergent property of the dynamic interplay of many repair proteins. As a consequence, the rate of DNA repair is a systems property that is controlled collectively by the expression levels of all repair factors. Given that transcription in living cells has similar dynamic features - rapidly exchanging components of the transcription machinery and slow bursts of mRNA synthesis - collective rate control might be a general property of chromatin-associated molecular machines.

Summary

DNA repair is indispensable for the intracellular protection against environmental and endogenous damaging agents. This is reflected in an increased susceptibility against cellular aging and cancer development as a consequence to impaired repair. Functional repair is carried out by enzymatic macromolecular complexes that assemble at specific sites on the chromatin fiber. How the rate of these molecular machineries is regulated by their constituent parts is poorly understood. Here we quantify nucleotide-excision DNA repair (NER) in mammalian cells and find that, despite the pathways' molecular complexity, repair effectively obeys slow first-order kinetics. Theoretical analysis indicates that these kinetics are not due to a singular rate-limiting step. Rather, first-order kinetics emerge from the interplay of rapidly and reversibly assembling repair proteins, stochastically distributing DNA lesion repair over a broad time period. Based on this mechanism, the model predicts that the repair proteins collectively control the repair rate. Exploiting natural cell-to-cell variability, we corroborate this prediction for the lesion-recognition factors XPC and XPA. Our findings provide a rationale for the emergence of slow time scales in chromatin-associated processes from fast molecular steps and suggest that collective rate control might be a widespread mode of robust regulation in DNA repair and transcription.

1 Introduction

1.1 DNA damage

1.1.1 Sources of DNA damage

1.1.2 Types of DNA damage

1.1.3 Variety of DNA damage repair mechanisms

1.2 Experimental research on NER

1.2.1 Key components of the NER pathway

1.2.2 Expression of fluorescent NER factors

comparison between CPDs and 6-4PP

1.3 Mathematical models of NER

1.3.1 Metabolic control analysis

2 Mathematical Modelling of Nucleotide Excision Repair (NER)

An essential advantage for the research on Nucleotide Excision Repair (NER) is the ability to inflict UV-induced DNA damages within a locally confined area in the cell nucleus [1]. This technique reformed the investigation of chromatin associated processes whose *in vivo* analysis was, until then, focusing on enzyme mobility and exchange dynamics at steady state performed with photobleaching experiments [2, 3]. In contrast, the enclosed distribution of single stranded DNA damages causes the local activation of a multi-protein repair machinery which in turn allows to study the *de novo* assembly and dissociation kinetics of single repair factors. Exploiting this standard experimental setup we acquired a comprehensive data set comprising the accumulation and dissociation time series of seven individual repair enzymes. In addition, we were able to directly quantify newly synthesized DNA measuring the incorporation of a fluorescently modified nucleotide analog.
Based on this data we developed a mathematical model that is able to explain the connection between the dynamic exchange of individual repair factors at the chromatin on one hand and the overall time scale of the repair process on the other. The exact parametrization of the model was adapted iteratively in correspondence with a profile likelihood estimation leading to identifiable parameters and hence an increased predictive power of the model. As a consequence our model can reliably predict the behavior of an experimentally not accessible DNA intermediate within the repair process.

- clarify who did what...

2.1 Fluorescent time-laps imaging of the NER process

2.1.1 Locally inflicted UV-induced DNA damage

To observe the dynamic behavior of NER factors upon local infliction of DNA damage we applied the experimental setup developed by Moné *et al.* (2001) [1]. Therein, NER competent cells are covered with a polycarbonate UV filter containing pores with a given diameter (cf. Figure 2.1A). For our purposes cells were grown on uncoated 24 mm coverslips

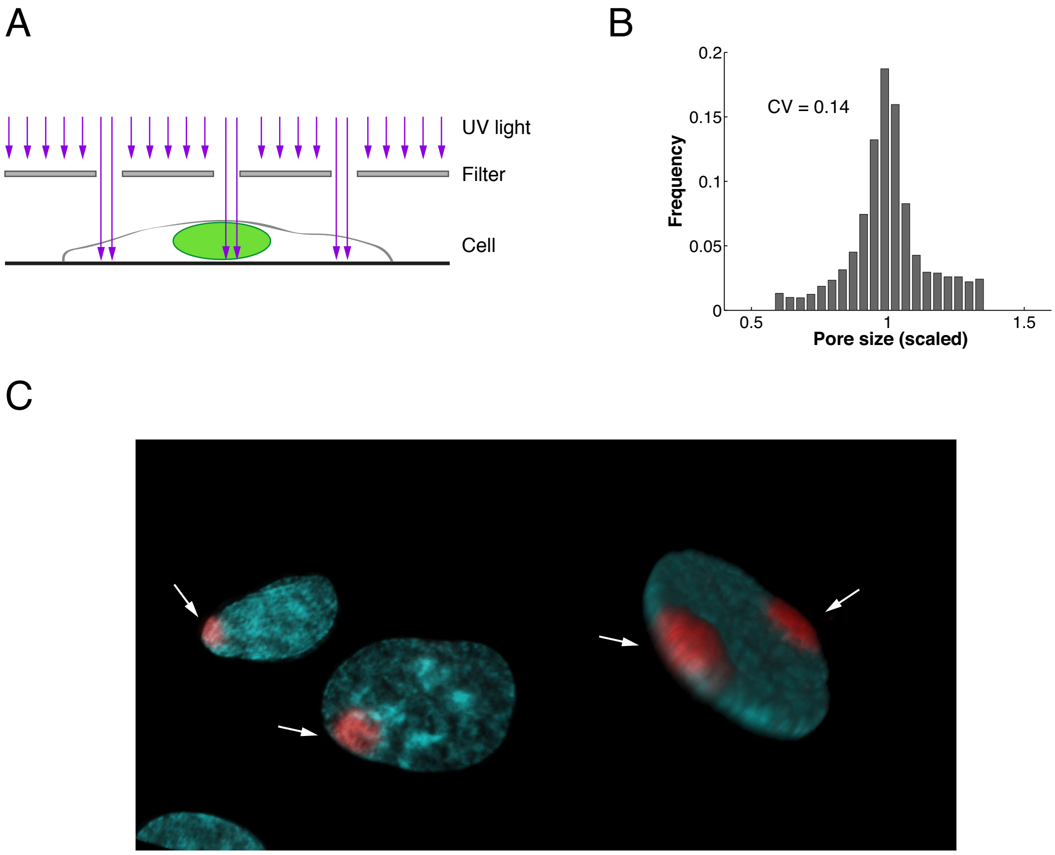
A
UV light
Filter
Cell
B
0.2
0.15
0.1
0.05
0
Frequency
CV = 0.14
0.5
1
1.5
Pore size (scaled)
C

Figure 2.1: Local irradiation leads to spatially confined DNA damage. A) Schematic illustration of the experimental setup. UV-light (purple arrows) is transmitted through a UV filter containing pores with 5 µm diameter. Local irradiation of the chromatin occurs, if a pore is located above a cell nucleus. (**reprint of**) B) Distribution of the pore sizes within the UV irradiation mask (n=5008) - **ask paul how he did it** C) Partially UV-irradiated nuclei of cultivated mammalian cells were made visible with a blue-fluorescent DNA stain. Red fluorescence depicts incorporation of nucleotides due to Nucleotide Excision Repair of damaged DNA.

before overlaid with a mask including pores with 5 µm diameter [4]. Figure 2.1B illustrates the realistic deviation of pore diameters due to manufacturing inaccuracies. The actual size distribution has a very small CV of 0.14 indicating a negligible effect on the filter's transmissivity. Filter-covered cells were immediately irradiated with a dose of 100 J/m^2 of UV-C with a fluency of 3.85 W/m^2. Due to the specific pore density (4 x 10^5 pores/cm^2) we could observe nuclei containing either one damage spot or no spot at all. Very rarely two spots per cell were present (cf. Figure 2.1C, white arrows indicate damage spot). Consequently, with this technique one can measure NER in damaged nuclei as well as undamaged control nuclei in the same experiment.

2.1.2 Repair factor accumulation and dissociation occur on different time scales

The ability to locally inflict DNA damage with a discrete dose of UV-C allows to study the accumulation and exchange behavior of fluorescently tagged repair proteins under different experimental conditions. In the following we describe the comprehensive dataset acquired by Luijsterburg *et al.* (2010) [5], which represents the basis of our quantitative analysis of the NER process. In total, the kinetics of seven repair factors were measured: i) XPC, the lesion recognition factor; ii) TFIIH, the helicase responsible for DNA unwinding; iii) XPA and RPA, which bind and thereby protect single stranded DNA against cleavage iv) the exonucleases XPF/ERCC1 and XPG performing the incision of the damaged DNA strand and v) PCNA, which loads the DNA polymerase and hence indirectly provides insight into the DNA repair-synthesis kinetic.

Each repair factor is tagged with a green fluorescent protein (GFP) (or its 'enhanced' derivative EGFP) and expressed at physiological levels within the cell nucleus. Before the cells are UV-irradiated at time t = 0 the repair proteins are homogeneously distributed. Immediately afterwards the repair factors accumulate at the damaged chromatin sites, which leads to a higher visible fluorescence intensity (cf. Figure 2.2A). For quantification of the fluorescence intensity image analysis was done by using ImageJ software (NIH Bethesda, MD). Accumulated repair-factor concentrations were determined by multiplying the nuclear reference concentrations (cf. Table 2.1) with the fraction of bound proteins at the damaged DNA:

$$\text{Bound fraction} = (I_{\text{LD}} - I_{\text{outspot}})A_{\text{LD}}/(I_{\text{nucleus}} - I_{\text{background}})A_{\text{nucleus}} \qquad (2.1)$$

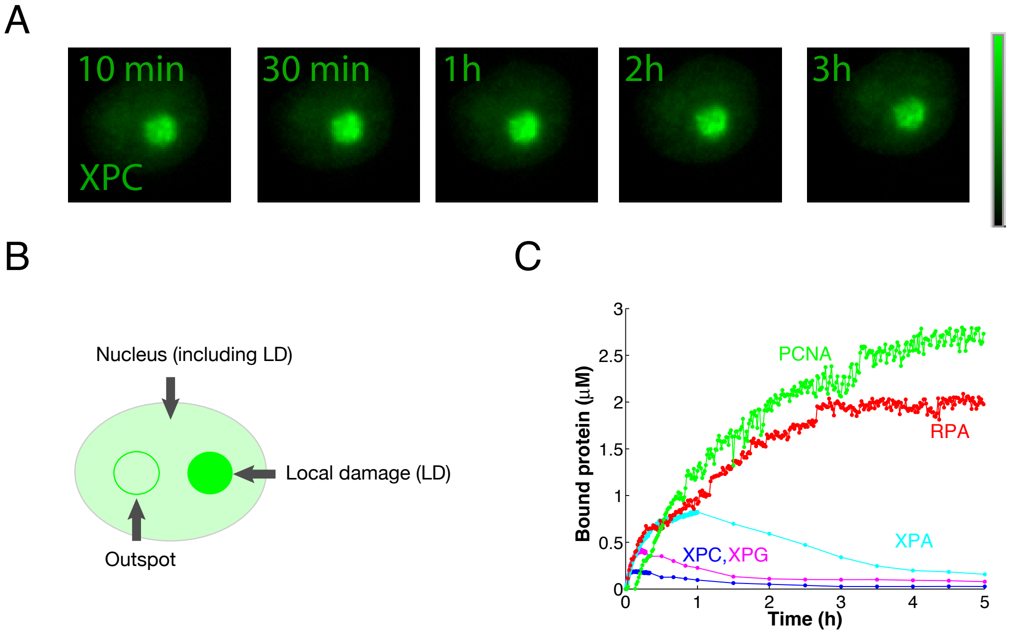

Figure 2.2: Fluorescently labeled NER factors accumulate at locally irradiated nucleus. A) XPC-eGFP accumulation stably expressed in XP-C cells at different time points after local irradiation with UV-C (100 Jm^{-2} through 5-μm-diameter pores). B) Scheme of a locally UV-irradiated nucleus. Depicted are all regions relevant for signal quantification. C) Time courses of XPC-eGFP (n = 12), XPG-eGFP (n = 5), eGFP-XPA (n = 7), eGFP-PCNA (n = 5) and RPA-eGFP (n = 5) showing their accumulation at the local damage (LD) spot. For consistency only cell nuclei with one single damage spot were used.

where I_{LD}, I_{outspot}, I_{nucleus} and $I_{\text{background}}$ represent the average fluorescence intensities within the locally damaged spot, an equally sized area in the non-damaged nucleus, the whole nucleus and the background, respectively (cf. Figure 2.2B). A_{LD} and A_{nucleus} depict the damaged area and the size of the nucleus. Finally, the concentrations of accumulated protein are calculated assuming a damaged nuclear volume of 0.03 pL.

Protein	Concentration	Bound fraction
XPC	0.140 µM	13%
TFIIH	0.360 µM	10%
XPG	0.440 µM	9%
XPA	1.110 µM	7%
XPF/ERCC1	0.170 µM	7%
RPA	1.110 µM	15%
PCNA	1.110 µM	20%

Table 2.1: Nuclear concentrations of NER factor in (in µM) All nuclear quantities are based on published data or on previous estimates [3, 6, 7]. The nuclear concentrations are calculated assuming a nuclear volume of 0.3 pL. The bound fractionis determined by Eqn. 2.1 [5].

During the timespan of DNA repair (cf. Section 1.2) NER factors accumulate towards and then gradually decrease from their plateau levels at different rates (cf. Figure 2.2C). For example the half-life $t_{1/2}$ for XPC- and XPG-EGFP is $\sim$1 h whereas $\sim$2.5 h for XPA-EGFP [5]. Moreover, PCNA and RPA stay present in the damage spot even after lesion removal. These result show that NER factors engage for hours in the repair process with temporal changes in the molecular composition.

To characterize the interaction between repair proteins and DNA intermediates dwell times were determined in fluorescence loss in photobleaching (FLIP) experiments [5]. Thereby, a large part of the nucleus, away from the local damage spot, is continuously bleached at 100 % laser power (cf. 2.3A, white rectangle). At the same time fluorescent proteins are probed at low laser power elsewhere within the nucleus. Repair proteins dissociating from the local damage spot have a high probability to be bleached before rebinding due to their large diffusivity. Accordingly, for the dwell time of the repair proteins binding instead of diffusion seems to be rate limiting [5].

Still, compared to their long overall presence in the range of hours, all NER factors dissociate very quickly from damaged DNA with half-lives of 20 s (RPA), 25 s (XPC), 50 s (TFIIH, XPG, ERCC1/XPF) and 80 s (RPA) (cf. Figure 2.3B). For PCNA the dissociation is strongly biphasic with half-lives of 10 s and 225 s respectively. To analyze, whether the dwell time of slowly accumulating NER factors changes throughout the repair process resynthesis was stalled by addition of the drugs hydroxyurea (HU) and cytosine-β-arabinofuranoside. While NER factors reaching their plateau level earlier were not affected [5], omitting the repair progression at this late stage mainly slowed down the dissociation of PCNA and RPA (cf. Figure 2.3C). In contrast, XPA's half-life decreased by twofold indicating its higher affinity to repair synthesis intermediates. This shows that the dwell times of NER factors change as repair progresses and suggests that their affinity towards damaged chromatin is defined by the state of the DNA substrate.

