DOI: 10.1002/cbic.200800397

Novel Diazirine-Containing DNA Photoaffinity Probes for the Investigation of DNA-Protein-Interactions

Malte Winnacker, Sascha Breeger, Ralf Strasser, and Thomas Carell*[a]

An investigation of the precise interactions between damaged DNA and DNA repair enzymes is required in order to understand the lesion recognition step, which is one of the most fundamental processes in DNA repair. Most recently, photoaffinity labeling approaches have enabled the analysis of even transient protein-DNA interactions. Here we report the synthesis and evaluation of oligonucleotides that contain two photoaffinity "catcher moieties" next to incorporated DNA lesions. With these DNA constructs it is possible to analyze the interactions between DNA lesions and the appropriate repair enzymes. The probes labeled the repair protein efficiently enough to enable subsequent protein analysis by mass spectrometry.

Introduction

Genome maintenance is of fundamental importance for all organisms on Earth.[1] In order to remove DNA lesions, which are produced in significant amounts every day, from the genome, all organisms have evolved a set of DNA repair pathways including base excision repair (BER)[2] and nucleotide excision repair (NER).[3] BER recognizes structurally altered nucleobases.[4] BER enzymes cleave the glycosidic bond between the base heterocycle and the deoxyribose as the first step of a complex base replacement reaction.[5] NER is primarily responsible for the removal of bulky adducts from DNA. Bulky adduct lesions are formed upon UV irradiation of cells or upon reaction of the nucleobases with external carcinogens including metabolically activated aromatic amines or polycyclic aromatic molecules such as benz[a]pyrene. Currently the mechanism through which these repair enzymes achieve the recognition of the individual lesions within a largely undamaged genome is not well understood. In particular for the NER process, the lesion recognition step remains complex and enigmatic, and many interactions between repair factors and the damaged oligonucleotide substrate are of a transient nature.[6]

Here we report the development of photoaffinity probes (Figure 1) composed of DNA duplexes that contain an incorporated, site-specific DNA lesion (D) next to two photocrosslinkers (arrows in Figure 1). A biotin tag (Bio) and a fluorescein tag (Fl) were also introduced, to allow for analysis of the binding event and isolation of the protein responsible for lesion binding.[7] As repair enzymes are known to bind to weakened duplex structures and nucleobases, one question that we needed to address was: are the repair proteins specifically interacting with the lesion in these probes or are they also binding to the photoaffinity labels?

Lesion recognition, particularly by NER proteins, frequently requires reduced stability of the DNA around the lesion.[8] We designed novel photocrosslinker nucleotides that induce neither a destabilization nor a structural change of the duplex. In this way, we hoped to circumvent any unwanted binding of repair factors to the photocrosslinkers and instead keep the specific interaction with the damaged site. For this purpose we used 3-fluoromethyl-substituted diazirines, moieties that are known to be broadly applicable crosslinking species, as photocrosslinking units and attached them to the 5-position of the pyrimidine bases (see below).[9–11] When irradiated with near UV-light, these groups form reactive carbenes that are able to insert into C–H bonds of nearby molecules. Many other photocrosslinking species for the investigation of DNA-protein interactions have been described in literature in recent years, for example, azido- or halogen-substituted nucleobases or benzophenone derivatives.[12] Due to their chemical stability and comparatively straightforward synthesis, diazirines have also been used extensively in crosslinking experiments.[9–11]

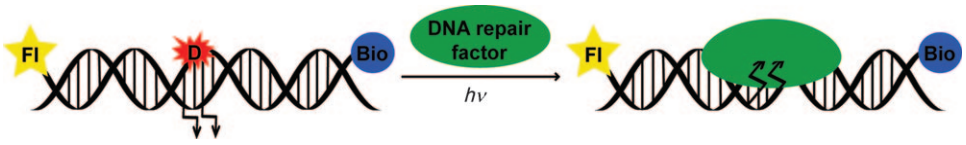


Figure 1. Schematic description of the photoaffinity probes designed for the investigation of interactions between damaged DNA and DNA repair factors by using photocrosslinking. D = DNA lesion, Fl = Fluorescein label, Bio = Biotin label. The arrows symbolize nucleotides that are modified with a photocrosslinking unit.

[a] *Dipl.-Chem. M. Winnacker, Dr. S. Breeger, Dipl.-Chem. R. Strasser, Prof. Dr. T. Carell*
Ludwig-Maximilians University Munich
Center for Integrated Protein Science (CiPS^M)
Department for Chemistry and Biochemistry
Butenandtstrasse 5-13, Haus F, 81377 Munich (Germany)
Fax: (+49) 89-2180-77756
E-mail: thomas.carell@cup.uni-muenchen.de

As a DNA lesion we initially incorporated a cyclobutane-pyrimidine-dimer (CPD), which is formed upon UV irradiation of DNA and is a typical substrate for DNA photolyase repair enzymes.[13] Then we investigated recognition of the major oxidative guanine lesion (8-oxo-dG), which is a substrate for BER.[14] The repair factors formamidopyrimidine glycosylase (Fpg/MutM) from *Lactococcus lactis* and the DNA photolyase from *Escherichia coli* were studied either in pure form or from cell extracts. We also investigated the yeast Rad14 protein as a representative member of NER machinery.

Results

Synthesis of the photoactive DNA probes

Our photoaffinity probes were designed to contain the DNA lesion in one strand and all the additionally needed modifications—the photocrosslinking nucleosides, the biotin and a fluorescence label—in the counter strand. We placed the crosslinking nucleotides close to the lesion to optimize the chance that these probes would covalently bind to proteins attached to the lesion (Figure 2).

For the synthesis of the DNA probes, we used a commercially available biotin carrier to incorporate the biotin molecule into DNA. A fluorescein phosphoramidite (FAM) was used as the final coupling partner to attach the fluorescent label to the DNA. These modifications and the 8-oxo-dG phosphoramidite coupling were carried out according to the procedure described in the Experimental Section. The synthesis and incorporation of the CPD (TT) lesion analogue were performed as reported elsewhere.[13,15,16] For synthetic simplicity, we used the formacetal-bridged T=T-dimer shown in Figure 2. In all previous studies this compound was a perfect mimic of the natural lesion.[13,16] To incorporate the diazirine crosslinker, we attached the diazirine to a deoxyuridine base through an alkyne spacer. In this way, a modified base was generated that did not disturb duplex geometry or change the stability of the duplex. The corresponding phosphoramidite **1** was synthesized according to Scheme 1. We achieved the synthesis of **1** starting with 5-iodo-desoxyuridine (**2**), which was first TBDMS-protected to give **3**. After Sonogashira coupling of **3** with 4-pentyne-1-ol (**4**), the substituted sugar **5** was obtained.[17] It was then transformed with diazirine iodide **6**, to give the diazirine substituted nucleoside **7**. The diazirine compound **6** was synthesized using an eight-step procedure according to Brunner et al.[18] After deprotection of **7** with TBAF, the unprotected nucleoside **8** was finally transformed into phosphoramidite **1** by reaction with dimethoxytritylchloride to give nucleoside **9**. Conversion of **9** with 2-cyanoethyl-*N,N,N′,N′*-tetraisopropylphosphorodiamidite[19] provided phosphoramidite **1**. Compound **1**, which was ready for solid-phase DNA synthesis, was obtained with an overall yield of about 25%.

The synthesis described above allowed the preparation of the crosslinking phosphoramidite **1** on a 100 milligram scale, which was a sufficient amount for solid phase DNA synthesis. The integrity of the diazirine unit of **1** was confirmed by UV spectroscopy and showed a significant 365 nm absorption, which is characteristic for the presence of the diazirine moiety.[9] Due to the modular synthesis, further variation and adjustment of spacer length and composition was readily possible. Incorporation of the phosphoramidite **1** into DNA oligonucleotides was achieved using standard conditions. Despite the bulk of the crosslinking phosphoramidite, the coupling of compound **1** proceeded with high efficiency after double coupling. Cleavage of the highly modified (with biotin, fluorescein, and photocrosslinkers) oligonucleotides from the solid support was performed using standard procedures. All DNA strands were subsequently purified by high performance liquid chromatography (HPLC). Typically, two HPLC purification steps were needed to reach a purity of >98% (see Figure 3). The integrity of the prepared DNA strands was confirmed by MALDI-TOF or ESI measurements and in all cases these measurements

