Comparative Nucleosomal Reactivity of 5-Formyl-Uridine and 5-Formyl-Cytidine

Leander Simon Runtsch, Michael Stadlmeier, Alexander Schön, Markus Müller, and Thomas Carell* [a]

In memory of François Diederich

Abstract: 5-Formyl-deoxyuridine (fdU) and 5-formyl-deoxyctydine (fdC) are formyl-containing nucleosides that are created by oxidative stress in differentiated cells. While fdU is almost exclusively an oxidative stress lesion formed from deoxythymidine (T), the situation for fdC is more complex. Next to formation as an oxidative lesion, it is particularly abundant in stem cells, where it is more frequently formed in an epigenetically important oxidation reaction performed by α-ketoglutarate dependent TET enzymes from 5-methyl-deoxycytidine (mdC). Recently, it was shown that genomic fdC and fdU can react with the ε-aminogroups of nucleosomal lysines to give Schiff base adducts that covalently link nucleosomes to genomic DNA. Here, we show that fdU features a significantly higher reactivity towards lysine side chains compared with fdC. This result shows that depending on the amounts of fdC and fdU, oxidative stress may have a bigger impact on nucleosome binding than epigenetics.

5-Formyl-deoxyuridine (fdU) is a major lesion produced by reactive oxygen species that attack genomic deoxythymidine (T) (Figure 1a).[1–3] Similarly, 5-formyl-deoxyctydine (fdC) is formed by oxidative stress from 5-methyl-deoxycytidine (mdC), but because only about 5 % of the deoxycytidines (dC) in the genome are methylated to mdC,[4] fdC is formed only in minor quantities. In contrast to fdU, however, fdC is also formed in an enzymatic oxidation reaction from mdC with the help of TET enzymes (Figure 1a). These enzymes oxidize mdC first to 5-hydroxymethyl-deoxycytidine (hmdC) and in a second step to fdC.[5–7] fdC is then removed from the genome by the action of the base excision repair (BER) enzyme thymine-DNA-glycosylase (TDG) to give an abasic site (Ap) that is finally replaced by an unmodified dC.[8] Alternatively, it is assumed that fdC can react directly to dC by a deformylation reaction.[9] fdU is in contrast exclusively removed from the genome by the BER glycosylase SMUG1, which also generates Ap sites that are replaced by T.[10]

Elevated levels of fdC were detected in stem cells and recently also in neurons.[11–14] In cells with high Tet activity, like stem cells, the fdC levels even exceed those of fdU.[12] In differentiated tissues however, fdU is the dominant formyl-containing modification.

It was recently shown by Li et al. and Ji et al. for fdC as well as Zou et al. for fdU that these bases can form a covalent Schiff base adduct with lysine side chains that are present at many positions in the histones of the nucleosome core particle.[15–17] This process could be important for nucleosome positioning and ultimately for forming compact and hence silenced chromatin structures.[15,16] Other studies show a distinct behavior of these adducts during replication and transcription.[17,19–22]
While both fdC and fdU can react with lysine side chains, it is unclear which of the two modifications is more reactive in nucleosomes. Balasubramanian and co-workers observed a higher reactivity of fdU towards nucleophiles with an α-effect,[23–25] suggesting that this trend could also hold true for the reaction with nucleosomal lysines. Here, we answer the question of whether the more abundant fdU influences genome stability by covalent adduct formation more strongly than fdC.

Reactivity differences between fdU and fdC can potentially arise from their very different conformational and H-bonding preferences (Figure 1b). While fdC exists predominantly in the syn-conformation, due to the formation of a strong intramolecular 6-membered ring H-bond,[23–25] fdU is known to prefer the anti-conformation.[23–26] A stabilizing intramolecular H-bond does not exist in fdU.

In order to examine the reactivity of fdU towards lysine side chains in direct comparison to fdC, we prepared the acetylated heptapeptide AcNH-IEAKGER-OH (1) containing a central lysine. This peptide 1 mimics a tryptic peptide with a charge distribution that is suitable for identification via MS. The amino group at the N-terminus was protected with an acetyl group to block its reactivity towards formyl groups. In consequence, the peptide has only one amino group that can react. For the reaction, an excess of fdC or fdU was added to a solution of the peptide 1.

