Volume 12 Number 24 1984 Nucleic Acids Research

The complete pattern of mutagenesis arising from the repair of site-specific psoralen crosslinks: analysis by oligonucleotide hybridization

Wilma A.Saffran and Charles R.Cantor

Department of Human Genetics and Development, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA

Received 14 November 1984; Accepted 23 November 1984

ABSTRACT

Psoralen crosslinks were site-specifically placed in plasmid pBR322 near the BamHI site in the tet gene by enzymatically inserting mercurated nucleotides and reacting at the target site with a sulfhydryl-containing psoralen. The damaged plasmid was repaired in SOS-induced E. coli cells. Mutants were detected by colony hybridization to oligonucleotides in the target region, and their sequences were determined. The mutations are all base substitutions, 80% transitions and 20% transversions, similar to the mutations previously identified by the loss of tetracycline resistance. However, the mutation sites detected by a physical method, unconstrained by phenotypic changes, follow a broader distribution than those identified genetically. They occur primarily at favored psoralen crosslinking sites, where T-T and T-C interstrand crosslinks can be formed. A majority of these mutations are silent.

INTRODUCTION

Different DNA damaging treatments result in characteristic patterns of mutation, which have been examined to gain information on the mechanisms of DNA repair and mutagenesis (1). UV irradiation results in single base substitutions at the sites of adjacent pyrimidines, (2-6). Damage with 4-nitroquinoline oxide produces mainly transitions at GC base pairs (2) while aflatoxin and benzo[a]pyrene cause GC to TA transversions (7, 8).

Mutants are generally identified by genetic methods: reversions of known mutations, appearance of amber or ochre mutations at well-characterized sites, or easily screened forward mutations. These methods all detect a subset of the total mutations, those that impair the function of the gene being studied. Mutations that do not affect the amino acid

sequence, result in conservative changes, or occur in non-essential portions of the gene are not measured.

Hybridization with oligonucleotides is a physical technique which is sensitive to even single base mismatches between the probe and sample DNA (9, 10). It has been widely used to identify the products of *in vitro* oligonucleotide-directed mutagenesis. Here we show that this method has the ability to identify mutations arising from the repair of DNA damage *in vivo*. This method has the advantage of detecting the complete spectrum of mutations within a localized region, including those which are genetically silent.

We previously studied the mutagenic repair in *E. coli* of psoralen damage placed site-specifically *in vitro* in the *tet* gene of plasmid pBR322 (11). Psoralen is a bifunctional reagent which, upon near UV irradiation, forms covalent bonds with pyrimidines, chiefly thymines (12). This results in the formation of monoadducts and interstrand psoralen crosslinks. The known affinity of psoralen for -T-A- sites predicts that these should be hotspots for psoralen mutagenesis (13). However, the preferred crosslinking sites were under-represented in the tetracycline sensitive mutants we isolated, and half the target site appeared to be free of mutations. In the present study we have screened for mutants by hybridizing colonies to oligonucleotides in the target region. The mutations found by this technique are evenly distributed in the target region and occur primarily at psoralen crosslinking sites.

MATERIALS AND METHODS

Site-specifically crosslinked plasmid

Psoralen crosslinks were placed near the BamHI site of pBR322 (14) and were transfected into irradiated repair competent *E. coli* as previously described (11).

Colony hybridization

Colonies were transferred with toothpicks onto 82 mm nitrocellulose filters (Millipore HATF 082) and grown at 37°C. Two replicas were made of each filter. The filters were

treated with 0.5 N NaOH for 4 minutes, 1 M Tris-HCl, pH 7.5 for 4 minutes, then shaken in 0.5 M Tris-HCl, pH 7.5, 1.5 M NaCl for 10 minutes. They were placed in 2 x SSC, 0.1% SDS and the cell debris gently wiped off. After a 5 minute wash in 2 x SSC, the filters were air dried, then baked at 80°C in a vacuum oven for 2 hours. The filters were prehybridized in 10% dextran sulfate, 0.9 M NaCl, 0.18 M Tris-HCl, pH 8.0, 6 mM EDTA, 0.1% lauryl sarcosinate, 5 x Denhardt's, 100 µg/ml denatured salmon sperm DNA at 20°C for 1 hour (15).