Probe 1 3′- TTACTGGTCA**TT**AGTTCATGAACTGAA- 5′
5′-**F1**-AATGACCAG***T***AA***T***CAAGTACTTGACTT-**Bio**-3′

Probe 2 3′- TTACTGGTCA**G**AGTTCATGAACTCGA- 5′
5′-**F1**-AATGACCAG***T***C***T***CAAGTACTTGAGCT-**Bio**-3′

Probe 3 3′- TTACTGGTCAGAGTTCATGAACTCGA- 5′
5′-**F1**-AATGACCAG***T***C***T***CAAGTACTTGAGCT-**Bio**-3′

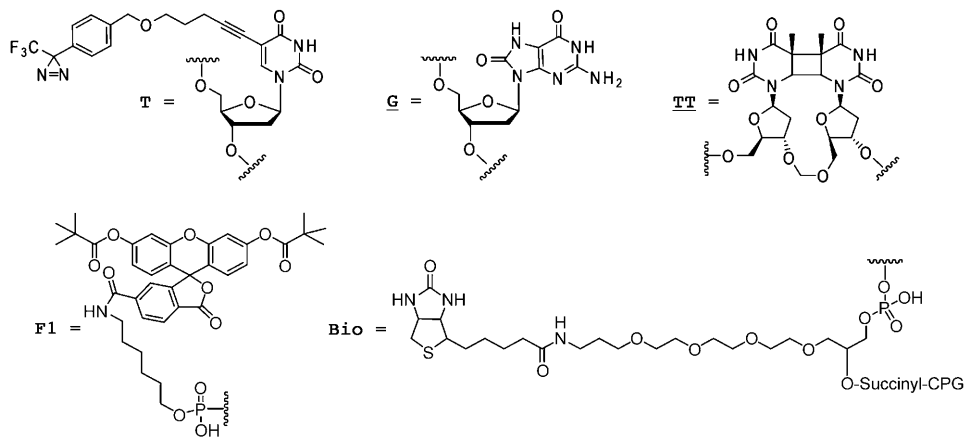
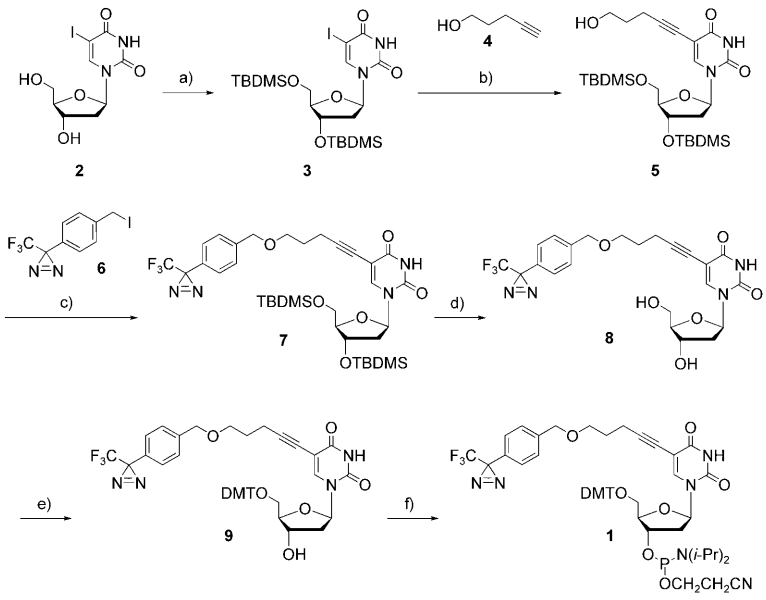


Figure 2. Synthesized DNA photoaffinity probes 1–3 and incorporated modifications. ***T*** = photocrosslinker, G = 8-oxo-dG lesion, TT = cyclobutane-pyrimidine-dimer (CPD) lesion, *Fl* = fluorescein tag (before deprotection), *Bio* = biotin tag (before deprotection).



Scheme 1. a) TBDMSCl, imidazole, DMF, 93%; b) 4-pentyne-1-ol (4), $Pd(PPh_3)_2Cl_2$, CuI, DMF, 75%; c) 6, NaH, THF, 76%; d) TBAF, THF, 95%; e) DMTCl, pyridine, 60%; f) P-($Ni Pr_2$)$_2$(OCH_2CH_2CN), diisopropylammonium tetrazolate, CH_2Cl_2, 84%.

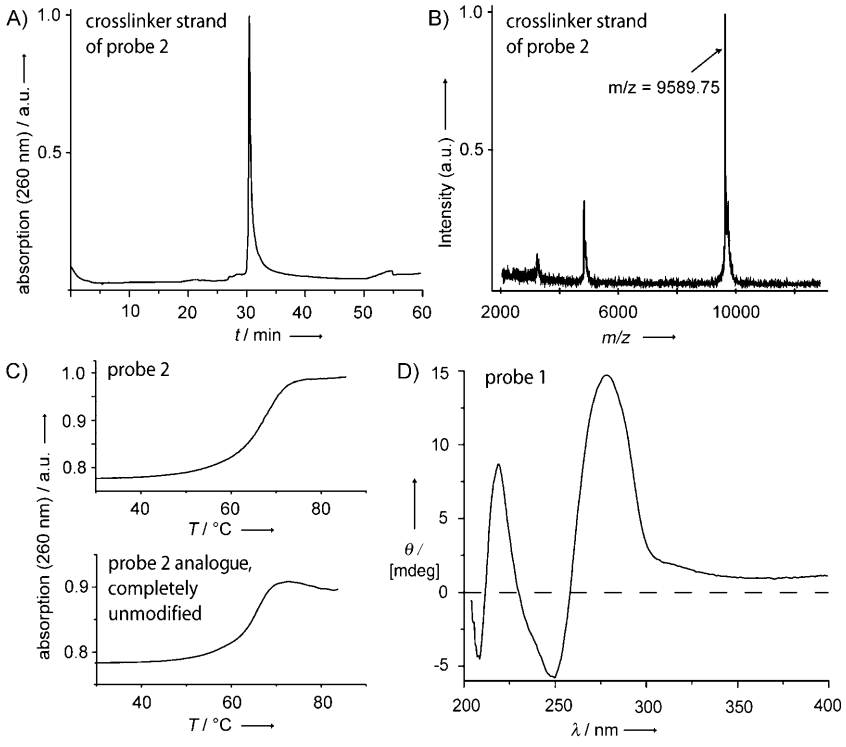


Figure 3. A) HPL-chromatogram of the purified crosslinker strand of DNA probe 2. B) Corresponding MALDI-TOF mass spectrum. C) Above: Melting curve of DNA probe 2 (conditions: 3 μM DNA, 150 mM NaCl, 10 mM Tris-HCl); Below: Melting curve of a DNA probe 2 analogue, completely unmodified (conditions: same as part B). D) CD spectrum of DNA probe 1.

showed the expected correct molecular weight. Further analysis was performed by UV spectroscopy of the DNA strands. In all cases, the presence of the 365 nm absorption band, which is characteristic for the intact trifluoromethyldiazirine moiety, was observed in agreement with a report from Marx et al.[10] Figure 3A shows the HPL-chromatogram of the purified crosslinker-containing strand of DNA probe 2; figure 3B shows the corresponding MALDI-TOF spectrum. The data show the correct molecular weight of the strand at $m/z = 9589.75$. In summary, all three modifications—the biotin and fluorescein labels and the crosslinking phosphoramidites—were successfully incorporated into the DNA single strand.

The photocrosslinker-containing oligonucleotides were subsequently hybridized in a 1:1 ratio (see Experimental Section) with the appropriate counter strand that contained either the 8-oxo-dG or the T=T lesion. To ensure that a double strand was formed despite the many modifications present in the duplex, UV melting curves of the DNA probes were measured. One of the obtained curves is depicted in Figure 3C (above, for probe 2). The data show perfect melting behavior. No hysteresis was observed. In order to answer the question of how much the crosslinking units influence the duplex stability, we measured the melting curve of a totally unmodified duplex (Figure 3C, below) in comparison to our DNA probes. The data show that the melting point is only marginally reduced (probe 2: 66 °C, corresponding completely unmodified probe: 65 °C). This is a very important finding as it proves that the crosslinking nucleotides do not disturb the helical structure. Figure 3D shows a circular dichroism (CD) spectrum of DNA probe 1.

CPD-lesion recognition by *E. coli* DNA photolyase

We started our photoaffinity labeling experiments with DNA probe 1, which contained the surrogate of the UV-induced lesion (CPD) and the purified *E. coli* CPD photolyase protein. This protein efficiently repairs this lesion in a light dependent reaction.[13] First, we incubated 50 pmol of DNA probe 1 (1 μM) with an increasing amount of the CPD photolyase protein (2, 5, or 10 μg) in a Tris-HCl buffer for 20 minutes at 0 °C and then irradiated the sample with near UV-light (365 nm) at 0 °C for 30 minutes. At this wavelength the diazirine is cleaved, yielding a reactive carbene, which is thought to insert into nearby C–H bonds of the protein. The solution was concentrated, SDS

loading buffer was added, and the resulting mixture was heated for five minutes at 95 °C to denature all proteins. This solution was finally loaded on a 12% SDS gel for gel electrophoretic separation. The result of the experiment is depicted in Figure 4A. The figure shows the fluorescence image of the

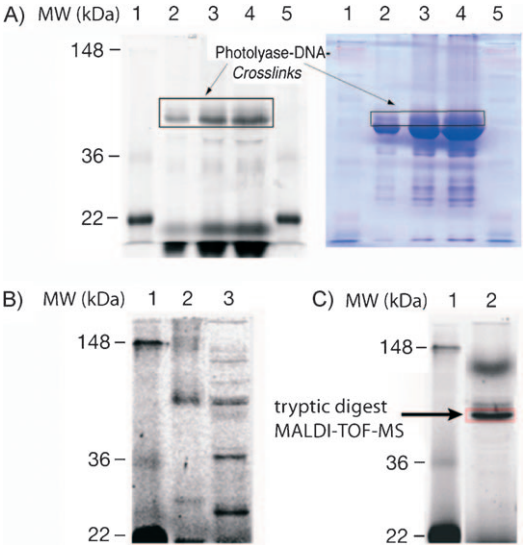


Figure 4. A) Fluorescence image of a gel (left) and Coomassie staining of the same gel (right) obtained after photoaffinity labeling experiments on an analytical scale (50 pmol DNA, 1 μM) with CPD-containing DNA probe 1 and *E. coli* CPD photolyase protein (2, 5 and 10 μg, lanes 2, 3, and 4, respectively). Lanes 1 and 5 show the molecular weight marker. Comparing the DNA-protein crosslinks (edge) with the non crosslinked DNA at about 10 kDa (left picture) or the crosslinked protein (edge) with the noncrosslinked protein (right picture) shows crosslinking yields of about 10–20%. B) Experiments performed on an analytical scale (50 pmol DNA, 1 μM) with the purified protein (lane 2) and with a cell lysate obtained from overexpressing the photolyase protein in *E. coli* cells (lane 3). Lane 1 shows the molecular weight marker. C) Experiment performed on a preparative scale (1 nmol DNA, 1 μM). The resulting DNA–protein crosslink was excised, the protein was digested with trypsin and analyzed by MALDI-TOF mass spectrometry. Lane 1 shows the molecular weight marker.

resulting SDS-polyacrylamide gel (left) and the Coomassie stained gel (right). Lanes 1 and 5 show the molecular weight marker, and lanes 2, 3 and 4 the crosslinking experiments with the CPD photolyase protein (2, 5, or 10 μg, respectively). It is evident that the DNA probe specifically crosslinked the photolyase protein; this shows that the crosslinking experiment functioned as expected. The detected bands were visible at around 54 kDa, which is in perfect agreement with the expected value of 44 kDa for the photolyase protein plus 10 kDa for the DNA single strand (Figure 4A, left, marked with an edge). Comparison with the bands of the non-crosslinked DNA at about 10 kDa using the image reader program showed crosslinking yields of about 10–20%. In the Coomassie stained gel, the free, non-crosslinked protein is visible at about 50 kDa, which is slightly below the DNA-protein crosslinks. Some bands with a lighter mass probably result from decomposition products. The fluorescence image shows no crosslinking of the decomposition products, which are obviously not able to bind to CPD damaged DNA anymore, suggesting that the selective labeling of the photolyase protein is linked to the activity of the complete protein.

