

Discovery and Mutagenicity of a Guanidinoformimine Lesion as a new Intermediate of the Oxidative Deoxyguanosine Degradation Pathway

Dimitrios Stathis, Ulrike Lischke, Sandra C. Koch, Christian A. Deiml, and Thomas Carell*

Center for Integrated Protein Science at the Department of Chemistry, Ludwig-Maximilians-Universität, Munich, Butenandtstraße 5-13, D-81377 Munich, Germany

Supporting Information

ABSTRACT: Oxidative degradation of DNA is a major mutagenic process. Reactive oxygen species (ROS) produced in the course of oxidative phosphorylation or by exogenous factors are known to attack preferentially deoxyguanosine. The latter decomposes to give mutagenic lesions, which under physiological conditions are efficiently repaired by specialized maintenance systems in the cell. Although many intermediates of the degradation pathway are today well-known, we report in this study the discovery of a new intermediate with an interesting guanidinoformimine structure. The structure elucidation of the new lesion was possible by using HPLC–MS techniques and



organic synthesis. Finally we report the mutagenic potential of the new lesion in comparison to the known lesions imidazolone and oxazolone using primer extension and pyrosequencing experiments.

INTRODUCTION

Guanine is the nucleobase with the lowest oxidation potential of the four canonical nucleobases.¹ Therefore, reactive oxygen species (ROS) degrade preferentially dG bases to give dGderived lesions.²⁻⁶ These lesions are mutagenic and typically repaired by dedicated DNA repair enzymes.7-12 Scheme 1 summarizes the main degradation pathway of 2'-deoxyguanosine. Most of the oxidation products are today known.¹³ The structures of certain putative intermediates along the dG degradation pathway and their mutagenic properties remain, however, to be uncovered. Particularly the reaction of the oxazolone lesion (dZ) to the aminoriboside, which hydrolyzes to give an AP site as the final product of the oxidative dG degradation, is an enigmatic transformation. One-electron oxidation of dG followed by reaction of the dG-radical cation $(dG^{\bullet+})$ with water and finally either oxidation or reduction of the 8-OH-dG radical give rise to 8-oxo-guanosine $(8-0x0-dG)^{14}$ or formamidopyrimidine (FaPy-dG) lesions,15-17 respectively (Scheme 1). Both are well studied oxidative lesions, which strongly contribute to the mutagenic effect of ROS. A second more complex degradation pathway involves deprotonation of $dG^{\bullet+}$ and trapping of the radical by $O_2^{\bullet-}$. The corresponding 5-OOH-dG intermediate predominantly leads to the imidazolone (dIz) lesion,^{5,18} known to rearrange to oxazolone (dZ).^{19,20} The formation of spiroiminohydantoin (dSp), guanidinohydantoin (dGh), and iminoalantoin (dIa), reported before,²¹ is assumed to take place via an intramolecular redox reaction in which the C5-peroxyl group participates in the oxidation of C8 to generate 8-hydroxy-dG as the precursor of these compounds. Oxazolone (dZ) is the last characterized intermediate before the occurrence of 1'-aminoriboside, which hydrolyzes to give an

abasic site. The mutagenic effect of ROS is caused by all of these degradation intermediates, which exist in damaged DNA in unknown relative amounts. They all have been implicated in human pathogenesis and disease.^{22–24} For example, 8-oxo-dG²⁵ and FaPy-dG lesions^{26,27} are known to produce G to T transversion mutations, while imidazolone lesions are assumed to induce G to C transversion mutations.²⁸

RESULTS AND DISCUSSION

In order to get more insight into the imidazolone/oxazolone based degradation pathway we performed HPLC-MS experiments with purified dIz lesion-containing DNA strands. For the study we prepared the oligonucleotide ODN1 (Figure 1A) and irradiated the DNA strand, containing a single dG base in the middle of the sequence, in the presence of riboflavin. The latter is a well-known dG photooxidant, which leads to the formation of imidazolone lesions.^{20,29} Indeed, irradiation of ODN1 gave a mixture of products with one major component. This oligonucleotide (dIz-ODN1) was isolated and the presence of the imidazolone lesion was confirmed by high resolution ESI mass spectrometry. Figure 1A shows the HPL-chromatogram of the purified dIz-containing ODN1 strand, together with the high resolution mass spectrum. The molecular peak for dIz-ODN1 appears at $m/z = 1367 ([M - 3H]^{3-})$, while the second major peak at m/z = 1374 is caused by the corresponding monosodium adduct. To further confirm the presence of the dIz lesion in the purified oligonucleotide, we digested the DNA