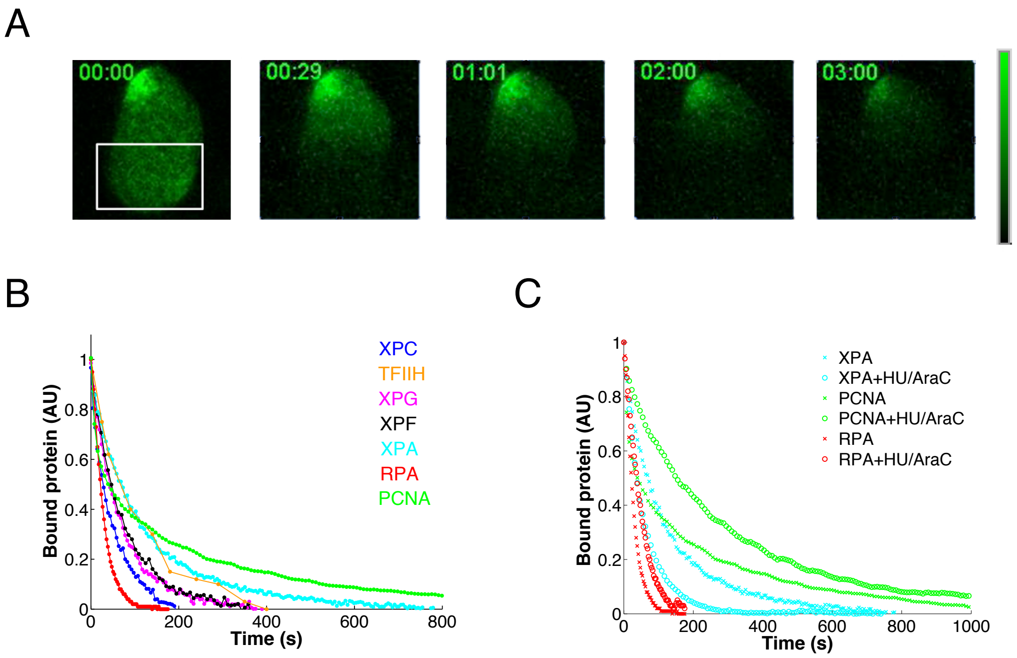

Figure 2.3: Rapid dissociation of NER factors from damaged DNA. A) FLIP experiment in XP2OS cells stably expressing eGFP-XPA after 2 h of local irradiation. The nucleus is continuously bleached in an undamaged region (white rectangle). Fluorescence loss is recorded at the sites of local UV-irradiation B) Amount of bound NER-factors (XPC-eGFP, TFIIH-eGFP, XPG-eGFP, XPF-GFP, eGFP-XPA, RPA-eGFP and eGFP-PCNA) monitored over time at LD. C) Quantification of FLIP experiments in the absence or presence of HU and AraC for stably expressed eGFP-XPA, RPA-eGFP and eGFP-PCNA.

2.1.3 Direct measurement of DNA resynthesis

In order to expand our quantitative inspection of the NER pathway in intact mammalian cells we established a protocol for the direct measurement of the repair synthesis process. DNA resynthesis reflects the kinetic of the post-incision repair process, complementing the measure of DNA lesion removal, which captures the systems behavior of the pre-incision steps (cf. section 1.2). To visualize newly incorporated DNA shortly before and after local UV irradiation cells were incubated in microscopy medium supplemented with 10 µM 5-ethynyl-2'-deoxyuridine (EdU). Due to its excessive presence in solution the DNA polymerase integrates EdU (a thymidine analog) into the new DNA strand instead of the endogenous thymidine (cf. Figure 2.4A). After incubation for the desired time cells were fixed stopping EdU incorporation and subsequently permeabilized. EdU is then tagged with the fluorescent azide (AlexaFluor-555, Life Technologies) forming a covalent bond by click chemistry [8]. Analogous to the quantification of NER factor dynamics EdU intensities were captured with a laser scanning microscope (Zeiss).

Accordingly, incorporated EdU is exclusively present at the locally confined area of damaged DNA, which coincides with the localization of immunostained XPC (cf. Figure 2.4B, upper row). The same result occurs for XP-C cells with stably transfected XPC-eGFP (cf. Figure 2.4B, lower row). Replicating cells could be excluded easily from the analysis due to there prevalent EdU incorporation distributed over the entire nucleus.

2.1.4 Repair rate follows first order rate kinetic

In contrast to the real time measurements for accumulation and dissociation of the NER factors the incorporation of EdU cannot be followed continuously. As mentioned in section 2.1.3 cells have to be fixed and permeabilized before the newly incorporated bases can be fluorescently labeled. Therefore, only the accumulated EdU incorporated in the time interval between UV-irradiation and fixation can be followed. By repeating this procedure for growing time intervals we acquired successively the repair kinetic for newly repaired DNA (cf. Figure 2.5A and B). For each time point we averaged over multiple cells.

We found that EdU incorporation essentially stops after 4 h, which coincides with the removal of 6-4PP (cf. [5] and Figure 2.5B red and blue curve, respectively). This agrees with the

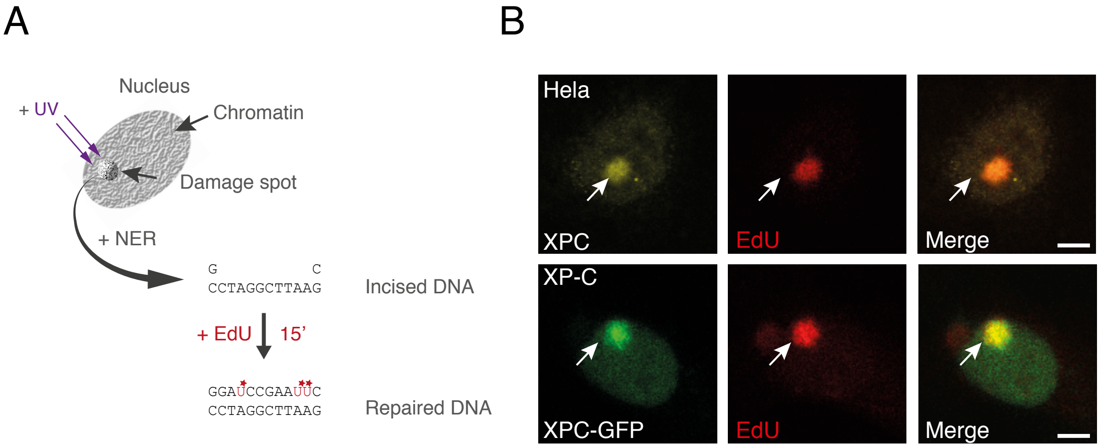
A
B
Nucleus
+ UV
Chromatin
Damage spot
+ NER
G C
CCTAGGCTTAAG
Incised DNA
+ EdU
15'
GGAUCCGAAUUC
CCTAGGCTTAAG
Repaired DNA
Hela
XPC
EdU
Merge
XP-C
XPC-GFP
EdU
Merge

Figure 2.4: EdU is incorporated at sites of local DNA damage. A) Schematic illustration depicting the experimental procedure of EdU incorporation of locally damaged cells. B) Endogenous XPC in Hela cells (upper panel) or stably express XPC-eGFP (lower panel) accumulate in the LD spot and co-localize with incorporated EdU.

observation that NER is not primarily engaged in the removal of cyclobutane pyrimidine dimers (CPD) which are repaired on a much longer timescale. To test whether the availability of EdU is rate limiting we measured the incorporation of EdU in discrete equidistant time intervals after UV-irradiation (cf. Figure 2.5C and Appendix XX). As it turns out the amount per time interval of EdU incorporation is indeed continuously declining and hence the rate of repair synthesis. Moreover the EdU kinetic follows the trajectory of PCNA accumulation as measured by Luijsterburg *et al.* [5] (cf. Figure 2.5D). PCNA in turn is thought to act as processivity factor for the DNA polymerase and remains bound to the DNA [5, 9, 10]. Taken together, these data establish EdU incorporation as a direct and quantitative measure of DNA repair synthesis in locally damaged cells.

The DNA repair time series characterized by incorporated EdU is fitted by a mono-exponential kinetic (cf. Figure 2.5E):

$$EdU(t) = EdU_{\text{max}}(1 - e^{\lambda t}), \tag{2.2}$$

where $EdU(t)$ and EdU_{max} depict the amount of incorporated EdU and its value at saturation, respectively. The time constant is λ=0.58 ($\pm$0.07) h^{-1}. This result indicates that, despite its molecular complexity, 6-4PP removal by NER is a slow first-order reaction with a half-time of 1.2 hours.

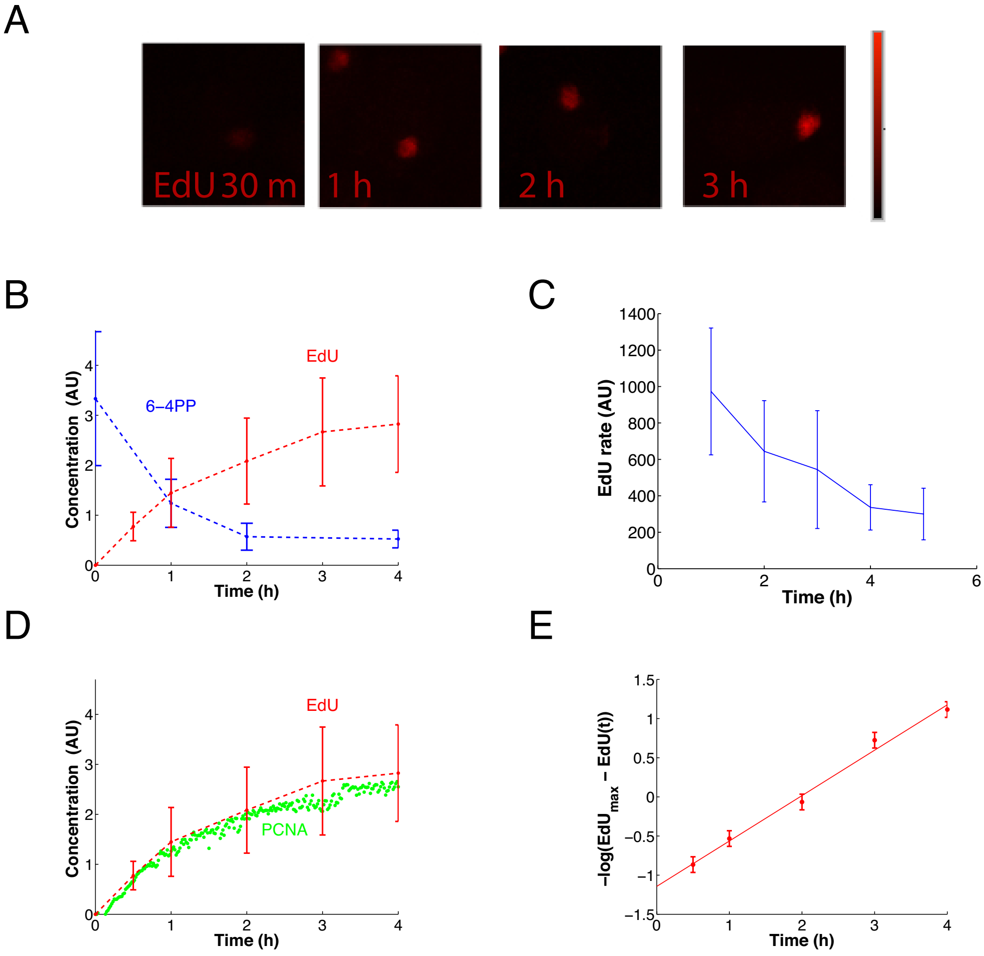

Figure 2.5: EdU incorporation reflects DNA repair synthesis quantitatively. A) EdU signal at various time points after local UC-C irradiation. B) Repair DNA synthesis (EdU, red curve) coincides with damage removal (6-4PP, blue curve) measured previously by Luijsterburg *et al.* [5]. The EdU trajectory represents mean $\pm$ SD (derived from three independent experiments) of DNA repair in n=150 locally damaged cells per time point. C) D) DNA repair synthesis follows PCNA accumulation as measured previously [5]. E) Mono-exponential fit of the EdU kinetic according to Eqn. 2.2.

2.2 Kinetic model of NER

2.2.1 Slow first-order reaction kinetic due to many fast interacting components

To investigate the mechanistic connection between fast NER factor exchange at the DNA template and the overall slow repair time it appears practical to apply an analytical approach. During her Ph.D. thesis Gesa Terstiege together with Thomas Höfer performed this analysis [4,11], considering the complex formation with a simplified model including N repair factors. They examined different scenarios distinguishing random and sequential protein assembly; reversible and irreversible protein binding or mixed mechanisms specific for each DNA repair intermediate. Under appropriate assumptions, the mean repair time of such a process can be generally expressed with:

$$\tau = \underbrace{\frac{1}{k}\sum_{i=0}^{N-1} A_i \left(\frac{l}{k}\right)^i}_{\text{first assembly}} + \underbrace{\frac{1}{\rho}\sum_{i=0}^{N} B_i \left(\frac{l}{k}\right)^i}_{\text{reassembly and reaction}} \tag{2.3}$$

where k denotes the pseudo-first order association rate constant of NER factors, l the dissociation rate constant of NER factors and ρ the rate of the repair reaction (cf. Figure **tbd**A, [4]). A_i and B_i differ between random assembly

$$A_i^{\text{rand}} = \sum_{j=1}^{N-i} \frac{1}{i+j} \frac{\binom{N}{j-1}}{\binom{N}{i+j}}, \quad B_i^{\text{rand}} = \binom{N}{i}, \tag{2.4}$$

and sequential assembly

$$A_i^{\text{seq}} = N - i, \quad B_i^{\text{seq}} = 1. \tag{2.5}$$

Sequential and random assembly occur on a similar time scale for a number of repair components smaller than 10 considering reversible protein binding ($l = k = 1$) [11], which coincides with the number of core NER factors [5]. The theoretical component number for equal repair times is even higher for increasing reaction rates ($\rho \gg k, l$). Given that all repair factor dwell times are in the order of ~1 min (k = l = 1 min^{-1}) the model was simulated for N = 9 components (cf. Figure **tbm**C). Remarkably, the resulting trajectory followed a single-exponential kinetic with a half time of ~1 hour in case of balanced reversibility. In contrast, for an irreversible process ($l/k = 0$) the repair kinetic resembles a sigmoid time course (cf. Figure**tbm**B). These results suggest that the rapidly exchanging NER factors naturally generate a mono-exponential repair kinetic whereas the slow overall time span can be explained by the stochastic distribution of repair events.