We next investigated the possibility of labeling DNA repair proteins in cell lysates. As repair proteins are typically present only in very low numbers, we expected that gel electrophoresis would not allow us to identify any crosslinking and that later direct coupling to a sensitive mass spectrometer would be required. In order to validate the method, we therefore turned first to cell lysates obtained from *E. coli* cells that were overexpressing the CPD photolyase (for the detailed protocol see the Experimental Section). The *E. coli* cells were harvested, lysed, and the cell extract was obtained after centrifugation. We added probe 1 (50 pmol, 0.5 μM) to the cell extract (100 μg) and irradiated the mixture at 0 °C for 30 min. The solution was concentrated and loaded onto the gel. The result of the experiment is depicted in Figure 4B. (Lane 1: molecular weight marker, lane 2: crosslinking experiment with the purified CPD photolyase protein as well as the experiments shown in Figure 4A as comparison, lane 3: crosslinking experiment with the cell lysate.) The presence of a strong protein band at 54 kDa is likely to be caused by the same DNA-protein crosslinking already detected with the purified protein.

In order to prove that indeed the correct protein was obtained from the cell lysate, we repeated the experiment on a preparative scale using 1 nmol (1 μM) DNA and 1 mg of *E. coli* extract (Figure 4C, lane 2). After incubation and irradiation of the resulting mixture, we added streptavidin-coated magnetic beads. After a one hour incubation at room temperature (slow shaking), the beads were removed from the solution, washed, and heated for five minutes at 95 °C in SDS loading buffer (see the Experimental Section) to cleave the biotin-streptavidin interaction. Thus, the biotin-modified DNA probes were released from the streptavidin-coated beads. The resulting solution was again analyzed by SDS gel electrophoresis. The crosslinked protein band was strong enough that we were able to excise it. The crosslink was extracted from the excised gel band according to a protocol described in the Experimental Section. To this solution we added trypsin to achieve a full tryptic digest of the protein-DNA crosslink. The peptide mixture that we obtained was finally analyzed by MALDI-TOF mass spectrometry (Figure 5). After calibration to two trypsin autoproteolysis fragments ($m/z = 2058.1$ and 2223.7, marked with a star), the peptide data that we obtained (marked with a diamond) were compared with data in the NCBI database using the MASCOT algorithm. The obtained data unequivocally identified the *E. coli* CPD photolyase as the parent protein (score value = 73 for identification number gi|1827916, middle mass aberrance = 0.48 kDa).

8-oxo-dG lesion recognition by the *L. lactis* Fpg (*Ll*Fpg) protein

We next extended our studies to the BER system and questioned whether we could also detect the recognition of the main oxidative DNA lesion 8-oxo-dG by the repair enzyme Fpg/MutM from *L. lactis* with our probes. Fpg is a bifunctional DNA glycosylase that excises oxidized purines such as 8-oxo-

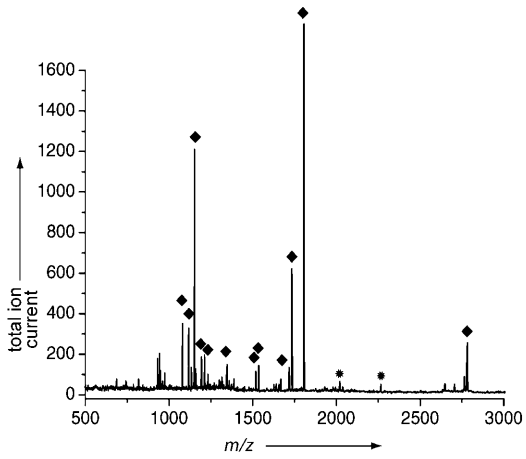


Figure 5. MALDI-TOF-MS spectrum obtained from the peptides resulting from the tryptic digest of the crosslinked protein (Figure 4C, lane 2). A data comparison with the NCBI database using the MASCOT algorithm confirmed the identity of the DNA CPD photolyase protein. The two peaks marked with a star show the trypsin proteolysis fragments, the peaks marked with a diamond show the peaks that resulted from the *E. coli* CPD photolyase digest and were searched against NCBI database.

dG and 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FaPydG) from damaged DNA. While the catalytic mechanism of Fpg has already been extensively studied, the question of how the protein discriminates between damaged and undamaged DNA is under intensive investigation.[14] For these experiments, DNA probe 2 containing the 8-oxo-dG damage was used (Figure 2). As a comparison, DNA probe 3 (Figure 2) was utilized as a duplex containing the same sequence as duplex 2 but without any lesion present. We carried out experiments with pure *Ll*Fpg protein and with an *E. coli* cell lysate in which the *Ll*Fpg protein was overexpressed. In each experiment, the DNA (1 μM) was incubated with the protein (5–10 μg) or the lysate (200–300 μg) in a Tris-HCl buffer for 30 min and then irradiated with near UV-light (365 nm) again for 30 min at 0 °C. After concentration and denaturation, the mixtures were separated by SDS gel electrophoresis and the bands were visualized with a fluorescence imager or by Coomassie staining. The results of these experiment are shown in Figure 6.

Figure 6A shows the fluorescence image of the resulting polyacrylamide gel. Clearly defined bands can be seen under fluorescence light after experiments performed both with the pure Fpg protein (lanes 2 and 3) and with cell lysates of *E. coli* cells overexpressing the Fpg protein (lanes 4 and 5). The group of bands in the area around 40–45 kDa shows the crosslinks between the Fpg protein (33 kDa) and the single crosslinker strand (about 10 kDa). The fact that the crosslinking produces more than just one band can be explained by the existence of different reaction sites. In fact, all visible bands on the gel result from proteins that are crosslinked to fluorescein-labeled oligonucleotides. In the experiment with cell lysate (lanes 4 and 5), we were also able to detect other specific bands that resulted from proteins crosslinked to our probes. Most importantly, however, is the comparison with the results obtained from the control experiments. Here, in lanes 3 and 5 (DNA probe 3), only slight bands are detectable in comparison to the lanes 2 and 4 (DNA probe 2); this clearly shows the higher affinity of the Fpg protein for damaged DNA. This result in itself is very important because the specificity of this repair protein for damaged DNA is still controversially discussed in the literature.[14] Our experiments clearly show selective binding of the Fpg protein (MutM) to lesion-containing DNA. As two additional controls, we performed an experiment in which we did not irradiate (lane 6, only marginal crosslinking), and an experiment with a heat-denaturated cell lysate (lane 7, no crosslinking products). In both cases no crosslinks were obtained. These additional control experiments show that the formation of crosslinks strictly requires UV activation of the photoactive oligonucleotides and the presence of properly folded, and hence active, proteins. Figure 6B shows the Coomassie staining of lanes 1–3. Due to the much lower sensitivity of this visualization method, the Fpg-DNA-crosslink appears, but is very weak (edge). Most importantly, however, is the fact that this method allows us to visualise the additional non-crosslinked proteins. In fact, the non-crosslinked Fpg protein appears at its expected molecular weight at around 33 kDa in the form of a rather thick band, showing that only a small portion of the protein was indeed crosslinked to our photoaffinity DNA. However, comparing the intensity of the crosslinks with the intensity of the free DNA in Figure 6A shows crosslinking yields of 10–20%. More crosslinking was observed when the amount of DNA probe or of the protein was increased (see also the discussion concerning Figure 4).

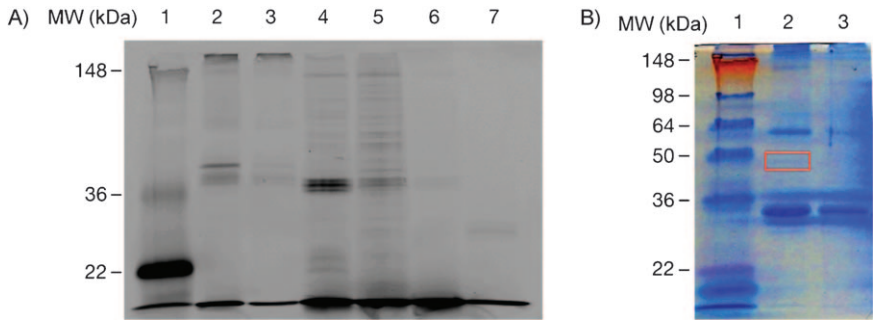


Figure 6. A) Photoaffinity labeling experiments with 8-oxo-dG lesion-containing DNA probe 2 (50 pmol 1 μM, lanes 2 and 4) and nonlesion-containing DNA probe 3 (50 pmol, 1 μM, lanes 3 and 5). Lane 1 shows the molecular weight marker. We used the purified *Ll*Fpg protein (lanes 2 and 3) and *E. coli* cell lysate with overexpressed *Ll*Fpg protein (lanes 4 and 5). Two control experiments (no irradiation and thus no activation of the diazirine moiety, lane 6, and denaturated proteins, lane 7) show the reliability of our model DNA system for crosslinking DNA repair factors. B) Part of the gel (lanes 1–3) after Coomassie staining.

