



**HAL**  
open science

# XPG and XPF Endonucleases Trigger Chromatin Looping and DNA Demethylation for Accurate Expression of Activated Genes

Nicolas Le may, Delphine Fradin, Izarn Iltis, Pierre Bougneres, Jean Marc Egly

► **To cite this version:**

Nicolas Le may, Delphine Fradin, Izarn Iltis, Pierre Bougneres, Jean Marc Egly. XPG and XPF Endonucleases Trigger Chromatin Looping and DNA Demethylation for Accurate Expression of Activated Genes. *Molecular Cell*, 2012, 47 (4), pp.622-632. 10.1016/j.molcel.2012.05.050 . hal-02329959

**HAL Id: hal-02329959**

**<https://hal.science/hal-02329959v1>**

Submitted on 29 Jan 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**XPG and XPF endonucleases trigger chromatin looping  
and DNA demethylation**

**Nicolas Le May<sup>1</sup>, Delphine Fradin<sup>2</sup>, Izarn Iltis<sup>1</sup>, Pierre Bougnères<sup>2</sup>  
and Jean-Marc Egly<sup>1</sup>**

<sup>1</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/Inserm/ULP, BP 163, 67404 Illkirch Cedex, C. U. Strasbourg, France.

<sup>2</sup>Department of Paediatric Endocrinology and U986 Inserm, Hôpital Bicêtre, Paris11 University, Paris France.

\*Correspondence: [egly@igbmc.fr](mailto:egly@igbmc.fr)

**Abstract**

Nucleotide-Excision Repair (NER) factors, initially characterized as part of DNA repair, were also shown to belong to the transcriptional process in absence of genotoxic attack. However, their molecular function at the promoter of activated genes still remains obscure. Here we show that XPG and XPF endonucleases are required for establishing the CTCF-dependent chromatin looping between the promoter and terminator of activated *RARβ2*. XPG promotes DNA break and demethylation around the promoter while XPF-ERCC1 acts at the terminator concomitantly to CTCF recruitment. Silencing or inhibiting XPG or XPF-ERCC1 abolishes gene looping, DNA break formation and DNA demethylation.

Our results contribute to the better understanding of transcriptional regulation by a timely orchestrated endonuclease activity of the NER factors XPG and XPF in the formation of the active chromatin hub that controls gene expression. This could also provide explanation for several clinical features of the patients with mutations in NER factors.

## **Introduction**

RNA synthesis is the result of a cascade of chronologically orchestrated events that requires several hundreds of proteins. Upon gene activation, a host of proteins including the RNA polymerase II (pol II), the general transcription factors (GTFs), coactivators, corepressors, and chromatin remodelers are assembled around the promoter and their enzymatic activities contribute to protein coding gene expression (Brivanlou and Darnell, 2002; Kornberg, 2007). Among these proteins necessary to initiate RNA synthesis, are the Nucleotide-Excision-Repair (NER) factors (Le May et al., 2010; Schmitz et al., 2009). These factors were shown to be required for optimal chromatin remodeling including histones post-translational modifications (PTMs) as well as DNA demethylation at the activated genes. These NER factors (XPC, CSB, TFIIH, XPA, XPG, XPF-ERCC1) were first characterized as part of the DNA repair machinery to eliminate lesions originated by exogenous or endogenous genotoxic attacks (For review, see (Nospikel, 2009)). Mutations in the genes coding for NER factors have been associated with the human genetic disorders Xeroderma Pigmentosum (XP), Trichothiodystrophy (TTD), Cockayne Syndrome (CS), XFE progeroid syndrome and cerebro-oculo-facio-skeletal syndrome (COFS) (Jaspers et al., 2007; Kraemer et al., 2007; Niedernhofer et al., 2006).

In the present study, we investigated the implication of XPG and XPF-ERCC1 in the transactivation of nuclear receptor (NR)-target genes. We found that these two factors recruited at the promoter and terminator of activated genes, are necessary to promote chromatin looping and DNA demethylation, the two very crucial steps in transcription.

## **Results**

### **XPG and XPF are involved in DNA looping upon RAR $\beta$ 2 activation**

We first attempted to investigate the presence of the transcription and NER factors along the RAR $\beta$ 2 activated gene. Analysis of the promoter region (Pro) including two RAR-responsive elements (RARE) and the TATA box as well as a region defined as the terminator (Ter) including the PolyA site, was performed using Chromatin Immunoprecipitation (ChIP) followed by quantitative PCR, at various times as indicated (Figure 1). Cells that either stably express targeted shRNA silencing XPG and ERCC1 (named SiXPF), XPA, or transitory express siRNA against CTCF as well as their respective corresponding controls (named SiCtrl1-3, see experimental procedure) were treated with all-trans retinoic acid (t-RA).

Silencing ERCC1 also silenced XPF (Figure 1, left panel, see also (Gaillard and Wood, 2001)).

In each silenced cell lines, we observed a significant defect in the *RARβ2* mRNA expression compared to the SiCtrl1-3 cells (Figure 1, left panels). We noticed that as a function of the cell line and the set of experiments, *RARβ2* mRNA synthesis peak either at 3 or at 6 hours (h). We noticed *RARβ2* mRNA synthesis peak either at 3 or at 6 hours as a function of respective cell lines and set of experiments. Therefore ChIPs were performed either at 0, 3 and 6 hours post-treatment.

In SiCtrl1 cells at 6 hours post-treatment, we observed the concomitant recruitment of RAR, pol II and TFIIB together with the NER factors indicated by the presence of XPA, XPG and XPF (Figure 1, A2 and A6) at Pro paralleling *RARβ2* mRNA synthesis (left panel). All of these proteins were also observed at Ter (panels A3 and A7). Such recruitment pattern was not detected at position 65 kb (-65kb) and 323 kb (+323kb), upstream and downstream to the transcription start site (TSS) respectively (panels A1, A5 and A4, A8). The recruitment pattern at Pro/Ter is different from the one found at exon 6 that lacks RAR as well as TFIIB (Figure S1). In SiCtrl2 and SiCtrl3, the transcription and the NER factors were similarly recruited at Pro and Ter, 3 hours post t-RA treatment (panels C1-C8 and F1-F8). Moreover, we noticed that the presence of the NER proteins was much less significant at the -65kb and +323kb regions (Figure 1).

Previous work has described the sequential arrival of the NER factors following the preinitiation complex (PIC) formation, with XPC, RPA and XPA preceding XPG and XPF-ERCC1 (Le May et al., 2010). In SiXPG cells, XPA together with XPF and in SiXPF cells, XPA with XPG was detected at Pro although to a lower extent comparing with their respective SiCtrl1-2 (compare panels B2, B6, with panels A2, A6 and panels D2, D6 with panels C2 and C6). In SiXPG cells, transcription and NER factors are still recruited at Ter, but to a lower extent in comparison to SiCtrl1 (compare panels B3 and B7 with panels A3 and A7). However, in SiXPF cells, the co-recruitment of transcription machinery with XPA and XPG, obvious at Pro, was less significant at Ter (compare panels D2, D6 and D3, D7 with panels B2, B6 and B3, B7). In SiXPA cells, neither XPG nor XPF were detected at Pro and Ter (panels E6, E7 and panels C6, C7). Knowing that XPG stabilizes TFIIH architecture (Ito et al., 2007), it is not surprising that the absence of XPG might also partially disturb the formation of the entire transactivation complex, explaining the low recruitment of PIC and consequently a drop in *RARβ2* mRNA synthesis.

We next found that in SiCtrl1-3, the CCCTC-binding factor (CTCF), a chromatin organizer (Phillips and Corces, 2009), was recruited at Pro and Ter upon t-RA induction (panels A6, A7, C6, C7, F6, F7). However CTCF was hardly detected on these regions in SiXPG, SiXPF and SiXPA cells (panels B6, B7, D6, D7 and E6, E7 respectively) when comparing with SiCtrl1-2.

In SiCTCF cells, RAR $\beta$ 2 synthesis was significantly reduced compared to SiCtrl3 but not totally abolished (Figure 1, left panel). The ChIP assay showed a concomitant detection of the NER factors with the transcriptional machinery at Pro but not at Ter (compare panels G2, G6 with panels F2 and F6 and panels G3, G7 with panels F3, F7). Although the presence of TFIIB as well as RAR was less pronounced than Pol II at Pro, it was interesting to notice their absence at Ter in SiCTCF (panel G3). Our data also suggest that the loop is not absolutely required for basal level of transcription, pol II being detected at exon 6 and at Ter in SiCTCF cells (Figure S1 and Figure 1G3; see also below).

The correlated recruitment to distant regions (around 200 kilobases) along the RAR $\beta$ 2 locus suggested the formation of gene looping. We thus performed quantitative Chromatin Conformation Capture assays (q3C) and analyzed interactions between several regions along the RAR $\beta$ 2 gene (Vernimmen et al., 2007). Digestion of crosslinked chromatin by *HindIII* resulted in restriction fragments equidistantly separated by 60kb and containing either -65 kb, Pro, Ter, +323 kb as well as an intronic region (M1) of the RAR $\beta$ 2 locus (Figure 2, upper scheme). Ter and M1 of RAR $\beta$ 2 gene were used as baits. In t-RA treated SiCtrl1 cells, we observed that Pro could specifically and significantly interact with Ter at 6 hours (Figure 2A), paralleling the RAR $\beta$ 2 mRNA synthesis (Figure 1, left panel) and the recruitment of the transcriptional apparatus (Figure 1, A2, A3, A6 and A7). Similar observations were made at 3 hours in SiCtrl2-3 (data not shown). By contrast, in SiXPG, SiXPF and SiXPA cells as well as in SiCTCF, no spatial proximity between Ter and Pro was revealed. As controls, no specific interactions were observed between the intronic M1 bait and Pro as well as between all the other analyzed fragments upon t-RA treatment (Figure 2B).