This mixture was incubated at 37 °C for 75 min to allow formation of the Schiff base adducts 2 and 3 (Figure 2a). To stabilize the adducts 2 and 3 for subsequent MS analysis, we reduced the imines with NaBH₄. This generates the stable secondary amine adducts 4 and 5. These adducts were next analyzed by HPLC-MS (see Supporting Information for detailed method). Although mass spectrometry is a priori not a quantitative method due to different ionization efficiencies of the analytes, the almost identical structures of the peptide adducts 4 and 5 allow to extract some quantitative information, assuming that the ionization properties of adducts 4 and 5 are comparable to each other. This assumption is based on previous studies which show that ionization efficiencies are mostly dependent on the molecular volume and the hydrophobicity of a species, which should both be very similar for the two adducts.[27–28] The measured ion current produced by the adducts 4 and 5 (Figures 2b and c), showed a much higher signal for the fdU-adduct, which led us to conclude that fdU is the more reactive nucleoside. Judging by the fact that we also detected the triply charged ion of the fdC-adduct, fdC seems to increase the effective charge of this species. Since this has been shown to positively correlate with ionization efficiency to a lesser extent than molecular volume and hydrophobicity,[27–28] we believe that we may even underestimate the difference in reactivity.

To investigate whether this reactivity difference persists when the fdU and fdC nucleosides are embedded in an oligonucleotide context, we prepared the DNA strands 6 and 7 by solid phase synthesis and again incubated them with peptide 1 (Figure 3a). In this experiment we used an excess of the peptide 1 in order to enable a gel electrophoretic analysis of the ssDNA and in order to detect also small amounts of adduct. After incubation overnight at r.t., the adducts were reduced with NaBH₄CN and then analyzed by denaturing urea polyacrylamide gel electrophoresis and subsequent staining with SYBR Green I. When we performed the overnight incubation at physiological pH (7.4), we saw in both cases (6 and 7, lanes c, d) no formation of any adduct (8, 9, Figure 3b), as the observed bands were on the same height as in the controls without the peptide (lanes a, b). This shows that the reactivity of both formyl-nucleotides in a ssDNA context is not sufficient to promote Schiff base formation towards this peptide to an extent that it can be detected on the gel. As Schiff base

![Figure 2.](image-url)
formation can be enhanced with a catalyst, we repeated the incubation of 6 and 7 with 1 under non-physiological conditions in the presence of p-anisidine at a lower pH (catalyst A).

Using this system, we now detected slight adduct formation for fdU (lane f), but clearly none for fdC (lane e) supporting a higher reactivity for fdU. Adduct formation in lane f is visible by the partial shift of fluorescence intensity to higher molecular weight bands for the fdU-adduct 9 on the denaturing gel (Figure 3b). Even better adduct formation of the fdU-containing strand was observed when we used the stronger catalyst p-diaminobenzene (B) (lane h). For fdC however, adduct formation was not observed even under these accelerated conditions (lane g).

These studies show that the reactivity of the two formyl-containing bases is quite low in solution, which is expected because in water the equilibrium between Schiff base and aldehyde is far on the aldehyde side. Additionally, it has been shown that the reactivity is dependent on the peptide sequence. Nevertheless, from these studies fdU appears to be more reactive than fdC. These results suggest that the previously observed reaction of fdC and fdU with nucleosomes may benefit greatly from the tight association of the nucleosome with the formyl-containing DNA which would increase the reactivity of the system entropically.