Next, we again wanted to characterize the crosslinked protein. We conducted the experiment at a much higher scale, which again allowed direct isolation of the DNA-protein crosslink through the biotin affinity tag. A subsequent nano-HPLC-ESI-MS/MS analysis of the digested

crosslinked protein indeed confirmed the identity of the *Ll*Fpg protein. In order to prove that Fpg from cell lysate would be bound, we isolated the crosslinks obtained after incubation of the DNA probe with the cell lysate. The results of this experiment are shown in Figure 7 A. Lanes 2 and 3 show crosslinking

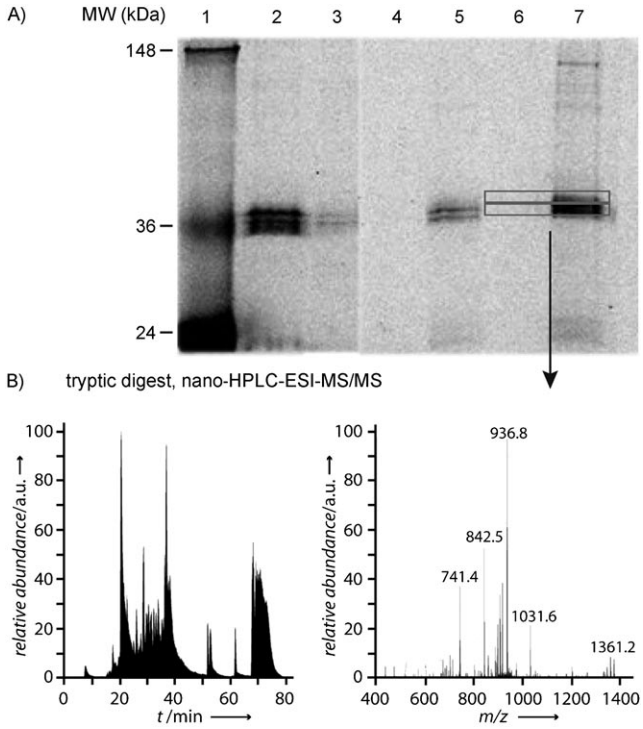


Figure 7. A) Photoaffinity labeling experiments with 8-oxo-dG lesion-containing DNA probe 2 (50 pmol, lane 2) and with nonlesion-containing DNA probes 3 (50 pmol, 1 μM, lane 3) using an *E. coli* cell lysate of cells overexpressing the *Ll*Fpg protein. Probe 2 was also used for a control experiment with a denaturated extract (lane 4). Probe 2 was then used for the large scale experiment (1 nmol DNA) shown in lane 7 and for comparison in lane 5. The experiment shown in lane 6 did not contain any DNA probe. B) Ion current (left) and peptide signals (26 min, right) of the ESI-MS/MS experiment of the tryptic digest of the broad bands shown in Figure 7 A, lane 7 (direct output of BioWorks 3.3.1. software).

experiments on an analytical scale (50 pmol, 1 μM) performed with DNA duplex 2 (damaged) and DNA duplex 3 (undamaged). It is again clearly evident that a thick band is obtained only in the presence of a lesion. We attribute the presence of more than one band either to degradation of the crosslinked protein or crosslinking to different sites. Lane 4 shows a control experiment with damaged duplex 2 and denaturated cell lysate, which shows no crosslinking as expected. Lanes 5, 6, and 7 show a similar experiment, however, on a 20-fold larger scale (1 nmol DNA, 1 μM). This time we used avidin–agarose beads for crosslink isolation. We preferred these particles to the streptavidin-magnetic particles described above because initial experiments pointed to their higher sensitivity and lower tendency for error when using an optimized protocol (see the Experimental Section). In lane 7, the gel electrophoresis data are depicted showing the thick band of crosslinked proteins. Lane 6 shows a negative-control experiment that we performed in the absence of DNA probe (for background protein information). Clearly, no crosslink was formed as expected. The two broad bands in lane 7 were excised. We also cut out a small piece of the gel of lane 6 as the negative control. We next reduced and alkylated any disulfide bridges in the putative proteins and performed a tryptic digest. The obtained set of peptides were analysed using a nano-HPLC-ESI-MS/MS system. The resulting ion current and the peptide signal at a certain time (here: 26 min, selected arbitrarily) are indicated in Figure 7 B, which contains the direct output of the Thermo Scientific BioWorks 3.3.1 software. In order to demonstrate the enrichment process, we took a small amount of the solution used for lane 7 prior to the isolation with the avidin–agarose beads (lane 5). Comparison of lanes 5 and 7 shows successful enrichment.

The obtained data were compared with data from the NCBI database using the SEQUEST algorithm (by using Thermo Scientific BioWorks 3.3.1 software) allowing us to unequivocally identify the protein with the following values: *P* value: 2.2×10^{-14}; score: 86.2; *m/z* 31.173 kDa; max. Xcorr: 4.09; number of peptides: 14.

Lesion recognition by NER proteins

We finally extended our studies to another lesion recognition problem in nature: the lesion detection and verification system of the NER pathway. Previous results from *Maltseva* et al.[20] have already shown that proteins of the NER pathway can be crosslinked to lesion-containing DNA. We focused our initial studies to the Rad14 protein, which is the yeast homologue of the human XPA protein. This protein is thought to be involved in lesion recognition but its lesion specificity is controversially discussed.[21] For these experiments DNA probe 1 was used. We overexpressed a N-terminal truncated version of the Rad14 protein (amino acids 37 to 400) and incubated the pure protein with DNA probe 1. After irradiation for 30 min at 0 °C and concentration, the solution was again loaded onto a SDS gel for separation. The fluorescence image of the gel shows defined bands at around 55 kDa (Figure 8 A, lane 2), which are in perfect agreement with the expected molecular weight (43 kDa for the Rad14 protein plus 10 kDa for the single strand

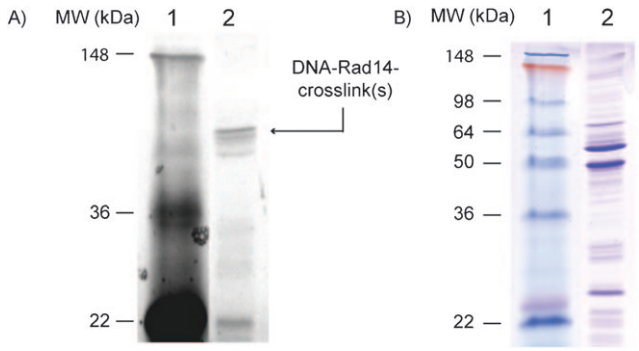


Figure 8. A) Fluorescence image of the gel showing the experiment with DNA probe 1 (50 pmol) and yeast Rad14 protein (5 μg) in lane 2 (lane 1: molecular weight marker). B) Coomassie staining of the same gel.

of DNA probe). The appearance of more than one band can again be explained by assuming either partial photolysis of the protein or different reaction sites. Comparatively high crosslinking yields were also obtained when we used undamaged DNA instead of probe 1 (data not shown).[22] This observation questions the assumption that Rad14 binds specifically to damaged DNA. However, this observation needs further experimental investigation with DNA that contains lesions known to be excellent substrates for the NER systems such as acetylaminofluorene adducts, the pyrimidine(6-4)pyrimidone photoproduct or *cis*-Pt-damaged DNA. Figure 8B shows the Coomassie stained gel. The crosslinks as well as the free protein appear as clearly defined bands between 45 and 64 kDa.

Conclusions

Photoaffinity labeling is a unique method that can be used to investigate the interactions between biological molecules. Some of the most convenient photocrosslinking reagents are 3-trifluoromethyl-phenyldiazirines. These molecules are stable under various chemical conditions and turn into reactive species when irradiated with near UV-light. In this study, we have developed DNA photoaffinity probes containing diazirine-modified photocrosslinking units plus additional defined DNA lesions as molecular "bait" for DNA repair enzymes. Due to the additional presence of a biotin and a fluorescein tag, these complex affinity probes were able to selectively label various repair enzymes, both in pure form and in cell extracts, in the presence of a background of other proteins. In the future this technique will allow us to define precisely which proteins bind to which lesions directly in cell lysates. Of particular importance is the result that repair proteins, which are designed by nature to detect modified bases, are seemingly slow to recognize the photocrosslinking unit itself; this was a major concern at the beginning of the study. However, at this point we can not exclude weak binding to the fluorescein tag for example. The fact that the crosslinking units do not disturb the helix geometry and do not destabilize the helix is seen as the key element for the successful labeling of repair factors.

Experimental Section

General: Solvents and chemicals were obtained from commercial sources and used in *puriss.*, p.*a.* or *purum* grade. Bidistilled water (impedance 18.2Ω) was generated by a Milli-Q Plus device from Millipore. Aqueous and buffered oligonucleotide solutions were concentrated using a SpeedVac Plus SC110A or a SpeedVac SPD 111V device from Savant (Thermo Life Sciences, Egelsbach, Germany). ^{1}H NMR spectra were obtained on a Varian Mercury-200 and Varian INOVA-400. The chemical shifts were referenced to $CHCl_3$ in $CDCl_3$. ^{13}C NMR spectra were obtained on a Varian INOVA-400 and Bruker AMX 600. ESI spectra were obtained on a Thermo Finnigan LT-FT ICR spectrometer.

