



Communication

5-Formyluracil targeted biochemical reactions with proteins inhibit DNA replication, induce mutations and interference gene expression in living cells

Guangrong Zou, Kaiyuan Zhang, Wei Yang, Chaoxing Liu, Zhentian Fang, Xiang Zhou*

College of Chemistry and Molecular Sciences, Key Laboratory of Biomedical Polymers of Ministry of Education, The Institute for Advanced Studies, Hubei Province Key Laboratory of Allergy and Immunology Wuhan University, Wuhan 430072, China

ARTICLE INFO

Article history:

Received 14 January 2021

Received in revised form 29 January 2021

Accepted 17 February 2021

Available online 19 February 2021

Keywords:

5-Formyluracil
Biochemical reaction
DNA replication
Mutation
Gene expression

ABSTRACT

Covalent DNA–protein cross-links are toxic DNA lesions that interfere with essential biological processes, which can cause serious biological consequences, such as genomic instability and protein misexpression. 5-Formyluracil (5fU) as an important modification in DNA, which is mainly from oxidative damage, exists in a variety of cells and tissues. We have reported that 5fU mediated DNA–protein conjugates could exist in human cells [Zhou *et al.* *CCS Chem.* 2 (2020) 54–63]. We now aimed to explore its potential biological effects *in vitro* and *in vivo*. In this paper, we firstly reported that 5fU intermediated DNA–peptide or DNA–protein conjugates (both were called DPCs) could inhibit different polymerases bypass or cause mutations. Then we further investigated the functional impacts caused by 5fU-mediated DPCs, which appeared in different gene expression components [in the promoter sequence or 5'-untranslated regions (UTR)]. These results together may contribute to a broader understanding of DNA–protein interactions as well as the biological functions associated with 5fU.

© 2021 Chinese Chemical Society and Institute of Materia Medica, Chinese Academy of Medical Sciences. Published by Elsevier B.V. All rights reserved.

Natural chemical modifications in DNA or RNA nucleobases profoundly influence a series of biological processes, such as 5mC was proved to typically associate with promoters of silent genes while 5hmC could mark actively transcribed genes [1,2]. 5-Formyluracil (5fU), which is widely distributed in cells and organisms [3,4], has drawn wide attention in many related fields. Most scientists hold the view that 5fU is one of the oxidative damage products of thymine and arisen by ultraviolet (UV) light [5], γ -irradiation [6] and so on. As a kind of well-known genotoxic lesion, it can introduce base mispairing and further perturbs the DNA function [6]. What is more, 5fU might be associated with some human diseases, such as thyroid carcinoma [7].

Genomic DNA is closely related to a range of architectural and regulatory proteins. Noncovalent DNA–protein interactions are vital for the normal cell function and responding to DNA damage. Their disruption is likely to result in serious biological consequences [8]. The close contact of genomic DNA with the cellular proteins may result in reactions, thus form conjugates even covalent DNA–protein cross-links (DPCs) upon exposure to unsaturated carbonyls, free radicals, heavy metals, UV radiation,

nitric oxide and toxic metals [9]. Considering the enormous size of DPCs compared with other nucleobase lesions as well as their ability to disrupt the recognition of DNA by specific proteins, covalent DPCs may interfere with DNA replication and transcription [10], induce mutations and toxicity [11]. Thus, cellular DPCs may contribute to human diseases including cancer, aging and neurodegenerative disorders [12]. Although multiple sites can be targeted for cross-linking, such as N7 of guanine, C-5 methyl group of thymine, exocyclic amino groups of guanine, cytosine and adenine [9]. Most studies need extra modification to form relatively stable DPCs *via* derivatives such as 7-deaza-dG [13], azide-functionalized C5-dT [14]. Besides, only rare studies explored their cellular effects, especially for inherent DPCs.

It was reported that 5fC based DNA–histone Schiff base sites existed in a cellular environment and played a role in controlling transcription [15]. As for the amount of 5fC and 5fU at the same quantitative level [16] and the higher reactivity of aldehyde group on 5fU than 5fC [17], we suspected that 5fU-based DPCs would also cause crucial biological effects. Besides, our group recently reported 5fU can yield regulable DPCs *via* chemical interaction between amino and aldehyde groups *in vitro* and *in vivo* [18]. However, to our knowledge, there are no correlational research that address the influence of 5fU based biochemical reactions associated with proteins. Therefore, it is quite necessary to look

* Corresponding author.