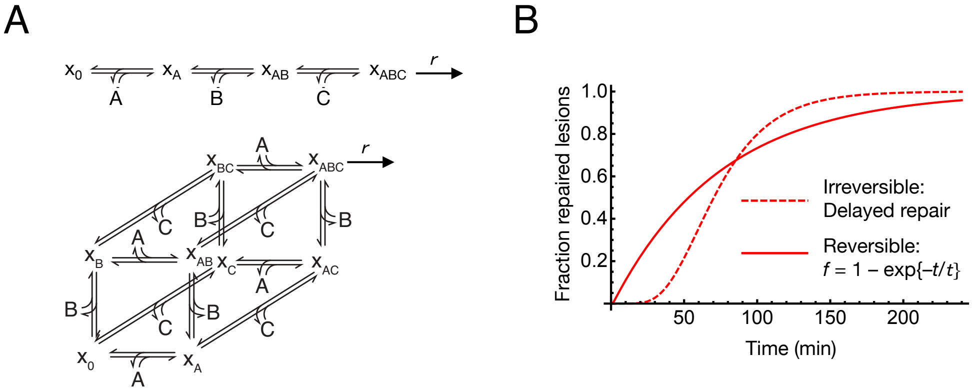

Figure 2.6: Simplified analytical model of the repair complex formation A) Sequential (above) and random (below) assembly scheme of the repair factors A, B and C to the DNA template x_i. x_0 denotes the empty lesion (e.g. damaged DNA). When the complex is fully assembled (x_{ABC}) it performs a catalytic repair step with rate r. B) Simulated repair time courses for random reversible (solid line) and irreversible (dashed line) repair factor assembly. For reversible binding ($k = l =$ 1 min$^{-1}, N =$ 9) the trajectory fits a mono-exponential repair kinetic with time constant τ whereas for irreversible binding the fit has sigmoidal shape ($l =$ 0$, k =$ 0.037 min$^{-1}, N =$ 9, chosen to get the same time constant).

2.2.2 Model structure and parametrization

To examine the relation between rapid repair factor exchange and the slow first-order reaction kinetic we extended this analysis on the basis of a realistic NER model. Conceptually, it follows a model worked out by Luijsterburg *et al.* (2010) [5]. According to this, NER factors bind transiently to DNA repair intermediates thereby forming catalytic complexes that, if complete, perform the next repair step. These are usually irreversible reactions embodying the sequential characteristics of this pathway (cf. Figure **tbm**).

The nature of these distinguishable repair intermediates is widely investigated in biochemical but also in *in vivo* studies [12–15]. Accordingly, the evolved model distinguishes five DNA repair intermediates: i) damaged DNA, unwound DNA, incised DNA, resynthesized DNA

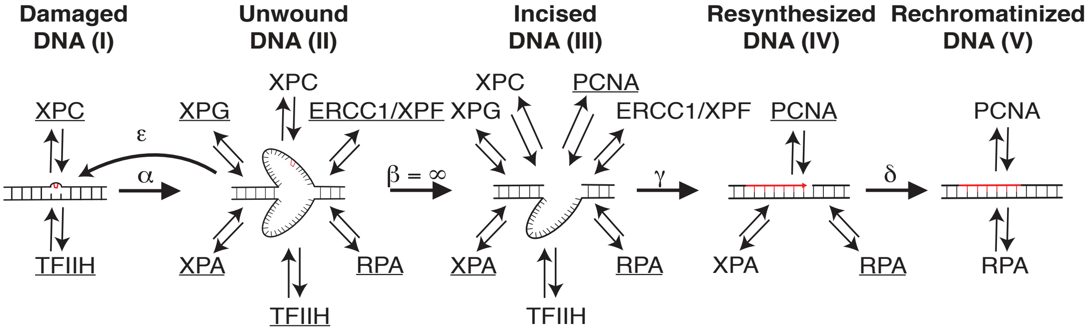

Figure 2.7: Schematic model description of the DNA repair mechanism. The model distinguishes five individual repair intermediates: Damaged (I), unwound (II), incised (III), resynthesized (IV) and rechromatinized DNA (V). As indicated, specific tuples of repair proteins bind reversibly to the intermediates. Catalytic reactions, denoted by Greek letters, occur when the particular reaction complex (underlined proteins) is assembled.

and rechromatinized DNA. The molecular state of these intermediates defines the binding affinity for a specific set of repair proteins (cf. Figure 2.7). Table 2.2 summarizes all repair intermediates and lists the corresponding repair factors that show affinity for a particular repair intermediate. Repair factors that catalyze the enzymatic reaction if assembled are indicated. As stated before, FLIP measurements indicate that diffusion is not rate limiting for protein binding and thus not included in the model (cf. Section 2.1.2 and [16, 17]).

A prominent role within the NER pathway belongs to the lesion recognition factor XPC. Theoretical results suggest that the detection of damaged DNA sites by a single element instead of multiple simultaneously is crucial for the efficient initiation of the repair process [18]. Beside this exception the model allows random, non-cooperative binding binding for all NER factors subsequently assembling at the chromatin.

The model structure introduced above and shown in Figure 2.7 was translated into an ordinary differential equation system assuming mass-action kinetics for all protein-DNA interaction processes. Each equation describes the concentration of a single DNA state y_π^R (cf. Eqn. 2.6) associated with a specific repair intermediate R = I, II,...,V (damaged (I), unwound (II), incised (III), resynthesized (IV), rechromatinized (V)). The index π represents a binary vector where each position displays the presence or absence of a repair protein p (p $\in$ C, T, G, A, F, R, P, where XPC (C), TFIIH (T), XPG (G), XPA (A), ERCC1/XPF (F), RPA (R) and PCNA (P)). $\pi(p) = 1$ if protein p is bound and $\pi(p) = 0$ if not. The length of $\pi(p)$ is defined according to the model structure (cf. Table 2.2) with: $\pi_{R=I} = \{\text{C,T}\}$; $\pi_{R=II} = \{\text{C,T,G,A,F,R}\}$; $\pi_{R=III} = \{\text{C,T,G,A,F,R,P}\}$; $\pi_{R=IV} = \{\text{A,R,P}\}$; $\pi_{R=V} = \{\text{R,P}\}$. The time course of this model is then described by:

$$\frac{d}{dt}\, y_\pi^R = \sum_{p \text{ in } R} \eta \, \Big((-1)^{\pi(p)}\, l_p^R\, y_\pi^R \,|_{\pi(p)=1} + \, (-1)^{1+\pi(p)}\, k_p^R\, C_p(t)\, y_\pi^R \,|_{\pi(p)=0} \Big) + E(y_\pi^R), \qquad (2.6)$$

where η represents a cooperativity ensuring the exclusion of all cooperative binding events:

$$\eta = \begin{cases} = 0 \quad \text{if} & R = \text{I} \wedge p = \text{C} \wedge \text{T} = 1, \\ & R = \text{I} \wedge p = \text{T} \wedge \text{C} = 0, \\ = 1 \quad \text{else.} & \end{cases}$$

The kinetics describing protein exchange at the repair intermediates are characterized by the binding constant k_p^R and the dissociation constant l_p^R. The free protein concentrations are denoted by $C_p(t)$ representing seven additional differential equations for $p \in \{\text{C,T,G,A,F,R,P}\}$:

$$\frac{d}{dt}\, C_p \;=\; \sum_{R=\text{I}}^{\text{V}} \sum_{\pi} \xi \left(\delta_{\pi(p)1}\, l_p^R\, y_\pi^R - \, \delta_{\pi(p)0}\, k_p^R\, C_p\, y_\pi^R \right) \qquad (2.7)$$

Analogous to η, ξ governs the sequential binding of XPC and TFIIH by:

$$\xi = \begin{cases} = 0 \quad \text{if} & R = \text{I} \wedge \text{C} = 0 \wedge \text{T} = 1, \\ & R = \text{I} \wedge p = \text{C} \wedge \text{C} = \text{T} = 1, \\ & R = \text{I} \wedge p = \text{T} \wedge \text{C} = \text{T} = 0, \\ = 1 \quad \text{else .} & \end{cases}$$

Repair intermediate	**Binding proteins**	**Catalyzed process** **Required proteins**	**Remarks**	**Ref.**
(I) Damaged DNA	XPC,TFIIH (3 states)	Unwinding (reaction α) XPC and TFIIH	Initiation by binding of XPC and subsequent recruitment of TFIIH.	[12] [19] [20] [16] [21]
(II) Unwound DNA	XPC,TFIIH, XPG, XPA, ERCC1/XPF, RPA (64 states)	Dual incision (reaction β) TFIIH,XPG, XPA, RPA and ERCC1/XPF	If the DNA becomes devoid of any protein, it will re-anneal (reaction ε). Dual incision requires the endonucleases XPG and ERCC1/XPF and is stimulated by TFIIH, XPA, RPA and possibly XPC.	[12] [22] [23] [24] [25]
(III) Incised DNA	XPC,TFIIH, XPG, XPA, ERCC1/XPF, RPA, PCNA (128 states)	Repair-synthesis (reaction γ) XPA, RPA and PCNA	PCNA binds to the free 3'-OH group generated by the ERCC1/XPF incision. DNA polymerase is also required (not measured).	[12] [24]
(IV) Resynthesized DNA	XPA, RPA, PCNA (8 states)	Rechromatinization (reaction δ) RPA and PCNA	Accumulation data imply that XPA binds to repaired DNA while the pre-incision proteins do not (Figure 2.2C).	[26] [27] [5]
(V) Rechromatinized DNA	RPA, PCNA (4 states)		RPA and PCNA associate with rep. intermediate I, as levels of bound EGFP-PCNA and EGFP-RPA are high up to at least 4h after UV irradiation while other repair proteins are no longer bound.	[19] [5]

Table 2.2: Model assumptions. Adapted from Terstiege *et al.* (2010) [11]

Finally, if an enzymatic complex has fully assembled at the DNA template it catalyzes the next repair step, which is represented in the model with the term $E(y_\pi^R)$. After damage recognition by XPC damaged DNA is unwound by the helicase TFIIH with the unwinding activity α, whereas XPC acts as a stabilizing/proof reading factor in parallel. Accordingly $E(y_\pi^R)$ translates into the following catalytic reactions for damaged DNA ($R = \text{I}$):

$$E(y_{00}^{\text{I}}) = \varepsilon\, y_{000000}^{\text{II}} \quad \text{and} \quad E(y_{11}^{\text{I}}) = -\alpha\, y_{11}^{\text{I}},$$

If all proteins fall off the DNA template due to false damage detection the DNA will re-anneal with the rate ϵ. Otherwise a complex formed by TFIIH, XPG, XPA, XPF, XPA and RPA will eventually promote the excision of the lesion DNA strand leading to the following catalytic reactions for unwound DNA ($R = \text{II}$):

$$\begin{array}{lcll} E(y_{000000}^{\text{II}}) & = & -\,\varepsilon\, y_{000000}^{\text{II}}\,, \quad E(y_{110000}^{\text{II}}) & = \alpha\, y_{11}^{\text{I}}, \\ \text{and} & & \phantom{-\,\varepsilon\, y_{000000}^{\text{II}}\,,} \quad E(y_{011111}^{\text{II}}) & = -\,\beta\, y_{011111}^{\text{II}}. \end{array}$$

Once the lesion strand is excised with incision rate β the remaining repair steps are irreversible. Incised DNA is resynthesized with the rate γ by the resynthesis complex XPA-RPA-PCNA. XPA is assumed to assemble at post incision repair intermediates as suggested by experiments with inhibited incision that showed accelerated dissociation FLIP kinetics for XPA [5]. This result is supported by a chromatin immunoprecipitation (ChIP) experiment with antibodies against XRCC1- Lig III showing the co-precipitation of XPA and RPA but not XPC and TFIIH [28]. Evidence for the importance of PCNA and RPA for the resynthesis reaction was shown by Shivji *et al.* (1995) [27]. This results into the following catalytic reactions for incised DNA ($R = \text{III}$):

$$E(y_{1111111}^{\text{III}}) \;=\; \beta\, y_{011111}^{\text{II}} \quad \text{and} \quad E(y_{0001011}^{\text{III}}) = \;-\gamma\, y_{0001011}^{\text{III}}.$$

In correspondence to the previously described accumulation measurements (cf. Figure 2.2) only RPA and PCNA stay bound during chromatin remodeling, the last modeled repair intermediate. This leads to the following enzymatic reactions for resynthesized DNA ($R = \text{IV}$)

and rechromatinize ($R = \text{IV}$):

$$\begin{array}{rclrl} E(y^{\text{IV}}_{111}) & = & \gamma\, y^{\text{III}}_{0001011}, & E(y^{\text{IV}}_{011}) = & -\delta\, y^{\text{IV}}_{011}, \\ \text{and} & & & E(y^{\text{V}}_{11}) = & \delta\, y^{\text{IV}}_{011}. \end{array}$$

Since all NER factors can bind independently the DNA template for each repair intermediate there are a total of 2^N repair states where N is the number of repair proteins assembling to the particular repair intermediate. Due to the sequential assembly of TFIIH after lesion detection of XPC the number of states for damaged DNA is reduced to $2^2 - 1 = 3$ states. For the remaining repair intermediates we derive $2^6 = 64$ for unwound DNA; $2^7 = 128$ states for incised DNA; $2^3 = 8$ states for resynthesized DNA and $2^2 1 = 4$ states for rechromatinized DNA. This results in a total of 214 states including seven differential equations for the free NER-factor protein concentrations.

Summing over all repair states and the respective intermediates associated to one repair factor we can simulate the its accumulation kinetic. As initial conditions serve the measured free protein concentrations denoted in Table 2.1 [5, 11] and the initial amount of inflicted damages whose concentration was estimated with $y^{\text{I}}_{00} = 3.33$ µM. The remaining states start at zero. To reproduce the FLIP kinetic for a specific repair factor all corresponding dissociation constants were set to zero at the time of maximal accumulation when the FLIP experiment started. Accordingly FLIP kinetics were acquired after 600 s for XPC and ERCCC1/XPF, 900 s for XPG and TFIIH and 2000 s for XPA and 7200 s for PCNA.