Oligonucleotides: Oligonucleotides were synthesized on an Amersham Oligoplot 900 system. Unmodified phosphoramidites were acquired from Samchully Pharm. Co., Ltd. (Seoul, South-Korea), 8-oxo-dG phosphoramidite, 3′-biotin-CPG, as well as 5′-fluorescein phosphoramidite were acquired from Glen Research GmbH (Sterling, USA). The diazirine-containing oligonucleotides were synthesized as described below. Phosphoramidites and activators were dissolved in acetonitrile (water content <0.1 ppm) from Riedel de Haen (Honeywell Company, Morristown, USA). For the other reagents, acetonitrile (water content <0.3 ppm) from Roth (Karlsruhe, Germany) was used. The synthesis of the oligonucleotides was accomplished according to standard protocols of the manufacturer. For the modified nucleotides the coupling time was doubled. Unmodified oligonucleotides were ordered from Metabion (Martinsried, Germany). The synthesized oligonucleotides were seperated from the carrier with NH_3(aq.)/EtOH (3:1) or, in the case of 8-oxo-dG-containing oligonucleotides, with NH_3(aq.)/H_2NMe (1:1) with 0.2 μM mercaptoethanol. The oligonucleotide concentrations were measured by using the absorption coefficients of the nucleobases and the modifications with a nanoDrop 3000 device from Peqlab (Erlangen, Germany).

HPLC conditions: Analytical HPLC was performed on a Waters system using 3 μm C18 reverse phase Nucleodur® columns from Machery–Nagel. Eluting buffers were buffer A (NEt_3 (0.1 M) and HOAc (0.1 M) in H_2O) and buffer B (NEt_3 (0.1 M) and HOAc (0.1 M) in H_2O/MeCN (20:80)). The gradient was 0–45% buffer B over 45 min with a flow of 0.5 mL min^{-1}. Preparative HPLC was also performed on a Waters system using Nucleodur® columns (C18ec, 250× 10 mm, 5 μm particle size) C18 reverse phase from Machery-Nagel. The gradient was also 0–45% B in 45 min with a flow of 5 mL min^{-1}. The elution was always monitored at 260 nm.

Melting points of the oligonucleotides: Melting points of the oligonucleotides were measured on a Varian Cary 100 Bio with temperature controller, transport unit, and multicell block. The temperature gradient was 0.5 °C/min or 1.0 °C/min. Five cool down curves (85 °C→12 °C) and five warm up curves (0 °C→85 °C) at 260 nm and 320 nm were recorded per measurement. The measurement of the temperature occurred in a reference cuvette. To avoid formation of a film over the cuvettes at low temperatures, the sample room was fluted by nitrogen during the measurement. Cuvettes from Hellma (Jena, Germany) were used with 4 mm inner diameter and 10 mm optical path. Oligonucleotides were (3 μM) dissolved in NaCl (150 mM) and Tris-HCl (10 mM, pH 7.4). The solutions in the cuvettes were overlayed with dimethylpolysiloxane to avoid vaporization of the samples. The analysis of the measurements was accomplished using Microcal (Northampton, USA). Therefore, the curves at 260 nm were averaged, and the averaged background measurement at 320 nm was subtracted. The generated curve was approximated with a ninth order polynomial, and the zero point of the second derivation demonstrated the corresponding melting point.

Mass spectrometry and bioinformatics: MALDI-TOF mass spectra of the oligonucleotides were recorded on a Bruker Autoflex II mass spectrometer using 3-hydroxypicolinic acid as the matrix. The measurements were arranged in the positive or the negative polarity mode and confirmed the correct mass of the oligonucleotides within MALDI measuring accuracy. The MALDI spectrum of the crosslinker strand of probe 2 (and thus for probe 3) is depicted in Figure 3B (calculated *m/z*: 9589.85; found: 9589.75). Other values were as follows: damaged strand of probe 2, calculated *m/z*: 7997.40, found: 7999.39; undamaged strand of probe 3, calculated *m/z*: 7985.30, found: 7985.80; crosslinker strand of probe 1, calculated *m/z*: 9902.92, found: 9902.74; damaged strand of probe 1: calculated *m/z*: 8234.46; found: 8236.27. For the MALDI-TOF spectra of the peptides, a mixture of 3,5-dihydroxybenzoic acid and α-cyanocinnamic acid was used for the matrix (10 mg/mL each in trifluoroacetic acid (0.1%)/acetonitrile (30:70)). The obtained peptide

data of the MS fingerprint of the tryptic peptides were searched using the MASCOT algorithm and the software Bruker BioTools. The following parameters were set: monoisotopic masses; max. aberrance: 0.1 Da; fixed modification: carbamidomethyl (C); variable modification: oxidation (M); search against the NCBI database.

For the nano-HPLC-ESI experiments, the tryptic digested petides were loaded onto a Dionex (Sunnyvale, USA) C18 Nano Trap Colum (100 μm) and subsequently eluted and separated by a Dionex C18 PepMap 100 (3 μm) column for analysis by tandem MS. This was followed by high resolution MS using a coupled Dionex Ultimate 3000 LC-ThermoFinnegan LTQ-FT MS system. By using the SEQUEST algorithm and the software BioWorks 3.3.1 (Thermo Fisher Scientific), the mass spectrometry data were compared to the NCBI database. The search was limited to only tryptic peptides with two missed cleavage sides and monoisotopic precursor ions with a peptide tolerance of <10 ppm. Filters were set to further refine the search results. The Xcorr versus charge state filter was set to Xcorr values of 1.5, 2.0, and 2.5 for charge states +1, +2 and +3, respectively. The number of different peptides had to be equal to two, and the peptide probability filter was set to <0.001. These filter values are similar to others previously reported for SEQUEST analysis.[23] Important values that were obtained from BioWorks 3.3.1. software for the confirmation of the identity of the *Ll*Fpg protein were the following: P-value: $2.2\cdot10^{-14}$; score: 86.2; mol. weight: 31.173 kDa; max. Xcorr: 4.09; number of peptides: 14 (for protein ID YP_808428.1).

Photoaffinity labeling experiments: For the photoaffinity labeling experiments on an analytical scale, DNA duplexes (50–100 pmol, probe 1 for the CPD-photolyase, probe 2 and probe 3 for the Fpg protein) were placed in a 0.5 mL PCR-reaction cup, and protein (10–20 μg) or cell lysate (200–300 μg) was added. The system was then diluted (to a concentration of 0.5–1 μM) with a special irradiation buffer established in our group. This buffer contains Tris-HCl (10 mM, pH 7.5), $MgCl_2$ (10 mM), KCl (50 mM), EDTA (1 mM), Nonidet P-40 (0.05%), BSA (0.2 μg mL^{-1}) and calf thymus DNA (50 ng μL^{-1} from Sigma). Then the preparation was incubated on ice for 30 min and then irradiated for 30 min at 365 nm with a UV lamp (VL-215L 2×15W–365 nm Tube Power: 60 W from LTF Labortechnik GmbH & Co. KG (Wasserburg, Germany) from a distance of 5 cm. (When working with the photolyase, the pipetting was done under red light due to its light-dependent activity). For the heat denaturation control experiments, the proteins were heated for 5 min at 95 °C (if proteins precipitate due to the heating, adding 1 μL 40% SDS solution is recommended) before incubation with the DNA probes. For the irradiation control experiments, the samples were not irradiated but just incubated on ice. Then the mixtures were filled into Biomax centrifugation filter colums from Millipore (100 μL, exclusion rate 10 kDa) and centrifuged (3500 rpm at 4 °C) down to a volume of 30 μL (30–45 min). SDS loading buffer (5×) was added, and the samples were heated at 95 °C for 5 min. Samples were placed on ice, loaded onto a gel (a 10–15 μL) and analyzed with SDS-PAGE (10% or 12.5%, respectively). The gels were visualized under a LAS3000-Imager from Raytest (Straubenhardt, Germany) and photographed. Then they were stained with Coomassie Blue.

In preparative scale experiments, the DNA probe (1 nmol) and cell lysate (1 mg) were used and diluted with the irradiation buffer described above to 1 mL (1 μM DNA concentration, 1 mg mL^{-1} protein concentration). Then the samples were incubated and irradiated as described above. The following procedure was dependent on the used beads for the biotin-affinity purification.