E-mail address: xzhou@whu.edu.cn (X. Zhou).

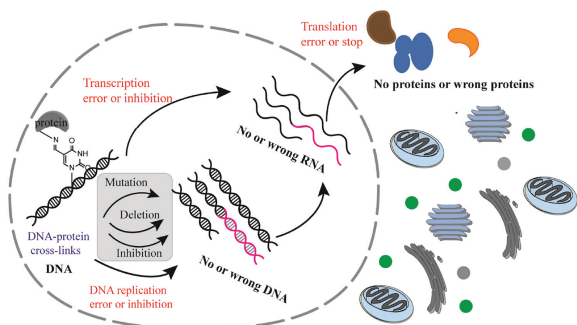
into the DPCs influence on DNA replication, transcription and translation caused by 5fU.

In this paper, we investigated the biological effects of 5fU based DNA–protein cross-links *in vitro* and in living cells. Except for researching the effects of DPCs on DNA polymerase mediated replication, we also explored the magnitude of the negative effect on transcription and translation due to the DPCs formation, which occurred at the promoter or 5'-untranslated regions (UTR) of the gene (Scheme 1).

To examine the influence of DPCs on DNA replication, we firstly chose two short peptides based on previous research [19]. One was 10-mer length of c-Myc protein analogue, in which the lysine was replaced by arginine; the other was 31-mer length of human β -endorphin peptide. The two peptides were incubated with 5fU-containing DNA respectively (Fig. 1A). Resulting conjugates were stabilized by reduction, then analyzed and purified by 20% denaturing polyacrylamide gel (Fig. 1B and Fig. S1A in Supporting information). The correctness of the conjugates was confirmed by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) (Fig. 1C and Fig. S1B in Supporting information).

Then we selected three frequently used DNA polymerases, Klenow fragment polymerases (KF, New England Biolabs, U.S.A), KOD dash (TOYOBO, Japan) and Bst polymerases (New England Biolabs, U.S.A) to investigate the influence of 5fU DNA–peptide conjugates on DNA replication *in vitro*. Compared with the template of 15 mer-5fU or 15 mer-T, the experiment results revealed that the cross-links of 5fU with 10-mer peptide completely or almost completely reduced the amount of full-length primer extension products (Fig. S2 in Supporting information).

For further study the negative effects of this DNA lesion, we selected translesion synthesis (TLS) polymerases κ and η to act on different size of DNA–peptide conjugates. As for there were some reports that bypass of modified bases based DNA lesions by TLS polymerases was almost conserved in bacteria, yeast and mammalian cells [20]. Under the same reaction conditions with TLS polymerases κ , even if the incubation time lasted as long as 60 min and some overextended products could be observed for the template 15 mer-5fU or 15 mer-T, not all primers were extended to full length with template 15 mer-5fU–10 mer peptide (Fig. 2A) or 15 mer-5fU–31 mer peptide containing 5fU–peptide conjugates (Fig. S3A in Supporting information). Besides, we found the larger the size of conjugates, the less the full-length extension product. Besides, at the same incubation time, the extension products for 15 mer-5fU–31 mer peptide were obvious less than 15 mer-5fU–10 mer peptide (Fig. 2A and Fig. S3A). More obvious phenomenon was obtained with TLS polymerases η (Fig. S4 in Supporting information).



Scheme 1. Biological effects of DNA–protein biochemical reactions at 5-formyluracil sites on DNA replication, transcription and translation.