2.3 Maximum likelihood approach together with PLE analysis identifies realistic model of NER

2.3.1 A maximum likelihood approach for efficient model fitting

To find a realistic parameterization for the temporal development of the repair states y^R_π (cf. Eqn. 2.6) the model was mapped to m observables via a function f_z:

$$z(t_i, \theta) = f_z(t_i, y(t_i, \theta), \theta). \tag{2.8}$$

The observables z are parameterized by θ and resemble experimentally derived quantities at time t_i. θ depicts binding, dissociation and catalytic constants adding up to a total of 45 model parameters. Each observable $z_k(t_i, \theta)$ corresponds to the measured data $z_k^\dagger(t_i)$ with intrinsic noise ϵ_{ki}. Its origin can be traced back to a mixture of measurement noise combined with the naturally occurring biological variability. Assuming additive, normally distributed noise leads to $z_k(t_i)^\dagger = z_k(t_i, \theta) + \epsilon_{ki}$ with $\epsilon_{ki} \sim N(0, \sigma_{ki}^2)$. To calibrate the measured data $z_k(t_i)^\dagger$ with the model observables $z_k(t_i, \theta) + \epsilon_{ki}$ we applied a maximum likelihood approach as distance measure which translates for normally distributed noise into:

$$L(z^\dagger|\theta) = \mathit{\Pi}_{k=1}^{m} \mathit{\Pi}_{i=1}^{d_k} \frac{1}{\sqrt{2\pi\sigma_{ki}^2}} \exp\left(-\frac{1}{2\sigma_{ki}^2}\left(z_{ki}^\dagger - z_k(t_i, \theta)\right)^2\right) \tag{2.9}$$

Here, d_k denotes the number of distinct experimental data sets $z^\dagger$ for each observable $k \sim 1...m$, measured at time points t_i with $i \sim 1...d_k$. σ_{ki}^2 are the variance components of the measurement noise of each data point. Instead of maximizing the likelihood it is equivalent and numerically more efficient to minimize its negative logarithm $-2log(L(z^\dagger|\theta))$ instead. In the following we will refer to it as X^2. For the minimization of the X^2 the choice for θ is controlled by σ_{ki}^2 (cf. Eqn. 2.9). As shown by Raue and colleagues for the reliable estimation of the model parameters θ the simultaneous approximation of σ_{ki}^2 together with the model dynamics leads to a statistically more accurate assessment of the model parameters than using noise estimations from preprocessed data [29]. Accordingly σ_{ki}^2 can be considered as parameterized function

$$\sigma_k(t_i, \theta) = f_{\sigma_k}(t_i, z(t_i, \theta), \theta). \tag{2.10}$$

wherein additional parameters are introduced representing the type and magnitude of the modeled noise. Analogous we applied an additive error model for each observable with the parameterized function $\sigma_k(t_i, \theta) = s_a$ and $\epsilon_{ki} \sim N(0, \sigma_{ki}^2)$ where s_a is included in θ.

For reasons of fitting-speed efficiency we implemented our model into the online-available D2D software environment [29] optimized for MATLAB (2011a, The Mathworks Inc., Natick, MA). The integrated fitting procedure applies the trust region algorithm LSQNONLIN, which is pre-implemented in MATLAB. The algorithm requires the derivatives of the objective function with respect to the parameters (cf. Eqn. 2.9). The inner derivatives $dy(t, /theta)/d\theta$, also called sensitivities, provide gradient information about the parameter landscape and thus are crucial to guide the optimization algorithm to the nearest optimum. The sensitivities can be passed in form of sensitivity equations

$$\frac{d}{dt}\frac{dy(t,\theta)}{d\theta} = \frac{\partial f_y}{\partial y}\frac{dy(t,\theta)}{d\theta} + \frac{\partial f_y}{\partial \theta}, \tag{2.11}$$

which represent additional ODEs that are solved in parallel to the original ODE system (cf. Eqn. 2.6) [30]. Applying sensitivity equations instead of a simple finite difference approximation proved to be numerically more accurate and computationally also faster [29]. Both, model and sensitivity ODEs were solved with the CVODES algorithm written in ANSI (**don't forget in the abbreviations list**) standard C [31].

To avoid terminating the optimization procedure in a local minimum we used a 'multi-start' approach by drawing the initial parameter vector systematically using Latin hypercube sampling (LHS) [32]. In contrast to a random sampling approach LHS provides a better coverage of the sampling space by maximizing the distance between successive parameter draws [29].

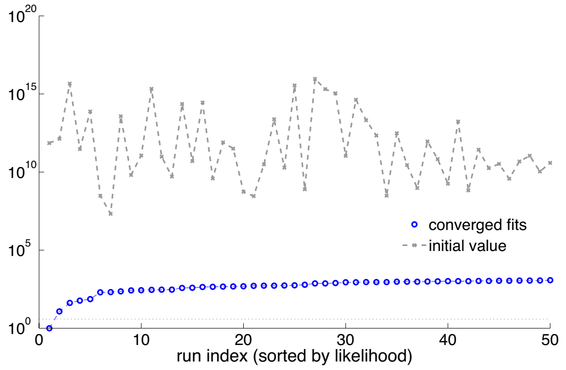

Figure 2.8: Quantitative NER model fits to accumulation and dissociation time courses.

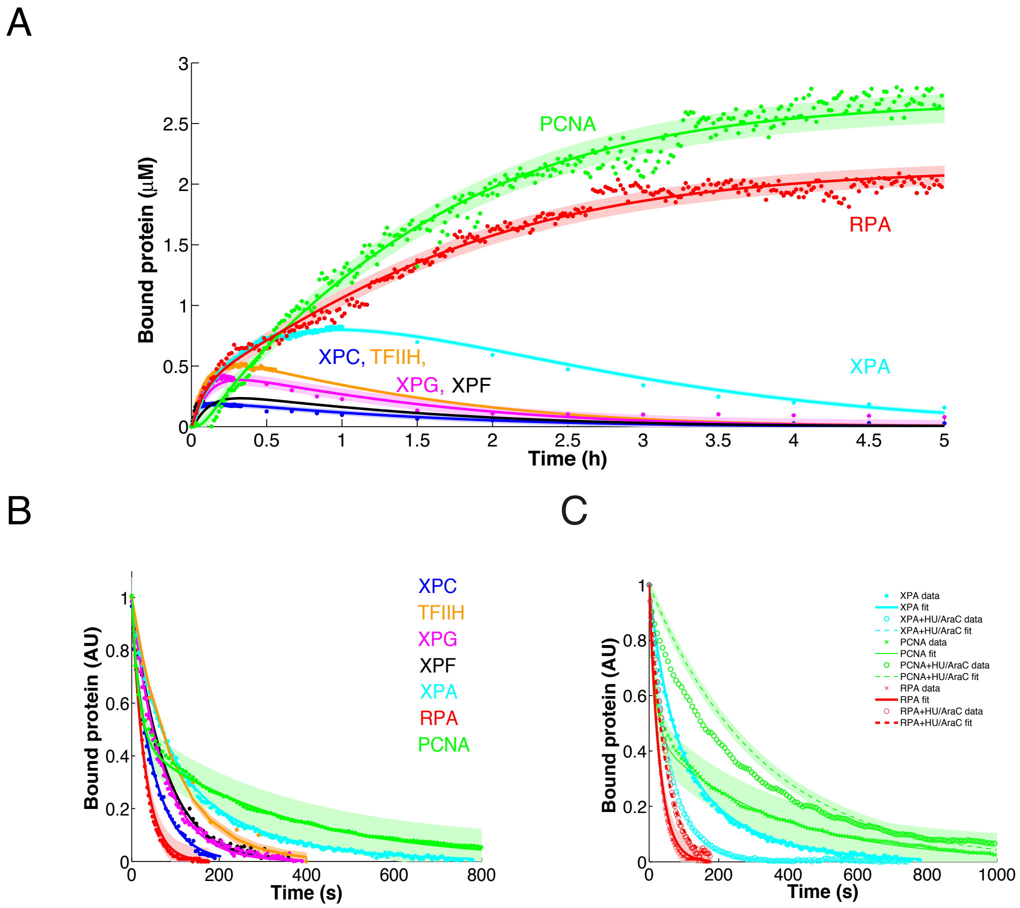

Figure 2.9: Quantitative NER model fits to accumulation and dissociation time courses. A and B) Simulation (lines) and measurement (dots) of the accumulation and dissociation kinetics for the repair factors XPC, TFIIH, XPG, XPA, XPF, RPA and PCNA. Estimated errors are depicted as shaded area. C) Fitted FLIP time courses of XPA, RPA and PCNA in the absence or presence of AraC and HU on locally irradiated cells.)

2.3.2 NER model fits accumulation, FLIP and repair synthesis measurements

To derive faithful fitting results we reiterated the optimization process for 250 times. For each iteration the starting parameters were redrawn by Latin hypercube sampling. Despite the size of the model, with respect to data points and the number of parameters, the majority of fits terminated relatively close to the global X^2 minimum (cf. Figure 2.8) give percentage of the difference). Nevertheless two distinct minimums can be distinguished suggesting a close local minimum, which 'hides' the rarely reached global minimum. For both parameter sets the model fits the experimental data set comprising accumulation, FLIP, perturbation and repair synthesis measurements with small estimated measurement errors (cf. Figure 2.9A, B, C). The accumulation kinetics (cf. Figure 2.9A) are depicted as concentrations scaled by the volume of the locally damaged area which was assumed to be 10% of the nuclear volume. We believe that this specification is more intuitively comprehensible compared to scaling by the whole nuclear volume as performed by Luijsterburg *et al.* [5]. Moreover it allows the realistic comparison between simulated and measured microscopy images of NER factor accumulation and EdU incorporation (cf. Figure 2.10).

2.3.3 Identifiability analysis

In the following section we want to quantify the quality of the model fit and determine whether the current model structure is competent for reliable predictions concerning experimentally unobserved system behavior. This capability depends on the structural and practical identifiability of the model, which can be influenced by functionally related parameters or limited amount and quality of the data, respectively [33, 34]. Both can be analyzed and tested numerically by a formalism called profile likelihood estimation (PLE) [35–37], where the multi-dimensional model uncertainty inflicted by an individual parameter is projected to a one-dimensional profile likelihood (PL)

$$PL(\theta_i) = \max_{\forall j \neq i}[L(z^{\dagger}|\theta_j)]. \tag{2.12}$$

Accordingly, one parameter θ_l with $l \in 1...N$ is gradually fixed along this dimension for different values of p. In each step $X^2_{\theta_l}(p)$ is minimized fitting all other parameters θ_k with $k \in 1...N; k \neq l$. Subsequently, the identifiability of a parameter θ_l can be determined by

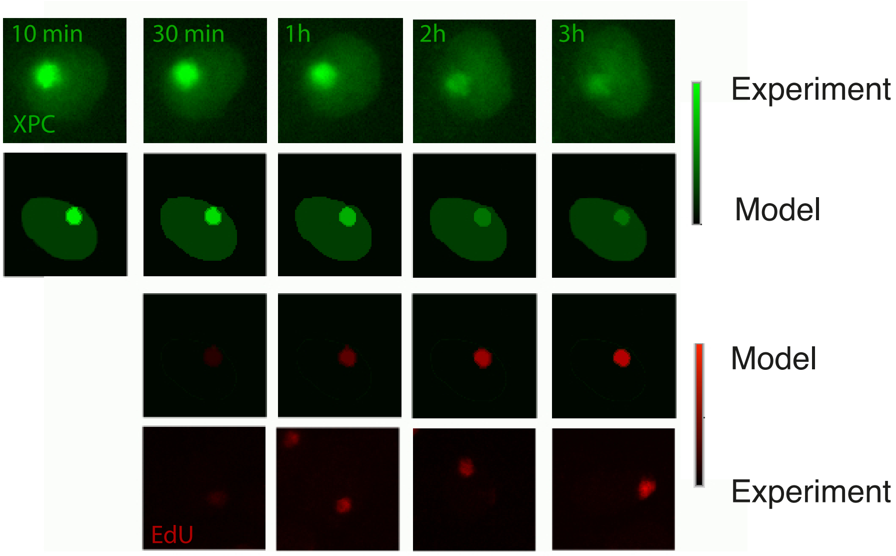
10 min
30 min
1h
2h
3h
XPC
Experiment
Model
Model
Experiment
EdU

Figure 2.10: Comparison between measured and simulated single cell microscopy images Simulated time courses (lines) and the associated microscopy images for GFP tagged XPC (two upper rows) and EdU incorporation (two lower rows). Simulated XPC expression intensities were normalized to the nuclear intensities of the microscopy images. For the EdU incorporation intensities were scaled according to the highest and lowest intensity values of the microscopy images.

$\Delta X^2_{\theta_l}(p)$, which describes the difference between the global X^2 minimum and the parameter dependent local minimum $X^2_{\theta_l}(p)$:

$$\begin{aligned}\Delta X^2_{\theta_l}(p) &= \min_{\{\theta_k|k=1,\dots,N;k\neq l\}} \left(X^2(\theta_1,\dots,\theta_{l-1},p,\theta_{l+1},\dots,\theta_N)\right) \\ &- \min_{\{\theta_k|k=1,\dots,N\}} \left(X^2(\theta_1,\dots,\theta_N)\right) \end{aligned} \tag{2.13}$$

Confidence bounds for the particular parameter depend on a threshold $Q_{X^2}(1-a,df)$, which represent the $(1-a)$ quantile of the X^2-distribution with df degrees of freedom. The associated pointwise confidence intervals are defined as

$$CI_{1-a}(\theta_l) = \{p|\Delta X^2_{\theta_l}(p) \leq Q_{X^2}(1-a,1)\}. \tag{2.14}$$

For one fixed parameter at a time and thus one degree of freedom we can derive the common confidence region $CI_{95\%}$, which corresponds to a X^2-distribution quantile of $Q_{X^2}(95\%,1) = 3.8$. A parameter θ_l is identifiable, if the confidence interval $CI_{1-a}(\theta_l)$ is finite, which can be determined directly from the graph of the profile likelihood $\Delta X^2_{\theta_l}(p)$ for different values of p (cf. Figure **tbm-appendix**).

We applied the identifiability analysis on our original NER model comprising 40 binding and dissociation parameters and 5 catalytic reaction constants. As it turns out, all binding and dissociation constants were identifiable with small bounds only under the assumption that the repair factor exchange at unwound and incised DNA is equal (cf. Figure 2.11). Besides the slow rate of rechromatinization δ,presumably identifiable due to the slow decrease in accumulated XPA (cf. Figure 2.9), all catalytic rates are fast. This is seen by the existence of lower bounds on the rate constants of the order of 1 s^{-1}. For the numerical values of the parameters see Table (**tbm Appendix**) and ... The K_d values which depict the protein binding affinities to the respective DNA repair intermediate fall into a physiological realistic range between ~100 nM and ~1 µM (cf. Table 2.3). Only XPC has a particular low affinity of 9 µM which is consistent with previous findings reporting that the time until DNA incision is mainly determined by the slow lesion recognition [5].

