The EMBO Journal vol.6 no.1 pp.63-67, 1987

Unique pattern of point mutations arising after gene transfer into mammalian cells

Janet Hauser, Arthur S.Levine and Kathleen Dixon

Section on Viruses and Cellular Biology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892, USA

Communicated by J.H.Miller

We have used a simian virus 40 (SV40)-based shuttle vector, pZ189, to analyze the sequence specificity of spontaneous point mutations that arise after transfection of this vector into monkey cells. The majority of the mutants which we studied had multiple base substitutions (mostly G·C→A·T transitions and G·C→T·A transversions) within the 160-bp region sequenced. Almost all of the mutations occurred in the right-hand G·C bp of one of the two following sequences, 5′-TC-3′: 3′-AG-5′ or 5′-CC-3′ : 3′-GG-5′. We postulate that these mutations result from DNA replication infidelity occurring during repair of the transfected DNA which has been damaged by cellular nucleases. The sequence specificity of the mutations suggests an effect of the following nucleotide on misincorporation wherein A (or less frequently T) is preferentially misincorporated opposite C when the next nucleotide inserted is A (or less frequently G). Our results support the utility of the shuttle vector as a model in studies on gene transfer and document the extreme plasticity of DNA transfected into mammalian cells.

Key words: mutagenesis/shuttle vector/transfection

Introduction

The ability to introduce foreign genes into living eukaryotic cells by DNA transfection has revolutionized the study of gene structure and expression. Much of our current knowledge concerning the identification and characterization of oncogenes, and the pathways of gene regulation in eukaryotic cells, has been derived from experiments involving DNA transfection. Moreover, there is now evidence that DNA transfection techniques may be useful in gene therapy for certain heritable human diseases (Weatherall, 1984). In all of these applications, it is important that the informational integrity of the input DNA is faithfully preserved. However, it is now becoming clear that DNA sequences that are introduced into eukaryotic cells by DNA transfection procedures are subject to extensive rearrangement, deletion, insertion and point mutagenesis (Calos *et al.*, 1983; Lebkowski *et al.*, 1984; Razzaque *et al.*, 1983, 1984; Miller *et al.*, 1984). The mechanisms by which such alterations occur are not well understood, although it is thought that these mutations are initiated by damage to the DNA (e.g. nuclease cleavage) occurring upon entry of the naked DNA into the cell (Razzaque *et al.*, 1984; Wake *et al.*, 1984; Miller *et al.*, 1984). Detailed analysis of the characteristics of these mutations should yield useful information on the mechanisms that are involved in the genetic instability of transfected DNA, and this knowledge may lead to the development of methods to control these mechanisms.

We have used a shuttle vector plasmid to analyze the point mutations that occur upon transfection and replication of transfected DNA in monkey cells. The shuttle vector plasmid pZ189 (Seidman *et al.*, 1985) contains the pBR327 replication functions and β-lactamase gene for growth and selection in bacterial cells, the simian virus 40 (SV40) early region for replication in mammalian cells, and the *Escherichia coli supF* suppressor tRNA gene as a mutagenesis target that can be easily selected and sequenced. We showed previously (Hauser *et al.*, 1986) that 'spontaneous' mutants arise during passage of the plasmid in monkey cells at a frequency of about 0.04% of replicated molecules, and that about half of these altered molecules carry point mutations in the *supF* gene. The other half carry extensive DNA alterations (additions and deletions). Here we offer an analysis of the specific DNA sequence changes observed in the spontaneous mutants described in our previous report and additional spontaneous mutants we have since isolated and sequenced.

Results

To isolate spontaneous mutants, the pZ189 plasmid DNA was introduced into CV-1 monkey cells by the DEAE dextran method (McCutchen and Pagano, 1968) and after 48 h, newly replicated vector DNA was recovered and purified (Hauser *et al.*, 1986). This DNA was then introduced into the bacterial indicator strain MBM7070 and transformed colonies carrying mutant plasmids were identified by color (see Materials and methods). Mutant plasmids were screened for insertions or deletions by gel electrophoresis and those exhibiting normal gel mobility were sequenced in the region of the *supF* gene.

