Normanly, J., Ogden, R., Horvath, S., & Abelson, J. (1986) *Nature 321*, 213-219.
Quigley, G. J., & Rich, A. (1976) *Science 194*, 796-806.
Reilly, R. M., & RajBhandary, U. L. (1986) *J. Biol. Chem. 261*, 2928-2935.
Reyes, V. M., & Abelson, J. (1988) *Cell 55*, 719-730.
Riesner, D., Pingoud, A., Boehme, D., Peters, F., & Maass, G. (1976) *Eur. J. Biochem. 68*, 71-80.
Rigler, R., Pachmann, U., Hirsh, R., & Zachau, H. G. (1976) *Eur. J. Biochem. 74*, 307-315.
Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F. C., & Klug, A. (1974) *Nature 25*, 546-551.
Roe, B., Sirover, M., & Dudock, B. (1973) *Biochemistry 12*, 4146-4154.
Romby, P., Moras, D., Bergdoll, M., Dumas, P., Vlassov, V., Westhof, E., Ebel, J. P., & Giegé, R. (1985) *J. Mol. Biol. 184*, 455-471.
Sampson, J. R., & Uhlenbeck, O. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1033-1037.
Sampson, J. R., DiRenzo, A. B., Behlen, L. S., & Uhlenbeck, O. C. (1989) *Science 243*, 1363-1366.
Schevitz, R. W., Padjarny, A., Krishnamachari, N., Hughes, J., & Sigler, P. (1979) *Nature 278*, 188-190.
Schoemaker, H. J., & Schimmel, P. R. (1976) *J. Biol. Chem. 251*, 6823-6830.
Schulamn, L., & Pelka, H. (1988) *Science 242*, 765-767.
Starzyk, R., Koontz, S., & Schimmel, P. (1982) *Nature 298*, 136-140.
Starzyk, R., Schoemaker, H., & Schimmel, P. R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 339-342.
Westhof, E., Dumas, P., & Moras, D. (1985) *J. Mol. Biol. 184*, 119-145.
Woo, N. H., Roe, B. A., & Rich, A. (1980) *Nature 286*, 346-351.

A Sensitive Genetic Assay for the Detection of Cytosine Deamination: Determination of Rate Constants and the Activation Energy†

Lisa A. Frederico,‡ Thomas A. Kunkel,§ and Barbara Ramsay Shaw*,‡

Department of Chemistry, Duke University, Durham, North Carolina 27706, and Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

Received August 14, 1989; Revised Manuscript Received October 19, 1989

ABSTRACT: Previously it has not been possible to determine the rate of deamination of cytosine in DNA at 37 °C because this reaction occurs so slowly. We describe here a sensitive genetic assay to measure the rate of cytosine deamination in DNA at a single cytosine residue. The assay is based on reversion of a mutant in the *lacZα* gene coding sequence of bacteriophage M13mp2 and employs *ung*⁻ bacterial strains lacking the enzyme uracil glycosylase. The assay is sufficiently sensitive to allow us to detect, at a given site, a single deamination event occurring with a background frequency as low as 1 in 200 000. With this assay, we determined cytosine deamination rate constants in single-stranded DNA at temperatures ranging from 30 to 90 °C and then calculated that the activation energy for cytosine deamination in single-stranded DNA is 28 ± 1 kcal/mol. At 80 °C, deamination rate constants at six sites varied by less than a factor of 3. At 37 °C, the cytosine deamination rate constants for single- and double-stranded DNA at pH 7.4 are 1×10^{-10} and about 7×10^{-13} per second, respectively. (In other words, the measured half-life for cytosine in single-stranded DNA at 37 °C is ca. 200 years, while in double-stranded DNA it is on the order of 30 000 years.) Thus, cytosine is deaminated ~140-fold more slowly when present in the double helix. These and other data indicate that the rate of deamination is strongly dependent upon DNA structure and the degree of protonation of the cytosine. The data suggest that agents which perturb DNA structure or facilitate direct protonation of cytosine may induce deamination at biologically significant rates. The assay provides a means to directly test the hypothesis.

The integrity of genetic information depends upon the inherent chemical stability of bonds in DNA. Reactions involving deamination of bases, imidazole ring opening, or hydrolysis of phosphodiester or glycosylic bonds can result in mutagenesis, cell transformation, and even cell death (Friedberg, 1984). In particular, deamination of cytosine and 5-methylcytosine in DNA generates the bases uracil and thymine. Replication of these deamination products will produce a C·G → T·A transition mutation (Coulondre et al., 1978; Duncan & Miller, 1980).

Previous analytical methods lacked the sensitivity to determine the rate of deamination of cytosine at 37 °C because this reaction occurs so slowly. Therefore, until now, the spontaneous *rate* of cytosine deamination had been estimated only from chemical measurements of the amount of uracil produced from DNA incubated either at high temperatures (Lindahl & Nyberg, 1974) or at extremes of pH (Wang et al., 1982; Ullman & McCarthy, 1973; Shapiro & Klein, 1966). Our continuing interest in understanding the chemical basis

† Parts of this work were suspported by Grant BC-617 from the American Cancer Society and Grant CA44709 from the NIH to B.R.S. B.R.S. was also supported by a Faculty Development Award from the American Cancer Society.
* Correspondence should be addressed to this author.
‡ Duke University.
§ National Institute of Environmental Health Sciences.