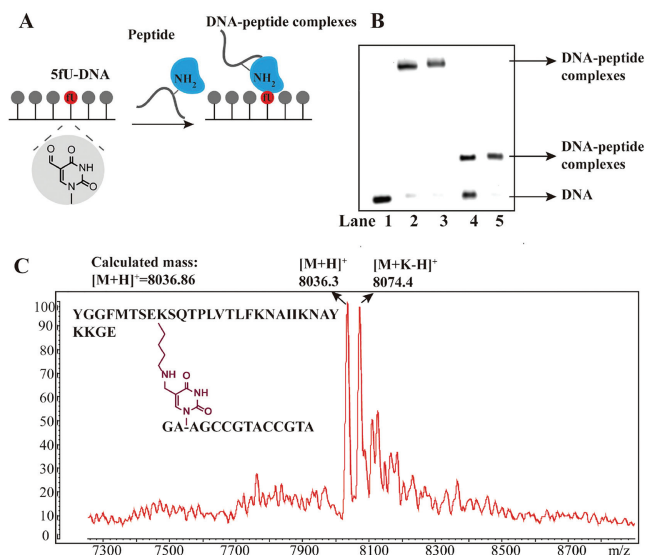


Fig. 1. The formation and analysis of DNA–peptide cross-links at 5fU sites in DNA. (A) Scheme of DPCs formation. (B) Denaturing PAGE analysis of the cross-links between 15 mer-5fU and 31-mer peptide or 10-mer peptide. Lane 1: 15 mer-5fU only. Lane 2: reaction mixture between 15 mer-5fU and 10-mer peptide. Lane 3: products of reaction between 15 mer-5fU and 10-mer peptide after purification. Lane 4: reaction mixture of 15 mer-5fU with 31-mer peptide. Lane 5: reaction products of 15 mer-5fU with 31-mer peptide after purification. (C) MALDI-TOF MS spectrum of the DPCs between 15 mer-5fU and 10-mer peptide.

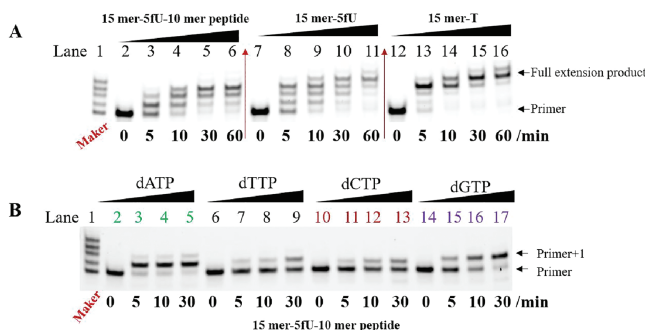


Fig. 2. Primer extension assays of TLS polymerases κ with different DNA–peptide cross-links, compared with 15 mer-5fU or 15 mer-T. (A) and (B) Denaturing PAGE analysis of the extension products generated from the template of 15 mer-5fU–10 mer peptide in the mixture of dNTPs (A) or individual dNTP (B).

We also performed the primer extension assay with individual dNTP to probe whether the 5fU–peptide cross-links could be accommodated in the active site of TLS polymerases and copied with high fidelity. For TLS polymerases η , almost only dATP could be incorporated into the primer extension products when the template was 15 mer-5fU–10 mer peptide (Fig. S5A in Supporting information). While for the template of 15 mer-5fU–31 mer peptide, none of dNTP could be involved into the primer (Fig. S5B in Supporting information). At the same time, we discovered that almost all of dNTP could be paired with 5fU (Fig. S5C in Supporting information) or T (Fig. S5D in Supporting information). Meanwhile, we also conducted the primer extension assay with individual dNTP in the presence of TLS polymerases κ . Unlike the TLS polymerases η , not only 15 mer-5fU–10 mer peptide, but also 15 mer-5fU–31 mer peptide, both dATP and dGTP could be paired by TLS polymerases κ (Fig. 2B and Fig. S3B), which indicating the possibility of mutation by 5fU based DPCs. Besides, for the 15 mer-5fU–10 mer peptide, there were extra bands above the primer for dTTP and dCTP (Fig. 2B). By the way, for the 5fU and T, almost every

kinds of dNTP could be incorporated into primer with the help of TLS polymerases κ (Figs. S6A and B in Supporting information). In short, for TLS polymerases η , the effects of DPCs made by 5fU and peptide mainly was inhibiting DNA replication. While for TLS polymerases κ , the influence of 5fU-peptide cross-links were prone to replication inhibition and mutation.