Oligonucleotides were synthesized on an Applied Biosystems synthesizer and purified by preparative gel electrophoresis on a 14% sequencing gel.

The oligonucleotide probes were 5' end labelled with $\gamma[^{32}P]$ATP and kinase. Unincorporated ATP was removed by chromatography on DEAE cellulose (16). The final DE-52 column fractions, containing labelled probe, were added to the filters in the prehybridization mix described above. Hybridization was continued at 20°C for 20-24 hours.

The filters were washed with 6 x SSC at 20°C until the wash supernatant was non-radioactive, generally 4 washes, then exposed to XAR film to obtain a record of the hybridization. They were then washed once with 6 x SSC at increasing temperatures, at intervals of 5°C, and exposed to X-ray film in between the washes.

Colonies were counted as mutant if they hybridized strongly at 20°C, but weakly at elevated temperatures. The first weak spots appeared at 55°C.

Sequencing

Plasmids were prepared from 100 ml saturated cultures, grown in L broth, by the method of Marko et al., (17) up to the proteinase K digestion step. About ten µg of each were purified on low gelling temperature agarose. The plasmids were linearized with HindIII and sequenced by the chain-termination method according to Wallace et al., (16), using a 15 base oligonucleotide primer consisting of pBR322 bases 309-323.

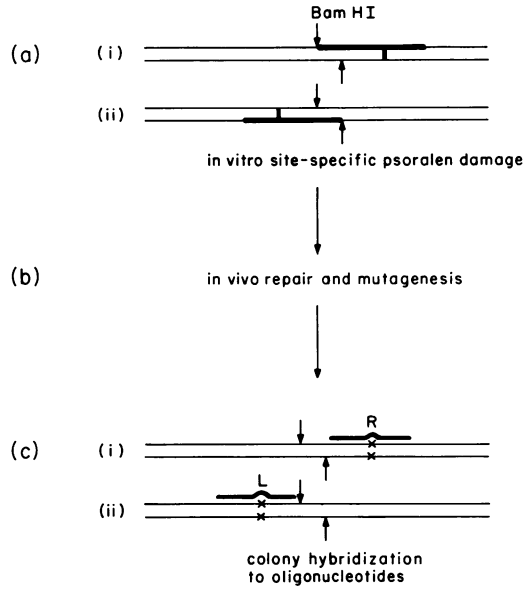

Figure 1. Experimental scheme. a) Site-specific psoralen crosslinking. Intact closed circular pBR322 was nicked by BamHI, in the presence of ethidium bromide, in the top strand (i) or the bottom strand (ii) in equal proportions. The nicked circular plasmid was nick-translated briefly with E. coli DNA polymerase I and dATP, dGTP, HgdCTP, and HgdUTP. This incorporates a patch of 20-30 mercurated nucleotides into either the top strand (i) or the bottom strand (ii). Sulfhydryl containing psoralen was added and the sample was phenol extracted, leaving only the psoralen bound specifically through Hg-S linkages. The plasmid was irradiated with near UV light, forming interstrand psoralen crosslinks, and crosslinked molecules were purified by electrophoresis and extraction from low gelling temperature agarose. _____, pBR322; bold face, mercurated regions; arrows, BamHI recognition sites.
b) In vivo SOS repair. Purified crosslinked pBR322 was transfected into E. coli AB1157 cells which had received far UV irradiation. The cells were grown for 3 hours in L broth, to allow repair and mutagenesis to occur, then plated onto 50 µg ampicillin to select cells containing repaired plasmid.
c) Screening for mutants by oligonucleotide hybridization. Plasmid-containing colonies were grown on nitrocellulose filters, on ampicillin plates. Colony hybridization was performed, on replicas, against oligonucleotide probes (i)R and (ii)L covering the psoralen modified regions to the right and left of the BamHI site. The filters were hybridized at low temperature and washed at increasing temperatures. Mutant plasmids containing mismatches within the probe sequences were identified by their poor hybridization at elevated temperatures. x, mutation sites; bold face, oligonucleotides.

