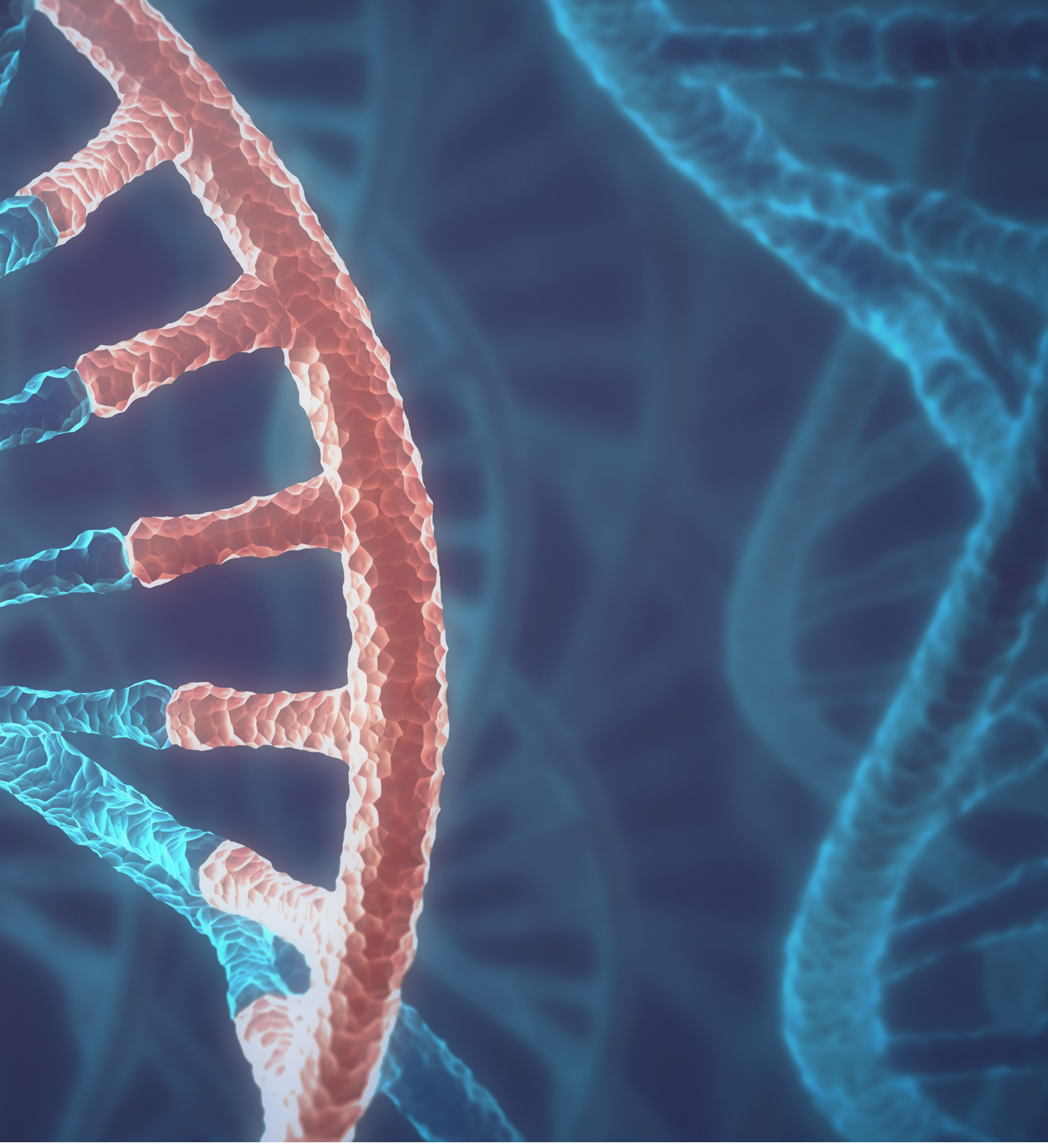


Convertible Nucleosides

To Detect Polymerase Bypass In DNA Lesions

GLEN RESEARCH

part of Maravai LifeSciences

Glen Research provides several unique, base-modified 2'-deoxynucleoside monomers for site specific modification of oligonucleotides. These synthons are known as Convertible Nucleosides, and the process for use in oligo modification is referred to as "Post-Synthetic Substitution".[1] Since their introduction they have proven useful in physical chemical studies and in studies of DNA replication, error and mutation potential. Here we provide an introductory overview of these interesting molecules and a brief update of the key chemistry milestones since their introduction. We also review a recent, thought-provoking research application using convertible deoxyguanosine (dG) that takes advantage of site-specific preparation of a DNA lesion to investigate translesion synthesis (TLS) properties of mammalian cell polymerases and the replication accuracy of two of the Y-family DNA polymerases, κ (kappa) and ι (iota).