Considering most of DNA were in the nucleus and proteins surrounding the DNA were mainly histones, we chose cross-linked products of 5fU-histone H3 as a model for investigating the effect of 5fU DNA-protein on polymerases. Firstly, we prepared the 5fU-histone H3 cross-links by 4%–12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) after the reaction between FAM-labeled 5fU DNA and histone H3. The purity of the product was > 97% (Fig. 3A). Before the purification, we applied the mixture of reaction between FAM-labeled 18-mer DNA and Histone H3 for isothermal amplification by TLS polymerases η and κ . The 20% denaturing polyacrylamide gel suggested that the amplification products were rare or almost none, especially for TLS polymerases η (Fig. 3B). To further prove the impact of DPCs, we used the purified 5fU-histone H3 conjugates template for amplification. As expected, no new band was visible to the naked eye for TLS polymerases η as well as κ (Fig. 3C). However, when the protein component of DPCs were digested with proteinase and pronase, the full length of extension products could be obtained again (Fig. 3D). This indicated that the 5fU-protein cross-links inhibited the DNA replication. In view of the size of DPCs, the degree of inhibition of 5fU-protein cross-links was greater than 5fU-peptide cross-links (Fig. 2 and Figs. S3–S5).

The process of transcription initiation in eukaryotes is rather complex and a variety of transcription factors (TFs) are required to assist the transcription process *in vivo*. These TFs can recognize specific DNA sequences and guide expression of the genome. Mutations in TFs binding sites can cause many human diseases and abnormal phenotypes. AP-1 (activator protein 1) and NF- κ B (nuclear factor kappa-B) were two normal transcription factors that are regulated by intracellular oxidation and reduction states [21]. As for DNA-protein crosslinks were mainly from oxidative damage, and intracellular oxidative damage could influence gene expression through transcriptional or post-translational pathways

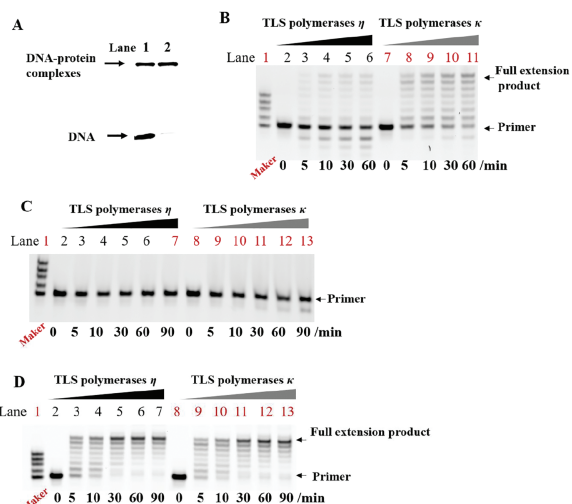


Fig. 3. The formation and DNA replication inhibition effect of DPCs between 5fU containing DNA and histones H3. (A) SDS-PAGE analysis of FAM-labelled 18-mer DNA containing 5fU and histone H3 cross-links. Lane 1 was the reaction mixture between DNA and protein. Lane 2 was the products of the reaction after purification. Figs. B–D were primer extension assays of TLS polymerases with templates of DNA-histone H3 cross-links mixture without purification (B), purified products (C), purified products treated with proteinase (D) by TLS polymerases η (black) and κ (grey) for different time.

[22]. We speculated that 5fU based DPCs in the consensus binding sequence of promoter element would modulate the gene transcription by influence the TFs' binding ability.