RESULTS

Site-specific psoralen damage

Plasmid pBR322 was site-specifically modified with psoralen by targeting a sulfhydryl-containing psoralen derivative to mercurated nucleotides inserted near a unique restriction site (Figure 1a) (14). The plasmid was first nicked with BamHI in the presence of ethidium bromide; under these conditions the two strands have equal probabilities of receiving nicks. A 20-30 base patch of mercurated nucleotides was inserted by nick-translation with DNA polymerase I, starting at the nick and either extending downstream, on the top (coding) strand, or upstream, on the bottom (non-coding) strand, from the BamHI site. The sulfhydryl-containing psoralen was added and, after removal of non-specifically bound psoralen, the sample was irradiated with near UV light to form psoralen adducts within the mercurated regions. Thus, half the population contained psoralen molecules to the left of the restriction site, and half to the right. Crosslinked molecules were purified by alkali denaturing the DNA and isolating the rapidly renaturing double-stranded molecules from non-denaturing agarose gels. Molecules without interstrand crosslinks run as single-stranded DNA in these gels.

Repair and mutagenesis in E. coli

Plasmids containing crosslinks site-specifically placed within the tetracycline resistance gene were transformed into E. coli cells, which had been previously irradiated with far UV light to induce the SOS repair system (Figure 1b). Prior work has shown that SOS induction is necessary for the repair of psoralen crosslinks in plasmid DNA (11). After a four hour repair period, colonies containing repaired plasmid were selected by plating onto 50 µg/ml ampicillin.

Plasmids mutated within the target region were detected by colony hybridization to oligonucleotide probes (Figure 1c). The probes used were 18-base sequences covering nucleotides 357-374 (L), to the left of the BamHI site, and nucleotides 381-398 (R), to the right of the restriction site. The BamHI site itself, nucleotides 375-380, was not covered since the

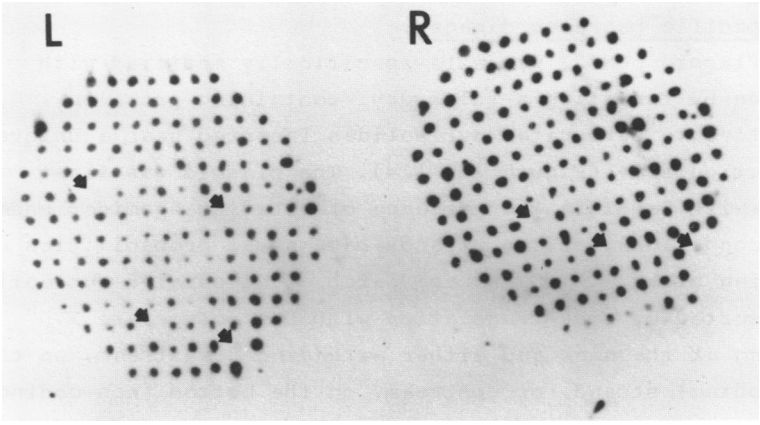

Figure 2. Colony hybridization to oligonucleotide probes L and R. Replica filters were hybridized to oligonucleotide L or R at 20°C and washed successively at 50°, 55°, 60°, and 65°. The autoradiogram after the 65° wash is shown. The indicated colonies (arrows) are mutants which hybridized well at 20° but lost signal after the high temperature washes.

nick-translation reactions on the opposite strands overlap in this region.