0006-2960/90/0429-2532$02.50/0

for mutational phenomena has led us to develop a sensitive genetic assay to detect and quantitate cytosine deamination in DNA incubated under a variety of conditions.

We describe herein an assay that allows one to accurately assess the rates of deamination of cytosine in both single- and double-stranded DNA under conditions of temperature and ionic strength that closely approximate those that exist in vivo. The detection system is based on reversion of a mutant of bacteriophage M13mp2, from a colorless to blue plaque phenotype, resulting from C → T transitions within the *lacZα* gene coding sequence. This approach is highly sensitive; deamination of a single cytosine residue can be detected.

The method is used here to determine the rate constant for cytosine deamination over a wide temperature range. These data are then used to calculate the activation energy for cytosine deamination.

Experimental Procedures

Materials

Escherichia coli strains NR8051 [Δ(*pro-lac*), *thi*, *ara*] and NR8052 [Δ(*pro-lac*), *thi*, *ara*, *trpE9777*, *ungl*] were used for transfections (Kunkel, 1985c). The *ung* phenotype of both strains was confirmed during each transfection experiment using normal and uracil-containing DNA controls (Kunkel, 1985c). *E. coli* strain CSH50 [Δ(*pro-lac*), *thi*, *ara*/F′(*tra*D36, *pro*AB, *lac*I^{q}ΔM15] and bacteriophage M13mp2 were from J. E. LeClerc (University of Rochester). *E. coli* strain NR9099 [Δ(*pro-lac*), *rec*A56, *thi*, *ara*/F′ (*pro*AB, *lac*I$^-$qΔ-M15] was from Roeland Schaaper (NIEHS). Mutant M13mp2 derivatives were from previously described mutant collections (Kunkel & Alexander, 1986; Kunkel, unpublished observations).[1] Homogeneous *E. coli* uracil DNA glycosylase was a kind gift from Bruce Duncan (Institute of Cancer Research, Philadelphia, PA). All other reagents were from previously described sources.

Methods

M13mp2 DNA substrates were prepared as described (Kunkel, 1985a; Kunkel & Soni, 1988). Briefly, bacteriophage M13mp2C141 was plated on minimal medium plates using *E. coli* NR9099 as a host strain. A single plaque was added to 1 L of 2 × YT medium (containing, per liter, 16 g of Bacto-Tryptone, 10 g of yeast extract, and 5 g of NaCl, pH 7.4) containing 10 mL of an overnight culture of *E. coli* NR9099. M13mp2-infected cells were grown overnight at 37 °C with vigorous shaking, and cells were harvested by centrifugation at 4000*g* for 30 min at 0 °C in a Beckman J-6M centrifuge with a 5.2 rotor, without application of the brake for deceleration. Phage in this supernatant were used for isolation of single-stranded DNA described below. The pellet was resuspended in 0.9% NaCl and used to prepare double-stranded DNA by the method of Birnboim and Doly (1979) (with two exceptions: after addition of LiCl, the RF DNA sample was incubated on ice for 15 min instead of at 60 °C in order to reduce deamination, and after the RNase A and T1 treatment, the DNA in 100 mM NaCl was simply phenol extracted, ethanol precipitated, and resuspended as described for the single-stranded DNA below). Single-stranded DNA was prepared from the first supernatant by precipitating phage over a period of 1–3 h at 4 °C by addition of one-fourth volume (200 mL) of a mixture of 15% poly(ethylene glycol) 8000 (PEG) and 2.5 M NaCl. The phage pellet, obtained by centrifuging in the Beckman J-6M centrifuge with the 5.2 rotor at 5000 rpm for 20 min at 0 °C, was resuspended in 7.5 mL of phenol extraction buffer (100 mM Tris-HCl, pH 8.0, 300 mM NaCl, and 1 mM EDTA) and extracted with an equal volume of phenol equilibrated with phenol extraction buffer and containing 0.08% 8-hydroxyquinoline. The top aqueous layer was reextracted with phenol, followed by two chloroform/isoamyl alcohol (24:1) extractions. The DNA was precipitated with ethanol and resuspended in TE buffer (10 mM Tris-HCl, pH 8.0, and 1 mM EDTA). These procedures yield 3–6 mg of single-stranded and 1–3 mg of double-stranded DNA per liter of culture.

Incubation of DNA. Either single-stranded or double-stranded M13mp2 DNA was sealed in a capillary tube with sample buffer B (0.1 M KCl, 0.05 M Hepes–KOH, 0.01 M $MgCl_2$, and 0.001 M EDTA, pH 7.4) of Lindahl and Nyberg (1974) at a concentration of 20 ng/μL. Typically, 150-μL samples were incubated for each time point. The samples were incubated in a water bath in the dark for the indicated time and at the desired temperature (±1 °C) and then stored at −70 °C until used for transfection.