Of the 27 point mutants sequenced, the majority had multiple base changes, some as many as six or eight within the 160 bp region sequenced. The DNA sequence changes observed in our collection of spontaneous point mutants are indicated in Figure 1, together with the normal sequence of the region examined. It can be seen that the mutations were found both within the bounds of the coding region for the tRNA itself, as well as within a region of base pairs surrounding the tRNA gene. All of the mutants had at least one mutational change within the tRNA coding sequence, accounting for their selection as *supF*$^-$. Our previous data (Hauser *et al.*, 1986; Protić-Sabljić *et al.*, 1986) and that of others (Bredberg *et al.*, 1986) indicate that single base changes in at least 40 different sites within the tRNA coding region can lead to the *supF*$^-$ phenotype. Thus, it is unlikely that our assay favors the selection of multiple mutations.

Interestingly, 73 out of the 74 base changes observed involve G·C base pairs. In addition, further inspection of the types of base changes present in the individual mutants revealed a striking pattern of site specificity. Almost all of these changes occurred in the right-hand base pair of either of the two following sequences, 5′-TC-3′:3′-AG-5′ or 5′-CC-3′: 3′-GG-5′ (Table I). It is clear from Table I that this sequence specificity is not due merely to a sequence bias in the target gene itself, since we observe very few mutations in Pur-C-Pur sequences, despite the

Fig. 1. Base changes in spontaneous *supF*$^-$ point mutants. Methods for selection and sequencing of mutants have been described previously (Hauser *et al.*, 1986). All figures were prepared with the use of the DNADRAW program developed by Shapiro and Senapathy (1986).

presence of a reasonable number of these sequences in the target, and we observe a large number of mutations in the Pyr-C-Pur sequence and very few in the Pur-C-Pyr sequences, despite the presence of many of both of these sequences in the target. Furthermore, each mutant generally had either changes of C exclusively or changes of G exclusively in the strand sequenced. This pattern is illustrated in Figures 2 and 3 for each of the spontaneous point mutants sequenced. All of the TC and CC dinucleotides are shown for each of the two DNA strands, along with the position number of the 3′ base in the DNA sequence. The specific base changes observed in the mutants are indicated beneath the altered base. Of the 74 mutational changes observed in our spontaneous mutant collection, only eight failed to conform to this pattern. Of these, three (not shown) were found in mutants with single mutations and the other five were in mutants with multiple mutations that otherwise conformed to the pattern.

We were interested to determine whether the multiple mutations observed earlier among a collection of u.v.-induced mutants (Hauser *et al.*, 1986; and unpublished results) might also follow this pattern. While most of the u.v-induced mutants we had sequenced had single or tandem double mutations, some contained three or more mutational changes. The frequency of these was low enough to suggest that they might reflect isolates from the 'spontaneous' background. As illustrated in Figures 2 and 3, almost all (46/49) of these multiple mutations also conform to the mutational pattern described above for the spontaneous mutants.

Discussion

We find that spontaneous mutants that arise as a result of passage of a shuttle vector plasmid in monkey cells contain multiple base substitutions, predominantly G·C→A·T transitions and G·C→T·A transversions. These occur almost exclusively at the 3′ positions of TC and CC dinucleotides (or at the 5′ positions of GA or GG dinucleotides). Miller *et al.* (1984) also observed the preferential occurrence of spontaneous point mutations at G·C base pairs in another shuttle vector system. They suggested that these mutations might result from deamination of cytosine and depurination of guanine. However, in that study, amber mutations were analyzed by genetic techniques and multiple mutations would not have been detected. We consider it unlikely that the multiple mutations we observe arise as a result of this type of base damage, since the possibility of sufficient damage to generate these multiple mutations is remote.

We assume that these spontaneous mutations arise as a result of nuclease damage to the transfected DNA. We further assume that repair of this damage includes gap filling and/or nick translation by cellular DNA polymerase(s). During this DNA synthesis we propose that unfaithful replication generates the mutations we observe. However, we cannot rule out the possibility that the mutations arise due to misincorporation opposite damaged bases during DNA replication. In any case, given the striking sequence specificity of the point mutations we observe, we can propose two general models for their origin. Either the DNA polymerase misincorporates a base (usually A) opposite a C (or damaged C) in the template strand and the frequency of this misincorporation depends on the following nucleotide, or the DNA polymerase misincorporates a base (usually T) opposite G (or damaged G) in the template strand and the frequency of this misincorporation depends on the preceding nucleotide. Evidence from prokaryotic systems suggests that the base pair formed by a nucleotide either preceding (5′ to) or following (3′ to) a misincorporated nucleotide can influence misincorporation frequency (Loeb and Kunkel, 1982).