Hybridization between an oligonucleotide and plasmids containing mutations within the probe results in mismatched base pairing. These imperfect hybrids are stable at low temperature, but have lower melting temperatures than the completely base paired hybrids formed with wild-type plasmid. The estimated T_m for both the L and R probes in 6 x SSC, calculated as T_m = 4°/G-C + 2°/A-T, was 60°C (18). The 5'[^{32}P] labelled L and R probes were colony hybridized at 20°C to replica filters containing the repaired plasmids. At this temperature, all the colonies hybridized strongly to both probes. The filters were then washed successively at 50°, 55°, 60°, and 65°, and examined for weakly hybridizing spots. The first light spots appeared after the 55° wash, and additional light spots were detected at 60° and 65°. The overall hybridization signal decreases considerably over this range, but only colonies hybridizing very weakly relative to the majority of colonies on the filter were scored as mutants. Figure 2 illustrates replica filters hybridized to the L and R

Table 1

Mutation frequency of site-specifically crosslinked pBR322

Treatment	Frequency L	R
mercuration, no crosslinking	0.002	0.001
site-specific crosslinking	0.014	0.009

Plasmid pBR322 was site-specifically modified near the BamHI site by nick-translation with mercurated nucleotides, followed by crosslinking with sulfhydryl-psoralen. Damaged plasmid was transformed into UV irradiated E. coli. Colonies containing repaired plasmid were hybridized to oligonucleotides L and R, and plasmids mutated within the probe regions were detected by their weak hybridization signals after washes above 55°C.

probes after the 65°C wash. On each filter, several spots are noticeably lighter than the rest. In each case, they hybridized with normal intensity at 20° and 50° to both probes, but to only one of the probes at 65°.

The mutation frequencies for the two regions,determined by this method, are given in Table 1. The frequencies for both probes are about 1%, slightly higher for L than for R. Control experiments done with intact pBR322, or plasmid which had been nick-translated with mercurated nucleotides but not reacted with psoralen, revealed no apparent mutants with either oligonucleotide.

Sequence analysis of mutants

A number of the plasmids with low hybrid melting temperatures were analyzed by sequencing in the target region. All of the clearly weakly hybridizing spots were found to have a base change within the identified probe region. The mutations are summarized in Figure 3; the mutants at each site are from independently derived colonies, grown in separate cultures.

The mutations are all single base changes, with no insertions or deletions. Most are transitions, either GC to AT or AT to GC in approximately equal numbers (Table 2). There are also smaller numbers of transversions, including GC to TA, GC to CG, and AT to TA changes. There are no AT to CG

Table 2

Sequence changes found in mutants	
Transitions	
GC→AT	10
AT→GC	16
Transversions	
GC→TA	2
GC→CG	1
AT→TA	4
AT→CG	0

transversions included among these mutants but, in view of the small number of transversions observed, this may not be significant. There were two instances of double mutations: a GC to AT at 360 along with an AT to GC at 378, and AT to GC changes at both 384 and 395.

The mutations were distributed across the entire probe regions. Base changes in the middle of the probe sequence are expected to be more destabilizing than changes at the extreme ends, and in fact the mutations in the middle were detected at lower wash temperatures. The changes at position 374, the extreme edge of L, were found only after the 65° washes. A mutation at position 381, the edge of R, previously detected by screening for tetracycline sensitivity (11) was not picked up by hybridization (unpublished experiments).

Psoralens bind preferentially to thymines and thus crosslinks are found primarily at -A-T- or -T-A- sequences. The two such sites, at positions 383-384 and 395-396, were found to be hot spots for mutation detected by the R probe. There are no such high affinity sites within the L probe, but there are a number of -A-C- and -C-A- sites, most of which had one or more mutations. The changes on this side are more evenly distributed, consistent with the presence of several lower affinity sites.