Uracil Glycosylase Treatment of DNA. In certain experiments, incubated DNA was treated with uracil DNA glycosylase prior to transfection. The reaction was carried out with 19 μL of a solution containing 50 ng/μL DNA (i.e., 20 nM M13mp2 DNA), 10 mM Tris-HCl (pH 8.0), and 1 mM EDTA to which was added 1 μL, i.e., 200 nM (a 10-fold molar excess), of *E. coli* uracil DNA glycosylase stored in 10 mM Tris-HCl, 0.1 mM EDTA, and 20% glycerol. The sample was incubated at 37 °C for 30 min. Reactions were placed on ice and transfected immediately.

Transfection and Plating. DNA samples were used to transfect competent NR8051 (*ung*$^+$) or NR8052 (*ung*$^-$) *E. coli* cells (Kunkel, 1984). Transfected cells were plated on minimal media plates containing 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-Gal)[2] and IPTG, and revertants were scored and, where indicated, sequenced (Kunkel, 1985a, Kunkel & Soni, 1988). The efficiency of transfection averaged 2×10^{-7} [50 plaques/ng (i.e., 2.5×10^8 molecules) for unheated control single-stranded DNA] and varied 2–5-fold. Double-stranded DNA had a transfection efficiency of about 3000 plaques/ng. Transfections were carried out by diluting the DNA into 75 mM $CaCl_2$ and adding 2 volumes of competent cells. The volumes used were determined by the ratio of DNA molecules to cells, which was kept constant for all variables in a single experiment, typically at 20:1. After a 40-min incubation at 0 °C, the mixture was heat-shocked at 42 °C for 3 min and then placed at 0 °C. Transfected cells were plated on minimal media plates containing 1.5% agar, 2% VB salts, and 0.24 mg of IPTG per plate. Plating was performed with a soft agar overlay containing 3 mL of melted 0.8% agar in 0.9% NaCl containing 1.5 mg of 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-Gal). Just before pouring, we added 0.2 mL of a log-phase culture of *E. coli* CSHF′50 cells containing 0.24 mg of IPTG and an appropriate amount (varying from 2 to 100 μL) of the transfection mixture to produce approximately 1000–5000 infected cells per plate. For quantitation of total plaques, smaller aliquots of the transfection mixture were used.

The host cell strains were tested for uracil glycosylase repair capacity during each transfection by using a mixture of wild-type uracil-containing DNA (Kunkel, 1985c) (which upon transfection and plating produces dark blue plaque phenotypes) and M13mp2C141 DNA[3] (which upon transfection produces

[1] T. A. Kunkel, unpublished observations.

[2] Abbreviations: IPTG, isopropyl β-D-thiogalactoside; X-Gal, 5 bromo-4-chloro-3-indolyl β-D-galactoside; wt, wild type.

colorless plaques). With this control mixture, the *ung*$^+$ cells should produce only colorless phenotype plaques while *ung*$^-$ should produce both wild-type blue and colorless phenotype plaques. In *ung*$^+$ cells, uracil residues (including any resulting from deamination of cytosine) are removed by uracil glycosylase to produce an abasic site; this abasic site gives rise to biologically inactive M13 DNA in the transfection assay, due to the subsequent cleavage of the abasic site by cellular repair endonucleases. The *ung*$^-$ cell strain (NR8052) will retain the deaminated cytosines and therefore the wild-type blue phenotype; conversely, the *ung*$^+$ cell strain (NR8051) will give rise solely to the colorless plaque phenotype. In this way, we can confirm the *ung* phenotype of the host cell strain.

The plating conditions were carefully controlled to ensure that loss of visualization of the wild-type phenotype would not occur. First, the maximum number of plaques per plate was determined by varying the amount of C141 phage while maintaining the wild-type plaques at approximately 100 per plate. By plotting the percentage of wild-type plaques observed versus the total plaques per plate, we found that the maximum number of plaques that may be plated without significant loss of visualization was 10000 plaques per plate. Therefore, we chose to perform all plating experiments with approximately 5000 plaques per plate. Second, the plating conditions of the experiment were tested to ensure that visualization of the revertant would be linear with increasing concentration of the mutant, as follows. Combinations of C141 and wild-type phage, ranging from a 1:1 to a 5000:1 ratio of C141 DNA to uracil-containing wild-type DNA, were plated at a maximum of 5000 plaques per plate. On a ln/ln plot comparing the percentage of wild-type plaques of uracil-containing wild-type DNA, the system gave a linear response with slope equal to 1.0. Thus, the plating conditions are not obscuring the detection of the wild-type phenotype plaques, which in our experiments would be equivalent to deaminated cytosine residues. These experiments demonstrated the ability of the assay to detect a small fraction of deaminated cytosines (dark blue phenotypes) among a large population of undeaminated (colorless plaque) molecules.