Background

Over time, experience gained with customers has led us to refine our convertible nucleoside monomer offerings to the highly reliable molecules shown in Figure 1: TMP-F-dU-CE Phosphoramidite (convertible F-dC), O4-Triazolyl-dU-CE Phosphoramidite (convertible dU), O6-Phenyl-dI-CE Phosphoramidite (convertible dA), and 2-F-dI-CE Phosphoramidite (convertible dG). These convertible names are indicative of the final nucleoside structure achieved after synthesis and, where applicable, subsequent post-synthetic modification of the convertible nucleoside.

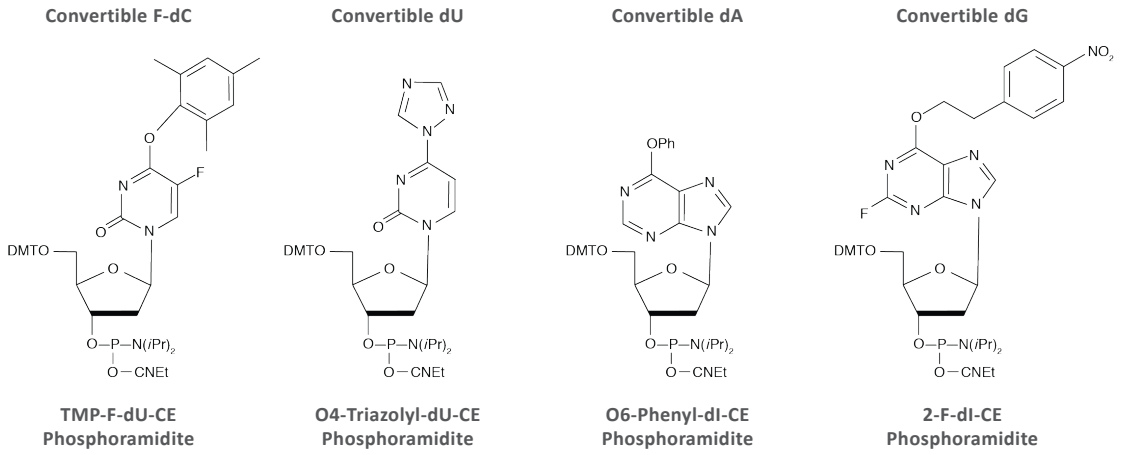


Figure 1. Convertible nucleoside monomers

Convertible F-dC readily permits preparation of a 5-Fluoro-dC residue without the need for additional post-synthetic nucleoside substitution since normal ammonium hydroxide treatment effects the conversion of F-dU to F-dC during workup.[2] In contrast, the special utility of the other three convertibles is flexibility to change ring substituents and enable ready access to a variety of base-modified structures useful in research. For example, convertible dU allows post-synthetic replacement of the O4-triazolyl group with nucleophilic substituents at the 4-position of the pyrimidine ring, including alkoxy-, amino-, or sulfur-based substituents. Convertible dA can be substituted on the adenosine ring 6-position using alkyl-amines to displace the O6-phenoxy group.[3] Convertible dG is of special interest because it facilitates substitutions at the 2-position on the dG ring.[4,5,6] The dG N^2-amino group is long known to be susceptible to alkylation in vivo by exogenous compounds, including mutagenic substances implicated in cancer.[7]

DNA Polymerases and Mutagenesis

DNA is susceptible to damage by high-energy electromagnetic radiation, radioactive isotopes, and endogenous and exogenous chemical agents.[8] When DNA damage has occurred and remains unrepaired there is the potential for errors in DNA replication, transcription, and translation which can have cytotoxic and mutagenic effects. Research has identified several pathways within cells that correct DNA damage and errors in DNA replication, including polymerase 3'-5'-exonuclease activity, base excision repair (BER), mismatch repair (MMR), and photodimer repair, among others.[9]

DNA lesions unrepaired by the above pathways possess two major mechanisms that permit cells to tolerate unrepaired DNA lesions: homologous recombination and translesion synthesis (TLS).[10] While TLS will be highlighted below, extensive research also has elucidated the variety of discrete DNA polymerases used in DNA replication with activities and DNA lesion tolerances that differ among prokaryotes and eukaryotes. The nature and activities of polymerases across species, typically grouped by sequence similarity, is beyond the scope of this note.