The phenotype of each sequenced mutant was determined by streaking the mutant colony onto tetracycline plates, and the

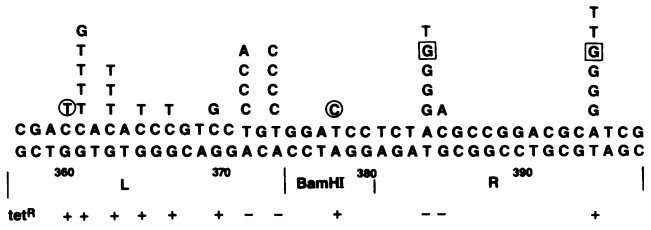

Figure 3. Sequences of mutant plasmids. Mutants identified by oligonucleotide hybridization were sequenced and the changes are shown, in the top strand only, for clarity. The 2 pairs of double mutations are indicated by circles or squares around companion changes. The tetracycline resistance of each mutant was determined, and indicated for each site on the bottom line of the figure. At each site where two different substitutions were found, both had the same phenotype.

results are shown in Figure 3. The changes at two thirds of the sites were silent, tetracycline resistant mutants.

More than one type of mutation was observed at many of the sites. At the high frequency positions 361, 372, 384, and 395, both transitions and transversions occurred. Transitions were more frequent than transversions at each site. This heterogeneity within mutation sites, and our observation of most of the possible changes suggest that, although there is a favored misincorporation at each site of psoralen damage, a unique nucleotide is not inserted at these lesions.

The mutations found with the L and R probes differ, in that those in L occur where there are pyrimidines in the coding strand, while those in R occur predominantly where there are purines in this strand. This may reflect the pattern of mercuration in the plasmid molecules. The nick translation reaction inserts mercurated pyrimidines into the non-coding strand in L, or into the coding strand in R. While the presence of mercury by itself does not lead to mutations, as shown in Table 1, these modified nucleotides may bias one or more steps in the repair process to one of the strands. For instance, while mercurated DNA is susceptible to degradation by many nucleases, it is resistant to the 5'→ 3' exonuclease activity of DNA polymerase I (19), and may have altered characteristics toward other DNA repair enzymes. The

fact that the same kinds of substitution occur in both L and R indicates that there is no strand specificity as such. In both regions, the change occurs where there is a pyrimidine, presumably the site of psoralen binding, in the non-mercurated strand.

DISCUSSION

By hybridizing oligonucleotide probes to plasmids mutagenized within a restricted region, we have detected mutations efficiently without regard to their phenotype. The results are strikingly different from those obtained in this system using a traditional genetic screening process. In a previous study, we used the loss of resistance to tetracycline to identify mutants within the psoralen target region (11). The mutations detected were primarily transitions, and a smaller number of transversions, similar to the substitutions we have found in the present study. However, the patterns of change in the two collections of mutants are different in several ways.

The tet^S mutations were unequally distributed, most falling within the BamHI site or the R probe region. The only mutations within L were at 374, at the right edge of the region. This contrasts with the much more even distribution of the present mutants between L and R. The reason for the difference is evident when we examine the phenotype of the mutants. Most of the changes within L do not diminish tetracycline resistance, and are therefore not detected by the genetic screening.

The preference for strong psoralen crosslinking sites was not apparent in the earlier study, in which some changes were found at positions which were not even possible psoralen crosslinking sites, and the favored -A-T- and -T-A- sites seemed to be under-represented. Again, many of the mutations at these sites are silent. Most of the changes identified by hybridization were found at likely crosslinking sites. Only one, a single GC to CG transversion at 370 within the sequence -T-C-C-, could not have arisen from a crosslink, since a psoralen monoadduct bound to the first C does not have a

pyrimidine on the opposite strand to complete the crosslink. More than one type of substitution was found at several of the mutation sites detected by hybridization, while only one kind of change was found at each site by genetic screening.

These differences illustrate the different pictures of mutagenesis presented by the two techniques. The mutations biased by the requirement for a particular phenotype led to the conclusion that psoralen mutagenesis occurs rarely at crosslinks, is unevenly distributed over the damaged region of DNA, and that psoralen lesions code exclusively for one kind of misincorporation. The more impartial physical method used here allows the conclusion that mutagenesis occurs mainly at preferred crosslinking sites and that more than one type of substitution is possible at the damage sites.