Isolation of the proteins via streptavidin-magnetic beads: M-270 Dynabeads® (1.0 mg) from New England Biolabs (Schwalbach, Germany) were released from their storage buffer, washed twice with magnetic bead binding buffer (500 μL, contents see protocol of the producer), and the DNA–protein suspensions were added to the beads. The mixture was slowly rotated for 2–3 h at room temperature, the magnetic particles were fixed magnetically, and the unbound supernatant was rejected. The beads were washed with buffer C (1 mL), which contained Tris-HCl (5 mM, pH 7.5), EDTA (0.5 mM), and NaCl (1 M), four times with buffer D (1 mL), which contained buffer C+5% (*w/v*) SDS+1% (*v/v*) Nonidet P-40, and one more time with buffer C (1 mL). Then the beads were resuspended in a 1:1 mixture of SDS loading buffer and ddH_2O (30 μL), and the bound proteins were eluted through heating at 95 °C for 5 min. After magnetically fixing of the beads the supernatant was then removed, cooled on ice and then analyzed by SDS-PAGE. Gel bands were isolated, washed and tryptically digested as described previously.[24]

Isolation of the proteins by using avidin–agarose beads: According to a procedure reported by Sieber et al.,[25] after irradiation and incubation the proteins were precipitated using an equal volume of prechilled acetone. Samples were stored on ice for 20 min and centrifuged at 13000 rpm for 10 min. The supernatant was discarded and the pellet washed two times with prechilled methanol (400 μL) and resuspended by sonication. Subsequently, the pellet was dissolved in PBS (1 mL) containing SDS (0.2%) by sonication and incubated under gentle mixing with avidin–agarose beads (50 μL) (Sigma–Aldrich) for 1 h at room temperature. The beads were washed three times with a solution of SDS (0.2%) in PBS (1 mL), twice with urea (1 mL of 6 M) and three times with PBS (1 mL). SDS loading buffer (50 μL, 2x) was added, and the proteins were released for preparative SDS-PAGE with a 5 min incubation at 95 °C. Gel bands were isolated, washed and digested with trypsin as described previously.[24]

5-(5-Hydroxy-pent-1-inyl)-5′,3′-O-*tert*-butyldimethylsilyl-2′-desoxyuridine (5): *tert*-Butyldimethylsilyl (TBDMS)-protected 5-iododesoxyuridine (**3**, 1.000 g, 1.72 mmol), obtained by standard TBDMS-protection of 5-iododesoxyuridine (**2**), was dissolved in DMF (1 mL) and the solution was degassed three times. Afterwards $PdCl_2(PPh_3)_2$ (241 mg, 0.34 mmol, 0.20 equiv), Hünig's base (597 μL, 3.44 mmol, 2.00 equiv), 4-pentyne-1-ol (4, 479 μL, 5.16 mmol, 3.00 equiv) and copper(I) iodide (130 mg, 0.69 mmol, 0.40 equiv) were added. The mixture was stirred overnight in the dark. Then it was diluted with ether (40 mL), washed once with saturated sodium bicarbonate solution and twice with saturated sodium chloride solution. After drying with sodium sulfate, the solvent was removed under reduced pressure. The brown product was purified through column chromatography (silica gel −60, dichloromethane/methanol (100: 1→50:1)). **5** was obtained as a colorless oil (540 mg, 75%). $R_f=0.37$ (chloroform/methanol 10:1); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta=9.05$ (s, 1H, N-H), 7.92 (s, 1H, C6-H), 6.28 (dd, 1H, $^3J=5.9$ Hz, $^3J=7.4$ Hz, C1′-H), 4.39 (td, 1H, $^3J=2.6$ Hz, $^3J=5.6$ Hz, C3′-H), 3.95 (m, 1H, C4′-H), 3.89 (dd, 1H, $^3J=2.4$ Hz, $^2J=11.4$ Hz, C5′-H_a), 3.76 (m, 3H, H_2C-OH, C5′-H_b), 2.49 (t, 2H, $^3J=6.9$ Hz, H_2C-C≡C), 2.29 (ddd, 1H, $^3J=2.8$ Hz, $^3J=5.9$ Hz, $^2J=13.1$ Hz, C2′-H_a), 2.01 (m, 1H, C2′-H_b), 1.80 (q, 2H, $^3J=6.0$ Hz, CH_2), 0.92 (s, 9H, *t*Bu-TBDMS), 0.88 (s, 9H, *t*Bu-TBDMS), 0.14 (s, 3H, H_3C-TBDMS), 0.12 (s, 3H, H_3C-TBDMS), 0.07 (s, 3H, H_3C-TBDMS), 0.06 ppm (s, 3H, H_3C-TBDMS); ^{13}C NMR(150 MHz, $CDCl_3$): $\delta=162.1$, 149.1, 141.5, 100.5, 94.6, 88.3, 85.6, 72.2, 62.9, 61.4, 41.9 (2C), 31.0, 26.0 (3C), 25.7 (3C), 18.4, 18.0, −4.7, −4.9, −5.4, −5.6 ppm; IR (ATR): $\tilde{\nu}=3416$ w, 3181 w, 3066 w, 2951 m, 2928 m, 2856 m, 1681 vs, 1627 m,

1575 w, 1461 m, 1404 w, 1360 w, 1322 w, 1278 m, 1251 s, 1190 w, 1099 m, 1065 s, 1028 s, 1005 m, 966 m, 937 w, 919 w, 885 w, 829 vs, 812 s, 775 vs, 670 s; HRMS (ESI^+): calcd *m/z* for $C_{26}H_{48}N_2O_6Si_2$ $[M+H]^+$: 539.2973; found: 539.2968.

1-[4-(*tert*-Butyl-dimethyl-silanyloxy)-5-(*tert*-butyl-dimethyl-silanyl-oxymethyl)-tetrahydro-furan-2-yl]-5-{5-[4-(3-trifluoromethyl-3*H*-diazirin-3-yl)-benzyloxy]-pent-1-inyl}-1*H*-pyrimidin-2,4-dione (7): NaH (50 mg, 60% suspension in oil, 1.25 mmol, 2.50 equiv) was added to a solution of **5** (270 mg, 0.50 mmol) in absolute THF (1 mL). After evolution of gas (20 min) **6** (204 mg, 0.60 mmol, 1.20 equiv) was added, and the solution was stirred under an inert atmosphere at room temperature for 3.5 h. The reaction was quenched by the addition of saturated ammonium chloride solution (1 mL), diluted with ether (20 mL) and washed with saturated NaCl solution. After drying with sodium sulfate and concentration in vacuo, the crude product was purified by using column chromatography (silica gel-60, dichloromethane/methanol 75:1). **7** was obtained as a colorless foam (280 mg, 76%). $R_f = 0.17$ (chloroform/methanol 30:1); ^{1}H NMR (600 MHz, $CDCl_3$): $\delta = 8.20$ (s, 1H, C6-H), 7.36 (d, 2H, $^3J = 8.2$ Hz, $H_{aromatic}$), 7.16 (d, 2H, $^3J = 8.2$ Hz, $H_{aromatic}$), 6.29 (dd, 1H, $^3J = 6.0$ Hz, $^3J = 7.5$ Hz, C1′-H), 4.52 (s, 2H, H_2C-Bn), 4.41 (m, 1H, C3′-H), 3.96 (d, 1H, $^3J = 2.2$ Hz, C4′-H), 3.89 (dd, 1H, $^3J = 2.2$ Hz, $^2J = 11.4$ Hz, C5′-H_a), 3.76 (dd, 1H, $^3J = 2.2$ Hz, $^2J = 11.4$ Hz, C5′-H_b), 3.59 (t, 2H, $^3J = 6.1$ Hz, H_2C-OBn), 2.50 (t, 2H, $^3J = 7.1$ Hz, H_2C-C≡C), 2.30 (ddd, 1H, $^3J = 2.6$ Hz, $^3J = 5.8$ Hz, $^2J = 13.1$ Hz, C2′-H_a), 2.12 (ddd, 1H, $^3J = 6.0$ Hz, $^3J = 7.5$ Hz, $^2J = 13.3$ Hz, C2′-H_b), 1.87 (dt, 2H, $^3J = 6.1$ Hz, $^3J = 7.1$ Hz, CH_2), 0.92 (s, 9H, *t*Bu-TBDMS), 0.89 (s, 9H, *t*Bu-TBDMS), 0.13 (s, 3H, H_3C-TBDMS), 0.12 (s, 3H, H_3C-TBDMS), 0.08 (s, 3H, H_3C-TBDMS), 0.07 ppm (s, 3H, H_3C-TBDMS); ^{13}C NMR (150 MHz, $CDCl_3$): $\delta = 161.4$, 149.0, 141.6, 140.4, 128.2, 127.7 (2C), 126.1 (2C), 122.5 (q, $^1J_{C\text{-}F} = 274$ Hz), 100.6, 94.2, 88.3, 85.6, 72.4, 72.1, 69.1, 63.0, 41.9 (2C), 28.6, 26.0 (3C), 25.7 (3C), 18.4, 18.0, 16.5, −4.6, −4.8, −5.4, −5.5 ppm; IR (ATR) : $\tilde{\nu} = 3182$ w, 3065 w, 2953 m, 2929 m, 2885 w, 2857 m, 2285 m, 1723 s, 1682 vs, 1625 m, 1519 w, 1461 m, 1404 w, 1361 w, 1344 m, 1323 w, 1278 m, 1253 s, 1232 m, 1182 s, 1153 s, 1103 s, 1067 m, 1028 m, 1005 w, 967 m, 937 s, 918 w, 885 m, 836 vs, 810 m, 776 vs, 733 m, 670 cm^{-1} s; HRMS (ESI^+): calcd *m/z* for $C_{35}H_{51}F_3N_4O_6Si_2Na$ $[M+Na]^+$: 759.3197; found: 759.3206.