The above data indicate that the t-RA induction could initiate CTCF-related long-range interaction between Pro and Ter of RAR $\beta$ 2 gene that was abolished in absence of either XPG or XPF underlining the role of these NER factors in chromatin reorganization.

### **The catalytic activity of both XPG and XPF are required for gene looping**

Consistent with the role of XPG and XPF-ERCC1 in NER (Gillet and Scharer, 2006; O'Donnovan et al., 1994; Sijbers et al., 1996), we were wondering about the requirement of

their endonuclease activity in chromatin looping upon transactivation. The SV40-immortalized XP-G (XPCS1RO) and XP-F (XP2YO) fibroblasts that were derived from a XP/CS and XP patient respectively (Ellison et al., 1998; Yagi and Takebe, 1983) stably transfected either with XPG/WT, XPG/E791A or with XPF/WT, XPF/D676A (Figure S2A). XPG/E791A and XPF/676A mutations were found to abolish the catalytic activity of these two endonucleases and consequently their ability to eliminate DNA damages when added in an *in vitro* NER assay ((Lalle et al., 2002; Staresincic et al., 2009) and Figure S2B).

RAR $\beta$ 2 mRNA synthesis was significantly inhibited in XPG/E791A and XPF/D676A cells as well as in the corresponding parental cells compared to the cells expressing either XPG/WT or XPF/WT respectively (Figure 3, panels A1, B1 and Figure S2, panels C and D). In XPG/WT and XPG/E791A cells, we observed a conserved recruitment of RAR, pol II, XPA, XPG and XPF both at Pro and Ter at 3 and 8 hours respectively (Figure 3, compare panels A2, A3 with panels A4, A5). In XPG/E791A cells, the optimal ChIP recruitment reflected the RAR $\beta$ 2 mRNA synthesis peak at 8 hours (panel A1). In XPF/D676A and XPF/WT cells, the recruitment patterns of transcription and NER factors were similar both at Pro and Ter (compare panels B4, B5 with panels B2, B3) contrary to what was observed at Ter in SiXPF cells (Figure 1, panels D3, D7). We also noticed that both XPG/E791A and XPF/D676A proteins were recruited with the other NER factors.

However, in spite of the presence of both endonucleases in XPG/E791A and XPF/D676 cells, q3C assay did not show any interaction between Pro and Ter while in both XPG/WT and XPF/WT cells, such interaction was observed at 3 and 6 hours upon t-RA induction respectively (Figure 3C and 3D). In XPG/WT and XPF/WT cells, CTCF was detected at both Pro and Ter concomitantly with gene looping (Figure 3, panels A2, A3 and B2, B3). In XPG/E791A or XPF/D676 cells, the inhibition of gene looping was coincident with the absence of CTCF (Figure 3, panels A4, A5 and B4, B5).

These data indicate that the t-RA induction could initiate CTCF-dependent long-range interactions between Pro and Ter of RAR $\beta$ 2 gene, which was obliterated in the absence of catalytically active XPG or XPF-ERCC1.

### **XPG endonuclease induces DNA breaks and DNA demethylation at promoter region.**

Chromatin rearrangement involving DNA twists might require DNA breaks (Ju et al., 2006). Using a Bio-ChIP assay which measures the incorporation of biotinylated dUTP within broken DNA, we observed a concomitant increase of DNA cleavage specifically both at Pro and Ter in all the five t-RA treated SiCtrl1, SiCtrl2, XPG/WT, XPF/WT and SiCtrl3 cells

(Figure 4, panel A1, B1, C1 and D1 respectively). In SiXPG and XPG/E791A cells, DNA breaks were not detected at Pro (Figure 4, panels A1, C1) whereas in SiXPF and XPF/D676 cells, DNA cleavage could only be observed at Pro (panels B1 and C1). However we noticed that in XPG deficient cells, there was a slight incorporation of bio-UTP at Ter (panel A1).

Several studies have documented some relationship between XPG and DNA demethylation upon transcription (Barreto et al., 2007; Jin et al., 2008; Schmitz et al., 2009). We thus investigated whether or not the endonuclease activity XPG and XPF was linked to DNA demethylation. Using an unmethylated DNA Immunoprecipitation (unMeDIP) approach, we measured the removal of the 5'-methyl-Cytosine (5mC) along the different regions of RAR $\beta$ 2. In all the five control cells, Pro and Ter were found unmethylated post t-RA treatment (Figure 4, panels A2, B2, C2 and D2). In SiXPG and XPG/E791A cells, Pro remained methylated while a DNA demethylation was observed at Ter when compared to their corresponding SiCtrl1 and XPG/WT cells (panels A2, C2). In SiXPF and XPF/D676 cells, Pro but not Ter was unmethylated (panels B2 and C2). Using the Methylated DNA IP (MeDIP) approach, no significant decrease was observed in DNA methylation at both Pro and Ter in t-RA treated XPG/E791A and XPF/D676A cells (Figure S3), which was found to be consistent with the UnMEDIP (see above) assays and pyrosequencing data (see below, panel E). In XPG/WT and XPF/WT, we observed a weak but significant decrease of DNA methylation. Interestingly in SiXPA, neither DNA breaks nor DNA demethylation was detected at Pro and Ter (panels B1 and B2 respectively), likely due to the absence of XPG and XP (Figure 1, panels E6 and E7).

Having demonstrated the connection between CTCF and the XPG and XPF endonucleases for chromatin looping, we then investigated whether the presence of CTCF is a prerequisite for DNA modifications. In SiCTCF, we repeatedly observed both DNA breaks and DNA demethylation at Pro but not at Ter when compared to SiCtrl3 (Figure 4, panels D1 and D2). These data strongly suggest a role for CTCF in the gene looping and Ter modification following DNA breaks and DNA demethylation of the Pro.

To further localize the DNA modifications that occur during the formation of the transactivation complex, the promoter of RAR $\beta$ 2 was pyrosequenced. Pyrosequencing analysis of Bio-ChIP DNA samples from t-RA treated XPG/WT and XPF/WT cells indicated a high frequency of DNA breaks on both strands localized 300 nucleotides downstream the RAR $\beta$ 2 TSS (Figure 4E, red arrows); most of them were detected on the transcribed strand (TS). In XPG/E791A cells, the frequency of DNA cleavage strongly decreased and the

detected breaks were differently localized (panel E, light blue arrows). Interestingly, in XPF/D676A, the cleavage pattern observed around the RAR $\beta$ 2 promoter is similar (panel E, dark blue arrow), demonstrating that XPF endonuclease inhibition does not prevent Bio-UTP incorporation at Pro (see also panels B1 and C1).

Genomic DNA was next bisulfite-converted and pyrosequenced to localize the demethylated 5mC at Pro in t-RA treated cells. Seventeen CpG dinucleotides, localized inside the proximal promoter and downstream the TSS, were analyzed (Figure 4, panel E). In XPG/WT, XPF/WT and XPF/D676A cells, CG1, CG3, CG15 and CG16 (closed dark blue circles) were significantly demethylated whereas only CG1 remained demethylated in XPG/E791A (Figure 4, panel E).

The above data underline the role of XPG and XPF endonucleases in both the formation of DNA breaks and active DNA demethylation at Pro and Ter respectively. Our data also show that XPG dependent-DNA breaks and DNA demethylation at Pro precede the recruitment of CTCF.

## Discussion

Upon t-RA treatment, RAR nuclear receptor as well as the transcriptional machinery including mediator, GTFs, pol II are recruited at the promoter of RAR $\beta$ 2 gene. Then, NER factors where XPA precedes XPG and XPF are sequentially recruited (Le May et al., 2010). Indeed silencing XPA abolishes the recruitment of XPG and XPF at the RAR $\beta$ 2 promoter (Figure 1), all of them being necessary for optimal gene expression (Barreto et al., 2007; Le May et al., 2010; Schmitz et al., 2009). We also observed a long-range juxtaposition of the promoter and the terminator of the activated RAR $\beta$ 2 gene, a situation that constitutes an “active chromatin hub” as previously described for  $\beta$ -Globin, *Igf2*, *BRCAl* loci and HIV-1 provirus (de Laat and Grosveld, 2003; Spilianakis and Flavell, 2004) (Perkins et al., 2008; Tan-Wong et al., 2008). This higher order folding of chromatin, already reported for inter- and intra-chromosomal interactions between specific subsets of NR-bound transcription units (Hu et al., 2008; Nunez et al., 2008), provides additional control mechanisms for regulated gene expression (Deng and Blobel, 2010; Nunez et al., 2008).

In the present study, we described how XPG and XPF might participate to optimal gene expression (see model Figure 5). We first show that the physical presence of both XPG and XPF is required to recruit the chromatin organizer CTCF and to participate in chromatin loop formation. Indeed, silencing either XPG or XPF is sufficient to inhibit the chromatin

rearrangement even in the presence of the other NER factors such as XPA (Figures 1 et 2). Moreover, mutations that affect the catalytic activity of XPG and XPF (XPG/E791A and XPF/D676A respectively) prevent CTCF recruitment, suggesting their functional involvement in the loop formation. In addition, we discovered that both chromatin looping and CTCF recruitment occur in parallel with the formation of DNA breaks. We especially observed that XPG/E791A mutation prevented DNA breaks formation at Pro (Figure 4). Since the XPF catalytic mutation does not affect those DNA modifications at Pro, it is tempting to suggest that upon t-RA transactivation and NER factors recruitment, XPG induces some DNA breaks at Pro independently of both the presence of CTCF and the status of XPF. Indeed, in SiCTCF, there are DNA breaks at Pro but not at Ter around which no NER factors were detected (Figure 4). Our results thus suggest that the action of XPF occurs once XPG has fulfilled its role at Pro and CTCF has been recruited.