Nonetheless, recently the Y-family of DNA polymerases, present in prokaryotes and eukaryotes, was revealed to significantly facilitate DNA replication across certain DNA lesions, where modifications of DNA were caused by chemical agents that covalently modify the standard nucleoside bases. These low-fidelity Y-family polymerases are specialized to bypass DNA lesions through TLS, which appears to be a major mechanism of DNA damage tolerance. The TLS that facilitates DNA replication across DNA lesions permits cell survival but at the cell's expense, with replication errors dependent on TLS polymerase fidelity.[11] The frequency at which polymerase adds the wrong nucleotide opposite a DNA lesion is an important factor impacting spontaneous mutations. Thus, there has been much recent interest in the synthesis properties of the Y-family polymerase kappa (Polk) and polymerase iota (Poli) members.

DNA Lesion Bypass Investigation

The Yinsheng Wang Lab (Institute for Integrative Genome Biology, University of California, Riverside) is investigating DNA lesions and their impact on DNA replication and mutation potential with a major focus on how the mammalian polymerases process these lesions. His team has used convertible dG to develop an efficient system to introduce site-specific N^2-carboxymethyl-dG lesions in DNA constructs.[12] A 2011 paper describes the method workflow and investigates the polymerases in mammalian cells required for the error-free TLS bypass of N^2-carboxyethyl-2'-deoxyguanosine (N^2-CEdG) and N2-carboxymethyl-2'-deoxyguanosine (N^2-CMdG) lesions (Figure 2).[13] The authors described extensive analysis of polymerase activity in human fibroblast cells and several mouse embryonic fibroblast (MEF) cell types. Here we focus on lesion bypass studies in three MEF cell types: wild-type (WT), Polk-deficient cells (Polk-) and Poli-deficient cells (Poli-).

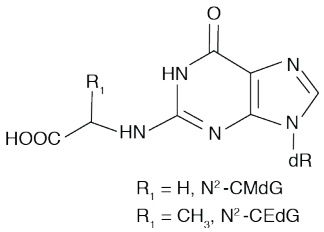


Figure 2. Carboxymethyl- and carboxyethyl- lesions at N2 of dG

N2-CEdG is the major stable DNA adduct formed from endogenous methylglyoxal (MG), a glycolytic metabolite. It is found in urine from healthy human subjects and observed more frequently in kidney and aortic cells of diabetic and uremic patients but is also detected in cultured human cells.[14] MG is mutagenic in vivo in bacterial cells and cultured human cells. Exposure to MG also comes via exogenous sources including cigarette smoke, soy sauce, coffee, and whiskey. In vivo, MG binds irreversibly to proteins, primarily on arginine (Arg) residues, and binds to DNA and other substrates with significant pathological impacts. Recently MG was found to be an abundant histone modification, similar in frequency to Arg methylation.[15] Note that N2-CEdG forms as two diastereomers with either an R- or S-enantiomeric center at the N2 carbon and that glyoxal (an exogenous agent similar to endogenous MG) reacts with dG to form the closely related N2-CMdG, a pair of adducts that allow the impact on TLS activity to be probed by quite subtly altered dG lesions (where a hydrogen is replaced by an methyl group). Although these two dG modifications are, chemically speaking, small compared to many other exogenous, mutation-promoting lesions, their stability and widespread occurrence raises a question about their intrinsic potential to cause mutations resulting from in vivo replication errors. Wang's research team investigated the details of replication accuracy using DNA constructs containing a single dG lesion with the convertible nucleoside approach to prepare plasmids as described below.

Oligodeoxynucleotide (ODN) Synthesis (References 13,14)

Preparation of lesion-modified plasmids using the convertible dG nucleoside method starts with synthesis of 17-mer ODN 5'-GCG CAA AXC TAG AGC TC-3' using the 2-fluoro-2'-deoxyinosine–modified phosphoramidite[16] as convertible dG in position X (Figure 3). To make N^2-CEdG, the ODN was cleaved from the support, partially deprotected with 0.1 M NaOH, and neutralized. The dried residue was mixed with diisopropylethylamine (DIEA) in excess, L- or D-alanine (45 equivalents) and DMSO and then stirred at 55°C for 48–72 h. To synthesize less bulky N^2-CMdG, glycine was used to displace the fluorine. After drying, further base deprotection with concentrated NH_4OH was carried out at 55°C for 6-8 h. The resulting crude O6-trimethylsilylethyl protected oligo was isolated, treated with 5% acetic acid, neutralized, and purified by reversed-phase HPLC to yield the fully-deprotected ODN.