Scoring Revertants and Determination of DNA Sequences. After 12–15 h of incubation at 37 °C, C141 revertant[3] plaques appear as dark blue plaques against a colorless or faint blue background. Revertant plaques are picked from the plate, diluted in 50 mM sodium borate, and replated as above to eliminate false positives. All revertant phenotypes are further confirmed by plating the dilution with wild-type M13mp2 phage. To determine the exact sequence of the revertants, revertant blue phenotype plaques are picked at random and sequenced with the chain terminator method (Sanger et al., 1977), using a 15-base primer (New England BioLabs) complementary to the coding sequence for amino acids 46–50 of the *lacZ* gene.

DNA viability (i.e., survival of plaque-forming ability) can be determined from the total number of plaques (as determined for calculation of the reversion frequency) obtained from transfecting a known amount of DNA into both *ung*$^-$ and *ung*$^+$ cells. Since the transfection efficiencies of the two cell strains can differ, they must be normalized; this is done by transfecting each cell strain with an unheated control DNA denoted as the 0 sample. For example (in Table I), the *ung*$^-$ cell strain transfected at approximately 70% the efficiency of the *ung*$^+$ cell strain. Therefore, the corrected number of plaques per milliliter at the 4-h time point for *ung*$^-$ cells can be obtained by multiplying the number of plaques at the 4-h time point by 1.4 (i.e., 100/70).

[3] Mutant M13mp2C141 has a cytosine in place of guanine at position 141 of the lac Zα gene, and it will be referred to simply as C141. Likewise, the other M13 mutants will be referred to first by the base and then by the position that differs from the wild type. Position 1 is the first transcribed base. See Kunkel (1984) for a description of the gene and its sequence.

Table I: Deamination of Single-Stranded C141 DNA at 80 °C

incubation time (h)	cell type transfected	plaques scored: total (×10^4)	plaques scored: blue	reversion frequency[c] (×10^{-5})
0	*ung*$^-$	220	25	1.1
0	*ung*$^+$	95	13	1.4
4	*ung*$^-$	6.5	95[b]	150
4	*ung*$^+$	4.5	4	8.9
4[a]	*ung*$^-$	4.1	6	15

[a] Treated with uracil DNA glycosylase prior to transfection. [b] Sequencing of 32 blue revertants showed that 17 were at position 141 and 15 were at position 142. [c] Reversion frequencies were determined as described under Experimental Procedures. A decreased survival after incubation at 80 °C for 4 h presumably results from depurination (Schaaper & Loeb, 1981).

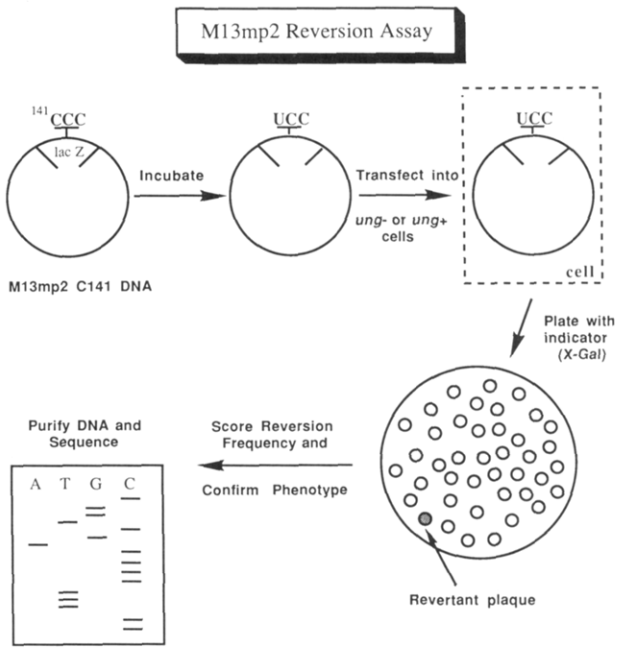

FIGURE 1: Experimental design of the genetic assay for cytosine deamination. See text for explanation. While the example shown here is for single-stranded M13mp2 DNA from mutant C141, the principle for detection of deamination events is the same when using DNA derived from other mutants.

Results

Principles of the Assay for Cytosine Deamination. Cytosine deamination is monitored by using mutant derivatives of the single-stranded DNA-containing bacteriophage M13mp2. As one example, we began with a mutant containing a G → C base change at position 141 of the *lacZ*α gene; the mutant is therefore designated C141. This creates a CCC (proline) codon for the 34th amino acid of the α-peptide, which yields a defective polypeptide unable to complement β-galactosidase activity in infected *E. coli* host cells. Thus, whereas cells infected with the normal wild-type M13mp2 generate a blue plaque phenotype when plated on X-Gal indicator plates, cells infected with C141 mutant phage yield plaques that are colorless.