In prokaryotic systems, an influence of the following nucleotide on base misincorporation is in some cases due to the 3′→5′ exonuclease proofreading function of the DNA polymerase (reviewed in Loeb and Kunkel, 1982). Although DNA polymerases, purified from a variety of eukaryotic cells, have a low replication fidelity and appear to lack 3′→5′ exonuclease proofreading activity, it has been speculated that accessory proteins may function in concert with these polymerases in the cell to provide the proofreading function and improve replication fidelity (Loeb and Kunkel, 1982). The data reported here on the sequence specificity of spontaneous point mutations could be explained on the basis of a reduced efficiency of proofreading. Both the following nucleotide and the preceding nucleotide can influence misincorporation by stabilization of the mismatched base pair due to base stacking interactions (Loeb and Kunkel, 1982; Topal *et al.*, 1980). The mutations we observe might be explained by the stabilization of a misincorporated base (opposite C) by a following purine

Table I. Distribution of spontaneous mutations in 5′-NCN-3′[a] trinucleotides in the sequenced portion of the *supF* gene of pZ189

5′ nucleotide	Incidence of mutations: 3′ nucleotide				Incidence of 5′-NCN-3′[a]: 3′ nucleotide			
	T	C	G	A	T	C	G	A
T	4	8	24	6	3	7	9	5
C	7	10	12	5	6	11	6	7
G	2	1	2	0	5	6	4	5
A	0	0	1	0	5	6	1	1

[a]N denotes any nucleotide.

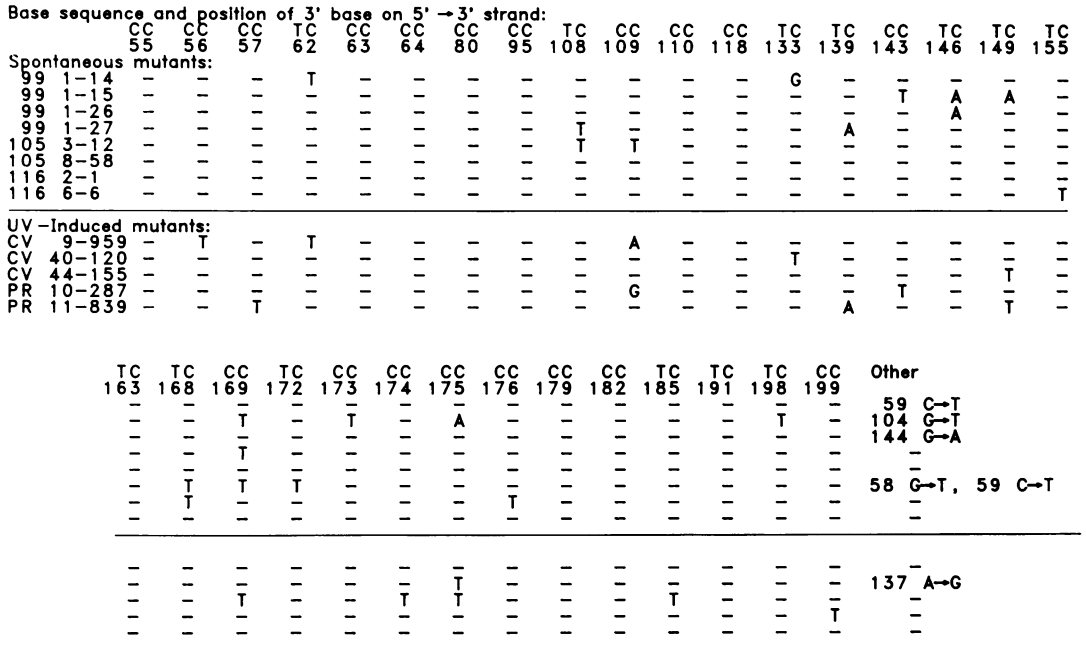

Fig. 2. Base changes in the 3′ base of TC and CC sequences of the 5′→3′ strand in mutants from the spontaneous and u.v.-induced collections.