The narrow confidence bounds on the parameters allow us to make valid computational predictions. For example, we can simulate the so far not observable fraction of incised

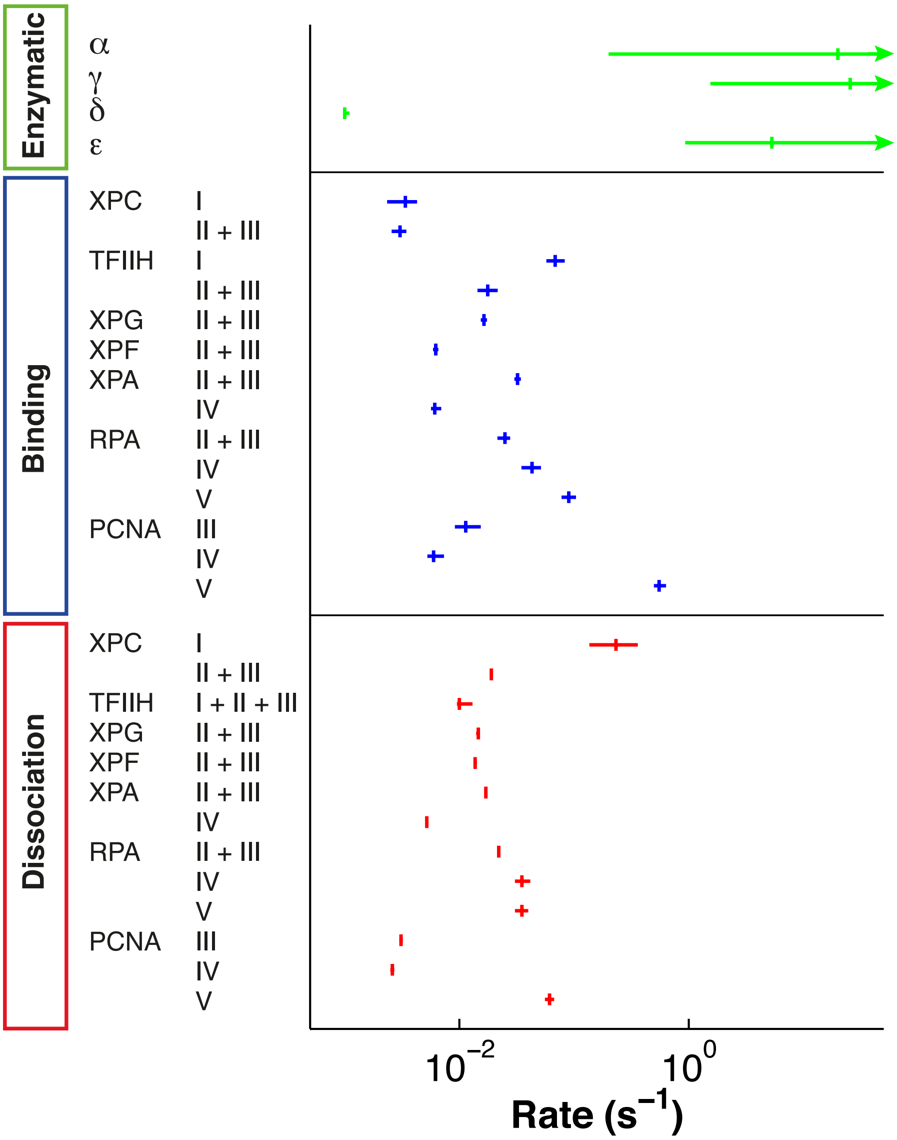

Figure 2.11: Parameterization identifies realistic model of NER. Mean values (vertical bars) and confidence intervals (horizontal bars) of catalytic (green), binding (blue) and dissociation constants (red) characterizing the dynamic assembly of NER factors at the successive repair intermediates (damaged DNA I, unwound DNA II, incised DNA III, resynthesized DNA IV, rechromatinized DNA V). Arrow heads indicate infinite confidence intervals.

Value	XPC	TFIIH	XPG	XPF	XPA	RPA	PCNA
Damaged DNA							
K_d(µM)	9.35 (3.46;16.09)	0.052 (0.044;0.072)	NA	NA	NA	NA	NA
Unwound DNA							
K_d(µM)	0.864 (0.635;1.006)	0.204 (0.163;0.259)	0.395 (0.373;0.419)	2.446 (2.344; 2.596)	0.147 (0.138;0.158)	1.222 (1.048;1.36)	NA
Incised DNA							
K_d(µM)	0.864 (0.635;1.006)	0.204 (0.163;0.259)	0.395 (0.373;0.419)	2.446 (2.344; 2.596)	0.147 (0.138;0.158)	1.222 (1.048;1.36)	0.388 (0.319;0.538)
Resynthesized DNA							
K_d(µM)	NA	NA	NA	NA	0.236 (0.222;0.27)	1.167 (0.924;1.521)	0.605 (0.531;0.747)
Rechromatinized DNA							
K_d(µM)	NA	NA	NA	NA	NA	0.538 (0.438;0.645)	0.154 (0.134;0.182)

Table 2.3: K_d Values. NA, not applicable. K_d ($k_{\text{off}}/k_{\text{on}}$) values are given for every repair protein and arranged in columns. Reference parameter set and 95% confidence intervals (in parentheses) are shown.

DNA (cf. Figure 2.12, green trajectory). As the EdU incorporation measurement shows, the damaged (blue) and repaired DNA (red) kinetics are tightly coupled and thereby omitting a higher accumulation of incised DNA.

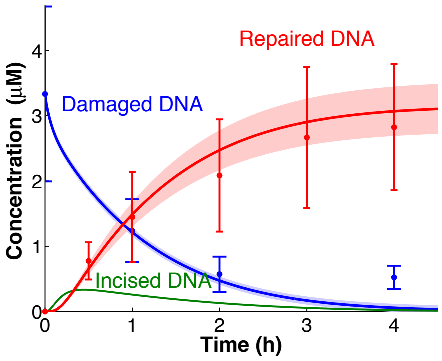

Figure 2.12: Short delay between DNA damage removal and DNA repair synthesis omits accumulation of incised DNA. Experimental (dots with error bars) and simulated (lines) time courses for damaged DNA (blue) and DNA repair synthesis (red). Simulated trajectories depict repair intermediate I+II for damaged DNA and repair intermediate IV+V for repaired DNA (cd. Figure 2.7). Estimated errors are depicted as shaded area. Model prediction for incised DNA (green) constitutes from DNA repair intermediate III.

3 Control analysis of the DNA repair rate

In the previous chapter we introduced the incorporation of EdU upon UV-irradiation as a quantitative readout for the DNA repair synthesis kinetic (cf. Figure 2.4). We could show that despite the pathway's complexity DNA repair follows a mono-exponential kinetic of first order. Moreover, the repair of single lesions is distributed over a broad time span with 50% of the lesions being repaired after a half-time of $\sim$1.2 h. Notably, measurements on the single cell level revealed that also the half-time and hence the rate of repair is highly variable (cf. Figure 2.5B-E). So far the origin of this variability is unknown but also if and how the cell faces this highly fluctuating environment.

In the following chapter we will apply the kinetic NER model and explore the nature of the repair rate variation. Simulating the effect of NER factor variation on the repair rate we find that the repair rate control is distributed among all NER factors. Exploiting the natural variability of expressed repair proteins we can experimentally corroborate the computationally derived finding for the NER factors XPC, TFIIH, XPA, XPF and RPA. Apart from the variability in protein expression the model identifies the initial amount of inflicted DNA damage as major contribution determining the repair rate distribution. Both sources together appear to be sufficient to explain the overall rate variability.

- clarify who did what...

3.1 Kinetic NER model predicts collective rate control

3.1.1 Response coefficients

The narrow confidence bounds derived by PLE analysis identified parameters in reasonable biological ranges (cf. Section 2.3.3) and allowed us to use the model for quantitative predictions. On the basis of the simplified model result, indicating that the fast and random enzyme exchange defines the slow first-oder rate kinetics (cf. Section 2.2.1), we wanted to test whether the concept of multi-protein rate control also applies on the realistic NER model. To determine the response of a system to changes in the environment one can calculate the response coefficients $\tilde{R}_i$ [38, 39]. Accordingly, we can quantify the relative change in the repair rate ν as a function of the relative change in the protein concentration C_i (i = XPC, TFIIH,...).

$$\tilde{R}_i = \frac{\partial \ln \nu}{\partial \ln C_i}. \tag{3.1}$$

The inverse of ν is referred to as characteristic time τ so that Eqn. 3.1 can be rewritten as

$$\tilde{R}_i = \frac{C_i}{\tau^{-1}} \frac{\partial \tau^{-1}}{\partial C_i}. \tag{3.2}$$

τ, in turn, can be directly approximated from the distribution of repaired DNA states $y_\pi^R(t)$, which include the modeled DNA intermediates for resynthesized and rechromatinized DNA. By taking the ratio between the first ($\mu^{(1)}$) and the zeroth ($\mu^{(0)}$) central moment of the distribution we derive the reaction-specific mean time

$$\tau_R = \frac{\mu(1)}{\mu^{(0)}}, \tag{3.3}$$

with

$$\mu^{(m)} = \int t^m \, y_\pi^R(t) \, dt. \tag{3.4}$$

To ensure the convergence of the integral we subtract all repair synthesis states (IV + V) from the initial amount of damages $y_{00}^{\mathrm{I}} = 3.33$ µM at t = 0.

$$\tau_{\text{syn}} = \frac{\int (y_{00}^I(0) - (\sum_\pi y_\pi^{IV}(t) + \sum_\pi y_\pi^V(t)) \, dt) \cdot t \, dt}{\int y_{00}^I(0) - (\sum_\pi y_\pi^{IV}(t) + \sum_\pi y_\pi^V(t)) \, dt} \tag{3.5}$$

Using τ_syn in Equation 3.2 we derive the response coefficients for the repair synthesis rate, which are uniformly small $\sim$0.3 and below (cf. Figure 3.1). This result implies that that there is no single repair protein whose effect on the repair speed could be interpreted as rate-limiting. A similar result is obtained for the rate of re-synthesis response coefficients (cf. Appendix), with the characteristic time

$$\tau_{\text{inc}} = \frac{\int_0^\infty t \sum_{x=I}^{\mathrm{III}} \sum_\pi y_\pi^x(t) \, dt}{\int_0^\infty \sum_{x=I}^{\mathrm{III}} \sum_\pi y_\pi^x(t) \, dt}. \tag{3.6}$$

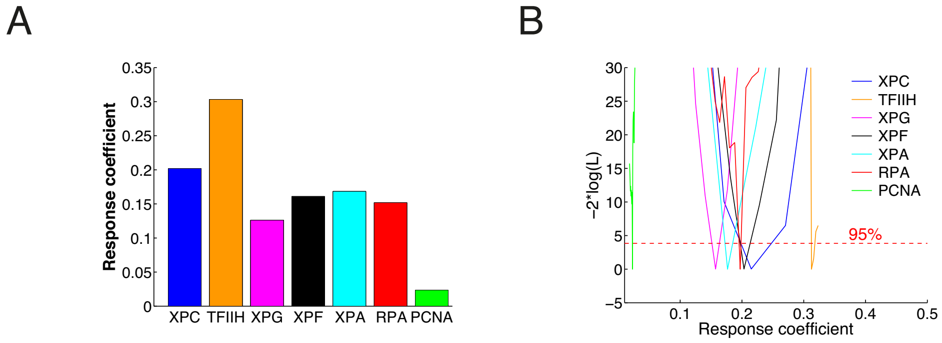

Figure 3.1: Collective control of the repair rate. A) Response coefficients for seven repair factors (XPC, TFIIH, XPG, XPA, XPF, RPA, PCNA) are small and uniformly distributed. B) Prediction profile likelihood indicate small prediction confidence bounds for each response coefficient. 95% threshold is given by the X^2 distribution with one degree of freedom.

Prediction profile likelihood To obtain realistic confidence bounds for the response coefficients we performed a prediction profile likelihood estimation [40, 41]. Thereby the predicted response coefficients $\tilde{R} = F(D_{\text{pred}}, \theta)$ are considered as model outcome F for a predicted experimental design $D_{\text{pred}} = (z_{\text{pred}}, t_{\text{pred}}, u_{\text{pred}})$. D_{pred} specifies a prediction observable z_{pred} at time point t_{pred} given the externally controlled stimulation u_{pred}. As defined in Eqn. 2.8 $z_{\text{pred}}(y(t), \theta)$ comprises a model simulation that can be mapped to experimentally observable quantities. Analogous to Eqn. 2.12 the prediction profile likelihood

$$PPL(\tilde{R}) = \max_{\theta^* \in \{\theta | F(D_{\text{pred}}, \theta) = \tilde{R}\}} L(z^\dagger | \theta^*, \tilde{r}) \tag{3.7}$$

is obtained by maximization over the model parameters satisfying the constraint that the model response $F(D, \theta^*)$ after fitting is equal to the considered value $\tilde{r}$ for the prediction $\tilde{R}$ with respect to the measured data $z^\dagger$. This procedure is repeated for continuous variations of $\tilde{R}$. The model response can then be expressed as

$$\Delta X^2_{\theta_l}(\tilde{r}) = \min_{\{\theta, \tilde{r} \in \tilde{R}\}} \left(X^2(\theta, \tilde{r}) \right) - \min_{\{\theta\}} \left(X^2(\theta) \right), \tag{3.8}$$

which describes the difference between the global X^2 minimum and the best fit with r included into the objective function. Similar to Eqn. 2.14 we can determine prediction confidence bounds

$$PCI_{1-a}(D_{\text{pred}}, z^{\dagger}) = \{\tilde{r} | \Delta X^2_{\theta_l}(\tilde{r}) \leq Q_{X^2}(1-a,1)\}, \tag{3.9}$$

which include the set of predictions $\tilde{R} = F(D_{\text{pred}}, \theta)$ for which -2log(PPL) is below the threshold given by the X^2-distribution. The PPLs were computed within the d2d-framework [29] using the CVODE package [31] for the numerical integration of the ODEs and the sensitivity equations (cf. Section 2.3.1). By applying the PPL analysis for each response coefficient we find that all of them have small confidence bounds indicating how well the model predictions are determined by the data (cf. Figure 3.1B).

The moderate response predicted by the response coefficients also holds true for larger variations in the repair factor concentration (cf. Figure 3.2). The linear approximation (on which the response coefficients are based) yields a reasonable description for about two-fold concentration decreases or increases (corresponding to a knockdown or overexpression experiment), while for very large decreases the repair rate drops eventually to zero (corresponding essentially to a gene knockout). From this result we can conclude that kinetic NER model predicts a collectively NER factor control of the repair rate. Consequently, the repair pathway appears robust against natural fluctuations in repair protein expression.

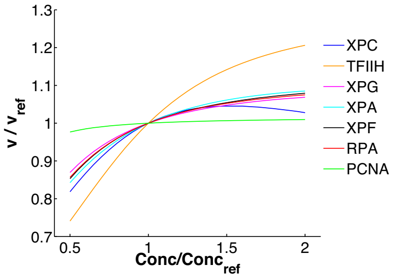

Figure 3.2: Moderate repair rate response against natural NER factor expression variability. Repair rate as a function of concentration changes in individual repair factors.