Received: December 19, 2011 Published: February 13, 2012 Scheme 1. Oxidative Degradation of 2'-Deoxyguanosine by Direct Electron Abstraction and the Corresponding Base Modifications



with P1 nuclease and subsequently incubated the reaction mixture with alkaline phosphatase and snake venom phosphodiesterase. The obtained nucleoside mixture was again analyzed by HPLC-MS. Figure 1B shows the HPLC signals of the natural nucleosides present in the digest, as well as a new signal which appears with a retention time of 9.6 min. This signal gives an m/z value of 229.0930, which is in perfect agreement with the calculated molecular weight of the dIz nucleoside. These results confirm that the first, major reaction product, formed during riboflavin induced dG oxidation, is indeed the imidazolone lesion.

While handling the dIz-containing ODN1 strand, we noticed that the imidazolone lesion is rather unstable. The lesion degraded quickly to give further reaction products. Figure 2 shows the HPL-chromatogram of the formerly pure dIz-containing DNA strand after standing of the solution for 24 h at room temperature. It is clearly visible that at least two additional compounds were formed within the DNA. To investigate the stability in more detail, we dissolved the dIz-containing DNA in different buffers at various pH values and monitored the decomposition reaction by LC–MS.

Indeed, the imidazolone lesion is unstable and rapidly degrades ($t_{1/2} = 18$ h at pH = 7.0) to give two further products with a rate that depends on the salt concentration, the temperature, and more importantly, the pH of the buffer. It is most stable when the DNA solution is kept cold (<10 °C) at pH = 7.0. Changing the pH value (Supplementary Figure S2) or increasing the temperature of the solution, e.g., to 37 °C, caused rapid decomposition.

As previously reported,³⁰ analysis of the dIz decomposition reaction by HPLC–MS showed that the dIz-ODN1 initially forms the oxazolone product dZ, which is detectable at m/z =686.27 (z = 6). If a solution of the dIz-ODN1, for example, was allowed to stand for 24 h at room temperature at pH 7.0, the solution contained finally only residual amounts of dIz-ODN1. Instead, the dZ-ODN1 dominated (>90%) and another, so far







Figure 1. (A) HPLC of the purified dIz-ODN1 and the corresponding high resolution mass spectrum $([M - 3H]^{3-}_{caled} = 1367.5547, [M - 4H + Na]^{3-}_{caled} = 1374.8820)$. (B) HPLC–MS spectra obtained from the total enzymatic digestion of the dIz-ODN1. dIz $[M + H]^+_{caled} = 229.0931$, dC $[M + H]^+_{caled} = 228.0979$, and dT $[M + H]^+_{caled} = 243.0975$.



Figure 2. HPLC of dIz-ODN1 after standing for 24 h at room temperature and the corresponding high resolution mass spectrum. dIz-ODN $[M - 6H]^{6-}_{calcd} = 683.2737$, dZ-ODN $[M - 6H]^{6-}_{calcd} = 686.2755$, and dGf-ODN $[M - 6H]^{6-}_{calcd} = 678.9438$.

Journal of the American Chemical Society

unknown reaction product, appeared. Analysis of the DNA by mass spectrometry after 48 h led to the surprising discovery of a new compound that accumulates with time to significant amounts (>25%). Certainly the new compound is more stable than the original imidazolone lesion.

For this so far unknown compound (dGf) formed in ODN1, we determined a molecular weight only 26 mass units lighter than the original dIz-ODN1 (Supplementary Figure S3). Because the dGf-ODN1 accumulated only after formation of the oxazolone lesion dZ, we reasoned that the new compound, dGf, must be a further dZ-derived degradation product that builds up in solution and that could be the unknown link between the dZ lesion and the amino-riboside. Based on the molecular weight and the fact that dGf is formed likely from dZ, we speculated that the compound forms by decarboxylation and therefore proposed the guanidinoformimine structure shown in Scheme 2.