Several studies have underlined the influence of DNA binding proteins such as activators and/or silencers in modifying chromatin environment in a DNA methylation dependent manner driving to specific loops formation and distinct gene expression as it was demonstrated for the imprinted *Igf2/H19* (Murrell et al., 2004) and *GATA-4* loci (Tiwari et al., 2008).

Using different approaches (UnMEDIP, MEDIP and pyrosequencing), we observed a relationship between the DNA demethylation at Pro and Ter and the presence of CTCF necessary for the loop formation of the RAR $\beta$ 2 activated gene. In particular we observed that demethylation at Pro is not sufficient to attract CTCF when Ter is not demethylated: XPG and XPF defects abolish DNA demethylation at Pro and Ter respectively and consequently CTCF recruitment (Figure 4). It thus is likely that following DNA breaks and DNA demethylation at Pro, CTCF recruitment is initiated with the assistance of XPF at Ter (Figure 5).

Mechanisms underlying the removal of methyl groups from genomic DNA are very controversial and have implicated DNA repair processes either through the deamination followed by excision of the 5mC or by direct removal of the methyl moiety from the base (Gehring et al., 2009; Ma et al., 2009). Our results do not allow the identification of the precise causal relationships between DNA demethylation and the effects of XPG and XPF endonucleases. They however support that XPG endonuclease and may be to a lower extent XPF, participate to the transcription process (Ito et al., 2007; Lee et al., 2002), either directly in the DNA demethylation itself and/or in the relaxation of chromatin conformation to drive gene activation (Barreto et al., 2007; Schmitz et al., 2009). Indeed, we found that the DNA breaks induced by XPG were located almost exclusively on the transcribed strand and nearby

the demethylated CG dinucleotides in the vicinity of RAR $\beta$ 2 TSS (Figure 4E). Whether or not DNA cuts are a prerequisite for DNA demethylation remains to be shown. However, the pattern of DNA cuts and demethylated 5mC is strongly disturbed when XPG and XPF endonucleases are inactivated.

Such localized roles of XPG at Pro and XPF at Ter as well as their cooperation in promoting loop formation and consequently optimal RAR $\beta$ 2 gene expression are reminiscent to the NER process in which XPF and XPG excise at the 3' and 5' of the DNA lesion, respectively in a cooperative manner (Tapias et al., 2004b)(Araujo et al., 2000; Staresincic et al., 2009). In NER, XPG mutations completely abolished the incision while XPF mutation still allow XPG cuts at the 5' side of the DNA damage (Lalle et al., 2002).

As chromatin loop formation is abolished upon deficiency of kin28 the yeast counterpart of human cdk7 kinase (O'Sullivan et al., 2004), it would be legitimate to say about the possibility of some other inter-connecting links in this cascade of events followed by the action of XPA, XPG, XPF and CTCF. In this later case, it is likely that defect in TFIIH that disturb either pol II (O'Sullivan et al., 2004) or NR phosphorylation (Compe et al., 2007) as well as mutation in the mediator that weaken interaction between activator and the basal transcription machineries (Hashimoto et al., 2011; Wang et al., 2005) indirectly disturb the proper recruitment and function of XPG/XPF endonucleases participating in DNA demethylation and DNA breaks formation and/or CTCF in chromatin loop formation.

The present study underlines the essential role of XPG, XPF and CTCF recruitment in the chromatin loop organization required for optimal expression of activated genes. Any kind of events that results in inaccurate transactivation complex formation as observed in XPA, XPG and XPF deficient cells, will result in altered gene expression that might at least partially explain the broad range of clinical features such as those observed within XP, TTD and XP/CS patients.

## **Experimental Procedure**

### **Cell Culture:**

HeLa Silencix cells (Tebu-Bio, provided by D. Biard) were used including SiCtrl (SiCtrl2, BD690), SiXPF and SiXPA cells. SiXPG cells and the corresponding control (SiCtrl1, SiLuc) were provided by R. Tanaka. SiCTCF and SiCtrl3 cells were obtained by respectively transfecting HeLa cells with SMART pool siRNA targeting CTCF or scrambled siRNA (Dharmacon) at a final concentration of 100nM. XPG/WT and XPG/E791A, as well as XPF/WT and XPF/D676A cells (gift from O. Scharer and W. Vermeulen) were obtained by the stable transfection of the corresponding wild type or mutated XPG and XPF constructs (Staresinic et al., 2009) in SV40-immortalized XP-G (XPCS1RO) and XP-F (XP2YO) (Ellison et al., 1998; Yagi and Takebe, 1983).

All cells were cultured in appropriate medium. Cells were incubated with red phenol-free medium containing 10% charcoal treated Fetal Calf Serum (FCS) and 40µg/ml gentamycin. Cells were treated with 10µM all-trans-retinoic acid (t-RA) into the same medium.

### **Antibodies:**

RNA pol II (7C2), RAR (9A6), XPA (1E9), TBP (3G3) and XPG (1B5) antibodies were from IGBMC antibody facilities. TFIIB (C-18), XPF (H-300) and Biotin (33) antibodies were obtained from Santa-Cruz Biotechnology,  $\beta$ -tubulin from Millipore and CTCF (ab70303) from Abcam.

### **Reverse Transcription and Quantitative PCR**

Total RNA was isolated from several cell lines using a GenElute Mammalian Total RNA Miniprep kit (Sigma) and reverse transcribed with SuperScript II reverse transcriptase (Invitrogen). The quantitative PCR was done using the Lightcycler 480 (Roche). The primer sequences for RAR $\beta$ 2 and Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) genes used in qPCR are available upon request. The RAR $\beta$ 2 mRNA expression represents the ratio between values obtained from treated and untreated cells normalized against the housekeeping GAPDH mRNA.

### **Chromatin Immunoprecipitation:**

Cells were cross-linked at room temperature (RT) for 10 min with 1% formaldehyde. Chromatin was prepared and sonicated on ice 30 min using a Bioruptor (Diagenode) as

previously described (Le May et al., 2010). Samples were immunoprecipitated with antibodies at 4°C overnight and protein G-sepharose beads (Upstate) were added, incubated 4 hours at 4°C and sequentially washed. Protein-DNA complexes were eluted and DNA fragments were purified using QIAquick PCR purification kit (QIAGEN) and analyzed by quantitative PCR.

**Biotin-ChIP (Bio-ChIP):**

Cross-linked cells through a 1% formaldehyde treatment for 10 min at RT, were permeabilized with cytonin (Active Motif) 30 min at RT. After extensive washes with Phosphate Buffer Salt (PBS), Terminal deoxynucleotidyl Transferase (TdT) reaction was performed using Biotin-16-dUTP (Roche) and 60 units of recombinant enzyme rTdT (Promega). TdT reaction was stopped with specific stop buffer (Chemicon International) 15 min at RT. After extensive washes with PBS, the resulted samples were sonicated on ice 20 min (40 cycles: pulse 10s, pause 20s) using a Bioruptor (Diagenode) and immunoprecipitated using anti-Biotin antibodies and treated as described in the ChIP protocol. DNA fragments were purified using QIAquick PCR purification kit (QIAGEN) and analyzed by quantitative PCR using sets of primers,

**Quantitative Chromosome Conformation Capture (q3C):**

The quantitative chromosome conformation capture assay was performed as previously described (Vernimmen et al., 2007). Briefly, cells were cross-linked at RT for 10 min with 2% formaldehyde. Cross-linked chromatin was digested in the appropriate restriction buffer by 400 units of enzyme *HindIII*. The restriction enzyme mixture was stopped, diluted in ligation buffer and incubated with the highly concentrated T4 DNA ligase (Roche) for 4 hours at 16°C. The cross-linking was heat-reversed and DNA fragments were purified. Undigested DNA or digested DNA without ligation step was used as negative controls. Moreover, we used as internal positive control the endogenous *xpb* locus that has been reported to adopt the same spatial conformation in different tissues (Vernimmen et al., 2007). All q3C results were normalized by data from *xpb* analysis (see Figure S2G), controlling for changes in nuclear size, chromatin density and cross-linking efficiency. Primers and probes were designed as follows: a universal sequence-specific Taqman probe and corresponding reverse primer on a fixed restriction fragment (Ter or M1) in combination with different forward primers specific of other restriction fragments (see upper panel Figure 2). Quantitative 3C templates (200ng)

were used for Taqman/PCR reaction using the universal PCR Master Mix and the Lightcycler 480 from Roche.

#### **Unmethylated DNA Immunoprecipitation (unMeDIP):**

Genomic DNA was extracted using GenElute Mammalian Genomic DNA Mini-prep Kit (Sigma). Unmethylation of 5mC on the RAR $\beta$ 2 locus was measured by digesting 2 $\mu$ g genomic DNA with 10 units of *MseI* (Fermentas) and by using the UnMethylcollector kit (Active Motif). Unmethylated DNA Immunoprecipitation kit (UnMeDIP) is based on the affinity of the three zinc-coordinating CXXC domains, localized in chromatin-associated proteins such as DNA methyltransferase 1 (DNMT1) or mixed lineage leukemia (MLL), that specifically bind nonmethylated CpG sites. The resulted samples were selected using magnetic beads conjugated with CXXC domains, extensively washed and analyzed by quantitative PCR.