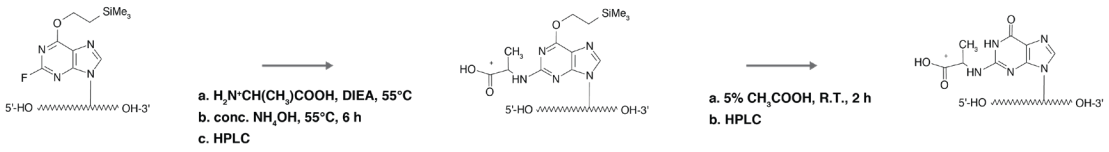


Figure 3. Post oligomerization incorporation of N2-dG modifications

Electrospray ionization mass spectrometry (ESI-MS) showed the presence of an N^2-CEdG at the expected 72 Da higher than unmodified 17-mer. Capillary HPLC ESI-MS/MS product ion spectrum of the [M – 4H]4– ion (m/z 960.6) exhibited multiple peak patterns consistent with the presence of an N^2-CEdG at the 7-position of the ODN.

Preparation of Site-Specific, Lesion-Bearing Plasmids

Key to Wang's in vivo method of TLS analysis was the use of recombinant DNA methods and the lesion-containing ODN to create a single lesion–containing double-stranded DNA (ds DNA) to be used for replication within mammalian cells. Plasmids were prepared and isolated as summarized below (Figure 4). The group used shuttle vectors and single-strand nicking endonucleases (nickases) to prepare a gapped vector and then inserted site-specific dG lesion-containing ODN into bacterial shuttle plasmids for replication (Ref. 13). Subsequently these plasmids were used in mammalian cells where the resulting replication products could be assessed for the identity and number of mutations at the lesion position on the original plasmid.

Thus, shuttling plasmids pTGFP-Hha10 with lesion modifications and an SV40 origin were replicated in the SV40-transformed human 293T and POLH-deficient cells. POLH expresses Pol η, also known as DNA polymerase eta, one of the four specialized human Y-family polymerases and POLH- cells are useful for studying the DNA repair activities of the other specialized polymerases.[17] Note that Wang's team also inserted the (S)-N^2-CEdG, (R)-N^2-CEdG, and N^2-CMdG, or dG-control containing ODNs into pMTEX4 to be replicated in MEF cells.

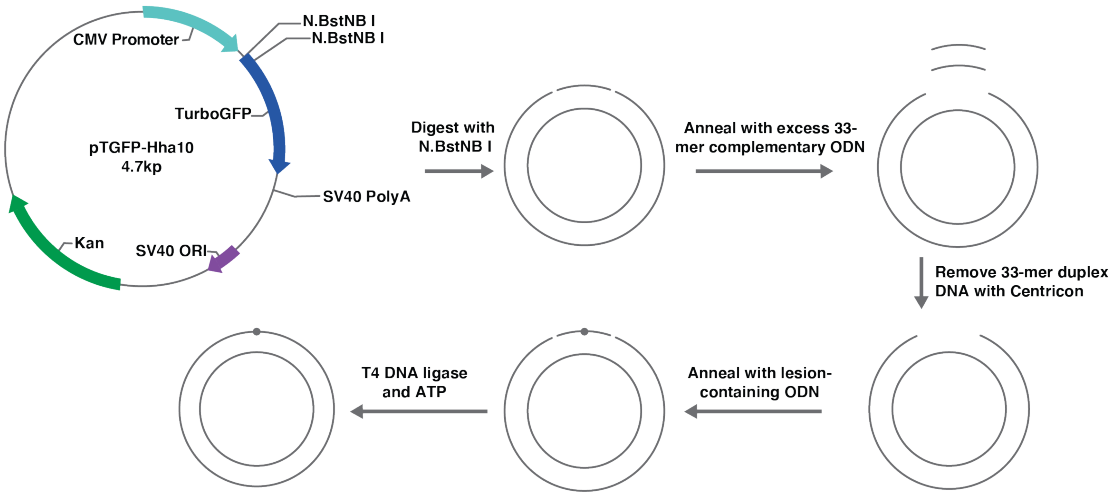


Figure 4. Preparation of lesion-containing pTGFP-Hha10

Evaluation of Replication Bypass Efficiency and Mutation Frequency

To assess the roles of various TLS DNA polymerases in replication accuracy, WT, Polk-, and Poli- MEF cells were compared for their replication of lesion-containing and control vectors. This involved PCR amplification of the replicated vector, excising the small section containing the lesion site, and analyzing the replication efficiency of top- and bottom-strand products, and visualizing the presence of mutations by polyacrylamide gel electrophoresis (PAGE).

The PCR amplification and digestion into fragments readily separated by PAGE analysis (Figure 5) shows the isolated vector amplified by PCR as a 3,950-bp product. After purification, the amplicon was digested with SacI and alkaline phosphatase (shrimp). The 5'-termini were labeled with γ-^{32}P-ATP, and the resulting DNA fragments were further digested with FspI. This process yields ds 10-mer fragments, where AC and TN refer to the two 10-mer single-stranded DNA products formed from the respective replication of the top- and bottom-strands of the original plasmid where N is potentially any base incorporated by inaccurate replication across the DNA lesion. Accurate replication, of course, yields d(CAAGCTTTGC) (10-mer-AC) and d(CTAGCTTTGC) (10-mer-TC).