Schaaper et al., (20) have reported that E. coli SOS repair of apurinic sites results predominantly in the incorporation of A opposite these noncoding lesions, and in vitro synthesis past apurinic sites of pyrimidine dimers is also reported to favor A (21-23). We see no such preference for A over G, although most of the substitutions, presumably opposite psoralen-adducted pyrimidines, involve purine insertions.

The ratio of transitions to transversions (4:1) is similar to that seen for UV mutagenesis (2, 5). Psoralen mutagenesis of rII in bacteriophage T4 also produces primarily transitions (24). However, angelicin, a monofunctional analog of psoralen incapable of forming crosslinks, induces mainly transversions in lacI (S. Miller and E. Eisenstadt, personal communication).

ACKNOWLEDGEMENTS

We wish to thank Lex van der Ploeg for advice and instruction in colony hybridization. This research was supported by grant GM 14825 from the USPHS.

REFERENCES

1. Miller, J.H. (1982) Cell 31, 5-7.
2. Coulondre, C., and Miller, J.H. (1977) J. Mol. Biol. 117, 577-606.

3. Todd, P.A., and Glickman, B.W. 91982) Proc. Natl. Acad. Sci. (1982) 4123-4127.
4. Brash, D.E., and Haseltine, W.H. (1982) Nature 298, 189-192.
5. LeClerc, J.E., and Istock, N.L. (1982) Nature 297, 596-598.
6. Wood, R.D., Skopek, T.R., and Hutchinson, F. (1984) J. Mol. Biol. 173, 273-291.
7. Foster, P.L., Eisenstadt, E., and Miller, J.H. (1983) Proc. Natl. Acad. Sci. 80, 2695-2698.
8. Eisenstadt, E., Warren, A.J., Porter, J., Atkins, D., and Miller, J.H. (1982) Proc. Natl. Acad. Sci., 79, 1945-1949.
9. Wallace, R.B., Shaffner, J., Murphy, R.F., Bonner, J., Hirose, T., and Itakura, K. (1979) Nucl. Acids Res. 6, 3543-3557.
10. Zoller, M.J., and Smith, M. (1982) Nucl. Acids Res. 10, 6487-6500.
11. Saffran, W.A., and Cantor, C.R. (1984) J. Mol. Biol., in press.
12. Hearst, J.E. (1981) Ann. Rev. Biophys. Bioeng. 10, 69-86.
13. Peckler, S., Graves, B., Kanne, D., Rapoport, H., Hearst, J.E., and Kim S.-H. (1982) J. Mol. Biol. 162, 157-172.
14. Saffran, W.A., Goldenberg, M., and Cantor, C.R. (1982) Proc. Natl. Acad. Sci. 79, 4594-4598.
15. Wallace, R.B., Schold, M., Johnson, M.J., Dembek, P., and Itakura, K. (1981a) Nucl. Acids Res. 9, 3647-3656.
16. Wallace, R.B. (1981b) Gene 16, 21-26.
17. Marko, M.A., Chipperfield, R., and Birnbaim, H.C. (1982) Anal. Biochem. 121, 382-387.
18. Dalbadie-McFarland, G., Cohen, L.W., Riggs, A.D., Morin, C., Itakura, K., and Richards, J.H. (1982) Proc. Natl. Acad. Sci. 79, 6409-6413.
19. Dale, R.M.K., and Ward, D.C. (1975) Biochemistry 14, 2458-2469.
20. Schaaper, R.M., Kunkel, T.A., and Loeb, L.A. (1983) Proc. Natl. Acad. Sci. 80, 487-491.
21. Sagher, D., and Strauss, B. (1983) Biochemistry 22, 4518-4526.
22. Kunkel, T.A., Schaaper, R.M., and Loeb, L.A. (1983) Biochemistry 22, 2378-2384.
23. Rabkin, S.D., Moore, P.D., and Strauss, B.S. (1983) Proc. Natl. Acad. Sci. 80, 1541-1545.
24. Drake, J.W., and McGuire, J. (1967) J. Vir. 1, 260-267.