3.1.2 Exploiting natural variability in protein expression to quantify rate control

To corroborate the model prediction of a distributed repair rate control, we developed an experimental setup for the investigation of single repair factors and their quantitative influence on the repair rate. In particular, we asked whether there is a measurable response in the repair rate due to the natural occurring variability in protein expression. By capturing the integrated nuclear fluorescence intensity of antibody-stained repair proteins we derived the expression values for XPC, TFIIH, XPA, XPF and RPA. (cf. Figure 2.2B). For these five individually measured NER factors the expression variability was quantified by calculating the coefficient of variation (CV; standard deviation divided by the mean), which was on average ∼0.37 (cf. Figure 3.3A-E). Table 3.1 depicts mean and standard deviation representing at least three biologically independent measurements of the expressed protein cell-two-cell variability.

	XPC	**TFIIH**	**XPA**	**XPF**	**RPA**
CV:	0.34 ± 0.05	0.33 ± 0.02	0.33 ± 0.03	0.4 ± 0.04	0.44

Table 3.1: Mean and standard deviation of the variability in nuclear XPC, TFIIH, XPA and XPF expression. Distributions for nuclear protein expression were acquired in n_b(XPC) = 5, n_b(TFIIH, XPA,XPF) = 3 and n_b(RPA) = 2 biological independent replicates. Within each measurement between n=250 and n=572 with an average of n=477 cells were analyzed.

To distinguish whether the measured variability is due to differences in nuclear expression or rather superimposed by measurement noise we co-analyzed XPC-eGFP stably expressed in XP-C fibroblasts together with immunofluorescently labeled XPC in the same cell. Both quantities are strong positively correlated (cf. Figure 3.3F) suggesting a large natural variability compared to much lower measurement noise. To quantitatively validate this observation we performed a principal component analysis [42]. Thereby both quantities, XPC-eGFP and antibody-stained XPC intensities, are orthogonally transformed into a new coordinate system were the new transformed variables are linearly uncorrelated. These new variables are referred to as principal components. In a two-dimensional case the variances of both principal components define an error-ellipse as illustrated in Figure 3.3F. From this we estimated a relative measurement error of antibody labeling of ∼11% showing that the technique is suitable for quantification of nuclear XPC, TFIIH, XPA, XPF and RPA concentrations.

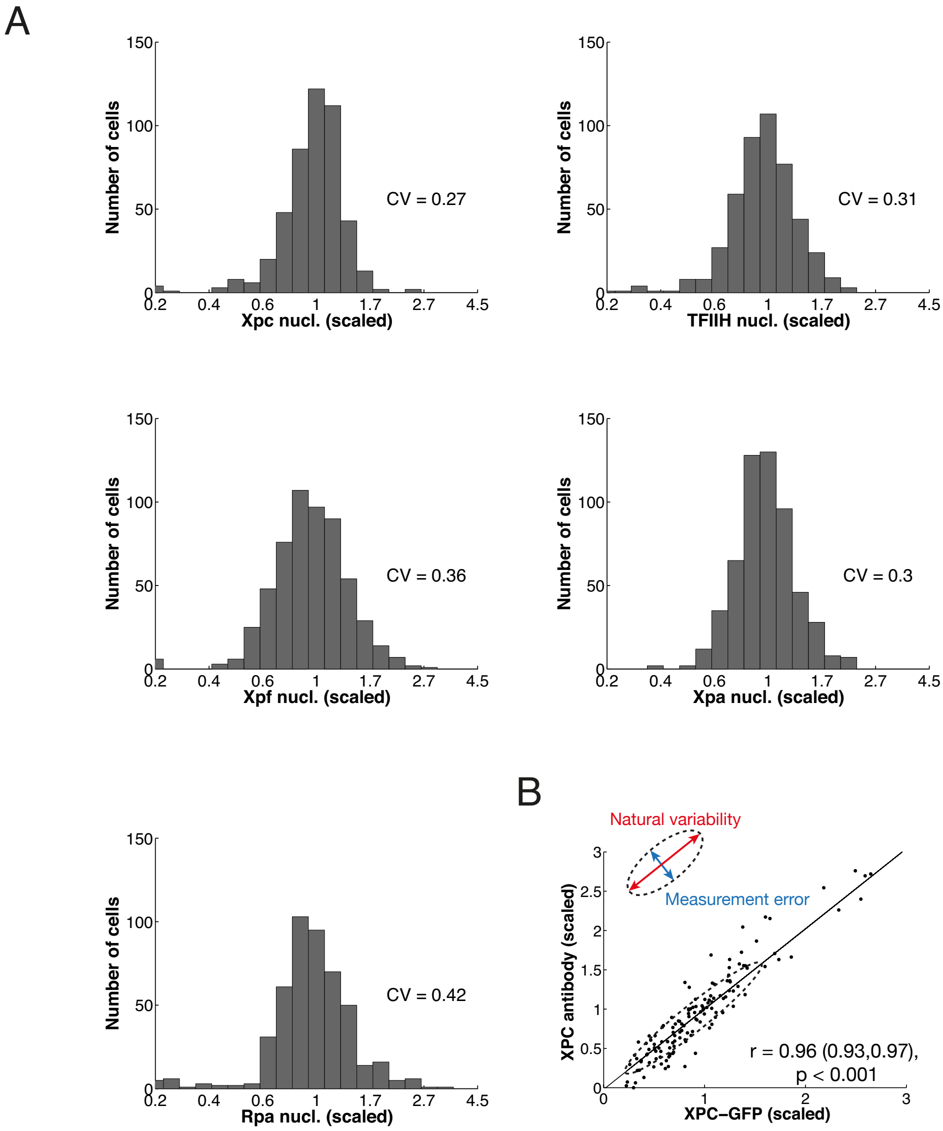

Figure 3.3: Natural variability in nuclear NER factor expression is significantly larger than measurement error A-E) Histograms of nuclear protein concentrations of the antibody stained NER factors XPC, TFIIH, XPF, XPA and RPA (**Bildunterschrift: Nuclear und cpt. letters**/n nachschauen). Scatter plot of antibody recognized XPC against XPC-eGFP stably expressed in XP-C XPC-eGFP cell lines. The dashed error-ellipse illustrates the proportion of natural variability and measurement noise.

Before exploring the direct relation between protein expression variability and the speed of repair we tested whether higher protein amounts in the nucleus correlate with the accumulation of NER factors in the locally damaged area. In correspondence to previous findings for XPG [5] there is a significant positive correlation between the nuclear XPC-eGFP expression and its local accumulation at the DNA lesions (cf. Figure 3.4A) thirty minutes after UV irradiation. The same holds true for XPA [4], TFIIH, XPF and RPA in narrow confidence bounds (cf. Appendix **tbm**). Consequently, we conclude that the DNA lesions are not saturated so that a higher NER factor concentration could potentially accelerate the repair rate.

In fact, the nuclear protein concentrations for all five antibody-stained repair factors and the amount of incorporated EdU after one hour were significantly correlated (cf. Figure 3.4B-F). Remarkably, the dependency between both quantities, characterized by the slop of the regression line is in the same order of magnitude as predicted computationally by the response coefficients (cf. Figure 3.1A). Also among each other, correlation and regression slope fell into the same range. This agrees with the *in silico* finding that the kinetic control of the measured repair factors is distributed among them and that the rate of repair synthesis is robust against natural variations in the repair protein expression.

Correlation analysis The correlation analysis was performed in MATLAB (2012a, The Mathworks Inc., Natick, MA) using the inbuilt corrcoef function, which also returns p-values indicating the significance of the correlation coefficient. The p-values are determined by a t-test, where the null-hypothesis states that the true correlation between two variables equals zero and that for its observed value r the quantity

$$t = \frac{r}{\sqrt{\frac{(1-r^2)}{N-2}}} \tag{3.10}$$

follows approximately a t-distribution for $N-2$ degrees of freedom and a sample size N [43, 44]. A low p-value suggests the rejection of the null-hypothesis and thus, demonstrates the significance of the correlation.

- confidence bounds were derived by bootstraping ...

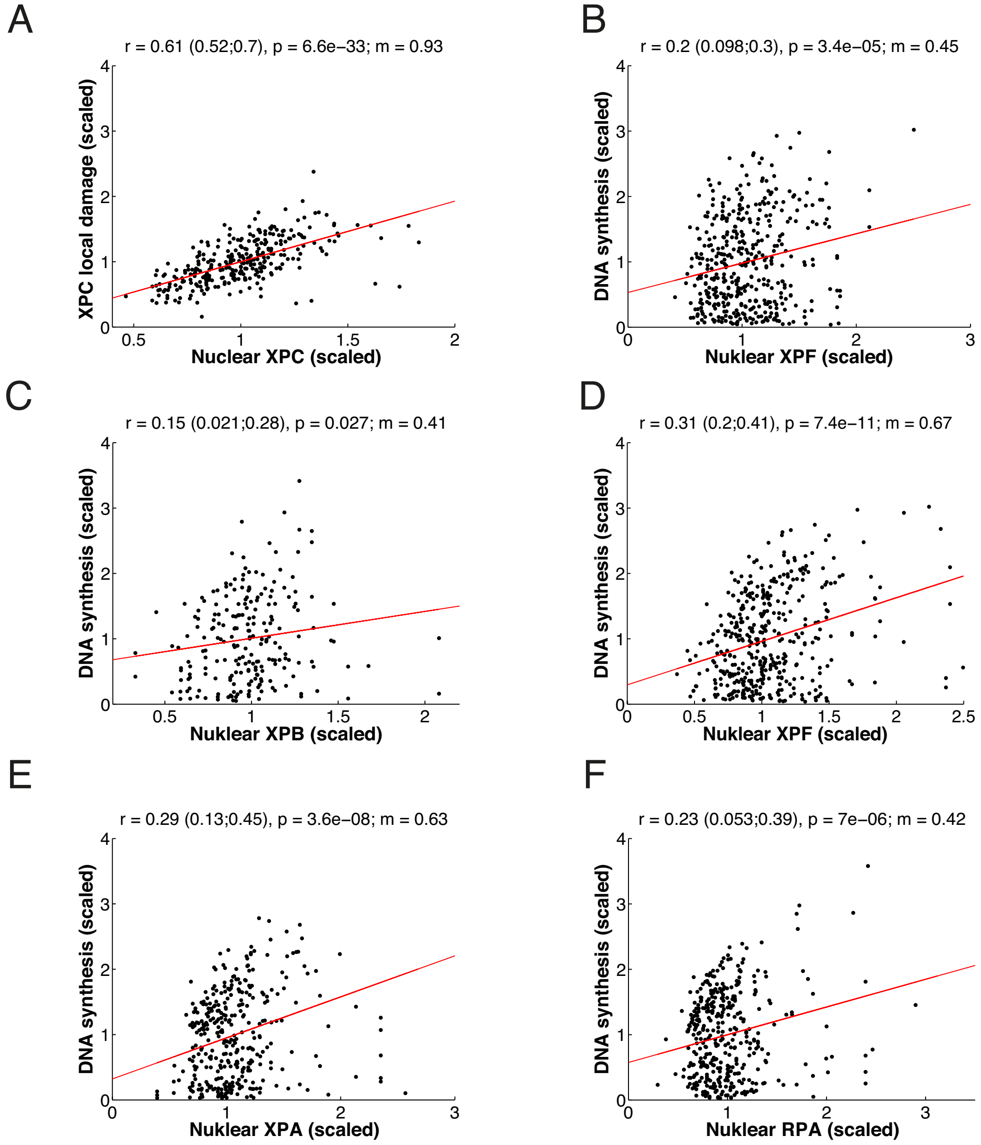

Figure 3.4: Blubb. A) B)

3.1.3 NER Variability comparable between complemented patient cell lines and native repair

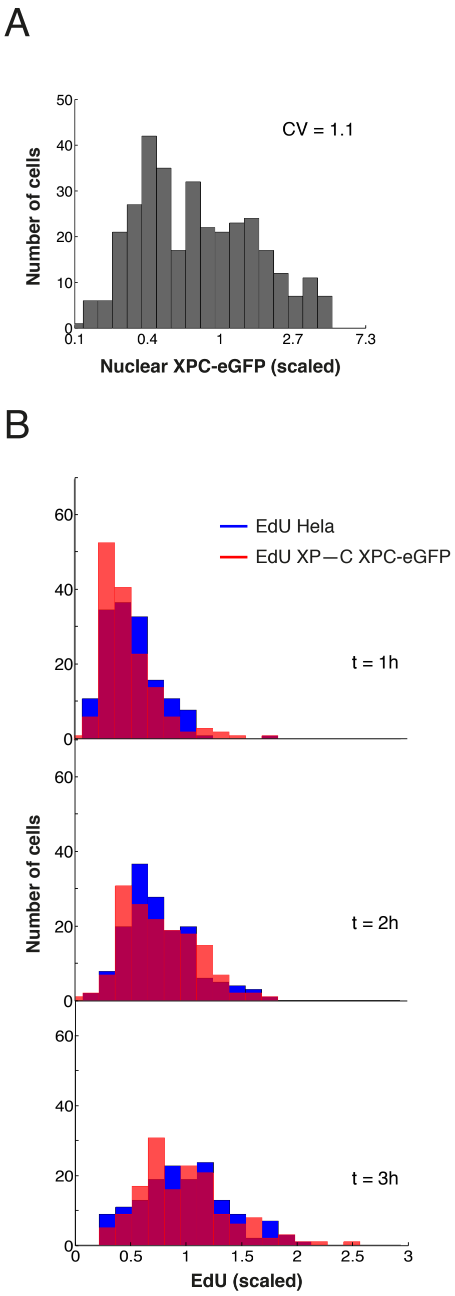

Figure 3.5: Blubb. A) B)

3.2 Sources for repair rate variability

3.2.1 Variable NER factor expression and inflicted lesions account for the distribution of repair rates

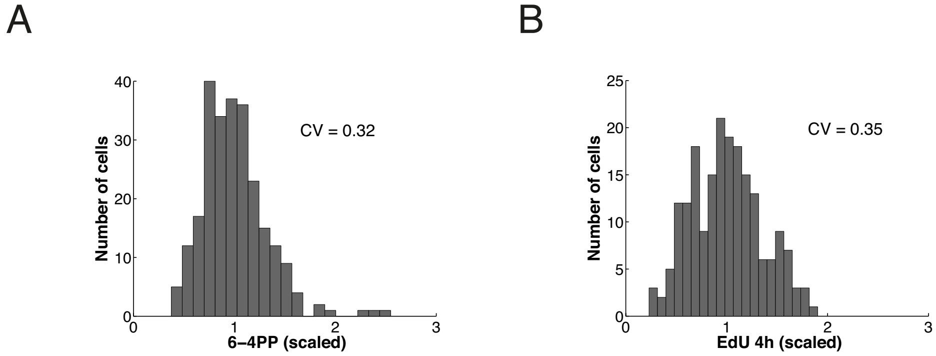

Figure 3.6: Blubb. A) B)