#### **Pyrosequencing**

**Sequencing analyses.** Pyrosequencing was performed on a Pyromark Q96 ID platform using the PSQ Gold SQA reagent kit (Qiagen). Briefly, 20 $\mu$ l amplified DNA products were mixed with 2 $\mu$ l Streptavidin Sepharose beads (Amersham Biosciences AB), 38 $\mu$ l binding buffer (10 mM Tris-HCl, pH 7.6, 2 M NaCl, 1 mM EDTA, and 0.1% Tween-20) and 20 $\mu$ L H<sub>2</sub>O, followed by shaking for 10 min. Then, the immobilized biotinylated PCR products—Streptavidin Sepharose beads complex was captured using the Qiagen Pyromark Q96 Work Station. Single strand DNA purification was achieved by sequential washes with 75% ethanol, 0.2 M NaOH and washing buffer (10 mM Tris-acetate, pH7.6). The unbiotinylated strand was dissociated and discarded. The immobilized single biotinylated strands were released to a 96-well microtiter plate, which was pre-added with 38 $\mu$ l annealing buffer and 2 $\mu$ l complementary sequencing primer. The plate was incubated at 80°C for 2 min, followed by slow cooling to room temperature. The processed mixture was loaded onto the PyroMark ID system equipped with PyroMark ID software. The resulting pyrograms and associated sequences were generated and analyzed automatically using PSQ 96 SQA software (Qiagen).

**Methylation analyses.** Pyrosequencing was performed using a PyroMark Q96 ID Pyrosequencing instrument (Qiagen). Pyrosequencing assays were designed using MethPrimer (<http://www.urogene.org/methprimer/index1.html>). 200ng of genomic DNA was treated with EZ DNA Methylation-Gold Kit (Zymo Research Corporation) and amplified using 1.25 U Platinum Taq DNA Polymerase (Invitrogen) and 1 $\mu$ M each forward and reverse

primers in a 50 $\mu$ L reaction volume with 0.2mM dNTP and 1.5mM MgCl<sub>2</sub>. PCR conditions were 94°C for 3 min, then 40 cycles of 94°C for 30 sec, T<sub>m</sub> for 30 sec and 72°C for 30 sec, followed by a 9 min extension at 72°C. Biotin-labeled single stranded amplicons were isolated using the Qiagen Pyromark Q96 Work Station and underwent pyrosequencing with 0.5  $\mu$ M primer. The percent methylation for each of the CpGs within the target sequence was calculated using PyroQ cpG Software (Qiagen). All methylation analyses were performed in duplicate.

All the sets of primers and probes targeting different regions of RAR $\beta$ 2 gene and used for the qPCR analysis are available upon request.

## REFERENCES

- Araujo, S.J., Tirode, F., Coin, F., Pospiech, H., Syvaioja, J.E., Stucki, M., Hubscher, U., Egly, J.M., and Wood, R.D. (2000). Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes Dev* 14, 349-359.
- Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Doderlein, G., Maltry, N., Wu, W., Lyko, F., *et al.* (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* 445, 671-675.
- Brivanlou, A.H., and Darnell, J.E., Jr. (2002). Signal transduction and the control of gene expression. *Science* 295, 813-818.
- Compe, E., Malerba, M., Soler, L., Marescaux, J., Borrelli, E., and Egly, J.M. (2007). Neurological defects in trichothiodystrophy reveal a coactivator function of TFIIH. *Nat Neurosci* 10, 1414-1422.
- de Laat, W., and Grosveld, F. (2003). Spatial organization of gene expression: the active chromatin hub. *Chromosome Res* 11, 447-459.
- Deng, W., and Blobel, G.A. (2010). Do chromatin loops provide epigenetic gene expression states? *Curr Opin Genet Dev* 20, 548-554.
- Ellison, A.R., Nospikel, T., Jaspers, N.G., Clarkson, S.G., and Gruenert, D.C. (1998). Complementation of transformed fibroblasts from patients with combined xeroderma pigmentosum-Cockayne syndrome. *Exp Cell Res* 243, 22-28.
- Gaillard, P.H., and Wood, R.D. (2001). Activity of individual ERCC1 and XPF subunits in DNA nucleotide excision repair. *Nucleic Acids Res* 29, 872-879.
- Gehring, M., Bubb, K.L., and Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science* 324, 1447-1451.
- Gillet, L.C., and Scharer, O.D. (2006). Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem Rev* 106, 253-276.
- Hashimoto, S., Boissel, S., Zarhrate, M., Rio, M., Munnich, A., Egly, J.M., and Colleaux, L. (2011). MED23 mutation links intellectual disability to dysregulation of immediate early gene expression. *Science* 333, 1161-1163.
- Hu, Q., Kwon, Y.S., Nunez, E., Cardamone, M.D., Hutt, K.R., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G., *et al.* (2008). Enhancing nuclear receptor-induced

- transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. *Proc Natl Acad Sci U S A* *105*, 19199-19204.
- Ito, S., Kuraoka, I., Chymkowitz, P., Compe, E., Takedachi, A., Ishigami, C., Coin, F., Egly, J.M., and Tanaka, K. (2007). XPG Stabilizes TFIIH, Allowing Transactivation of Nuclear Receptors: Implications for Cockayne Syndrome in XP-G/CS Patients. *Mol Cell* *26*, 231-243.
- Jaspers, N.G., Raams, A., Silengo, M.C., Wijgers, N., Niedernhofer, L.J., Robinson, A.R., Giglia-Mari, G., Hoogstraten, D., Kleijer, W.J., Hoeijmakers, J.H., *et al.* (2007). First reported patient with human ERCC1 deficiency has cerebro-oculo-facio-skeletal syndrome with a mild defect in nucleotide excision repair and severe developmental failure. *Am J Hum Genet* *80*, 457-466.
- Jin, S.G., Guo, C., and Pfeifer, G.P. (2008). GADD45A does not promote DNA demethylation. *PLoS Genet* *4*, e1000013.
- Ju, B.G., Lunyak, V.V., Perissi, V., Garcia-Bassets, I., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (2006). A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. *Science* *312*, 1798-1802.
- Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. *Proc Natl Acad Sci U S A* *104*, 12955-12961.
- Kraemer, K.H., Patronas, N.J., Schiffmann, R., Brooks, B.P., Tamura, D., and DiGiovanna, J.J. (2007). Xeroderma pigmentosum, trichothiodystrophy and Cockayne syndrome: a complex genotype-phenotype relationship. *Neuroscience* *145*, 1388-1396.
- Lalle, P., Nospikel, T., Constantinou, A., Thorel, F., and Clarkson, S.G. (2002). The founding members of xeroderma pigmentosum group G produce XPG protein with severely impaired endonuclease activity. *J Invest Dermatol* *118*, 344-351.
- Le May, N., Mota-Fernandes, D., Velez-Cruz, R., Iltis, I., Biard, D., and Egly, J.M. (2010). NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. *Mol Cell* *38*, 54-66.
- Lee, S.K., Yu, S.L., Prakash, L., and Prakash, S. (2002). Requirement of Yeast RAD2, a Homolog of Human XPG Gene, for Efficient RNA Polymerase II Transcription. Implications for Cockayne Syndrome. *Cell* *109*, 823-834.
- Ma, D.K., Guo, J.U., Ming, G.L., and Song, H. (2009). DNA excision repair proteins and Gadd45 as molecular players for active DNA demethylation. *Cell Cycle* *8*, 1526-1531.
- Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat Genet* *36*, 889-893.
- Niedernhofer, L.J., Garinis, G.A., Raams, A., Lalai, A.S., Robinson, A.R., Appeldoorn, E., Odijk, H., Oostendorp, R., Ahmad, A., van Leeuwen, W., *et al.* (2006). A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* *444*, 1038-1043.
- Nospikel, T. (2009). DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility. *Cell Mol Life Sci* *66*, 994-1009.
- Nunez, E., Kwon, Y.S., Hutt, K.R., Hu, Q., Cardamone, M.D., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G., *et al.* (2008). Nuclear receptor-enhanced transcription requires motor- and LSD1-dependent gene networking in interchromatin granules. *Cell* *134*, 189.
- O'Donovan, A., Davies, A.A., Moggs, J.G., West, S.C., and Wood, R.D. (1994). XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* *371*, 432-435.
- O'Sullivan, J.M., Tan-Wong, S.M., Morillon, A., Lee, B., Coles, J., Mellor, J., and Proudfoot, N.J. (2004). Gene loops juxtapose promoters and terminators in yeast. *Nat Genet* *36*, 1014-1018.