In order to compare the reactivity of fdC and fdU in the context of nucleosomes, we generated two longer dsDNA sequences (11, 12) derived from a 145 bp region of the Widom 601 DNA sequence (10), which is known for its capability to form distinct mononucleosomes. These strands were prepared by a PCR reaction in which either TTP was completely replaced by fdUTP or dCTP fully exchanged against fdCTP. This provides 145mer DNA duplexes in which, except for the primer sequences, either all dCs are replaced by fdCs (11) or all Tss are exchanged against fdUs (12, Figures 4a and S3). As a control, we used the unmodified version of the sequence (10). For the PCR reaction of the modified strands, we had to carefully adjust various parameters in order to generate the 145mer containing either 54 fdUs or 69 fdCs. We achieved this by using the optimized polymerase blend KOD XL and lengthening the extension time (see Supporting Information), similar to what we described earlier. We then used these double strands together with the four nucleosome-constituting histones H2A, H2B, H3 and H4, which we acquired commercially as H2A/H2B dimer and H3/H4 tetramer, to assemble nucleosomes using the dilution method reported by the manufacturer. Importantly, in these experiments we did not add any additional catalysts to ensure physiological conditions. Also, since these histones were shipped in TRIS buffer (which features a primary amine), we exchanged this buffer against HEPES by careful dialysis, to avoid interference of the buffer with the reaction. We then analyzed the successful nucleosome formation by separating them from free DNA using native gel electrophoresis (Figure 4b). Interestingly, to visualize the reconstituted fdU-containing nucleosomes, we could not stain the gels with SYBR Green I. With this reagent it was impossible to visualize the nucleosome assembly with strand 12. We assume that this lack of detectability is caused by the structure of the heavily fdU-modified DNA. The structural effects generated by the high density incorporation of fdU may prevent proper intercalation of the fluorophore into the duplex, as also a single fdU modification has been shown to affect DNA structure. To circumvent this problem, we used a forward primer containing the fluorescent dye Cy3 covalently attached to its 5’-end in all PCR reactions. This generates the DNA duplexes 10–12 with a Cy3 dye label as the readout signal. Analysis of the nucleosome assembly by native PAGE using this Cy3 detection reveals that the fdU containing 601 sequence gives a slightly better assembly of the nucleosomes as compared to fdC. The amount of free DNA is clearly decreased in the assembly reaction with fdU compared to fdC and dC. We believe that this is already an effect of the higher reactivity of fdU towards the lysine amines.

To investigate the reactivity of fdC and fdU, we allowed the reconstituted nucleosomes to react for 24 h at 37°C. In order to stabilize the Schiff bases, we next added NaBH₄CN to reduce the imines and analyzed the product using denaturing SDS-PAGE. This should disrupt unstably linked complexes. The data for these experiments are shown in Figures 4c and d. The samples with the control strand 10 (lanes a, b) show only a...
band corresponding to free DNA, meaning that without formyl groups, no adducts can form. This changes for the formyl-containing strands 11 and 12: even without reduction (lanes c, h), we detected a few distinct bands at high molecular weight, which likely resemble stable fdC- and fdU-DNA-nucleosome complexes that do not denature on the gel. This is surprising, given the reversible nature of the non-reduced Schiff bases. We postulate that these bands are likely the result of multiple adducts forming simultaneously at various positions of the histones, which collectively counteract the reversible nature of imine adduct formation. Indeed, when we heated the non-reduced complex to 95°C for 5 min in sample buffer before loading onto the gel (lanes d, i), these bands mostly disappeared supporting this interpretation.

Informative, however, is a comparison of the products that result from heating between the fdC and fdU sample. For fdC (lane d) mainly the intensity of the band representing free DNA increases, arguing that the nucleosomal complex dissociates completely. In contrast, preheating of the fdU nucleosome (lane i) leads to many additional diffuse bands in addition to an increased intensity of the free DNA band. Indeed, when we heated the non-reduced complex to 95°C for 5 min in sample buffer before loading onto the gel (lanes d, i), these bands mostly disappeared supporting this interpretation.

