Sequence Specificity of Mutagenesis in the cI Gene of Bacteriophage Lambda

by T. R. Skopek,* R. D. Wood,† and F. Hutchinson†

Studies of DNA base sequence alterations have shown that for every agent the mutagenic process is specific with respect to the types of base changes induced and the location of the changes in the DNA. Analysis of the types of mutations produced by mutagenic agents can provide insight into the mechanism of mutation and can suggest which DNA lesions may be involved in the actual mutagenic event.

We have developed a system for the analysis of chemically induced base sequence alterations in the cI repressor gene of bacteriophage lambda using DNA sequencing techniques. To illustrate the utility of this type of analysis, we present the results obtained with ultraviolet light (UV). Irradiation of target DNA with UV alone, or UV followed by photoreactivating light (which removes dimers), produces mostly transitions at pyrimidine–pyrimidine sites. Conversely, irradiation with 313 nm light plus acetophenone (which produces only thymine dimers) produces mostly transversions at low efficiency. This and other evidence suggests that the actual premutagenic UV lesion in *E. coli* may not be pyrimidine–pyrimidine dimers, but rather pyr(6-4)pyo photoproducts.

Introduction

The advent of rapid DNA cloning and sequencing techniques has made it possible to analyze directly DNA base alterations induced by mutagenic agents. These alterations represent an unequivocal and permanent record of the final endpoint in the mutagenic pathway. Their analysis is extremely useful in understanding the mechanism of mutation, since it sets constraints on the possible DNA lesions involved in mutagenesis and on the enzymatic reactions acting on them.

Mutagenesis is not a random process occurring along the DNA helix. When viewed at the nucleotide level, the most striking feature of induced mutations is their specificity. Mutagenesis is specific in terms of the types of base change induced and the location of these changes in the DNA. All mutagens (and spontaneous mutagenic processes) produce characteristic patterns of changes in a gene; these patterns are like molecular "fingerprints," each reflecting an agent's activity in the DNA.

Most of the work done to date concerning mutagen specificity has been carried out using *E. coli* and bacteriophage, and we review some of the key systems in this paper. We will also describe a system we have developed in the cI gene of bacteriophage lambda for the analysis of mutagen specificity. Data on the sequence specificity of ultraviolet light mutagenesis are presented to illustrate the utility of this type of analysis

---

*Chemical Industry Institute of Toxicology, Research Triangle Park, NC 27709.

†Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511.

in understanding the mechanism of mutation and in identifying the DNA lesions involved.

Background

The sequence specificity of spontaneous and induced mutations was first demonstrated in the rII locus of *E. coli* bacteriophage T4 by Benzer (1). Benzer isolated a large number of spontaneous and chemically induced rII mutants and located them to within a few base pairs in the rII gene by genetic crosses with deletion mutants in the region. With this system, Benzer made two important discoveries concerning mutagen specificity. First, for a given agent certain sites in the rII gene displayed higher frequencies of mutagenesis than others. Benzer termed these sites mutagenic "hotspots." Second, the distribution of these hotspots as a function of base-pair position along the gene was different for spontaneous mutations and for mutations induced by several chemical agents, which in turn varied among themselves. The rII system is still in use, with the study of mutations being extended by reversion analysis and direct DNA sequencing.

The most productive system for the study of mutagen specificity has been the *E. coli* lac I system developed by Miller and his colleagues (2). In one application of this system nonsense mutations in the lac I gene are selected and mapped to within a few base-pairs with a set of known deletion mutants (3). Since the lac I base sequence is known (and hence every site that can give rise to a nonsense codon), it is possible to infer the actual base-pair change that occurred to generate a given nonsense mutant based on its map location and pattern of

suppression. This approach possesses the tremendous advantage that hundreds of nonsense mutations can easily be analyzed due to the elegant genetics engineered by Miller. The approach has the disadvantage that a limited number of sites in a gene (i.e., those which can mutate to a nonsense codon) can be analyzed for mutation. However, in another application, any induced mutation in lac I can be determined directly by DNA sequence analysis, but at the cost of more labor per mutant (4).

A great deal of what is known about mutagen specificity was learned by investigators using the lac I system (5).