The first two cytosine residues of the CCC codon are the targets for scoring cytosine deamination. (In this codon, all bases at the third position code for the same amino acid, and

thus all changes are phenotypically silent.) The experimental outline is shown in Figure 1. C141 single-stranded DNA is extracted from intact virus particles, and double-stranded DNA is purified from cells infected with this same mutant. Either form of DNA is incubated in vitro for the desired time and at the indicated temperature. Deamination of cytosine to uracil at either position 141 or position 142 produces a wild-type codon, resulting in a blue plaque. The DNA is used to transfect a strain of *E. coli* defective in repair of uracil (an *ung*⁻ strain). Since the uracil cannot be removed, it codes like thymine during the first round of DNA replication in vivo, eventually generating, within the infected cell, progeny that contain either a TCC or a CTC codon (due to C → T transitions at positions 141 or 142, respectively). These encode serine and leucine, respectively, both of which result in a functional protein and produce a blue plaque phenotype. Thus, cytosine deamination events (i.e., revertants) are scored as blue plaques in the field of colorless plaques. Molecules that have deaminated elsewhere but not at either of the two detectable sites will give colorless plaques.

Deamination at 80 °C. The assay was first used to detect cytosine deamination in single-stranded C141 DNA at 80 °C, since these conditions were previously found to produce detectable deamination using chemical procedures (Lindahl & Nyberg, 1974). In order to facilitate the comparison of rate constants, the C141 DNA was incubated in the same buffer B as in their study. The results are shown in Table I. In *ung*⁻ cells, which are unable to remove uracil due to a defect in the uracil glycosylase gene, the reversion frequency[5] of single-stranded DNA incubated for 4 h at 80 °C was 150×10^{-5}, a 140-fold increase over the unincubated control. When an aliquot of the same DNA sample was used to transfect *ung*⁺ cells, the reversion frequency was 17-fold lower and only slightly above the background, indicating removal of uracil residues by endogenous uracil glycosylase. Furthermore, when an aliquot of the same DNA was treated with purified uracil glycosylase and then used to tranfect *ung*⁻ cells, the reversion frequency was decreased 10-fold. The observation that >90% of the mutagenesis can be reversed by uracil glycosylase treatment either in vivo or in vitro strongly suggests that the lesion responsible for the mutations is the uracil which results from incubation of the DNA at 80 °C. Consistent with this interpretation, DNA sequence analysis of 32 blue revertants recovered from transfection of *ung*⁻ cells with the heat-treated DNA showed that all revertants contained a single C → T transition mutation expected from deamination mutagenesis. Seventeen were at position 141, and 15 were at position 142.

Since the reversion frequency in the *ung*⁻ cell strain increased linearly with time of incubation at 80 °C (data not shown, but see below for 37 °C incubation), the rate constant (k) for cytosine deamination can be calculated from the equation

$$k = \frac{\ln C}{t} = \frac{-\ln (1 - F)}{t} \quad (1)$$

where t is time of incubation and F is the corrected reversion frequency (obtained by subtracting the background reversion frequency of the unincubated control from the reversion frequency of the heat-incubated sample). Also, since there are two detectable cytosine sites for deamination using the C141 mutant, the corrected frequency is divided by 2 to calculate a normalized rate constant per cytosine residue. The single-stranded deamination rate constant at 80 °C, calculated from the data in Table I, is 4.2×10^{-8}/s. This is in reasonable agreement with chemically determined 80 °C rate constants [3.3×10^{-8}/s and 1.1×10^{-8}/s for cytosine deamination in poly(dG)·poly(dC) and denatured *E. coli* DNA, respectively, as reported by Lindahl and Nyberg (1974)].

[4] L. A. Frederico, T. A. Kunkel, and B. Ramsay Shaw, manuscript in preparation.

[5] Reversion frequency is defined as the number of blue plaque revertants divided by the total number of plaques.

Table II: Reversion Frequency at Five Additional Cytosine Targets[a]

cytosine position[b]	surrounding sequence[c]	plaques scored: total (×10⁴)	plaques scored: blue	reversion frequency (×10⁻⁵)
−36	GCCTT	19	209	110
−35	CTCTA	18	220	120
87	ACCGG	14	190	140
112	ACCTA	11	104	95
147	GCCGG	14	77	55

[a] The indicated mutant single-stranded DNA was incubated for 4 h at 80 °C and then used to transfect *ung*⁻ host cells as described under Experimental Procedures. The reversion frequencies obtained from transfection of the unincubated control DNAs varied from 0.6×10^{-5} (for C147) to 4.5×10^{-5} (for C-36). [b] Position 1 is the first transcribed base of the *lacZα* gene. For the entire wild-type M13mp2 *lacZα* sequence, refer to Kunkel (1985b). [c] The target cytosine is underlined.

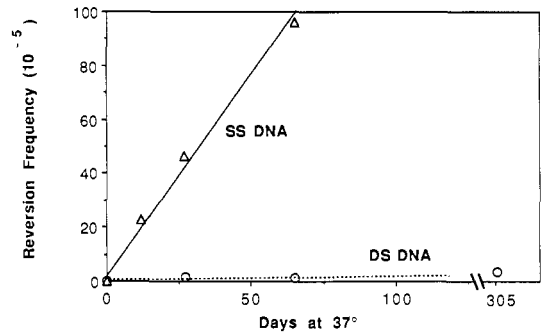

FIGURE 2: Deamination of single- and double-stranded C141DNA at 37 °C in *ung*⁻ cells. The results with single-stranded DNA (Δ) represent the average of at least three experiments for each time point. The results with double-stranded DNA (O) (Frederico-Zuraw, 1988) represent data points with at least 120 000–1 300 000 total plaques and from 2 to 9 revertant plaques per data point.