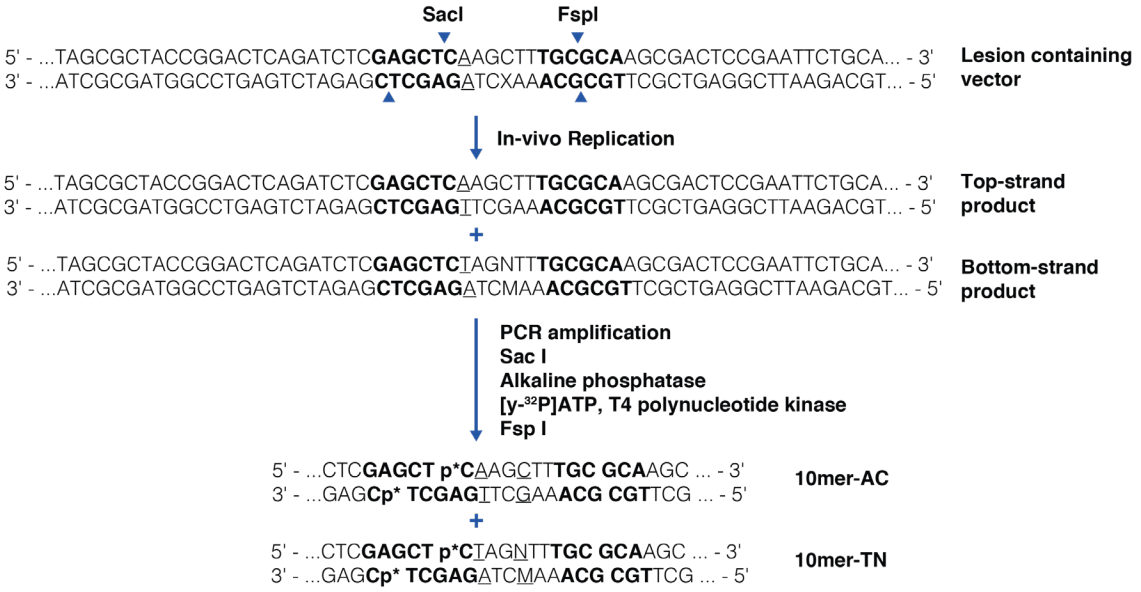


Figure 5. Amplicon preparation, fragmentation and labeling

At last, fragments are ready for polymerase lesion bypass analysis. The short fragments from cultured MEF-cell plasmid replications were analyzed by 30% PAGE (Figure 6) allowing high resolution of the fragment mixture. In each analysis of either WT, Polk-/- deficient, and Poli-/- deficient cells, the first 5 lanes contain labeled synthetic oligonucleotide standards containing the indicated two nucleotides at the mismatch position and the lesion position, respectively. The next four lanes reveal the fragment bands evident in mouse cell replication of control plasmid, and then the lesion containing plasmids containing the (R)-N^2-carboxyethyl-dG and (S)-N^2-carboxyethyl-dG and N^2-carboxymethyl-dG in lanes 6, 7, 8 and 9, respectively.

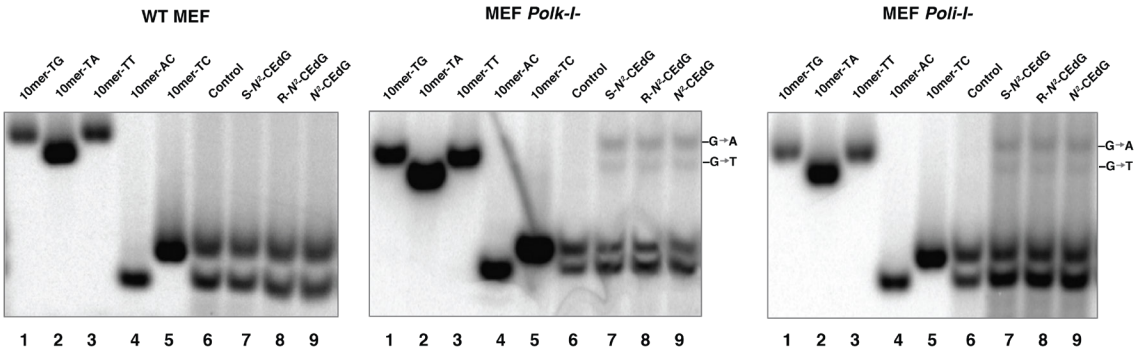


Figure 6 PAGE analysis of plasmid fragment

WT MEF cells replicated control and lesion-containing plasmids (left PAGE) accurately, generating the expected 10-mer AC and 10-mer-TC bands (lanes 4 and 5) with no extraneous, additional fragments present. In each case the two rapidly migrating fragments showed similar intensities, indicating equivalent strand amplification efficiency. Although the MEF Polk-/- cells (middle PAGE) similarly