Lambda cI System

We have developed a lambda phage/*E. coli* system that can be used to determine changes in base sequence induced by various mutagenic agents (6). The genetic target in this system is the cI repressor gene. The cI gene codes for a DNA-binding control protein which suppresses the expression of phage lytic functions during lysogeny. Mutants in cI gene cannot form stable lysogens and can be identified by their clear-plaque morphology. Wild type phaques are turbid due to the establishment and growth of stable lysogens in the plaque. Mutations in cI can be mapped genetically by crosses with known deletion mutants, and the appropriate DNA restriction fragment sequenced to determine any changes in base order. This procedure detects any forward mutation that leads to loss of function of the cI gene product.

The use of a lambda system for studying mutagenesis in *E. coli* has several advantages. First, whether growing lytically or as a lysogen, lambda uses *E. coli* polymerases and repair enzymes. Therefore the response seen should be characteristic of that occurring in the *E. coli* genome. Second, the lambda system affords several different modes of exposure of the target gene to the mutagenic agent being studied. Lambda can be exposed while in the *E. coli* genome as a lysogen or can be exposed separately as either intact phage or naked DNA. This permits separate exposure of the host cells to mutagenic treatments in order to induce the SOS response required for mutagenic expression of many compounds (7), thus providing greater flexibility in experimental design.

To illustrate the utility of the system we present results obtained in our studies with ultraviolet light. First, however, a brief discussion of the factors affecting mutagen specificity should be included.

Factors Affecting Specificity of Mutagenesis

A first consideration in analyzing mutagenic spectra is the fact that not all base pair changes can be detected as a change in cell phenotype. Mutation at codon wobble positions will go undetected since no change in amino acid will occur. Also undetected will be those mutations

producing amino acid substitutions which result in wild-type enzyme activity. The remaining class of mutations will adversely affect enzyme activity. The fraction of these mutations which will be scored will depend on the stringency of selection; that is what level of enzyme activity will define the boundary between mutant and wild-type phenotype.

These factors can be observed in the cI gene. Most of the base pair substitution mutations observed in cI (approximately 90%) occur in the first 240 bases of the 720 bases of the gene. This region of the gene codes for the amino terminus of the protein, which is the portion of the molecule which actually binds to the DNA to carry out its function (8). The carboxy terminus is responsible only for protein–protein interactions between cI molecules (which usually exist as a dimers). Therefore, it makes sense that substitutions in the amino terminus would be more likely to result in cI inactivation. Polar mutations such as frameshifts and large insertions, which result in gross changes in protein structure, can be seen throughout the gene (6,9).

The final spectrum of mutation for a given agent results from its chemical specificity which operates to induce its unique subset of observable mutations in a gene. This specificity can be attributed to several factors. First, different agents will obviously produce different types of chemical alterations in DNA. It is not surprising that mutagens forming bulky adducts, such as benzo(a)pyrene or aflatoxin, behave differently than alkylating agents. Even similar agents will have different affinities for the various moieties in DNA. MNU and MMS, two methylating agents, produce greatly different patterns of methylated bases; for example the ratios of O^6-methyl guanine to N^7-methyl guanine is 9.4% for MNU and only 0.4% for MMS (10).

The reaction of an agent to produce a specific modified base may also be dependent on surrounding bases. For example the rate of formation of the pyr(6-4)pyo photoproduct at a given 5'-pyrimidine–pyrimidine-3' sequence varies more than an order of magnitude from site to site (11).

The rate of repair of a given type of modified base may also depend on the surrounding sequence. Todd and Glickman (12) have compared the sequence specificity of UV mutagenesis in the lac I gene in an excision proficient and deficient strain of *E. coli*. The authors observed that the two spectra were very similar except for one site, amber 24, which was approximately 50 times less mutable in the excision-proficient strain. This suggests that the lesion responsible for the amber 24 mutation is extremely sensitive to the presence of excision repair.