Scheme 2. Decomposition of the Oxazolone Containing Oligonucleotide via the New Lesion As an Intermediate to 1'-Aminoriboside



The high resolution mass spectrum shown in Figure 2 provides in total three sets of signals caused by dIz-, dZ-, and dGf-ODN1. The calculated value of the dGf-containing DNA strand (m/z = 678.9438, z = 6) is in excellent agreement with the experimentally determined value of m/z = 678.9436 (z = 6), strongly supporting the decarboxylation idea.

In order to gain further support for the structure, we next performed a total enzymatic digest of dIz-ODN1, which was allowed to stand in solution for 2 days, and analyzed the digest by HPLC-MS (Figure 3). The experiment proved the simultaneous presence of DNA containing dIz, dZ, and dGf lesions. Again, the molecular weight of dGf, now as the nucleoside, with m/z = 203.1140, was in excellent agreement with the proposed structure. Furthermore, we were able to observe a signal corresponding to the putative dGf-aglycon (Gf) in an MS/MS experiment. The fragment has the molecular weight of m/z = 87.0668.

In order to investigate if the lesion composition changes in double-stranded DNA, we repeated the study with a dIzcontaining DNA double strand. The same decomposition products dIz, dZ, and dGf were observed at a similar ratio as in single-stranded DNA. However, the decomposition process was slightly slower in the duplex, probably due to partial shielding of the lesions from the aqueous environment.

To finally prove that the dGf product possesses the proposed guanidinoformimine structure, we next wanted to show that the key fragmentation product with m/z = 87.0668 is indeed the deribosylated fragment. To this end we synthesized the 3',5'-bisacetyl-dG derivative (ac_2dG) .^{31,32} After 8 min of irradiation of the compound in the presence of riboflavin, we isolated the bisacetyl-protected ac_2dIz nucleoside in 90% yield. An aqueous solution of the ac_2dIz derivative was allowed to stand at room temperature for several days. Analysis of the solution during this time showed that within 12 h the ac_2dIz nucleoside fully converted to the corresponding ac_2dZ nucleoside. After about 3





Figure 3. Enzymatic digest and HPLC–MS/MS of the dIz-ODN1 after 2 days. The experimental exact masses are in agreement with the calculated values. dT $[M + H]^+_{calcd} = 243.0975$, dC $[M + H]^+_{calcd} = 228.0979$, dIz $[M + H]^+_{calcd} = 229.0931$, dZ $[M + H]^+_{calcd} = 247.1037$, dGf $[M + H]^+_{calcd} = 203.1139$, and Gf $[M + H]^+_{calcd} = 87.0665$.

days, a new compound with a molecular weight (m/z = 287.1349) calculated for ac_2dGf was formed in about 20% yield. In the bisacetylated form, the reaction to the dGf compound is hence significantly slower than in DNA. The ac_2dGf compound was next used for MS/MS experiments, which allowed us to detect again the fragment at m/z = 87.0668. This proves that the new intermediate is the deribosylated species with the proposed guanidinoformimine structure (Supplementary Figure S6). In order to collect additional data which support the structure, we also performed NMR studies, although this was possible only for the obtained compound mixture. In support of the structure, we detected a signal at 7.78 ppm that corresponds to the amidinium-CH (Supplementary Figure S7). The signal grew with time in line with increasing amounts of dGf in solution.

From our study we conclude that the first dG degradation product is the dIz lesion, which rearranged within a few hours to give the dZ lesion as the dominating species. Subsequently decarboxylation occurs giving the new dGf intermediate. This compound is already detectable after 1 day and accumulates while the amount of the dZ lesion decreases (20 mM in water, rt). Finally 1'-amino-ribose is detected as the hydrolysis product of the dGf compound. The aminoriboside hydrolyses thereafter to form the abasic site. In this scenario, the dIz lesion is only able to contribute to the mutagenic effect of ROS in the first hours after lesion formation. Thereafter the dZ lesion will contribute dominantly followed by the newly discovered dGf lesion.