- Perkins, K.J., Lusic, M., Mitar, I., Giacca, M., and Proudfoot, N.J. (2008). Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. *Mol Cell* 29, 56-68.
- Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. *Cell* 137, 1194-1211.
- Riedl, T., Hanaoka, F., and Egly, J.M. (2003). The comings and goings of nucleotide excision repair factors on damaged DNA. *Embo J* 22, 5293-5303.
- Schmitz, K.M., Schmitt, N., Hoffmann-Rohrer, U., Schafer, A., Grummt, I., and Mayer, C. (2009). TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. *Mol Cell* 33, 344-353.
- 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., *et al.* (1996). Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* 86, 811-822.
- Spilianakis, C.G., and Flavell, R.A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol* 5, 1017-1027.
- Staresincic, L., Fagbemi, A.F., Enzlin, J.H., Gourdin, A.M., Wijgers, N., Dunand-Sauthier, I., Giglia-Mari, G., Clarkson, S.G., Vermeulen, W., and Scharer, O.D. (2009). Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO J* 28, 1111-1120.
- Tan-Wong, S.M., French, J.D., Proudfoot, N.J., and Brown, M.A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. *Proc Natl Acad Sci U S A* 105, 5160-5165.
- Tapias, A., Auriol, J., Forget, D., Enzlin, J., Scharer, O., Coin, F., Coulombe, B., and Egly, J. (2004a). Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *J Biol Chem*, in press.
- Tapias, A., Auriol, J., Forget, D., Enzlin, J.H., Scharer, O.D., Coin, F., Coulombe, B., and Egly, J.M. (2004b). Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *J Biol Chem* 279, 19074-19083.
- Tiwari, V.K., McGarvey, K.M., Licchesi, J.D., Ohm, J.E., Herman, J.G., Schubeler, D., and Baylin, S.B. (2008). PcG proteins, DNA methylation, and gene repression by chromatin looping. *PLoS Biol* 6, 2911-2927.
- Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J* 26, 2041-2051.
- Wang, G., Balamotis, M.A., Stevens, J.L., Yamaguchi, Y., Handa, H., and Berk, A.J. (2005). Mediator requirement for both recruitment and postrecruitment steps in transcription initiation. *Mol Cell* 17, 683-694.
- Yagi, T., and Takebe, H. (1983). Establishment by SV40 transformation and characteristics of a cell line of xeroderma pigmentosum belonging to complementation group F. *Mutat Res* 112, 59-66.

### Acknowledgements

We thank C. Braun for her technical expertise, D. Vernimmen for his help in designing q3C experiment, D. Biard, O. Scharer and W. Vermeulen for providing the different cells and the IGBMC cell culture facilities. We also are grateful to F. Coin, E. Compe, A. Singh and L. Dubs for fruitful discussions and critical reading of the manuscript. This study was supported by an ERC Advanced grant (to J.M.E.), l'Agence Nationale de la Recherche (N#ANR-

08MIEN-022-03), l'Association de la Recherche contre le Cancer, the Institut National du Cancer (INCA-2008-041) and l'Association de Recherche sur le Diabète (ARD grant 2009-01 to PB) and research grants from NovoNordisk France ("PharmacoEpigenetics of Growth" 2009-2011 to PB) and Pfizer France ("Candidate Gene Search for Shortness Gene Variations" GETC-2010-2028326 to PB). N.L.M. is recipient from the INSERM; I.I. is granted by the l'Association de la Recherche contre le Cancer; DF is the recipient of a young investigator support grant from INSERM.

### **Figures Legends:**

#### **Figure 1: XPG, XPF and CTCF are necessary for concomitant recruitment of transcription machinery at promoter and terminator upon RAR $\beta$ 2 induction.**

Schematic representation of the RAR $\beta$ 2 gene with the indicated amplicons designed at the -65kb, Pro, Ter and +323kb regions (Upper panel).

Overtime relative RAR $\beta$ 2 mRNA expression monitored by qPCR from t-RA treated SiXPG, SiXPF, SiXPA and siCTCF HeLa cells, and the corresponding SiCtrl 1, SiCtrl 2 and SiCtrl 3 control cells (left panels). Error bars represent the standard deviation of three independent experiments. Western-blotting analysis of XPG, XPF, XPA CTCF and TBP were performed from chromatin extracts of SiXPG, SiXPF, SiXPA and SiCTCF HeLa cells and their corresponding controls (left panels). TBP is used as a positive nuclear marker.

ChIP monitoring the t-RA-dependent occupancy of RAR, pol II, TFIIB (A1-4, B1-4, C1-4, D1-4, E1-4, F1-4, G1-4), XPA, XPG, XPF and CTCF (A5-8, B5-8, C5-8, D5-8, E5-8, F5-8, G5-8) on RAR $\beta$ 2 locus from SiXPG (B1-8), SiXPF (D1-8), SiXPA (E1-8), SiCTCF (G1-8) and the corresponding controls (SiCtrl1, SiCtrl 2, SiCtrl3) chromatin extracts at 0, 3 and 6 hours upon t-RA induction as indicated. A1-4, A5-8 --- G5-G8 are aligned under the corresponding region of the RAR $\beta$ 2 gene being probed. Each series of ChIP is representative of at least two independent experiments and values are expressed as % of the input.

#### **Figure 2: XPG and XPF is involved in CTCF-induced RAR $\beta$ 2 gene looping.**

Schematic representation of the quantitative Chromatin Conformation Capture (q3C) for the analysis of the intra-chromosomal interactions along the RAR $\beta$ 2 locus (upper panel). Two probes were designed at Ter (grey square) and M1 (green square) of RAR $\beta$ 2 gene. These probes were used to investigate the associations between the different elements including upstream (-65kb, Pro), intronic (M1, M2) and downstream (Ter, +257kb, +323kb, +385kb) regions as indicated by the black (Ter probe) and red (M1 probe) arrows.

q3C assays were performed using crosslinked and *HindIII*-digested chromatin from SiCtrl1, SiXPG, SiXPA, SiXPF and siCTCF HeLa cells at 0, 3 and 6 hours post t-RA treatment (10  $\mu$ M). The bar chart (Y axis) shows the enrichment of PCR product (%) normalized to the enrichment within the human *xpb* gene (=100%), as illustrated in Figure S2G. Each PCR was performed at least three times and averaged as indicated by standard deviation. Signals were normalized to the total amount of DNA used, estimated with an amplicon located within a *HindIII* fragment in RAR $\beta$ 2 gene (see experimental procedure).

**Figure 3: Endonuclease activity of XPG and XPF are required for gene looping.**

Relative RAR $\beta$ 2 mRNA expression monitored by quantitative PCR from XPG/WT (A1), XPG/E791A (A1), XPF/WT (B1), XPF/D676A (B1) cells at 0, 1, 3, 6, 8, 10 hours post t-RA treatment. Error bars represent the standard deviation of three independent experiments.

ChIP monitoring the t-RA-dependent occupancy of RAR, pol II, XPA, XPG, XPF and CTCF at Pro and Ter regions of RAR $\beta$ 2 gene from XPG/WT (A2-3), XPG/E791A (A4-5) and XPF/WT (B2-3), XPF/D676A (B4-5) chromatin extracts as indicated. Each series of ChIP is representative of two independent experiments.

Intra-chromosomal interactions between Ter and upstream/downstream elements of RAR $\beta$ 2 gene: Q3C assays were performed using crosslinked, *HindIII*-digested chromatin from t-Ra treated XPG/WT or XPG/E791A (C) and XPF/WT or XPF/D676A (D) fibroblasts at 0, 3 and 6 hours. These independent graphs represent a measure of the association between the upstream -65kb, Pro, the intronic M1 and M2 elements, the downstream +257 kb, +323kb and +385kb regions to either the terminal region (Ter Probe) or the intronic region (M1 Probe, see Figure S2, panels E and F) as negative control. Each PCR was performed at least three times and averaged as indicated by standard deviation. Signals were normalized to the total amount of DNA used, estimated with an amplicon located within a *HindIII* fragment in RAR $\beta$ 2 gene.

**Figure 4: XPG and XPF are required for DNA breaks and DNA demethylation at Pro and Ter respectively.**

Detection of DNA break at -65kb, Pro, Ter and +323kb of the RAR $\beta$ 2 locus from SiCtrl1, SiXPG (A1), SiCtrl2, SiXPA, SiXPF (B1), XPG/WT, XPG/E791A (C1) and XPF/WT, XPF/D676A (D1) t-RA treated cells. DNA break are detected through firstly the incorporation of Biotin-dUTP via a Terminal deoxynucleotidyl Transferase (TdT) reaction and secondly a ChIP approach. DNA fragments containing DNA breaks are immunoselected

by biotin antibodies and analyzed by qPCR. Each serie of Bio-ChIP is representative of three independent experiments and values are expressed as % of the input.

Analysis of the unmethylation of DNA (UnMedIP) at -65kb, Pro, Ter and +323kb of the RAR $\beta$ 2 locus from SiCtrl1, SiXPG (A2), SiCtrl2, SiXPA, SiXPF (B2), XPG/WT, XPG/E791A (C2) and XPF/WT, XPF/D676A (D2) t-RA treated cells. UnMedIP was performed using the UnMethylCollector kit (Active Motif). *MseI* digested genomic DNA was selected using magnetic beads conjugated with CXXC domains and the resulted samples containing unmethylated DNA were analyzed by qPCR. Each serie of UnMedIP is representative of two independent experiments and values are expressed as % of the input.

(E) Schematic representation of the methylated (open circle) or unmethylated (close circle) CpG dinucleotides (in red) and DNA break (close triangle) along the RAR $\beta$ 2 promoter for WT (red), XPG/E791A (light blue) and XPF/D676A (dark blue) cells. DNA break was localized at Pro by pyrosequencing the samples generated by Bio-ChIP. The methylated status of CpG dinucleotides at Pro was measured using bisulfite converted genomic DNA from XPG/WT, XPG/E791A and XPF/WT, XPF/D676A fibroblasts and pyrosequencing.

**Figure 5: A role for XPG and XPF in the CTCF-dependent gene looping for optimal gene expression.**

Upon the ligand induction, Nuclear Receptor (NR) targets its responsive element (RE). The transactivation complex is then formed at the promoter (PRO) including the arrival of GTFs (IIA, IIB, IIE, IIF, IIH) and pol II. After the formation of the PIC, NER factors (XPA, XPG and XPF/ERCC1) are sequentially recruited. XPG (red arrow) initiates DNA break (doted DNA) and demethylation (meC--C) at PRO, thus favouring the CTCF recruitment and gene looping between PRO and the terminator TER; Concomitantly, XPF (red arrow) induces DNA break (doted DNA) and DNA demethylation at TER.