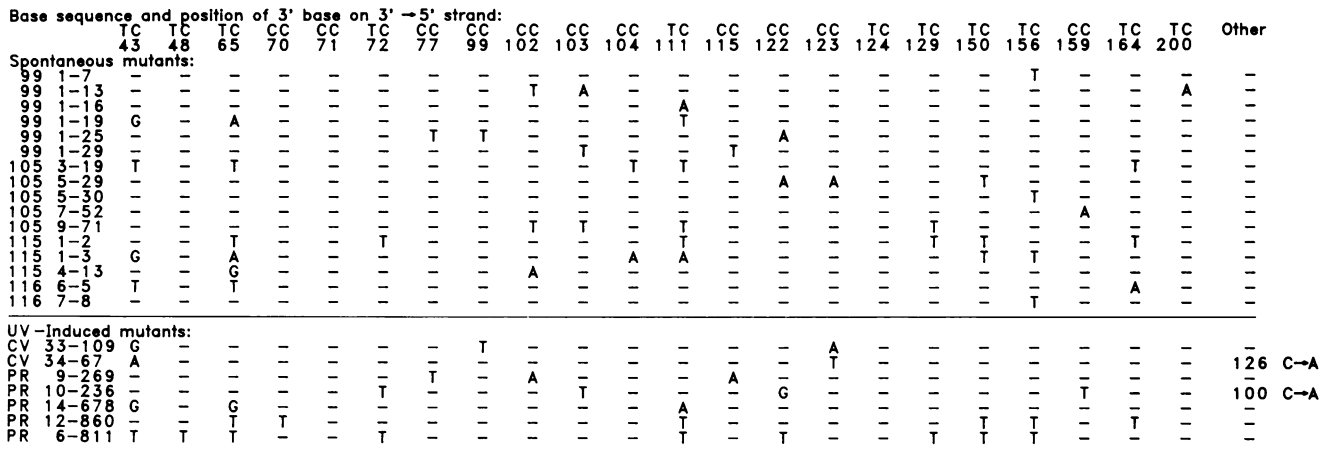

Fig. 3. Base changes in the 3′ base of TC and CC sequences of the 3′→5′ strand in mutants from the spontaneous and u.v.-induced collections.

nucleotide. Stabilization of a misincorporated base (opposite G) by a preceding pyrimidine would be less likely since interactions involving pyrimidines are not as strong (Topal *et al.*, 1980).

On the basis of the above discussion, we favor a model for misincorporation in our system in which a mispaired nucleotide is inserted opposite C in the template DNA; the frequency with which this leads to permanent misincorporation is determined by the next nucleotide. Misincorporation is favored when the next nucleotide is A (or G).

The majority of base changes we observe are G·C→A·T. If misincorporations occur opposite C, this would mean that the insertion of A would be favored. We have already discussed the idea advanced by several groups (Miller, 1985; Rabkin *et al.*, 1983) that the spectrum of u.v.-induced mutations may be due in part to a preferential insertion of A opposite u.v. photoproducts in the DNA template (Hauser *et al.*, 1986). Preferential insertion of A opposite u.v. photoproducts and abasic sites in DNA has been observed with several purified prokaryotic DNA polymerases and viral DNA polymerases *in vitro* (Loeb, 1985; Rabkin *et al.*, 1983). Although similar studies have not been reported for DNA polymerases purified from mammalian cells, Kunkel (Kunkel, 1985a; Kunkel, 1985b; Kunkel and Alexander,

Table II. Base changes (per cent) observed with purified eukaryotic DNA polymerases compared with our data

Base change	Purified DNA polymerase[a]		Our data
	α	β	
G·C→A·T	43	47	64
G·C→T·A	23	7.6	28
G·C→C·G	16	3.4	6.8
A·T→G·C	8.1	31	0
A·T→C·G	1.9	6.9	0
A·T→T·A	8.5	4.1	1.3

[a]Data compiled from Kunkel and Alexander (1986). Primed phage M13mp2 DNA (containing the *lacZα* gene) was used as a template for DNA synthesis. *LacZ* α-mutants were isolated and sequenced. Although each single base change was monitored separately in these studies, the base changes have been combined here for ease of comparison.

1986) has reported DNA sequence analyses of the base misincorporations observed after replication of a natural template by purified eukaryotic DNA polymerases. The results of his studies are shown in Table II for comparison. Although Kunkel observed a wider range of base substitutions than we find, he also noted a predominance of the G·C→A·T transition resulting largely from misincorporation of A opposite C (Kunkel and Alexander, 1986). In that study, the G·C→T·A transversion was also observed to occur relatively frequently, mostly due to misincorporation of A opposite G. In contrast, we have postulated that the G·C→T·A transversions we observe are due to misincorporation of T opposite C.