Cytosine Deamination at Five Additional Sites. To determine whether this result was specific for cytosines at positions 141 and 142 or typical of cytosines residing within a variety of different neighboring DNA sequences, we examined the rate of cytosine deamination at five additional sites, located at positions −36, −35, 87, 112, and 147 within the *lacZα* gene. These sites were chosen because a cytosine at these positions creates a very faint blue plaque phenotype and because C → T transitions (in each case at only a single position) can be scored as dark blue revertants. Thus, the principles of the assay remain as shown in Figure 1.

Incubation of mutant single-stranded DNAs at 80 °C for 4 h followed by transfection of *ung*⁻ cells resulted in substantial increases in the reversion frequencies over the unincubated controls (Table II). As before, only small increases were observed upon transfection of *ung*⁺ cells (data not shown). The reversion frequencies in *ung*⁻ cells differed by less than 3-fold at the six sites examined, suggesting that there is no strong site or sequence specificity to cytosine deamination mutagenesis in single-stranded DNA at 80 °C. Yet, it is possible that a preference might be revealed upon a more extensive examination of the site and temperature dependence of deamination, and in experiments in progress, we are examining whether there are site-to-site variations.

Deamination at 37 °C. The sensitivity of the genetic reversion assay permits a determination of the cytosine deam-

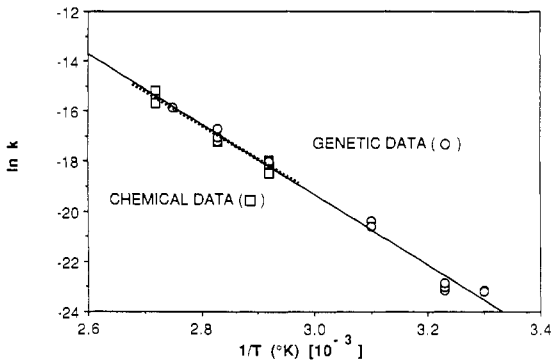

FIGURE 3: Arrhenius plot for determination of the activation energy for cytosine deamination. Data from the genetic assay (in Table IV) are plotted as circles, while the published high temperature chemical rate constant data [obtained with poly(dG)·poly(dC), poly(dC), and dCMP by Lindahl and Nyberg (1974)] are indicated by squares. The activation energy calculated from the slope of the genetic data (—) is 28 kcal/mol. The activation energy from the chemical data (---) is also 28 kcal/mol (which differs slightly from the 29 kcal/mol reported by Lindahl and Nyberg for the same data).

ination rate for DNA incubated at 37 °C in pH 7.4 buffered solution. The reversion frequencies for single- and double-stranded C141 DNA as a function of incubation time at 37 °C are shown in Figure 2. For single-stranded DNA, the reversion frequency increases linearly with time, such that, after 60 days, the frequency is >40-fold above the background reversion frequency. As before, transfection of an aliquot of this DNA into *ung*$^+$ cells produced a reversion frequency only slightly above the background.

In contrast to the substantial mutagenesis observed with single-stranded DNA, transfection of an *ung*$^-$ cell with double-stranded DNA incubated at 37 °C, even for almost a year, produced only a slight (less than 5-fold) increase in reversion frequency (Figure 2). To eliminate the possibility that cytosine deamination had been followed by mismatch correction of the uracil in the resulting U·G mispair prior to replication within the transfected cell, we performed control transfections (not shown) with defined mismatched heteroduplexes.[4] These experiments are similar to that described by Kunkel and Soni (1988) using mismatched heteroduplexes (in which each strand coded for a different shade of blue) and transfecting these heteroduplexes into *ung*$^-$ cells. The results clearly demonstrated that, were deamination to occur in double-stranded DNA, it should easily be detected upon transfection. However, it is possible that some of the deaminations could be repaired after tranfection by processes like mismatch correction or nick translation. If this were to happen, the derived rate constant for deamination in double-stranded DNA could be somewhat (albeit less than 2-fold) higher than that measured. To summarize, cytosine deamination at 37 °C occurred much more rapidly in single-stranded DNA than in double-stranded DNA. The actual rate constants for single-stranded and double-stranded DNA calculated from the reversion frequencies are 1×10^{-10}/s and $\sim 7 \times 10^{-13}$/s, respectively.