Our experiments were performed with artificial nucleosomes, in which fdU and fdC occupy different positions within the complex (see Figure S3). This could obscure a direct reactivity comparison. It should however be noted that fdU occupies fewer sites on the nucleosome in the 601 context. Despite this, we see that the DNA-nucleosome adducts disappear, leaving just a diffuse DNA-only band behind. This confirms the involvement of the histone proteins in complex formation. Judging from the sharper band obtained from the released fdC-containing DNA 11 (lane g) compared to the fdU-containing DNA 12 (lane l), one can conclude that the digestion of the fdC-complex is more complete. This further strengthens the idea that the fdU-containing complexes are more stable and hence more compact, which leads to reduced accessibility for proteolysis.
buffer we used for nucleosome reconstitution (Figure S6). This revealed that while the unmodified strand 10 is present as B-DNA, the fdC-containing strand 11 adopts a different conformation. This conformation features similar CD properties to what has been reported for multiple fdC modifications contained in shorter DNA, termed F-DNA.[13] In contrast, the fdU-containing strand 12 exhibits a CD spectrum that resembles A-DNA.[34] These findings suggest that changes in DNA conformation, induced by the clustered fdC or fdU modifications, likely also play a role in the observed difference of adduct formation in nucleosomes.

In summary, we confirm the previously reported Schiff base adduct formation between fdC- and fdU-containing DNA and histones.[15–17] We report here that fdU is forming these adducts more readily than fdC, which is seen in all our experiments comprising different levels of complexity. In nucleosomes, this seems to be due to a combination of a higher reactivity of the formyl group of fdU as well as a difference in DNA conformation. Because fdU is the more abundant oxidative stress base in differentiated tissues, this is an important result that shows how oxidative stress can potentially have a stronger influence on transcriptional activity than epigenetic processes. Stable Schiff base formation in an aqueous environment is surprising, given that the equilibrium of the reaction of an amine with an aldehyde to a Schiff base is far on the hydrolyzed aldehyde side under normal aqueous conditions. This is confirmed in our experiments with the peptide 1 and fdC and fdU nucleoside or short fdC- and fdU-containing ssDNA. It seems that the simultaneous formation of multiple Schiff bases is beneficial to stabilize the adducts. This would mean that predominantly clustered fdC as suggested to occur by Balasubramanian and co-workers[34] can impact transcriptional activity by compaction. Our data therefore indicate that formation of genomic islands containing multiple fdC may indeed mediate epigenetic control.

The fact that fdU forms Schiff base adducts so efficiently could establish an additional mechanism of how oxidative stress and clustered oxidative lesions in particular can induce epigenetic processes[25] by influencing the tightness of the DNA-nucleosome association. Similar to the aforementioned clustering of fdC, the formation of clustered oxidative lesions has indeed been proposed as well.[19]

Based on previous reports, it is furthermore likely that Schiff base formation affects the accessibility of fdU and fdC,[17,19,20,36] which would attenuate their repair. This would be a mechanism of how fdU and, in an epigenetic context also fdC, can form stable sites in the genome that escape repair. For fdC such a permanent or semi-permanent character was indeed already reported.[15,29]

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (DFG) for financial support via CRC1309 (Project ID 325871075) and CRC1361 (Project ID 393547839). This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program under grant agreement No 741912 (EpiR) and Marie Sklodowska-Curie grant agreement No 765266 (LightDyNAmics). We are grateful for additional funding from the Volkswagen Foundation. L.S.R and M.S. thank the Fonds der Chemischen Industrie for predoctoral fellowships. We thank Iacovos N. Michaelides and Franziska R. Traube for help with the solid phase synthesis and Eleonora Nigliato for assisting with gel shift experiments. Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: DNA damage · DNA histone adducts · DNA repair · epigenetics · nucleosomes