3.2.2 Imbalanced contribution

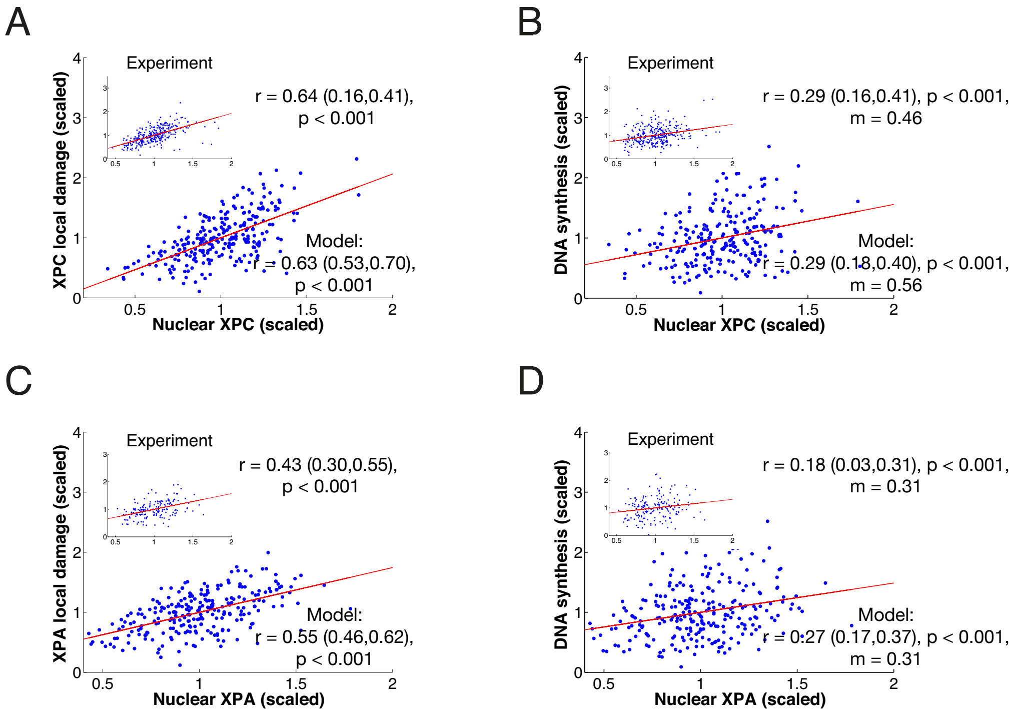

Figure 3.7: Blubb. A) B)

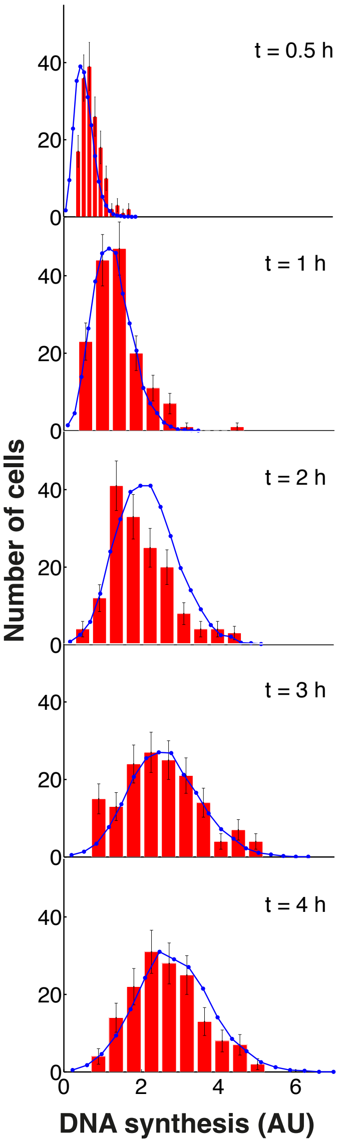

Figure 3.8: Blubb. A) B)

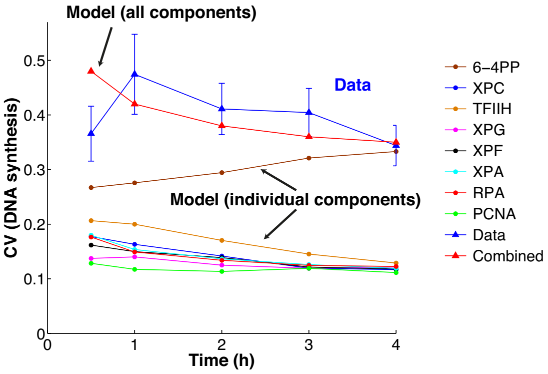

Figure 3.9: Blubb. A) B)

4 Co-expression of NER repair factors

The previously described model-aided analysis of the DNA repair process revealed the link between the emergent phenomenon of rapidly exchanging and transiently interacting NER components with the experimentally observed slow first-order kinetics of repair. An important functional consequence of this kinetic design is that the control of the repair rate is shared by all repair factors. This manifests in the mathematical prediction of uniformly distributed response coefficients, which quantify the relative change of the repair rate in answer to changes in the nuclear repair protein concentration. Exploiting the natural variability in NER factor expression, their moderate control on the rate of repair could be experimentally resolved for different repair phases. These findings, however, were made under the assumption of a functional independence of the individual NER components. Effects resulting from co-regulated NER factor expression or protein degradation were so far neglected as a first approximation.

To test this assumption, in the following chapter we will investigate a potential cross-correlation among the involved repair factors experimentally. Surprisingly, we find that the nuclear expression of five pairwise measured repair factors (XPC, TFIIH, XPA, XPF and RPA) is indeed strong positively correlated, whereas there is no correlation with the repair-independent cell cycle marker **Ki67**. This result suggests an additional control mechanism orchestrating NER factor expression on the transcriptional or translational level.

- sentence about who did what

4.1 Nuclear expression of NER factors is strongly correlated

4.1.1 Microscopy analysis of co-staining experiments

Using fluorescence microscopy for the quantitative analysis of the NER process proved to result in accurate measurements of the nuclear expression and UV-induced repair dynamics of the individual repair factors at a high signal to noise ratio. This applies for the stably transfected fluorescently tagged repair protein expression as well as for an immunocytochemistry approach with an indirect antibody-labeling of the measured NER factors. The latter technique has the advantage that once a specifically binding antibody is established it can be applied on all cell types containing the desired antigen. On this basis, indirect labeling of the primary antibody with a secondary antibody allows for the combinatorial tagging of multiple antigens restricted only by cross-reactions between antibodies originating from the same host species [45, 46]. Following such a protocol we established five single cell double stainings (cf. Table 4.1 using the primary and secondary antibodies listed in Table **tbm**.

	XPC	**TFIIH**	**XPA**	**XPF**	**RPA**
XPC	X	X	X	X	X
TFIIH		-	-	X	-
XPA			-	X	X
XPF				-	-
RPA					-

Table 4.1: Mean and standard deviation of the variability in nuclear XPC, TFIIH, XPA and XPF expression.

Additional two cross-correlation measurements were possible because of a directly labeled monoclonal mouse XPC-antibody. After an initial indirect labeling step for one of the other four repair factors no further secondary antibody is needed. Hence, there is no cross-recognition for primary antibodies from the same species. Moreover, the direct labeling permits a double staining for XPC with two individual fluorophors, which offers a precise estimate of the measurement error (cf. Section 3.1.2).

The cross-correlation analysis was performed in human diploid female fibroblasts which were grown to confluency on coverslips. Analogous to the description in Section 2.1.1 cells were UV-irradiated locally with a dose of 100 J/m^2. Pre-incubated with serum-free medium containing 10 µM EdU cells were allowed to repair for 60 minutes in an incubator. After the subsequent direct or indirect antibody labeling of two selected repair factors cells were additionally incubated with a DAPI solution, which visualizes the cell's chromatin and thus the contour of the nucleus. For each double staining microscopic 3-dimensional imaging was conducted on a Leica TCS SP5 II confocal microscope. All images were analyzed following the protocol presented in Section 2.1.2. Segmentation of the nucleus was performed on the DAPI signal, due to the large signal-to-nose ratio in this channel (cf. Appendix **Appendix** and Figure 4.1A). The segmented region was then used to quantify the signal emitted by the secondary antibodies at 488 and 647 nm (cf. Figure 4.1B and C). For the same reason segmentation of the locally UV-irradiated chromatin was done with the EdU signal and then projected onto the signal of accumulated protein (cf. Figure 4.1D).

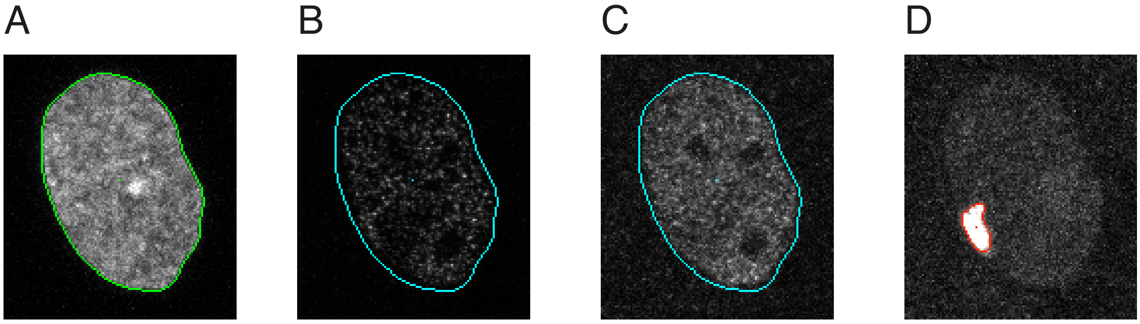

Figure 4.1: Blubb. A) B)

To test whether the antibody co-staining experiment is suitable for the cross-correlation analysis we measured XPC with a directly labeled antibody as well as an indirect labeling. We correlated both signals originating from the entire nucleus including the UV-irradiated region (cf. Figure 4.2). Signal quantification and PCA were performed as described in Section 2.1.2 and 3.1.2, respectively. Accordingly, the measurement error was determined with ..., which lies in the same range as the combined approach, with one protein stably expressed and the other indirectly labeled. Hence, also this approach is sufficiently accurate to exploit the natural variability in protein expression for the cross-correlation analysis.

As it turns out, all pairwise correlations of the measured co-staining experiments (cf. Table 4.1) are strong positively correlated with correlation coefficients between 0.74 and 0.88 (cf. Figure 4.2B-D and Appendix ...). Notably, the result is indifferent whether the acquired fluorescence signal is taken from the whole cell nucleus including the locally damaged area or only from the undamaged chromatin region (cf. Figure 4.2B-D and 4.2E-F). This suggests that the correlation of nuclear NER factors is independent of the ongoing repair. As a consequence, we suspect the regulatory mechanism determining the protein concentrations on a preceding level like transcription or translation.

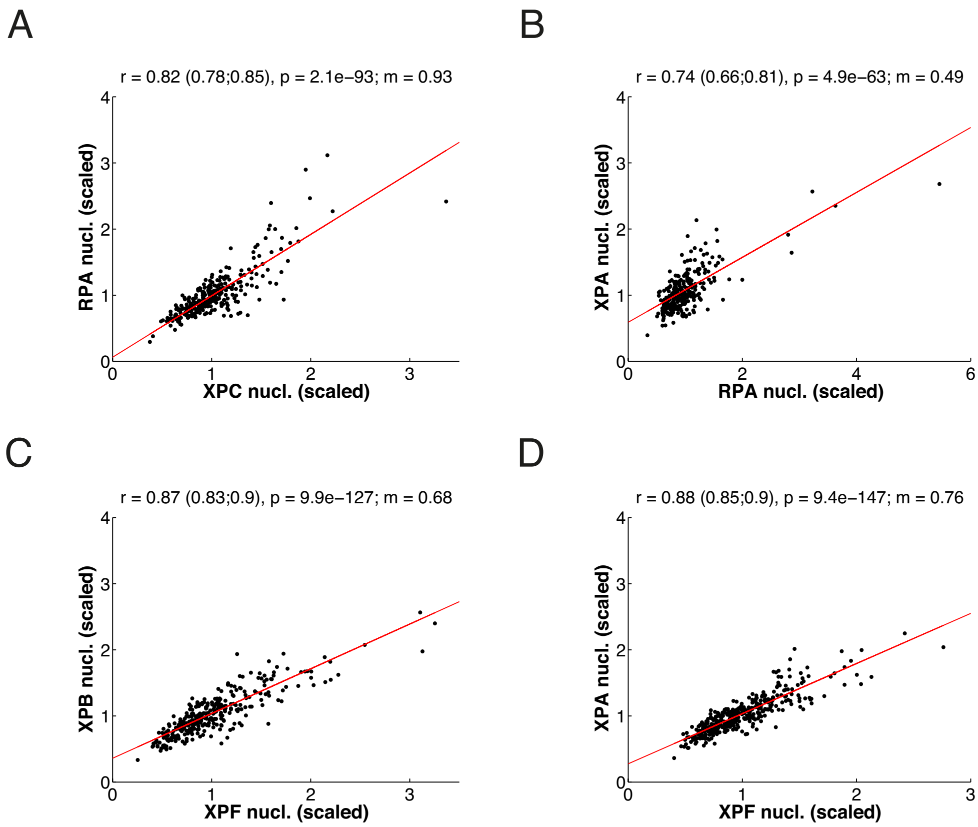

Figure 4.2: Blubb. A) B)

4.1.2 Flow cytometry

4.1.3 Calibration of the response coefficients

The positive pairwise correlation of the nuclear repair factor concentration was conclusively shown by two independent experimental approaches. This implies that the abundance of a protein is dependent on the other repair components, which has to be considered when determining the control on the rate of repair. Including this functional interdependence the relation between the measured EdU response (cf. Figure 3.4) and the response coefficients extends to

$$\frac{C_i}{\nu}\frac{d\nu}{dC_i} = \frac{C_i}{\nu}\frac{\partial\nu}{\partial C_i} + \sum_{j\neq i}\left(\frac{C_j}{\nu}\frac{\partial\nu}{\partial C_i}\right)\left(\frac{C_i}{C_j}\frac{\partial C_j}{\partial C_i}\right). \tag{4.1}$$

Apart from the unknown response coefficients

$$R_i = \frac{C_i}{\nu}\frac{\partial\nu}{\partial C_i}$$

we directly measured the remaining derivatives

$$A_i = \frac{C_i}{\nu}\frac{d\nu}{dC_i}$$

as EdU incorporation for varying NER factor concentrations after one hour and

$$B_{ij} = \frac{C_i}{C_j}\frac{\partial C_j}{\partial C_i}$$

from the regression of repair factor concentration i against repair factor j. Hence, for the calculation of the response coefficients we have to solve a linear equation system:

$$R_i + \sum_{j\neq i} B_{ij}R_j = A_i, \tag{4.2}$$

for $i \in$ 1,...,N. We approximate B_{ii} = 1 and derive the matrix notation $B\vec{R} = \vec{A}$. However, even if we measured all pairwise NER factor dependencies, due to the strong correlations the matrix B would not be invertible. - Invertierbarkeit...