The rate of base misinsertion at a given location in the DNA may also be dependent upon the surrounding sequence. For example in the cI gene bromouracil induces only AT → GC transitions; however, this change is extremely sequence-specific. Mutation at four 5'-ACGC-3' sequences accounts for approximately two-thirds of the observed mutations. The sequence 5'-AC-3' appears approximately ten times more mutable than

AG, AT, or AA. The sequence ACG does not occur in the gene except in the ACGC sequences (6).

The exact distribution of induced mutations occurring in a gene may also be dependent on the dose delivered to the organism. If there is more than one mechanism operating to convert modified bases into mutations in a gene, and one or more of these are saturable processes, then a different distribution of mutations will be observed when none of the processes are saturated than when one or more are saturated.

These factors operating together result in a unique spectrum of mutation for a given agent.

Example of Sequence Specificity: UV Mutagenesis

Ultraviolet light has been the most widely studied mutagenic agent in molecular biology. The principal lesion formed in DNA is the pyrimidine–pyrimidine cyclobutyl dimer, which has long been implicated as a premutagenic lesion involved in UV mutagenesis. The reasons for this are (1) the dimers are the major type of UV photoproduct; and (2) E. coli treatment with photoreactivation light (which activates the E. coli photoreactivation enzyme) removes dimers from the DNA and abolishes the mutagenic effect of UV. Recently, Brash and Haseltine (11) have suggested that another lesion, the pyr(6-4)pyo photoproduct, may be the actual premutagenic lesion in UV irradiated DNA.

We studied the base alterations induced in the cI gene of lambda irradiated with 30 J/m² UV (13). Irradiated phage were used to infect an excision deficient E. coli cell that had previously been irradiated with 3 J/m². The irradiation of the host cell is necessary for the induction of the SOS system in E. coli, and is required to observe UV mutagenesis (7).

The types of mutations observed following this treatment are summarized in Table 1. Approximately two-thirds of the mutations were transitions, equally divided between GC → AT and AT → GC. The remaining one-third was about equally divided among transversions, frameshifts, and double mutation events. Double mutation events include alterations involving two base-pair changes, or a base change and a frameshift.

The sequence specificity of the observed transitions and transversions is presented in Table 2. One can see from these data that there is a strong preference for mutations at pyrimidine–pyrimidine sequences in the DNA; few mutations are seen at pyrimidine–purine–pyrimidine sequences. Furthermore, there is a strong preference for mutation at the 3'-base of pyrimidine–pyrimidine sequence. Of the possible combinations of thymine and cytosine in a pyrimidine–pyrimidine sequence, only the 5'-CT-3' sequence appeared to be less sensitive to mutation. Mutations have been observed at 5'-CT-3' sequences with other agents, demonstrating that there is no systematic bias against these mutations in the cI gene.

Phage were also irradiated with 313 nm light in the presence of acetophenone and then used to infect in-

Table 1. Spectrum of mutation induced by ultraviolet light in the cI gene of bacteriophage lambda.[a]

	Region of cI gene			
	Operator-promoter	Base-pairs 1–240	Base-pairs 240–714	Totals
Mutants mapped	8	71	19	98
Mutants sequenced	5	50	7	62
Types of mutations				
Transitions				
GC → AT	3	20	0	23
AT → GC	0	16	2	18
Transversions	0	7	1	8
Frameshifts				
− 1 base pair	1	0	1	2
+ 1 base pair	0	3	1	4
+ 2 base pair	0	1	0	1
Double events				
2 base changes	1	2	0	3
1 base change and frameshift	0	1	2	3

[a] Data are from Wood et al. (13).

duced host cells. This irradiation treatment yields mostly TT dimers in the phage and no pyr(6-4)pyo photoproducts. Phage were irradiated to produce approximately the same number of dimers as in the experiment described above. Of the 22 mutants analyzed, only one was a transition; of the remaining 21, 18 were transversions and 3 were frameshifts. Therefore it can be concluded that TT cyclobutyl pyrimidine dimers do not efficiently induce transitions, the major class of UV-induced mutations.