In order to determine the mutagenic effect of the lesions (dIz, dZ, and dGf), we performed primer extension experiments using one high (Klenow exo⁻) and two low fidelity polymerases (hPol κ and scPol η , respectively) and lesion-containing oligonucleotides X-ODN2 (X = dG, dIz, dZ, or

Journal of the American Chemical Society

dGf). The polymerases were allowed to extend a complementary primer strand along the lesion containing template in the presence of all four canonical triphosphates. In order to estimate the bypass efficiencies exhibited by the polymerases mentioned above we performed primer extension studies and followed the reaction using polyacrylamide gel electrophoresis. The experiment revealed that all lesions can be bypassed although the bypass efficiency was found to be limited (Figure 4A). We discovered that the efficiencies vary significantly depending on the polymerase and the lesion. Pol κ is almost completely stalled and is hardly able to incorporate even a single nucleotide opposite the lesions. The Klenow fragment in contrast incorporates one nucleotide relatively efficiently, after which it stalls to a large extent. Pol η has the highest bypassing capabilities. For all three polymerases the dGf lesion constitutes the most severely blocking lesion. Despite the low bypass efficiency, still significant amounts of the primers were fully extended. To characterize and quantify the mutations associated with the bypass, we sequenced the fully extended strands by pyrosequencing.^{33,34} The results of the primer extensions and of the pyrosequencing experiments are summarized in Figure 4.

As a control, all studies were also performed with a dG-ODN2 template strand, always resulting in quantitative incorporation of dC opposite the dG residue, independently of the polymerase used. The Klenow fragment, as a model for a high fidelity polymerase, inserts only dC opposite dG into the primer strand. The dIz lesion is preferentially base paired with dC (46%), but also dA (21%) and dG (33%) are inserted. The dZ lesion, in contrast, instructs the polymerase to insert a dA (59%) with remarkable efficiency. The dGf compound induces the incorporation of dC, dA, and dG with very similar efficiencies of about 30%. The low fidelity polymerase Polk pairs the lesions dIz and dZ preferentially with dC (58%, 41%). However, incorporation of dG is the second best option opposite dIz and the counterbase of preference for the dGf lesion. Pol η , in contrast, base pairs all lesions preferentially with dG (\approx 60%) followed by dC (\approx 40%). Finally, we observed that none of the polymerases performs any base pairing with dT, which is very surprising.

The results show that all investigated polymerases are able to bypass the three lesions to a certain extent. Pol η is the polymerase that will generate dG to dC transversion mutations with the highest rate in vitro. In vivo studies have, however, shown that in cells small oxidative lesions are likely bypassed with a high fidelity polymerase³⁵ and by the low fidelity polymerase Polk.³⁶ Bypass by Poln in vivo is consequently unlikely.³⁷ We observe that in the cases of dZ and dGf the high fidelity Klenow fragment incorporates the correct base dC in roughly only 30% of all bypass reactions. Polk shows a comparable selectivity. Overall, Polk reduces the error-rate during bypass only marginally. The unusually high incorporation of dA, by both Klenow fragment and Pol κ , can be partially explained by the A-rule of polymerases.^{38–41} We explain the small differences between our data and data in the literature also with the different method that we utilized here to investigate the error-rate during bypass.^{42–44} In our case only fully elongated primers are sequenced to provide data. Primers, in which the polymerase inserted a base opposite the lesion which are not fully extended, are consequently not analyzed.

Quite interesting is the observation that the new dGf intermediate is bypassed by all polymerases despite its open structure. dG incorporation opposite this lesion is slightly



Figure 4. (A) Polyacrylamide gel electrophoresis of the primer extension experiments and (B) graphical representation of the mutagenesis assay results. X-ODN2 (where X = dG, dIz, dZ, or dGf) is annealed to a biotinylated primer, which is extended *in vitro* by the polymerase of interest. The corresponding biotinylated strands, containing the lesion's counter-base Y, are annealed to a reverse primer and analyzed by pyrosequencing. The mutagenic "fingerprint" of each dG degradation product is directly visualized. A dG-containing control strand was used for each polymerase. Here incorporation of dC was always 100%. Incorporation of dT was not observed during the analysis.