Sequences of the AP-1 and NF- κ B oligonucleotides containing AP-1 or NF- κ B transcription factor recognized consensus binding sites (in bold) were shown in Table S1 (Supporting information) except that the 5fU modification was at a different thymine in each substrate. Considering there may be a large steric hindrance caused by 5fU-DNA or peptide, we chose the 10-mer length of peptide to react with 5fU modified DNA, AP-1-5fU, AP-2-5fU, NF- κ B-5fU to get products AP-1-5fU-10pep, AP-2-5fU-10pep and NF- κ B-5fU-10pep, respectively. We used 20% native PAGE to measure the degree that whether these single strands were annealed into duplex strands with DNA AP-Complex or NF- κ B-Complex. As shown in Fig. S7 (Supporting information), compared with Lane 4, these upper bands in Lanes 1–3 and Lane 5 in the native polyacrylamide gel electrophoresis proved AP-1-5fU-10pep, AP-2-5fU-10pep and NF- κ B-5fU-10pep were annealed into double strands successfully. Then, end-labelled duplexes were incubated with human recombinant c-jun protein based on previous electrophoretic mobility shift assay [23]. For transcription factor AP-1 (Fig. 4A and Fig. S8A in Supporting information), our native PAGE and gray scale analysis demonstrated that the protein/oligonucleotide complexes resulted by duplex strand containing 5fU-peptide crosslinks was almost none, while we could observe relatively distinct complexes from the sample treated with unmutated double sequences. To explore whether the site of the 5fU-peptide crosslink residue was determinative for inhibition of transcription factor binding, AP-2-5fU-10pep oligonucleotides were used and similar results were obtained (Fig. 4B and Fig. S8B in Supporting information). Proving that 5fU based DPCs could cause the inhibition of transcription factor binding and this phenomenon was sequence independent. Besides, this phenomenon was reappeared with 8 μ g nuclear extracts of HeLa cells. Considering the interference of nonspecific binding, 200-fold unlabeled duplex strands as competitor were added into the reaction mixture and same essential results were achieved (Figs. S8C and D, S9A and B in Supporting information). To further verify this conjecture, we also explored another transcription factor NF- κ B (Fig. 4C and Fig. S8E in Supporting information). As expected, there was also no apparent DNA-protein interaction products for NF- κ B-5fU-10pep duplex strands, while for the undamaged double-stranded DNA, partial binding of NF- κ B transcription factor could be observed. By the

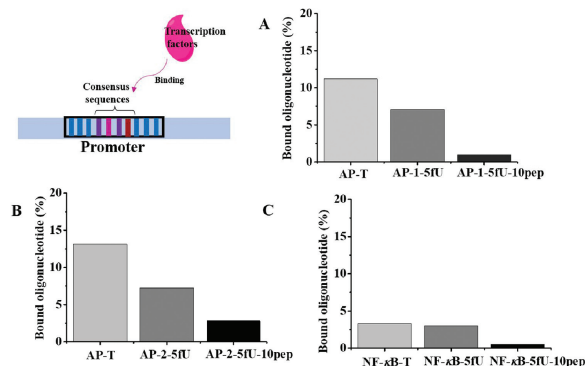


Fig. 4. Binding inhibition effects of 5fU based DPCs damage in promoter elements on transcription factor AP-1 or NF- κ B. Histogram analysis of the gray scale from native PAGE with binding efficiency by consensus binding sites and transcription factor AP-1 or NF- κ B. (A) DNA with different modifications (AGCGGTTXGACTCAGG-GGATT, X = T, 5fU or 5fU-10 pep) incubated with human recombinant c-jun protein. (B) DNA (AGCGGTTGACXC-ACGGGGATT, X = T, 5fU or 5fU-10 pep) was binding with human recombinant c-jun protein (B). (C), Nuclear extracts incubated with DNA (AGTTGAGGGACXITCCCAGGC, X = T, 5fU or 5fU-10 pep).