replicated control plasmid and generated equivalent amounts of the two expected strands, when lesion dG–containing plasmids were replicated, there were two slower-migrating fragment products corresponding to the migration positions of the 10-mer-TG,-TA, and-AC synthetic ODNs. The slower fragments did not form in equivalent amounts: the slowest migrating fragment was more intense than the next faster migrating fragment. In comparison, the MEF Poli-/- plasmid replication products (right PAGE) show similar PAGE results, with slower migrating products in unequal amounts.

To identify the mutations present, fragment mixtures were prepared without radiolabel by treating the 238-bp PCR products simultaneously with SacI, FspI, and shrimp alkaline phosphatase followed by extraction (phenol/chloroform/isoamyl alcohol) and the aqueous portion was dried, resuspended in water and subjected to LC-MS/MS. Quantitative analysis found that replication of S-N^2-CEdG-, R-N^2-CEdG-, and N^2-CMdG-containing genomes in Polk-/- cells resulted in G A transition at frequencies of about To identify the mutations present, fragment mixtures were prepared without radiolabel by treating the 238-bp PCR products simultaneously with SacI, FspI, and shrimp alkaline phosphatase followed by extraction (phenol/chloroform/isoamyl alcohol) and the aqueous portion was dried, resuspended in water and subjected to LC-MS/MS. Quantitative analysis found that replication of S-N^2-CEdG-, R-N^2-CEdG-, and N^2-CMdG-containing genomes in Polk-/- cells resulted in G A transition at frequencies of about 23% and G T transversion at frequencies of about 15%, respectively. Mutational frequencies were slightly lower for the Poli-/- cells.

To identify the mutations present, fragment mixtures were prepared without radiolabel by treating the 238-bp PCR products simultaneously with SacI, FspI, and shrimp alkaline phosphatase followed by extraction (phenol/chloroform/isoamyl alcohol) and the aqueous portion was dried, resuspended in water and subjected to LC-MS/MS. Quantitative analysis found that replication of S-N^2-CEdG-, R-N^2-CEdG-, and N^2-CMdG-containing genomes in Polk-/- cells resulted in G A transition at frequencies of about 23% and G T transversion at frequencies of about 15%, respectively. Mutational frequencies were slightly lower for the Poli-/- cells.

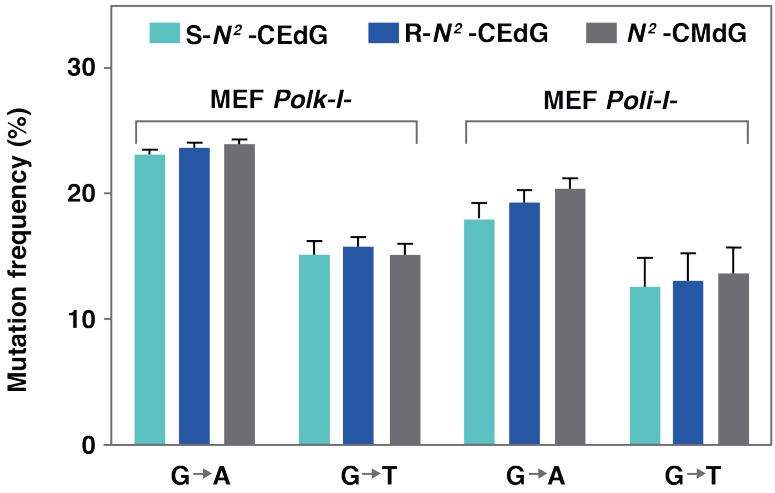
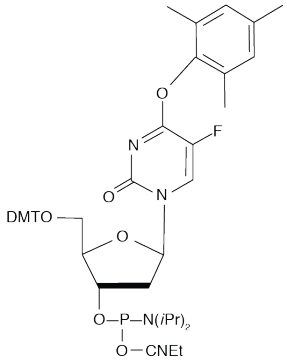


Figure 7. Mutation frequencies in Polk-/- and Poli-/- MEF cells

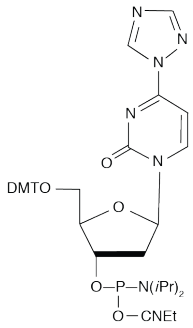
The authors concluded that in mammalian cells the polymerase κ and polymerase ι-mediated nucleotide incorporation opposite N^2-CEdG and N^2-CMdG is error-free. Results also supported that polymerase ι, a Y-family polymerase present only in higher eukaryotes, accurately bypasses these lesions.