What could cause the loss of replication fidelity? We have postulated that the spontaneous point mutations we observe arise from a localized loss of fidelity in DNA synthesis occurring during the repair of transfected DNA damaged by cellular nucleases. It is possible that the transfection conditions *per se* induce a mutagenic state in the cells that reduces the fidelity of repair synthesis. For example, a generalized effect on the intracellular nucleotide pools or ionic environment might play a role. Indeed, the dGTP pool appears to be rate limiting for DNA replication (Leeds *et al.*, 1985) so that a reduction in this pool might cause misincorporation of bases opposite C, resulting in the observed preference for mutations at G·C base pairs. Alternatively, a lack of supercoiling or the absence of the normal nucleosomal structure in the transfected DNA might influence the fidelity of DNA replication. A final possibility is that the reduced fidelity we observe here is not unusual for mammalian cells, but that misincorporated bases are normally removed by an active mismatch repair system (Hare and Taylor, 1985). Since the transfected DNA would lack the normal mammalian methylation patterns, methyl-directed mismatch repair would not be selective for the newly replicated DNA strand. Further experimentation will be required to determine whether these mutational characteristics are general for all types of transfected DNA or whether they relate more specifically to the particular type of plasmid used here. For example, DNA length, circularity or extent of supercoiling may influence the susceptibility of transfected DNA to base substitution mutagenesis. Furthermore, subsequent replication of the transfected DNA as an extrachromosomal element may further influence the level and type of spontaneous mutagenesis.

It is of interest to compare the spontaneous point mutations described here with those induced by u.v. radiation in the same shuttle vector system and described earlier (Hauser *et al.*, 1986). The u.v.-induced mutations occurred predominantly at G·C base pairs but about 20% of the changes were at A·T sites. In addition, among the u.v.-induced point mutants 75% had either single (58%) or tandem double (17%) mutations, whereas only 33% of the spontaneous mutants had single mutations and only one tandem double mutation (with no additional mutations) was found. Only 5% of the u.v.-induced mutants contained three or more mutations, whereas almost 50% of the spontaneous mutants had three or more changes. It is clear that the characteristics of the u.v.-induced mutants are quite different from those of the spontaneous mutants, consistent with their induction as a consequence of the presence of u.v. photoproducts in the DNA. However, both types of mutations are predominantly G·C→A·T transitions and in both cases we ascribed this to a tendency of the DNA polymerase to insert A under conditions of reduced fidelity either during replication of u.v.-damaged DNA or during repair of transfected DNA. Further investigation of the mechanism of mutagenesis of transfected DNA using this shuttle vector system should yield more insight into the nature and function of eukaryotic DNA polymerases, into the mechanisms of induced mutagenesis, and into the reasons for the extreme plasticity of transfected DNA.

Materials and methods

Cells and plasmids

The TC7 clone (Robb and Huebner, 1973) of the African Green Monkey kidney cell line CV-1 was grown in Dulbecco's modified Eagle's medium (GIBCO Laboratories, Grand Island, NY) supplemented with antibiotics with 5% fetal calf serum. *E. coli* MBM7070 (Seidman *et al.*, 1985) has the genotype F^-, *lacZam*CA7020, *lacY*i, *hsd*R^-, *hsd*M^+, *araD*139, Δ(*araABC-leu*)7679, *galU*, *galK*, *rpsL*, *thi*. In the presence of isopropyl-β-D-thio-galactoside, an inducer of the *lac* operon, and 5-bromo-4-chloro-3-indolyl-β-D-galactoside, an artificial substrate for β-galactosidase, strain MBM7070 forms blue colonies if it contains an active *supF* suppressor tRNA gene and white colonies if the suppressor is inactive. The plasmid pZ189 (Seidman *et al.*, 1985) was purified by CsCl equilibrium sedimentation by Loftstrand Labs Ltd, Gaithersburg, MD.