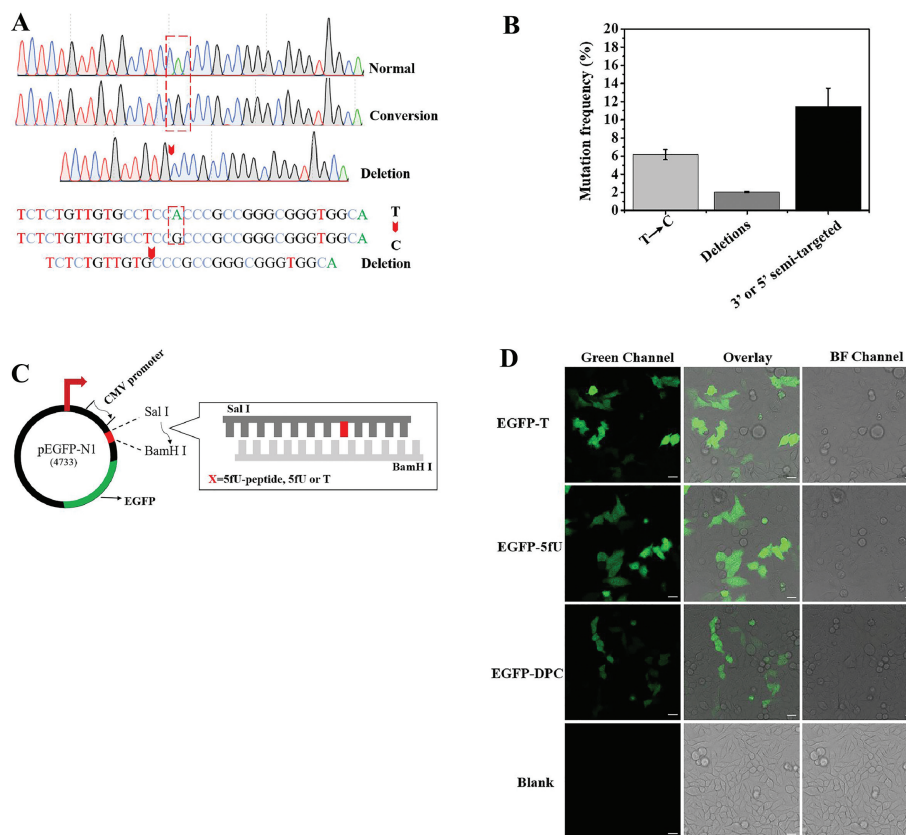


Fig. 5. 5fU based conjugates could cause errors in DNA replication and interference the translation *in vivo*. (A) Typical Sanger sequencing results of TA cloning assay with 80 bp 5fU-peptide crosslinks duplex. (B) Types and frequencies of mutations induced by the 80 bp 5fU-10-mer peptide cross-links in *Escherichia coli*. The experiment was performed thrice with 50 monoclonal colonies chosen for sequencing each time. (C) Scheme of the site-specific 5fU-containing DPCs plasmids for EGFP reporter assay. X referred to a thymine residue, 5fU residue or 5fU-peptide cross-links at the template strand. (D) Fluorescent images of HeLa cells for EGFP expression. Each experiment was independently repeated for three times. The first column was the EGFP fluorescent images; the middle column was the merging of the fluorescent images and bright field images; the third column was the bright field images.

way, we found that 5fU was different from 8-oxodG that a single 8-oxodG was sufficient to inhibit transcription factor binding to AP-1 and it had no effect on binding to NF- κ B [23]. A single mutation of 5fU at transcription factors AP-1 or NF- κ B recognized conserved sequences only partly or substantially uninhibited binding to TFs.

To further investigate the biological impacts of 5fU based DPCs *in vivo*; we attempted to incorporate 5fU-peptide conjugates into plasmids. These plasmids were transfected into prokaryotic or eukaryotic cells to research the influence of 5fU-based DPCs on replication or translation. In order to determine the DNA replication situation accurately *in vivo*, we employed the TA cloning approach with 80 bp length of double-stranded DNA, in which each of strand contained one 5fU site. After the reaction between the 5fU-DNA and 10-mer or 31-mer length of peptide, we ligated these products into vector and analyzed the DNA replication in prokaryotic cells by Sanger sequencing. Just as expected, about 20%–29% progeny plasmids contained mutants, including about 6% T → C conversions, 2%–4% deletions, 12%–18% 3' or 5' semi-targeted mutations after the 80 bp duplex DNA being treated with 10 or 31-mer length of peptides (Figs. 5A and B, Fig. S10 in Supporting information), which was consistent with the previous *in vitro* experimental results (Fig. 2). These data proved that 5fU based DPCs could induce a series mutation *in vitro* and *in vivo*.

In order to comprehend the 5fU-based DPCs' biological effect on gene expression, we constructed DNA-peptide cross-links into the enhanced green fluorescent protein (EGFP) reporter containing plasmid at 5'-UTR as shown in Fig. 5C. As illustrated in Fig. 5D of

the confocal images obtained from HeLa cells, compared with 5fU-peptide cross-links plasmids, we observed brighter green fluorescent signals from the groups of either unmutated or 5fU contained plasmids (positive controls). The similar results could also be received from flow cytometry results. The fluorescence intensity of the experimental group (cells treated by EGFP-DPC) was lower than that of the control group (cells treated by EGFP-T or EGFP-5fU) (Fig. S11 in Supporting information). Both the confocal images and flow cytometry results proved that 5fU based DPCs could disturb gene expression in living cells, which were consistent with our previous conjecture.