TMP-F-dU-CE Phosphoramidite



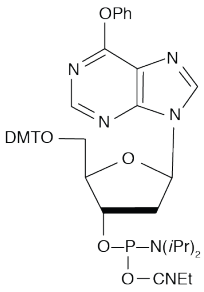
TMP-F-dU is used to introduce F-dC into oligonucleotides The conversion from F-dU to F-dC occurs during deprotection with ammonia. F-dC influences DNA structure and inhibits methyltransferases.

O4-Triazolyl-dU-CE Phosphoramidite



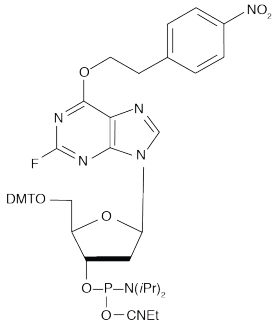
O4-Triazolyl-dU can be used to introduce a variety of modifications at the convertible position, including N, O and S modifications.

O6-Phenyl-dI-CE Phosphoramidite



After incorporation into an oligo, reaction of the O6-phenyl on the inosine base with a primary amine displaces the oxygen atom, and converts the nucleotide into a N6-substituted dA.

2-F-dI-CE Phosphoramidite



2-F-dI is a convertible nucleoside for the preparation of 2'-dG derivatives following the displacement of the 2-fluorine by primary amines.

Table 1. Ordering Information

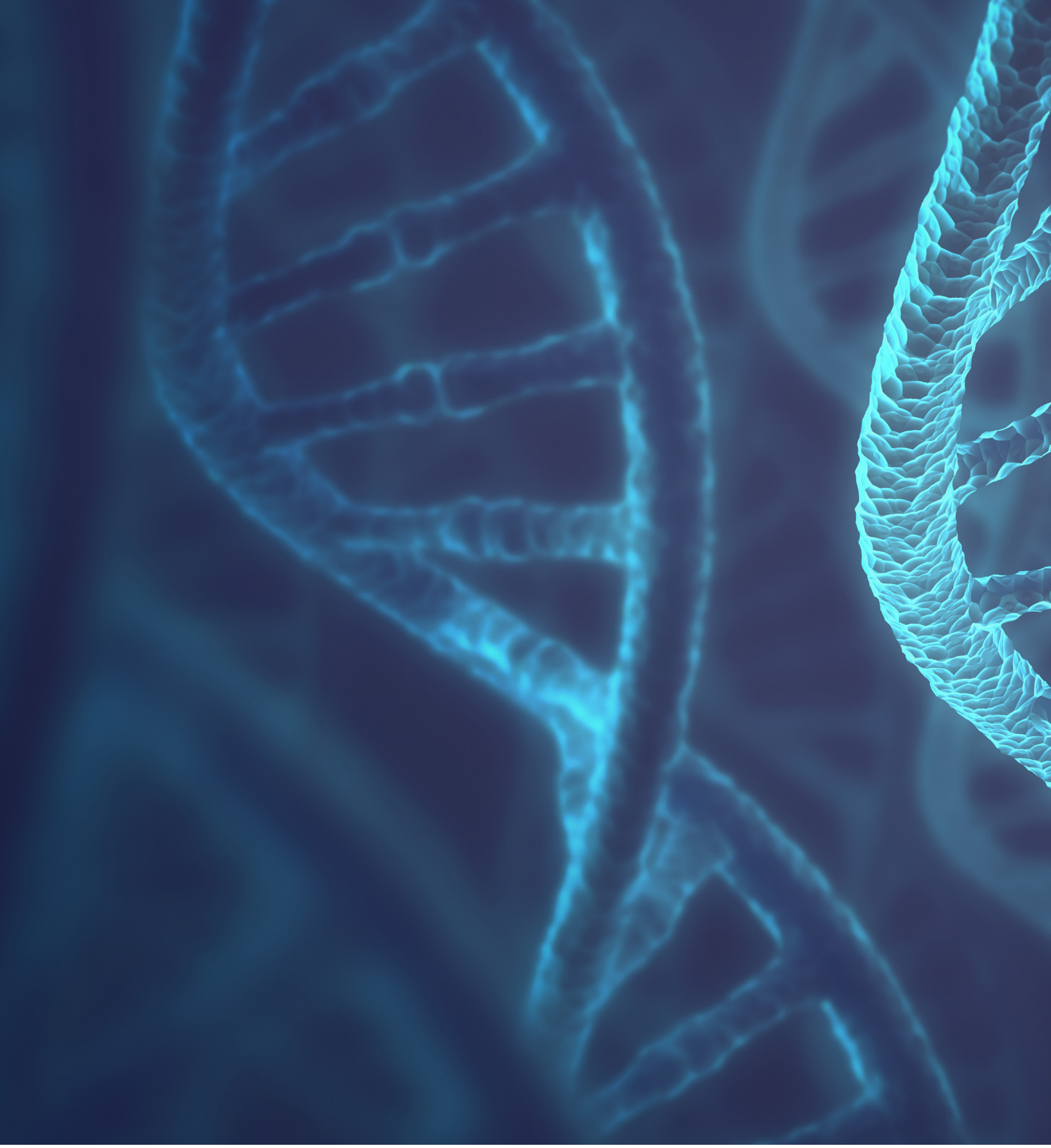
Item	Pack Size	Catalog No.
TMP-F-dU-CE Phosphoramidite	0.25g	10-1016-02
	100µmol	10-1016-90
O4-Triazolyl-dU-CE Phosphoramidite	0.25g	10-1051-02
	100µmol	10-1051-90
O6-Phenyl-dI-CE Phosphoramidite	0.25g	10-1042-02
	100µmol	10-1042-90
2-F-dI-CE Phosphoramidite	0.25g	10-1082-02
	100µmol	10-1082-90
	50µmol	10-1082-95