favored in the low fidelity systems, while the high fidelity Klenow fragment incorporates dG, dC, and dA with similar probability. We believe that the non-instructive character of the new lesion is explainable by its small size, its flexibility, and possibly the existence of different tautomers. The molecules themselves should exist in cyclic fully conjugated quasi "homoaromatic" structure^{45–47} due to the presence of an

internal H-bond, known to be quite strong in similar systems. It seems that this H-bond allows the dGf lesion to adopt a quasiheterocyclic structure, with the result that the dGf structure can function as a standard nucleobase (Figure 5), which explains



Figure 5. Possible base pairs formed between dGf and dG, dA, and dC.

the efficient bypass. In this cyclic structure, the dGf compound can, however, exist in different tautomeric forms, which could be responsible for the flexible coding properties as depicted in Figure 5.

CONCLUSION

In summary we report the discovery of a new intermediate along the oxidative dG degradation pathway. The new dGflesion constitutes a significant barrier to replication *in vitro*, but depending on the polymerase, substantial bypass is observed as well. The flexible base pairing properties of the compound direct polymerases to pair it in a severely error-prone manner, thus rendering it a highly mutagenic lesion. This property is explained by us with a rigid, cyclic structure of the lesion held together by an intramolecular hydrogen bond and by the existence of different tautomeric states.

ASSOCIATED CONTENT

S Supporting Information

Experimental procedures, supplementary figures, and detailed mass spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

Thomas.Carell@cup.uni-muenchen.de

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (SFB749 and CA275/8-4) and the Volkswagen Foundation. We thank Toni Pfaffeneder for his help with the HPLC-MS/MS experiments and Dr. Markus Müller for critical reading of the manuscript.

REFERENCES

- Steenken, S.; Jovanovic, S. V. J. Am. Chem. Soc. 1997, 119, 617.
 Cadet, J.; Douki, T.; Gasparutto, D.; Ravanat, J.-L. Mutat. Res.,
- Fundam. Mol. Mech. Mutagen. 2003, 531, 5.
 (3) Cadet, J.; Courdavault, S.; Ravanat, J.-L.; Douki, T. Pure Appl.
- Chem. 2005, 77, 947.
- (4) Gimisis, T.; Cismas, C. Eur. J. Org. Chem. 2006, 1351.
- (5) Pratviel, G.; Meunier, B. Chem.-Eur. J. 2006, 12, 6018.