## SUPPLEMENTAL INFORMATION

### Supplemental figures

#### **Figure S1. Presence of NER factors actors at exon 6 of RAR $\beta$ 2:**

ChIP monitoring the occupancy of RAR, pol II, TFIIB, XPA, XPG, XPF and CTCF at RAR $\beta$ 2 exon 6 from SiXPG (B), SiXPA (D), SiXPF (E), SiCTCF (G) and the corresponding controls SiCtrl1 (A), SiCtrl 2 (C), SiCtrl3 (F) chromatin extracts at 0, 3 and 6 hours post t-RA treatment as indicated. Each series of ChIP is representative of two independent experiments and values are expressed as % of the input.

#### **Figure S2. Endonuclease activity of XPG and XPF; negative control for q3C for XPG and XPF cells.**

(A) Steady-state level of XPG, XPF and b-tubulin in XP-G, XPG/WT, XPG/E791A (left panel) and XP-F, XPF/WT, XPF/D676A (right panel) cells. Western-blotting analysis was performed from chromatin extracts of the different cells. Detection of XPG, XPF, and b-tubulin (as a positive marker) was tested with appropriate antibodies.

(B) In vitro NER assay (Lalle et al., 2002; Tapias et al., 2004a) using nuclear extracts (NE) from XPF/WT (lanes 1 to 4) and XPF/D676A (lanes 5 to 8) cells expressing either XPF/WT or XPF/D676A and supplemented by either recombinant XPF or XPF/WT as indicated, (lanes 4 and 8). As a control, the NER assay was realized either with HeLa nuclear extracts (NE) or recombinant XPF alone (lanes 9 and 10 respectively).

Relative RAR $\beta$ 2 mRNA expression monitored by quantitative PCR from the parental XP-G, XPG/WT or XPG/E791A (panel C) and the parental XP-F, XPF/WT or XPF/D676A (panel D) cells treated overtime (0, 1, 3, 6, 8, 10 hours) with t-RA (10  $\mu$ M).

Intra-chromosomal interactions between internal M1 regions of RAR $\beta$ 2 gene and upstream/downstream elements as indicated figure 3 from t-RA treated XPG/WT or XPG/E791A (panel E) and XPF/WT or XPF/D676A cells (panel F).

(G) *xpb* level was monitored in XPG/WT, XPG/E791A, XPF/WT and XPF/D676A cells untreated (T0) or treated (T(h)) with t-RA (10  $\mu$ M) by PCR.

#### **Figure S3. DNA methylation measured by MedIP in XPG/WT, XPG/E791A, XPF/WT and XPF/D676A fibroblasts.**

Analysis of the DNA methylation (MedIP) at Pro and Ter from untreated (T0) or t-RA treated (T(h)) XPG/WT, XPG/E791A, XPF/WT, XPF/D676A chromatin extracts. MedIP was performed using the MethylCollector ultra kit (Active Motif). *MseI* digested genomic DNA was selected using His-tagged MBD2b/MDB3L1 protein complexes and nickel-coated magnetic beads. The resulted samples containing methylated DNA were analyzed by qPCR. Each serie of MedIP is representative of two independent experiments and values are expressed as % of the input.

#### **Figure S4. CpG methylation in XPG and XPF cells.**

**(A)** Specific primers sequences using for pyrosequencing from XPG/WT and XPG/E791 and XPF/WT and XPF/D676A cells.

**(B)** Histograms show the percentage of methylation of the different cytosine along the promoter for XPG/WT or XPF/WT (grey bar) and XPG/E791 XPF/D676A (black bar) cells lines. Each serie of pyrosequencing is representative of two independent experiments.

#### **Methods**

##### **In vitro NER assay:**

Reconstituted dual incision reactions (Riedl et al., 2003) were carried out in a buffer containing 50 mM HEPES–KOH pH 7.6, 20 mM Tris–HCl pH 7.6, 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 0.5 mM DTT, 0.5 mM EDTA, 10% glycerol, 0.02% NP40 and 2 mM ATP. NER factors were purified as described (Araujo et al., 2000) and used in saturating amounts: XPC-HR23B (10 ng), TFIIH HAP (50 ng), XPA (30 ng), RPA (200 ng), XPG (50 ng) and XPF-ERCC1 (10 ng) per reaction. The NER factors were incubated with 40 fmols of either free or immobilized damaged DNA substrate, for 40 min at 30°C in the presence of ATP. Reactions were stopped by boiling for 5 min after addition of 9 ng 32 nt oligonucleotide complementary to the excised DNA fragment with a 5'-extension of four extra G residues. After annealing the oligonucleotide to the excised DNA fragment, excision products were radiolabelled by extension with 0.15 U of Sequenase version 2.0 polymerase (U.S.B.) and 2 Ci of [<sup>32</sup>P]dCTP (3000 Ci/mmol). Labelled excision products were separated on a denaturing 14% polyacrylamide gel and visualized by autoradiography. Dual incision reactions with HeLa NE were performed in the presence of 15 M wortmannin, 60 M aphidicolin and 2% DMSO; reactions were pre-incubated for 10 min at 30°C before DNA substrate addition and further incubated at 30°C for 40 min or indicated times.

**Methylated DNA Immunoprecipitation (MeDIP):**

Genomic DNA was extracted using GenElute Mammalian Genomic DNA Mini-prep Kit (Sigma). Methylation status of CpG dinucleotides localized on the RAR $\beta$ 2 locus was investigated by digesting 2 $\mu$ g genomic DNA with 10 units of *MseI* (Fermentas) and was performed using the Methylcollector ultra kit as described by Active Motif. Methylated DNA Immunoprecipitation kit (MeDIP) is based on the high affinity of the Methyl-CpG binding proteins (MBD) MBD2b and its binding partner MBD3L1 for the methylated DNA fragments. The resulted samples were selected using His-tagged MBD2b/MBD3L1 protein complexes and nickel-coated magnetic beads, extensively washed and analyzed by quantitative PCR using sets of primers, available upon request, targeting different regions of RAR $\beta$ 2 gene.