Conclusions

Glen Research provides convertible dG as the protected deoxyinosine monomer 2-F-dI-CE amidite suitable for post-oligomerization synthesis of lesion-containing, ds DNA. Using this monomer, a wide array of N^2-dG lesion structures is readily accessible. O4 Triazolyl-dU and O6-Phenyl-dI monomers also are available to help facilitate various site-specific modifications of dU at C4 and dA at C6 of the pyrimidine ring.[18] Additional information about these post-synthesis modifications is available on our website.

References

1. Glen Report **6.1**, *Post-Synthetic Substitution – Convertible Nucleosides*
2. Glen Report **6.26**, *New Minor Bases*
3. Glen Report **7.15**, *New Minor Bases*
4. Glen Report **15.14**, *More and Still More Novel Minor Bases*
5. Glen Report **16.13**, *Minor Base and Related Novel Phosphoramidites*
6. Technical Bulletin, **TB_2-F-dI-1**, *Use of 2-F-dI to Produce N2-Modified dG Derivatives*
7. A. Dipple, *Carcinogenesis*, 1995, **16**, 437-441
8. T. Lindahl, *Nature*, 1993, **362**, 709-715
9. E.C. Friedberg, G.C. Walker, W. Siede, R.D. Wood, R.A. Schultz, T. Ellenberger, *DNA Repair and Mutagenesis*, Second edition. 2006, ASM Press, Washington, D.C.
10. P.P. Ghodke, P.I. Pradeepkumar, *Eur. J. Org. Chem*., 2020, **44**, 6831-6844
11. J.E. Sale, A.R. Lehmann, R. Woodgate, *Nat. Rev. Mol. Cell Biol*., 2012, **13**, 141–152
12. C. You, Y. Wang, *Nat. Protoc.* 2015, **10**, 1389-1406
13. B. Yuan, C. You, N. Andersen, Y. Jiang, M. Moriya, T.R. O'Connor, Y. Wang, *J. Biol. Chem.*, 2011, **286**, 17503-17511
14. H. Cao, Y. Jiang, Y. Wang, *J. Am. Chem. Soc.*, 2007, **129**, 12123-12130
15. J.J. Galligan, J.A. Wepy, M.D. Streeter, P.J. Kingsley, M.M. Mitchener, O.R. Wauchope, W.N. Beavers, K.L. Rose, T. Wang, D.A. Spiegel, and L.J. Marnetta, *Proc. Natl. Acad. Sci. U.S.A.*, 2018, **115**, 9228-9233
16. Conditions for F-substitution and subsequent TMS removal used by these researchers differ from reaction conditions found optimal using the p-nitroethyl protected convertible dG sold by Glen Research. Both of these convertible G structures permit efficient replacement of F on the incipient dG-residue.
17. R. Barnes and K. Eckert, *Genes*, 2017, **8**, 1-19
18. N. Andersen, P. Wang, Y. Wang, *Chem. Res. Toxicol.*, 2013, **26**, 1730-1738



22825 Davis Drive | Sterling, Virginia 20164
703-437-6191 | support@glenresearch.com | glenresearch.com

GLEN RESEARCH
part of **Maravai LifeSciences**