In summary, the current study firstly explored that the reversible 5fU-based DNA-peptide or protein crosslinks had profound impacts on gene replication, transcription even translation. Future researches were needed to identify the exact locations of endogenous 5fU-based DPCs as well as their repair and processing pathways at the cellular level. In all, this study opened up a broader view in the field of DNA-protein interactions, especially for the inherent DNA modifications. We also believed that this research would enrich our understanding of 5fU and therefore encourage us for further investigating the 5fU related biological processes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank National Natural Science Foundation of China for their financial support (Nos. 91753201, 21721005).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.ccllet.2021.02.036>.

References

- [1] C. Luo, P. Hajkova, J.R. Ecker, *Science* 361 (2018) 1336–1340.
- [2] A. Hofer, Z.J. Liu, S. Balasubramanian, *J. Am. Chem. Soc.* 141 (2019) 6420–6429.
- [3] M.J. Booth, E.A. Raiber, S. Balasubramanian, *Chem. Rev.* 115 (2015) 2240–2254.
- [4] T. Carell, M.Q. Kurz, M. Müller, et al., *Angew. Chem. Int. Ed.* 57 (2018) 4296–4312.
- [5] W. Hirose, K. Sato, A. Matsuda, *Angew. Chem. Int. Ed.* 49 (2010) 8392–8394.
- [6] C. Liu, Y. Chen, Y. Wang, et al., *Nano Res.* 10 (2017) 2449–2458.
- [7] H.P. Jiang, T. Liu, N. Guo, et al., *Anal. Chim. Acta* 981 (2017) 1–10.
- [8] N.Y. Tretyakova, A. Groehler, S. Ji, *Acc. Chem. Res.* 48 (2015) 1631–1644.
- [9] T. Nakano, H. Terato, K. Asagoshi, et al., *J. Biol. Chem.* 278 (2003) 25264–25272.
- [10] H. Ide, M.I. Shoulkamy, T. Nakano, et al., *Mutat. Res. Fund. Mol. Mech. Mutagen.* 711 (2011) 113–122.
- [11] A.M. Sanchez, I.G. Minko, A.J. Kurtz, et al., *Chem. Res. Toxicol.* 16 (2003) 1019–1028.
- [12] M. Dizdaroglu, E. Gajewski, *Cancer Res.* 49 (1989) 3463–3467.
- [13] P. Pande, S. Ji, S. Mukherjee, et al., *Chem. Res. Toxicol.* 30 (2017) 669–677.
- [14] S. Wickramaratne, E.J. Boldry, C. Buehler, et al., *J. Biol. Chem.* 290 (2015) 775–787.
- [15] E.A. Raiber, G. Portella, S. Martínez Cuesta, et al., *Nat. Chem.* 10 (2018) 1258–1266.
- [16] Z.J. Liu, S. Martínez Cuesta, P. van Delft, et al., *Nat. Chem.* 11 (2019) 629–637.
- [17] R.E. Hardisty, F. Kawasaki, A.B. Sahakyan, et al., *J. Am. Chem. Soc.* 137 (2015) 9270–9272.
- [18] G. Zou, C. Liu, W. Zeng, et al., *CCS Chem.* 2 (2020) 54–63.
- [19] S. Ji, H. Shao, Q. Han, et al., *Angew. Chem. Int. Ed.* 56 (2017) 14130–14134.
- [20] J.B. Stafford, R.L. Eoff, A. Kozekova, et al., *Biochemistry* 48 (2009) 471–480.
- [21] C.K. Sen, L. Packer, *FASEB J.* 10 (1996) 709–720.
- [22] H. Sies, *Exp. Physiol.* 82 (1997) 291–295.
- [23] R. Ghosh, D.L. Mitchell, *Nucleic Acids Res.* 27 (1999) 3213–3218.