Synthesis, Characterization and Biological Consequences of DNA-protein Cross-links

Shaofei Ji and Natalia Tretyakova

Department of Chemistry, Department of Medicinal Chemistry, University of Minnesota
Minneapolis, Minnesota 55455

DNA-protein cross-links (DPCs) commonly form upon exposure to various endogenous, environmental and chemotherapeutic agents such as nitrogen mustards, platinum compounds and mitomycin C. Because of their considerable size and helix-distorting nature, DPCs can interfere with the progression of replication and transcription machineries leading to toxicity and mutations.

The goal of present work is to construct DNA substrates containing model DPCs structurally analogous to lesions found in cells and to elucidate their biological outcomes using biochemical and cellular assays. In this work, we recently find that 5fC bases in DNA readily form Schiff base conjugates with Lys or Arg side chains of nuclear proteins, forming covalent DNA-protein conjugates. These covalent protein-DNA interactions are reversible, suggesting that they may play a role in transcriptional regulation and chromatin remodeling. We also have studied the *in vitro* mutagenicity of the 5fC induced DNA-protein or DNA-peptide cross-links on DNA replication and transcription using recently developed *in vitro* adducts bypass assays. We found that DNA-protein cross-links completely block DNA and RNA polymerase *in vitro*. In contract, short DNA-peptide cross-links can be bypassed by TLS polymerase and T7 RNA polymerase with reduced efficiency. LC-MS/MS based strategies showed that the replication bypass of 11-mer peptide cross-linked to 5-methyl-dC by TLS polymerases induced large numbers of C to T mutations and deletions. We also employed RT-PCR and LC-MS/MS base strategy to investigate the transcriptional mutagenesis of the DNA-peptide cross-links by T7 RNA polymerase. Our result showed that DNA-peptide cross-links (11-mer, 31-mer and 57-mer) can also be bypassed by T7 RNA polymerase with reduced efficiency, resulting a large number of C to T mutations.