1-(4-Hydroxy-5-hydroxymethyl-tetrahydro-furan-2-yl)-5-{5-[4-(3-trifluoromethyl-3*H*-diazirin-3-yl)-benzyloxy]-pent-1-inyl}-1*H*-pyrimidin-2,4-dione (8): **7** (530 mg, 0.72 mmol) was dissolved in tetrabutylammoniumfluoride (1 M) in THF (20 mL) and stirred at room temperature for 3 h (TLC control). The solvent was removed in vacuo. The oily residue was dissolved in ether (30 mL) and successively washed with saturated ammonium chloride solution, twice with water and once with saturated NaCl solution. After drying with sodium sulfate, the solvent was removed under reduced pressure. **8** was obtained as a slightly orange oil (315 mg, 95%). $R_f = 0.21$ (chloroform/methanol 10:1); ^{1}H NMR (600 MHz, $CDCl_3$): $\delta = 7.85$ (s, 1H, C6-H), 7.37 (d, 2H, $^3J = 8.2$ Hz, $H_{aromatic}$), 7.17 (d, 2H, $^3J = 8.1$ Hz, $H_{aromatic}$), 6.18 (t, 1H, $^3J = 6.6$ Hz, C1′-H), 4.57 (m, 1H, C3′-H), 4.52 (s, 2H, H_2C-Bn), 4.02 (m, 1H, C4′-H), 3.92 (dd, 1H, $^3J = 2.9$ Hz, $^2J = 11.7$ Hz, C5′-H_a), 3.82 (dd, 1H, $^3J = 2.8$ Hz, $^2J = 11.7$ Hz, C5′-H_b), 3.60 (t, 2H, $^3J = 6.1$ Hz, H_2C-OBn), 2.51 (t, 2H, $^3J = 7.1$ Hz, H_2C-C≡C), 2.35 (m, 2H, C2′-H), 1.87 (qi, 2H, $^3J = 6.6$ Hz, H_2C) ppm; ^{13}C NMR (150 MHz, $CDCl_3$): $\delta = 161.5$, 149.3, 142.9, 140.5, 128.5, 128.0 (2C), 126.8 (2C), 122.4 (q, $^1J_{C\text{-}F} = 273$ Hz), 101.0, 94.5, 87.2, 87.0, 72.3, 71.6, 69.2, 62.4, 40.9, 20.8, 18.2, 16.6, 14.2 ppm; IR (ATR) : $\tilde{\nu} = 3403$ br, 3058 w, 2961 m, 2928 w, 2859 w, 2090 w, 1681 vs, 1625 m, 1519 m, 1461 m, 1344 m, 1320 w, 1278 m, 1259 s, 1231 m, 1180 m, 1149 s, 1078 vs, 1050 vs, 956 w, 937 s, 869 w, 799 vs, 733 m, 662 w cm^{-1}; HRMS (ESI^-): calcd *m/z* for $C_{23}H_{23}{}^{35}ClF_3N_4O_6$ $[M+Cl]^-$: 543.1258; found: 543.1264.

1-[4-(Hydroxy-5-(4′,4′′-dimethoxytrityl)oxymethyl-tetrahydro-furan-2-yl]-5-{5-[4-(3-trifluoromethyl-3*H*-diazirin-3-yl)-benzyl-oxy]-pent-1-inyl}-1*H*-pyrimidin-2,4-dione (9): **8** (315 mg, 0.61 mmol) was disolved in dry pyridine (2.0 mL) and molecular sieves (4 Å) were added. After stirring for one hour at room temperature 4,4′-dimethoxytritylchloride (252 mg, 0.73 mmol, 1.20 equiv) was added, and the solution was stirred for two additional hours. Then the reaction was stopped by adding methanol (300 μL). The solvent was removed in vacuo and the crude product was purified by column chromatography (dichloromethane/methanol/pyridine 50:1:0.1). **9** was obtained as a white compact foam (248 mg, 60%). $R_f = 0.46$ (chloroform/methanol 10:1); ^{1}H NMR (600 MHz, $CDCl_3$): $\delta = 8.94$ (s, 1H, HN), 8.01 (s, 1H, C6-H), 7.43 (d, 2H, $^3J = 7.4$ Hz, $H_{aromatic}$), 7.34–7.32 (m, 4H, $H_{aromatic}$-DMT), 7.30–7.28 (m, 4H, $H_{aromatic}$-DMT), 7.20–7.17 (m, 1H, $H_{aromatic}$-DMT), 7.14 (d, 2H, $^3J = 8.1$ Hz, $H_{aromatic}$), 6.83 (d, 4H, $^3J = 8.8$ Hz, $H_{aromatic}$-DMT), 6.32 (dd, 1H, $^3J = 6.0$ Hz, $^3J = 7.5$ Hz, C1′-H), 4.51 (m, 1H, C3′-H), 4.37 (s, 2H, H_2C-Bn), 4.07 (m, 1H, C4′-H), 3.76 (s, 3H, H_3CO-DMT), 3.75 (s, 3H, H_3CO-DMT), 3.41 (dd, 1H, $^3J = 3.0$ Hz, $^2J = 10.7$ Hz, C5′-H_a), 3.37 (t, 2H, $^3J = 6.1$ Hz, H_2C-OBn), 3.32 (dd, 1H, $^3J = 3.5$ Hz, $^2J = 10.7$ Hz, 5′-H_bC), 2.48 (ddd, 1H, $^3J = 2.5$ Hz, $^3J = 5.6$ Hz, $^2J = 13.5$ Hz, 2′-H_aC), 2.29–2.21 (m, 3H, C2′-H_b), C≡C-H_2C), 1.55 (qi, 2H, $^3J = 6.6$ Hz, C≡C-CH_2-H_2C); ^{13}C NMR (150 MHz, $CDCl_3$): $\delta = 161.7$, 158.5, 149.2, 144.4, 141.5, 140.4, 135.4, 129.9 (2C), 128.0 (2C), 127.9 (2C), 127.8, 127.6 (4C), 126.9 (2C), 126.4 (2C), 123.0 (q, $^1J_{C\text{-}F} = 273.1$ Hz), 113.2 (4C), 100.9, 94.4, 86.9, 86.4, 85.5, 72.3, 71.8, 71.0, 69.0, 68.1, 63.4, 55.1 (2C), 41.4, 38.7, 29.6, 28.3, 16.3; IR (ATR) : $\tilde{\nu} = 3467$ w, 3181 w, 3066 w, 2954 m, 2930 m, 2869 w, 2837 w, 2084 w, 1688 vs, 1607 s, 1592 m, 1507 s, 1439 m, 1404 w, 1343 m, 1278 m, 1247 s, 1175 vs, 1150 vs, 1097 s, 1052 m, 1031 s, 1002 w, 988 w, 937 s, 914 w, 872 w, 825 m, 810 m, 791 w, 772 w, 752 m, 726 w, 701 vs, 657 cm^{-1} w; HRMS (ESI^+): calcd *m/z* for $C_{44}H_{41}F_3N_4O_8{}^{23}Na$ $[M+Na]^+$: 833.2774; found: 833.2773.

1-[4-(β-Cyanoethyl-*N,N*-diisopropylaminophosphanoxy)-5-(4′,4′′-dimethoxytrityl)oxymethyl-tetrahydro-furan-2-yl]-5-{5-[4-(3-trifluoromethyl-3*H*-diazirin-3-yl)-benzyloxy]-pent-1-inyl}-1*H*-pyrimidin-2,4-dione (1): **9** (200 mg, 0.25 mmol) was solved in absolute dichloromethane (1.5 mL) and degassed three times. Diisopropylammonium tetrazolate (22 mg, 0.13 mmol, 0.50 equiv) and β-cyanoethyl *N,N,N′,N′*-tetraisopropylphosphordiamidite (119 μL, 0.38 mmol, 1.50 equiv) were added, and the solution was stirred for 4 h at room temperature under argon. The solvent was removed in vacuo and the residue was purified by column chromatography (deactivated silica gel-60, dichloromethane/methanol/pyridine 75:1:0.1). **1** was obtained as a yellowish film (240 mg, 84%). $R_f = 0.42$ (chloroform/methanol 20:1); ^{1}H NMR (600 MHz, $CDCl_3$): $\delta = 8.94$ (*s*, 1H, HN), 8.05 (s, 0.5H, 6C-H), 8.01 (s, 0.5H, 6C-H), 7.43 (d, 2H, $^3J = 7.8$ Hz, $H_{aromatic}$C), 7.34–7.32 (m, 4H, $H_{aromatic}$-DMT), 7.30–7.28 (m, 4H, $H_{aromatic}$-DMT), 7.20–7.17 (m, 1H, $H_{aromatic}$-DMT), 7.14 (d, 2H, $^3J = 7.8$ Hz, $H_{aromatic}$C), 6.83 (d, 4H, $^3J = 7.8$ Hz, $H_{aromatic}$-DMT), 6.32–6.28 (m, 1H, C1′-H), 4.58 (m, 1H, C(3′-H), 4.37 (s, 1H, H_2C-Bn), 4.35 (s, 1H, H_2C-Bn), 4.20–4.14 (m, 1H, HC(4′), 3.76 (s, 6H, 2H_3CO-DMT), 3.66–3.49 (m, 4H, H_2C-O-P, 2HC-N-P), 3.43 (dd, 1H, $^3J = 2.5$ Hz, $^2J = 10.7$ Hz, C5′-H_a), 3.34 (m, 2H, $^3J = 6.1$ Hz, H_2C-OBn), 3.29 (dd, 1H, $^3J = 3.5$ Hz, $^2J = 10.7$ Hz, C5′-H_b), 2.61 (t, 1H, $^3J = 6.3$ Hz, H_2C-CN), 2.58–2.51 (m, 1H, C(2′-H_a), 2.44–2.42 (dt, 1H, $^3J = 6.4$ Hz, $^2J = 11.0$ Hz, H_2C-CN), 2.20 (m, 2H, C≡C-H_2C), 1.52–1.47 (m, 2H, C≡C-CH_2-H_2C), 1.28–1.06 (m, 12H, 4H_3C-CH-N-P); ^{13}C NMR (150 MHz, $CDCl_3$): $\delta = 161.8$, 158.8, 149.3, 144.3, 141.8, 140.7, 135.8, 130.2 (2C), 129.1 (2C), 128.2 (2C), 127.9 (2C), 127.1 (2C), 126.7 (4C),

124.0, 121.9 (*q*, $^{1}J_{C\text{-}F}=273.3$ Hz), 113.5 (4C), 101.2, 94.7, 87.2, 86.3, 85.8, 72.1, 71.3, 68.4, 69.3, 63.3, 58.5, 55.4 (2C), 45.6, 43.5, 40.8, 39.0, 30.6, 28.9 (*q*, $^{2}J_{C\text{-}F}=30.1$ Hz), 28.5, 24.8, 23.2, 20.4, 16.5, 14.3, 11.2 ppm; ^{31}P NMR (81 MHz, $CDCl_3$): $\delta=149.9$, 149.5; IR (ATR): $\tilde{\nu}=$ 3189 w, 3065 w, 2966 m, 2931 m, 2871 w, 1693 vs, 1608 s, 1583 w, 1508 s, 1456 s, 1398 w, 1364 m, 1344 m, 1278 s, 1247 vs, 1177 vs, 1153 s, 1106 m, 1077 m, 1031 s, 975 s, 937 m, 914 w, 896 w, 879 w, 826 m, 809 m, 791 m, 772 w, 754 m, 726 m, 700 m, 641 cm^{-1} w; HRMS (ESI$^-$): calcd *m/z* $C_{53}H_{57}F_3N_6O_9{}^{31}P$ $[M-H]^-=1009.3882$; found: 1009.3916.