Figure 1

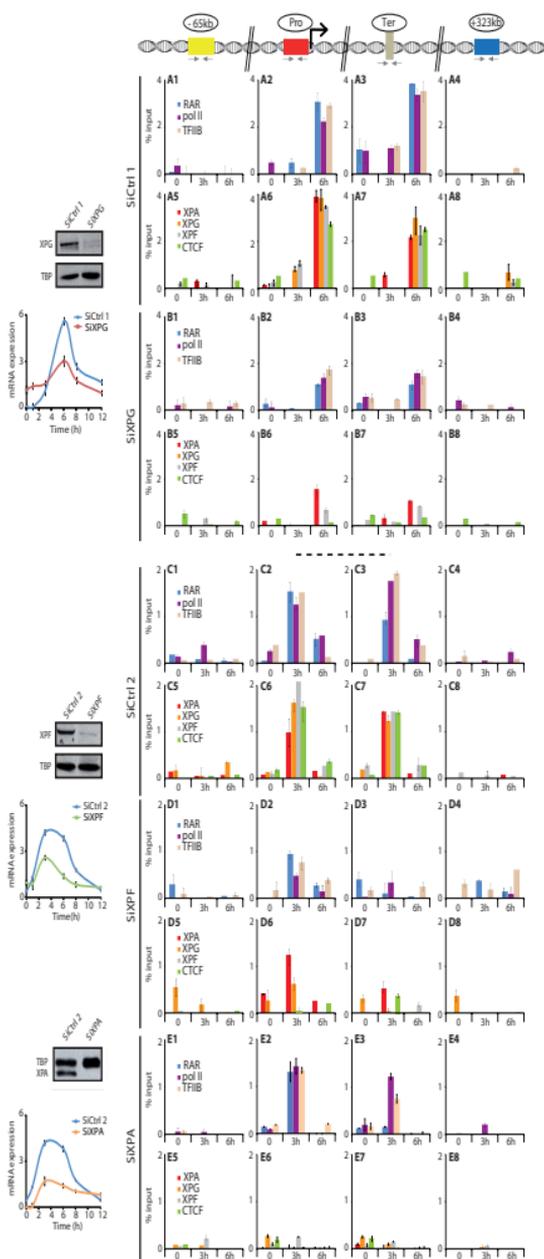


Figure 1 suite

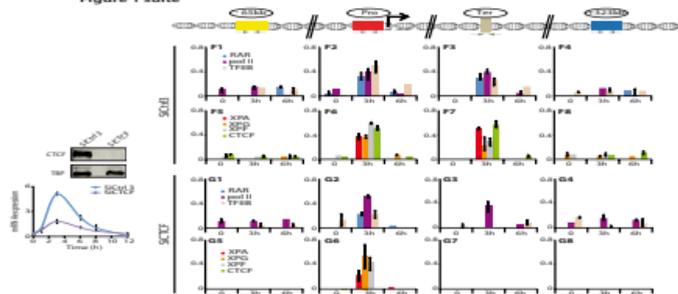


Figure 2

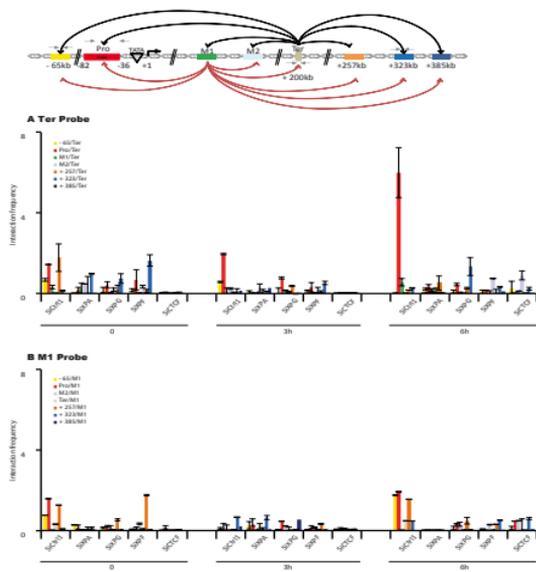


Figure 3

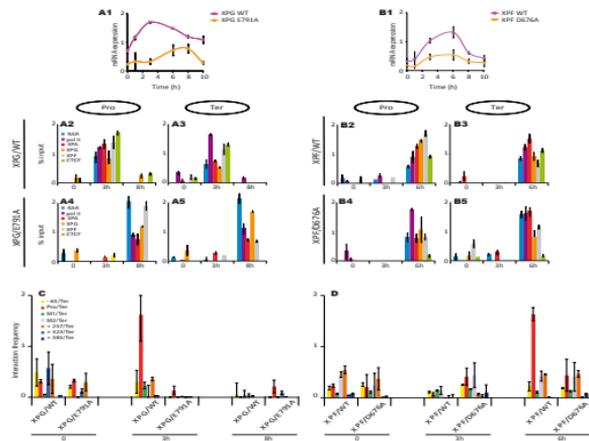




Figure 5

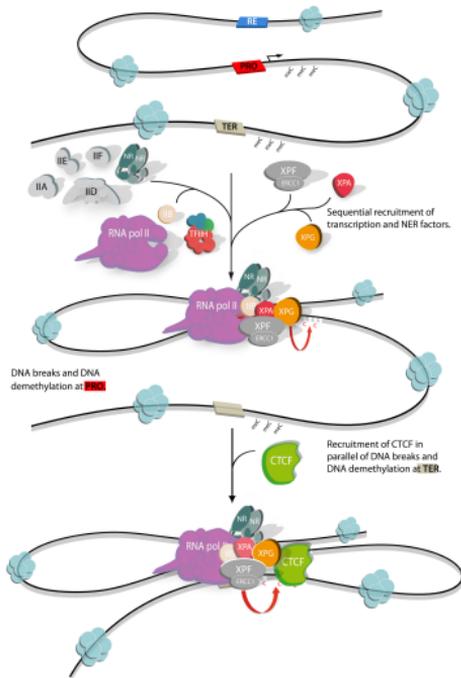


Figure S1

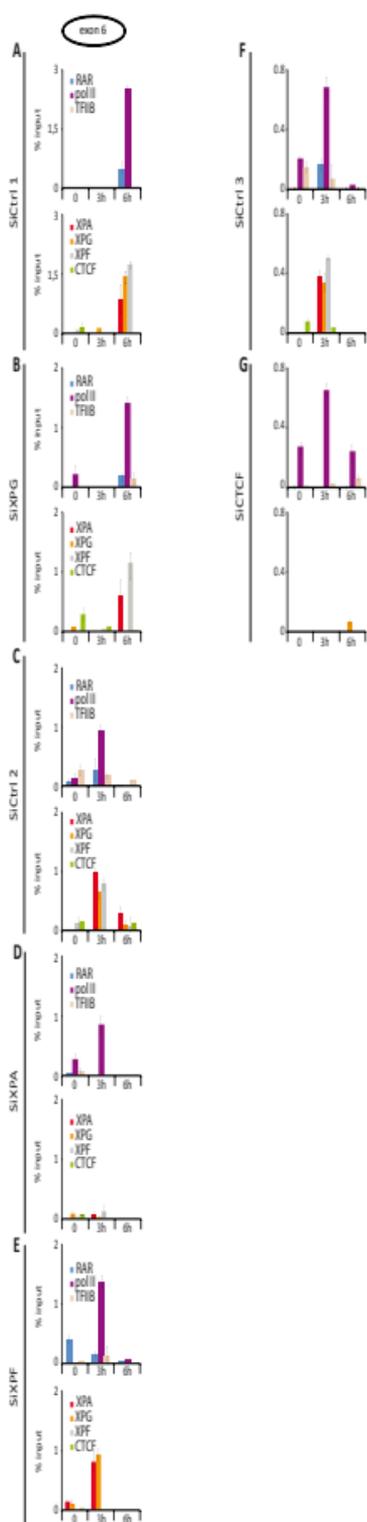


Figure S2

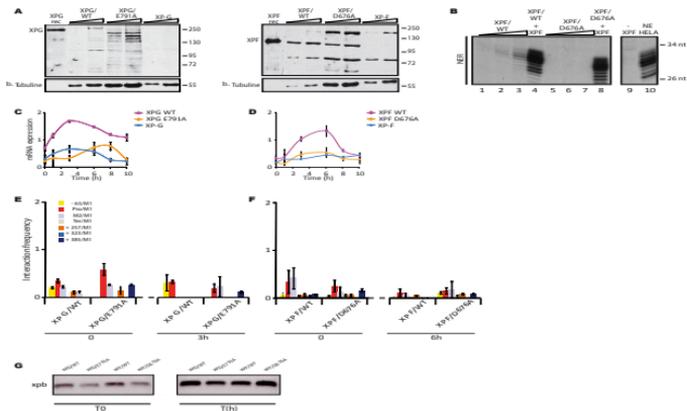


Figure S3

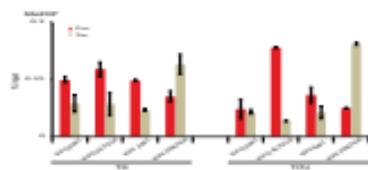


Figure S4

**A**

Primers		XPG/WT	XPG/B791A
F1	CGAGC TDTT TGAGGAC TGGGATCCGAGAAAGCGG	CGAGC TDTT TGAGGAC TGGGATCCGAGAAAGCGG	CGAGC TDTT TGAGGAC TGGGATCCGAGAAAGCGG
F2	CTD TGGGCTTGGATCTGGGAAGCATTCGGAGGG TTTTTCAGAGCA	CTD TGGGCTTGGATCTGGGAAGCATTCGGAGGG TTTTTCAGAGCA	CTD TGGGCTTGGATCTGGGAAGCATTCGGAGGG TTTTTCAGAGCA
F3	CGCCCCGGTGGCA	CGCCCCGGTGGCA	CGCCCCGGTGGCA
F4	CGAAAGGCGAGCGATCCGAGCAAGGGT TCTCTGGGACAGCT	CGAAAGGCGAGCGATCCGAGCAAGGGT TCTCTGGGACAGCT	CGAAAGGCGAGCGATCCGAGCAAGGGT TCTCTGGGACAGCT
F5	GC TGGAT TGGCCCGAGCAAGGCT TGGAAAATGCAARTT	GC TGGAT TGGCCCGAGCAAGGCT TGGAAAATGCAARTT	GC TGGAT TGGCCCGAGCAAGGCT TGGAAAATGCAARTT
B1	CTATGATCCGAATTAAGCCAACTGACACCGGGGCGAGGGGGGDTTCCCCAGCTAA	CTATGATCCGAATTAAGCCAACTGACACCGGGGCGAGGGGGGDTTCCCCAGCTAA	CTATGATCCGAATTAAGCCAACTGACACCGGGGCGAGGGGGGDTTCCCCAGCTAA
B2	CCGAATGCGTTCGGGATCC TACCCC	CCGAATGCGTTCGGGATCC TACCCC	CCGAATGCGTTCGGGATCC TACCCC
B3	CGGATC	CGGATC	CGGATC
B6	AGATCCCAAGCATCTCTCTCC TAAGCTAAGTGT TTAGCAA	AGATCCCAAGCATCTCTCTCC TAAGCTAAGTGT TTAGCAA	AGATCCCAAGCATCTCTCTCC TAAGCTAAGTGT TTAGCAA
Primers		XPF/WT	XPF/B47A
F1	CGAGC TDTT TGAGGAC TGGGATCCGAGAAAGCGG	CGAGC TDTT TGAGGAC TGGGATCCGAGAAAGCGG	CGAGC TDTT TGAGGAC TGGGATCCGAGAAAGCGG
F2	CTD TGGGCTTGGATCTGGGAAGCATTCGGAGGG TTTTTCAGAGCA	CTD TGGGCTTGGATCTGGGAAGCATTCGGAGGG TTTTTCAGAGCA	CTD TGGGCTTGGATCTGGGAAGCATTCGGAGGG TTTTTCAGAGCA
F3	CGCCCCGGTGGCA	CGCCCCGGTGGCA	CGCCCCGGTGGCA
F4	CGAAAGGCGAGCGATCCGAGCAAGGGT TCTCTGGGACAGCT	CGAAAGGCGAGCGATCCGAGCAAGGGT TCTCTGGGACAGCT	CGAAAGGCGAGCGATCCGAGCAAGGGT TCTCTGGGACAGCT
F5	GC TGGAT TGGCCCGAGCAAGGCT TGGAAAATGCAARTT	GC TGGAT TGGCCCGAGCAAGGCT TGGAAAATGCAARTT	GC TGGAT TGGCCCGAGCAAGGCT TGGAAAATGCAARTT
B1	CTATGATCCGAATTAAGCCAACTGACACCGGGGCGAGGGGGGDTTCCCCAGCTAA	CTATGATCCGAATTAAGCCAACTGACACCGGGGCGAGGGGGGDTTCCCCAGCTAA	CTATGATCCGAATTAAGCCAACTGACACCGGGGCGAGGGGGGDTTCCCCAGCTAA
B2	CCGAATGCGTTCGGGATCC TACCCC	CCGAATGCGTTCGGGATCC TACCCC	CCGAATGCGTTCGGGATCC TACCCC
B3	CGGATC	CGGATC	CGGATC
B6	AGATCCCAAGCATCTCTCTCC TAAGCTAAGTGT TTAGCAA	AGATCCCAAGCATCTCTCTCC TAAGCTAAGTGT TTAGCAA	AGATCCCAAGCATCTCTCTCC TAAGCTAAGTGT TTAGCAA