Although we are not able to derive a specific result for each response coefficient it is possible to make an estimation of their average size. Consider ν as a function of the concentration of the repair factors C_i and the initial amount of DNA lesions L:

$$\nu = Lf(C_1, ...C_N). \tag{4.3}$$

Using, once more, the standard law for the propagation of uncertainty analog to Eqn. ...

$$\sigma(\nu) = \sqrt{\sum_i \left(\frac{\partial \nu}{\partial C_i}\sigma(C_i)\right)^2 + \left(\frac{\partial \nu}{\partial L}\sigma(L)\right)^2 + \left(\frac{\partial \nu}{\partial C_i}\frac{\partial \nu}{\partial C_i}\mathrm{cov}_{ij}\right)^2} \tag{4.4}$$

and introducing the response coefficients

$$R_i = \frac{C_i}{\nu}\frac{\partial \nu}{\partial C_i}$$

5 Discussion

- discuss difficulties during antibody staining

6 Bibliography

[1] Moné, M. J., Volker, M., Nikaido, O., Mullenders, L. H., van Zeeland, a. a., Verschure, P. J., Manders, E. M. & van Driel, R. Local UV-induced DNA damage in cell nuclei results in local transcription inhibition. *EMBO reports* **2**, 1013–7 (2001).

[2] Houtsmuller, A. & Vermeulen, W. Macromolecular dynamics in living cell nuclei revealed by fluorescence redistribution after photobleaching. *Histochemistry and cell biology* **115**, 13–21 (2001).

[3] Moné, M. J., Bernas, T., Dinant, C., Goedvree, F. a., Manders, E. M. M., Volker, M., Houtsmuller, A. B., Hoeijmakers, J. H. J., Vermeulen, W. & van Driel, R. In vivo dynamics of chromatin-associated complex formation in mammalian nucleotide excision repair. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15933–7 (2004).

[4] Verbruggen, P., Heinemann, T., Manders, E., von Bornstaedt, G., van Driel, R. & Höfer, T. Robustness of DNA Repair through Collective Rate Control. *PLoS Computational Biology* **10**, e1003438 (2014).

[5] Luijsterburg, M. S., von Bornstaedt, G., Gourdin, A. M., Politi, A. Z., Moné, M. J., Warmerdam, D. O., Goedhart, J., Vermeulen, W., van Driel, R. & Höfer, T. Stochastic and reversible assembly of a multiprotein DNA repair complex ensures accurate target site recognition and efficient repair. *The Journal of Cell Biology* **189**, 445–63 (2010).

[6] Araújo, S. J., Nigg, E. A. & Wood, R. D. Strong Functional Interactions of TFIIH with XPC and XPG in Human DNA Nucleotide Excision Repair , without a Preassembled Repairosome. *Molecular and cellular biology* **21**, 2281–2291 (2001).

[7] Houtsmuller, a. B. Action of DNA Repair Endonuclease ERCC1/XPF in Living Cells. *Science* **284**, 958–961 (1999).

[8] Limsirichaikul, S., Niimi, A., Fawcett, H., Lehmann, A., Yamashita, S. & Ogi, T. A rapid non-radioactive technique for measurement of repair synthesis in primary human fibroblasts by incorporation of ethynyl deoxyuridine (EdU). *Nucleic acids research* **37**, e31 (2009).

[9] Essers, J., Theil, A. F., Baldeyron, C., Wiggert, A., Cappellen, V., Houtsmuller, A. B., Kanaar, R. & Cappellen, W. A. V. Nuclear Dynamics of PCNA in DNA Replication and Repair. *Molecular and cellular biology* **25**, 9350–9359 (2005).

[10] Sporbert, A., Gahl, A., Ankerhold, R., Leonhardt, H. & Cardoso, M. C. DNA Polymerase Clamp Shows Little Turnover at Established Replication Sites but Sequential De Novo Assembly at Adjacent Origin Clusters. *Molecular cell* **10**, 1355–1365 (2002).

[11] Terstiege, G. Mammalian DNA Repair: a Model for Multi-protein Complex Assembly on Chromatin. Ph.D. thesis, University of Heidelberg (2010).

[12] Evans, E., Moggs, J. G., Hwang, J. R., Egly, J.-m. & Wood, R. D. Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO Journal* **16**, 6559–6573 (1997).

[13] Mu, D., Hsu, D. S. & Sancar, A. Reaction Mechanism of Human DNA Repair Excision Nuclease. *Journal of Biological Chemistry* **271**, 8285–8294 (1996).

[14] Polo, S. E., Roche, D. & Almouzni, G. New histone incorporation marks sites of UV repair in human cells. *Cell* **127**, 481–93 (2006).

[15] Tapias, A., Auriol, J., Forget, D., Enzlin, J. H., Schärer, O. D., Coin, F., Coulombe, B. & Egly, J.-M. Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *The Journal of biological chemistry* **279**, 19074–83 (2004).

[16] Rademakers, S., Volker, M., Hoogstraten, D., Nigg, A. L., Mone, M. J., Zeeland, A. A. V., Hoeijmakers, J. H. J., Houtsmuller, A. B. & Vermeulen, W. Xeroderma Pigmentosum Group A Protein Loads as a Separate Factor onto DNA Lesions. *Molecular and cellular biology* **23**, 5755–5767 (2003).

[17] Zotter, A., Luijsterburg, M. S., Warmerdam, D. O., Ibrahim, S., Nigg, A., van Cappellen, W. a., Hoeijmakers, J. H. J., van Driel, R., Vermeulen, W. & Houtsmuller, A. B. Recruitment of the nucleotide excision repair endonuclease XPG to sites of UV-induced dna damage depends on functional TFIIH. *Molecular and cellular biology* **26**, 8868–79 (2006).

[18] Politi, A., Moné, M. J., Houtsmuller, A. B., Hoogstraten, D., Vermeulen, W., Heinrich, R. & van Driel, R. Mathematical modeling of nucleotide excision repair reveals efficiency of sequential assembly strategies. *Molecular cell* **19**, 679–90 (2005).

[19] Riedl, T., Hanaoka, F. & Egly, J.-M. The comings and goings of nucleotide excision repair factors on damaged DNA. *The EMBO journal* **22**, 5293–303 (2003).

[20] Yokoi, M., Masutani, C., Maekawa, T., Sugasawa, K., Ohkuma, Y. & Hanaoka, F. The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA. *J Biol Chem* **275**, 9870–9875 (2000).

[21] Volker, M., Moné, M. J., Karmakar, P., van Hoffen, a., Schul, W., Vermeulen, W., Hoeijmakers, J. H., van Driel, R., van Zeeland, a. a. & Mullenders, L. H. Sequential assembly of the nucleotide excision repair factors in vivo. *Molecular cell* **8**, 213–24 (2001).

[22] O'Donovan, A., Davies, A. A., Moggs, J. G., West, S. C. & Wood, R. D. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371**, 432–435 (1994).

[23] Sijbers, A. M., de Laat, W. L., Ariza, R. R., Biggerstaff, M., Wei, Y. F., Moggs, J. G., Carter, K. C., Shell, B. K., Evans, E., de Jong, M. C., Rademakers, S., de Rooij, J., Jaspers, N. G., Hoeijmakers, J. H. & Wood, R. D. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86**, 811–822 (1996).

[24] Winkler, G. S., Sugasawa, K., Eker, a. P., de Laat, W. L. & Hoeijmakers, J. H. Novel functional interactions between nucleotide excision DNA repair proteins influencing the enzymatic activities of TFIIH, XPG, and ERCC1-XPF. *Biochemistry* **40**, 160–5 (2001).

[25] de Laat, W. L., Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N. G. & Hoeijmakers, J. H. DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev* **12**, 2598–2609 (1998).

[26] Moser, J., Volker, M., Kool, H., Alekseev, S., Vrieling, H., Yasui, A., van Zeeland, A. a. & Mullenders, L. H. F. The UV-damaged DNA binding protein mediates efficient targeting of the nucleotide excision repair complex to UV-induced photo lesions. *DNA repair* **4**, 571–82 (2005).

[27] Shivji, M. K., Podust, V. N., Hübscher, U. & Wood, R. D. Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry* **34**, 5011–5017 (1995).

[28] Moser, J., Kool, H., Giakzidis, I., Caldecott, K., Mullenders, L. H. & Fousteri, M. I. Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Mol Cell* **27**, 311–323 (2007).

[29] Raue, A., Schilling, M., Bachmann, J., Matteson, A., Schelker, M., Schelke, M., Kaschek, D., Hug, S., Kreutz, C., Harms, B. D., Theis, F. J., Klingmüller, U. & Timmer, J. Lessons learned from quantitative dynamical modeling in systems biology. *PloS one* **8**, e74335 (2013).

[30] Leis, J. R. & Kramer, M. A. The Simultaneous Solution and Sensitivity Analysis of Systems Described by Ordinary Differential Equations. *ACM Transactions on Mathematical Software* **14**, 45–60 (1988).

[31] Hindmarsh, A. C., Brown, P. N., Grant, K. E., Lee, S. L., Serban, R., Shumaker, D. A. N. E. & Woodward, C. S. SUNDIALS : Suite of Nonlinear and Differential / Algebraic Equation Solvers. *ACM Transactions on Mathematical Software* **31**, 363–396 (2005).

[32] Owen, A. B. A Central Limit Theorem for Latin Hypercube Sampling. *Royal Statistical Society* **54**, 541–551 (2014).

[33] Cobelli, C. & DiStefano, J. J. I. Parameter and structural identifiability concepts and ambiguities: a critical review and analysis. *The American Physiological Society* **239**, 7–24 (1980).

[34] Swameye, I., Müller, T., Timmer, J., Sandra, O. & Klingmüller, U. Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 1028–1033 (2003).

[35] Venzon, D. J. & Moolgavkar, S. H. A method for computing profile-likelihood-based confidence intervals. *Applied Statistics* **37**, 87–94 (1988).

[36] Murphy, S. A. & van der Vaart, A. W. On Profile Likelihood. *Journal of the American Statistical Association* **95**, 449–485 (2000).

[37] Raue, a., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U. & Timmer, J. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics (Oxford, England)* **25**, 1923–9 (2009).

[38] Hofmeyr, J.-H. S. & Cornish-Bowden, A. Quantitative assessment of regulation in metabolic systems. *European Journal of Biochemistry* **200**, 223–236 (1991).

[39] Fell, D. A. Metabolic Control Analysis : a survey of its theoretical and experimental development. *Biochemical Journal* **286**, 313–330 (1992).

[40] Kreutz, C., Raue, A. & Timmer, J. Likelihood based observability analysis and confidence intervals for predictions of dynamic models. *BMC systems biology* **6**, 120 (2012).

[41] Hinkley, D. Predictive Likelihood. *The Annals of Statistics* **7**, 718–728 (1979).

[42] Pearson, K. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine* **2**, 559–572 (1901).

[43] Kendall, M. G. & Stuart, A. The Advanced Theory of Statistics. McMillan, London, 4th ed. (1979).

[44] Fisher, R. Statistical Methods for Research Workers. Hafner, 13th ed. (1958).

[45] Burry, R. W. Controls for immunocytochemistry: an update. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* **59**, 6–12 (2011).

[46] Giepmans, B. N. G., Adams, S. R., Ellisman, M. H. & Tsien, R. Y. The fluorescent toolbox for assessing protein location and function. *Science (New York, N.Y.)* **312**, 217–24 (2006).

7 Appendix

Table SI. **Values of binding and dissociation rate constants**

Value	**XPC**	**TFIIH**	**XPG**	**XPF**	**XPA**	**RPA**	**PCNA**
Concentration (µM)	0.140	0.360	0.440	1.110	0.170	1.110	1.110
Damaged DNA							
k_{on} (µ$M^{-1}s^{-1}$)	0.025 (0.013;0.036)	0.2 (0.152;0.288)	NA	NA	NA	NA	NA
k_{off} (s^{-1})	0.231 (0.136;0.36)	0.01 (0.009;0.012)	NA	NA	NA	NA	NA
K_d (µM)	9.35 (3.46;16.09)	0.052 (0.044;0.072)	NA	NA	NA	NA	NA
Unwound DNA							
k_{on} (µ$M^{-1}s^{-1}$)	0.022 (0.016;0.025)	0.049 (0.04;0.06)	0.037 (0.036;0.039)	0.0056 (0.0054;0.0059)	0.116 (0.109;0.125)	0.018 (0.016;0.02)	NA
k_{off} (s^{-1})	0.019 (0.019;0.019)	0.01 (0.009;0.012)	0.0146 (0.0140;0.0149)	0.0137 (0.0134; 0.0141)	0.017 (0.0167;0.0172)	0.022 (0.0213;0.0224)	NA
K_d (µM)	0.864 (0.635;1.006)	0.204 (0.163;0.259)	0.395 (0.373;0.419)	2.446 (2.344; 2.596)	0.147 (0.138;0.158)	1.222 (1.048;1.36)	NA
Incised DNA							
k_{on} (µ$M^{-1}s^{-1}$)	0.022 (0.016;0.025)	0.049 (0.04;0.06)	0.037 (0.036;0.039)	0.0056 (0.0054;0.0059)	0.116 (0.109;0.125)	0.018 (0.016;0.02)	0.008 (0.0066;0.0111)
k_{off} (s^{-1})	0.019 (0.019;0.019)	0.01 (0.009;0.012)	0.0146 (0.0140;0.0149)	0.0137 (0.0134; 0.0141)	0.017 (0.0167;0.0172)	0.022 (0.0213;0.0224)	0.0031 (0.003;0.0032)
K_d (µM)	0.864 (0.635;1.006)	0.204 (0.163;0.259)	0.395 (0.373;0.419)	2.446 (2.344; 2.596)	0.147 (0.138;0.158)	1.222 (1.048;1.36)	0.388 (0.319;0.538)
Resynthesized DNA							
k_{on} (µ$M^{-1}s^{-1}$)	NA	NA	NA	NA	0.022 (0.021;0.025)	0.032 (0.025;0.037)	0.004 (0.0038;0.005)
k_{off} (s^{-1})	NA	NA	NA	NA	0.0052 (0.0051;0.0053)	0.035 (0.031;0.042)	0.0026 (0.0025;0.0027)
K_d (µM)	NA	NA	NA	NA	0.236 (0.222;0.27)	1.167 (0.924;1.521)	0.605 (0.531;0.747)
Rechromatinized DNA							
k_{on} (µ$M^{-1}s^{-1}$)	NA	NA	NA	NA	NA	0.065 (0.056;0.075)	0.396 (0.359;0.457)
k_{off} (s^{-1})	NA	NA	NA	NA	NA	0.035 (0.031;0.04)	0.061 (0.056;0.067)
K_d (µM)	NA	NA	NA	NA	NA	0.538 (0.438;0.645)	0.154 (0.134;0.182)

Table SII. **Values of the enzymatic rate constants**

Enzymatic rate		k_{cat}
		s^{-1}
Unwinding α	19.9	(>0.2)
Resynthesis γ	25.5	(>1.5)
Rechromatinization δ	0.001	(0.001;0.0011)
Reannealing ϵ	5.3	(>0.9)