Bacterial strains, plasmids, culture conditions: The *E. coli* CPD photolyase (CPDP) gene was amplified by polymerase chain reaction of the genomic DNA (*E. coli K12*) using AccuPrime™ *Pfx* DNA polymerase (Invitrogen) with the appropriate primers (5′-CAC CAT GAC TAC CCA TCT GGT CTG-3′ and 5′-TTA TTT CCC CTT CCG CGC C-3′) for the TOPO reaction. The resulting PCR product was cloned into the vector pENTR™/D-TOPO® (Invitrogen) by the TOPO reaction. After sequencing, the vector was transferred into pDest007[26] through the LR reaction (Gateway cloning kit, Invitrogen, Carlsbad, USA), which provided pExp007-*cpdp*. For expression, the plasmid pExp007-cpd photolyase was transformed into *E. coli* TUNER™ (Novagen). The cells were incubated at 37 °C until A_{600} of 0.6 was reached, then anhydrotetracycline was added (2 nM) and incubation was continued for 4 h at 37 °C. The *Ll*fpg gene was obtained by PCR amplification of the genomic *L. lactis* DNA using AccuPrime™ *Pfx* DNA polymerase (Invitrogen) with the two primers, 5′-AAT GCC AGA GTT ACC AGA AGT TGA AA-3′ and 5′-TCC CTT TTT GCT GAC AGA ATG GGC AA-3′. The resulting *fpg* gene PCR product was cloned into the entry vector pENTRY-IBA10 by using the StarGate® method (IBA, Göttingen, Germany). The resulting vector was sequenced and mixed with the acceptor vector pPSG-IBA3 (IBA); this yielded the desired destination vector pPSG-IBA3-*fpg* which was transformed into *E. coli* BL21 Star™ (DE3) (Invitrogen). Cells were incubated in Luria broth at 37 °C until they reached an A_{600} of 0.8. After induction with isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM) and addition of $ZnCl_2$ (10 μM) incubation was continued for 4 h at 30 °C. The N-terminal truncated version of Rad14 was obtained by amplification of the genomic DNA of *S. cerevisiae* using the following primers: 5′-AAT GGA GGC TAA CAG GAA ATT AGC AAT AG-3′ and 5′-TCC CAA TGT CAA TTT CTT CAG TTT CTA GCC-3′. The resulting gene product was cloned into the Entry Vector pENTRY-IBA10 via the Stargate® method (IBA). After verification of the appropriate sequence the vector was mixed with the Destination vector pPSG-IBA3 (IBA). The resulting vector was transformed into *E. coli* BL21 Star™ (DE3) (Invitrogen). Incubation and induction were performed as described for *Ll*Fpg.

Protein purification of recombinant proteins: All purification steps were carried out at 4 °C and the purification was monitored by SDS-PAGE. For purification of the CPD Photolyase, the cells were resuspended in Buffer E (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA) with protease inhibitor mix (Roche). Cells were lysed in a French press and the cell debris was removed by centrifugation before it was applied to the Strep-Tactin® column (IBA). The purification step was performed according to the manufacturers protocol (IBA). The procedures of the *Ll*Fpg purification followed the once described above. After concentrating *Ll*Fpg was loaded onto a Source S column (GE) equilibrated with buffer F, which contained HEPES/NaOH (10 mM, pH 7.6), NaCl (100 mM), glycerin (5% *v/v*) and β-mercaptoethanol (5 mM). The protein was eluted with a gradient of 5 column volumes with the same buffer containing NaCl (600 mM). Prior to buffer exchange to buffer F the protein was concentrated. The procedure of the purification of Rad14 followed the once described for *Ll*Fpg with the exception that following buffer was used: Tris-HCl (50 mM, pH 7.5), NaCl (100 mM), glycerin (10% *v/v*), and 5 mM β-mercaptoethanol.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (SFB-749 and SFB-646) as well as the Excellence Cluster CiPS^M for financial support. Additional funding from Novartis, Bayer–Schering Pharma and the Fonds of the German Chemical Industry is gratefully acknowledged.

Keywords: chemical proteomics · DNA damage · DNA repair · DNA · photoaffinity labeling

[1] T. Lindahl, *Nature* **1993**, *362*, 709–715.
[2] B. Sun, K. A. Latham, M. L. Dodson, R. S. Lloyd, *J. Biol. Chem.* **1995**, *270*, 19501–19508.
[3] A. Sancar, *Annu. Rev. Biochem.* **1996**, *65*, 43–81.
[4] a) D. O. Zharkov, G. Shoham, A. P. Grollman, *DNA Repair* **2003**, *2*, 839–862; b) W. K. Hansen, M. R. Kelley, *J. Pharmacol. Exp. Ther.* **2000**, *295*, 1–9.
[5] W. Yang, *Cell Research* **2008**, *18*, 184–197.
[6] L. C. J. Gillet, O. D. Schärer, *Chem. Rev.* **2006**, *106*, 253–276.
[7] a) R. Aebersold, M. Mann, *Nature* **2003**, *422*, 198–207; b) M. J. Evans, B. F. Cravatt, *Chem. Rev.* **2006**, *106*, 3279–3301; c) S. A. Sieber, B. F. Cravatt, *Chem. Commun.* **2006**, 2311–2319.
[8] a) U. Camenisch, R. Dip, S. B. Schumacher, B. Schuler, H. Naegeli, *Nat. Struct. Mol. Biol.* **2006**, *13*, 278–284; b) U. Camenisch, R. Dip, M. Vitanescu, H. Naegeli, *DNA Repair* **2007**, *6*, 1819–1828.
[9] R. A. Smith, J. R. Knowles, *J. Am. Chem. Soc.* **1973**, *95*, 5072–5073.
[10] M. Liebmann, F. Di Pascale, A. Marx, *ChemBioChem* **2006**, *7*, 1965–1969.
[11] a) A. Blencowe, W. Hayes, *Soft Matter* **2005**, *1*, 178–205; b) M. Hashimoto, Y. Hatanaka, *Eur. J. Org. Chem.* **2008**, 2513–2523; c) U. K. Shigdel, J. Zhang, C. He, *Angew. Chem.* **2008**, *120*, 96–99; *Angew. Chem. Int. Ed.* **2008**, *47*, 90–93.
[12] a) K. M. Meisenheimer, T. H. Koch, *Crit. Rev. Biochem. Mol. Biol.* **1997**, *32*, 101–140; b) J. Zhang, P. V. Chang, S. J. Lippard, *J. Am. Chem. Soc.* **2004**, *126*, 6536–6537; c) O. I. Lavrik, R. Prasad, R. W. Sobol, J. K. Horton, E. J. Ackerman, S. H. Wilson, *J. Biol. Chem.* **2001**, *276*, 25541–25548.
[13] A. Mees, T. Klar, P. Grau, U. Hennecke, A. P. M. Eher, T. Carell, L. O. Essen, *Science* **2004**, *306*, 1789–1793.
[14] A. Banerjee, W. L. Santos, G. L. Verdine, *Science* **2006**, *311*, 1153–1157.
[15] J. Butenandt, R. Epple, E.-U. Wallenborn, A. P. M. Eker, V. Gramlich, T. Carell, *Chem. Eur. J.* **2000**, *6*, 62–72.
[16] F. Brueckner, U. Hennecke, T. Carell, P. Cramer, *Science* **2007**, *315*, 859–862.
[17] L. A. Agrofoglio, I. Gillaizeau, Y. Saito, *Chem. Rev.* **2003**, *103*, 1875–1916.
[18] J. Brunner, H. Senn, F. M. Richards, *J. Biol. Chem.* **1980**, *255*, 3313–3318.
[19] W. Bannwarth, A. Trzeciak, *Helv. Chim. Acta* **1987**, *70*, 175–186.
[20] E. A. Maltseva, N. I. Rechkunova, L. C. Gillet, I. O. Petruseva, O. D. Schärer, O. I. Lavrik, *Biochim. Biophys. Acta* **2007**, *1770*, 781–789.
[21] S. Prakash, L. Prakash, *Mutat. Res.* **2000**, *451*, 13–24.
[22] S. N. Guzder, P. Sung, L. Prakash, S. Prakash, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 5433–5437.
[23] S. P. Mirza, B. D. Halligan, A. S. Greene, M. Olivier, *Physiol. Genomics* **2007**, *30*, 89–94.
[24] S. A. Sieber, S. Niessen, H. S. Hoover, B. F. Cravatt, *Nat. Chem. Biol.* **2006**, *2*, 274–281.
[25] T. Böttcher, S. A. Sieber, *Angew. Chem.* **2008**, *120*, 4677–4680; *Angew. Chem. Int. Ed.* **2008**, *47*, 4600–4603.
[26] M. Ober, H. Müller, C. Pieck, J. Gierlich, T. Carell, *J. Am. Chem. Soc.* **2005**, *127*, 18143–18149.

Received: June 12, 2008
Published online on November 14, 2008