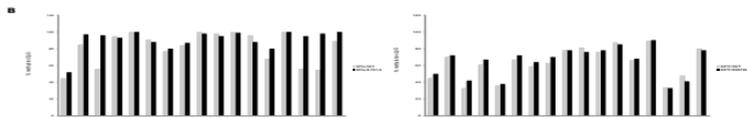


Figure S1

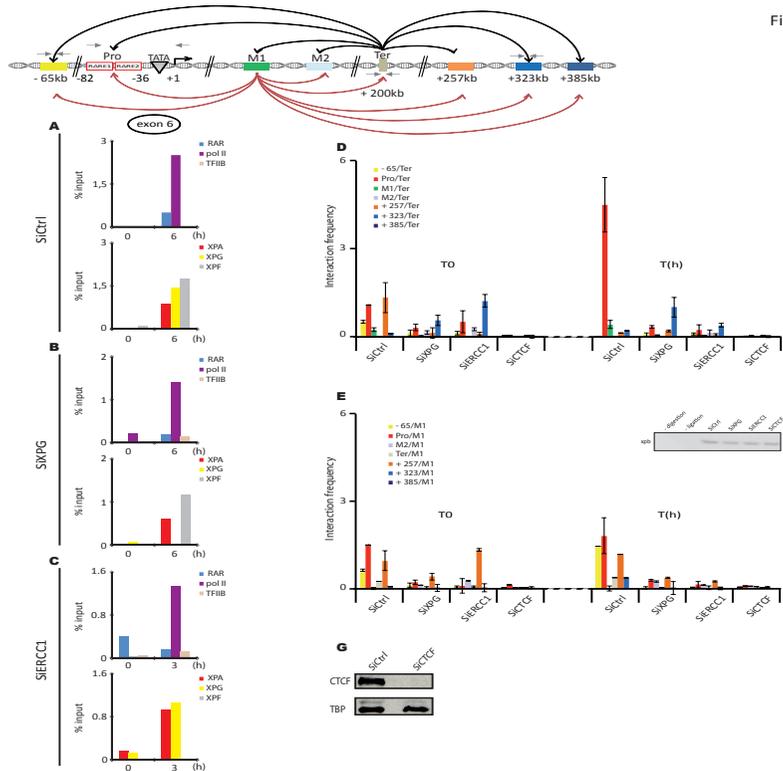
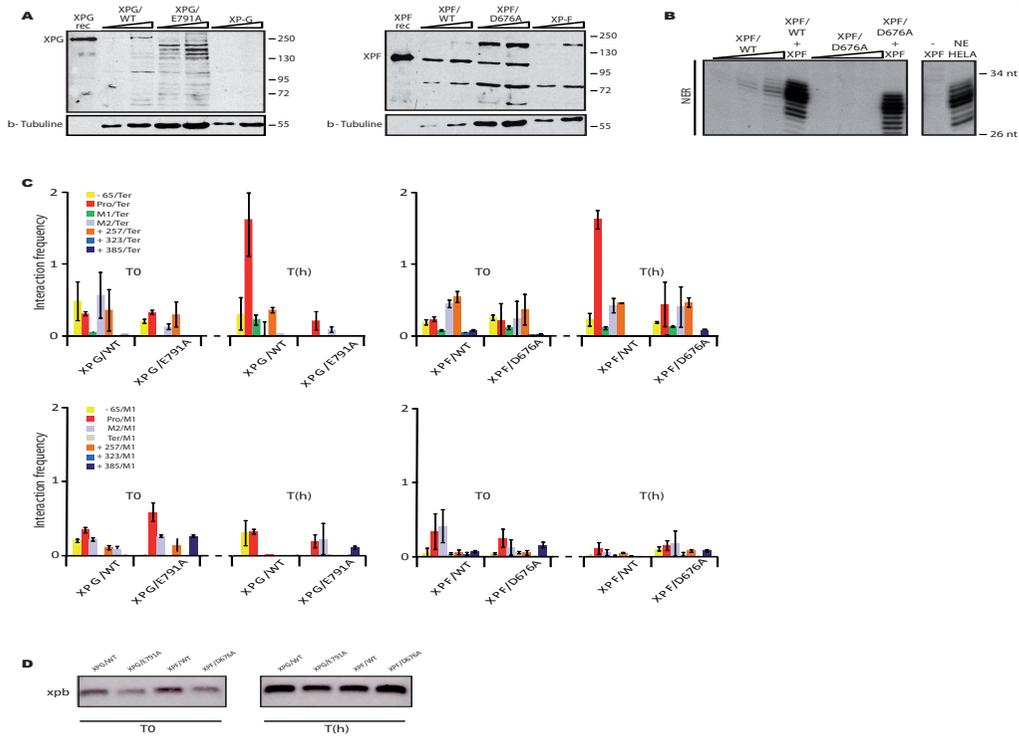
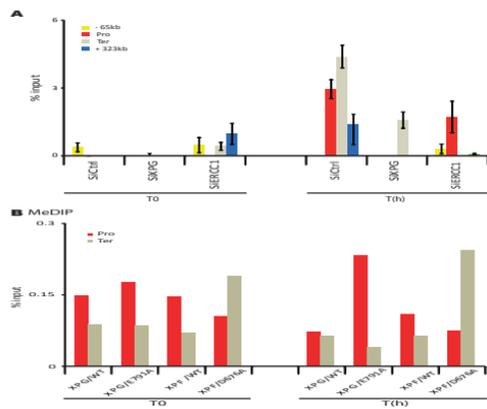


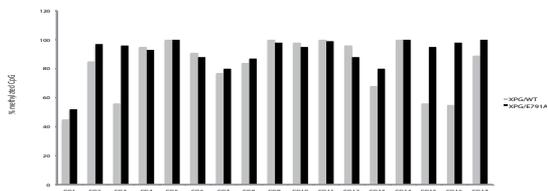
Figure S2





A

Primers	XPG/WT	XPG/E791A
F1	CGAGCTGTTTGAGGACTGGGATGCCGAGAACGCG	CGAGCTGTTTGAGGACTGGGATGCCGAGAACGCG
F2	CGTCGGGGTAGGATCCGGAACGCATTCGGAAGGCTTTTGC AAGCA	CGTCGGGGTAGGATCCGGAACGCATTCGGAAGGCTTTTGC AAGCA
F3	CGCCCCGGCTGGA	CGCCCCGGCTGGA
F7	CGAGAACGCGAGCGATCCGAGCAGGGTTTGCTGGGCACCGT	CGAGAACGCGAGCGATCCGAGCAGGGTTTGCTGGGCACCGT
F8	GCTGGATTGGCCGAGCAAGCCTGGAAAAATGACAATTG	GCTGGATTGGCCGAGCAAGCCTGGAAAAATGACAATTG
R1	CATGGATCCGAACTAAGCCAAGTGAACCGGGCAGGGGGTTACCCAGCTAA	CGGCCAATCCAGCCAGGGGACGGGGGGTTCCCA
R2	CCGAATGCGTTCGGATCCTACCCC	CCGAATGCGTTCGGATCCTACCCC
R3	CGGATC	CGGATC
R6	AAGATCCCAACGATTCTCCTTCCTAAGCTAAATGCTTAGCAA	AAGATCCCAACGATTCTCCTTCCTAAGCTAAATGCTTAGCAA



B

Primers	XPF/WT	XPF/D676A
F1	CGAGCTGTTTGAGGACTGGGATGCCGAGAACGCG	CGAGCTGTTTGAGGACTGGGATGCCGAGAACGCG
F2	CGTCGGGGTAGGATCCGGAACGCATTCGGAAGGCTTTTGC AAGCA	CGTCGGGGTAGGATCCGGAACGCATTCGGAAGGCTTTTGC AAGCA
F3	CGCCCCGGCTGGA	CGCCCCGGCTGGA
F7	CGAGAACGCGAGCGATCCGAGCAGGGTTTGCTGGGCACCGT	CGAGAACGCGAGCGATCCGAGCAGGGTTTGCTGGGCACCGT
F8	GCTGGATTGGCCGAGCAAGCCTGGAAAAATGACAATTG	GCTGGATTGGCCGAGCAAGCCTGGAAAAATGACAATTG
R1	CATGGATCCGAACTAAGCCAAGTGAACCGGGCAGGGGGTTACCCAGCTAA	CATGGATCCGAACTAAGCCAAGTGAACCGGGCAGGGGGTTACCCAGCTAA
R2	CCGAATGCGTTCGGATCCTACCCC	CCGAATGCGTTCGGATCCTACCCC
R3	CGGATC	CGGATC
R6	AAGATCCCAACGATTCTCCTTCCTAAGCTAAATGCTTAGCAA	AAGATCCCAACGATTCTCCTTCCTAAGCTAAATGCTTAGCAA