[1] S. Frelon, T. Douki, J.-L. Ravanat, J.-P. Pouget, C. Tomabene, J. Cadet, Chem. Res. Toxicol. 2004, 17, 1002–1010.
[2] J.-P. Pouget, S. Frelon, J.-L. Ravanat, I. Testard, F. Odin, J. Cadet, Radiat. Res. 2002, 157, 589–595.
[3] S. Frelon, T. Douki, A. Favier, J. Cadet, J. Chem. Soc. Perkin Trans. 1 2002, 2868–2870.
[4] D. Globisch, M. Münzel, M. Müller, S. Michalakis, M. Wagner, S. Koch, T. Brückl, M. Biel, T. Carell, PLoS One 2010, 5, e151637.
[5] M. Tahliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Bruhdo, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, Science 2009, 324, 930–935.
[6] S. Ito, A. C. D’ Alessio, O. V. Taranova, K. Hong, L. C. Sowers, Y. Zhang, Nature 2010, 466, 1129–1133.
[7] S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, Y. Zhang, Science 2011, 333, 1300–1303.
[8] A. Maiti, A. C. Drohat, J. Biol. Chem. 2011, 286, 35334–35338.
[9] K. Iwan, R. Rahimoff, A. Kirchner, F. Spada, A. S. Schröder, O. Kosmatchev, S. Ferizaj, J. Steinbacher, E. Parsa, M. Müller, T. Carell, Nat. Chem. Biol. 2018, 14, 72–78.
[10] A. Masaoka, M. Matsubara, R. Hasegawa, T. Tanaka, S. Kurisu, H. Terato, Y. Ohyama, N. Karino, A. Matsuda, H. Ide, Biochemistry 2003, 42, 5003–5012.
[11] T. Paffeneder, B. Hackner, M. Truss, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeier, T. Carell, Angew. Chem. Int. Ed. 2011, 50, 7008–7012; Angew. Chem. 2011, 123, 7146–7150.
[12] T. Paffeneder, F. Spada, M. Wagner, C. Brandmayr, S. K. Laube, D. Eisen, M. Truss, J. Steinbacher, B. Hackner, O. Kotljarov, D. Schuermann, S. Michalakis, O. Kosmatchev, S. Schiesser, B. Steigenberger, N. Raddaoui, G. Kashiwazaki, U. Müller, C. G. Spruijt, M. Vermeulen, H. Leonhardt, P. Schar, M. Müller, T. Carell, Nat. Chem. Biol. 2014, 10, 574–581.
[13] M. Bachman, S. Urbe-Lewis, X. Yang, H. E. Burgess, M. Iurlaro, W. Reik, A. Murrell, S. Balasubramanian, Nat. Chem. Biol. 2015, 11, 555–557.
[14] M. Wagner, J. Steinbacher, T. F. Kraus, S. Michalakis, B. Hackner, T. Paffeneder, A. Perera, M. Müller, A. Giese, H. A. Kretzschmar, T. Carell, Angew. Chem. Int. Ed. 2015, 54, 12511–12514; Angew. Chem. 2015, 127, 12691–12695.
[15] F. Li, Y. Zhang, J. Bai, M. M. Greenberg, Z. Xi, C. Zhou, J. Am. Chem. Soc. 2017, 139, 10617–10620.
[16] S. Ji, H. Shao, Q. Han, C. L. Seiler, N. Y. Tretyakova, Angew. Chem. Int. Ed. 2017, 56, 14130–14134; Angew. Chem. 2017, 129, 14318–14322.
[17] G. Zou, C. Liu, W. Zeng, W. Yang, K. Zhang, Y. Xie, C. Chen, X. Zhou, CCS 2020, 2, 54–63.
[18] E. A. Raiber, G. Portella, S. Martinez Cuesta, R. Hardisty, P. Murat, Z. Li, M. Iurlaro, W. Dean, J. Spindel, D. Berardi, Z. Liu, M. A. Dawson, W. Reik, S. Balasubramanian, Nat. Chem. 2018, 10, 1258–1266.
[19] S. Ji, I. Fu, S. Naldga, H. Shao, A. K. Basu, S. Brodey, N. Y. Tretyakova, Nucleic Acids Res. 2018, 46, 6455–6469.
[20] S. Ji, D. Park, K. Kropachev, M. Kolbanovskyi, I. Fu, S. Brodey, M. Essawy, N. E. Geacintov, N. Y. Tretyakova, J. Biol. Chem. 2019, 294, 18387–18397.