The role of dimers in UV mutagenesis was further studied by irradiating lambda DNA with UV light, and then removing dimers by treatment with photoreactivating enzyme and photoreactivating (14). The pyr(6-4)pyo lesion is not photoreversible (Brash, personal communication). The removal of dimers was confirmed by treating the DNA with T4 UV endonuclease and analyzing on an agarose gel. The treated DNA was then packaged in vitro to form viable phage and used to infect induced E. coli hosts. The survival of phage that were UV-irradiated and then photoreactivated was greatly enhanced relative to controls that were UV irradiated but not photoreactivated. This suggests that dimers are toxic lesions in E. coli. However, the number of cI mutants per surviving phage remained approximately constant, suggesting again that dimers are not the major premutagenic lesion in the DNA. When analyzed at the nucleotide level, the spectrum of mutations observed in UV-irradiated, photoreactivated phage appeared very similar to that observed in the original UV-induced population, with approximately two-thirds transitions occurring in pyrimidine–pyrimidine locations (Table 3). Thirteen of the mutants, observed at eight different sites, had previously been seen in non-photoreactivated phage.

Taken collectively, the data suggest that dimers are not the major premutagenic lesion formed in DNA fol-

Table 2. Base sequence specificity of UV mutagenesis in the cI gene (bases 1–240).

Sequence 5'–3'[b]	Transitions			Transversions		
	Total sites[c]	Sites mutated by UV	No. of UV mutants	Total sites[c]	Sites mutated by UV	No. of UV mutants
Pur-Pyr*-Pur	33	1	1	47	0	0
Pur-Pyr*-Pyr	40	1	1	51	2	2
Pyr-Pyr*	85	20	34	108	5	5
Total	158	22	36	206	7	7
T-C*	16	7	11			
C-C*	15	5	8			
T-T*	38	8	15			
C-T*	16	0	0			

[a] Data are from Wood et al. (13).

[b] Asterisk (*) indicates the pyrimidine undergoing mutagenesis.

[c] The number of sites between base-pairs 1 and 240 in the cI gene at which a transition or transversion changes the amino acid.

Table 3. Sequence specificity of transition mutations produced by UV irradiation followed by photoreactivation.[a]

5'–3' sequence[b]	No. mutants observed
Pu-Py*-Pu	1
Pu-Py*-Py	2
Py-Py*	13
T-T*	5
T-C*	5
C-C*	2
C-T*	1

[a] Data are from Wood (14).

[b] Asterisk (*) indicates the base undergoing transition mutation. A total of 28 mutants were sequenced (16 transitions, 10 transversions, 1 frameshift, and 1 double mutation).

lowing UV irradiation. As suggested by Brash and Haseltine (11) the pyr(6-4)pyo lesion appears to be a more likely candidate and is the consistent in many respects with our observations in cI: This lesion can only form at pyr–pyr sites, which is where the majority of UV mutations occur. This lesion does not form readily at 5'-CT-3' locations (11), which is consistent with the low level of mutation at these sites. The bond between the bases in the pyr(6-4)pyo adduct destroys the base-pairing moieties of the 3' base while leaving those of the 5' base intact; this is consistent with the observation that most mutations occur at the 3' base. The pyr(6-4)pyo photoproduct is not photoreversible and should still exist to give rise to mutations in UV-irradiated and photoreactivated DNA.

Earlier results obtained in vivo in E. coli have shown that photoreactivation reverses the mutagenic effect of UV. The results presented here suggest that the effect of this treatment may not be the removal of premutagenic lesions, but rather the removal of inducing signals (dimers) for the SOS system. The use of the lambda/E. coli system allows the dissection of the two processes.

Future Directions

Most of the work concerning mutagen specificity has been performed in E. coli. The wealth of information

concerning the mechanism of mutation in E. coli derived from this data is impressive, and is an indication of the value of the approach.

Virtually nothing is known about the detailed molecular mechanism of mutation in human and other mammalian cells. Initial attempts at assay development in this field have quickly demonstrated that such systems will not be trivial to develop. Calos et al. (15) and Razzaque et al. (16) have shown that in mammalian cells autonomously replicating plasmids that contain a target gene for mutagenesis are subject to extremely high levels of mutation (10). This undoubtedly involves a mechanism quite distinct from the mutagenesis occurring in the genome.