Mutant isolation

Subconfluent TC7 cells growing in 100-mm plastic tissue culture dishes were transfected with purified plasmid DNA (5–25 ng per dish) by the DEAE–dextran method (McCutchen and Pagano, 1968). Forty-eight hours later, plasmid DNA was harvested from the cells by the Hirt (1967) method, purified and treated with *Dpn*I (Bethesda Research Laboratories or New England Biolabs) to remove all nonreplicated DNA (Peden *et al.*, 1980). After additional purification, plasmid DNA was used to transform the bacterial tester strain, MBM7070. DNA from over 20 separate transfections was used to obtain the collection of spontaneous mutants described here. Transformation of bacteria, selection of colonies containing the *supF*$^-$ mutant plasmid, and purification of plasmid DNA from bacteria have been described previously (Hauser *et al.*, 1986). Double-stranded plasmid DNA from the *supF*$^-$ mutants was sequenced according to a modification (Zagursky *et al.*, 1985) of the Sanger *et al.* (1977) dideoxyribonucleotide sequencing method.

Acknowledgements

We thank Drs L.A.Loeb, T.A.Kunkel, and N.Nossal for their criticisms and comments on the manuscript. We thank F.Yamada for his assistance in preparing the figures with the DNADRAW program. The expert typing assistance of Evelyn Burrell is gratefully acknowledged.

References

Bredberg,A., Kraemer,K.H. and Seidman,M.M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8273–8277.

Calos,M.P., Lebkowski,J.S. and Botchan,M.R. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3015–3019.

Hare,J.T. and Taylor,J.H. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7350–7354.

Hauser,J., Seidman,M.M., Sidur,K. and Dixon,K. (1986) *Mol. Cell Biol.*, **6**, 277–285.

Hirt,B. (1967) *J. Mol. Biol.*, **20**, 365–369.

Kunkel,T.A. (1985a) *J. Biol. Chem.*, **260**, 5787–5796.

Kunkel,T.A. (1985b) *J. Biol. Chem.*, **269**, 12866–12874.

Kunkel,T.A. and Alexander,P.S. (1986) *J. Biol. Chem.*, **261**, 160–166.

Lebkowski,J.S., DuBridge,R.B., Antyell,E.A., Griesen,K.K. and Calos,M.P. (1984) *Mol. Cell. Biol.*, **4**, 1951–1960.

Leeds,J.M., Slabaugh,M.B. and Mathews,C.K. (1985) *Mol. Cell. Biol.*, **5**, 3443–3450.

Loeb,L.A. (1985) *Cell*, **40**, 483–484.

Loeb,L.A. and Kunkel,T.A. (1982) *Annu. Rev. Biochem.*, **52**, 429–457.

McCutchen,J.A. and Pagano,J.S. (1968) *J. Natl. Cancer Inst.*, **41**, 351–357.

Miller,J.H. (1985) *J. Mol. Biol.*, **182**, 45–68.

Miller,J.H., Lebkowski,J.S., Griesen,K.S. and Calos,M.P. (1984) *EMBO J.*, **3**, 3117–3121.

Peden,K.W.C., Pipas,J.M., Pearson-White,S. and Nathans,D. (1980) *Science*, **209**, 1392–1396.

Protić-Sabljić,M., Tuteja,N., Munson,P.J., Hauser,J., Kraemer,K.H. and Dixon,K. (1986) *Mol. Cell. Biol.*, **6**, 3349–3356.

Rabkin,S.D., Moore,P.D. and Strauss,B.S. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1541–1545.

Razzaque,A., Chakrabarti,S., Joffe,S. and Seidman,M.M. (1984) *Mol. Cell. Biol.*, **4**, 435–441.

Razzaque,A., Mizusawa,H. and Seidman,M.M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3010–3014.
Robb,J.A. and Huebner,R.J. (1973) *Exp. Cell Res.*, **81**, 120–126.
Sanger,F., Nicklen,S. and Coulson,A. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5476.
Seidman,M.M., Dixon,K., Razzaque,A., Zagursky,R. and Berman,M.L. (1985) *Gene*, **28**, 223–237.
Shapiro,M.B. and Senapathy,P. (1986) *Nucleic Acids Res.*, **14**, 65–73.
Topal,M.D., DiGuiseppi,S.R. and Sinha,N.K. (1980) *J. Biol. Chem.*, **255**, 11717–11724.
Wake,C.T., Gudewicz,T., Porter,T., White,A. and Wilson,J.H. (1984) *Mol. Cell Biol.*, **4**, 387–398.
Weatherall,D.J. (1984) *Lancet*,**ii**, 1440–1444.
Zagursky,R.J., Berman,M.L., Baumeister,K. and Lomax,N. (1985) *Gene Anal. Tech.*, **2**, 89–94.

Received on 29 July 1986; revised on 28 October 1986