Determination of Activation Energy for Cytosine Deamination. The reversion frequencies were measured for single-stranded C141 DNA at five additional temperatures ranging from 30 to 90 °C (Table III). The results were used to calculate the cytosine deamination rate constant at each temperature (Table III) in order to produce the Arrhenius plot shown in Figure 3. For comparison, Figure 3 shows both our genetically determined results as well as the chemically determined results of Lindahl and Nyberg (1974) for poly(dC), poly(dG)·(dC), and dCMP. The activation energy, calculated from our genetic data, is 28 kcal/mol. This value is similar to the value of 29 kcal/mol obtained by Lindahl and Nyberg (1974) using the chemically determined rate constants over a smaller temperature range and at higher temperatures, i.e., 70, 80, and 95 °C.

Table III: Reversion Frequencies and Cytosine Deamination Rate Constants Determined at Six Temperatures[a]

incubation temp (°C)	time of incubation	plaques scored: total ($\times 10^4$)	plaques scored: blue	reversion frequency ($\times 10^{-5}$)	deamination rate constant (per site) (s^{-1})
30	63 days	32	334	100	0.83×10^{-10}
	132 days	17	335	200	0.88×10^{-10}
37	12 days	59	68	23	1.2×10^{-10}
	26.8 days	36	152	42	1.0×10^{-10}
	65 days	9.6	92	96	0.88×10^{-10}
50	6.5 h	19	14	7.4	1.1×10^{-9}
	16 h	21	38	18	1.4×10^{-9}
	42 h	22	80	37	1.1×10^{-9}
70	5 h	17	86	51	1.5×10^{-8}
80	13 min	78	60	7.6	3.9×10^{-8}
	130 min	38	162	43	3.3×10^{-8}
	240 min	6.5	95	150	5.6×10^{-8}
90	1 h	6.4	60	94	1.3×10^{-7}

[a] After incubation of single-stranded C141 DNA at the indicated temperatures, reversion frequencies were determined by transfection of *ung*$^-$ cells as described under Experimental Procedures. Each reversion frequency represents the average value of three or more separate transfection experiments, with the exception of the data at 30 °C for 63 days which is the average of two experiments. Because the number of plaques in each experiment could differ, the reversion frequency in the table is close to, but does not necessarily correspond to, that which can be obtained from dividing the total revertants by the total plaques. The background reversion frequency of the unincubated control DNA (averaged for 10 separate experiments) was 1.1×10^{-5} for the samples at 30, 37, and 80 °C; the corresponding background reversion frequency (averaged for four separate experiments) was 3.7×10^{-5} for the 50, 70, and 90 °C samples and 5.3×10^{-5} for the 30 °C samples. The appropriate background frequency was subtracted from the measured reversion frequency for calculation of rate constants, as described in the text. Note that two cytosines in the codon can deaminate; the rate constants given have been normalized for deamination of one cytosine.

Discussion

We have developed a sensitive assay to determine the rate of cytosine deamination in DNA. The sensitivity of the method is limited only by the background reversion frequency of the phage DNA preparation (ca. 2×10^{-5}). The method has been used to measure deamination at several different sites in a gene, at neutral pH, and at several different temperatures.

At high temperatures, the rate constants for cytosine deamination determined here agree closely with previous measurements for the rate of cytosine deamination averaged over all cytosines in the molecule. Our study extends the observations to include deamination at single cytosine residues and at physiologically relevant temperatures. Thus, we have experimentally determined macroscopic first-order rate constants for deamination of both single-stranded and double-stranded DNA at 37 °C, as well as the activation energy for cytosine deamination. The single-stranded deamination reaction at 37 °C has a rate constant of 1.0×10^{-10}/s, i.e., a half-life of about 200 years. The double-stranded deamination reaction at 37 °C has a rate constant on the order of 7×10^{-13}/s, i.e., a half-life of about 30 000 years.

The previous estimate for the cytosine deamination rate constant in double-stranded DNA at 37 °C ($<1 \times 10^{-12}$/s) was extrapolated by using high-temperature data and the activation energy for single-stranded DNA obtained at temperatures from 70 to 95 °C (Lindahl & Nyberg, 1974; Lindahl, 1979). Our genetically derived activation energy for cytosine deamination in single-stranded DNA in this study was

determined from experiments performed over a temperature range of 60 °C, and extending down to 30 °C (Figure 3). The value obtained, 28 kcal/mol, agrees well with the previous, chemically derived value of 29 kcal/mol obtained over a 25 °C temperature range (Lindahl & Nyberg, 1974). Our data thus extend the Arrhenius plot into the physiological temperature range and validate the relationship between temperature and rate constant for this system. Because there is no discontinuity in the Arrhenius plot, and since the activation energy above 70 °C has been used along with the known rate constant to closely predict the rate constant of deamination in double-stranded DNA that we find at 37 °C, we deduce that the mechanism of deamination of double-stranded DNA must be similar to single-stranded DNA. Therefore, deamination of cytosine in double-stranded DNA most likely involves a single-strand intermediate.