Tindall et al. (17) have constructed a CHO line with a single stable copy of the E. coli gpt gene in its genome. They have used restriction fragment analysis to study gross alterations in the gpt gene following spontaneous and chemically induced mutation. At this level of analysis, the foreign gene seems to behave much like the native hgprt gene. The success of their initial approach suggests that more sophisticated systems may be developed which involve the construction of mammalian cell lines with readily retrievable target genes permitting detailed sequence analysis.

The development of systems for the analysis of mutagen specificity in mammalian cells may one day reveal: the important mutagenic events in these cells, whether or not these events are targeted at the site of an altered base, and what types of altered bases are involved. Knowledge of the types of mutations which occur spontaneously and which are induced by chemicals may one day allow the determination of the origin of mutants present in individuals. Also, with the tremendous growth of our understanding of oncogenes and the genetic alterations that are required to activate them, knowledge of the sequence specificity of mutagens/carcinogens may shed light on their mechanism of action and the factors governing their species and tissue specificity.

REFERENCES

1. Benzer, S. On the topography of the genetic fine structure. Proc. Natl. Acad. Sci. (U.S.) 47: 403–415 (1961).

2. Coulondre, C., and Miller, J. H. Genetic studies of the lac repressor: IV. Mutagenic specificity in the lac I gene of *Escherichia coli*. J. Mol. Biol. 117: 577–606 (1977).

3. Miller, J. H., Coulondre, C., and Farabough, P. J. Correlation of nonsense sites in the lac I gene with specif codons in the nucleotide sequence. Nature 274: 770–775 (1978).

4. Calos, M. P., and Miller, J. H. Genetic and sequence analysis of frameshift mutations induced by ICR-191. J. Mol. Biol. 153: 39–66 (1981).

5. Miller, J. H. Mutational specificity in bacteria. Ann. Rev. Genet. 17: 215–238 (1983).

6. Skopek, T. R., and Hutchinson, F. DNA base sequence changes induced by bromouracil mutagenesis of lambda phage. J. Mol. Biol. 159: 19–33 (1982).

7. Walker, G. C. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. Microbiol. Rev. 48: 60–93 (1984).

8. Sauer, R. T., Pabo, C. O., Meyer, B. J., Ptashne, M., and Bachman, K. C. Regulatory functions of the lambda repressor reside in the aminoterminal domain. Nature 279: 296–400 (1979).

9. Skopek, T. R., and Hutchinson, F. Frameshift mutagenesis of lambda prophage by 9-aminoacridine, proflavin, and ICR-191. Mol. Gen. Genet. 195: 418–423 (1984).

10. Singer, B., and Grunberger, D., Eds. Molecular Biology of Mutagens and Carcinogens. Plenum Press, New York, 1983, p. 68.

11. Brash, D. E., and Haseltine, W. A. UV-induced mutation hotspots occur at DNA damage hotspots. Nature 298: 189–192 (1982).

12. Todd, P. A., and Glickman, B. W. Mutational specificity of UV light in *Escherichia coli*: indications for a role of DNA secondary structure. Proc. Natl. Acad. Sci. (U.S.) 79: 4123–4127 (1982).

13. Wood, R. D., Skopek, T. R., and Hutchinson, F. Changes in DNA base sequence by targeted mutagenesis of lambda phage by ultraviolet light. J. Mol. Biol. 173: 273–291 (1984).

14. Wood, R. D. Mutations in ultraviolet-irradiated lambda phage DNA after photoreactivation. Submitted.

15. Calos, M. P., Lebkowski, J. S., and Botchan, M. R. High mutation frequency in DNA transfected into mammalian cells. Proc. Natl. Acad. Sci. (U.S.) 80: 3015–3019 (1983).

16. Razzaque, A., Mizusawa, H., and Seidman, M. M. Rearrangement and mutagenesis of a shuttle vector plasmid after passage in mammalian cells. Proc. Natl. Acad. Sci. (U.S.) 80: 3010–3014 (1983).

17. Tindall, K. R., Stankowski, L. F., Machanoff, R., and Hsie, A. W. Detection of detection mutations in pSVgpt-transformed cells. Molec. Cell. Biol. 4: 1411–1415 (1984).