- (6) Cadet, J.; Douki, T.; Ravanat, J.-L. Free Radical Biol. Med. 2010, 49, 9.
- (7) Qi, Y.; Spong, M. C.; Nam, K.; Banerjee, A.; Jiralerspong, S.; Karplus, M.; Verdine, G. L. *Nature* **2009**, *462*, 762.
- (8) Coste, F.; Ober, M.; Carell, T.; Boiteux, S.; Zelwer, C.; Castaing, B. J. Biol. Chem. **2004**, *279*, 44074.
- (9) Leipold, M. D.; Workman, H.; Muller, J. G.; Burrows, C. J.; David, S. S. *Biochemistry* 2003, 42, 11373.
- (10) Scharer, O. D. Angew. Chem., Int. Ed. 2003, 42, 2946.
- (11) Serre, L.; De, J. K. P.; Boiteux, S.; Zelwer, C.; Castaing, B. EMBO J. 2002, 21, 2854.
- (12) Fromme, J. C.; Verdine, G. L. Nat. Struct. Biol. 2002, 9, 544.
- (13) Burrows, C. J.; Muller, J. G. Chem. Rev. 1998, 98, 1109.
- (14) Cadet, J.; Douki, T.; Ravanat, J.-L. Acc. Chem. Res. 2008, 41, 1075.
- (15) Dizdaroglu, M.; Kirkali, G.; Jaruga, P. Free Radical Biol. Med. **2008**, 45, 1610.
- (16) Ober, M.; Linne, U.; Gierlich, J.; Carell, T. Angew. Chem., Int. Ed. 2003, 42, 4947.
- (17) Tudek, B. J. Biochem. Mol. Biol. 2003, 36, 12.
- (18) Neeley, W. L.; Essigmann, J. M. Chem. Res. Toxicol. 2006, 19, 491.
- (19) Raoul, S.; Berger, M.; Buchko, G. W.; Joshi, P. C.; Morin, B.; Weinfeld, M.; Cadet, J. J. Chem. Soc., Perkin Trans. 2 **1996**, 371.
- (20) Cadet, J.; Berger, M.; Buchko, G. W.; Joshi, P. C.; Raoul, S.; Ravanat, J.-L. J. Am. Chem. Soc. **1994**, 116, 7403.
- (21) Luo, W.; Muller, J. G.; Burrows, C. J. Org. Letters 2001, 3, 2801.
- (22) Tudek, B.; Winczura, A.; Janik, J.; Siomek, A.; Foksinski, M.; Olinski, R. *Am. J. Transl. Res.* **2010**, *2*, 254.
- (23) Sedelnikova, O. A.; Redon, C. E.; Dickey, J. S.; Nakamura, A. J.; Georgakilas, A. G.; Bonner, W. M. *Mutat. Res., Rev. Mutat. Res.* **2010**, 704, 152.
- (24) Radak, Z.; Boldogh, I. Free Radical Biol. Med. 2010, 49, 587.
- (25) Hsu, G. W.; Ober, M.; Carell, T.; Beese, L. S. Nature 2004, 431, 217.
- (26) Delaney, M. O.; Wiederholt, C. J.; Greenberg, M. M. Angew. Chem., Int. Ed. 2002, 41, 771.
- (27) Ober, M.; Mueller, H.; Pieck, C.; Gierlich, J.; Carell, T. J. Am. Chem. Soc. 2005, 127, 18143.
- (28) Kino, K.; Sugiyama, H. Chem. Biol. 2001, 8, 369.
- (29) Kino, K.; Saito, I.; Sugiyama, H. J. Am. Chem. Soc. 1998, 120, 7373.
- (30) Gasparutto, D.; Ravanat, J. L.; Gerot, O.; Cadet, J. J. Am. Chem. Soc. **1998**, 120, 10283.
- (31) Matsuda, A.; Shinozaki, M.; Suzuki, M.; Watanabe, K.; Miyasaka, T. *Synthesis* **1986**, *1986*, 385.
- (32) Wei, G.; Loktionova, N. A.; Pegg, A. E.; Moschel, R. C. J. Med. Chem. 2005, 48, 256.
- (33) Ronaghi, M.; Uhlén, M.; Nyrén, P. Science 1998, 281, 363.
- (34) Munzel, M.; Lischke, U.; Stathis, D.; Pfaffeneder, T.; Gnerlich,
- F. A.; Deiml, C. A.; Koch, S. C.; Karaghiosoff, K.; Carell, T. Chemistry 2011, 17, 13782–13788.
- (35) Neeley, W. L.; Delaney, J. C.; Henderson, P. T.; Essigmann, J. M. J. Biol. Chem. 2004, 279, 43568.
- (36) Prakash, S.; Johnson, R. E.; Prakash, L. Annu. Rev. Biochem. 2005, 74, 317.
- (37) Hubscher, U.; Maga, G.; Spadari, S. Annu. Rev. Biochem. 2002, 71, 133.
- (38) Bernard S, S. DNA Repair 2002, 1, 125.
- (39) Sagher, D.; Strauss, B. Biochemistry 1983, 22, 4518.
- (40) Kunkel, T. A.; Schaaper, R. M.; Loeb, L. A. Biochemistry 1983, 22, 2378.
- (41) Boiteux, S.; Laval, J. Biochemistry 1982, 21, 6746.
- (42) Duarte, V.; Gasparutto, D.; Jaquinod, M.; Cadet, J. Nucleic Acids Res. 2000, 28, 1555.
- (43) Kino, K.; Sugiyama, H. Mutat. Res., Fundam. Mol. Mech. Mutagen. 2005, 571, 33.
- (44) Kino, K.; Sugasawa, K.; Mizuno, T.; Bando, T.; Sugiyama, H.; Akita, M.; Miyazawa, H.; Hanaoka, F. *ChemBioChem* **2009**, *10*, 2613.