At 80 °C, a temperature at which the DNA of the mutational target sequence should have little secondary structure, cytosine deamination occurs at approximately equal rates at different sites (Table II). The small (2–3-fold) site variation in Table II may reflect subtle differences in deamination rates or simply experimental fluctuations; a more extensive analysis would need to be carried out to confirm any such differences. Yet, the finding that deamination at 80 °C is similar for all sites is validated by the observation that, for the 80 °C incubation of the C141 mutant DNA, C → T transition revertants were equally distributed between the two detectable C141 and C142 cytosine residues *in the same codon* (Table I). In that particular codon, deamination of either cytosine can give rise to a blue plaque. In effect, the two cytosines in the same codon serve as internal standards for each other, and hence the standard deviation is simply a function of the number of plaques that can be sequenced.

The approach described here was initiated by our interest in investigating the challenge to the integrity of genetic information posed by the instability of chemical bonds in DNA. The genetic assay is sensitive enough to permit examination of the effects of various perturbations on deamination at physiological temperatures. We have shown that the rate of deamination of a single cytosine residue is 2 orders of magnitude greater for single-stranded DNA than for double-stranded DNA at 37 °C. The difference is likely attributed to the increased accessibility of cytosine in single-stranded DNA to protons and hydroxyl ions. The rate of deamination of free cytosine monomers in solution is dependent upon the concentration of protonated cytosine or OH^- ions (Garrett and Tsau, 1972; Shapiro, 1981). Therefore, any means by which protonation of the N3 position can be facilitated (such as by a loss of a base pairing partner or by an increase in the pK of the cytosine) [e.g., see Sowers et al. (1987, 1989)] would increase the probability of deamination. Likewise, any process that would facilitate OH^- attack on the C4 residue of cytosine would also increase the probability of deamination. Thus, with our assay it is now possible to readily examine the sequence dependence of deamination with respect to buffer composition, pH, dielectric constant, or the addition of agents which damage DNA. One interesting possibility is to determine whether an O^6-alkylguanine or a mismatched base on one strand induces a "cross-strand deamination" [see Sowers et al. (1987) and Williams and Shaw (1987)] of the complementary cytosine. The sequence dependence of deamination can also be determined, to ask, for example, if the rate of cytosine deamination differs in supercoiled DNA or in unusual DNA structures [reviewed in Wells (1988)] such as Z DNA, bent DNA, or a triple helix.

Acknowledgments

We thank Drs. P. Modrich and K. Tindall for critical comments on the manuscript.

Registry No. Cytosine, 71-30-7.

References

Birnboim, H. C., & Doly, J. (1979) *Nucleic Acids Res. 7*, 1513–1523.

Coulondre, C., Miller, J. H., Farabaugh, P. J., & Gilbert, W. (1978) *Nature 274*, 775–780.

Duncan, B. K. (1985) *J. Bacteriol. 164*, 689–695.

Duncan, B. K., & Miller, J. H. (1980) *Nature 287*, 560–561.

Fix, D. F., & Glickman, B. W. (1986) *Mutat. Res. 175*, 41–45.

Frederico-Zuraw, L. (1988) Ph.D. Thesis, Duke University.

Friedberg, E. C. (1984) *DNA Repair*, W. H. Freeman, New York.

Garrett, E. R., & Tsau, J. (1972) *J. Pharm. Sci. 61*, 1052–1061.

Kunkel, T. A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1494–1498.

Kunkel, T. A. (1985a) *J. Biol. Chem. 260*, 12866–12874.

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

Kunkel, T. A. (1985c) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488–492.

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

Kunkel, T. A., & Soni, A. (1988) *J. Biol. Chem. 263*, 4450–4459.

Lindahl, T. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 3649–3653.

Lindahl, T. (1979) *Prog. Nucleic Acid Res. Mol. Biol. 22*, 135–192.

Lindahl, T., & Nyberg, B. (1972) *Biochemistry 11*, 3610–3618.

Lindahl, T., & Nyberg, B. (1974) *Biochemistry 13*, 3405–3410.

Messing, J., Groenenborn, B., Muller-Hill, B., & Hofschneider, P. H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3642–3646.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Schaaper, R., & Loeb, L. A. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 1773–1777.

Sowers, L. S., Shaw, B. R., Veigle, M. L., & Sedwick, W. D. (1987) *Mutat. Res. 177*, 201–218.

Sowers, L. S., Sedwick, W. D., & Shaw, B. R. (1989) *Mutat. Res. 215*, 131–138.

Shapiro, R. (1981) in *Chromosome Damage and Repair* (Seeberg, E., & Kleppe, K., Eds.) pp 3–18, Plenum Press, New York.

Shapiro, R., & Klein, R. S. (1966) *Biochemistry 5*, 2358–2362.

Ullman, J. S., & McCarthy, B. J. (1973) *Biochim. Biophys. Acta 294*, 396–404.

Wang, R. Y., Gehrke, C. W., & Ehrlich, M. (1980) *Nucleic Acids Res. 8*, 4777–4790.

Wang, R. Y., Kuo, K. C., Gehrke, C. W., Huang, L. H., & Ehrlich, M. (1982) *Biochim. Biophys. Acta 697*, 371–377.

Wells, R. D. (1988) *J. Biol. Chem. 263*, 1095–1098.

Williams, L. D., & Shaw, B. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1770–1783.